

1 **The *how* and *why* of polysemy: A pragmatic account**¹

2 Ingrid Lossius Falkum, University of Oslo

3

4

5

6 **Published in *Lingua* 157 (2015), pp. 83-99.**

¹ This research was supported by a postdoctoral fellowship from the Research Council of Norway (project no. 205513). I would like to thank Robyn Carston for many valuable and inspiring discussions of the topic of the present paper, and three anonymous reviewers for their thoughtful and constructive comments which helped improve the paper.

7

The *how* and *why* of polysemy: A pragmatic account

8 *Abstract*

9 A large number of word forms in natural language are polysemous, that is,
10 associated with several related senses (e.g., *line*, *run*, *tight*, etc.). While such
11 polysemy appears to cause little difficulty in verbal communication, it poses a
12 range of theoretical and descriptive problems. One concerns its very existence:
13 What is it about our language systems that make them so susceptible to
14 polysemy? In this paper I discuss two approaches to polysemy with different
15 answers to this question: (i) A code-based approach that treats polysemy in
16 terms of the operation of lexicon-internal generative rules, and (ii) an inference-
17 based approach that takes polysemy to be governed by pragmatic inferential
18 processes applying at the level of individual words. After evaluating how each of
19 these accounts fares with respect to some empirical data, I look more broadly at
20 their implications for the emergence and development of polysemy. I conclude
21 that, overall, the pragmatic approach provides the most promising basis for a
22 unified account of the role of polysemy in several domains, and for explaining
23 what motivates its proliferation natural language.

24

25 *Keywords:*

26 Polysemy, pragmatics, rule-based accounts, lexical underdeterminacy, Relevance
27 Theory

28

29 **1. Introduction**

30 A large number of natural language word forms are *polysemous*, that is,
31 associated with several related senses. Here are some illustrations:

32

33 (1) Kate *began a book*.

34 (2) Jane Austen wrote *good* books.

35 (3) a. There was *rabbit* all over the highway.

36 b. Steven had *rabbit* for dinner.

37 c. The model wore *rabbit* on the catwalk.

38 (4) John is *a lion*.

39 (5) *The ham sandwich* is getting impatient.

40

41 In (1), the VP *began a book* is compatible with several readings, each involving a
42 different event (e.g., reading/writing/mending/dusting/ripping up... etc.). It is
43 common to refer to constructions of this type as 'logical metonymy'
44 (Pustejovsky, 1995), where the argument of the verb in syntax is different from
45 that argument in 'logical form': In (1), the idea would be that the NP *a book*
46 denotes part of an event, which is used to stand for the event as a whole (e.g., the
47 VP 'reading a book'). The interpretation of the adjective *good* in (2) requires a
48 specification of its conceptual content (e.g., 'good reads'), which would be
49 different from the one it has in, e.g., *good knife/football player/weather/child*, and
50 so on (Katz, 1964). The alternations between the different senses of the noun
51 *rabbit* in (3) – 'rabbit remains', rabbit meat', 'rabbit fur' – is standardly analysed
52 as a form of systematic (or regular) polysemy (Apresjan, 1974) where the
53 related senses of a word are predictable on the basis of the ontological category
54 of its denotation (cf. other animal terms; *crocodile, seal, mink, lamb*, etc.). Finally,
55 in (4) and (5), we have a metaphorical use of the noun *lion* ('strong, courageous,
56 takes risks', etc.), derived on the basis of properties associated with the animal
57 denotation (e.g., Glucksberg, 2001), and a metonymic use of the NP *The ham*
58 *sandwich* ('the person who ordered the ham sandwich'), based on a contextually
59 salient association between the ham sandwich and the person who ordered it
60 (Nunberg, 1979). Such metaphorical and metonymical extensions are usually
61 taken to be prime sources of polysemy in language.

62 Already Bréal (1924 [1897]) noted that when talking to each other we
63 rarely get confused by the multiplicity of meanings that a word can have.
64 Sometimes, of course, speakers may exploit the polysemous potential of a lexical
65 item to create confusion or a humorous effect (e.g., by use of a pun), but
66 generally, polysemy causes little difficulty for users of a language; it is something
67 that we handle effortlessly and unconsciously, most of the time. As speakers, we
68 can trust hearers to quickly and reliably figure out the lexical meanings we
69 intend to communicate on a given occasion.

70 In contrast, polysemy raises a host of theoretical and descriptive
71 problems.² A first issue concerns the delimitation of the polysemy phenomenon.
72 How – if at all – should it be distinguished from the accidental multiple encoding
73 we find in homonymy (e.g., *bank, coach*), on the one hand, and contextually
74 modulated senses (e.g., *good weather/good student/good book*), derived from a
75 single encoded meaning, on the other hand?³ Several tests for distinguishing
76 between ambiguity (homonymy/polysemy) and non-specificity (monosemy)
77 have been proposed (e.g., Goddard, 2000; Kempson, 1977; Quine, 1960), but as
78 Geeraerts (1993) has meticulously shown, different tests may not always agree
79 with each other and by manipulating the context, they can be made to yield
80 inconsistent results.

81 A second taxing issue is how polysemous lexical items are represented in
82 the mental lexicon.⁴ According to ‘sense enumeration lexicons’, all the different
83 senses of a lexical item are stored under a single entry, and comprehension
84 involves selection of the contextually appropriate sense among the list of
85 candidates (e.g., Brugman & Lakoff, 1988; Katz, 1972). Given the proliferation of
86 polysemy, a problem for such fully encoding lexicons is that they would have to
87 store indefinitely many semantic distinctions for each lexical item. So-called

² This discrepancy has led some scholars to talk of a ‘polysemy paradox’ (Ravin & Leacock, 2000; Taylor, 2003).

³ In lexicography, this translates into a methodological issue of determining which cases of multiple encodings that should be listed as distinct entries (homonymy) and which should be listed under a single entry (polysemy). For any single entry, there are also decisions to be made concerning which senses are established (hence should be listed) and which are not (hence should not be listed).

⁴ By ‘mental lexicon’, I refer to individual speakers’ stable mental representations of words, which include information regarding their semantic properties (in the form of ‘meanings’ or ‘senses’), as well as phonological and syntactic properties, which are accessed when a word is encountered in discourse.

88 'core meaning approaches' take polysemy to be represented in terms of a single,
89 maximally general meaning, from which the appropriate one is contextually
90 derived (e.g., Caramazza & Grober, 1976; Ruhl, 1989). A challenge for this kind of
91 approach is to determine the appropriate level of abstraction for defining the
92 core meaning of a lexical item, in order to capture what is common to all its
93 possible, sometimes radically different uses. In between these two extreme
94 views we find several 'middle-ground approaches', assuming only a limited
95 number of stored senses of a lexical item and the rest to be derived in context
96 (e.g., Carston, 2002; Tyler & Evans, 2003), sometimes combined with a proposal
97 for a principled set of criteria for distinguishing between those senses that are
98 stored in the lexicon and those that are constructed on-line in language use
99 (Tyler & Evans, 2003). This kind of approach, involving a differential
100 representation of polysemous senses (some may be stored, some may be
101 contextually derived) is supported by experimental evidence (e.g.,
102 Klepousniotou, 2007; Klepousniotou, Titone, & Romero, 2008; Pyykkänen, Llinás,
103 & Murphy, 2006), although the results are to some extent conflicting (cf. Klein &
104 Murphy, 2001).

105 A third issue is *how* lexical meanings get extended into several different
106 meanings. Assuming that some (possibly many) senses of polysemous lexical
107 items are derived during on-line processing, what is the nature of the processes
108 or mechanisms involved? In other words, what kind of phenomenon is
109 polysemy? Is it mainly a result of the operation of lexical rules for sense
110 extension (e.g., Copestake & Briscoe, 1995; Ostler & Atkins, 1992; Pustejovsky,
111 1995)? Is it a direct reflection of how our cognitive categories are structured
112 more generally (e.g., Brugman, 1988; Brugman & Lakoff, 1988; Lakoff, 1987)? Or
113 does it arise through pragmatic processes operating over underspecified lexical
114 meanings and contextual knowledge (e.g., Carston, 2002; Sperber & Wilson,
115 1998)?

116 A fourth issue that the proliferation of polysemy in natural language
117 raises is the fundamental question of *why* it exists. Why are word meanings
118 extended in this way? What is it about our language systems – specifically their
119 lexical component – that make them so susceptible to polysemy?

120 The focus of this paper will be these two last issues: the *how* and *why* of
121 polysemy. I will consider two opposing views regarding the nature of the
122 polysemy phenomenon – rule-based vs. ‘radical’ pragmatic approaches – with
123 very different answers to the questions of how linguistic polysemy arises, and
124 what its underlying motivation may be.⁵ After evaluating how each of the
125 accounts fares with respect to explaining the polysemy data exemplified in (1)
126 through (5), I discuss their implications for the emergence and development of
127 polysemy. I conclude that, overall, the radical pragmatic account provides the
128 most promising basis for a unified account of the role of polysemy in several
129 domains, and for explaining what motivates its proliferation in natural language.
130 As regards the first two issues concerning the delimitation of polysemy and its
131 representation in the mental lexicon, I will assume, based on extant experimental
132 evidence, a differential representation of polysemy (where, depending on their
133 degree of conventionalisation, some senses may be stored in our mental lexicons,
134 some may be contextually derived), and operate with an intuitive distinction
135 between conventional (i.e., ‘encoded’ or ‘semantic’) polysemy and contextually-
136 derived polysemy, acknowledging that there may be no clear-cut way of drawing
137 this distinction.⁶

138 **2. Two approaches to polysemy**

139 A fundamental difference between rule-based and pragmatic approaches to
140 polysemy lies in their radically different conceptions of what a language is.
141 Underlying rule-based approaches is the view that language provides an
142 information-rich code that enables speakers and hearers to encode and decode
143 their thoughts in much detail, with pragmatics as a useful add-on to this

⁵ The cognitive linguistic hypothesis about underlying conceptual mappings being the source of linguistic polysemy (e.g., Lakoff, 1987) will not be considered further in this paper. Although the hypothesis may indeed be correct for some, even many, cases of conventional polysemy, it leaves open the question of how new polysemous senses are constructed as a result of communicative interactions between speakers of a language, which is the main concern of this paper.

⁶ One reviewer pointed out that this seems to imply that there is a continuum rather than a proper distinction between conventional and contextual polysemy, and that this would present a problem for the view that there is a distinction between encoded word meanings and unencoded (inferred) senses. One solution might be to abandon the assumption that words encode concepts and instead see them as encoding underspecified (non-conceptual, abstract schematic) meanings, a position which is currently being pursued by Carston (2012, 2013) within the relevance-theoretic framework. While I am generally sympathetic to this view, it would exceed the scope this paper consider it further here.

144 linguistic capacity, operating primarily when some interpretation other than the
145 linguistic default was intended. By contrast, radical pragmatic accounts see the
146 role of the linguistic system as being that of providing a minimal input or clue – a
147 ‘sketch’, or ‘blueprint’ of the speaker’s meaning – which the pragmatic inferential
148 system uses as evidence to yield hypotheses about occasion-specific, speaker-
149 intended meanings. In this sense, we may call the first a code-based approach,
150 and the second an inference-based approach. In what follows, I will consider
151 each of them in turn.

152 2.1 Code-based approaches: polysemy as lexical rules

153 Before Grice ([1967] 1989), virtually all theories of communication were based
154 on the so-called *code model*. The guiding assumption of this model is that
155 communication is a matter of encoding and decoding of messages in the form of
156 signals (e.g., Peirce, 1955; Saussure, 1974; Shannon & Weaver, 1949; Vygotsky,
157 1986, and many others). On this approach, linguistic communication proceeds by
158 a speaker encoding a thought into a sentence of a language – where a language is
159 seen as a code that pairs phonetic and semantic representations of sentences –
160 and by the hearer decoding the uttered sentence into an identical thought.
161 A well-known problem with the code-model of communication is that linguistic
162 utterances typically contain context-sensitive and/or linguistically ambiguous
163 expressions, as illustrated by (6) and (7):

164

165 (6) I lost my bat yesterday.

166 (7) John is a fine colleague.

167

168 An understanding of (6) requires assigning the appropriate referents to the
169 indexical expressions (*I, my, yesterday*) and disambiguating the homonymous
170 noun *bat* (‘wooden instrument’ vs. ‘flying rodent’) and the conventionally
171 polysemous verb *lost* (‘deprived of’ vs. ‘deprived of through death’). In
172 (7), the hearer must, in addition to assigning the appropriate referent to the
173 proper name *John*, form a hypothesis about the meaning of *fine* in the NP *fine*
174 *colleague* (‘considerate’, ‘diligent’, ‘hard-working’, etc.). In cases such as these it
175 is widely agreed that the hearer must consider information beyond that which is

176 linguistically encoded in order to derive the speaker-intended meaning, as the
177 context-sensitive expressions make reference to the discourse situation in which
178 they were uttered. So, code-models of communication need some way to account
179 for how hearers assign contextually appropriate meanings to the expressions
180 used (see Sperber & Wilson, 1986/1995, for a more extensive critique of such
181 communication models).

182 Rule-based approaches to polysemy can be seen as modern, sophisticated
183 versions of the code-model of communication as described above, in which a
184 considerable amount of the context-sensitivity of lexical meanings is built into
185 the linguistic system. More specifically, polysemy is seen as being generated by a
186 set of lexicon-internal generative rules, which operate over information-rich
187 semantic representations to yield default interpretations (Asher, 2011; Asher &
188 Lascarides, 2003; Copestake & Briscoe, 1992, 1995; Lascarides & Copestake,
189 1998; Ostler & Atkins, 1992; Pustejovsky, 1995). To illustrate, consider the
190 'logical metonymy' in (1), repeated below as (8), analysed by rule-based
191 approaches in terms of a lexicon-internal mechanism forcing a non-conventional
192 reading of the complement (Pustejovsky, 1995) ⁷.

193

194 (8) Kate *began a book*.

195

196 Constructions of this kind are seen as involving a verb that subcategorises for an
197 NP or a progressive VP syntactically, but which semantically requires a
198 complement with an event interpretation. In cases where this requirement is not
199 satisfied by the surface syntactic structure, a coercion mechanism changes the
200 denotation of the NP from an entity into an event consistent with eventive
201 information stored as part of the lexical representation of the noun. In (8), the
202 appropriate event would be provided by the so-called *telic* role of the noun:

⁷ More explicitly, Pustejovsky's (1995: 111) suggestion is that each expression α may have a set of shifting operators available to it, which he calls $\Sigma\alpha$, which may operate over an expression to change its type and denotation:

FUNCTION APPLICATION WITH COERCION (FAC): If α is of type c , and β is of type $\langle a, b \rangle$, then,

- (i) if type $c = a$, then $\beta(\alpha)$ is of type b .
- (ii) if there is a $\sigma \in \Sigma\alpha$ such that $\sigma(\alpha)$ results in an expression of type a , then $\beta(\sigma(\alpha))$ is of type b .
- (iii) otherwise a type error is produced.

203 'books are for reading' (or alternatively, by its *agentive* role: 'books come about
204 as a result of a process of writing'). In this way, the interpretation 'Kate began
205 *reading* a book' (or 'Kate began *writing* a book) is generated by the linguistic
206 system by default.

207 This type of approach was originally proposed by Pustejovsky (1991,
208 1995), whose main aim was to provide a more explanatory account of polysemy
209 than a mere listing of senses in the lexicon (cf. Katz, 1972). A more promising
210 approach, he argued, which captures how word senses may partially overlap and
211 be logically related to each other, is a lexicon where items are decomposed into
212 information-rich templates, so-called *qualia structures*, combined with a set of
213 generative mechanisms for the composition of lexical meanings.

214 An advantage of this rule-based theory of the processing of logical
215 metonymy is that it accounts for clear interpretive tendencies in uninformative
216 contexts (in the absence of any further contextual clues the preferred or 'default'
217 interpretation of (8) would be that 'Kate began reading a book' (or writing it),
218 and not that she, e.g., began ripping it up). The availability of such preferred
219 readings is often taken as evidence of a linguistic-semantic process. The claim is
220 that if the lexicon does not propose such a sense it is unclear how it can arise
221 since it is not otherwise indicated by the context (Asher, 2011; Copestake &
222 Briscoe, 1995).

223 Notwithstanding its intuitive appeal, the problems with this approach are
224 many (Asher, 2011; Blutner, 2002; Bosch, 2007; de Almeida, 2004; de Almeida &
225 Dwivedi, 2008; Falkum, 2007; Fodor & Lepore, 2002). First, there seem to be
226 many cases where a verb makes a demand on a complement that its lexical entry
227 does not satisfy. In the generative lexicon account, such cases are seen as
228 semantically ill-formed: Pustejovsky (1998) gives as an example the VP *enjoy the*
229 *rock*, which does not have a default interpretation due to the lack of a telic role
230 defined for the noun *rock*. Exceptions arise, according to Pustejovsky, when the
231 object is construed relative to a specific activity, as in *The climber enjoyed that*
232 *rock*, where *rock* acquires telicity on the basis of the semantics of the subject NP.
233 However, consider the arguably well-formed utterances in (9) and (10).

234

235 (9) Peter enjoyed the nice weather.

236 (10) *Karen enjoyed the children.*

237

238 Assuming that the intended interpretations here are that ‘Peter enjoyed being
239 outside in the nice weather’ and ‘Karen enjoyed playing with the children’, it is
240 unclear how they could be generated when there seem to be no telic information
241 in the lexical representations of the nouns *weather* and *children* that the coercion
242 mechanism could take as input to the compositional process. It is also difficult to
243 see how they could acquire telicity on the basis of the of the subject proper
244 nouns. Thus, it seems that the generative lexicon theory would either make no
245 interpretive predictions for cases such as (9) and (10), or wrongly predict that
246 they are ill-formed. Second, it is difficult to see how the rule-based account can
247 avoid making wrong predictions about many compositional interpretations. For
248 instance, the VPs *begin a car* and *begin a thermometer* should be interpreted as
249 ‘begin driving a car’ and ‘begin measuring the temperature’, due to the telic roles
250 encoded by the complement nouns (cars are for driving; thermometers are for
251 measuring temperatures) (Fodor & Lepore, 2002). It is unclear what would
252 prevent such clearly infelicitous interpretations from being constructed as
253 ‘default’. Third, by modelling the processing of logical metonymy entirely in
254 terms of a lexicon-internal process, the rule-based account is unable to account
255 for the interpretive flexibility that is arguably involved in these constructions.
256 Although preferred readings in uninformative contexts clearly exist (e.g., the
257 tendency to interpret the VP *begin a book* as ‘begin reading a book’ or ‘begin
258 writing a book’), more specific contextual information can easily point the hearer
259 toward a ‘non-default’ interpretation, which would have to override the ‘default’
260 reading (e.g., ‘begin dusting a book’, ‘begin mending a book’, ‘begin designing a
261 book’, ‘begin ripping up a book’, etc.). As a consequence, the rule-based account
262 must allow the compositional interpretations generated by the linguistic system
263 to be defeasible, but if this is so, some justification has to be given for why such
264 defeasible semantic rules are necessary in lexical interpretation, when
265 defeasibility is widely agreed to be one of the hallmarks of our pragmatic
266 capacity.

267 Many rule-based approaches give a similar analysis of the adjectival
268 specification exemplified by (2) above, repeated here as (11).

269

270 (11) Jane Austen wrote *good* books.

271

272 On the Pustejovskyan (1995) approach, the meaning of an evaluative adjective
273 like *good* is generated in linguistic context by a process of ‘selective binding’,
274 which enables an adjective to make available a selective interpretation of an
275 event expression contained in the lexical representation (or the ‘qualia
276 structure’) for the head noun. In (11), the idea would be that *good* selectively
277 modifies the event description given by its *telic* role (‘books are for reading’),
278 giving rise to the interpretation ‘good reads’. This analysis runs into the same
279 sort of problems as that for ‘logical metonymy’ above. It has little to say about
280 uses where there is arguably no telic or eventive information for the adjective to
281 selectively modify but where the compositional process seems to proceed as
282 usual (e.g., *good children*, *good weather*). It must also appeal to pragmatics for an
283 explanation of non-default interpretations, for instance, the number of other
284 context-dependent ways in which a book could be good (e.g., ‘entertaining’, ‘easy
285 to read’, ‘intellectually challenging’, ‘beautifully designed’, ‘useful to kill flies
286 with’, etc.)

287 A seemingly stronger case for a lexical rule-based analysis is the type of
288 polysemy that patterns with the syntactic count-mass distinction, often referred
289 to as ‘systematic polysemy’ (cf. Apresjan, 1974). Here the related senses of a
290 word can be predicted from the ontological category of its denotation, and are
291 linguistically marked by the count or mass syntax of the NP in which it occurs.
292 Consider again the examples in (3) above, repeated here as (12).

293

294 (12) a. There was *rabbit* all over the highway. (‘rabbit stuff’)

295 b. Steven had *rabbit* for dinner. (‘rabbit meat’)

296 c. The model wore *rabbit* on the catwalk. (‘rabbit fur’)

297

298 Computational semantic approaches have influentially argued that systematic
299 polysemy is generated by an inventory of lexical inference rules, where the effect
300 of the rules is to change the value of a [+COUNT] or [+MASS] feature in the lexical

301 representation of the noun, thereby altering its denotation accordingly (e.g.,
302 Copestake & Briscoe, 1992, 1995; Ostler & Atkins, 1992). One such lexical
303 inference rule is the UNIVERSAL GRINDER (originally proposed by Pelletier, 1975),
304 which creates from a count noun denoting a physical object a mass noun with
305 properties for an unindividuated substance, yielding the ‘rabbit stuff’ sense in
306 (12)a. above. In addition, the lexicon is thought to contain a set of
307 conventionalised sub-cases of this rule, including a specialised rule of MEAT-
308 GRINDING that forms food-denoting mass nouns from animal-denoting count
309 nouns, and a rule of FUR-GRINDING that forms fur-denoting mass nouns from
310 animal-denoting count nouns, yielding the ‘meat’ and ‘fur’ senses of *rabbit* in
311 (12)b. and (12)c. above. The idea is that this wholly linguistic account avoids a
312 listing of predictable senses in the lexicon and provides an explanation of how
313 such sense alternations can be extended productively to any new members of a
314 category targeted by the rules.

315 Again, one of the main problems with this rule-based approach is its lack
316 of interpretive flexibility. Even this kind of ‘systematic’ polysemy seems to
317 exhibit a considerable degree of context-sensitivity. Consider the following
318 examples, where the linguistic alternation between count and mass uses of
319 nouns yields senses that go beyond those that are normally taken to be
320 generated by linguistic rules.

321

322 (12) d. Will a hamster bite if it senses *rabbit* on my hands? (‘rabbit odour’)

323 e. [Biology teacher]: *Rabbit* is smaller than hare. (‘rabbit faeces’)

324 f. [Hunter]: This time of year I prefer using *rabbit* (‘electronic rabbit
325 calls’).

326 g. Last winter, we discovered *rabbit*, moose and fox in our garden. (‘rabbit
327 tracks’)

328

329 The contextually appropriate interpretation of each of the uses of *rabbit* in
330 (12)d.-g. should be easily inferable from the situation of utterance, but their one-
331 off character makes it seem unlikely that any of them can be generated by a
332 lexical rule. So it seems that lexical rules, even if they could be shown to be real,
333 would only be able to account for a subset of the interpretations that the

334 alternation between count and mass uses of nouns may give rise to. But if a
335 considerable number of senses are derived pragmatically, we may question the
336 motivation for the sense extension rules in the interpretive system: what do they
337 add to a theory of polysemy comprehension by way of explanation?⁸

338 Another problem is that the rules inevitably overgenerate. For instance,
339 the utterance *Sam enjoyed but later regretted the rabbit* (Copestake & Briscoe,
340 1995: 42), whose ‘default’ interpretation seems to be that Sam enjoyed but later
341 regretted *eating* the rabbit, would, given the universal grinder and the rules of
342 animal meat-grinding and animal fur-grinding, be three-ways ambiguous, and it
343 is not clear how hearers determine when one rule has prevalence over the
344 others. If we have to appeal to some sort of pragmatic mechanism to do this,
345 which indeed seems likely, it leaves us again with the question of what role the
346 lexical rules are playing in the interpretive process.

347 Turning to the examples of metaphorical and metonymic sense
348 extensions in (4) and (5) above, repeated here as (13) and (14), it is widely
349 agreed that consideration of discourse context is required for their
350 interpretation (e.g., Gibbs, 1994; Glucksberg, 2001; Nunberg, 1979; Sperber &
351 Wilson, 2008; Wilson & Carston, 2006).

352

353 (13) John is *a lion*.

354 (14) *The ham sandwich* is getting impatient.

355

356 However, rule-based accounts have suggested that sense extension rules may be
357 involved in some metonymic and metaphorical processes as well. For instance, it
358 has been proposed that since the metaphorical extension from animals to
359 humans with some particular characteristic(s), exemplified in (13), appears to be

⁸ Two reviewers pointed out that an analysis in terms of ellipsis of the NP head might also be possible for (12d-g) – where, for example, *rabbit* in (12g) might be an ellipsis for ‘rabbit tracks’, in a case where the prior discourse makes this interpretation available – with the consequence that such cases of specialised interpretations should not be seen as equivalent to the conventional interpretations in (12a-c). However, an ellipsis analysis might in principle also be possible for the conventional polysemy patterns in (12a-c), where, for instance, *rabbit* in the utterance *Steven had rabbit for dinner* might be seen as an ellipsis for ‘rabbit meat’. While I do not think that this is the correct way to account for these examples, the fact that an ellipsis analysis might be possible for all the examples in (12), depending on the information given by the prior discourse, suggests to me that we have to do with a single phenomenon here.

360 productive (*John is a lion/pig/lamb*, etc.), it can be (partly) expressed in terms of
361 a lexical rule, although the properties ascribed to the human by use of the animal
362 term would not be encoded in its lexical representation (Briscoe & Copestake,
363 1991; Copestake & Briscoe, 1995).⁹ Similarly, metonymic extensions such as that
364 from the ham sandwich to the person who ordered the ham sandwich in (14) are
365 analysed as involving a basic sense extension rule PHYSICAL OBJECT → HUMAN
366 (Copestake & Briscoe, 1995) or, as in Asher's (2011) recent proposal, a lexical
367 semantic process of coercion, where a type conflict requires an adjustment of
368 predication in order to satisfy a type presupposition (e.g., the predicate *is getting*
369 *impatient* presupposes an external argument of the type AGENT).

370 Given that consideration of pragmatic factors is clearly required in order
371 to derive the contextually appropriate meaning in both these cases, it is unclear
372 what is gained by introducing lexical rules here. For one thing, in view of the
373 considerable context-dependence of the processes of metaphor and metonymy,
374 the rules would only be able to account for a (very) small subset of cases. For
375 another, the pragmatic mechanism(s) that allow(s) us to construct the range of
376 metaphorical and metonymic meanings that are clearly *not* rule-governed,
377 should also enable us to derive the senses in (13) and (14).

378 In this section, I have discussed rule-based approaches to polysemy. On
379 the basis of a set of standard examples, I have argued that in spite of making
380 accurate predictions in a number of 'default' situations, rule-based accounts
381 leave much work for the pragmatic system to do, both in overriding 'default'
382 interpretations in contexts where another non-default interpretation was clearly
383 intended, and in constructing unpredictable (non-rule-governed)
384 interpretations, for instance, in cases of metaphor and metonymy. But if
385 pragmatics can do this work, it seems likely that it can also do the part of the
386 interpretative work that rule-based accounts do adequately. In the next section, I
387 consider how the polysemy data can be analysed within a wholly pragmatic-
388 inferential account of utterance interpretation.

⁹ This kind of regularity might also be captured in a conceptual metaphor approach (Lakoff & Johnson, 1980), in which (13) could be seen as a linguistic instantiation of the underlying conceptual metaphor HUMAN PERSONALITY TRAITS ARE ANIMALS.

389 *2.2 An inference-based approach: polysemy as pragmatics*

390 The challenge for code-based accounts of interpretation, even of the highly
391 sophisticated kind proposed by Pustejovsky and others, is that what is conveyed
392 by linguistic communication – both at the implicit ('what is implicated') and the
393 explicit ('what is asserted') levels – generally goes well beyond what can be
394 coded, and does so in a highly flexible way. Most pragmatic contextualist
395 accounts of verbal utterance understanding therefore follow Grice ([1967] 1989)
396 in his view that communication is first and foremost an inferential process, and
397 involves a kind of mind-reading: By using a verbal utterance, a speaker provides
398 evidence of her intention to communicate something to the hearer, and the
399 hearer recovers this intention by an inferential process using the evidence
400 provided. A particularly influential pragmatic theory that builds on this insight is
401 relevance theory (Carston, 2002; Sperber & Wilson, 1986/1995; Wilson &
402 Sperber, 2004, 2012). I will now discuss an alternative approach to polysemy
403 based on this framework.

404 According to relevance theory, human information processing “tends to
405 be geared to the maximisation of relevance” (Sperber & Wilson, 1986/1995:
406 260), where relevance is seen a potential property of inputs to cognitive
407 processes (e.g., verbal utterances, gestures, facial expressions, etc.), and is
408 assessed in terms of the amount of effort used to process the input and the
409 ‘positive cognitive effects’ the individual may derive from it (where a positive
410 cognitive effect can be described broadly as a ‘worthwhile difference to the
411 individual’s representation of the world’). Other things being equal, the more
412 cognitive effects an input yields to an individual and the less effort it takes to
413 process it, the more relevant it is to the individual at that particular time.
414 Further, relevance theory takes verbal utterances to constitute a special kind of
415 input, communicating “a presumption of [their] own optimal relevance” (ibid.).
416 By requesting the addressee’s attention, the communicator conveys that her
417 utterance is more relevant than alternative stimuli competing for his attention at
418 the time. An optimally relevant utterance is one that achieves enough cognitive
419 effects to make the utterance worth processing, while avoiding causing the
420 hearer any unnecessary effort in achieving those effects. The hearer’s goal in

421 communication is to find an interpretation of the speaker's utterance that meets
422 the expectations of relevance raised by the utterance itself.

423 In this framework, the distinction between linguistic semantics and
424 pragmatics is seen as corresponding to different processes involved in utterance
425 comprehension: (i) decoding of the linguistic material into a 'logical form', and
426 (ii) pragmatic inference. A logical form is seen as a structured set of concepts: a
427 'template' or 'schema' for a range of possible propositions, which contain slots
428 that have to be filled – a process that requires pragmatic inference (Carston,
429 2002). In this way, the relevance-theoretic approach to verbal understanding
430 distinguishes itself sharply from rule-based accounts in that most of the
431 interpretive work is performed not by lexicon-internal generative mechanisms
432 but by pragmatic processes operating over underspecified semantic
433 representations.

434 On this account, then, the speaker-intended event associated with the VP
435 in instances of 'logical metonymy' would be derived entirely by means of a
436 pragmatic process. Consider again the example in (1), repeated here as (15).

437

438 (15) *Kate began a book.*

439

440 Let us assume that (15) has the following logical form (cf. de Almeida & Dwivedi,
441 2008):

442

443 (16) KATE BEGAN [_{VP} [_{v⁰} e] [_{NP} a BOOK]]

444

445 As shown by (16), the syntactic structure of sentences containing a logical
446 metonymy can be seen as containing an extra VP with an empty verbal head. The
447 verbal gap that remains in the logical form of such constructions will have to be
448 saturated using information from the discourse context. Although the process
449 itself will be linguistically mandated and consist in supplying a missing
450 constituent to the proposition expressed, the relevant event associated with the
451 VP will be supplied by a wholly pragmatic process. Imagine the following
452 context:

453

454 (17) Kate, John and Sue work as book conservators at the British Museum. At
455 the moment they are working on restoring a collection of medieval books,
456 all of which are in a poor condition after having been stored on the
457 shelves for many years. Because they are completely covered in dust, each
458 book has to be carefully dusted before being rebound. One day, after
459 hours of hard work, John asks if they should all take a break and go for
460 coffee. Sue has just finished her pile and is ready to follow John to the
461 coffee bar in the Great Court, when Kate utters: 'Hang on a minute! I've
462 just *begun a huge old book*.'

463

464 The most relevant interpretation of the last part of Kate's utterance is clearly
465 that she has just 'begun *dusting* a huge old book', not reading it, which would be
466 the default interpretation predicted by the Pustejovskyan account. This would be
467 the one that satisfies the hearer's context-specific expectations of relevance: it is
468 the one that requires the least processing effort and offers satisfactory effects, in
469 the form of an adequate explanation for the content of Kate's previous utterance
470 of 'Hang on a minute!'. The prediction is that in (17) the hearer will go straight
471 for this interpretation without the prior computation – and subsequent
472 cancellation – of a 'default' interpretation.¹⁰

473 Now consider again VPs such as *begin a car* and *begin a thermometer*, for
474 which the rule-based, Pustejovskyan approach discussed in 2.1 above made
475 wrong interpretive predictions ('begin *driving* a car', 'begin *measuring* the
476 temperature'). The pragmatic account proposed here would, of course, come
477 with no such interpretive predictions, but it would equally require that an event
478 be supplied when the VPs are embedded within an utterance in a context.
479 Imagine the context of a garage where Bill is employed as a mechanic. Here it is
480 easy to imagine the most relevant interpretation (i.e. the least effort demanding,
481 yielding the expected sort of cognitive effect(s)) of an utterance of *Bill began a*
482 *car* being that 'Bill began *repairing* a car'. However, a speaker using the VP *begin*

¹⁰ Pustejovsky (1995) acknowledges that in some cases coercion can lead to different eventive interpretations depending on the aspect of the qualia structure that is modified. For instance, *begin a book* has two possible default interpretations: 'begin reading a book', derived from the telic quale of *book*, and 'begin writing a book', derived from the agentive quale.

483 *a car* to describe a situation in which someone began *driving* a car would (in
484 most cases, at least) not be optimally relevant, as the choice of this expression
485 instead of the more conventional *start a car*, would, due to the extra effort of
486 processing it would induce, send the hearer off searching for additional effects,
487 which would not be part of the speaker's intended meaning.

488 On the other hand, rule-based accounts are no doubt correct in their
489 assumption that some interpretations come more readily to mind in
490 uninformative contexts, and could therefore be said to have a 'default' character.
491 However, the claim that this is evidence of a linguistic semantic process
492 considerably underestimates the fact that hearers rarely come to the
493 interpretation process 'empty handed', as it were; utterances are not understood
494 in a vacuum. If, instead of being objective and linguistically given, context is seen
495 as a psychological construct – a subset of the hearer's assumptions about the
496 world – which may include assumptions derived from the observation of the
497 physical environment, encyclopaedic knowledge, memories and beliefs as well as
498 the preceding linguistic context (cf. Sperber & Wilson, 1986/1995), there would
499 not be any entirely context-free interpretations. When the assumptions that may
500 be derived from the discourse context are scarce, for instance, if (15) above were
501 to be interpreted in isolation, the hearer will have to rely more on information
502 stored in his long-term memory in interpreting the utterance. Given this, any
503 interpretive preferences observed for logical metonymies in the absence of
504 further context might stem not from lexically stored information but from highly
505 accessible real-world knowledge about the denotations of the lexical concepts in
506 the utterance. For instance, a person reading a book may be regarded as a
507 stereotypical event, which may be stored in encyclopaedic memory as a chunk
508 and accessed as a single unit of information. Retrieving this information from
509 encyclopaedic memory during the interpretation of (15), as a result of the
510 decoding and activation of the lexical concept *BOOK*, would require little
511 processing effort compared to other possible interpretations (e.g., dusting,
512 designing, mending, ripping up, etc.) which would involve accessing several units
513 of information and thus be more costly in processing terms. In this way, we may
514 account for why certain interpretations are often favoured over others without

515 being committed to the view that these are always computed as a result of
516 default inferences generated by the lexicon.¹¹

517 In relevance theory, lexical interpretation is seen as typically involving
518 the construction of *ad hoc* concepts – occasion-specific senses – which may be
519 narrower or broader than the linguistically encoded senses (Carston, 2002;
520 Wilson & Carston, 2006, 2007; Wilson & Sperber, 2012). A mentally-represented
521 concept, a constituent of the ‘language of thought’ (Fodor, 1975, 2008), is seen as
522 an address or entry in memory that may give access to different kinds of
523 information, including (i) *lexical* information connected with the linguistic form
524 that encodes the concept (i.e. its phonological and syntactic properties), and (ii)
525 a set of assumptions, or *encyclopaedic* information, about the denotation of the
526 concept, that is, conceptually represented assumptions and beliefs, including
527 stereotypes and culture-specific information, and also, in many cases, imagistic
528 and/or sensory-perceptual representations (Sperber & Wilson, 1986/1995: 86).
529 For example, the concept BOOK may give access to assumptions such as ‘Books
530 can be read, they are physical objects, are often entertaining, can be intellectually
531 challenging, ...’ and so on. Lexical interpretation involves taking the encoded
532 concept and its associated encyclopaedic information, together with a set of
533 contextual assumptions, as input to the inferential process of constructing a
534 hypothesis about the speaker’s intended meaning. Consider again the adjectival
535 specification in (2) above, repeated here as (18).

536

537 (18) Jane Austen wrote *good* books.

538

539 On this account, one pragmatic sub-task for the hearer in interpreting the
540 utterance in (18) would be a specification or narrowing of the concept
541 linguistically encoded by *good* in the NP *good books* (e.g., ‘good reads’) by a
542 process of *ad hoc* concept construction, taking as input encyclopaedic
543 information associated with the other lexical concepts in the utterance. The

¹¹ The different predictions about the processing of logical metonymy made by rule-based and pragmatic theories have been subject to some experimental testing (see, for instance, de Almeida, 2004; de Almeida & Dwivedi, 2008; McElree, Frisson, & Pickering, 2006; McElree et al., 2001; Pickering, McElree, & Traxler, 2005; Traxler, Pickering, & McElree, 2002). However, the results are inconclusive and to some extent conflicting, so more research would be needed to settle this debate.

544 adjective *good* can be seen as encoding a very general concept, which, on most
545 occasions of use, will have to be pragmatically adjusted in order for the hearer to
546 arrive at the speaker-intended interpretation. Often, as in (18), the
547 encyclopaedic information associated with the head noun will play a key role in
548 this process. The pragmatic process of *ad hoc* concept construction is
549 considerably more flexible than the rule-based generative mechanism discussed
550 in 2.1 above, in that not only linguistically-specified information but *any*
551 activated encyclopaedic or situation-specific assumption can be used in deriving
552 the communicated concept, as long as the hearer's occasion-specific expectations
553 of relevance are satisfied. In this way, the relevance-theoretic account predicts
554 that *good book* may communicate different occasion-specific senses (e.g., a good
555 book could be one that is 'entertaining', 'easy to read', 'intellectually challenging',
556 'beautifully designed', 'useful to kill flies with', etc.). It also provides an account of
557 the problem cases for the rule-based approach, namely, examples such as *good*
558 *children* (e.g., 'well-behaved', 'diligent', 'responsible', 'caring', etc.), *good weather*
559 (e.g., 'sunny', but could be used to mean 'overcast' if in a fishing context, etc.),
560 *good time* (e.g., 'enjoyable', 'fun', 'relaxing', etc.), and so on (with no linguistically-
561 specified purpose encoded by the head noun that could serve as input to lexical
562 rules), all of which would involve the construction of a different *ad hoc* concept
563 on the basis of activated encyclopaedic and/or situation-specific information,
564 constrained by the hearer's expectations of relevance.

565 Similarly, the uses of *rabbit* in (3) above, repeated here as (19), can be
566 analysed in terms of pragmatic narrowing where the concept communicated has
567 a more specific denotation than the concept linguistically encoded.

568

- 569 (19) a. There was *rabbit* all over the highway. ('rabbit remains')
570 b. Steven had *rabbit* for dinner. ('rabbit meat')
571 c. The model wore *rabbit* on the catwalk. ('rabbit fur')

572

573 Let us assume that when encountering a mass use of the noun *rabbit*, the output
574 of linguistic decoding – due to the presence of mass syntax – will be a concept

575 that is constrained to unindividuated entities (RABBIT STUFF).¹² This is in line
576 with the so-called *Cognitive Individuation Hypothesis* (Wisniewski, Lamb, &
577 Middleton, 2003), according to which the speaker's use of a count or mass
578 expression leads the hearer to construe the entity referred to as individuated or
579 unindividuated respectively.¹³ This provides a highly underspecified input to
580 pragmatic processing. Then, the pragmatic system will construct a narrower *ad*
581 *hoc* concept ('rabbit remains', 'rabbit meat', 'rabbit fur') on the basis of the
582 decoded concept, highly activated encyclopaedic information associated with it
583 (e.g., rabbits are animate creatures of flesh and blood, are edible, have fur, etc.),
584 often in combination with other contextual information derived from the
585 utterance situation (e.g. the knowledge that a convoy of trucks has just passed on
586 the highway, that Steven is a gourmet chef, etc.).

587 There are several advantages to this pragmatic account compared with a
588 standard rule-based account of systematic polysemy. First, it provides the
589 necessary interpretive flexibility for these constructions, allowing for a different
590 *ad hoc* concept to be constructed in each of (19)a.-c., in the same way as in the
591 more creative examples in (19)d.-g. below.

592

- 593 (19) d. Will a hamster bite if it senses *rabbit* on my hands? ('rabbit odour')
594 e. [Biology teacher]: *Rabbit* is smaller than hare. ('rabbit faeces')
595 f. [Hunter]: This time of year I prefer using *rabbit* ('electronic rabbit
596 calls').

¹² Here I take an approach to the grammatical count-mass distinction on which there is no lexical or semantic difference between count and mass expressions, and that all differences follow from the syntactic structure in which they occur (cf., e.g., Allan, 1980; Borer, 2005; Bunt, 1985; Pelletier, 2012). On this view, it is NPs, not nouns as such, that are the bearers of (syntactic and semantic) count and mass properties. Nouns themselves (and by extension the concepts they encode) are *underspecified* with respect to their count and mass properties. Further, assuming that the conceptual distinction between individuals and unindividuated entities is independent of count-mass syntax, our intuitions about the count or mass properties of many concepts (e.g., that the concept DOG is count but WATER is mass) may arise from mentally stored encyclopaedic (or real-world) knowledge about their denotations. This could also explain, at least in part, why some mass syntax embeddings seem less acceptable, or more 'marked', than others (e.g., ? 'I don't want *book* in this room' vs. 'I don't want *books* in this room').

¹³ There is some experimental evidence that collective mass expressions such as *furniture* are perceived as denoting individuals (e.g., Bale & Barner, 2009; Barner & Snedeker, 2005, but cf. ; Wisniewski, Imai, & Casey, 1996). This suggests that the grammatical distinction may not be a direct or perfect reflection of the conceptual distinction: sometimes there may be a conflict between linguistic conventions and real-world properties (e.g., *furniture*), or the real-world properties of a concept's denotation may allow for more than one possible construal (cf. French *meubles*).

597 g. Last winter, we discovered *rabbit*, moose and fox in our garden ('rabbit
598 tracks')

599

600 Second, while on the rule-based account an utterance such as *Sam enjoyed but*
601 *later regretted the rabbit* came out as being three-ways ambiguous as a result of
602 the operation of the universal grinder, the rules of animal meat-grinding and
603 animal fur-grinding (Copestake & Briscoe, 1995), such overgeneration does not
604 arise on the pragmatic account, where only interpretations that are consistent
605 with the hearer's expectations of relevance (i.e. achieves enough implications, at
606 a low enough processing cost) will be computed.

607 There is no doubt considerable regularity involved in polysemy of the
608 kind exemplified in (19)a.-c., where the related senses of the noun can be
609 predicted from the ontological, or 'real world', category of its denotation (e.g.,
610 animals), and the polysemy extends productively to any new members of that
611 category. This is also one of the main arguments in favour of a rule-based
612 account, where the lexical rules are seen as capturing language users' knowledge
613 of these regularities. But if such sense alternations are not generated by an
614 inventory of lexical rules, as assumed by the pragmatic account, how can we
615 explain this regularity?

616 There is clearly a tight connection between 'regular' sense alternations
617 and real-world regularities, to the extent that it seems reasonable to assume that
618 the sense alternations (whether they are regarded as linguistic or not) have their
619 origin in a number of highly regular and predictable states of affairs in the world
620 (Fodor & Lepore, 2002; Rabagliati, Markus, & Pylkkänen, 2011).¹⁴ Our general
621 knowledge of the world tells us, for instance, that there is an inherent relation
622 between an animal and its meat (or fur), and we can easily infer, upon
623 encountering a new kind of animal, that the relation also applies to this instance.
624 It seems likely that the same sort of inference would be made easily accessible to
625 us when we encounter an animal term with mass syntax (e.g., 'John loves *rabbit*'),
626 or when it occurs without a specification of its count or mass properties (e.g.,
627 'John regretted *the rabbit*'), as a result of the activation of encyclopaedic

¹⁴ Though which real-world regularities form the basis for conventional sense alternation patterns in a given language or language community may be subject to some arbitrariness.

628 knowledge associated with the concept in question (e.g., the concept RABBIT),
629 making the meat (or fur) sense easy to access or construct. Further, what may
630 start out as an *ad hoc* concept in its initial uses (e.g., the narrowing of a mass
631 occurrence of an animal-denoting noun into the meat sense) may become
632 stabilized or conventional over time within a language community as a result of
633 frequent adjustment of the lexical meaning of the word in a specific direction. In
634 such a case, the construction of the *ad hoc* concept may become progressively
635 more routinized, and a 'pragmatic routine' or inferential shortcut develop (cf.
636 Vega-Moreno, 2007), which is triggered by the activation of the concept in the
637 appropriate context (e.g., MASS OCCURRENCE OF ANIMAL TERM --> ANIMAL MEAT). Such
638 routinized inference patterns might be useful procedures in comprehension, by
639 increasing the accessibility of certain interpretations and thereby contributing to
640 a reduction of hearers' processing effort and thus to the overall relevance of the
641 utterance.¹⁵ However, rather than being part of the linguistic system, these
642 inferential short-cuts have a pragmatic basis and can easily be cancelled out by
643 contextual information (linguistic or otherwise) pointing to a different
644 interpretation. Frequent activation of these inferential routines might lead to
645 further conventionalisation of senses, and finally, in some cases, to lexicalisation.
646 An example of this may be the mass occurrence of the noun *chicken* in English,
647 whose meat sense seems conventional to the extent that it may have acquired a
648 conceptual address of its own. Thus, in this case, it is possible that we have to
649 with two linguistically encoded senses of the noun, where one has developed as a
650 result of frequent pragmatic adjustment of the other in a specific direction.
651 Notice that this is quite different from claiming that certain groups of nouns are
652 associated with lexical rules for sense extension. On this account, where the
653 development of a pragmatic routine might be one step on the way towards a new
654 lexically-stored sense, the conventional nature of many sense alternations is
655 given a wholly pragmatic explanation.

656 Finally, turning to the metaphorical and metonymical extensions in (4)
657 and (5) above, repeated here as (20) and (21), these are prime examples of
658 pragmatic processes on the relevance-theoretic account:

¹⁵ Though the current account makes no predictions regarding exactly which sense alternation patterns that end up being conventional in a given language or language community.

659

660 (20) John is a lion.

661 (21) The ham sandwich is getting impatient.

662

663 Relevance theory analyses the metaphorical use of *lion* in (20) as an instance of
664 *ad hoc* concept construction, but where the outcome would be a *broader* concept
665 than the one linguistically encoded (cf. Sperber & Wilson, 2008; Wilson &
666 Carston, 2006). The decoding of the noun *lion*, resulting in the activation of the
667 concept LION, will cause the hearer to access encyclopaedic information stored
668 about its denotation (e.g., a lion is a large cat, is tawny-coloured, is a skilled
669 hunter, is strong, courageous, takes risks, etc.). Suppose that the most
670 contextually obvious referent for *John* is the speaker's colleague (JOHN_x), who is
671 manifestly not a lion in any literal sense, but known to be a high-risk climber.
672 The encyclopaedic assumptions associated with the concept LION that are likely
673 to be added to the context in the interpretation of (20), then, would be a subset
674 of those that can be applied equally to humans (e.g., lions are strong, courageous,
675 take risks, etc.) and which contribute to the relevance of the interpretation. So
676 the hearer may broaden the concept encoded by *lion* to an *ad hoc* concept LION*
677 (paraphrasable as 'strong, courageous, takes risks, etc.'), which would denote
678 actual lions as well as those humans who possess these properties. This
679 interpretation would be a result of the hearer's mutually adjusting tentative
680 hypotheses about explicit content (JOHN_x IS A LION*), implicated premises (A LION*
681 IS STRONG, COURAGEOUS, TAKES RISKS, etc.), and implicated conclusions (JOHN_x IS
682 STRONG, COURAGEOUS, TAKES RISKS, etc.), which are incrementally modified against
683 the background of the hearer's context-specific expectations of relevance. (For
684 discussion of the mutual adjustment process in lexical interpretation, see, e.g.,
685 Carston, 2002; Sperber & Wilson, 1998; Wilson & Carston, 2006, 2007).

686 The metonymic use of *the ham sandwich* to refer to 'the person who
687 ordered a ham sandwich' in (21) can, from a relevance-theoretic perspective, be
688 seen as an instance of reference substitution based on a highly accessible
689 contextual assumption activated by the utterance situation, constrained by the
690 hearer's occasion-specific expectations of relevance. Imagine (21) being uttered
691 at a café by Jane the waitress to Sam the waiter during lunchtime, a very busy

692 time of the day. The waiters are running around trying to serve customers their
693 correct orders in time. Against this background, an anticipated conclusion of
694 Jane's utterance would be that whoever among the customers is getting
695 impatient should be served his or her food as quickly as possible. The
696 linguistically specified concept HAM SANDWICH would provide additional
697 activation to an already highly accessible contextual assumption about ham
698 sandwiches being possible orders at this café, and by a process of spreading
699 activation, about customers having ordered ham sandwiches. Let us say that at
700 the time of utterance of (21) there is only one customer waiting for his order of a
701 ham sandwich. The encoded meaning of *ham sandwich* would then activate the
702 contextual assumption 'customer *a* has ordered a ham sandwich'. The
703 interpretation of *the ham sandwich* as communicating 'the person who ordered a
704 ham sandwich' allows Sam to identify customer *a* as the referent of the
705 expression, and warrants the implicated conclusion (implicature) that customer
706 *a* should be served his food as quickly as possible. The overall inferential process
707 leading to the derivation of this implicature severely constrains the range of
708 possible associative relations that the encoded concept HAM SANDWICH may enter
709 into, and which may form the basis for the metonymic reference substitution in
710 (21).

711 The regularity associated with many metonymic uses (cf. Apresjan, 1974),
712 for instance, the development of a convention of referring to customers via their
713 food orders among the employees of a café, provides an important motivation for
714 many rule-based analyses of the phenomenon. From a relevance-theoretic
715 pragmatic point of view, this can be seen as cases where a repeated use of a
716 linguistic metonymy that links different concepts together has set up a pattern of
717 conceptual activation, or a 'pragmatic routine', which gives rise to a sense of
718 regularity (other examples may be PRODUCT FOR PRODUCER, BUILDING FOR
719 INSTITUTION, DIAGNOSIS FOR PATIENT, etc.).¹⁶

¹⁶ Such pragmatic routines have similar characteristics as the structures that cognitive linguists call 'conceptual metonymies' (first discussed by Lakoff & Johnson, 1980). An important difference between the two approaches, however, is that the pragmatic account takes the systematic conceptual correspondences to arise for communicative purposes, rather than as surface reflections of underlying conceptual metonymies. See Wilson (2009) for a discussion of this issue in connection with metaphor.

720 In this section, I have discussed an inferential, relevance-theoretic
721 approach to polysemy. On the basis of the same set of examples used to evaluate
722 rule-based accounts in section 2.1, I have argued that the pragmatic account
723 seems capable not only of handling the most context-dependent cases of
724 polysemy where rule-based accounts must appeal to pragmatics, but also that
725 part of the interpretive work that they do adequately. It also avoids many of the
726 problems associated with rule-based accounts, in particular with respect to
727 overgeneration and interpretive inflexibility. Given this, I think it remains for
728 proponents of rule-based accounts of polysemy to explain what makes the rules
729 necessary, and what is to be gained by deriving some senses in one way (via
730 lexical rules) and others in a distinct way (via pragmatics). At least, it seems that
731 considerations of theoretical economy would favour a unitary pragmatic
732 approach.

733 **3. Why polysemy?**

734 So far I have discussed two different approaches to the question of what the
735 nature of the mechanisms or processes involved in the derivation of polysemy
736 may be, and argued, on the basis of a set of standard polysemy examples, that the
737 pragmatic-inferential account seems to provide the most promising alternative
738 for a unified treatment of the data. But why do we find such proliferation of
739 polysemy in our languages in the first place? What is it about our language
740 systems, specifically their lexical component, that makes them so susceptible to
741 polysemy? In this section, I address this issue, and show how the rule-based and
742 pragmatic-inferential accounts may come up with different answers to the
743 question of what the underlying motivation for polysemy in natural language
744 may be.

745 As we have seen in the two previous sections, rule-based and pragmatic-
746 inferential accounts propose different solutions to the problem of linguistic
747 underdeterminacy: how addressees bridge the gap between (surface) linguistic
748 meanings (i.e., underspecified meanings) and speaker meanings (i.e.,
749 contextually enriched meanings). While rule-based accounts build a lot of
750 context-sensitivity into the lexicon, postulating a set of generative lexical
751 mechanisms that operates over information-rich lexical entries to yield default

752 interpretations, the pragmatic-inferential approach takes linguistic utterances to
753 provide no more than clues to speaker-intended meanings, which have to be
754 inferred on the basis of contextual evidence. It is possible to see these different
755 solutions as reflecting two distinct views on the nature of linguistic
756 underdeterminacy discussed by Carston (2002): one in which it is a form of
757 'convenient abbreviation' and another that takes it to be *essential*. I will discuss
758 these two positions and their implications for the question of polysemy
759 motivation in what follows.

760 On the first, 'convenient abbreviation' view, linguistic underdeterminacy
761 is a matter of effort-saving convenience for the hearer, possibly resulting from a
762 convention of linguistic usage or a natural drive towards communicative
763 efficiency (Carston, 2002: 29). Although sentence meaning more often than not
764 underdetermines the proposition expressed by it, a sentence that fully encodes
765 the speaker's meaning *could* in principle be supplied.¹⁷ Consider the utterance in
766 (6) again, repeated here as (22)a., and a suggestion for a fully encoding
767 counterpart in (22)b.

768

769 (22) a. She lost her bat yesterday.

770 b. Susan Thompson caused Susan Thompson to be deprived of the
771 wooden implement with a handle and solid surface used for hitting the
772 ball in baseball that belonged to Susan Thompson, between 2pm and 4pm
773 on 14 October 2013 somewhere in the Islington area, London, UK.

774

775 In order to save herself the effort of having to express a long, complex sentence
776 such as that in (22)b., the speaker can choose to use a sentence that does not
777 fully encode her intended meaning, and rely on the hearer using his pragmatic
778 capacity to turn it into a fully propositional representation. On this approach, our
779 pragmatic ability would be a useful add-on to our language capacity but would
780 not be strictly essential in enabling us to express ourselves and communicate the

¹⁷ This lies close to the view held by Quine (1960) and Katz (1972), who argued that every proposition expressed by a natural language sentence was describable in terms of a context-independent 'eternal' sentence. Katz termed this the *principle of effability*.

781 way we do. Some version of this view of linguistic underdeterminacy is likely to
782 underlie many rule-based approaches to natural language processing.

783 The second, essentialist view takes linguistic underdeterminacy to be an
784 essential feature of the relation between linguistic meanings and speaker
785 meanings (ibid.). Given the complexity and fine-grainedness of the thoughts that
786 speakers can entertain and communicate to each other (e.g., their containing
787 private references to time, space, people events and so on), they generally do not
788 lend themselves to a full encoding by natural language sentences. So although
789 the sentence in (22)b. comes closer to encoding the speaker-intended meaning
790 than the one in (22)a., any attempt to be fully explicit is bound to fail (for
791 extensive discussion of the rationale behind this position, and a defense of the
792 'ineffability of thought', see Carston, 2002: Chapter 1). On this approach, our
793 ability to make pragmatic inferences about speaker-intended meanings would
794 provide the essential foundation for our expressive and communicative abilities.
795 This is the view of linguistic underdeterminacy that underlies the pragmatic,
796 relevance-theoretic account of verbal understanding discussed in Section 2.2
797 above.

798 Both these positions provide a basis for polysemy. On the convenient
799 abbreviation view, polysemy could be motivated by a goal of economy of
800 expression, representing an effort-saving strategy for the speaker and
801 contributing to communicative efficiency. Instead of going through the laborious
802 task of fully encoding the lexical sense she has in mind – which she could do if
803 she wanted to – the speaker will often choose a more economical form of
804 expression, trusting the hearer to pragmatically infer her intended lexical
805 meaning. Although rule-based approaches do not explicitly adhere to this
806 position, it provides them with a plausible explanation for why polysemy is such
807 a pervasive phenomenon in natural language, given their postulation of
808 information-rich codes and the comparatively restricted role allocated to
809 pragmatic reasoning. However, with respect to polysemy, rule-based approaches
810 make a stronger claim about effability in that they *do* see the language as being
811 fully encoding in the canonical case, with all the information – as well the
812 procedures for manipulating it – necessary for the hearer to arrive at the
813 speaker-intended meaning being built into the lexicon, though reserving a role

814 for pragmatic reasoning in cases where an interpretation other than the default
815 one was intended. In this way, lexical sense extension rules could be seen as
816 contributing to communicative efficiency by minimising (surface) linguistic
817 complexity, requiring the hearer to use his linguistic capacity to generate default
818 compositional interpretations. In those cases where a non-default interpretation
819 is intended, it will be derived pragmatically on the basis of discourse context,
820 through a process of cancellation and substitution, but the main interpretive
821 work is done linguistically on the basis of lexical rules.

822 On the essentialist view of underdeterminacy, however, where the
823 linguistic codes that speakers make use of are not capable of fully encoding their
824 thoughts (i.e. speaker meanings) and must be supplemented by pragmatic
825 inference, polysemy would follow as a natural consequence. If the vocabularies
826 of our languages are not capable of encoding the range of concepts we can
827 entertain and communicate, polysemy – understood as the ability of words to
828 express different meanings in different contexts – would be a *necessity*.¹⁸ If this
829 view of linguistic underdeterminacy is correct, providing a full account of
830 polysemy in terms of the workings of the linguistic system should, in principle,
831 not be feasible. At least, polysemy would not have to be entirely linguistically
832 generated if communicators possess a powerful enough pragmatic-inferential
833 capacity. However, although the essential nature of underdeterminacy would be
834 the ultimate motivation for polysemy on this position, it is still compatible with
835 the idea that polysemy often represents an effort-saving strategy for
836 communicators. The proliferation of polysemy in natural languages suggests that
837 language users may find it easier to take an already existing word and extend it
838 to a new sense than to invent an entirely new word. One reason may be that the
839 stabilisation of a new word in a language is a relatively slow process that has to
840 be coordinated over a large group of individuals over time. But the typically
841 pairwise coordination involved in any given communicative act is a less
842 elaborate affair (for discussion, see Sperber & Wilson, 1998). Given our

¹⁸ It should be noted that on an account such as Pustejovsky's (1995), polysemy is an essential aspect of language, which arises as a result of generative mechanisms operating over underspecified, albeit informationally-rich, lexical representations (in the form of qualia structures). However, this is quite a different position from that of the inferential account, in which polysemy is taken to be an essential aspect of *communication*.

843 pragmatic ability to form hypotheses about speaker meanings on the basis of
844 linguistic utterances and contextual information, there would, in most cases, be
845 no need for a new word to describe something that may just as well be described
846 by using an already existing word with an extended meaning.

847 Although we find a basis for the existence of polysemy in both accounts
848 discussed here, it seems to me that it has a stronger motivation on the
849 pragmatic-inferential account, where it arises as a natural consequence of lexical
850 concepts being unable to fully encode speaker-intended concepts rather than
851 optionally as part of communicators' striving toward brevity. In the final section
852 of this paper, I will consider of some further implications of a fully pragmatic
853 approach to polysemy.

854 **4. A fully pragmatic approach: perspectives and implications**

855 In modern pragmatic theory, and relevance theory in particular, the capacity to
856 infer speaker meanings on the basis of the evidence provided is taken to be
857 reliant on the more general *theory of mind* capacity, that is, the ability to infer
858 and attribute contentful mental states to others (e.g., Baron-Cohen, 1995;
859 Premack & Woodruff, 1978; Wimmer & Perner, 1983). The claim is that the
860 theory of mind capacity provides the foundation for the kind of ostensive-
861 inferential communication that humans engage in (Sperber & Wilson, 2002); it is
862 what enables communicators to bridge the gap between characteristically
863 underspecified linguistic meanings and intended speaker meanings, including
864 inferring speaker-intended concepts from lexically-encoded concepts. As I have
865 argued in the previous section, this view provides a strong basis for polysemy,
866 with our pragmatic inferential ability playing a fundamental role in its
867 development and proliferation in verbal communication. Below I consider some
868 of its implications for acquisition, diachrony and non-verbal forms of
869 communication.

870 *4.1 Polysemy in acquisition*

871 It is widely agreed that young children's word learning requires an early capacity
872 for intention reading (e.g., Akhtar, Carpenter, & Tomasello, 1996; Baron-Cohen,
873 Baldwin, & Crowson, 1997; Bloom, 2000; Clark, 1997). Children also rely early

874 on their pragmatic ability in interpreting polysemy-related phenomena such as
875 class extensions, where a word is used in a novel lexical category (e.g., *Can you*
876 *yellow the circle?*) (Bushnell & Maratsos, 1984; Clark, 1982), metaphor (e.g.,
877 Deamer, 2013; Özçaliskan, 2005) and metonymy (Author et al., revised and
878 resubmitted)¹⁹. One study that specifically investigated children’s ability to cope
879 with systematic polysemy found an early emerging facility for using pragmatics
880 in sense resolution (Rabagliati, Marcus, & Pylkkänen, 2010). In this study, young
881 children showed an adult-like ability to make ‘licensed’ sense alternations, for
882 instance, to correctly interpret the conventional alternation between the movie
883 sense and the physical object sense of the word *DVD*, but were also more willing
884 than adults to accept ‘unlicensed’ senses, for instance, the physical object sense
885 of *movie* in *The movie is shiny*.²⁰ This suggests that in the early stages of language
886 learning, children may actually be more ‘pragmatic’ than adults in the sense that
887 they show an even greater degree of flexibility in interpretation, accepting
888 senses for words that go beyond those that adults find appropriate (or relevant).
889 (A weakness of Rabagliati et al.’s study is that the sentences containing the
890 systematic polysemy were presented to the participants in isolation, without a
891 supporting context. It is possible that if given contexts that clearly biased the
892 ‘unlicensed’ senses, adults too might have been more likely to accept these
893 senses).

894 Rabagliati et al. (2010) suggest that a rule-based account could explain
895 these results in terms of children initially possessing a broad range lexical rules,
896 generating both licensed and unlicensed senses (cf. Copestake & Briscoe, 1995),
897 some of which are ‘unlearned’ over time as a result of increasing exposure to

¹⁹ It should be mentioned here, though, that several studies have shown that children have difficulties deriving so-called ‘scalar implicatures’ (e.g., Noveck, 2001), and typically do not pass standard false-belief tasks until around the age of four (e.g., Doherty, 2000); this has led many scholars to take both pragmatics and theory of mind to be relatively late-emerging abilities. Recent studies, however, have suggested that children’s apparent pragmatic difficulty with scalar implicature may have more to do with the complexity of the tasks used or with the acquisition of various other cognitive skills than with the development of the pragmatic capacity *per se* (Barner, Brooks, & Bale, 2011; Papafragou & Tantalou, 2004; Pouscoulous, Noveck, Politzer, & Bastide, 2007). Also, recent work with non-verbal versions of the false-belief tasks (e.g., Baillargeon, Scott, & He, 2010) suggests that infants are already able to attribute false beliefs, and may have an early theory of mind ability.

²⁰ However, whether this is actually an instance of an ‘unlicensed’ sense seems questionable. For instance, as one reviewer pointed out, sentences such as *If a redbox movie is scratched, can you replace it for free?*, evoking a physical object sense of *movie*, are readily retrieved from the internet.

898 their language, a process which would involve ruling out infrequent senses from
899 their lexical repertoire. For instance, if children never hear the word *movie* used
900 in a physical object sense, they might discount the probability of there being a
901 rule that creates physical objects from abstract entities in their language. I find
902 this hypothesis very implausible, given the range of other polysemy-related
903 phenomena that children are able to cope with from a very early age, which are
904 unlikely to require the presence of any lexical rules for interpretation. This early
905 pragmatic competence should enable children to cope with cases of systematic
906 polysemy too, although it may sometimes overgenerate. Furthermore, one would
907 have to explain how such initially too broad rules develop quite independently of
908 the child's conceptual knowledge, given the tight connection that clearly exists
909 between systematic sense alternations and real-world regularities in the
910 language of adults. Another possible rule-based explanation could be that
911 children start out by being radically pragmatic, but acquire the lexical rules at a
912 later stage in development as a result of exposure both to their language and to
913 the real-world relations on the basis of which the lexical rules arise. While this
914 hypothesis seems more plausible, one would have to offer a reason for why the
915 'rules' must be part of the linguistic system and not rather an artefact of
916 conceptual organisation, reflecting highly predictable and regular states of
917 affairs in the world, where children's developing ability to handle systematic
918 polysemy could emerge as a by-product of acquiring a more adult-like
919 conceptual organisation and pragmatic competence.

920 *4.2 Polysemy in diachrony*

921 If the fully pragmatic account is on the right track, all synchronic instances of
922 polysemy should, in principle, be traceable back to the operation of a pragmatic
923 process. A prevalent hypothesis about semantic change is that its main driving
924 force is pragmatic, being motivated by speaker-hearer interactions and
925 communicative strategies (e.g., Traugott & Dasher, 2002). It is also widely
926 thought that semantic change must go through a stage of polysemy, in which
927 related meanings of a word that emerged at historically different periods coexist
928 over time in a language, both in individual speakers and in language
929 communities (Hopper, 1991). While rule-based accounts generally have little to

930 say about the historical development of polysemy, these hypotheses about
931 semantic change are quite compatible with the pragmatic-inferential account of
932 polysemy proposed here. For instance, what may start out as an *ad hoc* concept
933 in one stage may become stabilised or conventional over time for individual
934 speakers or within a language community, as a result of frequent adjustment of a
935 lexical meaning in a specific direction (e.g., a narrowing of the mass sense of
936 *chicken* into a meat sense, cf. Section 2.2). In such a case the construction of the
937 *ad hoc* concept may become progressively more routinized until an inferential
938 shortcut (or ‘pragmatic routine’) develops (cf. Vega-Moreno, 2007). A possible
939 development from there might be a lexicalisation of the new sense, and
940 potentially, that the original *ad hoc* concept takes over from the originally
941 encoded concept.²¹

942 At the synchronic level, then, individual speakers may differ with regard
943 to which senses of a word they have stored in their mental lexicons. For instance,
944 for some speakers of English, the broadened, metaphorical sense of *lion*
945 (discussed in Section 2.2) may be conventional and thus lexically stored.
946 Recognising this concept as the one intended by a speaker on a given occasion of
947 use would then be a matter of disambiguation rather than *ad hoc* concept
948 construction. For other speakers, even within the same speech community, *lion*
949 may have only one encoded sense (the animal sense), and the broadened,
950 metaphorical sense would be derived pragmatically through *ad hoc* concept
951 construction. So the construction of a particular *ad hoc* concept may be an
952 occasional, or even first time, affair for one communicator and a routine pattern
953 for another. In this way, pragmatic inference serves an important function in
954 compensating for such differences among members of a language community,

²¹ One reviewer objected that if all polysemy is a result of a process of meaning extension, this would predict that there should always be a gap in history between when different senses of a word enter a language, but that this is likely not to be the case. For instance, in the case of the noun *door*, which has both a physical object sense (‘He painted the door’) and an aperture sense (‘He walked through the door’), the different senses may have arisen at precisely the same times. While it is an empirical question exactly when specific senses arise in a language, I agree that cases such as *door* (and similar cases such as a *window*, *entrance*, *jalousie*, *portal*, etc.) are special in the sense that it is difficult to see how one sense could have been derived from the other. However, I still think that these (apparent) sense alternations can be given a pragmatic explanation, in terms of increased activation of certain aspects of encyclopaedic knowledge associated with the denotations (cf. Langacker’s 1984 notion of ‘active zones’). If something like this is true, there may be more than one route by which senses of a word may arise.

955 enabling them to end up with the same lexical senses but in many cases via
956 distinct routes (see Wilson & Carston, 2007, for further discussion).

957 *4.3 Non-verbal polysemy?*

958 Another implication of treating polysemy as a fundamentally communicative
959 phenomenon in this way is that, in principle, we should expect not just words but
960 any simple ostensive stimulus, for instance, manual and facial gestures, to be
961 susceptible to polysemy. One example might be the use that car drivers make of a
962 single flash of their headlights to another driver, which could (in England), mean
963 at least the following: (a) 'go ahead in front of me'; (b) 'thanks for giving way to
964 me'; (c) 'watch out – there is traffic police up ahead'. The general meaning of this
965 signal might be something like 'friendliness indication', with the more specific
966 meanings derivable on the basis of context (e.g., positioning of the cars on the
967 road, whether or not one of the drivers has already flashed his/her lights, etc.).²²
968 Another example might be smiles, which, depending on the context, can
969 communicate a range of related feelings: amusement, affection, sympathy, and so
970 on (cf. Wharton, 2009). An example from preverbal communication might be a
971 toddler extending his or her hands upwards, which, depending on the toddler's
972 position, could convey several different meanings: [*on floor*] 'I want you to pick
973 me up'; [*in high chair*] 'Help me down'; [*under object of interest*] 'Get me to that
974 object up there', and so on. The gesture itself could be seen as having a general
975 meaning along the lines of 'take me from one place to another', with the more
976 specific actions intended by the toddler being contextually-determined.

977 **5. Final remarks**

978 The central topic of this paper has been whether the aspects of meaning that are
979 involved in the construction of polysemy have a primarily linguistic or non-
980 linguistic basis, and the extent to which its proliferation and development in
981 natural language can be explained given each of these views. More specifically,
982 the question has been whether polysemy results primarily from the operation of
983 lexicon-internal processes or from pragmatic-inferential processes applying at
984 the level of individual words. Some people may argue that the difference

²² Thanks to Robyn Carston for this example.

985 between these two accounts is one of degree only, and simply concerns how
986 large a role pragmatics should play in the theory: While rule-based approaches
987 maintain that a considerable amount of linguistic knowledge is involved in the
988 generation of polysemy, with pragmatics taking over when an interpretation
989 other than the linguistic default meaning is intended, the wholly pragmatic
990 approach downplays the linguistic aspect and suggests that polysemy arises
991 mainly as a result of the operation of pragmatic processes over underspecified
992 lexical meanings, taking contextual information and encyclopaedic assumptions
993 about conceptual denotations as input to the inferential process. However, I
994 think that the difference between these two approaches is much more
995 fundamental than this, and involves two radically different conceptions of what a
996 language is, and the role it plays in the communication process. As we have seen,
997 rule-based approaches treat the language as providing a rich code that enables
998 communicators to encode and decode their thoughts in much detail, and
999 pragmatics merely as a useful add-on to this capacity. By contrast, the pragmatic
1000 account takes the role of the linguistic system to be that of providing a minimal
1001 input or clue, which the inferential system – seen as the essential foundation for
1002 our expressive and communicative abilities – uses as evidence to yield
1003 hypotheses about occasion-specific, speaker-intended senses. I have argued
1004 throughout this paper that the assumption that a large part of the interpretive
1005 work involved in the processing of polysemy should be attributed to the
1006 linguistic system itself requires independent justification, given that we have an
1007 independently motivated pragmatic interpretation system capable of rapidly
1008 generating new senses in contexts. The pragmatic-inferential approach predicts
1009 that lexical items are used to express a variety of occasion-specific senses, which
1010 include but go far beyond the default senses predicted by rule-based approaches.

1011 Finally, I have suggested that the proliferation of polysemy appears to
1012 have a stronger motivation on the pragmatic-inferential account, where it arises
1013 as a natural consequence of lexical meanings not being able to fully encode
1014 speaker-intended senses. Viewing polysemy as a fundamentally communicative
1015 phenomenon in this way allows us to provide a unified account of its role in
1016 several domains, including acquisition, diachrony and non-verbal forms of
1017 communication.

1018 **References**

- 1019 Akhtar, N., Carpenter, M., & Tomasello, M. (1996). The role of discourse novelty
1020 in early word learning. *Child Development*, 67, 635-645.
- 1021 Allan, K. (1980). Nouns and countability. *Language*, 56(3), 541-567.
- 1022 Apresjan, J. D. (1974). Regular polysemy. *Linguistics*, 14(2), 5-32.
- 1023 Asher, N. (2011). *Lexical Meaning in Context*. Cambridge: Cambridge University
1024 Press.
- 1025 Asher, N., & Lascarides, A. (2003). *Logics of Conversation*. Cambridge: Cambridge
1026 University Press.
- 1027 Author et al. (revised and resubmitted). 'The moustache sits down first': On the
1028 acquisition of metonymy.
- 1029 Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants.
1030 *Trends in Cognitive Science*, 14(3), 110-118.
- 1031 Bale, A. C., & Barner, D. (2009). The interpretation of functional heads: Using
1032 comparatives to explore the mass/count distinction. *Journal of Semantics*,
1033 26, 217-252.
- 1034 Barner, D., Brooks, N., & Bale, A. (2011). Assessing the unsaid: The role of scalar
1035 alternatives in children's pragmatic inference. *Cognition*, 118, 84-93.
- 1036 Barner, D., & Snedeker, J. (2005). Quantity judgments and individuation:
1037 evidence that mass nouns count. *Cognition*, 97(1), 41-66.
- 1038 Baron-Cohen, S. (1995). *Mindblindness: An Essay on Autism and Theory of Mind*.
1039 Cambridge, Massachusetts: The MIT Press.
- 1040 Baron-Cohen, S., Baldwin, D. A., & Crowson, M. (1997). Do children with autism
1041 use the speaker's direction of gaze strategy to crack the code of language?
1042 *Child Development*, 68(1), 48-57.
- 1043 Bloom, P. (2000). *How Children Learn the Meaning of Words*. Cambridge,
1044 Massachusetts: The MIT Press.
- 1045 Blutner, R. (2002). Lexical semantics and pragmatics. In F. Hamm & T. E.
1046 Zimmermann (Eds.), *Semantics* (Vol. 10, pp. 27-58). Hamburg: Helmut
1047 Buske Verlag.
- 1048 Borer, H. (2005). *In Name Only*. Oxford: Oxford University Press.
- 1049 Bosch, P. (2007). Productivity, polysemy and predicate indexicality. In H. Zeevat
1050 & B. ten Cate (Eds.), *Proceedings of the Sixth International Tbilisi*
1051 *Symposium on Language, Logic and Computation*. Heidelberg/Berlin:
1052 Springer.
- 1053 Bréal, M. (1924 [1897]). *Essai de sémantique: science des significations*. Paris:
1054 Gérard Monfort.
- 1055 Briscoe, T., & Copestake, A. (1991). Sense extensions as lexical rules *Proceedings*
1056 *of the IJCAI Workshop on Computational Approaches to Non-Literal*
1057 *Language* (pp. 12-20). Sydney, Australia.
- 1058 Brugman, C. (1988). *The Story of over: Polysemy, Semantics, and the Structure of*
1059 *the Lexicon*. New York: Garland.
- 1060 Brugman, C., & Lakoff, G. (1988). Cognitive topology and lexical networks. In S.
1061 Small, G. Cottrell & M. Tannenhaus (Eds.), *Lexical Ambiguity Resolution:*
1062 *Perspectives from Psycholinguistics, Neuropsychology and Artificial*
1063 *Intelligence* (pp. 477-507). San Mateo, California: Morgan Kaufman.
- 1064 Bunt, H. (1985). *Mass Terms and Model-Theoretic Semantics*. Cambridge:
1065 Cambridge University Press.

- 1066 Bushnell, E. W., & Maratsos, M. P. (1984). "Spooning" and "basketing": Children's
 1067 dealing with accidental gaps in the lexicon. *Child Development*, 55, 893-
 1068 902.
- 1069 Caramazza, A., & Grober, E. (1976). Polysemy and the structure of the subjective
 1070 lexicon. In C. Rameh (Ed.), *Semantics: Theory and Application* (pp. 181-
 1071 206). Washington: Georgetown University Round Table on Languages and
 1072 Linguistics.
- 1073 Carston, R. (2002). *Thoughts and Utterances: The Pragmatics of Explicit*
 1074 *Communication*. Oxford: Blackwell Publishers.
- 1075 Carston, R. (2012). Word meaning and concept expressed. *The Linguistic Review*,
 1076 29(4), 607-623.
- 1077 Carston, R. (2013). Word meaning, what is said and explicature *What is Said and*
 1078 *What is Not* (pp. 175-204). Stanford: CSLI Publications.
- 1079 Clark, E. V. (1982). The young word maker: A case study of innovation in the
 1080 child's lexicon. In E. Wanner & L. Gleitman (Eds.), *Language Acquisition:*
 1081 *the State of the Art* (pp. 390-425). Cambridge: Cambridge University
 1082 Press.
- 1083 Clark, E. V. (1997). Conceptual perspective and lexical choice in acquisition.
 1084 *Cognition*, 64, 1-37.
- 1085 Copestake, A., & Briscoe, T. (1992). Lexical operations in a unification-based
 1086 framework. In J. Pustejovsky & S. Bergler (Eds.), *Lexical Semantics and*
 1087 *Knowledge Representation, Proceedings of the First SIGLEX Workshop* (pp.
 1088 101-119). Berlin: Springer-Verlag.
- 1089 Copestake, A., & Briscoe, T. (1995). Semi-productive polysemy and sense
 1090 extension. *Journal of Semantics*, 12(1), 15-67.
- 1091 de Almeida, R. G. (2004). The effect of context on the processing of type-shifting
 1092 verbs. *Brain and Language*, 90(1-3), 249-261.
- 1093 de Almeida, R. G., & Dwivedi, V. D. (2008). Coercion without lexical
 1094 decomposition: Type-shifting effects revisited. *Canadian Journal of*
 1095 *Linguistics*, 53(2/3), 301-326.
- 1096 Deamer, F. (2013). *An investigation into the processes and mechanisms underlying*
 1097 *the comprehension of metaphor and hyperbole*. PhD dissertation:
 1098 University College London.
- 1099 Doherty, M. (2000). Children's understanding of homonymy: Metalinguistic
 1100 awareness and false belief. *Journal of Child Language*, 27, 367-392.
- 1101 Falkum, I. L. (2007). Generativity, relevance and the problem of polysemy. *UCL*
 1102 *Working Papers in Linguistics*, 19, 205-234.
- 1103 Fodor, J. A. (1975). *The Language of Thought*. Cambridge University Press:
 1104 Harvard University Press.
- 1105 Fodor, J. A. (2008). *LOT2: The Language of Thought Revisited*. Oxford: Oxford
 1106 University Press.
- 1107 Fodor, J. A., & Lepore, E. (2002). *The Compositionality Papers*. Oxford: Clarendon
 1108 Press.
- 1109 Geeraerts, D. (1993). Vagueness's puzzles, polysemy's vagaries. *Cognitive*
 1110 *Linguistics*, 4(3), 223-272.
- 1111 Gibbs, R. W. (1994). *The Poetics of Mind*. Cambridge: Cambridge University Press.
- 1112 Glucksberg, S. (2001). *Understanding Figurative Language: From Metaphors to*
 1113 *Idioms*. New York: Oxford University Press.

- 1114 Goddard, C. (2000). Polysemy: a problem of definition. In Y. Ravin & C. Leacock
 1115 (Eds.), *Polysemy: Theoretical and Computational Approaches* (pp. 129-
 1116 151). Oxford: Oxford University Press.
- 1117 Grice, H. P. ([1967] 1989). Logic and conversation *Studies in the Way of Words*.
 1118 Cambridge, Massachusetts: Harvard University Press.
- 1119 Hopper, P. J. (1991). On some principles of grammaticization. In E. C. Traugott &
 1120 B. Heine (Eds.), *Approaches to Grammaticalization, Vol. 1* (pp. 17-35).
 1121 Amsterdam: John Benjamins.
- 1122 Katz, J. J. (1964). Semantic theory and the meaning of 'good'. *The Journal of*
 1123 *Philosophy*, 61(23), 739-766.
- 1124 Katz, J. J. (1972). *Semantic Theory*. New York: Harper & Row.
- 1125 Kempson, R. M. (1977). *Semantic Theory*. Cambridge: Cambridge University
 1126 Press.
- 1127 Klein, D. E., & Murphy, G. L. (2001). The representation of polysemous words.
 1128 *Journal of Memory and Language*, 45(2), 259-282.
- 1129 Klepousniotou, E. (2007). Reconciling linguistics and psycholinguistics: On the
 1130 psychological reality of linguistic polysemy. In M. Rakova, G. Petho & C.
 1131 Rakosi (Eds.), *The Cognitive Basis of Polysemy: New Sources of Evidence for*
 1132 *Theories of Word Meaning* (pp. 17-46). Frankfurt: Peter Lang Verlag.
- 1133 Klepousniotou, E., Titone, D., & Romero, C. (2008). Making sense of word senses:
 1134 the comprehension of polysemy depends on sense overlap. *Journal of*
 1135 *Experimental Psychology: Learning, Memory, and Cognition*, 34(6), 1534-
 1136 1543.
- 1137 Lakoff, G. (1987). *Women, Fire, and Dangerous Things: What Categories Reveal*
 1138 *about the Mind*. Chicago, Illinois: The University of Chicago Press.
- 1139 Lakoff, G., & Johnson, M. (1980). *Metaphors We Live by*. Chicago, Illinois: The
 1140 University of Chicago Press.
- 1141 Langacker, R. W. (1984). Active zones. *Proceedings of the annual meeting of the*
 1142 *Berkeley Linguistics Society*, 10, 172-188.
- 1143 Lascardes, A., & Copestake, A. (1998). Pragmatics and word meaning. *Journal of*
 1144 *Linguistics*, 34, 387-414.
- 1145 McElree, B., Frisson, S., & Pickering, M. J. (2006). Deferred interpretations: Why
 1146 starting Dickens is taxing but reading Dickens isn't. *Cognitive Science: A*
 1147 *Multidisciplinary Journal*, 30(1), 181-192.
- 1148 McElree, B., J., Matthew, Traxler, M. J., Pickering, M. J., Seely, R. E., & Jackendoff, R.
 1149 (2001). Reading time evidence for enriched composition. *Cognition*, 78,
 1150 17-25.
- 1151 Noveck, I. (2001). When children are more logical than adults: experimental
 1152 investigations of scalar implicature. *Cognition*, 78, 165-188.
- 1153 Nunberg, G. (1979). The non-uniqueness of semantic solutions: Polysemy.
 1154 *Linguistics and Philosophy*, 3(2), 143-184.
- 1155 Ostler, N., & Atkins, B. T. S. (1992). Predictable meaning shift: Some linguistic
 1156 properties of lexical implication rules. In J. Pustejovsky & S. Bergler (Eds.),
 1157 *Lexical Semantics and Knowledge Representation, Proceedings of the First*
 1158 *SIGLEX Workshop* (pp. 87-100). Berlin: Springer.
- 1159 Özçaliskan, S. (2005). On learning how to draw the distinction between physical
 1160 and metaphorical motion: is metaphor and early emerging cognitive and
 1161 linguistic capacity? *Journal of Child Language*, 32, 291-318.
- 1162

- 1163 Papafragou, A., & Tantalou, N. (2004). Children's computation of implicatures.
 1164 *Language Acquisition*, 12(1), 71-82.
- 1165 Peirce, C. S. (1955). Logic as semiotic: The theory of signs. In J. Buchler (Ed.),
 1166 *Philosophical writings of Peirce* (pp. 98-119). New York: Dover Press.
- 1167 Pelletier, F. J. (1975). Non-singular reference: some preliminaries. *Philosophia*,
 1168 5(4), 451-465.
- 1169 Pelletier, F. J. (2012). Lexical nouns are both +MASS and +COUNT, but they are
 1170 neither +MASS nor +COUNT. In D. Massam (Ed.), *Count and Mass Across*
 1171 *Languages* (pp. 9-26). Oxford: Oxford University Press.
- 1172 Pickering, M. J., McElree, B., & Traxler, M. J. (2005). The difficulty of coercion: A
 1173 response to de Almeida. *Brain and Language*, 93, 1-9.
- 1174 Pouscoulous, A., Noveck, I., Politzer, G., & Bastide, A. (2007). A developmental
 1175 investigation of processing costs in implicature production. *Language*
 1176 *Acquisition*, 14(4), 347-375.
- 1177 Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind?
 1178 *Behavioral and Brain sciences*, 1, 515-526.
- 1179 Pustejovsky, J. (1991). The generative lexicon. *Computational Linguistics*, 17(4),
 1180 409-441.
- 1181 Pustejovsky, J. (1995). *The Generative Lexicon*. Cambridge, Massachusetts: The
 1182 MIT Press.
- 1183 Pustejovsky, J. (1998). Generativity and explanation in semantics. Reply to Fodor
 1184 and Lepore. *Linguistic Inquiry*, 29, 289-311.
- 1185 Pyllkkänen, L., Llinás, R., & Murphy, G. L. (2006). The representation of polysemy:
 1186 MEG evidence. *Journal of Cognitive Neuroscience*, 18(1), 97-109.
- 1187 Quine, W. V. O. (1960). *Word and Object*. Cambridge, Massachusetts: The MIT
 1188 Press.
- 1189 Rabagliati, H., Marcus, G. F., & Pyllkkänen, L. (2010). Shifting senses in lexical
 1190 semantic development. *Cognition*, 117, 17-37.
- 1191 Rabagliati, H., Marcus, G. F., & Pyllkkänen, L. (2011). Rules, radical pragmatics
 1192 and restrictions on regular polysemy. *Journal of Semantics*, 28, 485-512.
- 1193 Ravin, Y., & Leacock, C. (2000). Polysemy: An overview. In Y. Ravin & C. Leacock
 1194 (Eds.), *Polysemy: Theoretical and Computational Approaches* (pp. 1-29).
 1195 Oxford: Oxford University Press.
- 1196 Ruhl, C. (1989). *On Monosemy: A Study in Linguistic Semantics*. Albany, New York:
 1197 State University of New York Press.
- 1198 Saussure, F. de. (1974). *Course in General Linguistics*. Translated from the French
 1199 (1916) by Wade Baskin. Glasgow: Collins.
- 1200 Shannon, C. E., & Weaver, W. (1949). *The Mathematical Theory of Communication*.
 1201 Urbana, Illinois: University of Illinois Press.
- 1202 Sperber, D., & Wilson, D. (1986/1995). *Relevance: Communication and Cognition*.
 1203 Oxford: Blackwell.
- 1204 Sperber, D., & Wilson, D. (1998). The mapping between the mental and the public
 1205 lexicon. In P. Carruthers & J. Boucher (Eds.), *Language and Thought* (pp.
 1206 184-200). Cambridge: Cambridge University Press.
- 1207 Sperber, D., & Wilson, D. (2002). Pragmatics, modularity and mind-reading. *Mind*
 1208 *& Language*, 17, 3-23.
- 1209 Sperber, D., & Wilson, D. (2008). A deflationary account of metaphor. In R. W.
 1210 Gibbs (Ed.), *The Cambridge Handbook of Metaphor and Thought* (pp. 84-
 1211 108). Cambridge: Cambridge University Press.

- 1212 Taylor, J. R. (2003). Polysemy's paradoxes. *Language Sciences*, 25(6), 637-655.
- 1213 Traugott, E. C., & Dasher, R. B. (2002). *Regularity in Semantic Change*. Cambridge:
1214 Cambridge University Press.
- 1215 Traxler, M. J., Pickering, M. J., & McElree, B. (2002). Coercion in sentence
1216 processing: evidence from eye-movements and self-paced reading.
1217 *Journal of Memory and Language*, 47, 530-547.
- 1218 Tyler, A., & Evans, V. (2003). *The Semantics of English Prepositions*. Cambridge:
1219 Cambridge University Press.
- 1220 Vega-Moreno, R. (2007). *Creativity and Convention: The Pragmatics of Everyday*
1221 *Figurative Speech*. Amsterdam: John Benjamins.
- 1222 Vygotsky, L. (1986). *Thought and Language*. Translated from the Russian (1934)
1223 by Alex Kozulin. Cambridge, Massachusetts: The MIT Press.
- 1224 Wharton, T. (2009). *Pragmatics and Non-Verbal Communication*. Cambridge:
1225 Cambridge University Press.
- 1226 Wilson, D. (2009). Parallels and differences in the treatment of metaphor and
1227 Relevance theory and Cognitive Linguistics. *Studies in Pragmatics*, 11, 42-
1228 60.
- 1229 Wilson, D., & Carston, R. (2006). Metaphor, relevance and the 'emergent
1230 property' issue. *Mind & Language*, 21(3), 404-433.
- 1231 Wilson, D., & Carston, R. (2007). A unitary approach to lexical pragmatics:
1232 Relevance, inference and ad hoc concepts. In N. Burton-Roberts (Ed.),
1233 *Pragmatics* (pp. 230-259). London: Palgrave.
- 1234 Wilson, D., & Sperber, D. (2004). Relevance theory. In L. R. Horn & G. L. Ward
1235 (Eds.), *The Handbook of Pragmatics* (pp. 607-632). Oxford: Blackwell.
- 1236 Wilson, D., & Sperber, D. (2012). *Meaning and Relevance*. Cambridge: Cambridge
1237 University Press.
- 1238 Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and
1239 constraining function of wrong beliefs in young children's understanding
1240 of deception. *Cognition*, 13(1), 103-128.
- 1241 Wisniewski, E., Imai, M., & Casey, L. (1996). On the equivalence of superordinate
1242 concepts. *Cognition*, 60, 269-298.
- 1243 Wisniewski, E., Lamb, C., & Middleton, E. (2003). On the conceptual basis for the
1244 count and mass noun distinction. *Language and Cognitive Processes*, 18(5-
1245 6), 583-624.
- 1246
- 1247
- 1248