Plausibility and Syntactic Ambiguity Resolution

Neal J. Pearlmutter
Dept. of Brain & Cognitive Sciences, E10-109
Massachusetts Institute of Technology
Cambridge, MA 02139
neal@psyche.mit.edu

Maryellen C. MacDonald
Hedco Neurosciences Building
University of Southern California
Los Angeles, CA 90089-2520
mcm@gizmo.usc.edu

Abstract

Different theories of human syntactic parsing make conflicting claims concerning the role of non-syntactic information (e.g. semantics, real world knowledge) on on-line parsing. We address this debate by examining the effect of plausibility of thematic role assignments on the processing of syntactic ambiguities. In a self-paced reading experiment, ambiguous condition reading times were longer than unambiguous condition times at the point of syntactic disambiguation only when plausibility cues had supported the incorrect interpretation. Off-line measures of plausibility also predicted reading time effects in regression analyses. These results indicate that plausibility information may influence thematic role assignment and the initial interpretation of a syntactic ambiguity, and they argue against parsing models in which the syntactic component is blind to plausibility information.

Introduction

An important issue in psycholinguistic research concerns the type of information that the sentence processor or *parser* uses in deciding which interpretation(s) to pursue for a syntactic ambiguity. The most prominent position on this issue has been that the parser consults only syntactic information during the initial parse and is blind to all non-syntactic information (Frazier, 1987; Rayner, Carlson & Frazier, 1983; Ferreira & Clifton, 1986). This position is accompanied by two other claims: 1) that the parser can pursue only one interpretation at a time, and 2) that the parser must make its choice as soon as the ambiguity is encountered. A parser with these features could make many costly mistakes, and in order to increase the chances of making correct choices, the parser could rely on parsing strategies. The best known of these is Minimal Attachment (Frazier, 1987): The parser chooses the "simplest" interpretation, building only the syntactic (phrase structure) representation with the fewest nodes. On this view, semantic and/or discourse information may affect the parse only after the parser has begun building a structure; the semantic information aids in reanalysis when the parser's first analysis fails (Rayner et al., 1983; Ferreira & Clifton, 1986).

In this paper, we argue that the parser is not as limited in the resources available to guide decisions about ambiguities as Frazier's "garden path" model suggests. We present evidence that some non-syntactic information can guide the parser in the initial interpretation of syntactic ambiguities. Specifically, we claim that lexical or discourse information affects decisions about *thematic roles* that are assigned to noun phrases in a sentence, and that these role assignments can in turn affect the choices the parser makes when confronted with a syntactic ambiguity.

A thematic role is an abstract semantic relation between a verb and one of its noun phrase (NP) arguments; these roles include the *Agent* of the action described by the verb, the *Theme* (sometimes called *Patient*) of the action, the *Instrument* of the action, and others. For example, the verb *cut* has two different argument structures; it may appear with the three arguments Agent, Theme, and Instrument, as in *Mary cut the rope with a knife*, and it may also omit the Instrument role, as in *Mary cut the rope*. Tanenhaus and Carlson (1989) have hypothesized that when a verb is encountered in the input, its alternative argument structures are activated in parallel, and tentative thematic role assignments are attempted for the NPs that have been encountered to that point. Possible assignments are evaluated by comparing information in the noun's lexical entry with the verb's argument structure. For example, Agents must be animate, and so an inanimate NP will not be assigned an Agent thematic role. If one

argument structure produces compatible assignments for the NPs encountered to that point and other structures do not, then this argument structure will be adopted by the parser. The parser will then adopt a phrase structure representation of the input that is compatible with the chosen argument structure.

Trueswell, Tanenhaus & Garnsey (1992) tested these claims in a study measuring reading times in sentences containing "main verb/reduced relative" ambiguities, in which it is temporarily ambiguous whether a verb in the sentence is the main verb (e.g., *examined* in *The defendant examined the document*) or is modifying a noun in a reduced relative clause (as in *The defendant examined by the lawyer was nervous.*). All critical sentences in the study were resolved with the reduced relative interpretation; the key manipulation was the animacy of the subject noun, as in the pair *The [defendant/evidence] examined by the lawyer turned out to be unimportant.* The noun *defendant* is animate and so could be the Agent of *examined* (permitting the simple main verb structure, in which the defendant examined something), but *evidence* is inanimate and thus cannot receive an Agent thematic role, thereby eliminating the simple main verb interpretation as a candidate for this ambiguous sentence. In the animate (*defendant*) condition, which permits the incorrect main verb interpretation, reading times were longer at the disambiguation (*by the...*) compared to unambiguous controls. In the inanimate *evidence* condition, however, there was no ambiguity effect at the disambiguation, suggesting that subjects had used animacy information early in the parse to guide thematic role assignment (assigning the Theme role to *evidence*), which in turn guided them to the correct interpretation of the ambiguity.

The Trueswell et al. results suggest that thematic role assignments, rather than strategies like Minimal Attachment, guide the parser from the outset. If so, an obvious question is how the parser evaluates the possible thematic role assignments as it is constructing a syntactic representation of the input. The Trueswell et al. data point to the importance of animacy, but other data suggest that some non-syntactic information (e.g., real world knowledge such as who is likely to send or receive flowers, Rayner et al., 1983) does not influence early parsing. One reason for this difference might be that animacy can be represented in nouns' lexical entries, independent of context: *evidence* is inherently inanimate and the Agent role for *any* verb must be assigned to an animate noun. Perhaps the feature ±Animate can influence thematic role assignment, but real world plausibility information, such as the likelihood of someone sending flowers, cannot. However, the Rayner et al. (1983) stimuli differ from the Trueswell et al. (1992) stimuli not only in the real world *vs.* lexical (animacy) nature of the information, but also in the timing of the constraints and the ambiguity. The animacy constraints arrive at the ambiguous verb (e.g. the conjunction of *evidence* and *examined*), but Rayner et al.'s real world constraints appear well after the ambiguity is introduced and so might arrive too late to influence the initial parse. If we are to examine further the role of non-syntactic information in parsing, these factors must be unconfounded so that a plausibility cue is available at the point of ambiguity.

The study below manipulates features of the subject NP in a main verb/reduced relative ambiguity, as in Trueswell et al. (1992), but instead of a lexical feature like ±Animate, we manipulate the *plausibility* of a thematic role assignment within animate NPs. For the verb *applauded*, for example, *musician* is more plausible as a Theme than as an Agent, and *audience* is a more plausible Agent than Theme. NPs that are better Themes than Agents should be implausible in the main verb interpretation and will therefore promote the reduced relative interpretation, whereas NPs that are better Agents than Themes will promote the main verb interpretation. When the disambiguation reveals that the reduced relative interpretation is correct, readers should have difficulty (evidenced by slowed reading times) only when a Poor Theme had promoted the incorrect interpretation; Good Theme sentences should not be problematic. This pattern of results would indicate that real world plausibility information can influence thematic role assignment and parsing and would be difficult for Frazier's model to explain. If plausibility does not guide the parser, however, this would be evidence that only basic lexical information like animacy can influence thematic role assignment, but that plausibility information cannot.

Method

Ratings. To develop a set of stimuli manipulating plausibility, preliminary ratings were conducted on a large pool of NP-verb combinations such as *prisoner captured, student taught, audience applauded, musician applauded,* etc. We obtained ratings of the NP as both a subject (Agent role assigned, main verb reading) and as an object (Theme role assigned, relative clause reading) of a particular verb. We also obtained a comparative preference rating between these two readings. Because raters were directly comparing the two readings, both readings were conveyed with sentence fragments, as in *The audience applauded the...* and *The audience that was applauded by the...*, for the subject and object rating respectively. Each fragment was rated on a scale from 1=good to 7=bad; subjects were instructed to base their ratings on whether the fragment made sense and described something that could happen in real life. The comparative rating was also performed on a 7-point scale, with 1 corresponding to a very strong subject preference and 7 a very strong object preference. Four

rating lists were prepared so that subjects would see a given verb in only one pair of fragments. Forty-five subjects participated in the ratings.

Materials. Using the pool of rated fragments, 24 stimulus items were constructed. Each item was constructed from two pairs of rated fragments that had contained the same ambiguous verb. The NP in one fragment pair had received a very low comparative rating (such that the NP was strongly preferred as an Agent, and thus was a Poor Theme, e.g. *The audience applauded the...* was preferred to *The audience that was applauded by the...*), and the NP in the other pair had received a very high comparative rating (such that this NP was strongly preferred as a Theme, e.g. *musician applauded*). Introductory phrases and completions were constructed for each item, as were unambiguous controls, so that each stimulus item could appear in four versions by crossing two independent variables: Ambiguity (ambiguous (A) *vs.* unambiguous (U)) and Role Plausibility (Good Theme (GT) *vs.* Poor Theme (PT)) as shown in Table 1. Sixty-one filler items with a variety of syntactic structures were also prepared, as were yes/no comprehension questions for all items.

Subjects. Fifty-two MIT students who had not participated in the ratings pretest were paid for their participation. All were native speakers of English.

Procedure. Subjects read the sentences on a CRT in a Moving Window display (Just et al., 1982) in which they pressed a key to see each word of the sentence. A trial began with a display of dashes indicating all nonspace characters of the sentence. When the subject pressed a computer terminal key, the first word appeared, replacing its dashes. With each successive keypress, the visible word reverted to dashes and the next word appeared. Pressing the key after the last word removed the sentence and displayed a comprehension question. The subject responded by pressing a key labeled "Yes" or "No" and did not receive feedback on accuracy. Subjects saw five practice items, followed by the experimental and filler items in pseudo-randomized order; at least one filler preceded any experimental item.

Table 1--Example sentence set

The producer said that the live broadcast went smoothly, and	
Good Theme (GT)	
Ambig.	the musician applauded by the host enjoyed the show immensely from start to finish.
Unambig.	the musician that was applauded by the host enjoyed the show immensely from start to finish.
Poor Theme (PT)	
Ambig.	the audience applauded by the host enjoyed the show immensely from start to finish.
Unambig.	the audience that was applauded by the host enjoyed the show immensely from start to finish.

Results

Reading time. Reading times at each word position were trimmed for each subject, replacing data points more than 3 SD over the relevant mean with the 3 SD cutoff value, affecting less than 3% of the data. The experimental sentences were divided into four regions for the reading time analysis, as shown in Figure 1. Region 1 consisted of the Good or Poor Theme NP. Region 2 contained the ambiguous verb, as well as *that was/were* in the unambiguous conditions. Region 3 consisted of the by-phrase, and Region 4 contained the first two words of the main verb phrase (the disambiguation). The remainder of the sentence was analyzed separately. Region 4 is considered the disambiguation because the by-phrase in Region 3 does not necessarily disambiguate the sentence. The preposition "by" is ambiguous between an agentive phrase (which would disambiguate) and a locative (e.g., *by the sea*, which would not disambiguate if the preceding verb had an intransitive interpretation, as some stimuli did have; see MacDonald, 1992; Pritchett, 1989). Only the last word of Region 3 suggests the correct interpretation, and a locative interpretation is never ruled out.

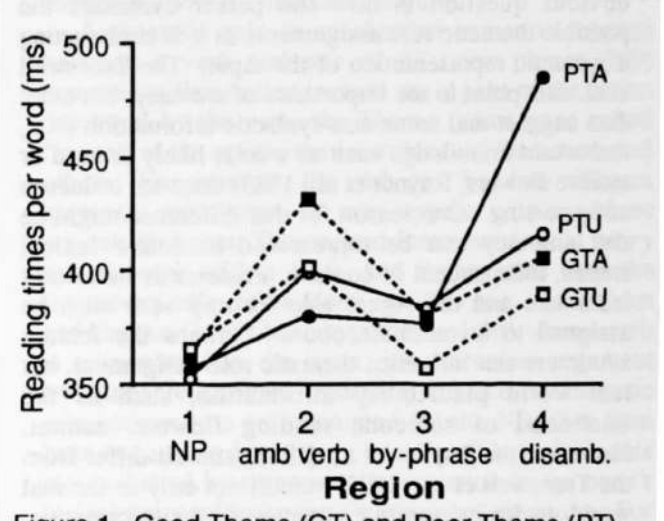

Figure 1. Good Theme (GT) and Poor Theme (PT) Ambiguous (A) and Unambiguous (U) reading times

An omnibus ANOVA revealed the predicted interaction of Role, Ambiguity and Region [$F_1(3,153) = 5.71, p = .001, F_2(3,69) = 2.87, p < .05$], justifying examination of effects at each Region separately.

As expected, there were no significant effects in Region 1, before the ambiguity was introduced. In Region 2, however, reading times were identical in the two unambiguous conditions (401 ms and 399 ms for the Good Theme (GT) and Poor Theme (PT) conditions, respectively), but there was an effect of Role on the Ambiguous items. The PT Ambiguous reading times were relatively fast (380 ms), while GT Ambiguous times were relatively slow (432 ms), a significant difference [$t_1(51) = -2.58, p < .05; t_2(23) = -3.14, p < .005$]. This result replicates MacDonald's (1992) finding that reading times in an ambiguous region are faster when the available information promotes a simple interpretation (here, the PT condition, which is plausible in the simple main verb interpretation), than when information promotes a more complex interpretation (the GT condition here, which is more plausible in the reduced relative structure than in the main verb structure). In other words, the plausibility information appears to be influencing reading times (and parsing) at the earliest possible point, and thus does not appear to be limited to later backtracking.

Analyses in Region 3, containing the by-phrase, revealed a Role x Ambiguity interaction [$F_1(1,51) = 4.28, p < .05; F_2(1,23) = 4.15, p = .05$]. The source of this interaction was that the GT Unambiguous condition (358 ms) was at least marginally faster than the other three conditions (all 377-384 ms), while these latter three did not differ from one another.

The analysis for Region 4, containing the disambiguation, also revealed a Role x Ambiguity interaction [$F_1(1,51) = 4.68, p < .05; F_2(1,23) = 2.22, p = .15$]. Further analysis revealed that in the PT conditions, reading times were significantly slower in the ambiguous sentences (485 ms) than in the unambiguous controls (416 ms) [$t_1(51) = 2.81, p < .01; t_2(23) = 2.58, p < .05$], but there was no effect of ambiguity in the GT conditions (405 *vs.* 390 ms in ambiguous and unambiguous conditions, $t's < 1$). In other words, subjects had no difficulty when plausibility favored a Theme assignment for the first NP, but they were slow when a Poor Theme had promoted the incorrect main verb interpretation. In addition, GT Unambiguous reading times were marginally faster than PT Unambiguous reading times [$t_1(51) = 1.90, p < .10; t_2(23) = 2.10, p < .05$], indicating an effect of plausibility even in the absence of ambiguity.

Analyses of the remaining words showed small effects of Role, but neither the Ambiguity factor nor its interaction with Role was robust. There was also no reliable effect of any factor on comprehension question accuracy.

In sum, the reading time data show a clear difference between the ambiguity effects for the Good and Poor Theme conditions. In the PT condition, containing no helpful plausibility information, reading times at the ambiguous verb were faster than in the unambiguous condition, but they increased substantially over unambiguous times at the disambiguation. Subjects appear to parse these sentences as the garden path model predicts, initially interpreting the ambiguous verb as the main verb of the sentence, and then showing surprise at the disambiguation. By contrast, the GT Ambiguous condition produced slower reading times at the ambiguous verb, suggesting that subjects are already considering the (correct) relative clause interpretation at this point. Later, at the disambiguation, subjects show no significant increase in reading times over the unambiguous condition. This pattern suggests that either subjects are using the plausibility information to choose an interpretation well before they reach the disambiguation, or, if Minimal Attachment is acting first and the plausibility information is affecting backtracking later, then this sequence of events must all be taking place very early, during the reading of the ambiguous verb in the GT condition, because subjects seem to have completely settled on the correct interpretation by the disambiguation. While it is conceivable that misanalysis and reanalysis are both occurring during the reading of a single word, the simpler explanation is that the plausibility information is guiding thematic role assignment and parsing from the outset. Additional evidence is provided by the relationship between the reading time and rating data, which we consider next.

Ratings. The reading data demonstrated that the sentences in the GT conditions did not show the same ambiguity effects as those in the PT conditions; the ratings will investigate whether the plausibility constraints operate along a continuum. The ambiguous and unambiguous reading times and ambiguous-unambiguous difference scores in the four regions were correlated with the object ratings we obtained in our initial norming procedures. The object ratings were those in which subjects rated the goodness of the NP as a direct object in unreduced relative clause fragments such as *The audience that was applauded by the....* The object ratings therefore reflect the plausibility of the eventual disambiguation of the ambiguity in the experimental items.

Because the stimuli were selected with the aid of comparative ratings (in which subjects indicated which interpretation was more plausible), it is important to examine the range of object ratings at each level of Role to be sure that these ratings are not merely recreating the dichotomous Role variable. There was a small range of object ratings in the Good Theme condition (1.07-2.27) and a wide range of ratings within the Poor Theme condition (1.95-5.27). The distributions of

ratings in the two conditions thus overlap, with some Good Theme items having poorer ratings than some Poor Theme items. Given this range, correlations between ratings and reading times should be informative. It should also be useful to examine the correlations within each level of Role, though only the strongest relationships could be identified, because the range of ratings and statistical power are reduced when only half of the items enter into the analysis.

The rating/reading time correlations for the entire stimulus set are shown in Table 2. As expected, correlations for Region 1 (before the ambiguity was introduced) were all non-significant, but in Region 2 (the ambiguous verb), the negative correlation with ambiguous reading times was marginally significant, indicating that ambiguous verbs following better Theme NPs tended to be read more slowly than when the NP was a less plausible Theme. In other words, subjects read the ambiguous verb more slowly when plausibility information provided some reason to expect a complex relative clause structure, and they read more quickly when plausibility information promoted the simple main verb interpretation. As mentioned above, this effect is important, as it suggests that the plausibility of alternate interpretations is affecting the *initial* interpretation of the sentence. Moreover, the correlation was present within the items in the Good Theme condition [$r = -.36, p < .10$]. Thus the best of the Good Theme items were read more slowly at the ambiguity than were slightly poorer Good Theme items.

In Region 3 (the by-phrase), ambiguous reading times were unrelated to the ratings, but unambiguous reading times decreased as the ratings improved. These results show that within a few words after encountering the Region 2 verb, the language processor is sensitive to the anomaly of assigning a Theme role to a poor Theme in an unambiguous relative clause. These effects on unambiguous reading times were also present, though smaller, within the Poor Theme items considered separately [$r = .26, p < .25$] and within the Good Theme items [$r = .38, p < .10$].

Table 2--Rating/Reading time correlations

	Region			
	1 NP	2 amb vrb	3 by-phr.	4 disamb
Amb	-0.10	-0.27*	-0.05	0.49***
Un	-0.22	0.07	0.34**	0.28**
A-U	0.07	-0.23	-0.29**	0.26*

A-U = Amb-Unamb Difference score. Ratings are 1=good to 7=bad; positive correlations therefore indicate that better ratings produce faster reading times or smaller A-U differences.

*$p < .10$. **$p < .05$. ***$p = .0005$.

In Region 4 (the disambiguation), both the ambiguous and unambiguous reading times correlated well with the ratings, though the effect was much stronger for the ambiguous items, suggesting that plausibility accounts for some effects in all sentences regardless of ambiguity, but that it has an additional effect on the ambiguous items. The correlation between the ratings and the difference scores indicates that as the initial NP becomes a better Theme, the ambiguity effect decreases at the point of disambiguation: Subjects expect the reduced relative structure when the plausibility information favors it, and not when plausibility promotes a main verb structure.

In sum, the correlations confirm and extend the major effects in the reading times: As the Region 1 NPs become poorer Themes and so promote the incorrect simple interpretation, reading times on the ambiguous verb in Region 2 become faster, and reading times in the disambiguating Region 4 become slower. As mentioned above, it is difficult to ascribe these effects to backtracking and revision; it appears that the early plausibility information supporting a relative clause encourages subjects to adopt that structure well before the definitive disambiguation is reached. Moreover, the fact that meaningful correlations can be found within the Good and Poor Theme items considered separately demonstrates the graded nature of the plausibility cues investigated here. It is not only dichotomous lexical features like ±Animate that influence syntactic parsing: Our graded plausibility effects also have a rapid influence on ambiguity resolution.

Discussion

The results of the present experiments are consistent with the Trueswell et al. (1992) findings but contrast with the Rayner et al. (1983) results that suggested that plausibility information could not help to determine the course of first-pass parsing. The differences across studies suggest that sufficiently strong plausibility cues present at the onset of an ambiguity can aid in initial syntactic parsing, as in the present experiments, but that weaker and late-occurring plausibility cues in the Rayner et al. study could not influence early processing, so that subjects still had difficulty in those sentences at the disambiguation. Our work in progress pursues these issues in two ways: 1) manipulation of the information between the ambiguity and the disambiguation (which was always a by-phrase here, cf. MacDonald, 1992; Rayner et al., 1983), and 2) examination of the role of working memory capacity (Just & Carpenter, 1992) on the ability to compute the plausibility information that guides thematic role assignment and ambiguity resolution.

Our results to date suggest that probabilistic plausibility cues do not aid parsing in an all-or-nothing

manner, but rather that the ease with which the parse is handled is directly related to the strength of the plausibility cue. These findings do not lend support to a two-stage model like the garden path model (Frazier, 1987), in which plausibility information can influence only the second stage of analysis (backtracking), unless, as we have noted, the first and second stages can take place during the processing of a single word, a modification that substantially weakens such models. By contrast, our results are easily explained by the class of models discussed by Tanenhaus and Carlson (1989), in which the syntactic parser builds its syntactic interpretation based on the best-supported of the available verb argument structures. Trueswell et al. (1992) showed that the "best-supported" argument structure could be influenced by relevant lexical semantic information such as noun animacy; the current data indicate that plausibility information arising from the combination of noun and verb meanings can also influence this determination.

Our results are also compatible with Altmann & Steedman's (1988) "weakly interactive" model of parsing in which the parser proposes multiple structural analyses in parallel, and pragmatic factors determine which interpretation will be pursued by the parser. What is not yet clear from any of this research is the level of representation that is mediating the plausibility effects: It may be the discourse representation of the sentence, incorporating real world knowledge, or it may be that activation of a rich lexical/semantic representation of nouns and verbs like *audience* and *applauded* is sufficient to guide thematic role assignments. An additional possibility is that multiple sources of information constrain thematic role assignments, but over different time courses, so that rapidly computed information (e.g. animacy) has an earlier effect than more subtle information such as the plausibility information investigated here. Additional study of this issue, especially that focusing on the graded nature of the cues, will have important implications for theories of both syntactic parsing and lexical representation.

Acknowledgments

Much of this research was completed while M. MacDonald was at MIT, where this research was supported by BRSG 2 S07 RR07047-23 from the National Institutes of Health.

References

Altmann, G. & Steedman, M. 1988. Interaction with context during human sentence processing. *Cognition, 30*, 191-238.

Ferreira, F., & Clifton, C. 1986. The independence of syntactic processing. *Journal of Memory and Language, 25*, 348-368.

Frazier, L. 1987. Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading.* Hillsdale, NJ: Erlbaum. 559-586.

Just, M.A. & Carpenter, P.A. 1992. A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*, 122-149.

Just, M. A. , Carpenter, P. A., & Woolley, J. D. 1982. Paradigms and processes in reading comprehension. *Journal of Exp. Psychology: General, 3*, 228-238.

MacDonald, M. C. 1992. *Probabilistic constraints and syntactic ambiguity resolution.* Submitted.

Pritchett, B. L. 1988. Garden path phenomena and the grammatical basis of language processing. *Language, 64*, 539-576.

Rayner, K., Carlson, M., & Frazier, L. 1983. The interaction of syntax and semantics during sentence processing. *Journal of Verbal Learning and Verbal Behavior, 22*, 358-374.

Tanenhaus, M. K., & Carlson, G. 1989. Lexical structure and language comprehension. In W. Marslen-Wilson, (Ed.), *Lexical Representation and Process.* Cambridge, MA: MIT Press.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. 1992. *Semantic influences on parsing: Use of thematic role information in syntactic disambiguation.* Submitted.

The Time Course of Metaphor Comprehension

Phillip Wolff and Dedre Gentner
Department of Psychology
Northwestern University
2029 Sheridan Road, Evanston, IL 60208
wolff@ils.nwu.edu, gentner@ils.nwu.edu

Abstract

This research investigates the process by which people understand metaphors. We apply processing distinctions from computational models of analogy to derive predictions for psychological theories of metaphor. We distinguish two classes of theories: those that begin with a *matching* process (e.g Gentner & Clement, 1988; Ortony, 1979) and those that begin with a *mapping* process (e.g. Gluckserg and Keysar, 1990). In matching theories, processing begins with a comparison of the two terms of the metaphor. In mapping theories, processing begins by deriving an abstraction from the base (or *vehicle*) term, which is then mapped to the target (or *topic*).

In three experiments, we recorded subjects' time to interpret metaphors. The metaphors were preceded by either the base term, the target term, or nothing. The rationale was as follows. First, interpretations should be faster with advanced terms than without, simply because of advanced encoding. The important prediction is that if the initial process is mapping from the base, then seeing the base in advanced should be more facilitative than seeing the target in advanced. Matching models predict no difference in interpretation time between base and target priming. The results generally supported matching-first models, although support for mapping-first models was found with highly conventional metaphors.

Introduction[1]

How are metaphors understood? For instance, on hearing *A surgeon is a butcher*, we apprehend the meaning of the metaphor to be something like "a surgeon is someone who cuts flesh." How do we derive this interpretation? Most current psychological theories of metaphor describe the process of comprehension in only very general terms. In this research we apply processing distinctions made in computational models of analogy to specify and test broad classes of processing models in psychology.

Approaches to Processing

From computational models of analogy we can derive three general classes of processing algorithms, which can be described as *matching*, *mapping*, and *matching-then-mapping* (Gentner, 1989).(See Figure 1.) In *matching* models, the commonalities of two representations are recognized by aligning their parts and structures (Winston, 1982). We might interpret the metaphor *A surgeon is a butcher*, for instance, by noting that both cut flesh. Ortony's (1979) salience imbalance theory of metaphors is primarily a matching theory. In *mapping* theories, processing begins by accessing or creating from the base term a higher-order category or abstraction which is then used to attribute properties to the target term (e.g., Burstein, 1983; Carbonell, 1983; Greiner, 1988; Kedar-Cabelli, 1985). Glucksberg and Keysar (1990) have recently proposed a theory of metaphor that fits into this processing framework. According to their model, a metaphor like *My job is a jail* is understood by accessing or deriving an abstraction or category associated with the *base* term, or *vehicle*, *jail*; this abstraction (e.g., confining institution) is then applied to the *target* term, or *topic*, *job*. We will refer to this particular instantiation of the mapping perspective as *Category-mapping*.

The third kind of processing can be termed *Matching-then-mapping* (Falkenhainer, Forbus, & Gentner, 1989; Hofstadter, Mitchell & French, 1987; Holyoak & Thagard, 1990). For example, in the

[1]This work was supported by NSF grant BNS 87-20301 and ONR grant N00014-89-J1272 awarded to the second author.