*A pragmatic solution to the polysemy paradox**

INGRID LOSSIUS FALKUM

Abstract

This paper investigates the phenomenon of polysemy: a single word with two or multiple related senses (e.g. *run* in *run a half marathon, run on gasoline, run a shop*, etc.). While it is largely unproblematic from the point of view of communication, polysemy poses a range of theoretical and descriptive problems. This has been described as the *polysemy paradox* (Ravin & Leacock 2000). In this paper, I claim that the source of the paradox lies with the traditional assumption that polysemy is primarily a fact about *language*, i.e. a phenomenon requiring an explanation in terms of the workings of the language system. Against this lexical semantic view, I argue that polysemy should be seen as a natural consequence of how *communication* works; it results from our capacity to infer the meanings/thoughts that speakers intend to communicate to us, taking the linguistic evidence as crucial but not fully determining of content. In this way, polysemy is rooted in the more general phenomenon of underdetermination of a speaker's meaning by the language, whose study belongs to the realm of pragmatics. This pragmatic approach, I argue, provides a more promising way of tackling the theoretical challenges posed by polysemy: one that contributes to resolving the paradox.

1 Introduction

Natural languages exhibit polysemy. Polysemy is the case where what we perceive to be a single word has two or multiple related senses. The word *run*, for instance, takes on different meanings depending on whether we're talking about *running a half marathon, running some water, running on gasoline, running on empty, running a shop, running away from responsibilities*, and so on. We may fly *over* a hill in a plane; we may walk *over* the same hill, or live *over* it: *over* expresses a different spatial relation in each case. If I talk about having *lamb* in curry sauce for dinner, I have the meat in mind, not the whole animal, like I do when I'm telling you about the *lambs* that gather every summer at the entrance of our mountain cabin, and that my father is desperately trying to chase away. The *position* I'm

*Many thanks to Robyn Carston for valuable discussion and very helpful comments on this paper.

sitting in when writing this is different from the *position* I may take on Norwegian foreign policy, and while I'm enjoying a *hot* cup of tea I may look forward to the Thai salad that I'm planning to prepare for dinner, which I'd like to be *hot* but not *hot* in the same way as my cup of tea is. (And certainly not *hot* the way my friend thinks her new colleague is *hot*!).

That words behave in this way is a well-known fact about language. Although polysemy may sometimes be a source of confusion, or a speaker may deliberately use a polysemous lexical item to create confusion or a humorous effect (for instance in a pun) speakers and hearers are rarely troubled by it. In normal circumstances, speakers can trust their audience to quickly and reliably figure out the meaning they intend to communicate when using a linguistic item that could take on a different meaning in a different context. On most occasions speakers and hearers are not even aware of the potential polysemy of the words they are using, and only upon reflection may they come to identify some of their other possible meanings.

From a theoretical point of view, however, polysemy has shown itself to present a real challenge. First, there is disagreement as to how polysemy should be defined and (if desirable) distinguished from homonymy (which has one linguistic form, two or *n* etymologically unrelated meanings), on the one hand, and pragmatic adjustment of lexical meaning (lexical modulation, resolution of pragmatic vagueness/generality of meaning), on the other hand. Second, there is the difficult question of how polysemy is represented in our mental lexicons. Are all the different senses of a polysemous word lexically stored, are only some senses stored, or is a single representation stored and the others pragmatically derived? Are sense relations represented in the lexicon, and if so, how are they represented? Polysemy also poses a problem in semantic applications, such as translation and lexicography. For instance, how is a polysemous lexical item to be listed in a dictionary? This discrepancy, between the theoretical problems that polysemy raises, and the relative ease with which it is produced and understood in communication, has been described as the *polysemy paradox* (Ravin & Leacock 2000, Taylor 2003a).

In this paper, I propose as a solution to the polysemy paradox a pragmatic inferential account that treats the existence and development of polysemy in natural language as being dependent on our capacity to infer speakers' meanings on the basis of decoded linguistic content and contextual information. This capacity, I claim, which allows for linguistic meaning to be underdetermining of the speaker's meaning, provides the motivation for polysemy, and may shed light on how it might have developed in the earliest manifestations of human language.

2 The polysemy paradox

As mentioned above, there are in particular two types of theoretical problem that arise in connection with polysemy, concerning (1) the definition of polysemy, and (2) its lexical representation. We will look at these in sections 2.1 and 2.2.

2.1 The problem of definition

A distinction is often made between polysemy and homonymy. The different senses of a polysemous lexical item are seen as being related in some non-trivial way, while in homonymy the different meanings encoded by a single linguistic form are unrelated and the multiple encoding thus accidental. However, this distinction between unrelated and related senses is not always easy to draw. A reason for this is that it depends on the criterion used to decide whether senses are related or not. One such criterion has been etymological relatedness. On this view, the linguistic form *file* is homonymous, as the sense 'folder or box for holding loose papers' originates from the French word *fil* and the sense 'tool with roughened surface(s)' comes from the Old English word *féol*.[1] That these two senses are associated with the same word form is thus a matter of historical accident. On the same view, the two senses of *letter*, 'symbol of the alphabet' and 'written, typed or printed communication' are polysemous by sharing etymological roots. However, this way of distinguishing homonymy and polysemy is problematic if we are concerned to characterise the linguistic knowledge of speakers and hearers, as relatedness of meaning seems to be both subjective and a matter of degree. Take the word *cardinal*, which could mean 'leader of the Roman Catholic Church' or 'a songbird'. Is it polysemous or homonymous? That the two senses are historically related (the 'songbird' sense presumably derives from the former by virtue of a resemblance between the colour of the bird and that of the cardinal's cassock) should indicate that they are polysemous; however, many people may not be aware of this connection, and to them the two senses may seem very much like homonymy. A word like *right* ('morally correct', 'the right side') should also be polysemous on an account that treats shared etymological origin as a diagnostic of polysemy, however, this may not mesh well with speakers' intuitions about the (un)relatedness of these senses. Still, it is possible that people may 'feel' that these two senses of *right* are more related that the senses of *cardinal* are, and that the semantic relation between the two senses of *letter* is easy to see. A second possibility, therefore, is to replace the criterion of 'etymological relatedness' with that of 'psychological relatedness', and make the distinction between polysemy and homonymy a matter of speaker intuitions. On this approach, a word is polysemous

---

[1] *The Oxford English Dictionary*.

if native speakers judge it to give access to a set of related semantic representations. If a linguistic form is judged to encode two or more unrelated semantic representations, we would have an instance of homonymy. A combination of considerations of etymology and speaker intuitions is surely what lexicographers rely on in the making of dictionaries, but it is unclear what role, if any, such speaker intuitions should play in a lexical semantic theory of homonymy and polysemy, given that it is doubtful whether they have any bearing on the way in which individuals use and understand words (Lyons 1977: 552).

It is possible to argue that the distinction between polysemy and homonymy is of no theoretical significance. For instance, Kempson (1977: 82) treats "cases of multiple meaning (…) in terms of a number of separate lexical items for each phonological word, each lexical item having a specific semantic value corresponding to its systematic contribution to the interpretation of all the sentences in which it occurs." Cognitive linguists (e.g. Taylor 2003: 107) have criticised this "maximisation of homonymy at the expense of polysemy" for reducing polysemy to "an unmotivated, arbitrary phenomenon", and rendering the study of what they take to be cross-linguistically recurring patterns of meaning extension theoretically and descriptively inaccessible. In the same breath, however, Taylor concedes that cognitive linguistics may have tended to err in the opposite direction, placing too much emphasis on polysemy at the expense of homonymy.

As with the distinction between homonymy and polysemy, the dividing line between polysemy and lexical modulation/pragmatic vagueness is not always easy to draw.[2] Take the word *good*. It has different meanings according to its use in describing different things, e.g., *good car*, *good job*, *good student*. It may also have different meanings in describing the same thing: a *good job* could be one that's well paid, one that offers interesting tasks, one that has an inclusive social environment, one that offers special benefits to the employees (e.g. extended maternity leave, holidays, gym membership), one that gives a certain social status, and so on (see Kempson 1977: 123-132 for discussion and further examples). Does this suggest that *good* should be analysed as being polysemous between its numerous possible meanings (e.g. Pustejovsky 1995) or as encoding a single general meaning which is made more specific in context as a result of the operation of the pragmatic process of lexical modulation (e.g. Carston 2002: Ch. 5)? A number of diagnostic tests have been proposed for distinguishing between polysemy and pragmatically adjusted lexical meaning (see Ravin & Leacock (2000: 3-4) for an overview of different types of tests), none of which seems particularly satisfactory. Geeraerts (1993) shows how the different tests may yield contradictory results, and that the outcome can be easily manipulated by context.

---

[2] As concerns lexical representation, this distinction equals the traditional distinction between monosemy and polysemy.

Reflecting this range of problematic issues concerning the definition of polysemy is a considerable terminological divergence between different accounts. For instance, the term 'polysemy' is sometimes used to describe instances of lexical modulation (c.f. Sperber & Wilson's (1998: 197) analysis of the verb *open*), sometimes to describe a putative encoding of related senses (e.g. Murphy 2002), while the term 'lexical ambiguity' is sometimes used to describe 'true' instances of homonymy (this is common practice in the relevance theory framework, cf. Sperber & Wilson 1986/1995), sometimes as a cover term for both polysemy and homonymy (e.g. Quine 1960; Kempson 1977; Cruse 1986).

2.2 The problem of representation

A second set of problems concerns the questions raised by polysemy about the organisation of our linguistic knowledge. Murphy (2002: 412) writes, "polysemy creates a real problem for any theory of word meaning because it can result in radically different meanings being picked out by the same word". He divides the problem into the following two questions: (1) How do we represent the meaning of a polysemous word (e.g. *run*) so that we are able to understand its different uses; and (2) Why and how is it that word meanings get extended to have these different meanings?

The first question, concerning the *representation* of polysemous words, is, together with the definitional problem, the one that lexical semanticists, and psychologists to some degree, have been mainly concerned with. Although many accounts are not explicit about this, this issue amounts to a specification of the kind of information that is linguistically encoded by a polysemous word. In this section, we will look at some approaches to this question, while Murphy's second question about polysemy will be the subject of section 2.3.

According to so-called 'sense enumeration lexicons' (e.g. Katz 1972, Katz & Fodor 1963), the different senses of a polysemous word are stored as a list under a single lexical entry (a so-called *dictionary entry*), and linguistic context works to disambiguate between them. (1) is the suggested dictionary entry for the word *bachelor* (adapted from Katz & Fodor 1963: 190).

(1) *bachelor*

1. Human, Male, Adult, who has never married.
2. Human, Male, Young, knight serving under the standard of another knight.
3. Human, who has the first or lowest academic degree.
4. Animal, Male, Young, fur seal when without a mate during the breeding time.

A problem with an account of polysemy that postulates a different sense for each conceptual difference is that it might lead to an indefinite proliferation of senses. The lexicon will have to store infinitely many semantic distinctions for each lexical item. Moreover, it is not clear that this approach really captures the semantic relatedness of polysemous senses (compare sense 1 and 4), and it contains no distinction between polysemy and homonymy (e.g. sense 1 and sense 2 are homonymous), as they are all stored under a single dictionary entry.

So-called core meaning approaches (Allerton 1979; Caramazza & Grober 1976; Ruhl 1989) take polysemous senses to share a common core meaning. For instance, what Allerton (1979: 51) takes to be the three 'main meanings' of *a paper*, 'newspaper', 'document', and 'academic lecture', are supposed to share the core meaning 'important written or printed material for public use'. This accounts for the relation between the senses. A problem with this approach is to determine what such a core meaning amounts to in cases where the senses of a word are radically different. For instance, what is the core meaning of the various meanings of the verb *run*? If such a core meaning cannot be identified, we would have an instance of homonymy, which in the case of *run* would seem highly counterintuitive.

In the cognitive linguistics tradition (e.g. Lakoff 1987; Langacker 1988; Rice 1992; Taylor 2003b), where polysemy has been a central domain of interest, it is common to treat the different meanings of a polysemous word to be related through lexical networks or 'meaning chains', as in (2):

(2) $A \rightarrow B \rightarrow C \rightarrow D$, etc.

In (2), meaning *A* is related to meaning *B* in virtue of some shared attribute(s), or other kind of similarity. Sense relations concern, in the first instance, adjacent members of a category, while members that are only indirectly connected in the meaning chain or network may be very different in semantic content, hence the term 'family resemblance category'[3] (Taylor 2003b: 108). So, the word *paper*, which can be used to refer to a substance made of wood pulp and also to the editorial policy of a newspaper, gives access to a family resemblance category. The chain of meanings leading to these two senses may be the following: From (i) the wood pulp substance to (ii) sheets printed from that substance, to (iii) a publication printed on those sheets, to (iv) the management that produces that publication, to (v) their editorial policy (Murphy 2002: 407). While this type of approach captures sense relations, it offers no help with the indefinite proliferation issue. The cognitive linguistic approach has been criticised for committing to the 'polysemy fallacy' (Sandra 1998; Tyler & Evans 2003), whereby the postulation of a

[3] Cf. Wittgenstein (1953).

multitude of sense distinctions for a word (explicitly or implicitly attributed to the mental representations of speakers) leads to a failure to appropriately distinguish between the aspects of meaning that are part of the word meaning proper, and those that result from its interaction with the context.

The generative semanticist solution to the problem of polysemy representation is to postulate a set of generative mechanisms that operate over underspecified lexical entries to yield compositional interpretations (Pustejovsky 1995, Pustejovsky & Boguraev 1996, Asher & Lascarides 2003). For instance, the meaning of the verb *bake*, assumed to be polysemous between a change of state sense (*bake a potato*) and a creative sense (*bake a cake*) is, on this account, dependent on the semantics of the complement noun as a result of the operation of a generative mechanism (so-called co-composition) over the VP, allowing for the information carried by the complement to act on the meaning of the governing verb (e.g. *bake* in *bake a cake* has a creative meaning as a result of *cake* being marked as an artefact). A problem with this account is that it seems too constrained with regard to the generation of compositional interpretations. For instance, the sentence *Mary baked the pizza* would have to mean that she baked it creatively, while the sentence is also compatible with a change of state reading of *bake*. Furthermore, like the cognitive linguistic account, the generative theory offers no way to distinguish what is linguistic knowledge from what is general world knowledge, as it requires, in order for the generative mechanisms to operate, that the semantic representation of a word contains a lot of (arbitrary) information.

As this indicates, there is nothing close to an agreement as to how polysemy should be analysed, or indeed, as to what polysemy really *is*, other than that it (somehow) involves different senses having as their source a single word.

2.3 The problem of motivation (and a hint at its solution)

As to Murphy's second question about polysemy, that is, why and how it is that word meanings get extended to have different senses, we need to distinguish the *why* question from the *how* question. The latter has been widely studied, in particular in the cognitive linguistics tradition (e.g. Lakoff & Johnson 1980; Lakoff 1987; Brugman 1988; Dirven & Pörings 2003; Evans 2009) and in the relatively new field of lexical pragmatics (Blutner 1998, 2004; Carston 1997, 2002; Wilson 2003, Wilson & Carston 2006; 2007).[4] However, the question *why* it is that word meanings get extended, or even more fundamentally, what is it about our language

[4] In particular, the relevance theory approach (Carston 1997; 2002; Wilson 2003; Wilson & Carston 2006; 2007) aims to explain what allows us to effortlessly and largely unconsciously arrive at the contextually appropriate sense of a lexically encoded concept in the course of communication.

systems, specifically their lexical component, that makes them so susceptible to polysemy, has not received nearly as much focus in the literature.

This question will be the main concern of this paper. What makes our languages exhibit polysemy? What is the motivation for it? Is it fundamentally linguistic, communicative or cognitive, or some combination of these? Why do we rather use the same word to describe a set of different things than have a distinct word for each thing?

I will take the 'polysemy paradox' as a starting point. If polysemy is rarely a problem for communication among people, shouldn't we expect our theories to reflect and explain this fact? Despite their different theoretical perspectives, it has been common among linguists to take polysemy to be a fact about *language*; that is, as a phenomenon requiring an explanation in terms of the functioning of the language system, with little or no importance assigned to how it arises in *communication*. In this paper, I hope to show that this focus on lexical semantics is, in large part, responsible for the existence of a 'paradox' in connection with polysemy. I will defend the view that polysemy should not be seen primarily as a fact about language but first and foremost as a natural consequence of how communication works; it results from our capacity to *infer* the meanings/thoughts that speakers intend to communicate to us taking the linguistic expression employed as crucial but not fully determining evidence. This pragmatic approach, I argue, provides a natural and straightforward solution to the problems of definition and representation as described above, and so, I claim, deflates the paradox.

3 A communication perspective on polysemy

As mentioned above, I think that many existing accounts of polysemy put too much emphasis on the contribution of the linguistic code to the comprehension of polysemy, thereby ignoring, or at least drastically underplaying, the very constructive role that pragmatic inference has to play within this process. The main concern in the literature has been to try to single out the relevant semantic aspects of polysemy (e.g. 'What's in the semantics of a polysemous lexical item that distinguishes it from homonymy?', 'What sort of semantic representation does a polysemous lexical item give access to when we encounter it in a discourse?', cf. the problems of definition and representation), while any possible pragmatic contributions to the comprehension of the phenomenon are usually only taken into account if considered relevant. This, I claim, implicitly commits them to an analysis of polysemy in terms of linguistic decoding, where the hearer resorts to pragmatics only when the linguistic context does not provide him with sufficient information to disambiguate between the polysemous senses. In this section, I argue that such an analysis overlooks the intrinsically pragmatic nature of

polysemy, and that we should take this, rather than its semantic properties, as a point of departure in our search for a proper account of polysemy.

3.1 Polysemy and decoding

On the lexical semantic analysis, polysemy is accounted for primarily in terms of the workings of the linguistic system. When we encounter a polysemous word in discourse, we proceed by decoding its content (e.g., a list of senses (Katz 1972); a sense network or a meaning chain (Lakoff 1987; Rice 1992; Taylor 2003); an underspecified semantic representation forming the input to a generative mechanism (Pustejovsky 1995; Asher & Lascarides 2003)) and selecting (or 'activating'/'generating', depending on theoretical framework) the contextually appropriate sense, primarily by reference to the surrounding linguistic context. To illustrate how this might work, consider a paradigmatic case of polysemy.

(3) a. The *chicken* pecked the ground. (= 'bird')
b. We ate *chicken* in bean sauce for dinner. (= 'meat')

A case could be made for the two meanings of *chicken*, the 'bird' sense and the 'meat' sense, being conventionalised. People seem to have very clear intuitions about the distinctness of these two meanings. A further argument is that the two senses may be seen as having different syntactic properties, one being a count noun ('bird') the other being a mass noun ('meat'). Thus (4) could only be used having the 'meat' sense in mind:

(4) I love *chicken.*

The disambiguation between the meanings of *chicken* in (3a) and (3b) could proceed simply by using information (syntactic and semantic) from the immediate linguistic context. The kind of polysemy associated with this word could thus be argued to be a wholly linguistic matter, where all the hearer has to do in order to arrive at the contextually appropriate sense is to decode the linguistic expressions.

This analysis fits well with the so-called *code* model of communication. A language is a code that pairs phonetic and semantic representations of sentences. On the code model, communication proceeds by speakers encoding a semantic content into a linguistic utterance (i.e. a code) which is decoded by hearers, who, as a result, come to represent more or less the same content.

However, as soon as we consider more examples, it becomes clear that even this apparently straightforward case of conventional polysemy cannot be analysed in terms of linguistic decoding alone. Consider (5):

(5) John went to the store and came home with a *chicken*. (= 'bird' or 'meat'?)

In (5) we can neither rely on syntactic properties to indicate the intended reference of *chicken*, nor on the overall linguistic context to do this, since it is presumably compatible with both meanings. It is compatible with a situation where John comes home with a living chicken, with a dead unplucked chicken (in which case we would most likely still consider *chicken* to refer to the bird), with a frozen plucked one, like the ones you find in any Sainsbury freezer cabinet (in which case we would take *chicken* to mean 'meat'), or even with a cooked chicken which is ready to be eaten, as from a delicatessen. In order to figure out which meaning of *chicken* the speaker has in mind, the hearer has to use extralinguistic contextual information in order to form a hypothesis about the speaker's intended meaning. Thus, the polysemy of *chicken* does not correspond to a linguistic distinction between count and mass nouns (although this may well be observed in a number of cases). In the next section, I will argue that even so-called conventional polysemy, like all cases of linguistic ambiguity, involves pragmatic inference processes geared to the recovery of the speaker's meaning, and that it is only in a (small) subset of cases that the hearer can rely on the linguistic context alone to indicate the speaker's intended sense, as in (3a) and (3b) above.

3.2 Polysemy and inferential communication

The trouble with the code model of communication is the systematic gap that is observed between the meaning of a sentence and the meaning the speaker intends to convey by use of that sentence on a given occasion, a phenomenon often referred to as *linguistic underdeterminacy* (Carston 2002: 29). Since Grice (1967/1989), it has been quite uncontroversial that the decoded sentence meaning may underdetermine the *implicit* content (what is implicated) of a speaker's meaning. There is now a growing consensus that the decoded sentence meaning may also underdetermine the *explicit* content (what is asserted) of an utterance (e.g. Carston 1988, 2002; Recanati 1993, 2004; Sperber & Wilson 1986/1995; Travis 1985), by containing indexical references and ambiguous expressions, and in yet many other ways. Thus, communication is rarely - many pragmatists would say never - only a matter of encoding and decoding of a semantic content.

Grice (1957/1989) was the first to hold the view that instead of being a matter of encoding and decoding of meanings, communication is *inferential*, and involves a kind of mind-reading: A speaker provides evidence for her intention to communicate something to the hearer, and the hearer recognises this intention on the basis of the evidence provided. The relevance theory framework (Sperber & Wilson 1986/1995) draws on this insight about communication as an inferential process. The goal is to provide a psychologically realistic explanation of how the

hearer infers the speaker's meaning on the basis of the evidence provided. Sperber and Wilson's central claim is that utterances create expectations of relevance, which are precise and predictable enough to guide the hearer toward the speaker's meaning. *Relevance* is a potential property of all types of input to cognitive processes, and may be assessed in terms of the amount of effort it takes to process the input and the cognitive effects (strengthening or elimination of existing assumptions, derivation of contextual implications) the individual may derive from it. Other things being equal, the more cognitive effects an input yields to an individual and the less effort it takes to process it, the higher the degree of relevance of that input to that individual at that time. On the basis of this definition, Sperber and Wilson (1986/1995: 260) propose the following two principles: (1) "Human cognition tends to be geared to the maximisation of relevance (= *The Cognitive Principle of Relevance*); and (2) "Every act of ostensive communication communicates a presumption of its own optimal relevance" (= *The Communicative Principle of Relevance*).

Inferential communication, or *ostensive-inferential communication* in relevance theory terms, involves two layers of intention: (1) An informative intention: The intention to inform the audience of something; and (2) A communicative intention: The intention to inform the audience of one's informative intention (Wilson & Sperber 2004: 611). To communicate this way, the communicator has to use an *ostensive stimulus*, that is, a stimulus "designed to attract an audience's attention and focus it on the communicator's meaning." (ibid.). Proper understanding depends on the fulfilment of the communicative intention, that is, on the hearer recognising the speaker's informative intention. In this way, utterance interpretation involves attributing several layers of intention, hence is a form of mind-reading. The hearer's goal in communication is to find an interpretation of the speaker's meaning that meets the expectations of relevance raised by the ostensive stimulus itself. To illustrate, consider the following utterance:

(6) We *lost* our crew!

On 27 January 1967, the command module of the *Apollo* 1 spacecraft was destroyed by a fire during a test and training exercise. The three astronauts aboard were killed. When one of the control centre employees called up another NASA astronaut to tell him about the tragedy that had taken place, he uttered (6). On his own account, the astronaut at the receiving end was at first unsure whether his colleague was telling him that the crew were missing, that they had been unable to

find them and so couldn't carry out the testing, or if he was in fact telling him that they had all died. [5]

The difficulty, of course, concerned the polysemy of the verb *lose*; whether his colleague in the control centre had intended by using it 'become unable to find' or 'be deprived of (someone) through death', both of which were compatible with the overall (linguistic and extralinguistic) context.

However, when his colleague continued 'There has been a fire!', in an agitated voice, it started to dawn on the astronaut what was being communicated. Although the contextual information was still compatible with an interpretation on which his colleague was informing him that the testing had been unsuccessful due to two problems, i.e. (a) they couldn't locate the crew and (b) there had been a fire, he inferred the causal-explanatory connection between the first and the second utterance and, taking the agitation in the speaker's voice into account, understood that what his colleague had intended to communicate by uttering (6) was that his three astronaut colleagues aboard the *Apollo* 1 had all died.

On the relevance theoretic view, the astronaut's interpretation of (6) proceeds according to a comprehension heuristic (Wilson & Sperber 2004: 613), which is applied automatically to verbal input. According to this heuristic, the hearer (1) takes the decoded linguistic meaning, follows a path of least effort in mutually adjusting explicit content, contextual assumptions and contextual implications, and (2) stops when the interpretation he arrives at satisfies his expectations of relevance. The interpretation that the astronaut arrived at was more relevant than the other possible interpretation. Given the available contextual assumptions (e.g. 'There had been a fire', 'The control centre employee was agitated', etc.) it was the most accessible interpretation (the least effort demanding), and it was the one that yielded the most cognitive effects (it provided a causal connection between the two utterances, and probably made the astronaut draw a great many contextual implications concerning the consequences for the crew's families, future Apollo missions, and so on).

There is thus no doubt that the sentence meaning in (6) falls well short of encoding the proposition that was explicitly communicated, which is something along the lines of (7):

(7) *Explicit content of* (6): NASA WAS DEPRIVED, THROUGH DEATH, OF THE CREW OF THE APOLLO 1 SPACECRAFT.

---

[5] Attested example from the British documentary film *In the Shadow of the Moon* (2007) directed by David Sington.

Arriving at this explicitly expressed proposition by the sentence in (6) (including an interpretation of the polysemous verb *lose*) clearly requires a pragmatic inference process geared to the recovery of the speaker's meaning.

Consider another example: Jane comes out of a meeting at her workplace. Peter is waiting outside the meeting room to take her to lunch. Jane utters:

(8) I was in a really awkward *position* in there.

The noun *position* is multiply polysemous. Two of its possible meanings are 'a particular way in which someone or something is placed or arranged', and 'a person's particular point of view or attitude toward something'. Thus, two possible ways in which Jane could continue her utterance in (8) are given in (9) and (10):

(9) … The room was packed with people and there weren't enough chairs so I had to squeeze myself in between the wall and the executive director.

(10) … Mary had told me not to say anything about our department's views on the new employment scheme but they kept asking me for my opinion.

The contextually appropriate interpretation may well be highly accessible to Peter. For instance, he may know that before going to the meeting, Jane was worried that her boss might ask her for her opinion on the new employment scheme, and so he would infer that the interpretation compatible with (10) was the one she intended him to derive. Also in the absence of this kind of information, Peter might draw this interpretation if it seemed adequately relevant to him in the given context (e.g. if this was the interpretation that yielded the most cognitive effects, and/or was the most accessible one). What's important, however, is that he could not do this by a mere decoding of the sentence. Pragmatic inference would be required as well.

Finally, consider the verb *open* in (11) below.

(11) a. Susan *opened* the door.
b. Mary *opened* a bottle of wine.
c. John *opened* a tin of beans.
d. Peter *opened* the curtains to let the light in.
e. The major *opened* the conference.

This word expresses different concepts in each of (11a.-e.). It may also express different concepts depending on the context each sentence is uttered in. For instance, *opened the door* in (11a) may be used to describe an instance of a 'normal' opening of a door, where Susan uses the door handle to open it, but it could also mean that Susan picked its lock, smashed it open, etc. (11c) could be used to express that John opened a tin of beans with a tin opener but it could also

mean that he opened it with a knife, which would be a different kind of process. Thus, *open* can be used to express a great number of different concepts, and it would be impossible for all of them to be listed in the lexicon (Sperber & Wilson 1998). When encountering this word in a discourse, the hearer would have to construct an occasion-specific sense. Sometimes, the intended concept may be jointly indicated by the verb and its object (e.g. in the 'normal' interpretation of *opened the door*), requiring less effort on the part of the hearer by way of pragmatic inference, while in other cases inferring the speaker's intended meaning may demand more of the hearer's pragmatic abilities.

4 Polysemy in an evolutionary perspective

Having argued that the comprehension of polysemy requires a constructive pragmatic inference process (not merely disambiguation) geared to the recovery of the speaker's meaning, I would like to show that there is a deeper reason for the claim that polysemy is essentially linked to pragmatics.

As already mentioned, the inferential view sees verbal comprehension as a type of mind-reading that involves attributing a communicative, and hence informative intention to the speaker, and making inferences about the content of her informative intention on the basis of the linguistic expressions she uses and the contextual information available. Thus, in modern pragmatic theory, and relevance theory in particular (Sperber & Wilson 1986/1995), this capacity to infer speaker meanings is taken to be closely tied to and reliant on the more general ability to infer and attribute contentful mental states to others, i.e. our 'theory of mind', or 'metarepresentational capacity' (Premack & Woodruff 1978; Wimmer & Perner 1983; Leslie 1987; Baron-Cohen 1995); according to Sperber & Wilson (2002), pragmatics is, in fact, a submodule of the mental system that is responsible for our ability to attribute mental states to each other. Arguably, this theory of mind capacity is a defining characteristic of human beings, distinguishing us from other animals. The language faculty is claimed to be another such defining characteristic. Sperber (2000) claims that of these two, the former is the more fundamental, and that given the dependence of verbal understanding upon the capacity to attribute contentful mental states to others, it is reasonable to assume that this capacity preceded and provided the basis for the emergence and development of language in human evolution (See Sperber 2000 for a detailed defence of this position). This position is supported by Carston (2002: 30), who writes:

> [P]ublic language systems are intrinsically underdetermining of complete (semantically evaluable) thoughts because they evolved on the back, as it were, of an already well-developed cognitive capacity for

> forming hypotheses about the thoughts and intentions of others on the basis of their behaviour.

The theory of mind capacity is, therefore, what makes ostensive-inferential communication possible, and is what allows for, indeed, provides the foundation for the systematic underdetermination of speakers' meanings by the language. This has important implications for a pragmatic analysis of polysemy.

4.1 Coded communication: A drift toward univocality?

Suppose for a moment that language use was entirely a matter of encoding and decoding. Could we expect polysemy to arise in such a situation? Most animal communication that we know about involves just encoding and decoding, given the inability of most animals to attribute mental states to others (with the possible exception of chimpanzees; Tomasello 2008). An example is the vervet monkey alarm calls. These monkeys are observed to produce acoustically different alarm calls in response to different kinds of predators, which in turn trigger specific escape responses among their conspecifics, presumably representing adaptive strategies for coping with the hunting behaviour of the predators: leopard alarm calls make them run up into trees, eagle alarms cause them to look up, run into the bush, or both, while a snake alarm call makes them look down at the ground around them (Struhsaker 1967). Recorded alarm calls have been shown to set off the same escape responses (Seyfarth, Cheney and Marler 1980). The initial interpretations of these behaviours took them as evidence that the vervets shared a referential (semantic) system of communication by which they could refer to (types of) objects they encountered in their environment, and receivers were able to extract referentially specific information from the alarm calls produced (ibid.). It was hypothesised that vervet alarm signals could be the evolutionary precursor for, or homologues of, human words (Cheney & Seyfarth 1990). More recent discussions have suggested that there are several reasons to question this proposed analogy between animal alarm calls and human words (Hauser, Chomsky & Fitch 2002), and that the vervet vocalisations are probably more appropriately analysed as an instance of innate, reflexive behaviour. As Seyfarth and Cheney (2003: 168) write, "Listeners acquire information from signalers who do not, in any human sense, intend to provide it." Much indicates that this is also the way to analyse the escape responses, i.e. that the receiver's response is automatically triggered by the cry and does not pass through any stage of semantic processing proper.

Now imagine there was another species, the *berbet* monkey, identical to the vervet monkey in all respects but one: it has only a single alarm call. The alarm call is used to warn conspecifics against an approaching or nearby predator. Like the vervet, the berbet monkey has three classes of natural enemies: leopards, eagles and

snakes. The alarm call is used to warn against the presence of any of the above enemies. Like the vervet monkey, the berbet monkey has no theory of mind capacity. Upon hearing the alarm call, how does it know what flight plan to activate? The immediate physical context may provide the berbet monkey with clues: its seeing or hearing, or otherwise "perceiving" the presence of an enemy may trigger the appropriate escape response. The decoding of the alarm call may function as a way to put the berbet monkey in a state of alert, enabling it to mobilise a quicker behavioural response upon encountering the predator. It is clear, however, that having three separate unambiguous alarm calls, like those found among vervet monkeys, would be highly advantageous compared to having a single all-purpose one. Instead of having to look for information in the environment to know which flight plan to activate as a result of decoding the signal, there would be a direct causal link between the decoding of the alarm signal and the behavioural response, which would result in a more efficient and probably more successful flight.

Imagine also a situation where the berbet's alarm signal was a response to the presence of leopards only, causing receivers to run up into trees, a result of there being originally only that one predator in its environment. Suppose at some stage a second kind of predator, snakes, enters into the berbet's environment making it imperative for survival that the berbet find some means of extending its alarm call repertoire. Could its existing signal evolve so as to become polysemous (sometimes expressive of leopard-presence, sometimes of snake-presence), or, alternatively, be broadened so as to express a more general alarm? In the absence of any pragmatic capacity, it is very difficult to see how either of these possibilities could arise since a polysemous signal would require disambiguation and a more general signal would require pragmatic narrowing (or specification) (e.g. Wilson & Carston 2007).[6] Rather, for a creature like the berbet, whose communication is confined to encoding/decoding, the solution would have to be that the code develops, if at all, in the direction of increasing differentiation and specialisation of signals. As Seyfarth and Cheney write:

> Natural selection has favoured callers who vocalise in ways that change the behaviour of listeners and ultimately return benefit to the caller. Simultaneously, selection has favoured listeners who detect the links between specific calls and particular events, thereby extracting whatever information may be relevant to them. In many group-living species, where callers and recipients have overlapping reproductive interests, selection has favoured callers who give acoustically different

[6] Compare with typical human alarm signals: 'Watch out!', 'Be careful', 'Run!', etc., all of which are underspecified with regard to the actual dangers they are used to warn against.

> vocalisations in different circumstances, thus allowing listeners to extract more specific information.
>
> (Seyfarth & Cheney 2003: 168).

So, it seems that in species with a code-like system of communication and no theory of mind ability the progress of evolution may involve a drift toward specificity and thus univocality; signals are shaped so as to become as unambiguous as possible, allowing for more efficient responses, to the benefit of both signalers and receivers. Thus, it appears that in the absence of a theory of mind capacity in the communicators, there is no motivation, or even possibility, for polysemy to evolve; in fact, natural selection seems to favour the opposite development.

This, I will try to show in the next section, seems to be pretty much the opposite of how cultural evolution has shaped the code used in human communication, which is intrinsically underdetermining of the thoughts communicated by use of it, as a result of its being coupled with a highly sophisticated metarepresentational theory of mind capacity.

4.2 Inferential communication: A drift toward polysemy?

So the question, then is: given a theory of mind capacity, how might polysemy have developed in human language? What follows in this section is an attempt to answer this question, an inevitably speculative suggestion about how polysemy might have evolved.

Suppose, in the very early days of language, that most communication proceeded in terms of holophrases, i.e. one-unit communicative acts, signals with no internal structure (Tomasello 2008). Assuming our ancestors had already developed an early form of metarepresentational theory of mind capacity, enabling a degree of intention-reading, such one-unit communicative acts would be quite rich in content (or speaker meaning).

The following scenario is borrowed from Dennett (1991: 197). It is originally a story about how autostimulation (inventing new paths of internal communication) might enhance cognitive organisation but it can be adapted to illustrate how polysemy might have developed as well. Dennett invites us to imagine a hominid one day idly drawing two parallel lines on the floor of his cave. “[W]hen he looked at what he had done, these two lines reminded him, visually, of the parallel banks of the river that he would have to cross later in the day, and this reminded him to take along his vine rope, for getting across” (ibid). Now, adapting this scenario, imagine the hominid, let us call him Tom, had a friend sitting next to him, watching him draw the two parallel lines. We might call her Jane. Suppose there was already a word, *pak*, for riverbank in Tom and Jane’s language. After having become aware

that the two lines reminded him of the riverbank, Tom looks at Jane and utters 'Pak', intending to remind Jane that they have to cross the riverbank later that day. Jane recalls their impending expedition and correctly attributes to Tom the intention to draw her attention to this fact, and perhaps she will also draw the inference that Tom thinks it's time to start preparing for the trip. Tom then repeats his utterance of 'Pak' in a more playful manner, together with a pointing at the drawing to communicate to Jane that it was this drawing that reminded him of the riverbank and so of their upcoming venture. Jane responds 'Pak' back at Tom communicating to him that she recognises the visual similarity between the drawing and the banks of the river. Tom might then proceed by drawing a vertical line that crosses the parallel ones while uttering the word for vine, with the intention to inform Jane that they have to bring with them a vine rope to cross the banks. (Notice how this communicative exchange is inferential and metarepresentational through and through).

However, this communicative exchange may have an additional consequence. Suppose that after having crossed the riverbank, having walked for a while, Tom and Jane decide (using holoprastic communication) to make camp for the evening. Tom gathers some wood to make a fire, which he throws rather carelessly into a pile on the ground. Jane, who doesn't think this is a good way to arrange the wood, sits down by the pile and starts arranging the logs into a stack. She takes a pair of logs, places them side-by-side with a distance between them, then takes another pair of logs and places them side-by-side on top of the previous two, in the alternate direction. She looks at Tom so as to capture his attention, imitates the distance between the logs with her hands and utters 'Pak'. Having Tom's previous use of the word fresh in mind (his showing her the visual analogy between the river banks and the parallel lines drawn on the floor of the cave, thereby communicating to her information about their crossing the river banks), assuming that Tom also remembers this use of the word, Jane intends to communicate to Tom that the logs should be arranged in parallel (alternating) lines when making a fire. John, who remembers their earlier communicative exchange (that is he, he remembers his previous use of the word and is simultaneously aware of Jane's familiarity with this use), infers that this is indeed what Jane intends to communicate to him.

This may give rise to a broadened use of the word *pak*, as a term describing not only the parallel banks of a river, but one that also has a more abstract sense, 'parallel lines'. The relation between *pak* (riverbank) and *pak* (parallel lines) will, of course, be transparent to Tom and Jane but it might be opaque to new learners who encounter the meanings in distinct contexts.

A similar story might be told about the evolution of the polysemy of words for animals, e.g., *chicken*. Suppose that the animal sense of *chicken* was the first to be acquired and conventionalised. Then we might imagine that it would extend to the meat of the chicken by way of the speaker using the word for chicken while making

it clear (perhaps by pointing at the meat or using other non-verbal ostensive cues) that she intends to refer to the meat. This interpretation will be easily accessible to the hearer who, taking his concept of the live animal as his point of departure, will be using his theory of mind capacity to infer the intended referent of the word. Then, as a result of frequency of use (and possibly other factors) this new sense would stabilise and become conventional in the language. What is crucial here is that unless the hearer is predisposed in some way to infer the speaker's intention - to refer to the meat when using the word for the animal in a given situation - communication would not be successful. As argued above, this simply could not occur within a wholly coded communication system.

5 Polysemy: convenience or necessity?

So I think there are good arguments in favour of a pragmatic account of polysemy, and for the assumption that our metarepresentational theory of mind capacity (and our ability to grasp speakers' meanings which derives from this ability), which allows for linguistic meaning to be systematically underdetermining of speakers' meanings, is the root of polysemy in language. However, as it stands, this claim only explains what makes polysemy possible, and not what causes its proliferation in human verbal communication. The question, then, is whether polysemy is a necessary consequence of the sort of linguistic systems that our minds/brains develop; whether the fact that we are endowed with such powerful pragmatic abilities has resulted in the development of public codes that are intrinsically underdetermining of the thoughts/propositions we can entertain and communicate, while at the same time providing just the expressive resources needed in order to achieve the kind of sophisticated verbal communication that we engage in all the time; or, whether polysemy arises from e.g., a convention of linguistic usage, where a polysemous lexical item is chosen at the expense of a fully encoding expression by virtue of being less effort demanding and/or more communicatively efficient. Carston (2002: 29), who has raised this question in connection with the broader issue of linguistic underdeterminacy, captures the two positions in the following claims about linguistic underdeterminacy:

1. Underdeterminacy is a matter of effort-saving convenience for the speaker, where, although sentence meaning more often than not underdetermines the proposition expressed by it, a sentence that fully encodes the speaker's meaning could always be supplied (= *The convenient abbreviation view*).
2. Underdeterminacy is an essential feature of the relation between sentence meaning and speaker meaning, and that "generally, for any given

thought/proposition, there is no sentence which fully encodes it". (= *The essentialist view*).

The 'essentialist view' has two versions: On the weak version, underdeterminacy is the rule but a full encoding may be possible in a small number of cases (perhaps in the case of analytic, i.e. context-free, sentences). The strong version takes underdeterminacy to be universal and denies that a full encoding is possible in any circumstances.

Both the convenient abbreviation and the essentialist views of linguistic underdeterminacy provide a basis for polysemy. On the first view, being a matter of effort-saving convenience for the speaker could be what motivates polysemy. Instead of having to go through the laborious task of fully encoding the sense she has in mind (which she could do, if she wanted to), the speaker can rely on the hearer's pragmatic abilities enabling him to arrive at the intended interpretation. However, on the essentialist view, where, given the complexity and fine-grainedness of our thoughts a fully encoding system would not be practicable - polysemy would be a *necessity*.

The hypothesis about the pre-existence of the capacity for mental state attribution to the capacity for language, and the dependency of human linguistic communication upon this theory of mind capacity certainly points to an essentialist position on linguistic underdeterminacy, although this position doesn't follow logically from it. Carston (2002: 30-42) argues against the effability (understood as encodability) of thought (and hence of communicated propositions) on the basis of (i) the plausibility of our thoughts (understood in the individualistic, psychological sense) containing private references, which do not seem amenable to linguistic encoding (Sperber & Wilson 1986/1995); (ii) the impossibility of providing so-called eternal (i.e. context-independent) sentences corresponding to the propositions expressed by sentences containing indexical, referential, and predicate expressions. For instance, she points to Wettstein's (1979) argument that the object referred to by an indexical expression can be picked out by a number of *non-synonymous* (i.e. truth-conditionally distinct) descriptions (e.g. *She* in '*She* came to the party' can be uniquely described as 'Mary James born on November 17th, 1979 in London', 'The brown-haired woman at the party at 26, Exmouth Market, London, on September 4th, 2009', etc.). Which of the possible descriptions should we take to be a constituent of the eternal sentence? Carston further endorses Recanati's (1987) claim that there simply cannot be context-independent reference (and thus no uniquely denoting descriptions). Recanati argues that the reference of a definite description always depends on the 'domain of discourse', which is "that with respect to which the speaker presents his or her utterance as true" (Recanati 1987: 62). Carston shows that this also includes the predication function of language. E.g., an utterance of 'France is *hexagonal*' is very unlikely to have an

eternal sentence counterpart, as the approximate value of the predicate *hexagonal* seems to be ineffable. Her conclusion is that there are no eternal sentences, and hence that linguistic underdeterminacy is essential in the strong sense.

5.1 Polysemy and effability

I think the arguments given by Carston (and others) referred to above are good reason to take linguistic underdeterminacy to be essential (without any commitment to the weak or strong version, however). The data from polysemy I think, add to the evidence in favour of this position. First, polysemous lexical items are, in a sense, intrinsically underdetermining of the concepts they are used to express. By definition, a word with two or *n* senses is underdetermining of its communicated sense in any given context (except, perhaps, in cases of 'mention'; e.g. in talk of 'the polysemous word *run*'). Second, the sense that is communicated by use of a polysemous lexical item in a given context will arguably be subject to the more general problems of ineffability of reference and predication. The examples discussed by Carston (2002: 41) of communicated concepts that depend on the contextually relevant standard of precision are good candidates for polysemy:

(12) a. France is *hexagonal*.
b. This steak is *raw*.
c. The fridge is *empty*.

For this type of predicates there may be considerable variation in the standards of precision required in a given context, and, even though there are cases where e.g., *hexagonal* is used to express 'a geometric figure with six equal sides', the predicate expressions in (12a.-c.) may be used to express a number of different concepts depending on the degree of precision that the context allows for, neither of which seems possible to express in words.[7]

Now consider (13a.-c.), involving the word *newspaper*, which is standardly analysed as being polysemous between a physical object, an information type and an organisation (Pustejovsky 1995):

(13) a. Knut spilled coffee on the *newspaper*. (physical object)
b. The *newspaper* told a story about an armed robbery that took place in London. (text/information type)

---

[7] Some philosophers of language, following Lewis (1979), will maintain that 'hexagonal' has the same content throughout and what varies is the way the standard of precision parameter is fixed in the context.

c. The *newspaper* announced major personnel changes among its senior editorial team. (organisation).

First, notice how replacing *newspaper* in (13a.-c.) with putatively uniquely denoting expressions such as 'the Norwegian daily publication *Aftenposten*', or 'the publication that has its offices in Biskop Gunnerius gate 14A, Oslo, Norway', preserves its polysemy. Second, if we try to provide more specific descriptions for *newspaper* in (12a.-c.), e.g., 'the copy of *Aftenposten* lying on the coffee table in the apartment on the 4th floor, entrance to the left, in Schwachs gate 1, Oslo, Norway, bought at Narvesen at Majorstua, Oslo, Norway, at 2:11pm on the 15th of April 2009' for (12a), we seem to face another problem: Is this really the thought that is *communicated* (or *asserted*) by use of *newspaper* in (12a), that is, does the speaker really use *newspaper* as a shorthand for this content? I think the answer is clearly 'no'. The representation of the object that the speaker is entertaining and communicating by use of *newspaper* does not need to involve this detailed information but may still be precise and fine-grained enough, although not amenable to linguistic encoding, and, given that the speaker and the hearer share a set of contextual assumptions (and are able to use pragmatics), the speaker can be pretty sure that the hearer will be able to assign the correct reference on the basis of her choice of linguistic expression, although his representation of it may not be exactly the same as the speaker's.

5.2 Polysemy: convenience *and* necessity?

While on the convenient abbreviation view of underdeterminacy pragmatic inference would be merely 'a useful add-on' to our language capacity (Carston 2002: 35), the essentialist view takes pragmatic inference to provide the basis for language use, and that the complexity and fine-grainedness of the thoughts we can entertain and communicate make them impossible to fully encode. The hypothesis is that our evolutionarily pre-existing pragmatic abilities made the development of a fully encoding language system unnecessary. As argued above, I think that the burden of the evidence supports the essentialist position on linguistic underdeterminacy, on which polysemy follows as a necessary consequence. However, I also think the essentialist position on linguistic underdeterminacy is fully compatible with the assumption that polysemy often *is* a matter of effort-saving convenience for the speaker, albeit in a different sense than above. First, the very high level of polysemy in languages suggests that it is apparently easier for people to take old words and extend them to new meanings than to invent new words. This may be because, as Sperber and Wilson (1998: 198) have pointed out, the stabilisation of a new word in a language is a "slow and relatively rare process" that has to be coordinated over a large group of individuals over time. The kind of

small-scale coordination involved in communication is a less elaborate affair, and with our pragmatic abilities enabling us to form hypotheses with great ease about the speaker's meaning on the basis of her linguistic utterances, there would, in most cases, be no need for a new word to describe something that may just as well be described by using an already existing word (with an extended meaning). Second, many cases of metaphorically generated polysemy have the feel of being shorthand for concepts that could in principle be paraphrased (albeit not in a fully encoding way). Consider (14a.-e.):

(14) a. My dog always holds a ball in its *mouth*.
b. If you blow across the *mouth* of a bottle you can often get a note.
c. The village is located near the *mouth* of the river.
d. They built a fire at the *mouth* of the cave.
e. A car accident happened at the *mouth* of the tunnel.

The 'aperture' sense of *mouth* in (14b.-e.) is derived from the sense in (14a) by way of metaphorical extension. Although clearly tokenings of this general sense, each of (14b.-e.) expresses a different concept that could be paraphrased; e.g., 'the opening for filling or emptying the bottle', 'the place where the river enters the sea', etc. However, the metaphorical sense of *mouth* may be used to communicate these concepts in a considerably more succinct and efficient way.

6 Implications of the pragmatic account

My argument has been that the comprehension of polysemy is dependent on our capacity to infer speakers' meanings, and not simply on our ability to decode the content of the linguistic signals they are using. This places the study of polysemy firmly within the realm of pragmatics. I think Sperber and Wilson (1998: 197) are right in their claim that "polysemy is the outcome of a pragmatic process whereby intended senses are inferred on the basis of encoded concepts and contextual information." This doesn't rule out the possibility that some instances of polysemy may be conventional (in a language community and/or in the lexicons of individual speakers); it only means that *all* instances, conventional or not, can be traced back to the operation of a pragmatic process. Sperber and Wilson note that such inferred senses may be ephemeral, stable, shared by few or many speakers or by a whole language community. There may also be individual differences; for instance, an inference pattern may be a first-time affair for one interlocutor and a routine pattern for another. Our pragmatic ability is the reason that, despite these differences, we manage to communicate with each other.

Where does this pragmatic account take us with respect to the 'polysemy paradox' we laid out at the beginning of this paper? To take them in the reverse order: I hope my proposal for an answer to the problem of motivation has been made sufficiently clear in the previous sections, and I will say no more about it here. As to the problem of representation, the pragmatic account predicts that there can be no uniform synchronic treatment of polysemy. In some cases, the senses may be conventional (e.g. *chicken*), and thus stored in the mental lexicons of speakers, while in other cases only one sense is stored and the others pragmatically derived (e.g. *open*). Which senses are stored and which are derived is an empirical question to be investigated by the methods of psychology, although the differences in the extent to which polysemous senses are stored or pragmatically inferred, and the likelihood of individual variations as to their representation do not make such investigations an easy undertaking (but see Klein & Murphy 2001, 2002; Klepousniotou 2002, 2008, Pylkänen et al. 2006 for very worthwhile attempts). However, regardless of the 'semantic' or 'pragmatic' nature of the polysemous senses of a word, a pragmatic inference process is required for the hearer to arrive at the sense that was intended by the speaker on a given occasion.

In lexical semantic accounts (Cruse 1986; 2004, Tuggy 1993), it has been suggested that the different senses of a polysemous word could be visualised as points on a continuum, with no clear-cut boundaries between them. We may extend this picture to the distinction standardly made between polysemy and homonymy, and see them as forming a 'continuum of lexical ambiguity', ranging from homonymy at the one end to contextually adjusted senses at the other end. The extent to which a word is 'semantically' or 'pragmatically' ambiguous would depend on its position on this continuum. A consequence of this picture for a theory of representation is that there would be little difference between homonymy and conventional polysemy in terms of their psychological representation. What differentiates conventional polysemy and homonymy is that while the multiple encoding in the latter case is a matter of historical accident, conventional polysemy has started out as a pragmatic phenomenon where one sense has been derived from another, and then this sense has become lexicalised. So although this does distinguish polysemy from homonymy, it means that the closer to the homonymy end of the continuum a polysemous word is placed the more its psychological representation might resemble that of homonymy. This is also suggested by experimental data on the homonymy/polysemy distinction (e.g. Klein & Murphy, 2001, 2002) although the results are far from conclusive.

This brings us to the problem of definition. It follows from the pragmatic account that the distinction between polysemy and pragmatic modulation of meaning is vacuous. Whether polysemous senses are linguistically encoded (i.e. conventional) or pragmatically derived is a matter of degree, and open to empirical investigation. The distinction between homonymy and polysemy, however, is real, and concerns

the history of the different senses of a linguistic form; in homonymy the multiple encoding is accidental, while polysemous senses can either be traced back to the operation of a pragmatic process or are pragmatically constructed on the spot.

Although I think the pragmatic account presented in this paper provides a promising way to tackle some of the difficult theoretical issues raised by polysemy, it leaves many open questions. For instance, what factors contribute to the lexicalisation of senses? Is this only a matter of 'routinised inferences' (Sperber & Wilson 1998: 197) or are there other factors that may contribute to this outcome as well? Theories of grammaticalisation (e.g. Hopper & Traugott 2003) may shed light on this issue. Another question concerns the role of processing effort in the production and comprehension of polysemy. To what extent does polysemy represent an effort-reducing strategy for speakers and hearers in communication? The cognitive linguistics tradition (Lakoff 1987, Brugman 1988, Sweetser 1990; Taylor 2003b) has emphasised that the processes involved in polysemy seem to be cross-linguistic recurring patterns of category structure. How are we to explain such regularities on the pragmatic account?

Finally, the pragmatic account leaves open the question of where our intuitions about sense relations come from. There might not be a single answer to this question. It is possible that some of these intuitions stem from our encyclopaedic knowledge about the concepts we are using. The *chicken* type polysemy may be an instance of this. Other intuitions about sense relations may be reflective. For instance, that we are able to perceive a relation between the two uses of *mouth* in *mouth of a dog* and *mouth of a river* may be because we know that the latter is metaphorically derived from the former. In other cases, being encoded by a single linguistic form may be what leads us to think there is a relation between two senses. In the case of homonyms, there seems to be a tendency to look for a relation between the different meanings. (I'm sure I'm not the only one who has been wondering if there really *is* a semantic relation between *bank* (financial institution) and *bank* (riverside) after all). As for polysemy, translation into a different language may put this point into view: While the English word *open* is seen as exhibiting polysemy in *open the door* and *open the curtains*, and hence the two senses are taken to be semantically related, a semantic relation between the Norwegian lexicalisations *åpne* and *trekke for* in the translations *åpne døra* ('open the door') and *trekke fra gardinene* ('open the curtains') is much harder to perceive.

7 Conclusions

I have argued that polysemy only arises in a linguistic system which is embedded in a pragmatic inferential capacity, allowing us to infer speakers' meanings on the

basis of the linguistic evidence provided. Like all utterance comprehension, the comprehension of polysemy is seen as being crucially dependent on this ability. By shifting the emphasis from its purely linguistic-semantic properties to the cognitive-communicative prerequisites for its existence and development in language, I have claimed that many of the theoretical challenges posed by polysemy, giving rise to the 'polysemy paradox', can be avoided. Polysemy, then, whether of the 'semantic' or 'pragmatic' type, can always be traced back to the operation of a pragmatic process. The nature of the pragmatic processes involved in the construction of polysemy has not been addressed in this paper. Within the relevance-theoretic framework, these have been thoroughly investigated by Carston (1997; 2002), and Wilson and Carston (2006; 2007). Their contribution to polysemy will be the subject of another paper.

References

Allerton, D. J. 1979. *Essentials of Grammatical Theory: A Consensus View of Syntax and Morphology*. London: Routledge & Keagan.

Asher, N. & Lascarides, A. 2003. *Logics of Conversation*. Cambridge: Cambridge University Press.

Baron-Cohen, S. 1995. *Mindblindness: An Essay on Autism and Theory of Mind.* Cambridge, Massachusetts: MIT Press.

Blutner, R. 1998. Lexical pragmatics. *Journal of Semantics* 15(2): 115-162.

Blutner, R. 2004. Pragmatics and the lexicon. In *The Handbook of Pragmatics*, Horn, L. & Ward, G. L. (eds.). Oxford: Blackwell, 488-514.

Brugman, C. 1988. *The Story of 'over': Polysemy, Semantics, and the Structure of the Lexicon.* New York: Garland.

Caramazza, A. & Grober, E. 1976. Polysemy and the structure of the subjective lexicon. In *Semantics: Theory and Application*, Rameh, D. (ed). Washington: Georgetown University Round Table on Languages and Linguistics, 181-206.

Carston, R. 1988. Implicature, explicature and truth-theoretic semantics. In *Mental Representations: The Interface between Language and Reality*, Kempson, R. (ed.). Cambridge: Cambridge University Press, 155-181.

Carston, R. 1997. Enrichment and loosening: Complementary processes in deriving the proposition expressed? *Linguistische Berichte* 8, 103-127.

Carston, R. 2002. *Thoughts and Utterances: The Pragmatics of Explicit Communication*. Oxford: Blackwell.

Cheney, D. L. & Seyfarth, R. M. 1990. *How Monkeys See the World: Inside the Mind of Another Species*. Chicago: University of Chicago Press.

Cruse, D. A. 1986. *Lexical Semantics*. Cambridge: Cambridge University Press.

Cruse, D. A. 2004. *Meaning in Language. An Introduction to Semantics and Pragmatics*. Oxford: Oxford University Press.

Dennett, D. 1991. *Consciousness Explained*. London: Penguin Books.

Dirven, R. & Pörings, R. (eds.). 2003. *Metaphor and Metonymy in Comparison and Contrast*. Berlin: Mouton de Gruyter.

Evans, V. 2009. *How Words Mean: Lexical Concepts, Cognitive Models and Meaning Construction*. Oxford: Oxford University Press.

Geeraerts, D. 1993. Vagueness's puzzles, polysemy's vagaries. *Cognitive Linguistics* 4(3); 223-72.

Grice, H. P. 1957. Meaning. In *Studies in the Way of Words*, Grice, H. P. 1989. Harvard: Harvard University Press, 213-223.

Grice, H. P. 1967. Logic and conversation. In *Studies in the Way of Words*, Grice, H. P. 1989. Harvard: Harvard University Press, 22-40.

Hopper, P. J. & Traugott, E. C. 2003. *Grammaticalisation*. Second edition. Cambridge: Cambridge University Press.

Katz, J. J. 1972. *Semantic Theory.* New York: Harper & Row.

Katz, J. J. & Fodor, J. A. 1963. The structure of a semantic theory. *Language* 39(2): 170-210.

Kempson, R. 1977. *Semantic Theory*. Cambridge: Cambridge University Press.

Klein, D. E. & Murphy, G. L. 2001. The representation of polysemous words. *Journal of Memory and Language* 45(2): 259-282.

Klein, D. E. & Murphy, G. L. 2002. Paper has been my ruin: conceptual relations of polysemous senses. *Journal of Memory and Language* 47(4): 548-570.

Kleposniotou, E. 2002. The processing of lexical ambiguity: Homonymy and polysemy in the mental lexicon. *Brain and Language* 81(1-3): 205-223.

Klepousniotou, E., Titone, D. & Romero, C. 2008. Making senses of word senses: the comprehension of polysemy depends on sense overlap. *Learning, Memory* 34(6), 1534-1543.

Lakoff, G. 1987. *Women, Fire and Dangerous Things: What Categories Reveal about the Mind.* Chicago: University of Chicago Press.

Lakoff, G. & Johnson, M. 1980. *Metaphors We Live by*. Chicago: University of Chicago Press.

Langacker, R. W. 1988. A view of linguistic semantics. In *Topics in Cognitive Linguistics*, Rudzka-Ostyn, B. (ed.). Amsterdam: John Benjamins, 49-90.

Langacker, R. W. 1991. *Foundations of Cognitive Grammar. Volume 2: Descriptive Application.* Stanford, California: Stanford University Press.

Leslie, A. 1987. Pretense and representation: The origin of 'theory of mind'. *Psychological Review* 94, 412-426.

Lewis, D. 1979. Scorekeeping in a language game. *Journal of Philosophical Logic* 8(1), 339-359.

Lyons, J. 1977. *Semantics. Vol. 2.* Cambridge: Cambridge University Press.

Murphy. G. L. 2002. *The Big Book of Concepts.* Cambridge, Massachusetts: MIT Press.

Premack, D. & Woodruff, G. 1978. Does the chimpanzee have a theory of mind? *Behavioural and Brain Sciences* 1: 532-526.

Pustejovsky, J. 1995. *The Generative Lexicon*. Cambridge: Cambridge University Press.

Pustejovsky, J. & Boguraev, B. (eds). 1996. *Lexical Semantics: The Problem of Polysemy*. Oxford: Clarendon Press.

Pylkkänen, L., Llinás, R. & Murphy, G. L. 2006. The representation of polysemy: MEG evidence. *Journal of Cognitive Neuroscience* 18(1): 97-109.

Quine, W van O. 1960. *Word and Object*. Cambridge, Massachusetts: MIT Press.

Ravin, Y. & Leacock, C. 2000. Polysemy: An overview. In *Polysemy: Theoretical and Computational Approaches*, Ravin. Y. & Leacock, C. (eds.). Oxford: Oxford University Press, 1-29.

Recanati, F. 1987. Contextual dependence and definite descriptions. *Proceedings of the Aristotelian Society* 87: 57-73.

Recanati, F. 1993. *Direct Reference.* Oxford: Blackwell.

Recanati, F. 2004. *Literal Meaning.* Cambridge: Cambridge University Press.

Rice, S. A. 1992. Polysemy and lexical representation: The case of three English prepositions. In *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Lawrence Erlbaum Associates, 89-94.

Ruhl, C. 1989. *On Monosemy: A Study in Linguistic Semantics.* Albany: State University of New York Press.

Sandra, D. 1998. What linguists can and can't tell us about the mind: a reply to Croft. *Cognitive Linguistics* 9: 361-378.

Seyfarth, R. M. & Cheney, D. L. 2003. Signalers and receivers in animal communication. *Annual Review of Psychology* 54(1): 145-173.

Seyfarth, R. M., Cheney, D. L. & Marler, P. 1980. Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. *Science* 210(4471): 801.

Sperber, D. 2000. Metarepresentations in an evolutionary perspective. In *Metarepresentations: A Multidisciplinary Perspective*, Sperber, D. (ed). Oxford University Press, 117-137.

Sperber, D. & Wilson, D. 1986/1995. *Relevance: Communication and Cognition*. Oxford: Blackwell.

Sperber, D. & Wilson, D. 1998. The mapping between the mental and the public lexicon. In *Language and Thought*, Carruthers, P. & Boucher, J. (eds.). Cambridge: Cambridge University Press, 184-200.

Sperber, D. & Wilson, D. 2002. Pragmatics, modularity and mind-reading. *Mind & Language*17: 3-23.

Struhsaker, T. T. 1967. Auditory communication among vervet monkeys (Cercopithecus aethiops). *Social Communication Among Primates*, 281-324.

Sweetser, E. E. *From Etymology to Pragmatics. Metaphorical and Cultural Aspects of Semantic Structure.* Cambridge: Cambridge University Press.

Taylor, J. 2003a. Polysemy's paradoxes. *Language Sciences* 25(6): 637-655.

Taylor, J. R. 2003b. *Linguistic Categorization.* Oxford: Oxford University Press.

Tomasello, M. 2008. *Origins of Human Communication*. Cambridge: Cambridge University Press.

Tuggy, D. 1993. Ambiguity, polysemy, and vagueness. *Cognitive Linguistics* 4(3), 273-290.

Tyler, A. & Evans, V. 2003. *The Semantics of English Prepositions. Spatial Scenes, Embodied Meaning and Cognition*. Cambridge: Cambridge University Press.

Wettstein, H. K. 1979. Indexical reference and propositional content. *Philosophical Studies* 36: 91-100.

Wilson, D. 2003. Relevance and lexical pragmatics. *Italian Journal of Linguistics* 15: 273-292.

Wilson, D. & Carston, R. 2006. Metaphor, relevance and the 'emergent property' issue. *Mind & Language* 21(3), 404-433.

Wilson, D. & Carston, R. 2007. A unitary approach to lexical pragmatics: relevance, inference and ad hoc concepts. In *Pragmatics*, Burton-Roberts, N. (ed.). Basingstoke: Palgrave Macmillian, 230-260.

Wilson, D. & Sperber, D. 2004. Relevance theory. In *The Handbook of Pragmatics*, Horn, L. & Ward, G. L. (eds.). Oxford: Blackwell, 606-632.

Wimmer, H. & Perner, J. 1983. Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* 13(1): 103-128.

Wittgenstein, L. 1953. *Philosophical Investigations*. Oxford: Blackwell.