A denotational account for phenomena of polysemy

King's College London Workshop on Word Meaning, Cumberland Lodge, Windsor
24-25 April, 2014

Peter Bosch
Institute of Cognitive Science
University of Osnabrück

The plan

1. introductory remarks on modularity and theory
2. language processing and its resources
3. rich and poor models of lexical meaning
4. indexicality and contextual modulations of "word sense"
5. kinds of contextual modulation
6. modelling contextual modulation as concept modulation
7. computing contextual concepts
8. some applications, examples, and details

1. Modularity and theory

I assume that linguistic knowledge, i.e., the knowledge a speaker has of his mother language, constitutes a module of cognition.

By this I mean that
1. There is an interesting set of regularities within linguistic knowledge that can be captured with the parameters of linguistic knowledge alone, and with no reference to other modules of cognition. I.o.w., the notion of a linguistic theory as a theory of linguistic knowledge makes sense.
2. A theory of linguistic knowledge is a theory of the linguistic resources required in human language processing. It is not itself a theory of language processing.
3. A theory of language processing requires a theory of linguistic knowledge, plus theories of conceptual representation, attention, perception, action, memory, etc.

2. Interactive processing of linguistic and visual information – Visual World eye tracking

German default descriptions:

- der (masc) gelbe Hut (neutral) the yellow hat
- die (fem) blaue Rakete (nom) the blue rocket
- der (masc) blaue Stern (neutral) the blue star
- das (neut) gelbe Hufeisen (nominative) the yellow horse shoe

Klicken Sie auf die blaue Rakete.

click on ... [followed by a def. determiner, adjective, and noun] Hartmann 2005
2. The role of linguistic knowledge

The denotation of the definite determiner is commonly modelled as a partial function, defined only for a domain in which its nominal complement has a unique denotation (Heim & Kratzer 1998):

\[ f \in D_{<e,t>} \land \exists y. f(y) = 1 \land f(x) = 1 \]

Suppose this is actually part of an entry in the mental lexicon. Then the determiner could not have any processing effect, as long as \( f \) is not known, i.e., not before the offset of the NP.

But some information about \( f \) is already available with the determiner: the gender of the basic level noun denoting \( f \), and the experiment shows that this information is used immediately.

2. Knowledge available to the processor

Let’s look at the information the processor has available:

Visual information from the display limits the referential choice to 4 objects, for each of which linguistic experience provides the knowledge of basic level common nouns, linguistic knowledge contributes the determiner denotation and the information that only one of the basic level nouns is gender-congruent with the determiner.

A lesson from this kind of observation is that lexical information, at least occasionally, contributes very little to the results of language understanding, simply because other “contextual” resources, are already available earlier, or simultaneously.

So, perhaps, we can make do with much more attenuate lexical meanings – provided they are more capable to interact with other knowledge resources?

2. Knowledge available to the processor

So the processing effect of the determiner may be explained by postulating

1. the knowledge the processor has of the German basic level common nouns for all objects in the display, incl. minimal lexical information:
   - Hut [gen:masc], [sem:Hut]
   - Rakete [gen:fem], [sem:Rakete]
   - Stern [gen:masc], [sem:Stern]
   - Hufeisen [gen:neut], [sem:Hufeisen]

2. the identification of exactly one display object as an instance of [Rakete], presupposing sufficient conceptual and visual information

3. and a lexical entry for the determiner *die* including gender information:
   \[ f \in D_{<e,t>} \land \exists y. f(y) = 1 \land g(BCN(y), fem) = 1 \land f(x) = 1 \]

But only some of this is linguistic knowledge!

1. the knowledge the processor has of the German basic level common nouns for all objects in the display, incl. minimal lexical information:
   - Hut [gen:masc], [sem:Hut]
   - Rakete [gen:fem], [sem:Rakete]
   - Stern [gen:masc], [sem:Stern]
   - Hufeisen [gen:neut], [sem:Hufeisen]

2. the identification of exactly one display object as an instance of [Rakete], presupposing sufficient conceptual and visual information

3. and a lexical entry for the determiner *die* including gender information:
   \[ f \in D_{<e,t>} \land \exists y. f(y) = 1 \land g(BCN(y), fem) = 1 \land f(x) = 1 \]

2. Immediate evaluation to the utterance context

In Visual World eye tracking we can observe that referential expressions are evaluated to the visual utterance context as soon as sufficient information is available, and before sentence meanings, or even constituent meanings, are computed.

The evaluation of referential expressions makes use of visual and linguistic input, linguistic knowledge, conceptual knowledge, world knowledge.

A surprise for the linguist may be that the lexical information that comes with the nominal constituent, *blaue Rakete*, in the case discussed, is actually completely redundant.

3. Lexical meaning - poorer, but more sociable

A lesson from this kind of observation is that lexical information, at least occasionally, contributes very little to the results of language understanding, simply because other "contextual" resources, are already available earlier, or simultaneously.

So, perhaps, we can make do with much more attenuate lexical meanings – provided they are more capable to interact with other knowledge resources?
3. Lexical meaning in printed dictionaries

1. "to open" (at Oxforddictionaries.com – 18 April, 2014):
   1.1 [object] Move (a door or window) so as to leave a space allowing access and vision: she opened the door and went in [no object, in imperative]: ‘Open up!’ he said
   1.2 Undo or remove the lid, cover, or fastening of (a container, package, letter, etc.) to get access to the contents: he opened a bottle inexpertly, spilling some of the wine as he opened the presents now?
   1.3 Part the lips or lids of (one’s mouth or eye): she opened her mouth to argue
   1.4 [no object] (Of the mouth or eyes) have the lips or lids parted: her eyes slowly opened
   1.5 [no object] Come apart; lose or lack its protective covering: old wounds opened and I bled a little bit
   1.6 [no object] (open on/to/into) (Of a room, door, or window) give access to: the kitchen opened into a pleasant sitting room
   1.7 Cause evacuation of (the bowels) of (an enterprise, meeting, or event) begin or be formally established: two new restaurants opened this week
   1.8 [with object] (Of a book or file) to read it: she opened her book at the prologue
   1.9 [no object] (open out) Become wider: the path opened out into a glade
   1.10 [with object] (open out) Become wider: the path opened out into a glade

3. Lexical meaning in printed dictionaries

2.3 [no object, with adverbial] (Of a prospect) extend into view: stop to marvel at the views that open out below

2.4 [with object] Nautical Achieve a clear view of (a place) by sailing past a headland or other obstruction: we shall open Torbay shortly

3. Make or become formally ready for customers, visitors, or business: [with object]: the shops didn’t open until 10

2.1 A Break the conducting path of (an electric circuit):

3.1 [with object] Ceremonially declare (a building, road, etc.) to be completed and ready for use: the Queen opened the power plant on 17 October 1956

3.2 [with object] Make possible access to or passage through: the President announced that his government would open the border

4. Formally establish or begin (a new business, movement, or enterprise): she began to teach and opened her own school

5. We opened a branch in Madrid

6. announced that his government would open the border

3. Lexical meaning – poor, but sociable

For the English transitive verb "to open" the semantic information is no more than what is in the SEM substructure of the following lexical entry:

3. Lexical meaning – poor, but sociable

4. Contextual modulation of "word sense"

The kind of modulation that I am interested in

- is a form of "sense" modulation that is made and understood automatically and with no effort,
- typically remains unnoticed by the language user, and
- is productive, i.e., it can yield infinitely many variants of arbitrary fine granularity

Typical cases:

The {watch / tap / lecture / …} is running.
Fred {read, published, found} a novel.

This is a fast {run / bike / solution /…}.

This is a fast {run / bike / solution /…}.

Fred {read, published, found} a novel.
I am here now.
You could have been there yesterday.
This is a computer.
It is raining.

Intuitively speaking, indexical expressions refer differently in each utterance context, but hold their meanings constant across contexts.

Sentences that contain indexicals have potentially different content, i.e., say something different, in each context, but do not change their meaning across contexts.

Kaplan’s sentence meanings contain index variables that are evaluated at the utterance context to yield the appropriate truth conditions for the utterance. This conception is primarily geared to account for the context-dependent variation in the reference of indexical expressions.

- Can the same model be used for context dependence quite generally?
- Can it also account for the context-dependent variation that we find in polysemy?
5. Productive modulation (ii)
where the semantic value of an implicit argument is sensitive to context

(1) I live nearby. place
(2) She saw an enemy. person
(3) The office is left of the entrance. perspective
(4) That was a fast run. comparison class
(5) It's raining. place, time

Open questions:
- How do we know which expressions have implicit arguments?
- How do we find the relevant argument in the context?

5. Productive modulation (iii)
where neither explicit nor implicit arguments seem involved: “free” contextual modulation

- Where is Fred?
  (1) He's working. WORK(fred) → \( \varphi \) (LOCATION(fred))
- How can Fred afford these expensive holidays?
  (2) He's working. WORK(fred) → \( \psi \) (WEALTH(fred))
- Can I speak to Fred, please?
  (3) He's working. WORK(fred) → \( \sigma \) (AVAILABILITY(fred))

5. Productive modulation (iii)
where neither explicit nor implicit arguments seem involved: “free” contextual modulation

The modulation is Inferentially and hence truth-conditionally relevant.
Nothing follows about Fred's location when Fred is working.
Fred is working.
Fred is working and so is Pete.
Fred is in his office and Pete can afford expensive holidays.

The modulation is stable within the utterance context:
Fred is working and so is Pete.
cannot be interpreted as, e.g.,
Fred is in his office and Pete can afford expensive holidays.

6. What is it that is being modulated in productive modulation?

- not lexical meanings (characters)
  - because the variation is productive and correlates with variation in the context
- but semantic values (contents, denotations)

For unsaturated expressions the denotations are

Contextual Concepts

(Cf. Frege’s notion that the values of “predicates” are concepts.)

6. Contextual Concepts

Contextual Concepts (CCs)
- are the values of character functions, applied to a set of context parameters as arguments
- they are themselves partial functions that are defined for all and only arguments in the intended domain (they "live" in their own context and do not survive it)
- they may be arbitrarily abstract or concrete and may contain context-specific and other contingent information
6. Contextual Concepts

**Contextual Concepts (CCs)**

- are necessarily stable within their context (because they are values of functions that only have contexts as arguments), and can be anaphorically resumed within that context (they are contextual entities).

Thus it is CCs – not meanings or senses – that remain constant in VP anaphora, VP ellipsis, or VP repetition:

Fred is working, and so is Pete.
Fred is working, and Pete Ø too.
Fred is working, and Pete is working.

Thus it is CCs that define the units of counting:

[talking of Fred, Pete, and a few others:]
I wonder how many of them are working?

and it is CCs that define coherence in Q-A relations:
Are Fred or Pete working? Pete is working.

6. CCs are denotations, not meanings

Note, again, that CCs are not meanings, but the denotations of linguistic expressions: the denotation of a VP may depend on referential expressions and their presuppositions, and it fails to denote a CC if no referent can be recovered:

Fred is working for her, and so is Jane.
Fred is working for his wife, and so is Jane.

The referent, and not just the meaning, of (her / his wife) must be the same for Fred and Jane; and if Fred has no wife, the reference fails and no CC can be construed.

The assignment of a compositional denotation of the VP depends on the referential success of its constituents. The compositional meaning of an expression could never depend on the referential success of its constituents.

6. Division of labour between lexicon and concept representation

Semantic entries in the lexicon only provide pointers to concepts

<table>
<thead>
<tr>
<th>lexical item</th>
<th>concept name*</th>
</tr>
</thead>
<tbody>
<tr>
<td>WORK</td>
<td>λx WORK(x)</td>
</tr>
</tbody>
</table>

 Lexicon

<table>
<thead>
<tr>
<th>Conceptual Representation</th>
</tr>
</thead>
<tbody>
<tr>
<td>WORK ⊆ WORK₁</td>
</tr>
<tr>
<td>WORK ⊆ WORK₂</td>
</tr>
<tr>
<td>etc.</td>
</tr>
</tbody>
</table>

∀x (WORK₁(x) ⊆ DEVICE (x) ...)
∀x (WORK₂(x) ⊆ HUMAN (x) ...)

etc.

* concept names are completely arbitrary and receive their content only from subsumption relations and axioms in the conceptual representation; see below.

6. Intermediate summary

I am claiming that all the effects of productive contextual modulation that we looked at

- modulation of a function by the denotation of a linguistically explicit argument, or of an argument by the denotation of a linguistically explicit functor,
- modulation of a function by an implicit argument,
- modulation of a function by any other form of contextual information

can be modelled by contextual concepts.

But not all in the same way...

6. Context dependence 1: Indexicality

The simplest case is the case of Kaplan indexicality.

The computation of the content of indexicals (in the narrow sense) rests on the assumption that it is a process of variable saturation: Lexically specified variables are evaluated by the context (speaker, listener, place, time, ...).

The exact same procedure is responsible for the interpretation of functor expressions, such as nearby, enemy, left, fast, which contain a lexically specified variable (location, person, perspective, comparison class, ...) that is evaluated to the current contextual value and may be bound.

In extension of the latter, we may also consider expressions that arguably also contain implicit index variables, albeit perhaps not lexically specified, such as it's raining.
6. Context dependence 1: Indexicality – shifting and binding

The evidence for implicit variables in any case is in shiftability of the value and bindability of the variable:

1. Shiftability

Even though my utterance of It's raining is by default interpreted as it's raining here and now, the sentence Both Graham and Peter said it was raining.

is plausibly interpreted as Graham having said it was raining in Washington and Peter in Berlin, when the context tells us that Graham was in Washington and Peter in Berlin.

2. Bindability

Since also binding by quantifiers is possible, as in Whenever I go out, it's raining.

Wherever I go, it's raining.

we must probably conclude that the character of it's raining contains unarticulated place and time variables.

But not all context dependence can be tied to variables.

6. Context dependence 2: Non-indexical modulation

Variable evaluation as it operates in the interpretation of indexicals is clearly not applicable to lexical items whose characters do not contain variables.

This includes the cases of

- modulation of a function by the denotation of a linguistically explicit argument; e.g., open {book, letter, door, bottle, buffet}
- modulation an argument by the denotation of a linguistically explicit functor, {read, publish, find, start} a novel
- and modulation of a functor denotation by other forms of contextual information that is not given by a variable or an explicit functor or argument expression; e.g., the various interpretations of work in replies to different questions.

6. Context dependence 3: “free” contextual modulation

- modulation of a functor denotation by other forms of contextual information that is not given by a variable or an explicit functor or argument expression; e.g., the various interpretations of work in replies to different questions.

The problem here is that

- there are arbitrarily many parameters with respect to which any two actual contexts may differ;
- there can be no criteria for the identity of contexts that are themselves context-independent.
- accordingly “contexts” are no respectable entities and cannot figure as arguments for any respectable function.

“Free” productive contextual modulation is a task for non-monotonic knowledge-based reasoning, but still compositional!

7. How are CCs computed?

- Differently for different information structure status

Anaphoric expressions are interpreted with reference to an established discourse representation. There is no modification and (next to) no lexical semantics involved.

Focal expressions are interpreted via their lexical entries, plus contextual modulation.

The lexical entry provides a grossly underspecified lexical concept (a node in the conceptual representation) and modulation operates on conceptual representations.
7. Anaphoric vs. focal occurrence

Is Fred working?

a. No, he isn’t working. He’s sick.

a’. No, he isn’t. He’s sick.

If both (a) and (a’) mean the same, and have, in this discourse, the same truth-conditions, then the lexical semantics of “working” in (a) can’t be making a semantic contribution: it just copies its denotation from the preceding question.

7. Computing values for anaphoric expressions

Anaphoric expressions select their denotation from a small and finite domain: the discourse representation (short term memory).

The information used in their evaluation is structural (incl. salience status) and high-level semantic class information (n-features, subcat), i.e. the mechanisms are those of anaphora resolution.

Typical case are anaphoric pronouns and VP anaphora.

7. Computing values for focal expressions

Focal expressions are mapped to their values in two steps:

i. the expression selects a lexical entry, which contains a pointer to a lexical concept.

----------- the semantics strictly ends here ---------------

The concept is part of a conceptual representation and is linked to other concepts via its internal structure, via subsumption and various axioms.

It is underspecified with regard to what may be needed in the semantic values.

ii. Conceptual processes and information from arguments, modifiers, and discourse, as well as non-linguistic sources complete the construction of semantic values.

8. Argument (Re-)Interpretation

(1a) Fred began reading a novel.

(1b) Fred began a novel.

A lexical entry for begin:

8. application to "coercion"

Argument (Re-)Interpretation

(1a) Fred began reading a novel.

(1b) Fred began a novel.

But where is the needed event argument for Fred began a novel?

also: 

\[
\text{reading a novel } \subseteq \text{ EVENT}
\]

\[
\begin{align*}
\text{The lexical concept} &
\text{[begin]} = \\
\chi y &\chi x \exists s \ [ \text{BEGINNING}(s) \land \text{AGENT}(x,s) \land \text{EVENT}(y,s) ]
\end{align*}
\]

\[
\text{[a novel ] } \subseteq \text{ EVENT}
\]
8. A lexical entry for "novel"

[novel] = \lambda x. NOVEL(x)

not very informative?

I want to make do with just one lexical entry and one underspecified lexical concept.

If we believe that a novel is a longish prose narrative, etc., then this is not lexical knowledge, but literary theory:

8. some polysemous senses of "novel"

The various "polysemous senses", i.e., CCs, are generated in the conceptual representation as different views*.

Views provide enriched subordinate concepts of a more abstract concept and inherit attributes from select superconcepts.

Summary (i)

A sketch for an account for contextual variation in predicate denotation

1. The methodology
   - assumptions of incrementality, immediacy, and multi-modality (to keep closer to processing)
   - compositionality (to account for productivity and learnability), but at the level of denotations
   - computability

Summary (ii)

2. Some of the features
   - Polysemy disappears from the lexicon and is relocated in the conceptual representation.
   - No productivity in the lexicon, (leaving aside questions of word formation, possibly lexical rules)
   - The observed productivity is a conceptual, not a linguistic phenomenon.
   - Semantic information in the lexicon is reduced to a concept link ("disquotational lexicon") and subcategorization (domain restriction on arguments).
   - Conceptual representation integrates constraints collected from subcategorization, predication, and (non-monotonic) inference from discourse and utterance context.

Thank you!
References (very incomplete, sorry)


References ctd.


