

Chapter 5

(Intercategorical) Polysemy in English

In Chapter 4, the notion of a mentally represented concept as the basis for lexical meaning was explored and critiqued on the basis of Shanon (1992). Shanon (1992: 30-32) explicitly uses the phenomenon of **polysemy** as an argument against the notion of representationalism. In Chapter 3, linguistic creativity was defined as essentially a trait of all human beings which involves the ‘making of new meaning’ along a continuum of, amongst other things, productivity and predictability. **Intercategorical polysemy** was postulated to be in the centre of this continuum because it is a productive formal pattern in English, but gives unpredictable and novel outcomes. In this chapter then, the empirical phenomenon of (intercategorical) polysemy will be investigated as a creative linguistic phenomenon. Two theoretical models which have been proposed to account for intercategorical polysemy will be presented and evaluated. In each case, the focus of the discussion will be on

- (a) the nature of the lexical knowledge that is proposed to form the basis for creativity,
- (b) the nature of the cognitive and linguistic processes that are meant to account for the creativity, and
- (c) whether the model can account for the online production and

interpretation of novel examples of intercategoryal polysemy.

It will be shown that these two models cannot account for the novel examples of intercategoryal polysemy, and thus cannot account for lexical creativity. There is therefore a need to explore alternative accounts on (intercategoryal) polysemy and linguistic creativity.

In Section 5.1, polysemy in general and intercategoryal polysemy in particular will be defined. Proposals that have been made for the mental representation of polysemy will be reviewed in Section 5.2. Section 5.3 will introduce intercategoryal polysemy as an empirical phenomenon, with a particular focus on some novel instances of this phenomenon. In Section 5.4, two theoretical models (the representational-derivational and the network-activation model) that have been proposed to account for intercategoryal polysemy will be reviewed and evaluated, particularly in terms of the question of whether they can account for the novel examples presented earlier.

5.1 Defining (intercategoryal) polysemy

Traditionally, polysemy refers to a lexical relation where a single linguistic form (i.e. a single phonological word form belonging to a single lexical category, i.e. word class, say either N or V) has different senses that are related to each other by means of regular shifts or extensions from the basic meaning (Allen 1986:147, De

Stadler 1989: 61-62, Taylor 1991: 99). This will be referred to as lexical polysemy from now on in this discussion. Lyons (1977: 550) states the following features of lexical polysemy in the form of criteria:

- (1) (a) There must be a clear derived sense relation between the polysemic senses of a word.
- (b) The polysemic senses of a word must be shown to be etymologically related to the same original source word.
- (c) Lexical polysemy is a sense relation within a particular syntactic category, i.e. lexical polysemy does not cut across syntactic word class boundaries.

For example, in the following expressions the word *school* belongs to the syntactic category noun, and although it has slightly different senses in each expression, all these expressions have the same etymological history.

- (2) (a) *The school is in Murray Street* ('the building').
- (b) *The boys love their school* ('the institution').
- (c) *The school will visit the old age home* ('the pupils').

- (d) *Working abroad is a hard school for anyone* ('opportunity for learning').
- (e) *The Prague School of Linguistics...* ('group of scholars propagating a particular theoretical approach').

More recently, Taylor (1991: 101-102) applied traditional semantic tests (or criteria) which were more typically used to distinguish between vagueness and ambiguity, to differentiate between monosemy and polysemy (cf. also Geeraerts (1989), Gouws (1989)). According to Taylor (1991: 101) a word is monosemous (i.e. it has only one sense) if it is vague, and it is polysemous (i.e. it has more than one sense) if it is ambiguous. These definitions and linguistic tests are, however, problematic in ways that will be discussed in the next paragraphs. For the sake of the argument, the typical traditional examples will be presented here, following Taylor (1991: 101), and will be critically evaluated after their presentation:

Test 1: An ambiguous sentence has more than one reading. It is therefore possible to assert one reading whilst denying the other. For example, in the sentence *I don't want a pig in the house*, it is possible to assert that there is a gluttonous person in the house, whilst denying that there is a farm animal in the house. *Pig* can therefore be regarded as polysemous in this sentence, referring to both a farm animal and metaphorically to a gluttonous person. In the sentence

There is a bird on the lawn, one could however not deny that there is a bird on the lawn (say a robin), whilst asserting that there is a bird on the lawn (say a sparrow). In this sentence *bird* is vague and therefore monosemous.

Test 2: Zeugma is the typical rhetorical device used to pun on the different senses of polysemous words. Cruse (1986: 73) gives the example *Arthur and his driver's licence expired on Thursday* where the verb *expire* has two distinct, but related, senses (i.e. 'died' and 'no longer valid'). If zeugma is possible, the relevant word is ambiguous and therefore polysemous.

Test 3: The phrase *to do so too* requires for its interpretation an antecedent with the same sense. In the sentence *I saw a bird on the lawn, and so did Jane*, the referent of *bird* is vague and therefore monosemous, because the sentence would be appropriate even if *Jane* and *I* saw different kinds of birds. However, in the sentence *I don't want a pig in this house, and neither does Jane* can not mean that *I* do not want a gluttonous person in the house and *Jane* does not want a farm animal in the house, as Taylor (1991: 102) says "punning aside".

Taylor (1991: 102) himself comes to the following conclusion:

Unfortunately, the results from ambiguity tests are frequently far from unambiguous themselves. The test sentences need to be constructed with great care...

There are several problems, not only with these linguistic tests themselves, but also with the idea of equating the distinction between monosemy and polysemy with the distinction between vagueness and ambiguity. The first set of problems are methodological problems. The test sentences are not only artificially constructed by semanticists to prove a point, they also rely on the semantic intuitions of the semanticist *alone*. The tests are also, in principle, contradictory, in that Test 2 in principle relies on punning, whereas in Test 3, punning has to be left aside for the test to work. The strategy of trying to clarify one distinction (that between monosemy and polysemy) with an inherently unclear and problematic distinction (that between vagueness and ambiguity) is counterproductive in the least.

The main problem, however, is a conceptual confusion in that these tests do not, in principle, test for a lexical property of a specific word, and hence speaker's knowledge of that word. They do, in fact, test for the property of a word *in a particular context*. For example, based on the examples given in these tests, the word *bird* should have the lexical property of vagueness, and it is therefore monosemous. Using Test 2, however, it can be shown that in a different context *bird* is ambiguous and therefore polysemous: *The birds on the lawn all have beautiful feathers and bikini's* (where *bird* can refer either to feathered animals or to women). Whether *bird* is vague or ambiguous in a particular context is a matter of

communicative intent by the speaker (and can be a valid and interesting issue in another context), but it does not provide the lexical semanticist with a valid basis for assigning an inherent lexical property to the word *bird*.

In addition to the linguistic tests discussed here, Geeraerts (1993) mentions a logical test and a definitional test. The logical test claims that a lexical item is polysemous if it can be both true and false for the same referent. For example, saying *He is not a pig* [animal], *but he is a pig* [disgusting person] would show that *pig* is polysemous. The definitional test claims that an item is polysemous if there is no minimally specified definition which covers the extension of the item. For example, no single definition can cover the senses of *port* as harbour, and *port* as blended sweet fortified wine. Geeraerts (1992: 230) states that “clashes between the definitions of polysemy can be found”. Dunbar (2001) comes to the conclusion that the definitional test is unreliable, but he finds some use in the logical and linguistic tests. Croft (1998: 171) claims that these tests indicate something about the constructions rather than about the lexical properties of the words, and Sandra (1998: 371) comes to the conclusion that “[l]ooking for better tests will not solve the problem, looking for another theory of meaning might”.

From this discussion it is clear that the tests that are meant to distinguish polysemy as a lexical property of a word are unreliable and unsatisfactory. I will argue in the

following sections that whether these tests are unreliable or useful is not the real issue. The really problematic issue here is equating the distinction between vagueness and ambiguity with the distinction between monosemy and polysemy. Allen (1986:148) and De Stadler (1989: 63) also identify the problem of differentiating between polysemy and both homonymy (defined as one linguistic form with more than one unrelated meaning, for example, *bank*₁ as 'financial institution' and *bank*₂ as 'the side along which a river flows'), and vagueness (i.e. multiple significances of the same sense in particular contexts), as the main issue in defining polysemy. Allen (1986: 153), Cruse (1986: 71) and De Stadler (1989: 68) come to the conclusion that the difference between homonymy, polysemy and vagueness are best seen as gradations on a continuum (cf. also Lehrer 1990: 208 and Geeraerts 1993). Tuggy (1993: 278) states that even though the traditional tests fail to give clear judgements, he sees the traditional view of the difference between vagueness and ambiguity as essentially correct, as long as the boundaries between these semantic categories are seen as flexible. More recently Brisard et al. (2001: 262) state the situation as follows:

There seems to be a growing consensus around the idea that polysemy occupies the middle position on a continuum with homonymy situated at one extreme and vagueness at the other (Tuggy 1993, Geeraerts 1993). In this model, a homonymous item displays two unrelated *meanings*, a polysemous item has one meaning with two or more *senses* [sic], and a vague item has one meaning with only one sense that can be slightly refined, yet not fundamentally altered, through semantic integration with additional contextual material. The notion of a continuum or cline implies

based on various unequal relationships. As was mentioned before, the relationships which form the basis of this classification are different in kind (momentary communicative-pragmatic vs systematic semantic-cognitive), and are therefore not reconcilable. Dunbar (2001: 9) makes a similar point in distinguishing between denotational and referential senses. Presumably the point of the concept 'semantic classes' is to assign words as consistently and unequivocally as possible to one of the three classes, or at least to a position somewhere on the continuum, with the intent that the classes (or positions on the continuum) will be handled differently in a model of the mental lexicon. But as the *bird* example shows, one word can be in both classes (both vague and polysemous) depending on the context in which it is used. I would argue therefore, that vagueness has, in fact, nothing to do with either homonymy, polysemy or monosemy. Vagueness is a distinct communicative and pragmatic issue, depending on the amount of information a speaker wishes to convey at any given time, and can never be a property of a lexical item and as such be used as a basis for the representation of lexical classes in the mental lexicon. Tuggy (1999: 363) states that only semantic/conceptual distinctions and clusterings which are *conventionalised* are "part of the mental structure of the language", thus in principle agreeing that vagueness and polysemy are unrelated issues. This study, however, disagrees with Tuggy's position that only *conventionalised* instances of polysemy are relevant to studies of lexical semantics for a very particular reason, namely that by focusing

on only conventionalised instances, language is assumed to be in a static state, which cannot make provision for conceptual and linguistic creativity. Since the main point of departure of this study is linguistic creativity, the focus here is on the unconventional, novel instances of polysemy. This does not mean that conventional polysemy is not a worthy empirical domain to study (cf. Zawada 1996), but the aim in this thesis is to extend this domain to explicitly include the notion of linguistic creativity.

Homonymy is a relationship of form between certain sets of words (like *bank*), but it is in a sense an accidental property of these two semantically unrelated words, due to the fact that there are fewer possible forms than meanings in a linguistic system (given that the set of potential new meanings is, in principle open-ended). True monosemy, where a word has only one distinct and unambiguous sense is, in fact, quite rare and possibly only occurs in highly-specialised scientific and technical terminology. Polysemy, on the other hand is, in fact, the 'normal' state of affairs in that it allows speakers to extend and adapt the limited formal lexical repertoire into multiple, and in principle, infinite meanings in various contexts and across domains.

Traditionally, the lexicographical approach has distinguished between homonymy and polysemy by giving separate lexical entries to homonymous items, and by

giving the various polysemic senses of a word as part of a single entry. Although it remains important to lexicographers to find a coherent and principled way to distinguish between homonyms and the polysemous senses of a single word, this is a technical lexicographical issue and will not be explored further in this thesis. Within the scope of this thesis, the issue of how one would want to distinguish between homonymy, polysemy and the multiple uses of a monosemous item, only seems to be an issue if one assumes that these (types of) phenomena have to be handled differently in terms of their mental representation and cognitive processing. This issue (i.e. whether these three semantic phenomena have to be handled differently in terms of their mental representation, and therefore have to be distinguished on principled theoretical grounds), will be taken up in Section 5.2. For the moment, the focus will remain on polysemy.

In recent years polysemy as a lexical or semantic relation has received much attention, in various formal approaches (cf., for example, Ruhl 1992 and Pustejovsky 1995) as well as cognitive approaches (cf., for example, Langacker 1987, Lakoff 1987, Lehrer 1990, Cuyckens and Zawada 2001, Nehrlich et al. 2003). With the advent of Cognitive Semantics, and its focus on the nature of categorisation, polysemous words have been re-interpreted as categories in which the various senses of the polysemous word have been regarded as elements in a radial network of senses (which together form the category that is represented by

the word). A typical example of such a radial network is presented in Figure 5.2. This radial network as well as the cognitive-semantic processes that motivate the extensions in the radial network, adapted from Dirven and Verspoor (1998: 32-35), serves simply as an illustration and will not be critiqued here (cf. however an extensive critique of (prepositional) network analyses in Sandra and Rice 1995).

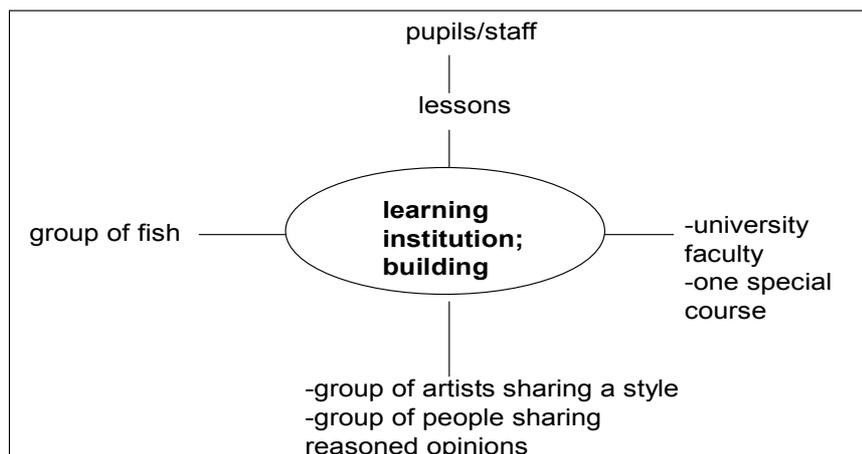


Figure 5.2 *Radial network of the senses of 'school'*

The following cognitive-semantic processes have been proposed to account for the extension of the various senses in this radial network. These processes include:

- (3) (a) Metaphorical extension: Metaphor is the cognitive process whereby a concept from a source domain is mapped to a concept in a target domain, and then named with an expression from the source

domain. For example, in the expression *the head of the department*, the person in that position is in a similar relationship to the department (the target domain), than the head of a body is to the rest of the body (the source domain) (Taylor 1991: 130-141, De Stadler 1989: 66-67). In Figure 5.2, metaphorical extension from the basic sense of *school* in the centre to the sense 'group of fish', the notion of a large group of small things together (as children in a school) is transferred to a large group of fish together.

- (b) Metonymy: Metonymy is the cognitive process whereby one concept can be named with a word referring to a concept that is conceptually related, or contiguous, to it. For example, a book can be named by its author, as in *Can you pass me the Shakespeare on the shelf?* where *Shakespeare* refers to a published book by an author named Shakespeare, and not the author itself (Taylor 1991: 121-130). A metonymic extension (based on physical contiguity) functions between the various senses of *school* that refer to 'building', 'lessons' and 'pupils' respectively in Figure 5.2.
- (c) Generalisation: Generalisation is when the original meaning of the word is made more inclusive or broader than the original meaning. For example, in the expression *my timing was wrong*, *timing* refers to the general concept of keeping time and doing things at the correct

time, rather than to a specific time such as *ten o' clock* (Dirven and Verspoor 1998: 32-35). In Figure 5.2, some of the sense extensions of *school*, i.e. 'learning institution' and 'group of people sharing opinions' is based on generalisation.

- (d) Specialisation: Specialisation is when the meaning of the word is narrowed down to a more specific meaning. For example, in the expression *to table the minutes*, the word *table* has a very specific meaning related to the procedures of meetings, rather than referring merely to the physical table around which people sit at meetings (Dirven and Verspoor 1998: 32-35). In Figure 5.2, the sense of *school* as 'learning institution' is narrowed down to 'one specific course' by means of specialisation.

Another process that is typically involved in meaning extension is chaining (which is not evident in the *school* example here). Chaining is where a polysemous category (i.e. the various senses of a polysemous word) consists of various related senses, but where the senses are not based on a unified common sense: Meaning A is related to meaning B based on some shared attribute, meaning B is related to meaning C based on some shared attribute, but meaning A and C have nothing in common (cf. Taylor's 1991: 106-109 discussion of *climb*, where a boy

climbing a tree, a train climbing a hill and a plane climbing into the sky).

The examples of *school* in (2) which are represented in the radial network in Figure 5.2 follow the traditional criterion that was specified by Lyons (1977), namely that polysemy only occurs within one and the same syntactic category or word class. However, the following examples, first introduced in Chapter 1, show that *school* which is standardly viewed as a noun can also be used as a verb, and that the meaning of *school* as a verb is still related to the notion of learning.

- (4) (a) *He schooled himself in the art of public speaking.*
 (b) *She schooled her horse for show-jumping.*

Traditionally words like *school_N* and *school_V* have been regarded as homonyms, i.e. as separate and unrelated words, merely because they belonged to different syntactic categories. This is in spite of the fact that they are clearly related both semantically and etymologically. Lehrer (1990: 208) supports the notion that these cases are part of the more general phenomenon of polysemy, by redefining **polysemy** as “all cases where there is phonological and orthographic identity of forms [and] there is a semantic relationship” (cf. also De Stadler 1989: 70).

In Cognitive Linguistics, it is assumed that *the systematic change in the meaning*

that is expressed by a word form, such as in the conventional lexical polysemy illustrated in the *school* examples in (2), should be regarded as a motivated *semantic phenomenon* that can be explained by general functional and cognitive principles (cf. for example Langacker 1999a). In the same way, *the systematic change in syntactic category* instantiated in polysemous pairs, such as *school_N* and *school_V*, should be regarded as a motivated *grammatical phenomenon* that can be explained by general functional and cognitive principles. That is, polysemous pairs (such as *school_N* and *school_V*, and *hammer_N* and *hammer_V*) display both these types of systematic change - a systematic change in meaning correlated with a systematic change in form (a change in word class), and as such, these pairs should be regarded as a kind or a subtype of polysemy (i.e. a type that has both a semantic and a grammatical aspect to it). The question arises whether the so-called traditionally lexical instances of polysemy are not, in fact, partially grammatical phenomena as well. A closer scrutiny of the examples of *school* in (2), reveals that each of the polysemous senses of *school* does, in fact, have specific syntagmatic or collocational patterns. These collocational patterns would have been viewed as selection restrictions in traditional formal approaches. For example, the sense of *school* as a building in the sentence in (2) (a), presupposes physical attributes such as size, colour, or location: *The school is large, The school is a red-brick building, The school is in Pretoria, Walk into the school and turn to your right at the steps*. Alternatively, the sense of *visit* in (2)(c) presupposes a

human agent in the subject position, and the noun *school* then takes on an interpretation of the humans that are typically associated with the school. Even more restrictive is the sense of *school* in (2) (e) (as a group of like-minded individuals) which only occurs in the following patterns: *school of N* or *N school*.

De Stadler (1989) points out that, just as there are paradigmatic lexical relations in the lexicon (such as synonymy, antonymy and hyponymy), there are syntagmatic lexical relations in the lexicon (such as the traditional selection restrictions, collocations and semantic roles). De Stadler exemplifies his claim with reference to such relationships in Afrikaans as *blonde – hare* ('blond – hair'), *tong – lek* ('lick – tongue'), *hond – blaf* ('dog – bark) and *veearts – diere* ('vet – animals'). De Stadler (1989: 67) defines the notion of syntagmatic lexical relations as follows:

... binne so 'n betrekking roep die leksikale item die konseptuele raamwerk op wat die domein vorm waarbinne die betekenis van die ander leksikale item verstaan moet word.

“... within such a relation the lexical item activates the conceptual framework that forms the domain in which the meaning of the other lexical item must be understood.” [my translation - BEZ]

The fact that these syntagmatic co-occurrence patterns exist, gives credence to the assumption that all instances of polysemy, even the traditionally purely lexical examples, have syntactic and grammatical consequences and should not be viewed as exclusively lexical or semantic phenomena at all.

The approach by Lehrer (1990) in which the focus in polysemy studies is on the semantic relationship between various linguistic elements, rather than on individual lexical items, has become virtually standardised in the cognitive-linguistic approach (Zawada (1996) and Cuyckens and Zawada (2001: xiv-xv)). In fact, in the cognitive-linguistic perspective the notion of polysemy has been broadened to include not only lexical items in the major word categories (as was discussed and illustrated in the previous section and is referred to as lexical polysemy), but to include, amongst others, grammatical functional elements such as prepositions (in various publications starting from the first case study of *over* reported in Lakoff (1987), including polysemous locative phenomena in other languages such as the German preposition *über* in Meex (2001), and the investigation of Southern Ndebele UPPER SPACE in Fleisch (2005), as well as grammatical morphemes (such as case markers) in Smith (2001) and grammatical constructions (such as the possessive) in Taylor (1991). In accordance with the view that the lexicon and the grammar are, in fact, related aspects of linguistic competence forming a continuum, polysemy has also been extended to include systemic polysemy (such as the noun class prefix system in the Bantu languages of Southern Africa) (Hendrikse 1996a and Hendrikse 2001). Following the continuum of the lexicon–grammar interface proposed in Hendrikse (1995), the continuum in Figure 5.3 can therefore be proposed for polysemy, which can be redefined as any systematic semantic relationship between any group of linguistic elements (ranging from the

been to re-introduce meaning into discussions of structure and form. For example, in Cognitive Grammar grammatical constructions are seen as the pairing of a specification of form with a specification of meaning (Taylor 1991: 198). It is, however, equally important that Cognitive Semantics includes issues related to form into their investigation of meaning. The danger of a purely and exclusively meaning-focused Cognitive Semantics is that important aspects of the form–meaning relationship are, yet again, lost. Finally, polysemy in general, and specifically intercategoryal polysemy, is certain to have implications for the nature of the mental representation and the cognitive processing of lexical items, particularly in as far as these mental representations form the basis for conceptual creativity. The issue of the mental representation of polysemy will be briefly discussed in Section 5.2.

5.2 The mental representation of polysemy

In Section 4.1.10 it was shown that one of the contributions shared by most approaches to lexical meaning is that the model of presenting lexical information in dictionaries in the lexicographical approach has almost become the standardised or orthodox model for the mental lexicon. In these views, concepts are viewed as well-defined, abstract, symbolic, determinate and static entities which form part of a unitary mental lexicon. As Geeraerts (1993: 259) puts it “meanings [are regarded] as things, prepackaged chunks of information that are contained in and carried

about by word bags". (cf. also Hendrikse 1996b on the container metaphor).

In Cuyckens and Zawada (2001: xvii) it is acknowledged that Cognitive Linguistics (as a typical psycholinguistic approach to linguistic meaning) is "meant to be a theory on the linguistic knowledge of the speaker, and about how this knowledge is embodied in the mind/brain of the speaker...". It is also stated that any claims by linguists regarding mental representation should be tested and supported by psycholinguistic experiments and other methods (cf. Cuyckens and Zawada 2001: xviii for a list of the types of evidence that can contribute to the notion of converging evidence in Cognitive Linguistics).

In an important debate regarding mental representation, and particularly the mental representation of polysemy between Sandra and Rice (1995), Croft (1998), Sandra (1998) and Tuggy (1999), the debate centres for the most part on two issues:

- The first issue is a methodological issue, namely whether linguists (as opposed to psycholinguists) are in a position to make claims about the mental representation of polysemy, particularly in view of the fact that (some) linguists tend to base these claims on introspection. Introspection (by the investigating linguist) is distinguished from intuition (judgements by a group of mother-tongue informants in a controlled environment).
- The second issue is a representational issue, namely a discussion of the

various models of polysemy that have been proposed. For example, the (dis)advantages of a network model (as discussed in Sandra and Rice (1995)) as opposed to the schematic continuum model (as proposed by Tuggy (1999)).

The methodological issue and the debate surrounding it is valid and has led to an awareness amongst Cognitive Semanticist that to make valid cognitive claims, more than introspection is necessary. Both methodological rigour and a broad empirical base are some of the basic requirements for a sound Cognitive Linguistics (cf. also Langacker (1999b) on the notion of 'converging evidence').

As far as the second issue is concerned, it is interesting to note that studies on the mental representation of polysemy often take as a basic point of departure the assumption that the mental lexicon is a fixed, canonical module with fixed, determinate entries (cf. Clarke and Clarke 1979, Sandra and Rice 1995, Murphy 1997, Croft 1998, Sandra 1998, Tuggy 1999, Brisard et al. 2001). In these studies, the issue to be investigated is whether the polysemic senses are stored separately from each other in the lexicon, and are derived from each other, or whether they are stored under one basic or schematic concept in one entry. As such, these kinds of studies assume the orthodox view of the mental lexicon as a static, fixed dictionary-like module with the various lexical items and their senses as fixed and

static entries (as in the representational-derivational model in Section 5.4.1).

Interesting other psycholinguistic experiments are those by Gibbs (1997), Gibbs and Matlock (2001) and Beitel et al. (2001) which do not test for the mental representation of polysemy as such, but rather for the role of embodied experience in a speaker's understanding and conceptualisation of the various polysemic senses of a word. In Chapter 4 it was shown that the traditional semiotic triangle (with a word, a concept and a real-world referent) is inadequate in various ways. Amongst other things it was shown that both our bodily experience of the world, as well as the representation of concepts in the embodied brain should be accounted for.

As was pointed out in Chapter 4, Classical dualism (separating the body and brain from an entity called 'mind') forms the basis of the psycholinguistic approaches, and largely informs the model of the mental lexicon as a separate mental module with fixed and abstract concepts. The extent to which the *mental* representation and processing of concepts, or of linguistic meaning in general, should, in fact be regarded as a *neural* representation and processing of concepts (i.e. the representational format), is not only an interesting issue, but is one which should be addressed in studies which focus on this issue (cf. Lamb 1998). This thesis does not focus on the issue of representational format, but instead seeks to

contribute to the debate on representational content in the following way: By questioning theoretical models that assume a mental lexicon as an intermediate mental module and concepts as fixed mental entities, this study will show, that these models are inadequate from a linguistic perspective that takes linguistic creativity into account.

Before I continue with the presentation and evaluation of some theoretical models that have been proposed for intercategoryal polysemy, I would, therefore, like to consider a sample of intercategoryal polysemy data that exemplifies linguistic creativity.

5.3 Creative intercategoryal polysemy

This study is not an empirical study as such. I will, however, use the extensive empirical studies that have been carried out in this domain (particularly Clark and Clark 1979, Dirven 1988, Lehrer 1990 and Zawada 1996) as a data base for the evaluation of the theoretical models in the following sections. It is important to note that the studies quoted above include a large number of conventionalised examples of intercategoryal polysemy (Zawada 1996 is based on a dictionary corpus which includes frequency measures), as well as novel and innovative examples (particularly Clark and Clark 1979). As Clark and Clark (1979: 769) put it, “by surveying both we will get a more complete picture of the process of

innovation”.

In this section, an additional set of specific examples of intercategoryal polysemy are presented in Table 5.1, which range from the more conventional in 1 to the more novel in 10 and 11. Even though this is a small set of data, it will suffice for its purposes, namely to test each one of the two theoretical models presented in Section 5.4 as well as the one in Chapter 6. The examples were taken from English radio cricket commentary on Radio 2000 in South Africa during the 1999 Cricket World Cup played in England, and are presented in Table 5.1. In each of the examples in Table 5.1, the polysemous word is underlined and categorised. Additional examples or related examples (not necessarily always in cricket commentary) are also given in some cases to illustrate complementary issues that will be relevant in the critical assessment of current approaches to intercategoryal polysemy. Since cricket commentary may not be known to all, relevant circumlocutions, collocations and an explanation of the context, as well as, in some cases additional comments and explanations are given. In the following section two theoretical models that have been proposed to account for the type of data in Table 5.1 will be presented and tested against this data set.

Table 5.1 *Examples of creative intercategoryal polysemy*

No.	Example	Circumlocution	Collocations, context and explanation
1	He was <u>dismissed</u> for 10 runs.	He lost his wicket and therefore cannot continue	

		to bat after having scored 10 runs.	
2	<i>He <u>dismissed</u>_V my ideas.</i>	My ideas were not taken seriously.	
3	<i>He was <u>dismissed</u>_V from his job.</i>	He was asked to leave his employment.	
4	<i>His <u>dismissal</u>_N was a big blow for Australia.</i>	The fact that he lost his wicket and could not continue to score was to the Australian team's disadvantage.	
5	<i>The journalist was <u>dismissive</u>_A of the coach's views on the pitch.</i>	The journalist did not take the coach's views on the pitch seriously.	
6	<i>The batsmen are <u>dismissed</u>_A.</i>	He lost his wicket and therefore cannot continue to bat.	
7	<i>The <u>dismissed</u>_A batsmen changed into their shorts.</i>	The batsmen who have finished batting changed into shorts.	
8	<i>The <u>dismissed</u>_N and the injured are lining up to shake hands.</i>	The batsmen who have finished batting or were injured come back onto the field and line up to shake hands with their opponents.	
9	<i>Kallis <u>middled</u>_V that one.</i>	Kallis hit the ball in/with the <u>middle</u> _N of the bat.	<i>middled</i> here means that the ball is hit by a batsman, in this case called Kallis, by using the middle of the bat and resulting in an effective shot, i.e. <i>middling the ball</i> in cricket is positive. Note that the expression is completely vague as to where the ball went and how many runs ('points' in cricket) were scored. The focus is entirely on the point of impact between bat and ball. The tone of the commentator gave a sense of satisfaction with this shot, even though the implication is there that some of the previous shots (as opposed

			to 'this one') had not been middled, i.e. not hit as cleanly. To experienced cricket supporters <i>middling</i> the ball will usually be accompanied by a very characteristic thud of the ball on the bat. Related terms in cricket describing the actions of the batsmen are, for example, <i>edged</i> , <i>nicked</i> , <i>sweep</i> , <i>drive</i> etc. all of which are examples of intercategory polysemy.
10	<i>He skied_v it!</i>	Klusener <i>skied_v</i> the ball. / Klusener hit the ball into the <i>sky_N</i> .	The ball is hit high into the sky by a batsman, <i>he</i> , in this case referring to Klusener. The tone of the (South African) commentator when uttering the original expression was of high excitement because the ball was hit high and far over the boundary of the field of play, and Klusener was awarded a <i>six</i> , i.e. 6 runs (the highest possible score for a shot in cricket). The exclamation in its original context was therefore a very positive one. When presented with the sentence <i>Klusener skied the ball</i> in a neutral written context, it was interesting to note that experienced cricket supporters find the expression <i>skied</i> to have negative connotations. Unless the ball is hit high and far, over the boundary of the field of play, as in this case, it is indeed not a positive thing to hit the ball into the air. In cricket, a ball that is hit into the air can easily be caught by the opposition, which means that the batsmen are dismissed.
11	<i>The umpire <i>wided_v</i> that delivery.</i>	The umpire decided to call that delivery a <i>wide_N</i> .	A ball (or delivery) that is bowled <i>wide_A</i> of the narrow strip of grass between the wickets is called a <i>wide_N</i> , and the decision by the umpire to call a <i>wide_N</i> is named as an action and therefore expressed as a verb. The commentator is questioning the decision by the umpire to call a wide. It is interesting to note that the agent of the action of bowling a ball in a way that makes it a <i>wide_A</i> ball is clearly the bowler. In this instance, the agency of making a delivery into a wide is transferred from the bowler onto the umpire; it is almost as if the commentator is holding the umpire responsible for the fact that the ball was bowled wide, whereas the actual action, and therefore the responsibility would normally fall on the bowler. In the context of a one-day match as this was the case, such a decision can cost a team a match. This was the first and only instance of this usage I have ever heard, whereas <i>middled</i> is fairly common, and I have heard <i>skied</i> since, mostly with a

			negative connotation.
--	--	--	-----------------------

In the context of this study, and the proposed continuum of linguistic creativity, *skied the ball* in 10 can be regarded as slightly less creative and more conventionalised than *wided that delivery*. For example, a Google search for *skied the ball* returned between 500 and 600 records, all from a sporting context, whereas *wided the delivery* (as well as *wided the ball*) returned no records. Cricket and soccer were about equally represented in the returns for *skied the ball*, with a few records from golf and baseball. The reason for including this example in the test sample, particularly as an example of creativity, is because 'skying the ball' in cricket would normally have a negative connotation, whereas this particular example has a very positive connotation. The context, as well as the tone (of voice) of the commentator, play an important role both in the creation of a different, positive meaning, as well as in the interpretation, and will need to be taken into account in the testing of the theoretical models.

5.4 Theoretical models of (intercategorical) polysemy

In this section, two of the theoretical models that have been proposed to account for intercategorical polysemy are presented. It is important to note that neither one of these two models subsumes any particular linguistic theory. These models are abstracted and idealised (in a theoretical sense) from various theoretical approaches to and claims that have been made about (intercategorical) polysemy and the mental representation and cognitive processing of linguistic knowledge. The focus in the following discussion will, therefore be on general principles, rather than on specific theoretical mechanisms. The critical criterion that will be applied to each model, will be whether the type of model can, in principle, account for the linguistic creativity in the data set. In particular, the focus will be on whether these kinds of models give an account of

- the lexical knowledge of speakers and its mental representation in a way that allows for that knowledge to form the **basis** of linguistic creativity, and
- the linguistic and cognitive **processes** involved in creating these examples of intercategorical polysemy in a way that allows for both their production and their interpretation.

It must be pointed out that both the theoretical models discussed in the next section are regarded as so-called psychologicistic (cognitive or mentalist) models of the knowledge and skills of speakers and hearers. In other words, these theoretical

models make claims about the cognitive processes and knowledge structures that play a role in (intercategorical) polysemy, and ultimately may play a role in mental representation and linguistic creativity.

Zawada (1996: 104-107) discusses three theoretical models of the mental lexicon which have been proposed to account for intercategorical polysemy. These three models are reduced to two models in the following sections. The first theoretical model in Zawada (1996) to account for N-V polysemy was the traditional **morphological** view in which the mental lexicon is viewed as a container with mini-container-like lexical entries (fully specified for form, meaning, syntactic category, subcategorisation frames and selection restrictions, and the word-formation rules that apply to it). According to the second model in Zawada (1996), referred to as the 'weak' cognitive view there, the lexical subsystem of the language competence of a speaker is divided into semantic fields and domains in which the lexical entries (which may contain encyclopaedic knowledge) are related to each other by means of cognitive principles (such as metaphor and metonymy) rather than formal WFRs. This is, in effect, a semantic version of the representational-derivational model, since the only difference between this model and the traditional morphological model is that semantic-cognitive principles are added to the basic morphological model, or that semantic-cognitive or pragmatic principles replace the word-formation rules. Both the morphological and the

semantic version of this model will therefore be dealt with under the representational-derivational model in Section 5.4.1.

The third model in Zawada (1996) is called the 'strong' cognitive model but will be referred to here as the **network-activation model**, and will be discussed in Section 5.4.2. In the following sections, the representational-derivational model and the network-activation model will not only be presented, but will be tested against the sample of creative intercategoryal polysemy in Table 5.1 in Section 5.3. It will be shown that neither of these models can adequately account for creative intercategoryal polysemy due to some principled shortcomings.

5.4.1 The representational-derivational model

As was pointed out earlier, in the traditional view of polysemy, intercategoryal polysemy was not regarded as an instance of polysemy. Intercategoryal polysemy was traditionally referred to as conversion, categorial alternation or functional shift and was regarded as the result of a morphological word-formation process of zero-derivation. The various terms that were traditionally used for this phenomenon stress slightly different aspects of the phenomenon. For example, conversion and categorial alternation focus on the idea that there is a change in the syntactic category, and there is an implicit acknowledgement of a semantic relationship. These two terms also imply that neither one of the senses (and associated

syntactic categories) is more basic than the other. Functional shift, on the other hand focuses not only on the change in syntactic category, but more particularly on the change as a 'shift' from one more basic sense (and category) to another. The 'functional' in 'functional shift' can either refer to semantic and/or to syntactic function. Intercategorial polysemy is referred to as *meerfunksionaliteit* in Afrikaans ('multiple functionality') by Kempen (1969) and De Stadler (1989: 70), and even though it appears to be closer to 'functional shift' as a term, it does not necessarily carry the implications of one sense being more basic than the other. Both Kempen's and De Stadler's discussion of the Afrikaans data does, however, support the notion that one sense in a particular pair is more basic than the other.

Even though the morphological process of zero derivation as a means of accounting for these polysemous pairs has been the standard view of this phenomenon, it was never really regarded as a satisfactory solution to this problem (cf. also Bauer 1983: 32-33). As the following discussion will show, this model has serious shortcomings. Take, for example, the various senses of *round* representing various word classes:

- (5) (a) *The ball is round_A / The round_A ball*
 (b) *To round_V the corner too fast is dangerous.*
 (c) *He enjoyed the round_N of golf on Saturday.*

Firstly, in a zero-derivation account, one sense of *round* will have to be regarded as the basic one so from which the other senses are derived. There is, however, no systematic and justifiable way to decide which sense is more basic than the others. However, to follow the traditional argument, let us accept a pre-theoretical notion of basicness that can be used intuitively to argue that *round* as an adjective is the most basic sense since *round* is typically an attribute of a physical entity. *Round* the verb would then have to be derived from *round* the adjective by means of a zero morpheme, zero₁. It would be necessary to postulate a zero₁ as opposed to a zero₂, since deriving a verb from an adjective, and deriving a noun from the adjective cannot be done by the same morpheme and/or the same WFR. Zero₁ as a morpheme would have to have a meaning along the lines of 'carrying out an action that follows /mimics the typical shape described by the A'. *Round* the noun would also be derived from *round* the adjective by means of a zero morpheme, zero₂. Zero₂, would have a meaning such as the following: 'an abstract entity in which someone follows a path from a starting point, along a set of intermediate points (which may, however, not necessarily be in the form of a perfect circle) and ends up at the same starting point'. Only for this one set of examples, we have to postulate two different zero morphemes each with its own meaning, as well as two different WFRs (cf. (6) (b) and (c)). For example:

- (6) (a) *round*_A (the basic sense, hence no derivation)
 (b) A + zero₁ ‘ V
 (c) A + zero₂ ‘ N

Bauer (1988: 31-32) explains this implausible situation as follows:

Let us assume that we can work out which *round* is the basic form of all these (this is feasible, but not immediately obvious). That basic type of *round* presumably has no zero morph on the end of it. The others, though, would have a zero morph. So a zero morph is contrasting with nothing at all, a lack of even a zero. What is more, all the other *rounds* must have different zero morphs, so that different zero morphs contrast with each other. Even [though] this state of affairs is possible ... it does not have much plausibility as an account of the way in which real speakers process language.

A purely morphologically-based account of intercategoryal polysemy is clearly implausible. In her study of polysemy, Lehrer (1990) proposes the following general hypothesis:

Polysemy (related multiple senses of words) can be predicted from very general functions and / or general cognitive principles. (Lehrer 1990: 211)

Lehrer's model, called the 'weak' cognitive model in Zawada (1996), is a combination of a traditional morphological approach, combined with some findings from Cognitive Linguistics. According to the model proposed by Lehrer (1990), the various senses of a word are derived from each other by means of general cognitive principles such as metaphor and metonymy (represented in (3) for

school). Lehrer (1990:210) then proposes the following principles of conventionality to constrain polysemy:

- (1) If a word already exists to express a meaning, use it; don't use or construct another one.
 - ...
 - (2) If a word lexicalizes a meaning (concept), don't use it to mean something else, even if that meaning would fit the patterns of the language.
 - ...
- These principles can, of course, be overridden when there is good reason for doing so, for example, to be entertaining, literary, shocking, or simply to be unconventional, or more commonly because the conventional word has a nuance that is inappropriate. (Lehrer 1990: 210)

Lehrer also comes to the following conclusion: "...[that her analysis] has a great deal of plausibility, but that it is so vague that it is hard to know how to falsify (confirm) it"; and, "[even though] noncentral senses of words cannot be predicted from central cases, they are not arbitrary; rather they are motivated by image-schema transformations and metaphorical models" (Lehrer 1990: 239-240).

In their account of intercategoryal polysemy, Clark and Clark (1979) make a distinction between idiomatic, well-entrenched and innovative examples of intercategoryal polysemy. They claim that the idiomatic and well-entrenched examples have fixed and ready-made senses which are retrieved from the mental lexicon, whereas for the innovative examples, speakers must "create new

meanings” (Clark and Clark 1979: 804). They come to the following conclusion:

The presence of innovations, near-innovations, and idioms sometimes processed as innovations offers a distinct challenge to most theories of comprehension and production. **These theories implicitly assume that all word meanings are available ready-made in the lexicon.** That assumption is clearly wrong. (Clark and Clark 1979: 807) [my emphasis – BEZ]

In the representational-derivational model, lexical entries are seen as containers (i.e. mental ‘things’) that are related in various ways, either by formal rules, such as WFRs or by cognitive principles, such as metaphor and metonymy. This kind of account may at first appear plausible in analyses that focus exclusively on lexical polysemy, and that contain the implicit claim that lexical polysemy has no formal grammatical consequences. This position was, however, shown to be untenable in that all forms of polysemy, including lexical polysemy, have formal grammatical consequences.

In the representational-derivational model, stable, finite and static linguistic knowledge (in the form of mini-container-like lexical entries in a container-like mental lexicon), plus stable and finite linguistic rules and/or cognitive mechanisms are meant to account for the linguistic knowledge of the speaker. The aim of these accounts is to account for the productivity of intercategoryal polysemy in English (i.e. the systematic and repeated creation of new senses based on existing words by using a fixed set of mechanisms), which is the only view of linguistic creativity

held in this model. However, according to Lehrer (1990), the principles she has suggested can be overridden in truly novel uses, which means that these principles cannot account for conceptual creativity.

The founding metaphor for the representational-derivational model is the basic mathematical algorithm. For example, $N + N \rightarrow N$ as in *black board*, or *bake_V + -er* \rightarrow *baker_N*, or *school₁ + 1* \rightarrow *school₂*. This model can be schematically represented in Figure 5.4.

$$9 + 1 = 6$$

Figure 5.4 *A diagrammatic representation of the representational-derivational model*

With reference to the ability of the representational-derivational model to account for the set of creative data in Table 5.1, the following: The morphological derivation of *dismissal* and *dismissive* from *dismiss* are perhaps the only examples that could be handled in this model, since *-al* and *-ive* are relatively productive suffixes to form abstract nouns and adjectives from a verb (Bauer 1983: 222 and 224). All the other examples of lexical polysemy and intercategory polysemy would have to be derived by zero derivation with complex semantic rules to make

the sense relations clear. The representational-derivational model can also not account for the following phenomena in the test data:

- § the fact that the sense extensions can bring about positive and / or negative connotations (as in *middled* and *skied*);
- § the fact that the context and the tone of the commentator (i.e. the speaker) caused a difference in these connotations (from the expected, usually negative to the very positive but unusual connotations in the case of *skied*);
- § the fact that extensive and detailed (extra-linguistic) background knowledge (in this case of cricket) is necessary to understand all the instances of lexical and intercategoryal polysemy, in spite of the fact that all the words used as examples here (i.e. *dismiss*, *middle*, *sky*, *wide*) are well-known and common words which can be assumed to have extensive lexical entries in any proposed standard mental lexicon; and
- § the notion that sense extensions can change agency and causality, and thereby affect the speech act of the speaker as is illustrated (in 11) where the umpire is blamed by the commentator. The novel use of *wide* (in 11) (with its particular connotation of questioning the decision of the umpire), can in no way be predicted by any standard WFR.

In my view then, the representational-derivational model can account only for the conventionally productive data in the data set, even though it must be considered

that phenomena such as tone and speech acts might, in fact, also affect the interpretation of the regular conventionalised examples. Clark and Clark (1979: 781) also come to the conclusion that a derivational account of novel expressions of intercategoryal polysemy “lead[s] to problems”. The derivational-representational model does, however, have the advantage of incorporating the notion of **production** into it, i.e. the idea that existing conventional units of language and knowledge can be used as a basis for creating new units of language and knowledge.

The disadvantages of the representational-derivational model can be articulated mainly on the non-incorporation of context, extra-linguistic background knowledge and speaker’s intention, particularly in the online production of novel intercategoryal polysemy.

5.4.2 The network-activation model

The network-activation model is similar to the ‘strong’ cognitive model as described by Zawada (1996) and is based on the following general hypotheses (repeated from Zawada 1996: 108):

- 1) There is no intermediate, conceptual-linguistic level in the mind, i.e. there is no container-like mental lexicon with mini-container-like lexical entries; rather linguistic forms are ‘empty’ (of meaning and of syntactic information) phonological and/or orthographic shells that link **directly** (without an intermediate, conceptual level) into our general knowledge base and cognition.

- 2) (Linguistic) meaning is not a 'thing' that resides in a container in our minds that can be described, but it is a process of construal that is repeated in each and every individual (linguistic) event or situation.

These general hypotheses are made more explicit by the following working hypotheses in Zawada (1996: 108):

- 1) Our general knowledge base is structured in the form of typical event frames (or Idealised Cognitive Models (ICM's), which are directly grounded in our experiential knowledge, i.e. in how we mentally and physically operate in the world and perceive / experience the world.
- 2) These event frames are not static knowledge configuration but are constantly adapted to fit particular events or situations or new experiences (mental or physical) by means of various cognitive processes, such as, for example, scanning mechanisms, metonymy and metaphor. (Zawada 1996: 108)

According to the network-activation model, lexical knowledge (both linguistic and encyclopaedic knowledge) is represented in the mind in a distributed fashion in the form of neural networks. Parts of the network are linked in various ways, by means of cognitive processes such as metaphor and metonymy for example, and can be activated to produce expressions. In the case of known expressions this activation is in the form of a (spreading) re-activation of previous linguistic and / or mental experiences. The founding metaphor for this model is the parallel distributed processing of connectionist models (Taylor and Taylor 1990: 24). This model which is overtly cognitive or mentalist, can be schematically represented as in Figure 5.5, where the blocks represent any kind of information, i.e. linguistic (including syntactic and semantic) and encyclopaedic information, which can be connected in

various ways, with the activation shown in bold lines. This particular version of such a model is based on Zawada (1996) which specifically attempts to account for intercategoryal polysemy by means of network activation and metonymy as a cognitive process. Other network models typically focus only on intracategoryal polysemy.



Figure 5.5 A diagrammatic representation of the network-activation model

In this model both form and meaning, as well as general cognitive processes (such as metaphor and metonymy) are taken into account. As was mentioned before, the empirical study in Zawada (1996) was based on entrenched and frequent examples of intercategoryal polysemy in a dictionary corpus. The focus in this

model is, therefore, on the motivation of existing, conventionalised and entrenched senses and relationships (including productive patterns which can be re-activated), rather than on the predictability of novel senses. In this model, a kind of representationalism has been retained, since the kinds of information that are activated (i.e. what is represented by the boxes in the diagram) is based on mutual conventionalised knowledge in the form of domains. This kind of model can readily account for the data in Table 5.1 based on *dismiss* as the senses can each be linked to the relevant domain (either TREATMENT OF PEOPLE AND THEIR RELATED THOUGHTS, IDEAS, ETC. or EMPLOYMENT or CRICKET) (in Cognitive Linguistics domains are typically indicated by capital letters). The regular morphological derivations (in examples 4 and 5 in Table 5.1) can be accounted for by the fact that morphological patterns, such as V-*al*_N and V-*ive*_A are part of the conventionalised knowledge of English mother tongue speakers and these patterns, together with the verb *dismiss* can, through spreading activation create *dismissal* and *dismissive*. Note that *dismissal* can be used in all three the domains above (e.g. *his dismissal from employment / his dismissal for a low score / his dismissal of my plan of action*), whereas *dismissive* can only be used in the TREATMENT OF PEOPLE AND THEIR RELATED THOUGHTS, IDEAS, ETC. (e.g. *he was dismissive of the coach*). In both the CRICKET and the EMPLOYMENT domains, the sense of *dismissive* remains as ‘not taking a person or their related ideas seriously’ (e.g. *the bowler was dismissive of the batsmen*

means that one individual denigrated another, not that the bowler bowled the batsmen out in a game of cricket; similarly *he was dismissive of a labourer* means that the employer denigrated the labourer, not that the labourer was dismissed from employment). The *V-ive_A* pattern is thus far more restricted than either the *V-a/_N* pattern, or even the intercategory pattern (in 1, 6, 7 and 8 in Table 5.1). These kinds of restrictions can also be accounted for in the network activation model. In the intercategory pattern (in 1, 6, 7 and 8 in Table 5.1) it is clear from the collocations which domain plays a role in determining the sense of *dismissed* (*runs, batsmen, injured* and *shake hands* are expressions typically associated with cricket). Alternatively, if *dismissed* is used as either verb, adjective or noun in the context of words such as *boss, employment, work* etc. all the senses will be interpreted in the domain of EMPLOYMENT, and if *dismissed* is used as either verb, adjective or noun in the context of words such as *plans, thoughts, ideas* or *opinions*, all the senses will be interpreted in the domain of TREATMENT OF PEOPLE AND THEIR RELATED THOUGHTS, IDEAS, etc.

From Lehrer's (1990:210) statements that these cognitive "principles can ... be overridden" and my own test data, it appears that the network-activation model has an element of randomness or a kind of unpredictability which can only be motivated in retrospect, rather than account for the online production of truly novel senses. This model can therefore account for both the interpretation of new senses

formed on productive and conventional patterns (such as *V-a_N* and *V-ive_A*, as well as V-N-A in English). It is possible, in retrospect, to construct a radial network with a suitable spreading activation pattern for completely novel senses such as *wided_V* (in 11 in Table 5.1), but it does not seem possible, in this model at least, to explain why the commentator was able to form that novel expression with that particular sense and connotation. A criticism that can therefore be levelled against the network-activation model is that the focus of the model is on the comprehension of novel expressions (i.e. the focus is on how a hearer can interpret a novel expression after it has been created) rather than on the production of novel expressions (i.e. on how a speaker can produce a novel expression). In this model, a hearer can create a relevant activation pattern of a novel expression, based on the context and the linguistic elements in the expression, and so comprehend a novel expression. The model can however, in principle, only give *random* activation as an explanation for the creation of the novel expression, such as *wided_V* (in 11 in Table 5.1) in the first instance. Whilst the random association of thoughts and patterns might indeed have a part to play in linguistic creativity, the patterns in the creation of these novel expressions seem very systematic and not random at all. Random activation as an **only** explanation for all instances of (systematic) intercategoryal polysemy does, therefore not seem justified.

The network-activation model, therefore, goes a step further than the

representational-derivational model. Not only can it give an account for conventional and productive senses and patterns in lexical items, but it can also explain, albeit in retrospect, why and how novel senses and patterns have been created and used with a particular sense and connotation. Why and how the speaker is able to produce a particular novel expression has, however, not been accounted for in this model. In a sense, the network-activation model has resolved the problems with the representational-derivational model that were identified at the end of Section 5.4.2 (i.e. the non-incorporation of context, extra-linguistic background knowledge and speaker's intention in the online production of novel expressions), but it has failed to retain the advantage of that model, namely the notion of actively producing novel expressions.

5.5 Conclusion

In this chapter, I have shown that intercategoryal polysemy is an interesting example of linguistic creativity, particularly because it has both a semantic and a grammatical aspect that has to be explained by theoretical models that intend to make claims about the mental representation and the cognitive processing of linguistic knowledge, particularly if they are to account for conceptual creativity. After defining **polysemy** in its broadest sense as a semantic relationship between any kind of linguistic element (be it lexical, grammatical, or systemic), a set of test data of intercategoryal polysemy, ranging from the conventional to the completely

novel, was introduced. The following theoretical models that give mentalist / cognitive accounts for (intercategorical) polysemy were presented: the representational-derivational model and the network-activation model. It was shown that each of these two models have both advantages and disadvantages. The main advantage of the representational-derivational model is that it, in principle, includes an element of production to account for the creation of new senses derived on the basis of conventional senses, forms and knowledge. The representational-derivational model cannot, however, account for genuinely novel creations since it does not include extra-linguistic background knowledge, context and the speaker's intention into the framework, even if the framework is extended to include cognitive-functional principles, such as metaphor and metonymy. The network-activation model addresses these problems in the representational-derivational model, but because it relies on the re-activation of existing knowledge, it does not account for the production of genuinely novel cases. From this discussion it is clear that only a theoretical model with, at least, the following characteristics will be able to adequately account for linguistic creativity as it is realised in the phenomenon of intercategorical polysemy:

- the model needs to include an active **process** of production which can account for both the production and the interpretation of genuinely novel instances
- the model needs to include the widest possible range of existing knowledge

(both linguistic and encyclopaedic) as the **basis** for the process, and

- the model needs to include context and speaker's intentions.

In Chapter 6, it will be shown how one such a mechanism that has been proposed in Cognitive Linguistics, namely conceptual integration, has these characteristics, and can therefore account for (intercategorical) polysemy as an example of linguistic creativity.