

## Metaphor in Discourse

Nicholas Asher

Department of Philosophy,  
University of Texas,  
Austin,  
Texas 78712,  
USA  
nasher@bertie.la.utexas.edu

Alex Lascarides

Centre for Cognitive Science,  
University of Edinburgh,  
2, Buccleuch Place,  
Edinburgh, EH8 9LW,  
Scotland, UK  
alex@cogsci.ed.ac.uk

### 1 Introduction

This paper focuses on metaphor and the interpretation of metaphor in a discourse setting. There have been several accounts put forward by eminent philosophers of language—Max Black, John Searle and Donald Davidson, among others—but none of them are satisfactory. They offer a few rules for metaphoric interpretation, but many of them are redundant, and they form a list without much coherence.

Many have thought that the principles of metaphorical interpretation cannot be formally specified. We'll attack this position with two claims. Our first claim is that some aspects of metaphor are productive, and this productivity can be captured by perspicuous links between generalisations that are specified in the lexicon, and general purpose circumscriptive reasoning in the pragmatic component. Indeed from a methodological perspective, we would claim the productive aspects of metaphor can give the lexicographer clues about how to represent semantic information in lexical entries.

Moreover, it is well known that domain knowledge influences metaphorical interpretation. Our second claim takes this further, and we argue that *rhetorical relations*—such as *Elaboration*, *Contrast*, and *Parallel*, among others—that connect the meanings of segments of text together, also influence this. Through studying these cases, we learn how to link lexical processing to discourse processing in a formal framework, and we give some preliminary accounts of how the link between words and discourse determine metaphor.

### 2 Metaphor in the Lexicon

But first, we consider an example of metaphoric productivity. The class of words concerned, is the 'normal' vs. the metaphorical use of verbs describing change of location (CoL) in English and in French:

- (1) a. Jean est passé de l'autre côté de la loi.
- b. Jean est allé (parti) à la guerre.
- c. He entered a blue funk/ came out of his blue funk.
- d. Sam crossed the line of permissible behavior.
- e. He deviated from the norm.
- f. Jean est à la guerre.
- g. He is outside (within) the law.
- h. He was from outer space.
- i. He was way off base.
- j. He skirted the boundaries of the permissible.
- k. She stayed right on target.
- l. You have now entered the Twilight Zone.

(1a–e,j–l) employ verbs of motion or motion complexes, that in the typology described in Asher and Sablayrolles (1995), describe changes of location (CoL). (1f–i) employ complexes describing a present or past location or position of the subject; these have a metaphorical behaviour very similar to the behaviour of CoL complexes, and so we'll loosely refer to these as CoL verbs too for the purposes of this abstract.

These verbs conventionally apply to spatial locations. Asher and Sablayrolles (1995) describe a taxonomy, in which the various CoL verbs are grouped into ten classes, which place restrictions on the kinds of arguments they take. For example *entrer* must have the interior of some object or place *l* as an argument and it must describe an event of moving from some location near *l* inside *l*. On the other hand, a verb like *arriver*, though similar to *entrer* in that it takes the interior of some object or place *l* as an argument and describes an event of moving inside *l*, differs from *entrer* in that the movement must have its source or starting point outside the contextually defined "nearby zone" of *l*. The verb *sortir* resembles *entrer* in that it describes a motion involving the inside of a "reference location" *l* and a point in the near outside, but the movement goes in the opposite direction; *l* is the source of this movement and the point in the neighborhood is the goal. All CoL verbs have in fact a reference location *l*, but their meaning differs in how this reference location is exploited in constraints on the source, goal and intermediate path of the eventuality described.

This kind of information can be represented in a typed feature structure (TFS) framework, and unification and

default unification can then be used to build phrases from the lexical entries, while ensuring that the arguments to verbs are of the right type, via the typing system that accompanies the TFSS. This is a model of lexical processing that has been suggested by many computational linguists (e.g., Boguraev and Pustejovsky 1990, Briscoe *et al* 1990, Evans and Gazdar 1989a, 1989b, Flickinger 1985, Lascarides *et al* 1994, and others), and in previous work, we have suggested how such a model can be linked to a theory of discourse processing, to handle disambiguation (Asher and Lascarides 1995a, Asher and Sablayrolles 1995) metonymy (Lascarides *et al* 1994, Lascarides 1995) and Zeugma (Lascarides, Copestake and Briscoe, 1994).

But these lexical entries for CoL verbs cannot be the ones that are used to represent the sentences in (1); at least not directly. For all the sentences in (1) violate the type constraints on the CoL verbs. Attempting to unify the TFSS in these cases will fail.

However, the sentences (1) expose a lexical generalisation, which should be represented in a formal and computationally tractable manner. The lexical generalisation is this: the essential structure of the CoL verbs—that they describe a change of location from a source to a goal via an intermediate path—is preserved in the metaphorical meaning. In other words, what distinguishes a particular CoL verb from other kinds of verbs in the conventional use, also distinguishes it from those verbs in their metaphorical use.

Ideally, we should capture this constraint as a generalisation in the lexicon: the lexicon should predict which aspects of a lexical entry can vary in metaphorical interpretation, and which can't. If we include in the semantic component of lexical entries a new feature—call it *differentiae*—which explicitly stipulates how that lexical entry is different from words in its subclasses and superclasses, then we can capture the constraint that the differences are maintained in a metaphorical setting, by a lexical rule, which is a function between TFSS in the spirit of Briscoe and Copestake (1991).

The Metaphor Lexical Rule will take as input those word classes that can be used metaphorically (this excludes auxiliaries, for example). The output will specify those elements of a word that can vary from conventional use to metaphor, and those that can't. The fact that the types of arguments can vary, can be captured by ensuring the output of the lexical rule assigns a general type ( $\top$ ) to the arguments. The fact that the *differentiae* cannot vary, can be captured by structure sharing or reentrancy between the *differentiae* of the input, and the *differentiae* of the output. So, in the case of CoL verbs, the Metaphor Lexical Rule will capture the fact that the *distinctions* among the CoL verbs and verb complexes, described in Asher and Sablayrolles (1995), are *preserved* in the metaphorical sense extensions. We assume the *differentiae* of CoL verbs stipulates the constraints on the source, goal and intermediate path that distinguishes it from other verbs. And so the reentrancy on the *differentiae* features will ensure that these constraints are maintained, but the types of the arguments the verbs take won't be constrained to **physical-space**

anymore.

In order to make any predictions about the sorts of examples we have seen in (1), we must apply the Metaphor Lexical rule not only to CoL verbs but also to their arguments. The Metaphor Lexical rule interacts with unification to constrain metaphorical meaning and to set certain limits on what metaphors are possible and what metaphors are not. The feature structures of the various lexical entries must unify together, to form the feature structure for the whole sentence, and this means, for instance, that the features of the arguments of a verb like *partir* must be compatible with the *differentiae* features of the verb class.

In the case of CoL verbs, this means that at the very least the arguments must have some sense extension that allows them to be interpreted as locations in some qualitative space. In concrete terms, how does this restrict the arguments of CoL verbs in a metaphorical interpretation? We are on less certain ground here than with the CoL verbs themselves, because we have not made an exhaustive study of these arguments. Nevertheless, we tentatively propose two generalizations. First, we claim that any common noun or NP  $\alpha$  that has a lexical antonym allows us to construct a spatial interpretation for an NP containing  $\alpha$  in which we have two spaces: one for the metaphorical denotation of  $\alpha$ ; and one for its antonym, which functions as a *far outside* zone. Second, if  $\alpha$  is part of a scalar system (e.g. good, bad, better, worse), then we claim that  $\alpha$  has a metaphorical, spatial interpretation in which the *near* neighbourhoods of reference locations are present.

So for instance, the Metaphor Lexical Rule and unification predict that (1b) is fine, because war has *differentiae* that distinguish it from another related lexical concept (an antonym) that allows us to have a metaphorical far outside (the opposite), and so the constraints on the metaphorical sense of *aller* are satisfied. *War* isn't a scalar concept and so does not necessarily support a spatial interpretation in which we can distinguish *nearby* neighborhoods. So this means that a CoL verb that requires that the goal be a near outside of the state in which one is in (peace) such as *sortir* cannot work metaphorically with the same arguments, which is indeed the case

- (2) \*Il est sorti à la guerre.  
La France est entré dans la guerre.

On the other hand, a scalar concept should be able to support the inner CoL verbs, if our generalizations are correct:

- (3) Il est sorti de son état de mauvaise humeur.  
Il est entré dans un état de mauvaise humeur.

The Metaphor Lexical Rule combined with our observations about the ontological status of certain locations also predicts that (1d) is fine, while (4) is not.

- (4) \*He has entered the line of permissible behavior.

A line, even one in a qualitative space defined by possible behaviors (a scalar concept again), is something that can be crossed but that cannot be entered, since lines in our commonsense geometry have no extension (for a

discussion see Aurnague and Vieu 1993). Thus, in (4) the unification process will fail.

Through studying certain aspects of metaphorical productivity, we have learnt about the kinds of features that we need in the semantics: representing *differentiae* would be useful for capturing lexical constraints on the degree of variation permitted in metaphorical interpretation.

### 3 Metaphor in Context

Having encoded constraints on metaphorical senses via the Metaphor Lexical Rule, we still have the problem of determining the metaphorical interpretation of a word in a discourse setting. Since the Metaphor Lexical Rule is only intended to represent productive constraints on metaphorical interpretation, it gives only an underspecified semantics to the lexical item, which must be fleshed out with pragmatic reasoning.

If the best metaphoric interpretation involves the minimal change to the conventional lexical entry, we can technically compute a minimal change in lexical meaning to get the fully specified metaphorical meaning, by *circumscribing* the changes, as in McCarthy (1986). One circumscribes the changes in the set of overall features and values in the lexical entries, while maintaining the constraints imposed by the Metaphor Lexical Rule on what must remain constant, and allowing features and values to vary so as to permit a successful unification.

Circumscription can be represented within the logical framework of Commonsense Entailment (CE) used to build up discourse representations in Asher's (1993) SDRT and Lascarides and Asher (1993) as follows: Let  $\Box(f_1 \rightarrow \phi)$  mean that feature  $f_1$  contains the information that  $\phi$ , while  $\Box(f_1 \rightarrow d : \phi)$  means that  $f_1$  has the information that  $\phi$  concerning its *differentiae*. Then we can determine the metaphorical sense of  $f_1$ —which is the FS  $M(f_1)$ —by exploiting the nonmonotonic consequence operation of CE with the following two axiom schemata ( $M$  is thus a function from FSS to FSS, constrained by these two schemata):

- $(\Box(f_1 \rightarrow \phi)) > (\Box(M(f_1) \rightarrow \phi))$
- $(\Box(f_1 \rightarrow d : \phi)) \rightarrow (\Box(M(f_1) \rightarrow d : \phi))$

These axiom schema will produce a feature structure  $f_2$  with the minimal number of changes compatible with the other constraints one knows about  $f_2$ . The combination of lexical rule and circumscription provides (a) an account of which features can vary and which can't in metaphorical interpretation, and (b) a defeasible reasoning mechanism for fleshing out the particular metaphorical interpretation, in a discourse setting.

One can also provide more specific lexical rules, that work on more specific word classes, to capture particular, more specific aspects of metaphorical productivity (cf. Briscoe and Copestake 1991). Ortony (1979) cites a further productive case of metaphor: adjectives which apply to physical objects, can be applied in a metaphorical sense to humans: *straight, bent, soft, hard, narrow, broad*, and so on. The licensing of these metaphorical uses of adjectives can be represented via a lexical rule, which permits the application of the adjective to an argument of type **human**.

But this lexical rule would not specify the semantics of the metaphorical sense in any detail. Pragmatics will again be needed to compute this.

Having captured this lexical generalisation, we can capture a further one: that physical objects can apply as predicates to humans, and they receive a metaphorical interpretation where the adjectives that apply to the physical object are now applied in their metaphorical sense to the human. So, for example, the lexical rule would predict that *rock* can take an argument of type **human**, and that in this metaphorical use, the adjectives that apply to the original **physical-object** meaning of *rock*, now apply in their metaphorically shifted sense to the **human**. So (5) is interpreted as *John is solid, heavy, hard to move*, and so forth.

(5) John is a rock.

It must be stressed, however, that since the metaphorical senses of the adjectives are under-specified in the lexicon, then so is the metaphorical sense of *rock*. The lexical rule licenses the predicate argument structure in (5), but gives few clues about the resulting meaning. Pragmatics—perhaps in the circumscriptive framework we suggested above—is needed for this task. This is why although the interpretation of (5) is relatively clear (because this is a fairly well-used, and arguably conventionalized use of *rock*), the metaphorical meaning of (6) is unclear, without further pragmatic information, even though the predicate-argument structure of the sentence is licensed by the lexical rule:

(6) Sam is a pebble.

This brings us to the role of discourse structure and pragmatics in the interpretation of metaphor. Although the meaning of (6) in isolation of any discourse context is hard to compute, this is ameliorated in (7):

- (7) a. John is a rock.  
b. But (compared to John) Sam is a pebble.

We would argue that the *Contrast* relation in (7) provides the information we need in order to calculate the metaphorical meaning of *pebble* at the discourse level. Note that it cannot simply be domain information about pebbles and rocks, since this was available for the interpretation of (6) and failed to provide clues. Rather, the clue lies in the *juxtaposition* of (7a) and (7b), and the fact that (7b) *Contrasts* with (7a).

To see how the *Contrast* relation determines metaphor in (7), we'll link the lexical reasoning described above, to the semantics of *Contrast* in Segmented Discourse Representation Theory (SDRT) (Asher, 1993). Using the techniques described in Asher (1993), we must check the coherence of the *Contrast* relation in (7), by ensuring that the discourse constituents are isomorphic both structurally and semantically, and that there is a contrasting theme between them. There is a default heuristic that one aims for as much structural and semantic isomorphism as possible. Maximum isomorphism is achieved, if *rock* is associated with *pebble* in the structural mapping between the constituents, and the semantics of these items contrast each other. Therefore, if we successfully interpret the metaphorical use of *rock* in (7a), via the

lexical and circumscriptive reasoning mentioned above, we can compute what the metaphorical meaning of *pebble* is in this context. If *rock* means *reliable*, then *pebble* must mean *unreliable*, for the contrasting theme to be as maximal as possible.

To see how the discourse component, the lexical rules and the pragmatic reasoning interact, let us look at this example in more detail. Because of the type clash in constructing the SDRS constituent  $\alpha(j)$  for (7a), we are forced to a metaphorical interpretation of *rock* using the more specific lexical metaphorical rule concerning **physical-objects** (POMLR). Now we process (7b) and attempt to construct a constituent  $\beta(s)$  for this sentence. Again unification fails and we must construct a metaphorical interpretation of *pebble*. This time, however, we do not have enough pragmatic information to get an informative constituent  $\beta'(s)$ ; as we saw above, information independent of the discourse context is not sufficient to give any determinate meaning to the metaphorical sense of *pebble* appropriate to human arguments. So now we attempt to exploit the discourse structure to add content to  $\beta'(s)$ . By the rules for attachment in SDRT,  $\alpha(j)$  is the only possible attachment point, and the use of *but* in the second sentence forces us to conclude *Contrast*( $\alpha(j), \beta'(s)$ ). But *Contrast* holds only if the two constituents yield contrasting themes, and there is not enough content in  $\beta'$  to construct the appropriate, contrasting theme (for details of the mechanisms for constructing themes, see Asher 1993).

We hypothesize that a strategy of accommodation comes into play in cases like these. Since there is sufficient information in the conventional properties on which the two sentences are derived to construct two contrasting themes (*x is a rock* vs. *x is a pebble*) for the respective constituents, we use these themes to verify the necessary coherence constraints on *Contrast*( $\alpha(j), \beta'(s)$ ). The presence of this *Contrast* relation now forces a revision of the truth conditional content of  $\beta'(s)$ ; the representation of the meaning of (7b) (again, for details see Asher 1993). According to this revision, we replace  $\beta'(s)$  with  $\neg\alpha(s) \wedge \beta'(s)$ , which gives us the intuitively right results: namely, that (7b) in this context means that Sam is a pebble, and not a rock.

We now consider an example where the discourse context actually *triggers* a metaphorical interpretation. Sentence (8), in the absence of information to the contrary, means that I climbed a (physical) greasy pole.

(8) I have climbed up that greasy pole.

There is no reason to believe that the sentence is metaphorical, because unification between the CoL verb *climb* and the **physical-space** *greasy pole*, succeeds. However, in (9)—which is a slightly expanded example from Searle (1979), that he puts in the mouth of Disraeli—the *Contrast* relation not only helps us determine the metaphorical interpretation of (9b), but unlike (7), the discourse context *triggers* the metaphor in the first place.

- (9) a. I have always despised politics.  
b. But I have climbed to the top of that greasy pole.

Again using the techniques in SDRT, one must check the coherence of the *Contrast* relation, and aim for maximum structural and semantic isomorphism between the constituents. Furthermore, the anaphoric expression *that greasy pole* must be resolved to an available antecedent from the discourse context. Given the discourse context, SDRT predicts there is only one candidate antecedent: *politics*. But the type hierarchy prevents *that greasy pole* from being identified with *politics*, unless it undergoes type coercion. Therefore, through calculating the anaphora resolution, with the aid of the discourse structure, we learn that *that greasy pole* has a metaphorical interpretation, and is identified with *politics* (and so *greasy* is assigned a metaphorical interpretation, which is ascribed to *politics*).

But now, the conventional CoL verb *climb* cannot combine with *that greasy pole* anymore. The default unification that is used to build the phrase (9b) from the conventional CoL verb and the resolved anaphoric expression fails, because the latter expression is not of the type **physical-space**, and so using a principle of Charity familiar from philosophical work on metaphor (Black 1962, 1979, Searle 1979), *climb* must be interpreted in a different way.

The constraints described in the Metaphor Lexical Rule, together with the circumscriptive reasoning, that as many properties of *climb* as possible are preserved in the metaphorical interpretation, predicts that all the properties of the source, intermediate path and goal of *climb* in (9b)—save that they are of type **physical-space**—are preserved. So, Disraeli starts at the bottom of politics, and climbs to the top. We must therefore assign *politics* a qualitative scale: the career hierarchy (clerk to prime minister) is a plausible candidate, and under this interpretation, (9b) means that Disraeli worked his way up the career hierarchy.

A particularly compelling metaphor may change the features that are present in a lexical entry, albeit temporarily. Arguably, this occurs in Romeo's metaphor for Juliet:

- (10) a. What light through yonder window breaks?  
b. It is the East, and Juliet is the sun!

*William Shakespeare, Romeo and Juliet, Act II, Scene 2*

Since (10a) is a question, by the constraints in SDRT it must form part of a question answer pair (*QAP*) (Asher and Lascarides, 1995b). An answer to this question is a proposition which asserts what light it is. Indeed, as we'll see below, (10b) gives us an answer, but not a direct one; we have to infer what the direct answer is.

One task is to resolve the anaphor *it* in (10b). Just as in (9), the discourse context constrains this to be identified with either the window or the light. Both invoke type violations, and will force a metaphorical interpretation. The identification of the anaphor with *window* is more plausible, because the mass term *light* doesn't determine a physical location to the extent that *window* does. Having supposed that the window is the East, it is now easy to infer an answer to the question (10a) at the metaphorical level. The metaphorical sun is rising in the

East (which, physically, is the window). And the second constituent in (10b) specifies what the metaphorical *sun* is in the physical location of the window: it's Juliet (so we've now inferred that Juliet is at the window). One could infer further content, using world knowledge of the sort Shakespeare might have assumed his audience would have at their disposal (like a roughly Ptolomaic view of the universe); for example, as Juliet is the metaphorical sun, she is at the centre of Romeo's world.

## 4 Conclusion

We have shown how to compute metaphorical interpretations from a combination of three mechanisms: lexical rules which specify the range of possible meaning shifts of classes of words; circumscriptive techniques for fleshing out metaphorical interpretation in different pragmatic contexts; and a theory of discourse structure, which provides mechanisms for revising the truth conditional content of metaphorically interpreted constituents, relative to the context in which they're uttered. We showed how this can explain data concerning: verbs involving change of location; the metaphorical shift of meaning of words that refer to kinds of physical objects when they are predicated of persons; and the dependence of metaphorical interpretation upon discourse structure. What we have done, of course, is very far from a comprehensive theory of metaphor. At best we have offered a proof of concept of an approach. But by using the modern logical tools of formal pragmatics and semantics, we hope that we and others can make progress on this difficult subject and that in turn a better understanding of metaphor will enhance our understanding of lexical meaning and lexical processes.

## 5 References

Asher, N. [1993] *Reference to Abstract Objects in Discourse*, Kluwer Academic Publishers.

Asher, N. and Lascarides, A. [1994] Intentions and Information in Discourse, *Proceedings of the 32nd Annual Meeting of the Association for Computational Linguistics*, pp34-41, Las Cruces.

Asher, N. and Lascarides, A. [1995a] *Lexical Disambiguation in a Discourse Context*, *Journal of Semantics*.

Asher, N. and Lascarides, A. [1995b] Questions in Dialogue, submitted to ACL95.

Asher, N. and Sablayrolles, P. [1995] A Typology and Discourse Semantics for Motion Verbs and Spatial PPs in French, *Journal of Semantics*.

Aurnague, M. and Vieu, L. [1993] Towards a Formal Representation of Space in Language: A Commonsense Reasoning Approach, *Proceedings of the IJCAI93 workshop on Spatial Reasoning*.

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

Black, M. [1979] More About Metaphor, in Ortony, A. (ed.) *Metaphor and Thought*, Cambridge University Press, 1979, pp19-45.

Briscoe, E. J., Copestake, A., and Boguraev, B. [1990] Enjoy the Paper: lexical semantics via lexicology, *Proceedings of the 19th International Conference on Computational Linguistics (COLING90)*, Helsinki, pp42-47.

Boguraev, B., and Pustejovsky, J. [1990] Lexical Ambiguity and the Role of Knowledge Representation in Lexicon Design, *Proceedings of the 19th International Conference on Computational Linguistics (COLING90)*, pp36-42, Helsinki.

Evans, R. and Gazdar, G. [1989a] Inference in DATR, *Proceedings of the 4th Conference of the European Chapter of the Association for Computational Linguistics (EACL89)*, Manchester, England, pp66-71.

Evans, R. and Gazdar, G. [1989b], The semantics of DATR, in Cohn, A. G. (ed.) *Proceedings of the Seventh Conference of the Society for the Study of Artificial Intelligence and Simulation of Behavior (AISB)* pp79-87, Pitman/Morgan Kaufmann, London.

Flickinger, D. *Lexical Rules in the Hierarchical Lexicon* PhD thesis, Stanford University.

Lascarides, A. and Asher, N. [1993] Temporal Interpretation, Discourse Relations, and Commonsense Entailment, *Linguistics and Philosophy* 16, pp437-493.

Lascarides, A., Briscoe, E. J., Asher, N. and Copestake, A. [1994] Order Independent and Persistent Default Unification, Acquilex Research report, and in Schöter, A. and Vogel, C. (eds.) *Non-Classical Feature Systems*, Edinburgh Working Papers, volume 10; also submitted to *Linguistics and Philosophy*.

Lascarides, A. [1995] The Pragmatics of Words, this volume.

Lascarides, A., Copestake, A., and Briscoe, E. J [1994] Word Puns: Ambiguity and Coherence, ACQUILEX research report.

Ortony, A. [1979] *Metaphor and Thought*, Cambridge University Press.

Pustejovsky, J. [1991] The Generative Lexicon, *Computational Linguistics* 17 4, pp409-441.

Searle, J. [1979] Metaphor, in Ortony, A. (ed) *Metaphor and Thought*, Cambridge University Press, pp92-123.

Searle, J. [1979] Metaphor, in Ortony, A. (ed) *Metaphor and Thought*, Cambridge University Press, pp92-123.