Feature Integration in Metonymy Resolution

Katja Markert
Language Technology Group
Division of Informatics
2 Buccleuch Place
Edinburgh EH8 9LW
email: markert@cogsci.ed.ac.uk

Abstract

In the past mostly three features of metonymies have been used to automatically recognise and interpret them, namely selectional restriction violations, world knowledge or lexical knowledge and conventionality of metonymic readings. I argue that these features are all inconclusive and not sufficient for propoer metonymy resolution. In particular, I argue that the discourse context as well as other features, including morphosyntax, language idiosyncrasies and aptness considerations play a vital part as well. I give an overview over these features and argue for extensive corpus studies looking into the reliability of the individual features and their combinations.

1 Introduction

Metonymies have a variety of properties or *features* that allow to recognise and interpret them. As examples look at the metonymies in Example 1 and 2.

(1) "**The ham sandwich** is waiting to pay."

(2) "I parked **the BMW** outside."

The most common property used at the moment for automatic *metonymy recognition* are selectional restriction violations (SRVs) (see Section 2). So in Example 1 the resolution mechanism might recognise that sandwiches cannot wait and that therefore the literal meaning of "ham sandwich" is excluded. Similar constraints hold for Example 2.

The most common features used in *metonymy interpretation*, i.e. deriving the underlying referent of the metonymic expression, are selectional restrictions and world or lexical knowledge (Section 3). Sometimes the fact that metonymies are often organised in conventional patterns is used as well. So metonymies like Example 2 occur for many producers and their products ("an IBM, a Ford"). We call these conventional patterns *schemata* — the terms *metonymic concept* (Lakoff & Johnson, 1980) and *lexical rules* (Lascarides & Copestake, 1998; Briscoe & Copestake, 1999) are used as well.

The purpose of this paper is threefold: First, I will give an overview over the mentioned features and describe how they are commonly used. This will shed light on often unquestioned assumptions about these properties, e.g. the assumption that all metonymies violate selectional restrictions. I will argue that these assumptions

are sometimes too readily accepted for metonymy resolution. In particular, I will argue that the indicated three features are not sufficient for adequate metonymy resolution. I will give examples that unlike Example 1 and 2 cannot be fully explained by the above properties.

Secondly, I will introduce other features that can account for the new data. In particular, I will describe the interactions of metonymies with the discourse context (see Section 4.

Thirdly, I will summarise more properties that are relevant for metonymy resolution and argue that corpus studies on the influence and frequencies of features are now paramount as otherwise automatic metonymy resolution relies on unevaluated assumptions (Section 5).

2 Selectional restrictions

Selectional restrictions are constraints predicates impose on their arguments, e.g. "park" imposes on its object that it is a kind of vehicle. Example 2 then violates this restriction as a firm — BMW— cannot be parked.

Therefore the violation of selectional restrictions can give clues as to then a metonymic reading is present. But this assumption is often taken further. So, many researchers regard it as mandatory that metonymies violate selectional restrictions. In these approaches metonymy resolution is triggered **only if** such a violation is present (Hobbs et al., 1993; Fass, 1997; Pustejovsky, 1995; Amghar et al., 1995; Bouaud et al., 1996; Copestake & Briscoe, 1992).[1] The underlying assumption is that metonymies always violate selectional restrictions. But this is not true as the Examples 3 and 4 show. In 3, taken from the British National Corpus and uttered by a speaker in the 20th century, the context and world knowledge make it clear that the works of Shakespeare are disliked, not the person. In Example 4 the discourse context indicates that *"caterpillar"* refers anaphorically to the butterfly mentioned previously and is therefore a metonymy. In both cases there is no SRV.

(3) "I don't really like **Shakespeare**."
(Author-for-Document metonymy, BNC)

(4) " I saw *this butterfly* fall. I said to myself: Similar is my destiny. [...] Like **this caterpillar** I have crawled around in the mud."
(Caterpillar-for-Butterfly metonymy, translation from a text by Zola, quoted by (Bonhomme, 1987))

These examples show that both the discourse context and world knowledge can override the SRV hypothesis. In a previous study we showed (Markert & Hahn, 1997) that only about 42% of metonymies in a corpus of 103 metonymies from 26 German magazine texts (information technology domain) are indicated by SRVs. Although this study is of a small scale it indicates the usefulness of corpus studies for automatic metonymy resolution as they can validate or question assumptions that have often been adopted without proof. It also indicates that metonymy recognition should not rely on SRVs alone, if it wants to cover the majority of metonymies.

3 Semantic knowledge and schemata

Metonymy interpretation usually relies on world knowledge or lexical knowledge (and selectional restrictions). Thus, in Example 1 the type constraint "wait" im-

[1] (Pustejovsky, 1995) also includes some s syntactic violations as well.

poses on its subject is PERSON, and thus *sandwich* stands metonymically for a person. Knowledge of the world then tells us that sandwiches are eaten by humans and that therefore the relationship between the metonymic expression and its underlying referent is an *eaten-by* relationship. Example 2 can be analysed in a similar way. In automatic metonymy resolution these relations are normally derived from detailed knowledge representation structures explicitly encoded in knowledge bases (Hobbs et al., 1993; Fass, 1997; Markert & Hahn, 1997; Bouaud et al., 1996) or lexica that are enriched with semantic information (Copestake & Briscoe, 1992; Briscoe & Copestake, 1999; Pustejovsky, 1995).

Metonymies are open-ended (Nunberg, 1978), so that any relationship can be exploited for metonymic expressions. So, for Example 5 the possible readings include *reading a work by Shakespeare*, *reading a paper about Shakespeare* etc.

(5) "I read **Shakespeare**."

To restrict the explosion of readings, many researchers employ schemata. Typical schemata are `Author-for-Document`, `Part-for-Whole`, `Producer-for-Product` and `Container-for-Contents`. Preference for schematic readings then selects the `Author-for-Document` reading for Example 5. Some researchers treat only schematic metonymies (Hobbs et al., 1993; Fass, 1997) whereas others prefer schematic ones over non-schematic ones, but will analyse non-schematic ones (Example 1) if no other interpretations exist (Markert & Hahn, 1997). The underlying assumptions are that schematic metonymies are more frequent than others and that the conventional reading is always preferred. There is corpus evidence for a limited range of metonymies indicating that conventional readings are much more common than unconventional ones (Briscoe et al., 1990; Verspoor, 1997) so that these assumptions are justified in a neutral context. But the situational context as well as the discourse context can again override this assumption as Example 6 and 7 show. In the first example the situational context makes clear that *"Shakespeare"* does not refer to a book by Shakespeare although this would be the conventional reading. In the second example, the discourse context makes clear that *"enjoyed your book"* does not have the conventional metonymic reading *enjoyed reading your book*.

(6) (uttered while looking at T-shirts with heads of English writers) "I particularly liked **the Shakespeare**."

(7) "John, my pet goat, loves eating things. He enjoyed **your book**."
Introduced first by (Lascarides & Copestake, 1998).

4 The influence of discourse

As the above examples show, the discourse context can be essential for metonymy recognition (see Example 4) — a feature neglected by most current metonymy research. Sometimes there are no intrasentential clues for metonymy recognition, whereas the discourse context makes clear that only a metonymic reading can be intended. Thus in Example 4 only a metonymic reading makes anaphora resolution possible and establishes discourse coherence. Therefore, anaphora resolution and metonymy recognition are interdependent in these cases.

The discourse context can be essential for metonymy interpretation as well. It can override conventional interpretations in favour of interpretations that establish discourse coherence/anaphoric relationships (see Examples 6 and 7). In addition, only the discourse context might allow a full interpretation of a metonymic reading even if a conventional reading is intended. In Example 8 the conventional reading of

a `Producer-for-Product` schema still holds, but only the previous discourse allows to establish the type of the product referred to (a motor-bike).

(8) "I had one of my motor-bikes repaired. **The BMW** made strange noises."

Different types of discourse coherence are affected by metonymy resolution (e.g. referential cohesion via anaphora resolution or coherence relations), the previous or the subsequent context can have an effect on metonymy resolution and the metonymy can interact with linear text or with pictures or tables in the text. The following table gives an example for each of these categories.[2] In general, the occurrence frequencies of these interactions are not known and other distinctions might exist as well, which can only come to the surface via extensive data analysis.

Anaphora	Coherence Relations
*...the butterfly.. **this caterpillar***	*eating things ... enjoyed **your book**.*
Previous context	Subsequent context
*He put on a record by Brahms as he wanted to save **Pergolesi** for later.*	*He wanted to save **Pergolesi** for later and put on another record.*
Linear Text	Deixis and Pictures
*...the butterfly ... **This caterpillar***	*(pointing to a picture of a butterfly) **This caterpillar***

Table 1: Different interaction types between discourse and metonymies

We have gathered data on phenomena in the first column (anaphoric relations in linear text, taking only the previous context into account) (Markert & Hahn, 1997). 26 product reviews in German magazine texts contained 606 utterances, containing 103 metonymies and 291 nominal anaphors. 56% of all metonymies were anaphoric and 19.9% of all anaphors were metonymic, showing the high interaction rate between the two phenomena. The inclusion of anaphora resolution interaction into our metonymy algorithm MetoTalk led to a resolution rate of 87% for anaphoric metonymies, which was not lower for metonymies without SRVs than for the ones with violations.

5 Other features

Although the inclusion of anaphora–metonymy interactions yielded substantially better results than a model relying on selectional restrictions, world knowledge and schemata alone, it is still not comprehensive enough as there are other features that still need to be included. It is easy to see that the inclusion of other discourse interaction features (Table 1) would further improve the model. But there are also features of a completely different type that are needed to both enhance and restrict metonymy resolution.

Aptness. A model allowing all kinds of metonymies is too unconstrained. Even schemata have exceptions — they are semi-productive (Copestake & Briscoe, 1995). The features mentioned up to now cannot explain why e.g. some parts are in general more likely to stand for a whole than others. So, "screen' is a more apt metonymy for a TV than "transistor", disregarding specific discourse contexts. These facts can

[2]The Pergolesi-example was pointed out to me by Bonnie Webber.

not be explained by textual properties alone. We have argued in (Hahn & Markert, 1997) that they are due to typicality properties of world knowledge relations and have given a formalisation of this concept. Thus, "screen" is a more typical part of a TV than a transistor and it is also a clearly visible part (see also (Tversky & Hemenway, 1984) for a psychological study on how humans distinguish between good and bad parts of objects.).

Blocking. Not every schema exception can be explained via language-independent aptness constraints. Some are also due to "Blocking" (Briscoe et al., 1995). Thus, the `Animal-for-Meat` schema is applicable for "lamb" or "rabbit" (*"I eat lamb/rabbit"*), but not for "pig", as the lexeme "pork" already exists. This is language-specific as it concerns the interaction of conventional metonymy with existing vocabulary.

Morphosyntax. Metonymy also interacts with morphosyntax. So metonymy can lead to changes in determination. Whereas e.g. proper names are mostly used without a determiner when used literally (*"Shakespeare was a great author"*), they are often used with a determiner when used metonymically (*"a Shakespeare"*). Other nouns typically affected are abstract nouns and nouns describing materials (see also (Bonhomme, 1987)). Metonymy and the valency frame of the verb can interact as well (Pustejovsky, 1995). Table 2 gives a summary of the interactions of metonymy and morphosyntax. These interactions can also be used for metonymy recognition, although they might be inconclusive. So, e.g. although a producer name with a definite determiner is mostly metonymic, there are cases where it is used literally (*"Even the mighty IBM had problems." (BNC)*).

Literal	Metonymic	Examples
Zero-determiner	definite or indefinite determiner	*"**a** real Meissen"*
definite or indefinite determiner in a particular genus	Genus change	*"**der** IBM"*(German)
Only definite or zero-determiner possible	indefinite determiner	*"**an** IBM", "**a** beauty", "**a** Paper"*
only singular	plural	***"beauties", "two BMWs"***
to- or -ing firm as object	noun as object	*"began **the book**"*

Table 2: Metonymy and morphosyntax

6 Conclusions

The discussion has shown that individual features are all inconclusive and that only an integration of as many features as possible can adequately explain metonymic usage. We have shown that e.g. the inclusion of the discourse context into a metonymy resolution model improves both metonymy recognition and anaphora resolution substantially.

Many features restricting and explaining metonymic usage are not yet sufficiently explored by corpus studies. For example, the frequency of determination changes is as yet unknown as is the frequency of SRVs in larger corpora. I am currently building a corpus of metonymies consisting of metonymic examples from the British National Corpus that will build the basis for further studies for metonymy resolution. This will shed more light on feature reliability for metonymy recognition and interpretation.

References

Amghar, Tassadit, Françoise Gayrat & Bernard Levrat (1995). *Table 10 left without paying the bill!* a good reason to treat metonymy with conceptual graphs. In Ellis, Levinson, Rich & Sowa (Eds.), *Conceptual Structures: Application, Implementation and Theory*, pp. 129–143. Berlin, Germany: Springer.

Bonhomme, Marc (1987). *Linguistique de la métonymie.* Bern, Switzerland: Peter Lang.

Bouaud, Jacques, Bruno Bachimont & Pierre Zweigenbaum (1996). Processing metonymy: A domain-model heuristic graph traversal approach. In *Proc. of the 16th International Conference on Computational Linguistics; Copenhagen, Denmark, 5–9 August 1996*, Vol. 1, pp. 137–142.

Briscoe, E.J., Ann Copestake & Branimir K. Boguraev (1990). Enjoy the paper: Lexical semantics via lexicology. In *Proc.of the 13th International Conference on Computational Linguistics on the Occasion of the 25th Anniversary of COLING and the 350th Anniversary of Helsinki University; Helsinki, Finland, 20–25 August 1990*, Vol. 1, pp. 42–47.

Briscoe, Ted & Ann Copestake (1999). Lexical rules in constraint-based grammar. *Computational Linguistics*, 25(4):487–526.

Briscoe, Ted, Ann Copestake & Alex Lascarides (1995). Blocking. In Patrick Dizier & Evelyne Viegas (Eds.), *Computational Lexical Semantics.* Cambridge, UK: Cambridge University Press.

Copestake, Ann & Ted Briscoe (1992). Lexical operations in a unification-based framework. In J. Pustejovsky & S. Bergler (Eds.), *Lexical Semantics and Knowledge Representations*, Chp. 8, pp. 101–119. Berlin, Germany: Springer.

Copestake, Ann & Ted Briscoe (1995). Semi-productive polysemy and sense extension. *Journal of Semantics*, 12:15–67.

Fass, Dan (1997). *Processing Metaphor and Metonymy.* Stanford, CA: Ablex.

Hahn, Udo & Katja Markert (1997). In support of the equal rights movement for literal and figurative language: A parallel search and preferential choice model. In *Proc. of the 19th Annual Conference of the Cognitive Science Society; Palo Alto, Cal., 7–10 August 1997*, pp. 609–614.

Hobbs, Jerry R., Mark E. Stickel, Douglas E. Appelt & Paul Martin (1993). Interpretation as abduction. *Artificial Intelligence*, 63:69–142.

Lakoff, George & Mark Johnson (1980). *Metaphors We Live By.* Chicago, Ill: Chicago University Press.

Lascarides, Alex & Ann Copestake (1998). Pragmatics and word meaning. *Journal of Linguistics*, 34(2):387–414.

Markert, Katja & Udo Hahn (1997). On the interaction of metonymies and anaphora. In *Proc. of the 15th International Joint Conference on Artificial Intelligence; Nagayo, Japan, 23–29 Aug 1997*, Vol. 2, pp. 1010–1015.

Nunberg, Geoffrey (1978). *The Pragmatics of Reference*, (Ph.D. thesis). City University of New York, New York.

Pustejovsky, James (1995). *The Generative Lexicon.* MIT Press, Cambridge, Mass.

Tversky, Barbara & Kathleen Hemenway (1984). Objects, parts, and categories. *Journal of Experimental Psychology: General*, 113(2):169–191.

Verspoor, Cornelia (1997). Conventionality-governed logical metonymy. In H. Bunt, L. Kievit, R. Muskens & N. Verlinden (Eds.), *Proc. of the 2nd International Workshop on Computational Semantics*, pp. 300–312. Tilburg, The Netherlands.