22. Polysemy

1. The notion of polysemy

The probably most widely accepted definition of polysemy is as the form of ambiguity where 2+ related senses are associated with the same word; consider the meanings of \textit{glass} in \textit{I emptied the glass} (‘container’) and \textit{I drank a glass} (‘contents of the container’). Ever since this notion was proposed by Bréal (1897), it has been puzzling researchers from many disciplines: linguists, lexicographers, psycholinguists, psychologists, computer scientists, etc. In the componential Classical Theory of Meaning (Katz and Fodor 1963; Katz 1967), (i) meanings\(^1\) of words were defined on the basis of necessary and sufficient conditions (or features/markers) without reference to contexts, (ii) therefore, a particular entity was either a full member of the category defined by a word or not, and (iii) the similarity of meanings of different words, or senses of the same word, could be quantified by counting the number of features/markers shared by meanings/senses. Thus, a word was ambiguous if it had more than one definition using such features (where no distinction between different kinds of ambiguity – homonymy and polysemy – was made).

Cognitive linguistics (CL), or cognitive semantics, drew on research in philosophy, anthropology, and cognitive psychology and adopted a perspective in which polysemy became an omnipresent property associated with lexical items but also morphemes, grammatical constructions, and whole grammatical classes. Section 1 sketches the devel-

\(^1\) I use \textit{meanings} for unrelated interpretations and \textit{senses} for related interpretations.
opment of polysemy in CL. Section 6 explores how polysemy was addressed in neighboring fields (psycholinguistics and corpus linguistics), and section 10 points out desiderata for future CL research on polysemy.

2. Polysemy in cognitive linguistics

In this section, I will discuss the “history” of polysemy in CL; as in most of CL, I will mostly focus on lexical semantics.

The treatment of polysemy in CL involves (i) a viewing meaning/sense as categorization, (ii) recognizing the importance of context for meaning/senses and that linguistic and encyclopedic knowledge are hard to keep separate, and (iii) incorporating prototype theory into linguistics. As for (i), meaning/sense is viewed as categorization such that, e.g., learning/recognizing that a sparrow is a bird amounts to establishing birds as a category of which sparrows are a member. That is, lexical items are the linguistically coded subset of all conceptual, mentally represented categories.

As for (ii), meanings of lexical items are difficult to pin down without considering both their context and encyclopedic real-world knowledge, an assumption from Fillmore’s (1975, 1982) Frame Semantics. An early example involves what Cruse (1995: 44) calls cooperative readings: The presence of zeugma in 2 appears to indicate that dissertation is polysemous with at least two senses (‘intellectual content’ vs. ‘physical object’), but the slight change to 2 results in an absence of zeugma, which does not support a similar polysemy (following Geeraerts 1993 and, ultimately, Norrick 1981):

(1) a. Judy’s dissertation is thought-provoking and yellowed with age
b. Judy’s dissertation is still thought-provoking although yellowed with age
(2) the splinter in my hand

In fact, Taylor (2012: 220 ff.) points out it is often unclear where in an utterance polysemy resides – in a lexical item or its context. Is 2 polysemous because of the polysemy of in or of hand or do both senses co-select each other? Similarly, Taylor (2012: 226) illustrates how the meaning of cut the lawn changes from the prototypical one to the meaning of ‘cut someone a piece of instant lawn (as cut someone a piece of cake)’ that it may have in an instant lawn business. Finally, Labov (1973) has shown that speakers presented with something that looks like something between a bowl and a cup prefer to call it bowl when it contains potatoes, and cup when it contains coffee.

As for (iii), CL has drawn on research in cognitive psychology (much of it by Heider/Rosch, e.g., Rosch 1975, 1978) that showed subjects/speakers do not categorize objects using necessary/sufficient features but by comparing their similarity to the prototype (see Taylor this volume) of the candidate category/categories. Specifically, prototypical members of a category are listed more often/earlier in experiments where subjects are asked to list members of a category, their category membership is verified more quickly, and they give more rise to generalizations about the category. The notion of a prototype has been defined/operationalized in different ways (see Lakoff 1987): the prototypical sense of a word may be the most frequent and/or salient and/or most concrete one, the
earliest attested one (historically or acquisitionally), the one from which most others can be derived best, but these criteria need not converge. I will consider a prototype (say, of *bird*) to be an abstract conceptual entity that combines attributes with a high cue validity for that category (‘flies’, ‘has feathers’, ‘lays eggs’, ‘has a beak’, etc.).

This perspective gave rise to the notions of (i) *radial categories* – categories with a central element combining many high-cue validity attributes and motivating the existence of less central members; the most-cited example is probably *mother* – and (ii) *family resemblance categories* – categories in which not all members share the same set of attributes but in which members are disjunctively related by sharing at least some attributes with each other; the usual example is Wittgenstein’s ([1953] 2001) *game*. That means that prototype effects and category structure can be found on the level of the individual senses, on multiple levels of more schematic elements subsuming similar senses, and on the level of the whole category of interrelated senses of an element; thus, “the semantic value of a word need no longer be a single, unitary structure, but rather, […] a set of interrelated senses” (Cuyckens and Zawada 2001: xiii). For example, Norvig and Lakoff (1986) discuss the structure of senses of the polysemous verb *take*. The prototype is exemplified by *John took the book from Mary* and different links are postulated to connect senses; for example,

- **profile shift**, relating the prototype to *John took the book to Mary*, which profiles the movement of the Agent, to the Recipient;
- **metaphoric links** (see Gibbs this volume): *John took the book to Mary* is connected to *John took the book at Mary* via the metaphor *APPLYING FORCE IS TRANSFERRING AN OBJECT*;
- **metonymic links** (see Barcelona this volume) and frame-addition links: *John took the book to Chicago* is connected to *John took Mary to the theater* via the metonymy *GOING TO (public establishment) D STANDS FOR DOING C (activity conventionally done at D)*.

Additional important types of links connecting senses are generalizations, specializations, and image-schema transformations. The latter is exemplified by *John walked over the bridge* (involving a *SOURCE* schema) being related via an image-schema transformation to *John lives over the hill* (involving an *ENDPOINT* schema). This also means that, ultimately, some relations between senses of a word are motivated by speakers’ conceptualizations of real-world events and concrete bodily/sensori-motor experience (cf. Lakoff and Brugman 1986, Gibbs and Matlock 2001).

### 2.1. Phase 1: extreme splitting

Considering studies of word senses on a continuum from extreme lumpers (strict monosemy analyses, e.g., Ruhl 1989) to extreme splitters (highly granular polysemy analyses), the initial phase of CL research on polysemy would be in the latter extreme. Beginning with Brugman’s (1981) analysis of the preposition *over* (cf. also Lakoff 1987: 416 ff. and Brugman and Lakoff 1988) and Lindner (1981) on *up* and *out*, cognitive-semantic studies involved the above theoretical notions and many minimally different senses in the so-called *full-specification approach*. For instance, in Brugman’s/Lakoff’s account
III. Central topics

of over, 3 and 3 constitute different senses since they differ with regard to whether the
trajectors (Sam, the bird) are in contact with the landmark (the wall) or not:

(3) The bird flew over the wall.
(4) Sam climbed over the wall.
(5) John walked over the bridge.
(6) John walked over the hill.

Similarly, 3 and 3 are considered different senses, because only 3 involves a landmark
(the hill) that is vertically extended (Lakoff 1987: 422). Brugman’s/Lakoff’s analysis
involves more than twenty senses arranged in a radial category around over’s prototypi-
cal sense, which they claim is exemplified by The plane flew over. Two crucial notions
of such analyses are those of cognitive/representational reality and motivation (of links
and senses). Regarding the former, many studies did not topicalize the ontological status
of the lexical networks of polysemous items, but some literature assumed some cognitive
reality: “a network-style mode of storage is cognitively real” (Brugman and Lakoff 1988:
477). Langacker offered less bold characterizations:

It is not suggested that a strong claim of psychological reality can be made for any particular
linguistic analysis as currently constituted. The description of a language is nevertheless a
substantive hypothesis about its actual cognitive representation. (1987: 56, see also p. 382)

Regarding the latter, motivation is situated between unpredictable arbitrariness and per-
fected predictability. For instance, if one extended the analysis of over to non-prepositional
cases – e.g., as a particle or prefix – one would encounter uses like sleep overpowered
him. If one wanted to express the concept ‘to overpower’ but did not know the verb
overpower, one might not predict there must be an English verb overpower. Nevertheless,
once one considers the prototypical spatial meaning of over and the independently-postu-
lated metaphor CONTROL IS UP, then a verb such as overpower “makes sense” (Lakoff

For ten to fifteen years, this approach was extremely influential. In fact, since CL (i)
viewed lexical items as categories, (ii) abandoned a strict separation between lexis and
syntax, and (iii) therefore, viewed constructions as categories, too, polysemy analyses
and Thornburg (2002) on -er nominalizations, Smith (2001) on Icelandic datives, Hen-
drikse (2001) on the Southern Bantu noun class system, Selvik (2001) on Setswana noun
classes, etc. However, the most far-reaching extension was to the semantics of syntactic
constructions. Goldberg’s work (e.g., 1992, 1995) on argument structure constructions
was particularly influential. First, it affected the decision where to localize polysemy:
instead of assuming that different intransitive verbs such as sneeze or cough are poly-
semous in having a caused-motion sense (e.g., Pat sneezed/coughed the napkin off the
table), she argued the syntactic pattern V-NP-PP itself has a meaning (here, ‘caused-
motion’) and that, say, verbs in a constructional verb slot elaborate the construction’s
meaning; for instance, the prototypical transfer-of- possesion sense of give elaborates
the prototypical ‘X causes Y to receive Z’ meaning of the ditransitive construction V-
NP-NP.

The second important extension was that constructions, just like words, were assumed
to have multiple senses related by polysemy links. For instance, apart from the prototypi-
Polysemy

Polysemy analyses became increasingly popular, scholars also began to discuss their shortcomings. One discussion was triggered by Sandra and Rice (1995); see also Rice (1996):

- how is the prototype defined? For over, Brugman/Lakoff postulated ‘above-across’ is the prototype, Tyler and Evans (2001) postulated ‘above’ to be central, Deane (2005) “characterized the preposition in terms of a trajector entity which intervenes between [an] observer and the landmark” (Taylor 2012: 236), etc.;
- how are different senses distinguished and is the fine level of resolution often adopted really warranted? Are 3 and 3 really different senses or can they be conflated into one? (Are there even different word senses?)
- what motivates the different representational formats (cf. Lewandowska-Tomaszczyk 2007: Section 4.2 for a comparison) and what is the ontological status of the proposed networks? Cognitive linguists often argued their analyses were compatible with, or stood for, some sort of cognitive reality, but how much do such linguistic analyses warrant psychological/psycholinguistic claims? (Cf. also Riemer (2005: Ch. 1)

Another discussion involved how much (cognitive) linguists can really say about mental representation (especially on the basis of something as volatile as introspection; cf. Nisbett and Wilson 1977). First, Croft (1998) argued that the typical introspective linguistic evidence – e.g., grammatical/semantic idiosyncrasies – can exclude more general models of mental representation (i.e., more schematic/moneismic models), but that, conversely, grammatical/semantic generality does not automatically support more general models – for that, additional experimental/observational evidence is required (e.g., sentence-sorting, sentence-similarity judgments, or [lack of] similar distributional behavior in corpora).

Sandra (1998) limited the purview of linguistic studies even more, arguing that “linguists have a very minor role to play when issues of mental representations are at stake [...] At most they can restrict the range of potential options” (1998: 361) Sandra also cautions CL to avoid the polysemy fallacy to automatically postulate very fine-grained sense distinctions (when more schematic sub-analyses might be sufficient) and to consider such analyses a rendering of the language user’s mental representation of the linguistic data. This view, which appears to exhibit a slightly old-fashioned and non-interdisciplinary division of linguists vs. non-linguists/psycholinguists as well as a lack of recognition of, say, Tuggy’s (1993) introduction of multiple levels of schematization, was addressed by Tuggy (1999). Tuggy points out shortcomings in Sandra’s characterization of Croft’s
positions and the polysemy fallacy, but also argues that introspective data are “extremely important evidence” because “[w]hen such intuitions line up impressively, they acquire a degree of objectivity” (1999: 352). This argument actually reinforces Sandra’s point since proper experimentation is a way to get intuitions by multiple speakers to “line up”. Also, Tuggy proposes additional polysemy diagnostics such as direct intuitions about sense relations, perceptions of puns, evidence from speech errors, and “holes in the pattern”, as when particular usages that should go with a particular form do not. (See Riemer 2005: Ch. 3 for discussion.)

Another point of critique involves the relation of the polysemies of words and/in constructions. One account discussed above with regard to 4 argued that constructions such as the ditransitive are polysemous just as the lexical items are. However, Croft (2003: 55) argued that the senses of, say, the ditransitive construction appear to be more due to the classes of verbs inserted into them: “It is not an accident that the verbs found with ditransitive sense E [‘X enables Y to receive Z’ from 4] are verbs of permission […]. That is, it seems that the different ‘senses’ of the ditransitive construction are very closely tied to the verb classes that each ‘sense’ occurs with” (2003: 56). Croft proceeds to make a case for verb-class specific constructions and even verb-specific constructions (cf. also Boas 2008), which testifies to the many difficulties of locating at which level(s) polysemy is situated.

2.3. Phase 3: newer developments

As a result of the research mentioned above, research on polysemy went, simplistically speaking, two different ways. First, new theoretical approaches were developed, most notably Tyler and Evans’s Principled Polysemy framework (but cf. also Kreitzer 1997); this approach will be discussed briefly in this section. Second, polysemy research turned to more diverse data, using psycholinguistic experimentation and corpus data, which is the topic of Section 6.

The Principled Polysemy approach (cf. Tyler and Evans 2001; Evans 2005) targeted the first of the two problem areas. First, they proposed criteria to determine when two usages constitute different senses by doing more justice to the role of context and distinguishing polysemy from vagueness; second, they proposed criteria to identify the prototype, or sanctioning sense, of a polysemous category. As for the former, for some usage to count as a distinct sense of $x$, it must contain additional meaning not apparent in other senses associated with $x$ (the meaning criterion) and it will feature unique or highly distinctive syntagmatic/collocational patterns (the concept elaboration criterion) and similarly distinctive structural dependencies (the grammatical criterion); the latter two criteria, thus, make an implicit reference to the study of corpus data. As for the latter, Evans (2005) lists four linguistic criteria (and mentions additional empirical evidence of the type discussed by Sandra and Rice [1995] or Croft [1998]): diachronic primacy, predominance in the lexical network, predictability regarding other senses, and – for time – a sense involving experience at the phenomenological level or – for over – relations to other prepositions.

This approach is promising as it is among the first to propose more rigorous “decision principles”; however, the concept elaboration and the grammatical criterion and many
of the prototype criteria (which, curiously, do not feature acquisitional primacy) are
gradable and may not converge. Nonetheless, these criteria help make decisions more
replicable especially as more empirical evidence guiding linguists’ decision is gathered.

3. Polysemy in neighboring fields

This section discusses how neighboring fields such as corpus linguistics and psycholin-
guistics have dealt with polysemy. This is essential because, as became apparent above,
CL regularly points to findings/methods in neighboring fields (without, however, really
integrating much of such work); cf. Cuyckens et al. (1997) for discussion. In general,
one can make a coarse distinction between (i) corpus-linguistic work, which by its very
nature is concerned more with associative/co-occurrence relations (cf. Section 6) and
psycholinguistic experimentation, which targets more semantic/categorical relations (cf.
section 8).

3.1. Corpus linguistic approaches

Corpus-linguistic work on polysemy within CL comes in three kinds: first, there are
studies where the corpus-linguistic component consists merely of using a corpus as a
source of examples – ideally, examples are not just cherry-picked to support a particular
point but also considered if they constitute counterexamples; given the limited role that
corpus methods other than mere retrieval play in such work, this will not be discussed
here. Second, there are analyses which involve the retrieval of, ideally, many examples
of the element to be analyzed, which are then annotated for various characteristics, which
are then analyzed statistically. Third, there are studies straddling the boundary between
corpus linguistics and computational (psycho)linguistics, which differ from the previous
kind of analyses in that many do not (i) involve (semi-)manual annotation and (ii) aim
at uncovering something about human language per se but rather test/evaluate computa-
tional models of linguistic data (with no pretense of cognitive realism).

The main assumption underlying the latter two approaches is what Miller and Charles
discuss as the co-occurrence approach, the assumption that distributional similarity is
correlated with functional (semantic, pragmatic, etc.) similarity, as expressed in Harris’s
(1970: 785 f.) famous dictum that “difference of meaning correlates with difference of
distribution”. Miller and Charles (1991) contrast this with the substitutability approach,
which essentially amounts to an experiment in which subjects fill gap in sentences with
one of several words whose similarity is tested. From this they argue for the notion of
a contextual representation of a word, which is

a mental representation of the contexts in which the word occurs, a representation that
includes all of the syntactic, semantic, pragmatic, and stylistic information required to use
the word appropriately. (1991: 26)

Correspondingly, different levels of statistical complexity can be distinguished in this
second approach. The earliest relevant corpus work is largely monofactorial in nature
and does not yet include the more advanced statistics characteristic of much of contem-
porary corpus linguistics; relevant examples include Schmid (1993) and Kishner and Gibbs’s (1996) work on how senses of *just* are correlated with different parts of speech in *just*’s environment.

More advanced analyses in “multidimensional corpus-based cognitive semantics” were then developed in particular in Gries’s and Divjak’s *Behavioral Profile* approach. This approach is similar in spirit to corpus-linguistic work such as Atkins (1987) and Hanks (1996, 2000) and has been applied to both polysemy and synonymy. It typically (i.e., not always) consists of four steps: (i) retrieving a sample of the expression(s) in question; (ii) annotating the concordance lines for a large number of features; (iii) converting these data into a table of percentage vectors that state how much of the data in percent exhibits a particular feature; (iv) statistically analyzing the data with exploratory tools (such as cluster analysis). Gries (2006) was the first to apply this approach in the study of polysemy, studying the verb *run* (cf. Glynn to appear for a replication) and showing how the correlations of percentage vectors helps decide where to locate senses in a network, whether to lump or split senses, what the prototypical sense may be. Berez and Gries (2009) use cluster analysis as a corpus-based equivalent to psycholinguistic sense-sorting experiments to explore what senses of *get* exhibit high inter-sense similarity. Divjak and Gries (2009) extend this approach to the senses of near-synonymous phrasal verbs in English and Russian, and other work has targeted near synonymy (Divjak and Gries 2006; Divjak 2006; Gries and Otani 2010). In addition, the BP approach has stimulated interesting extensions using different kinds of exploratory statistics and corpus data: Glynn (2010) applies what amounts to the BP approach to *bother* but instead of using cluster analysis he uses multiple correspondence analysis (MCA); Levshina (in prep.) uses an MCA to discover structure in the semantic field of seating furniture; Janda and Solovyev (2009) apply very similar methods to constructional similarities.

Finally, there are more computational approaches based on unannotated texts. Biber (1993) studies how the polysemy of the word *right* is reflected in the distribution of its collocates. More technical approaches involve NLP applications based on co-occurrence vectors in multi-dimensional space; cf. Schütze (2000) for a discussion of word, sense, and context vectors. However, much of this work is more concerned with ambiguity, not polysemy. Other similar work more concerned with psychological realism/applications is Burgess and Lund’s (1997) HAL or Landauer and Dumais’s (1997) LSA, which are both based on large co-occurrence matrices of words and have been used successfully in many domains, which may point to promising applications within cognitive semantics once the “symbol grounding problem” is addressed, possible via the notion of embodiment (cf. Traxler 2012: 89 f.).

Corpus data have been useful in cognitive semantics, but they usually do not allow researchers to make definitive statements about cause-effect relations or online processing. The studies discussed in the next section target these aspects.

3.2. Psycholinguistic approaches

Psycholinguistics was probably the field that CL related to first: Even when CL was still far from adopting experimental/observational approaches, there were attempts to inte-

---

2 BP analyses were first presented independently by Gries and Divjak in 2003 at the ICLC; a similar but otherwise unrelated approach is Speelman et al. (2003).
grate the psycholinguistic models/findings regarding into CL. Deane (1988), for instance, is an attempt to unite the theory of image schemas, Relevance Theory, and Anderson’s architecture of cognition. Geeraerts (1985) discusses how various characteristics of the human mind all motivate why the human mind should exhibit the type of conceptual organization around prototypes. But what about psycholinguistics proper?

Polysemy was not represented much in psycholinguistic research before the 1960s and some early work (Asch and Nerlove 1960 or Macnamara et al. 1972) was concerned with questions that may seem unrelated to CL work on polysemy. However, that is not entirely true. For instance, the former studied how children acquire and distinguish words denoting both a physical and a psychological quality, such as *hard*, *deep*, *bright*, etc., certainly a topic of current relevance. The latter study tests the hypotheses that speakers store meanings associated with a phonological form in a particular order, that this order is very similar across speakers, and that during comprehension speakers try meanings in that order. While this may seem far-fetched, given the lexical networks that have been developed, a usage-based approach that accords frequencies of words, senses etc. a primary role, implies at least some sort of rank-ordering of senses based on their frequencies. Indeed, the experimental results refuted that simplest rank-ordering hypothesis but also showed that (i) the first 1–2 senses named by subjects often coincided and (ii) context plays an important role in rapid online meaning disambiguation.

Such examples notwithstanding, most early work on the subject was concerned with ambiguity or homonymy and explored

- the time course of activation of word senses: are only relevant or relevant and irrelevant senses of words activated and how does the presence of multiple meanings or senses affect word recognition (cf. Hino and Lupker 1996 and Azuma and Van Orden 1997)?
- the importance of context for sense selection: does it have an effect at all, what exactly is it, and when does it kick in?
- the importance the frequency/dominance of senses plays for sense selection: less frequent meanings take longer to access (Hogaboam and Perfetti 1975);
- the interaction of the above.

That is, most earlier studies on lexical access/disambiguation neither included any systematic distinction between homonymy and polysemy in their designs/explanations; in fact, some psychological/psycholinguistic studies use *polysemy* to refer to cases such as *ball* (‘spherical object’ vs. ‘dance event’), which linguists would class as homonymy, others use *ambiguity* as meaning ‘polysemy’. (In fact, some recent introductions to psycholinguistics – e.g., Byrd and Mintz (2010) or Menn (2011) – do not feature *polysemy* or *ambiguity* as index entries). Therefore, some early work speaks directly to many questions of CL, but much is ‘only’ indirectly related to CL since, e.g., it hardly explores the nature of the relations between senses or adopt the same fine degree of sense granularity; cf. Gibbs and Matlock (2001) or Klein and Murphy (2001). While the exact nature of lexical access is still debated, there is evidence that

- all meanings of a word are accessed even if they are contextually inappropriate (semantic or syntactic context cannot constrain initial access);
- context both before and after the word can help make subordinate but contextually appropriate meanings more available for selection; also, context helps suppress con-
textually inappropriate meanings of homonyms within approximately 200 ms and can make reactions to dominant senses as fast as to unambiguous controls; – dominant and subordinate senses react differently to context (cf. Lupker 2007 for a detailed overview).

An early study that does speak to cognitive semanticists is Caramazza and Grober (1976). They first report results of three different tasks on the word line – acceptability judgments of concordance contexts, similarity judgments of such contexts, and a sentence production task – which produce highly interrelated results. Interestingly, they applied multidimensional scaling and cluster analysis to the similarity judgments and obtained a clear and interpretable 5-cluster/sense solution. On the whole, their model of how polysemy is represented in the mental lexicon, the Process Theory of Meaning, is similar to the monosemy approach and assumes a single or a small set of very general meanings. However, it also accords crucial roles to context and encyclopedic knowledge (cf. also Anderson and Ortony 1975 for similar conclusions): senses other than the central one are derived by extension/analogy, or “instruction rules”, and an encyclopedic dictionary, which stores all factual information a speaker has about a word, constrains the application of the instruction rules.

Another relevant study is Durkin and Manning (1989). For nearly 200 words (11 % of them homonyms), they collected typical senses from subjects and relatedness ratings of all senses to central ones. They find that senses of polysemous words are rated as more similar to each other than senses of homonymous words, but also that, while contexts boosts senses’ salience ratings, dominant senses enjoy a larger degree of salience even in contexts biasing subordinate senses. Also, the ease with which subordinate senses can be accessed differs considerably across words.

A classic study on the processing of homonymous meanings vs. polysemous sense is Frazier and Rayner (1990). Their eye-movement data indicate that, in the absence of a disambiguating context, fixation times for words with multiple meanings are longer while fixation times for words with multiple senses are not. They explain that as a consequence of having to immediately disambiguate such homonyms so as not to maintain inconsistent representations and selecting one meaning involves suppression of others, which requires extra processing time. Similarly, Pickering and Frisson (2001: 556) propose that, upon hearing a homonymous word in a non-disambiguating context, speakers make an early selection of a meaning whereas, upon hearing a polysemous word in such a context, the user “activates one underspecified meaning and uses context to home in on the appropriate sense”. Additional evidence in particular for the higher relatedness of senses of polysemous words (compared to the meanings of homonyms) comes from Williams (1992). He shows that senses of polysemous adjectives resulting in zeugma in the do so test prime contextually irrelevant related senses: “it does not appear to be possible to suppress the irrelevant meanings of a polysemous adjective in the same way as [those] of a homonym” (1992: 202). In addition, the time course of activation is similar to that of dominant properties of monosemous nouns. Finally, the direction of priming is significant, too: priming from non-central senses to central ones was signifi-

---

3 This approach is compatible with Frisson and Pickering’s (1999) study, which shows that both literal and metonymic senses can be accessed immediately, “perhaps through a single underspecified representation” (1999: 1366).
significant at all delays, but not the other way round, which Williams interprets as possibly related to category structure/similarity effects (e.g., prototype effects).

Similar differences were found in Brisard et al. (2001), who demonstrated significant priming effects for polysemous and vague adjectives, whereas homonymic adjectives did not. Also, consider Rodd et al. (2002, 2004). In the former study, they find that the unrelated senses of homonymous words (and their wider attractor basin) slow recognition down whereas the related senses and richer semantic representations (and deeper attractor basins) of polysemous words speed recognition up. In the latter, they propose a distributed-semantic-representation model of lexical knowledge that accommodates these effects by assuming that the related senses of polysemous words share overlapping regions in semantic space. Similarly, Beretta, et al. (2005) show that the meanings of homonymous words are accessed more slowly than senses of polysemous words; cf. also Azuma and Van Orden (1997).

On the other hand, in experiments similar to Light and Carter-Sobell’s (1970) and Bainbridge et al.’s (1993), Klein and Murphy (2001, 2002) show that both memory performance and sensicality judgments suggests that senses of paper are related but not similar, stored separately or at least functionally distinct. In fact, senses of one word may be excited and inhibited at the same time. Also, Klein and Murphy (2001: exp. 3) find no performance difference between homonyms and polysemous words but, in Klein and Murphy (2002: exp. 1–3), conclude that the similarity of polysemous senses is graded, and that polysemous senses share more connections than homonymous meanings (not unlike what a family resemblance approach would predict).

In sum, there is some agreement on some issues, but most hypotheses implicit in CL polysemous analyses are far from as unambiguously supported as many in the CL mainstream would hope for − a great deal of work lies ahead.

4. Desiderata

Given the huge amount of research on polysemy and ambiguity, this overview was selective and much interesting work could not be discussed. While psycholinguistic work has yielded some robust findings, many of the central questions of CL regarding senses’ distinctness, relatedness, representation, and their right level of granularity, remain largely unanswered. Across all three areas − CL, corpus linguistics, and psycholinguistics − a consensus is emerging to assume a multidimensional semantic space in which usages or senses are located such that their spatial proximity reflects distributional and/or semantic similarity; cf., e.g., Gries (2010) and Taylor (2012) for cognitive/corpus linguistics and, Rodd et al. (2004: 89) for psycholinguistics. Thus, while integral to early CL, the notion of distinct senses appears more of a descriptive device rather than a claim about psycholinguistic reality. This conception does justice to the fact that the same word/sense − i.e., region of semantic space − can be accessed or traversed at different levels of resolution and from different angles/trajectories. A simple example is shown in Figure 1, which represents the same usages (as dots) in semantic space from three different angles. The left panel suggests there is one group of relatively coherent usages, maybe corresponding to one sense. However, the other two panels show the same usages from different angles (e.g., from different semantic/discourse contexts), and these panels give rise to two or
four senses. That is, context facilitates sense recognition by imposing a particular view on, or trajectory through, stored exemplars, and the possibility of creativity is afforded by the speaker’s freedom to (i) approach the same region from different perspectives or (ii) see similarities between different point clouds in semantic space and exploit this linguistically by means of, say, analogy, or (iii) condense regions of space.

In what follows, I discuss a few areas for future research. First, corpus-based work needs to be developed further both in terms of scope (words, senses, and features included) and methodology. Current developments are promising and future work may evolve by including powerful new tools such as network models, fuzzy clustering, etc.

More importantly, CL must approach polysemy more interdisciplinarily. Many experimental studies in psycholinguistics on ambiguity should be replicated on the basis of CL analyses to shed more light on whether the fine-grained senses distinctions, the nature of links, etc. are supported. Similarly, we need better evidence on the role of prototypes in online processing and on how word senses interact with constructions and their senses.

With regard to language acquisition, there seem to be only few studies targeting polysemy from a CL perspective. The few studies that there are – e.g., Dowker (2003), Nerlich et al. (2003) on get, Rice (2003) on prepositional networks, Kidd and Cameron-Faulkner (2005) on with – have unfortunately not left enough of a mark on CL in spite of their relevance. Rice (2003: 275) makes the interesting suggestion that

> a lexical category for a young child does not start out as either monosemous or polysemous, but as potentially very homonymous. Additional senses do not emerge through extension. Rather, they may be integrated through some sort of schematization process at a much later date.

This is an area that needs more experimental/observational research, but also maybe computational modeling; cf., e.g., Parisien and Stevenson (2009) for a study of get.

Finally, neurolinguistics offers a completely new range of applications; cf. Coulson (2004) for an overview of EEG/ERP or Stringaris et al.’s (2006) fMRI study of semantic relatedness. Burgess and Simpson (1988) tested whether the brain’s two hemispheres respond identically to target words more associated with the dominant or the subordinate meaning of an ambiguous word and found that “the two hemispheres have opposite responses to subordinate meanings. The left hemisphere deactivates subordinate meanings, but the right hemisphere increases them over time” (Traxler 2012: 528 f.). Mason
and Just (2007) showed that the brain activity arising from processing lexically ambiguous words differs as a function of meaning dominance and working memory capacity (i.e., individual differences). Finally, CL has approached the polysemy of content and function words in the same way, but the two types of words seem to be lateralized differently (Bradley and Garrett 1983); in fact, Damasio and colleagues suggest that nouns vs. verbs and even different categories of concrete objects are represented in different neural regions, which has implications for polysemous words (cf. Lupker 2007: 169). Only by combining multiple approaches/tools will CL be able to develop polysemy analyses that are compatible with the cognitive commitment to make one’s account of human language accord with what is generally known about the mind and brain from disciplines other than linguistics.

5. References

Anderson, Richard C. and Andrew Ortony

Asch, Solomon and Harriet Nerlove

Atkins, Beryl T. Sue

Azuma, Tamiko and Guy C. Van Orden

Bainbridge, J. Vivian, Stephan Lewandowsky, and Kim Kirsner

Beretta, Alan, Robert Fiorentino, and David Poeppel

Berez, Andrea L. and Stefan Th. Gries

Biber, Douglas

Boas, Hans C

Bradley, Dianne C. and Merrill F. Garrett

Bréal, Michel
Brisard, Frank, Gert Van Rillaer, and Dominiek Sandra

Brugman, Claudia

Brugman, Claudia and George Lakoff

Burgess, Curt and Kevin Lund

Byrd, Dani and Toben H. Mintz
2010 *Discovering Speech, Words, and Mind*. Chichester: John Wiley.

Caramazza Alfonso and Ellen Grober

Coulson, Seana

Croft, William

Croft, William

Cruse, D. Alan

Cuyckens, Hubert, Dominiek Sandra, and Sally Rice

Cuyckens, Hubert and Britta Zawada

Deane, Paul D.

Deane, Paul D.

Divjak, Dagmar S.

Divjak, Dagmar S. and Stefan Th. Gries
Divjak, Dagmar S. and Stefan Th. Gries

Dowker, Ann

Durkin, Kevin and Jocelyn Manning

Evans, Vyvyan

Fillmore, Charles J.

Fillmore, Charles J.

Frazier, Lyn and Keith Rayner

Frisson, Steven and Martin J. Pickering

Geeraerts, Dirk

Geeraerts, Dirk

Gibbs, Raymond W. Jr. and Teenie Matlock

Glynn, Dylan

Glynn, Dylan

Goldberg, Adele E.

Goldberg, Adele E.

Gries, Stefan Th.
Gries, Stefan Th.  

Gries, Stefan Th. and Naoki Otani  

Hanks, Patrick  

Hanks, Patrick  

Harris, Zelig S.  

Hendrikse, A.P.  

Hino, Yasushi and Stephen J. Lupker  

Hogaboam, Thomas W. and Charles A. Perfetti  

Jackendoff, Ray  

Janda, Laura A. and Valery D. Solovyev  

Katz, Jerrold J.  

Katz, Jerrold J. and Jerry A. Fodor  

Kidd, Evan and Thea Cameron-Faulkner  

Kishner, Jeffrey M. and Raymond W. Gibbs Jr.  

Klein, Deborah E. and Gregory L. Murphy  

Klein, Deborah E. and Gregory L. Murphy  

Kreitzer, Anatol  

Labov, William  
22. Polysemy

Lakoff, George

Lakoff, George and Claudia Brugman

Landauer, Thomas K. and Susan T. Dumais

Levshina, Natalia

Lewandowska-Tomaszczyk, Barbara

Light, Leah L. and Linda Carter-Sobell

Lindner, Susan

Lupker, Stephen J.

Macnamara, John, Ann O’Cleirigh, and Thomas Kellaghan

Menn, Lise

Michaelis, Laura A. and Knud Lambrecht

Miller, George A. and Walter G. Charles

Nerlich, Brigitte, Zazie Todd, and David D. Clarke

Nikiforidou, Kiki

Nisbett, Richard E. and Timothy DeCamp Wilson

Norrick, Neal R.
Norvig, Peter and George Lakoff

Panther, Klaus-Uwe and Linda L. Thornburg

Parisien, Christopher and Suzanne Stevenson

Pickering, Martin J. and Steven Frisson

Rice, Sally

Rice, Sally A.

Riemer, Nick

Rodd, Jennifer M., Gareth Gaskell, and William D. Marlsen-Wilson

Rodd, Jennifer M., Gareth Gaskell, and William D. Marlsen-Wilson

Rosch, Eleanor

Rosch, Eleanor

Ruhl, Charles

Sandra, Dominiek

Sandra, Dominiek and Sally A. Rice

Schmid, Hans-Jörg
1993 *Cottage and Co., idea, start vs. begin*. Tübingen: Max Niemeyer.

Schütze, Hinrich

Selvik, Kari-Anne
Smith, Michael B.  

Speelman, Dirk, Stefan Grondelaers, and Dirk Geeraerts  

Stringaris, Argyris, Nicholas Medford, Rachel Giora, Vincent C. Giampietro, Michael J. Brammer, and Anthony S. David  

Taylor, John R.  

Taylor, John R.  

Tuggy, David.  

Tuggy, David  

Traxler, Matthew J.  

Tyler, Andrea and Vyvyan Evans  

Williams, John N.  

Wittgenstein, Ludwig  

*Stefan Th. Gries, Santa Barbara (USA)*

### 23. Space

1. Introduction
2. Spatial language defined
3. Spatial adpositions
4. Spatial demonstratives
5. Cross-linguistic differences and ‘linguistic relativity’
6. Conclusions
7. References

1. Introduction

Space has long been viewed as one of the fundamental building blocks in cognitive linguistics. For almost four decades it has been argued that both language and thought