LACUS FORUM XXVI
The Lexicon
University of Alberta
© 2009 The Linguistic Association of Canada and the United States (LACUS). The content of this article is from LACUS Forum 26 (published 2000). This article and others from this volume may be found on the Internet at http://www.lacus.org/volumes/26.

YOUR RIGHTS
This electronic copy is provided free of charge with no implied warranty. It is made available to you under the terms of the Creative Commons Attribution-NonCommercial license version 3.0 (http://creativecommons.org/licenses/by-nc/3.0/)

Under this license you are free:
- to Share — to copy, distribute and transmit the work
- to Remix — to adapt the work

Under the following conditions:
- Attribution — You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).
- Noncommercial — You may not use this work for commercial purposes.

With the understanding that:
- Waiver — Any of the above conditions can be waived if you get permission from the copyright holder.
- Other Rights — In no way are any of the following rights affected by the license:
  - Your fair dealing or fair use rights;
  - The author’s moral rights;
  - Rights other persons may have either in the work itself or in how the work is used, such as publicity or privacy rights.

Notice: For any reuse or distribution, you must make clear to others the license terms of this work. The best way to do this is with a link to the web page cited above.

For inquiries concerning commercial use of this work, please visit http://www.lacus.org/volumes/republication

Cover: The front cover of this document is licensed under the Creative Commons Attribution-No Derivative Works 3.0 license (http://creativecommons.org/licenses/by-nd/3.0/) and may not be altered in any fashion. The LACUS “lakes” logo and University of Alberta logo on the cover are trademarks of LACUS and the University of Alberta respectively. The University of Alberta logo is used here with permission from the trademark holder. No license for use of these trademarks outside of redistribution of this exact file is granted. These trademarks may not be included in any adaptation of this work.
1. truth value semantics. I will take my starting point in truth-value-based semantics in the tradition of Montague (1974). This so-called ‘logic’ paradigm has proved to be extremely useful in clarifying core issues in semantic processing, like for instance semantic composition, quantification, predication, the distinction of extensional versus intensional contexts, the grammar of focus or the role of utterance context in natural language understanding.

Yet, many linguists feel that the paradigm, in spite of its strengths, is inadequate because of its blind spot at the basis of semantic analysis: Word meaning. Word meanings are modelled by functional objects over the sets of individuals, times, the real and other possible worlds, according to real or counterfactual facts. Basically, the theory captures two aspects of word meaning:

• How would this word behave in combination with other words?
• Where would this word (perhaps in combination with other words) truthfully apply?

It has turned out that an answer to even these two simple questions will result in theories of considerable complexity. This very complexity brings another question into focus, a question that the approach leaves, programmatically, unanswered:

• How do speakers grasp meanings?

Truth value semantics takes it as a basic, unanalyzed fact that speakers do have something in their heads that will allow them to use words like meat, sponge, and or seek as if they had the corresponding intensional objects in their brains. Soberly refraining from any introspective meaning paraphrases, this notion of ‘meaning’ is based on the uncontroversial minimal requirement that two speakers are in command of the same meaning of a word only if they agree on the class of things that correctly can be referred to with that word1.

Linguists will eventually be doomed to accept some level of cognitive processing as ‘given input’ for semantic theory. Yet, truth value semantics seems to set that level unfruitfully high and, as a result, has to remain silent with respect to a range of interesting semantic questions:

• How do speakers learn, know, establish and change word meanings?
• How can distinct word meanings be ‘related’ (while other distinct meanings are definitively ‘unrelated’)?
• Why can speakers creatively use words in a shifted meaning?

In the present paper, I propose to build up word meanings from more basic ingredients of meaning. Specifically, I will integrate Gärdenfors’s recent proposal about the geometric nature of conceptual knowledge (Gärdenfors 2000) into truth value semantics, and apply the resulting theory to model certain cases of meaning change.

2. Meanings in the Head. Proper names are certainly the simplest instance of a link between word and meaning. The established bearer of a name is the one and only thing we need in order to know what the name means, as was argued convincingly in Kripke (1972) (a more recent reconstruction of the main arguments can be found in Devitt & Sterelny 1999). The resulting causal theory of reference assumes that the meaning of a name is established in an initial act of baptizing in which the name is ‘attached’ to an individual \( r \). In terms of intensional semantics, this leads to the so-called rigid designators: Let name be a proper name in the language under investigation:

\[
[[\text{name}]] = f: D_s \rightarrow D_e \text{ where } f(w) = r \text{ for all possible worlds } w.
\]

While this idea will work nicely for proper names, it evidently is too simple for the vast majority of words in natural languages that refer to classes of individuals, objects, events, states, processes etc. Putnam (1975) extended the account to natural kind terms (like gold, water, tiger) which he proposes to treat as proper names for natural kinds. The meaning of a natural kind term is again established in an initial act of baptizing with reference to some sample \( r \) of the natural kind in question. However, we want to name the entire kind, not only sample \( r \). This is captured by the definition in (2). Let kind be a natural kind term, let \( r \) be the sample that was pointed at in baptizing. Substidentity is the trans-world relation that holds between two lumps of matter exactly if they are of the same substance, or for two organic individuals exactly if they are of the same kind.

\[
[[\text{kind}]] = f: D_s \rightarrow D(e,t) \text{ where } f(w) = \{ b \mid b \text{ is a lump of matter in } w, \text{ and } b \text{ substidentical to } r \}.
\]

Note that (2), like (1), does not make reference to human categorization abilities. According to the received view, substidentity is a relation that is determined by reality itself, not imposed by human reflection of reality. The position of realism has earned its high reputation in philosophy by the fact that it can, better than any other position, explain the nature of scientific progress³. Yet, it is a consequence of this assumption that not very many words in our lexicon qualify as natural kind terms,
and are captured by (2). There seems to be good reason to ask for a more flexible, less externalist account.

The idea is certainly appealing that the class of objects that is denoted by a word \( \omega \) should somehow ‘cluster around’ those things which were first called ‘\( \omega \)’. We will leave it open whether these first uses were deliberate baptisms in the literal sense (like names of little babies) or whether the word \( \omega \) was established in a more implicit manner (like for instance the adoption of loanwords).

Consider the relation of substitentity in definition (2). Only reality (or, perhaps, God) knows which things are identical in that sense. Arguably, speakers don’t know it. I suggest to replace it by a binary similarity relation \( \text{sim} \). Loosely speaking then, the word \( \omega \) applies correctly to the first sample \( r \) which was referred to in baptizing, plus any other object \( x \) which resembles \( (\text{sim}) r \) in the appropriate sense. Some comments are in place.

Firstly, and most importantly, there is more than one respect in which things can be similar. Different choices of respect-in-which-things-are-compared will lead to different classes of things-which-are-similar-to \( r \). A classical place where this observation is discussed (but by no means the only one) is Goodman (1972:421–49). In establishing a new word \( \omega \), pointing at a sample object \( r \), we therefore need to specify the respect in which other referents of \( \omega \) should resemble \( r \). We will have to choose the intended one of a wide range of similarity relations \( \text{sim}_i, \text{sim}_j, \text{sim}_k \ldots \) This assumption, wild as it might look, is in fact supported by findings in experimental studies of word learning like Landau et al. (1998), summarized below.

Secondly, what is the formal nature of these similarity relations? I will adopt a recent proposal that has been made in cognitive science by Gärdenfors (1998, 2000). He argues that similarity is most naturally built on a notion of distance over the respective space of objects or events. An object \( b \) is similar to another object \( a \) if it is ‘close enough to \( a \)’ in terms of that distance. If the question is whether \( b \) is like \( a_1 \), or like \( a_2 \ldots \) or like \( a_n \), where these \( a_i \) are prototypical members in different classes which we already know, then we will say that \( b \) should go with the prototype \( a_j \) which is closest to it in terms of that distance. Nothing in the following will crucially depend on the formal nature of similarity. However, in view of the broad literature in the field it might be helpful to locate our approach at some definite point.

Thirdly, where do we determine which is the intended similarity relation in question? Sterelny (1983) observed that even the classical causal theory of reference needs to specify the taxonomic level at which ‘identity’ is to be determined (even assuming realism). Sterelny argues that this \textit{qua question} is determined by the context in which a baptizing takes place. I will adopt his proposal and assume that the utterance context of a baptize will fix the respect in which the sample is to be generalized. (‘Should Fido stand for the class of mammals, for dogs, for male dogs, for pets… ?’) Technically speaking, the \textit{qua question} is answered by context in pointing out some \textit{similarity relation} \( \text{sim}_i \). Thus, we come to the following modified theory of baptizing:
If word \( \omega \) has been introduced in baptizing \( \mathcal{c} \) on the basis of sample \( r \) then

\[
[[ w ]] = \text{a function } f \text{ from worlds to sets of objects such that }
\]

\[
f(w) = \{ x \mid x \text{ is an object in } \omega \text{ such that } \text{sim}(x,r) \text{ in } \omega \}\]

The similarity relation \( \text{sim} \) was salient in context \( \mathcal{c} \).

Note that this definition is the starting point for an epistemic interpretation of possible world semantics. It is no longer reality that cuts out the classes of objects denoted by a word. Extension and intension depend on speaker’s interests, as well as their knowledge at the time of introduction of \( \omega \). Similarity is not a ‘realistic’ notion in the philosophical sense. It is a conceptual notion, reflecting how speakers see the world.

In breaking down word meanings into core referents and similarity judgements, we have once more reached a level where some notions, namely similarity relations, are taken as basic givens. Humans, as well as animals, do have the ability to compare and match sensual experiences, if only to ensure successful interactions with their environment. It is not the task of the linguist to explain these cognitive abilities, yet the linguistic behaviour of speakers can provide additional evidence in the study of these abilities. For instance the careful study of meaning change, due to different ways to generalize a given set of samples, can be helpful in order to understand in how many ways things can be perceived as being ‘similar’. I will elaborate this idea in the following sections.

3. contexts of use. We’ve seen in the previous section that in a context in which we introduce a word \( \omega \), generalizing a given referent \( r \), this generalization will follow the most salient similarity relation \( \text{sim} \) suggested by that context. In the present section, I will attempt to give a clearer idea of this process.

Shape is probably the most prominent respect in which an initial sample can be generalized. The meaning of words like triangle or circle (in their everyday sense where also a somewhat shaky triangle counts as a triangle) are clearly based on the idea of being similar in shape to a prototype triangle, being similar in shape to a prototype circle. Shape also seems to be the default respect along which we generalize. It is the first respect used by children in categorization (see e.g. Keil (1994)). It is the respect both children and adults will use in absence of any other clues (see Landau et al. 1998). Finally, practically all words which refer to visible things have a reading that can be paraphrased as ‘looking like a core instance of…’, as exemplified by compounds like stone lion, plastic apple, wooden book or paper boat.

Function is often used as the alternative respect in which things can resemble each other (exemplified once more by Landau et al. 1998). Certainly, we will have to view function as a parametrized respect—if we know that a given sample is to stand for a larger class of objects in terms of function, we will have to observe our fellow humans in order to find out which function that sample usually fulfills. Consider the example of tin-openers. They are created for the purpose of opening tins, and everything which resembles core tin-openers in that respect will qualify as a tin-opener. An object which resembles my tin-opener in the functional respect of ‘being an instrument
with which one could severely hurt enemies’ will not qualify as a tin-opener (wrong function). An object which looks like a core tin-opener but is unsuitable for opening tins will only be a ‘tin-opener’ in the shape sense of the word.

We also have to count with mixed cases of function plus shape, or function plus a certain way of fulfilling that function. For example, modern engineering has created tools that fulfill the function of a saw, working with a very thin ray of water that is emitted at high pressure. Is this still a saw? Sometimes the interests of the marketing department may play a role, too. Or, why am I not allowed to call the ‘trackball’ in my laptop a ‘mouse’ in spite of it fulfilling exactly the same function? (I did, but was informed that I had something much fancier than a ‘mouse.’) We can observe that similarity judgements, in the more sophisticated case, certainly are subject to social negotiations and conventions as to who is licensed to classify a new artifact (as has been nicely shown in Bloom 1996).

Yet, it is possible to provisionally locate certain factors that influence the choice of $\text{sim}$ in an utterance context $c$ of word $\omega^i$:

(i) When does $c$ take place? — Time of $c$.
(ii) Who were speaker and addressee in $c$? — Speaker and hearer, and in a broader sense: language community they are part of.
(iii) What was the shared epistemic background of that language community at the time? — Knowledge in $c$.
(iv) What was the conversation about? — Interests of speaker and hearer in $c$.
(v) What was the referent $r$ of $\omega$ in $c$, if there was any?

We can at least argue that knowledge and interests contribute to the choice of $\text{sim}$ in a given context. Assume that we have two contexts $c_1$ and $c_2$ in which speakers with similar interests talk about some lump of matter $r$. Assume that these interests are such that the speakers are concerned with $r$ as a natural substance. However, assume that the speaker community in $c_1$ has the knowledge of ancient Greek, while the community in $c_2$ is in hold of modern chemistry and physics. The similarity relation chosen in $c_1$ will therefore be based on features like color, hardness, mass, etc. while the similarity relation in $c_2$ might use such sophisticated criteria as mass spectroscopy. This can (but need not necessarily) lead to the fact that the extension of a word $\omega_1$ for ‘$r$ and what is similar’, introduced in $c_1$, is different from the extension of a word $\omega_2$ for ‘$r$ and what is similar’, introduced in $c_2$. Thus, contexts with equal interest but different knowledge can lead to different concepts. This observation is complemented by similar findings in first language acquisition (see Keil 1989, 1994).

The opposite case consists of a pair of contexts which take place at the same time, in the same speaker community and under the same knowledge, but with different interests. Landau et al. (1998) describe that case and find that subjects indeed will use a new word differently, depending on whether the initial sample was shown to them without further comment or with a prolonged explanation as to what task one
would usually fulfill with that object. In the first case, subjects generalized by shape, in the latter by function.

Let me finally comment on the difference between speakers and speaker community. Evidently, we do not want to make word meanings dependent on the knowledge of arbitrary individual speakers. The idea of linguistic ‘division of labor’ was already introduced in Putnam (1975), and has been accepted without much objection. In that sense, single speakers in some given context c may not even themselves be fully aware of the kind of similarity relation in play. In the case of baptizing, there certainly are correlations between those who are considered to be experts in question and those who are authorized to introduce certain words into the language. Plainly speaking, usually only the biologist is licensed to establish the name of a new species.

In the next section, however, we will elaborate the idea that also contexts of use c of a word w are relevant for the further meaning of word w. In such contexts at latest, we have to face the possibility that laymen as well as experts might be speakers and hearers. But they still are part of a larger speaker community and in that sense participate in expert knowledge.

4. Words in Time. In Section 2, I proposed that word meanings are determined by two ingredients: Some core sample(s) r that are known to fall under the word in question and a similarity relation sim to generalize these first samples. These two factors can be traced clearly in baptising situations, that is in contexts where speakers want to establish a new word in their language. Yet, many ordinary utterance contexts of a word w will provide these two factors as well. Many utterance contexts offer referents r for words. All contexts of use take place between speakers and hearers that are part of larger linguistic communities. Many contexts of use are guided by specific interests of the communicating parties, and take place against shared background knowledge. Consequently, many contexts of use c of a word w provide us with everything that would be necessary to learn or establish some meaning for word w. Nothing, however, ensures that the context of use provides us with exactly those referents r and similarities sim that will create the ‘traditional’ meaning of word w. The context c might support a new meaning for w (although speaker and hearer will usually know and use w in its traditional meaning, of course).

While this is undangerous in the isolated case, it is the point where a great potential for language adaption, also called ‘meaning change’, is hardwired in our language faculty:

- If the most frequent core referents change over time, the resulting category may gradually follow its shifting core.
- If a word is used under a new perspective, or with new interests in the objects referred to, it may adopt a new meaning that reflects these changed interests.
- If background knowledge changes, speakers may also see the need to adapt their similarity judgements to advanced knowledge.
Let us look at some word histories that exemplify these three types of meaning change and extension.

The German word Karussell, which was taken over from French carrousel in the second half of the 17th century, originally denoted a form of tournament game where the task was to hit a ring with the lance while riding past it on a horse. In the 18th century, a variant of the original game became popular where the players had to hit the rings from a pivot mounting, sitting on wooden horses instead of real ones.

While this variant was still, in a sense, the old game, and while the participants’ interest still consisted in ‘having fun’, the kind of ‘fun’ taken out of the game changed: The original ‘Karussels’ were sportive competitions, the new version was already fun because one could travel in rounds on hobby horses. Following these new core exemplars, the category of ‘Karussell’ moved towards its present extension (like E caroussel). It was reported to me that in Holland, at least until 30 years ago, carousels had a bell attached at the roof. Riders who hit that bell could win a free ride. This version of caroussel still maintained an element of skill and competition.

In Germany today, the rings or bells to be hit have completely vanished from the ‘Karussell’, and the horses now come in the company of little cars, swans, bikes, closed carriages and all other sorts of vehicles which would have been unsuitable for a ‘Karussell’ in the sense of 1650.

The word Kreide ‘chalk’ has two meanings in contemporary German. In one reading Kreide₁, it means ‘calcium carbonate’ and is used as a name for a chemical substance. In a second reading Kreide₂ it denotes a writing tool, typically the small handy piece of Kreide₁ which can be used to write on blackboards. The two readings differ. Modern industry has created Kreide₂ not only coming in different colours, but also made from other materials which resemble Kreide₁ to a greater or lesser degree. The pieces of Kreide₂ which deviate most from Kreide₁ are Malkreide or Wachskreide (crayons) on the basis of wax (or, again, its modern substitutes) for kids to draw colourful paintings. It is easy to see how both readings are generalizations over the same core referents: Small pieces of a soft white stone that are good for drawing and writing.

Actually, Kreide₁ nowadays becomes fashionable in a second tool sense, namely as a white powder which, mixed with water, can be used to paint walls. Time will show if even a further reading Kreide₃ will develop on this basis which has an extension distinct from both, Kreide₁ and Kreide₂.

German Pilz, in its colloquial meaning, covers both eatable mushrooms and poisonous toadstools, as long as they exhibit the typical stem-and-cap shape, with a vague boundary towards species like morels and puff-balls. The word however acquired a second, botanical meaning where category boundaries are determined by similarity of reproduction mechanisms rather than similarity in shape. In this scientific sense, ‘Pilz’ extends to, for instance, kinds of mould or the fungus used in penicillin production. In cases like this, specific interests in taxonomy conspire with more sophisticated knowledge of the scientific language community to determine the technical sense of a word, which only afterwards spreads into the common speaker’s lexicon.
Note that the account is by no means restricted to nouns. A very nice example of meaning change in the case of verbs can be found in Verdaguer and Poch (1998) who observe that both in English and Catalan, verbs that denote the state of ‘being awake’ (OE wæccan) follow a meaning development over several steps, eventually turning into verbs that express visual attention or observation (ModE watch). Let us have a look at one step in detail, the one from watch = remain awake to watch = be watchful. Verdaguer and Poch claim that the latter watch consists in the former, plus an additional feature [+ mentally aware]. If that were so, we’d expect that be watchful should be a special form of remain awake. This is, however, not true. Specifically, remain awake presupposes that the agent is tired and inclined to sleep (or that there are other expectations or plans for him to fall asleep). Be watchful, on the other hand, does not carry any such presupposition.

It seems more appropriate to capture this meaning development as a case where core instances of ‘remaining awake’ are generalized in a new direction, different from the old one that led to the category of eventualities where someone remains awake: Why does one typically stay awake, instead of giving way and falling asleep? Apart from religious purposes, one typical situation is the one where one gets tired in hostile surroundings. Remaining awake under such circumstances has the immediate purpose of ‘keeping watch, being prepared to meet danger’. If we take into account this typical interest in staying awake, and generalize the activity of the watching agent in that direction, we immediately come to the concept of BEING WATCHFUL and KEEP WATCH. Moreover we will expect quite naturally that there will be instances of KEEP WATCH which are not instances of REMAIN AWAKE, and vice versa.

The other steps in the word history of watch as described by Verdaguer and Poch fall equally nicely into place, and the fact that similar meaning developments took place in English and Catalan can be explained by the simple observation that the same conflict between being tired and being afraid of imminent dangers can well arise in any culture, without having to resort to such grand labels as ‘being grounded in human cognition’.

Let me finally briefly indicate how these insights can be reconstructed in formal terms. Interestingly, the previously purely semantic modelling will have to be augmented by those aspects of actual history that influence the meaning of word ω⁷: A set of contexts of use C, a function time that yields, for each context c in C, the time at which c occurs, a function s that maps contexts to similarity relations sim, and a partial function ref that maps contexts c to the intended referent r in c (if there is any).

We can now define \( C(t) := \{ c \mid \text{time}(c) = t \} \), the set of all cases where ω was used at time t. Next, we can sort this full set into subclasses of contexts which all support the same similarity respect: \( \{ c \mid \text{sim} := \{ c' \mid c' \in C(t) \ \text{and} \ \sigma(c) = \sigma(c') \} \). The last step consists in collecting the referents in these contexts: \( \text{ref}(\{ c \mid \text{sim} \}) := \{ r \mid \exists c' \in \{ c \mid \text{sim} \} \text{ such that } \text{ref}(c') = r \} \). Eventually, one will want to reduce these actual referents to the class of those that were the most common referents, or that constitute a homogeneous class in themselves (see Geeraerts (1997) on the effect of prototypes in diachrony), but let me refrain here from this extra step.
Given these factors, the potential meanings of word $\omega$ at time $t+1$ can be characterized as follows:

$$f$$ is a potential meaning of $\omega$ at time $t+1$ if there is a $c \in C(t)$ such that
\[
f(w) = \{ x \mid \exists r \in \text{ref}[\{ c \}_{\text{sim}}] \text{ such that } \text{sim}(x, r) \text{ in } \omega \text{ holds true}\}
\]

Two comments are in place. Firstly, we have attempted to characterize the potential meanings of a word $\omega$ at time $t+1$, given that we know how $\omega$ was used at time $t$. This, of course, does not and can not tell us which of those potentials actually are installed in the language. Large sets $\{ c \}_{\text{sim}}$ will presumably be more likely to establish the corresponding meaning while small, isolated and idiosyncratic sets $\{ c \}_{\text{sim}}$ are less likely to alter language (note in that connection how mass media can multiply a single utterance into millions of contexts of use with different hearers each). Yet, the status of the speaker in $c$ will also play a role, as well as the question whether the respective hearer in $c$ has safe knowledge about the ‘traditional’ meaning of $\omega$ or actually uses $c$ as a clue to guess what word $\omega$ might mean.

Secondly, the account is evidently limited to those new meanings that arise through the processes under discussion in this paper. I do not intend to capture meaning extensions or shifts through reanalysis, folk etymology, metonymy or ellipsis. It is an open question as to how far the present approach can be generalized to metaphorical extensions (see Gaerdenfors 1996 and Gentner 1983).

---

1 ...and if they both combine the meaning of that word with other items in the same ways. We will in the following ignore these compositional aspects of meaning.

2 In accordance with the literature, I will adopt the following notational conventions: $D_e$ stands for the domain of individuals ($e \approx$ entity), $D_s$ for the domain of possible worlds. Individual concepts are represented by functions that map worlds onto individuals (abbreviated: $D_s \rightarrow D_e$).

3 In Devitt and Sterelny (1999); Eckardt (1999) critiques the concept of ‘natural kind term’.

4 Note that this set is redundant in that some of these factors depend on others. I chose to name all factors explicitly for the sake of clarity.

5 Unless the new word has a clear definition. Yet, I think that not many word meanings are introduced by means of a definition. A more common case is the one where the core samples are not given by ostension, but through a verbal description (e.g. unicorn). I see these examples as a more complex case of my account rather than as words that follow definitions.

6 The reader will have noticed that I do not distinguish between cases of literal meaning shift and cases where a word adopts a related additional meaning. Such cases of polysemy are often filed as ‘meaning change’ if the original first reading gets out of use.

7 The exposition is simplified in various respects. For one thing, I refrain from offering an explicit implementation of knowledge states, speaker communities and interests that are determinant factors in the mapping of contexts to similarity relations. Second, I have systematically omitted all formal discussions of properties of similarity relations. A third simplification consists in restricting the formalism to the history of only one single word in time. A fuller treatment is offered in Eckardt (1999).
REFERENCES