Finding Metaphor in Discourse: Pragglejaz and Beyond

GERARD STEEN

VU UNIVERSITY AMSTERDAM

ABSTRACT: This is a methodological paper which addresses three distinct ways in which metaphor can be found in discourse. The first approach concerns the Pragglejaz method for finding metaphorically used words, which involves the canonical case of metaphor identification in cognitive linguistics. The second approach concerns one way in which it is possible to go from words identified as metaphorically used to their related underlying conceptual structures, by means of a five-step procedure. And the third approach focuses on other linguistic forms of expression of metaphor as an underlying cross-domain mapping in conceptual structure, such as simile and analogy. All three approaches are discussed with reference to their application in empirical research on corpus data.

Keywords: metaphor identification, indirect language use, similarity, incongruity, method.

RESUMEN: Este artículo metodológico se centra en los diferentes modos en que el fenómeno de la metáfora puede aparecer en el discurso. El primer enfoque trata el método *Pragglejaz* para identificar palabras usadas metafóricamente, incluyendo el caso canónico de la identificación de metáforas en la lingüística cognitiva. El segundo enfoque describe, mediante un proceso de cinco pasos, una posible forma de acceder a la estructura conceptual subyacente a palabras cuyo uso se ha identificado como metafórico. El tercer enfoque se centra en otras formas lingüísticas de expresar la metáfora concebida como un mapeo subyacente entre dominios en la estructura conceptual, como son el símil o la analogía. Estos tres enfoques se analizan en torno a su aplicación a la investigación empírica en datos de corpus.

Palabras clave: identificación de metáforas, uso indirecto del lenguaje, semejanza, incongruencia, método.

1. Metaphor Identification in Discourse

The identification of metaphor has at least partly been regarded as a matter of finding indirect meaning by both Lakoff (1986, 1993) and Gibbs (1993, 1994). Thus, when somebody says *Sam is a gorilla*, and their utterance does not apply to a gorilla but to a human being, the word *gorilla* has been used indirectly to convey a meaning that differs from its basic, direct application. This is the case even though the contextual meaning of *gorilla* that we have to do with here has become so conventionalized that it has ended up in an advanced learners' dictionary like Macmillan's. This is a dictionary which is based on corpus research, suggesting that the metaphorical meaning may be found frequently enough for it to need description as a conventionalized meaning of the term. The use is analyzed as designating "a big man who seems stupid or violent".

Metaphor may hence be conventionalized to the degree that it becomes part of the language code, at least as this is reflected in cultural repositories such as dictionaries and grammars. Indeed, the conventional nature of linguistic metaphor has been one of the main points of cognitive linguistic research on the phenomenon, and numerous examples have been provided which show that metaphor is part and parcel of our language system and its use (e.g. Lakoff; Johnson, 1980, 1999). This is one of the interesting changes in linguistic metaphor research of the past 25 years, shifting metaphor from its time-honored position of novel and deviant language use to the conventional and the regular.

Conventionalization of metaphor does not mean that it cannot be distinguished from equally conventional non-metaphorical language. It is still possible to make a distinction between the direct and indirect application of a word, or more generally expression, in an utterance. Not many people will deny that *gorilla* has a basic sense which can be directly applied to one sort of referent, a type of ape, as opposed to a derived, metaphorical sense, which can only be indirectly applied to another sort of referent, human beings. Metaphor as indirect meaning and use also holds for other animal metaphors, like *pig* and *bitch*, and for all other metaphors that have been described under such rubrics as LOVE IS A JOURNEY, HAPPY IS UP, or BUSINESS IS WAR. This is the reason why Lakoff, Gibbs and others have adhered to a criterion of indirectness (or to the related notion of incongruity, as in, e.g., Cameron, 2003; Charteris-Black, 2004; cf. Steen, 2007).

Indirectness may be a good starting point for finding metaphor in language, but it is not sufficient. It is both too broad and too narrow. It is too broad because metaphor is also based on a salient distinction and contrast between the two semantic or conceptual domains involved in the expression, which then also needs to be bridged by some form of semantic transfer from the one domain to the other on the basis of similarity (cf. Cameron, 2003). Thus, *Sam is a gorilla* can be given a metaphorical analysis because it involves a contrast between the domain of gorillas and humans which may be bridged by constructing a similarity between the two. This is different than another form of indirectness, metonymy, where two domains may be contrasted but where the contrast is resolved by contiguity instead of similarity. Thus, in *The White House made the announcement yesterday*, there is a contrast between the domain of buildings and the

people that occupy them, causing a form of indirect meaning. But this is not resolved by metaphorical transfer, but metonymic transfer, via the contiguous relationship between houses and their occupants (for further discussion of metaphor versus metonymy, and similarity versus contiguity, see Dirven and Pörings, 2002).

The criterion of indirectness is also too narrow to capture all linguistic forms of expression of metaphor. If metaphor is defined as a conventional or less conventional mapping across two conceptual domains, as has become customary in cognitive linguistics, it is easy to show that such cross-domain mappings may also be realized by direct language use. Thus, simile and a lot of analogy employ their language in direct ways, in that the words are related to concepts which are directly connected to the intended referents in the text world. One illustration may be provided by the following line from a song by Bruce Springsteen ("I'm on fire"): *Sometimes it's like someone took a knife, baby, edgy and dull, /And cut a six-inch valley through the middle of my soul.* This is a form of a cross-domain mapping which is expressed directly when it comes to relations between words, concepts, and referents: as listeners, we do need to build a text world that contains a knife and a process of cutting in the soul. However, it is also clear that subsequent conceptual analysis has to be done to recover the intended meaning of this cross-domain mapping. Such figures do not use language indirectly but still express metaphorical mappings at a conceptual level of analysis. An inventory of these various forms of metaphor has been proposed by Goatly (1997) and their cognitive linguistic interpretation has been at the centre of attention in Conceptual Integration Theory (Fauconnier and Turner, 2002).

The identification of metaphor in language and its use is hence fraught with difficulties (Steen, 2007). In this paper I will discuss some of the issues involved, and report on some of the methodological work I have carried out in various contexts. I will begin with the development and application of the Pragglejaz procedure for finding metaphorically used words in natural discourse, called MIP, which caters for the most frequent expression of metaphor in conceptual structure by metaphorical language (Pragglejaz Group, 2007; Steen, 2002 *a*, 2005 *a*; Steen, *et al.*, in press). Then I will continue with one way in which analysts can become more precise in identifying the meaning of a metaphor *as* a conceptual cross-domain mapping (Steen, 1999, in press; Semino *et al.*, 2004). And finally I will consider some of the issues that arise when metaphors are not expressed indirectly but directly (Steen, 2007, in press). The latter two sections are two ways in which this paper goes beyond the Pragglejaz method, which explains the title of the paper.

2. The Pragglejaz Method for Finding Metaphorically Used Words

The Pragglejaz Group is an international collective of metaphor researchers who joined forces to examine whether it was possible to devise an explicit and precise method for canonical metaphor identification in discourse. Their name has been derived from the initial letters of their first names:

P   eter Crisp, Chinese University Hong Kong, China
R   ay Gibbs, University of California, Santa Cruz, USA
A   lan Cienki, VU University (Amsterdam) Netherlands
G   raham Low, University of York, UK
G   erard steen, VU University (Amsterdam) Netherlands
L   ynne Cameron, Open University (Milton Keynes), UK
E   lena Semino, Lancaster University, UK
J   oe Grady, Cultural Logic LLC (Washington DC), USA
A   lice Deignan, University of Leeds, UK
Z   oltán Kövecses, Eötvös Loránd University (Budapest), Hungary

The group has collaborated for six years and attempted to develop a tool for metaphor identification in natural discourse that is both reliable as indicated by statistical tests and valid in that it attempts to make explicit how it makes use of current empirical research in cognitive linguistics, discourse analysis, psycholinguistics, and applied linguistics. The group has published its procedure as Pragglejaz Group (2007; cf. Steen, 2002 *a*, 2005 *a*.

The Pragglejaz group started out on the basis of a preliminary theoretical conceptualization of the nature of the method (Steen, 2002 *b*). This conceptualization involved three issues, including making a choice for a particular theoretical framework, for which a combination was envisaged of the cognitive linguistic approach to metaphor with a broad view of discourse analysis. It also implied a decision about the model for metaphor within that theoretical framework, for which the Lakoff and Johnson view of metaphor as a cross-domain mapping was chosen. And a further decision had to be made about the unit of analysis to be adopted, for which the word (or more accurately, the lexical unit) in relation to concepts and referents was preferred (cf. Crisp, Heywood, *et al.*, 2002).

Having determined the theoretical framework for the methodological project, an attempt was made to formulate a procedure for metaphor identification. Tentative versions were used for application, testing, and revision, and the final version has now been reported, with a modest reliability test, in Pragglejaz Group (2007). The procedure looks like this:

1. Read the whole text or transcript to understand what it is about.
2. Decide about the boundaries of words.
3. Establish the contextual meaning of the examined word.
4. Determine the basic meaning of the word (most concrete, human-oriented and specific).
5. Decide whether the basic meaning of the word is sufficiently distinct from the contextual meaning.
6. Decide whether the contextual meaning of the word can be related to the more basic meaning by some form of similarity.

After a number of trial reliability tests over the previous years, the goal of publishing the procedure led to its independent application by six analysts to two pieces

of discourse of about 675 words each, one news text and one conversation from the British National Corpus. The reliability of the results was reasonable. About 85% of the words in the conversation and about 75% of the words in the news text were unanimously judged to be *not* metaphorical by all six analysts. Unanimous agreement between all six judges about metaphorical use was obtained for 4% of the words in the conversation, and 7% of the words in the news text. In all, then, there was unanimity about the analysis by six independently operating analysts for 89% and 82% of the cases. If criteria of success are relaxed to include cases where five independently operating analysts agreed with each other about metaphorical or non-metaphorical use, these percentages rise to 93.1 and 91.1, respectively. When these analysts get together to discuss the remaining cases of disagreement, the figures become even more positive.

An example of a stretch of discourse where there was unanimous agreement about non-metaphorical usage according to the criteria of the procedure is the following excerpt from the conversation:

A: So you deny all the studies that prove that...
B: No
A: ... conclusively?
B: And what I'm saying is that...
A: Do you deny those studies?
B: What I'm saying is that y— I probably do <unclear> deny those studies.

Each of these words is not used metaphorically in the sense defined above.

An example of a stretch of discourse where all judges agreed that most of the words were not metaphorically used, but two were, is the following (the number of positive identifications is included in brackets behind the relevant word):

What i— emerges(6) is depression(6) is a common condition which is under-diagnosed and under-treated.

When we apply the Macmillan dictionary to this excerpt, the verb *emerge* has a contextual meaning of "to become known", but a more basic meaning of "to come out of something or out from behind something". The latter is more basic because it is concrete, as opposed to the abstract meaning of the former. The two senses are distinct, as is reflected by their separate numbering in the dictionary. And they can be related by similarity: when an idea or fact becomes known to people, it is comparable to the physical emergence of a concrete entity. Similarly, *depression* in this context means "a feeling of being extremely unhappy", but its basic meaning designates "an area on a surface that is lower than the parts around it". Again, there is a contrast between the physical and the abstract, and this may be bridged by means of the mapping UNHAPPY IS DOWN.

It is interesting to note here that indirect, contextual meanings do not have to be less frequent than direct, basic meanings. For instance, the emotional sense of *depression* is listed as its first, most common sense by Macmillan. Another complication has to do with the register value of a word. Thus, the concrete basic sense of *depression* is listed

as being formal as opposed to general language use. However, the presence of both senses in the dictionary suggests that they are part of the current language system, albeit with different values, and their joint presence enables their juxtaposition as direct and indirect meanings, of which the latter is based on similarity and therefore metaphor.

An example of a stretch of discourse with less unanimity is the following:

> President Bush the elder's 'new world order(2)' led(6) to(4) the establishment, for the first time, of a Palestinian government, the Palestinian National Authority, on Palestinian soil, and the establishment of diplomatic relations between(3) Jordan and Israel.

Three words are seen as potentially metaphorical by less than five of the six judges: *to, between,* and *order.* Two of these concern prepositions, which are rather difficult when it comes to fixing their basic meanings without further theoretical discussion. Moreover, *to* is preceded by the heavily metaphorical verb *led*, which may have had an impact on its perception as metaphorically used by some analysts.

Less clear cases may be due to analytical error. But they may also be due to the complexities of metaphor and language use. They have proven to be extremely instructive for the improvement of the procedure and its theoretical shoring up. Methodological research is crucial for both theoretical as well as empirical work on metaphor, and it helps increase the reliability and validity of the findings.

The Pragglejaz procedure has been adopted in two research programs on metaphor in natural discourse at the VU University Amsterdam. The first program is called "Metaphor in discourse: Linguistic forms, conceptual structures, and cognitive representations", with four PhD researchers and myself, and runs from September 2005 through August 2010. In the first stage of this program we have analyzed four samples of 50,000 words from a publicly available sample from the British National Corpus, called BNC-Baby. The four samples involve conversation, news, fiction, and science texts.

The second program is called "Conversationalization of public discourse", has the same timing, and involves one other PhD researcher. In the first stage of the second program, two samples of in total 100,000 words were analyzed from two Dutch corpora in one coherent metaphor project. The two registers here are conversation and news. Both programs have employed the Pragglejaz procedure as part of more encompassing method for metaphor identification which we will touch upon in the rest of this paper.

The Pragglejaz procedure has turned out to provide a useful starting point for the corpus-linguistic work which we have begun to do in our two research programmes (Steen, Biernacka, *et al.*, in press). The procedure has shown to be generally applicable to large samples of British English and Dutch. We are also achieving high levels of reliability. But our practical experience has suggested one or two issues which we have had to solve in different ways than those proposed by the Pragglejaz Group.

One issue has to do with the definition of lexical units. The Pragglejaz Group have defined lexical units rather broadly. For instance, they do not make a distinction between the noun *squirrel* and the verb *squirrel* as separate lexical units. This is to be able to say that the verbal form of the word is a metaphorical manifestation of a basic sense that can be found in the noun. This can only be done if both senses relate to the same lexical unit. As a result, lexical units are defined in this broad way.

In our application of the method, we have limited the notion of lexical unit to the relevant grammatical category. We do not treat word forms as lexical units, but only consider as units those grammatical categories and subcategories which can be used to express the same type of referent in discourse, that is, grammatical word classes (verbs to indicate actions or process, nouns to express entities, and so on). This means that we cannot mark the verb *squirrel* as metaphorically used, because there is no more basic sense for that lexical unit with which the contextual verbal sense may be contrasted and compared. The same holds for the conventionalized adjectival use of the noun *key*, as in a *key variable*: this, too, cannot be marked as metaphorical in our approach.

It is important to understand what we are doing here. We are not denying that there is a metaphorical relationship between the two different manifestations of *squirrel* or *key*. What we *are* denying is that these are metaphorical relationships *in use*. To us, they are metaphorical relationships in the language system, which may be described by morphological analysis. Such morphological relationships may even have effects on language processing. However, they are not due to metaphor in use, which we (and the Pragglejaz Group) have defined as pertaining to the direct or indirect expression of a referent by a word. Since the basic meaning of the adjective *key* is simply "important", if its description in advanced language learners' dictionaries like Macmillan's is a reliable source, its referential application to an aspect of a variable is direct, not indirect, for there is no more basic meaning for *key* as an adjective than "important". Therefore, as a matter of lexical use, *key* is not metaphorically used if the lexical unit is restricted to the relevant grammatical category.

Another issue in our application of the Pragglejaz method has to do with the historical dimension of language and its role in determining what counts as the basic meaning of a word. In their definition of basic meanings, the Pragglejaz Group have listed concrete human-oriented experience in one breath with historically older meanings. Although this is a frequent combination, not all historically older meanings are also the more concrete ones. Thus, a word like *reinforce* exhibits a number of historically attested meanings (*Oxford English Dictionary*), of which two are most relevant here (Steen, Biernacka, *et al.*, in press): (1) to make a building, structure, or object stronger, and (2) to make a group of soldiers, police etc stronger by adding more people or equipment. It turns out that it is the latter, not the former, which is historically older, by almost one century. The two criteria of concrete physical meaning versus historically older meaning may hence yield different results for the analyst who needs to decide about what counts as the basic meaning of a word. In our work we have therefore emphasized a synchronic approach which privileges concrete, human-experience related meanings, without denying that the role of the history of language needs to be verified at a later stage.

There are other issues that also need to be commented on. For instance, the identification of the precise contextual meaning of a word may be rather problematic in conversations, which often become rather vague. By contrast, in science texts, the precise contextual meaning of a word is sometimes highly technical and specialized, which raises other questions. These and other issues are brought to the fore by the consistent and precise application of the Pragglejaz method to large samples of data. We are looking forward to uncovering many more of these details about metaphorically used

words in the future, which may be facilitated by doing the corpus work which we have undertaken at the VU University.

3. Beyond Pragglejaz (1): The Five Step Method

The cognitive linguistic approach to metaphor assumes that metaphors in language use like the ones we have discussed above are expressions of underlying cross-domain mappings which are part of the conceptual structure of language and discourse. The question that arises for the analyst of discourse, therefore, is how we can get from the linguistic expressions of metaphor in discourse as for instance uncovered by the Pragglejaz method to the presumed underlying conceptual structures. I have suggested that getting from the linguistic form of metaphor to its conceptual structure is a fundamental methodological problem and have proposed a five-step framework for addressing the issues involved (Steen, 1999, in press; cf. Semino, *et al.*, 2004). Finding metaphor in discourse is not just a matter of identifying metaphorically used words but also of identifying their related conceptual structures.

My attempt at an explicit procedure for the conceptual analysis of metaphor includes the following five steps:

1. Find the metaphorical focus
2. Find the metaphorical proposition
3. Find the metaphorical comparison
4. Find the metaphorical analogy
5. Find the metaphorical mapping

In this section I can only illustrate the basic mechanisms and assumptions that are at work for canonical metaphor identification. For more complex issues, see Steen (in press).

When Tennyson writes *Now sleeps the crimson petal,* it is obvious that the word *sleeps* has been used metaphorically (cf. Steen, 2002 *a*). The Pragglejaz method would say that it is not used in its basic meaning, which pertains to animate entities, but displays another meaning in this context, designating some action or state of the crimson petal which cannot be *sleep*. The indirect contextual meaning is analyzed by setting up some sort of contrast as well as similarity relation with the basic meaning. One candidate for facilitating that analysis is a cross-domain mapping between the domains or spaces of plants and animate beings. Thus, the analyst would have to find some sort of action or state for the crimson petal that corresponds with the situation where animate beings sleep. One possibility would be to say that the crimson petal is inactive.

Each of these comments serves to point to different aspects of the analytical process of deriving an underlying conceptual structure from the linguistic form of the metaphor. These aspects are now presented in more ordered and formalized fashion with reference to the five-step framework. Table 1 shows two columns, with the five steps displayed on the left, and their application to the textual materials on the right.

Table 1
Analysis of "Now sleeps the crimson petal"

Text	*Now sleeps the crimson petal*
1. Identification of metaphor-related words	Sleeps
2. Identification of propositions	P1 (SLEEP$_s$ PETAL$_t$) P2 (MOD P1 NOW$_t$) P3 (MOD PETAL$_t$ CRIMSON$_t$)
3. Identification of open comparison	SIM {∃F ∃a [F (CRIMSON PETAL)]$_t$ [SLEEP (a)]$_s$}
4. Identification of analogical structure	SIM {[BE-INACTIVE (CRIMSON PETAL)]$_t$ [SLEEP (HUMAN)]$_s$}
5. Identification of cross-domain mapping	SLEEP > BE-INACTIVE HUMAN > CRIMSON PETAL *inferences:* GOAL OF SLEEP > GOAL OF BE-INACTIVE: REST TIME OF SLEEP > TIME OF BE-INACTIVE: NIGHT

The first step concerns the identification of the metaphorically used words in the text, and I have shown how this can be done in the previous section. Even though the complete first utterance is the linguistic expression of a cross-domain mapping, or a metaphor, there is only one word that is metaphorically used, and that is *sleep*. In traditional terminology, it is the focus (Black, 1962) or vehicle (Richards, 1936) of the metaphor.

When step 1 identifies metaphorically used words, it identifies terms which express the focus, vehicle, or source domain of the metaphor. It does so by finding those words which are somehow indirect or incongruous in context (e.g., Cameron, 2003; Charteris-Black, 2004). Such words, like *sleep*, therefore form a potential threat to the coherence of the text (Steen, 2002 *b*). However, when it seems possible to integrate them into the overall discourse by some form of comparison or similarity which resolves the incongruity, the words are somehow metaphorical, or related to metaphor. Step 1 is hence explicitly based on the idea that metaphor is a form of indirect meaning that is based on correspondence or similarity.

Step 2 involves the transformation of the linguistic expressions of the text into conceptual structures in the form of a series of propositions. It makes explicit the assumption that metaphor is a matter of thought, not language. This type of conceptual structure for discourse is usually referred to as a text base, which has a linear as well as hierarchical quality (e.g. Kintsch, 1998). In order to indicate its conceptual instead of linguistic status, small capitals are used for its technical representation.

There are several formats for this structure, and discourse psychologists are rather practical about the ways in which text bases may be modeled to suit the purposes of research. In our case, we have added subscripts to the concepts related to the words to the effect that it is clear which concepts belong to the source domain versus the target domain. This preserves the linguistic analysis in step 1, which made a distinction between source domain and target domain language.

The third step transforms the single proposition with concepts from two distinct domains derived in step 2 into an open comparison between two incomplete propositions which each pertain to another conceptual domain. This can be done because we assume that there is some form of cross-domain mapping between the two conceptual domains framing the two sets of concepts distinguished in steps 1 and 2. Step 3 makes this explicit. It states that, for some activity F in the target domain and some entity a in the source domain, there is some similarity between the activity of the crimson petal on the one hand and the sleeping of some entity on the other hand. Moreover, labeling these two domains as target and source, respectively, suggests that the similarity has to be projected from the sleeping of the entity towards the activity of the petal. These assumptions lie at the basis of most metaphor analyses in the literature.

Several issues are implied by step 3. One involves the formal and conceptual separation of the two domains or spaces already involved in step 2. Another concerns the explication of the idea that was there from step 1, that we are indeed working on the assumption that there is some sort of similarity or correspondence between the two sets of concepts: hence the addition of the operator SIM. In addition, step 3 also postulates that we will need corresponding elements on both sides of the equation, to the effect that there is some activity or state needed for the petal in the target domain, and some agent for the activity of sleeping in the source domain; hence the addition of the open function and argument variables. These are natural additions if we want to align the two domains in order to reconstruct the correspondences between them. They are, moreover, minimal assumptions, in that no new conceptual elements are added to the comparison except the ones that are implied by the original proposition.

Step 4 turns the open comparison proposed by step 3 into a closed comparison which has the formal structure of an analogy (but in fact does not always need analogical interpretation). The open values indicated by F and a in step 3 have now been interpreted by the analyst. Step 4 thus makes explicit that analysts sometimes have to add new conceptual substance to the mapping between the two domains in order to make the mapping complete. This is often the crucial step of the analysis.

For this particular example, the fourth step also happens to be the least constrained of all steps. Thus, on the side of the source domain, there is one option to fill in the logically most encompassing candidate for the agent of sleeping, which would be "animal"; and there is another option to fill in the most obvious candidate from the perspective of human experience, which is "human". Since the rest of the poem also exploits personification and not animation, the example analysis has opted for the latter. However, this is just for expository purposes. If the analysis aims to capture the meaning of the text as it might function for a reader, then the first line might have to be interpreted in the broadest fashion possible, because readers do not know yet what the rest of the poem will do, and then the notion of "animal" might be preferable.

A similar story can be told for the interpretation of the open target domain value, but we will instead turn to the last step of the procedure. This step transforms the analogical structure derived in step 4 into a mapping structure between two separate domains or spaces. It explicates what has remained implicit in step 4, the precise correspondences between the separate elements in each of the conceptual domains. This does not seem to be problematic for our current example, but that is not always the case.

Step 5 can also add further correspondences which have remained in the background of the analogy until now. Implicit elements of the sleeping schema may be projected onto implicit elements of the crimson petal schema, such as the goal or function of sleeping (rest) which may be projected form source to target to infer that the petal is tired. Or the typical time of sleeping, night, may be projected from source to target to infer something about the time of the real action of the poem. These are examples of inferences which add minimal assumptions about the cross-domain mapping and, if they are accepted, enrich the information that may be derived from it for the meaning of the text.

With step 5 we have completed our sketch of the five-step method. We have moved from the identification of its linguistic form (step 1) through its propositionalization (step 2) to its transformation into an open comparison (step 3), which was then interpreted as an analogical structure (step 4) and fleshed out into a cross-domain mapping (step 5). This procedure explicates various aspects of what analysts do when they say that particular linguistic expressions in discourse are related to metaphorical mappings.

The method offers a framework for further development which may lead to similarly detailed procedures for the other four steps as the Pragglejaz method has offered for finding those metaphor foci in step 1 that are realized by indirectly used words. For instance, the analysis of propositions in step 2 involves an area of research that has received much attention in discourse psychology and linguistic forms of discourse analysis, and the variety of approaches is about as bewildering as the variety of approaches to linguistic metaphor identification which was addressed in the Pragglejaz project. Similarly, analogy, which plays a central role in steps 4 and 5, has been the subject of quite a few psychological and computational approaches which also require consideration before a suitable candidate or synthesis can be formulated. All of these aspects are on the agenda for future research.

4. Beyond Pragglejaz (2): Other Forms of Metaphor

Apart from indirect word use, there are other manifestations in discourse of metaphor defined as a mapping across two conceptual domains, such as simile, analogy, allegory, and so on. I have noted before that simile embodies a distinct linguistic form of metaphor in conceptual structure: it is not indirect language use but displays direct lexical indications that a cross-domain mapping underlies the meaning of the language. Finding metaphor in discourse does not stop at the border of finding metaphorically used words, morphemes, phrases, or constructions: if metaphor is defined as a conceptual cross-domain mapping and language usage is approached as grounded events of discourse, then there is still more metaphor to be found.

Consider the following world-famous sonnet XVIII by Shakespeare as a case in point (Steen; Gibbs, 2004):

Shall I compare thee to a summer's day?
Thou art more lovely and more temperate:
Rough winds do shake the darling buds of May,
And summer's lease hath all too short a date;
Sometime too hot the eye of heaven shines,
And often is his gold complexion dimmed;
And every fair from fair sometime declines,
By chance or nature's changing course untrimmed:
But thy eternal summer shall not fade,
Nor lose possession of that fair thou ow'st,
Nor shall death brag thou wandrest in his shade,
When in eternal lines to time thou grow'st.
So long as men can breathe or eyes can see,
So long lives this, and this gives life to thee.

The first line sets up a cross-domain mapping by evoking and contrasting two distinct mental spaces, but does not use metaphorical language to do so. The words activate the concepts of I, THEE, COMPARE, and SUMMER'S DAY, and each of these concepts has a direct role in designating their respective referents in the world of the text. The referents in the rest of the poem belong to two distinct conceptual domains or spaces of discourse. One domain or space pertains to the addressee of the sonnet, and the other to summer's days. The point, however, is that both are directly expressed as text topics in their own right. The reader is explicitly invited to set up and compare the elements of the one topic to the elements of the other. This is a cross-domain mapping in usage which does not exhibit indirect meaning as intended by Lakoff (1986, 1993) and Gibbs (1993, 1994): it is not the language that is being used indirectly, but there is one topic which is used to talk about another topic in an indirect way.

As a result, lines 2 through 8 use language that, as a rule, may also be deemed directly expressive of their subject: line 2 uses words that directly express the personal characteristics of the addressee, whereas lines 3 through 8 directly express the properties of a summer's day. The lines do not contain metaphorical language in the sense of being meaningful indirectly, as is the basis of the cognitive linguistic definition of metaphorical language. Instead, they work as non-metaphorical expressions. They are direct instructions for setting up conceptual structures in the domain of the beloved and the domain of a summer's day, respectively, and these conceptual structures require cross-domain mapping by some form of comparative inferencing in order to achieve textual coherence. If the analysts (and the reader) do not carry out these cross-domain mappings, expressions like "more lovely and more temperate" turn incomplete while "rough winds do shake the darling buds of May" lose their point. It may hence be concluded by the analyst that such metaphorical mappings are part of the intended conceptual structure of the text.

It should be noted, though, that the first seven lines also exhibit expressions that do deviate from the locally dominant semantic field. Consider *lease* and *date* in line 4 and

eye in line 5: these are indirectly meaningful when it comes to integrating them into the local discourse topic of a summer's day, which is dominant in these lines. To spell this out for *eye*, the word activates the concept EYE which does not designate a referent "eye" in the text world, for the text world does not deal with eyes but with summer's days. Instead, *eye* is indirectly meaningful; its semantic function for the complete text has to be resolved by some form of analogizing in which the sun in the sky is compared to (or, more generally, related to) the eye in the face of a person. If this does not happen, the part of the text containing *eye* becomes incoherent. Words like *eye* and *lease* and *date*, therefore, can be considered as local linguistic metaphors in the context of a more global topic, summer's day, which in turn functions as the non-metaphorical expression of the source domain that the poem stages for conceptual mapping onto the target domain.

The incidence and interaction between these various forms of metaphor in discourse is not restricted to poetry, although the intricacies of Shakespeare's text may be quite exceptional. In general, however, cross-domain mappings by means of non-metaphorical language are typical of other types of discourse as well, such as education and science (e.g. Gentner, 1982; Gentner and Jeziorski, 1993; Mayer, 1993). To give just one illustration, consider the following scientific text from the early nineteenth century, discussed by Gentner and Jeziorski (1993: 454):

1. According to established principles at the present time, we can compare with sufficient accuracy the motive power of heat to that of a waterfall. Each has a maximum that we cannot exceed, whatever may be, on the one hand, the machine which is acted upon by the water, and whatever, on the other hand, the substance acted upon by the heat.
2. The motive power of a waterfall depends on its height and on the quantity of the liquid; the motive power of heat depends also on the quantity of caloric used, and on what may be termed, on what in fact we will call, the *height of its fall*, that is to say, the difference of temperature of the bodies between the higher and lower reservoirs.

The first section presents the cross-domain mapping by combining the two domains within each of the various discourse units; the second section follows the opposite strategy, and discusses each of the domains in its own terms and orders them from source to target. Opposite orders, from target to source, may of course also be found. The problems that these factors may create for metaphor identification by the analyst, let alone for metaphor processing by the language user, have not been studied in any depth.

One fundamental question for all researchers of metaphor in discourse that is involved here is the question of the unit of metaphor. Several researchers have pointed out that this is a problem which requires more attention (e.g. Charteris-Black, 2004; Crisp, *et al.*, 2002; Goatly, 1997; Kittay, 1987; Musolff, 2004; White, 1996). It may now be clear that this is because units of metaphor can be defined at the linguistic level as well as the conceptual level, and both can happen in several ways. These are different venues to operationalizing metaphor in discourse, and they affect the nature and number of metaphors found in language.

Consider Croft and Cruse's (2004: 213) examples of what they call simile-within-metaphor:

(1) a. Bizarre, angry thoughts *flew through* my mind like *a thousand starlings*.
    b. She was standing there, her eyes *fastened to* me like *steel rivets*.
    c. Grief *tumbled out of* her like *a waterfall*.
    d. This is really *twisting* my brain like *a dishrag*.

From a linguistic point of view, each of these four sentences displays the same pattern: their main verb is used indirectly to convey some sort of action or process between two entities, and the adverbial adjuncts of comparison are used directly to specify the manner of that action or process. The latter takes place by comparing it with the way in which another entity than the one that is the topic of the discourse would typically perform the action or process that is indirectly used. The linguistic analysis would show that there are always these two parts of the cross-domain mapping, with the first part always being a metaphorically used verb. This has to be opposed to the second part, which does not display metaphorically used language but does express a cross-domain mapping. The conceptual analysis would integrate both linguistic parts within one conceptual mapping, with one source domain containing the verbal and the adverbial elements, which would have to be mapped onto the target domain.

There might hence be three ways of counting the metaphors in these data:

1. only the verbs (metaphor as indirect language use)
2. both the verbs and the adverbial adjuncts, but as combination of two distinct linguistic categories (metaphor as indirect as well as direct linguistic expressions of conceptual cross domain mappings; Croft and Cruse's category of simile-within-metaphor)
3. the concepts relating to both the verbs and the adjuncts as belonging to one conceptual structure (metaphor as cross domain mapping)

Analysts of metaphor in usage will have to explain which of these three options they follow.

In sum, metaphor does not have to be expressed by indirect language use at all. Goatly (1997) and Fauconnier and Turner (2002) are helpful sources for cognitive-linguistically inspired discourse analysts who wish to explore this area of research, but they have to take on board more general considerations of discourse analysis, for instance pertaining to the identification of units of discourse at various levels of measurement (cf. Steen, 2005 *b*). With corpora of conversations, news, fiction, and science, we aim to make a beginning with a systematic inventory of the phenomena involved, going beyond Pragglejaz in yet another way.

5. Concluding Comments

Metaphor identification in discourse may be pursued in various ways. In this paper I have sketched three.

The typical approach to metaphor identification in cognitive linguistics has focused on metaphorically used words. I have suggested that the Pragglejaz method may offer a

good tool for cognitive linguists who wish to make their results open for independent comparison, and that its application in large scale corpus work has revealed several issues that need to be addressed.

Another typical concern in cognitive linguistics with metaphor in discourse is the relation between metaphorically used words on the one hand and cross-domain mappings in conceptual structure on the other. Here I have suggested that the five-step method may offer a promising framework for methodological study and application.

A third way in which metaphor may be found in discourse has to do with less typical expressions of metaphor, by analogies and other figures. In cognitive linguistics this area has above all been addressed by Fauconnier and Turner, but their work has only begun to reveal some of the relevant aspects of the phenomena. Further theoretical and methodological work is needed here to make progress that is consistent with the other forms of metaphor identification discussed above.

In all, then, finding metaphor in discourse is an exciting and rapidly changing field of enquiry. My attention to the methodological problems that are part and parcel of this field has only one motivation: to improve the quality of our empirical research. For the question is: when we say that we have found a lot about metaphor in language, are we all talking about the same thing? If we do, we ought to be able to demonstrate this in simple reliability tests where analysts come up with the same findings after they have been given the same instruction. In my experience, this is an extremely hard but worthwhile pursuit.

Works cited

BLACK, M. (1962): *Models and Metaphors*, Ithaca, Cornell University Press.

CAMERON, L. (2003): *Metaphor in Educational Discourse*, London/New York, Continuum.

CHARTERIS-BLACK, J. (2004): *Corpus Approaches to Critical Metaphor Analysis,* London, Palgrave MacMillan.

CRISP, P.; J. HEYWOOD; G. J. STEEN (2002): "Identification and Analysis, Classification and Quantification", *Language and Literature* 11 (1): 55-69.

CROFT, W.; A. CRUSE (2004): *Cognitive linguistics*, Cambridge, Cambridge University Press.

DIRVEN, R.; R. PÖRINGS (2002): *Metaphor and Metonymy in Comparison and Contrast*. Cognitive linguistics research, Berlin, Mouton de Gruyter.

FAUCONNIER, G.; M. TURNER (2002): *The Way We Think: Conceptual Blending and the Mind's Hidden Complexities*, New York, Basic Books.

GENTNER, D. (1982): "Are Scientific Analogies Metaphors?" in MIALL, D. (ed.) (1982): *Metaphor: Problems and Perspectives*, Brighton, Harvester. 106-132.

GENTNER, D.; M. JEZIORSKI (1993): "The shift from metaphor to analogy in Western science" in ORTONY, A. (ed.) (1993): *Metaphor and Thought*, second edition, Cambridge, Cambridge University Press. 447-480.

GIBBS, R. W., JR. (1993): "Process and Products in Making Sense of Tropes" in ORTONY, A. (ed.) (1993): *Metaphor and Thought*, second edition, Cambridge, Cambridge University Press. 252-276.

— (1994): *The Poetics of Mind: Figurative Thought, Language, and Understanding,* Cambridge, Cambridge University Press.

GOATLY, A. (1997): *The Language of Metaphors*, London, Routledge.

KINTSCH, W. (1998): *Comprehension: A Paradigm for Cognition*, Cambridge, Cambridge University Press.

KITTAY, E. F. (1987): *Metaphor: Its Cognitive Force and Linguistic Structure*, Oxford and New York, Oxford University Press.

LAKOFF, G. (1986): "The meanings of literal", *Metaphor and Symbolic Activity*, 1(4): 291-296.

— (1993): "The contemporary Theory of Metaphor" in ORTONY, A. (ed.) (1993): *Metaphor and Thought*, second edition, Cambridge, Cambridge University Press. 202-251

LAKOFF, G.; M. JOHNSON (1980): *Metaphors We Live By*, Chicago, Chicago University Press.

— (1999): *Philosophy in the Flesh: The Embodied Mind and its Challenge to Western Thought*, New York, Basic Books.

MAYER, R. E. (1993): "The Instructive Metaphor: Metaphoric Aids to Students' Understanding of Science" in ORTONY, A. (ed.) (1993): *Metaphor and Thought*, second edition, Cambridge, Cambridge University Press. 561-578

MUSOLFF, A. (2004): *Metaphor and political discourse: Analogical reasoning in debates about Europe*, Houndmills, Basingstoke, Palgrave Macmillan.

PRAGGLEJAZ GROUP (2007): "MIP: A Method for Identifying Metaphorically Used Words in Discourse", *Metaphor and Symbol* 22(1): 1-39.

RICHARDS, I. A. (1936): *The Philosophy of Rhetoric*, New York, Oxford University Press.

SEMINO, E.; J. HEYWOOD; *ET AL.* (2004): "Methodological Problems in the Analysis of Metaphors in a Corpus of Conversations about Cancer", *Journal of Pragmatics*, 36(7): 1271-1294.

STEEN, G. J. (1999): "From Linguistic to Conceptual Metaphor in five Steps" in GIBBS, R. W., JR.; G. J. STEEN (eds.) (1999): *Metaphor in Cognitive Linguistics*, Amsterdam, John Benjamins. 57-77

— (2002 *a*): "Metaphor Identification: A Cognitive Approach." *Style*, 36(3): 386-407.

— (2002 *b*): "Towards a Procedure for Metaphor Identification", *Language and Literature*, 11(1): 17-33.

— (2005 *a*): "What Counts as a Metaphorically Used Word? The Pragglejaz Experience" in COULSON, S.; B. LEWANDOWSKA-TOMASZCZYK (eds.) (2005): *The Literal-nonliteral Distinction*, Frankfurt am Main, Peter Lang. 299-322

— (2005 *b*): "Basic Discourse Acts: Towards a Psychological Theory of Discourse Segmentation" in RUIZ DE MENDOZA IBANEZ, F. J.; M. S. PEÑA CERVEL (eds.) (2005): *Cognitive Linguistics: Internal Dynamics and Interdisciplinary Interaction*, Berlin/New York, Mouton de Gruyter. 283-312

— (2007): *Finding Metaphor in Grammar and Usage: A methodological Analysis of Theory and Research*, Amsterdam: John Benjamins.

— (In press): "From Linguistic Form to Conceptual Structure in five Steps: Analyzing Metaphor in Poetry" in Brône, G.; J. Vandaele (eds.) *Cognitive Poetics*, Berlin/New York, Mouton de Gruyter.

Steen, G. J.; E. Biernacka; *et al.* (In press): "Pragglejaz in Practice: Finding Metaphorically Used Words in Natural Discourse" in Low, G.; L. Cameron; A. Deignan; Z. Todd (eds.) *Metaphor in the Real World*. In preparation.

Steen, G. J.; R. W. Gibbs, jr. (2004): "Questions about Metaphor in Literature", *European Journal of English Studies,* 8: 337-354.

White, R. M. (1996): *The Structure of Metaphor*, Oxford, Blackwell.