

# A Computational Treatment of Generalized Reference

Roland Hausser

Universität Erlangen-Nürnberg (em.)  
rrh@linguistik.uni-erlangen.de

## Abstract

Medieval logic defined reference as a relation between language and objects in the world. Recently, however, the term “representational token” has been used instead of language (Reimer and Michaelson 2014). This allows for reference with and without language. In a similar vein, Database Semantics (DBS) has implemented concept-based reference as a matching between two contents. If a content is attached to a language surface it is called the literal meaning<sub>1</sub> of the surface.

Referring with a content (as a representational token), regardless of whether or not it is attached to a surface, leads to a generalized notion of reference (6.3). An example of reference without language is identifying a current nonlanguage recognition with something seen before. Another example is identifying a nonlanguage recognition with an earlier language content, e.g. something read (for example, in a guide book) or heard about.

In addition to the concept-based reference mechanism of (i) symbols (Sect. 6), natural language uses the reference mechanisms of (ii) indexicals (Sect. 7) based on pointers, and of (iii) names (Sect. 8) based on acts of generalized baptism. A fourth kind of reference is co-reference (Sect. 9), based on identity implemented by means of an address; it occurs as a variant of referring with indexicals and symbols, and is the foundation of name-based reference.

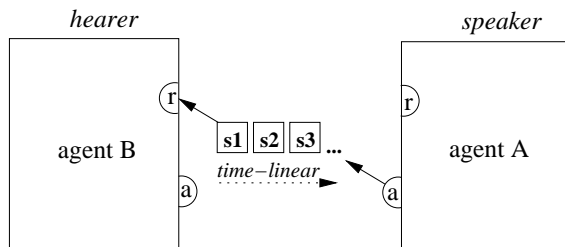
This paper systematically reconstructs the mechanisms of reference as they function with and without language in an agent-based computational framework. Language-dependent surfaces play a role only in the automatic word form recognition of the hear mode and automatic word form production of the speak mode. In conclusion, the agent-based reconstruction of reference is applied to the medieval distinction between *de dicto* and *de re* (Sect. 10).

**keywords:** reference based on matching, pointing, baptism, and identity, hear mode, speak mode, time-linear derivation order, talking robot

## 1 Agent-Based Ontology

When two agents communicate with each other by means of a natural language, the speaker uses its external action interface to produce a sequence of language surfaces while the hearer uses its external recognition interface to identify the elements of the sequence. The sequence is time-linear in the sense that it is linear like time and in the direction of time. In accordance with the Western writing convention, the progression of time is shown in the direction from left to right:

## 1.1 PHYSICAL FRAMEWORK OF COMMUNICATION



The recognition and action interfaces of the agents are indicated by half circles marked with *r* and *a*, respectively. The language surfaces are represented by boxes containing *s1*, *s2*, *s3*, ... . As agent-external modality-dependent sound waves (speech), dots on paper (writing), or gestures (signing) the surfaces may be measured and described with the methods of the natural sciences, but have no meaning and no grammatical properties whatsoever.

The first surface leaving the speaker is the first to reach the hearer. The last surface leaving the speaker is the last to reach the hearer.<sup>1</sup> All other aspects of language communication are agent-internal, modality<sup>2</sup>-independent, and cognitive.

Modality-independence may be illustrated by the basic operations of arithmetic, i.e., addition, subtraction, multiplication, and division. They exist at a level of abstraction which may be realized equivalently as the operations (i) of a human, (ii) a mechanical calculator, or (iii) an electronic computer.<sup>3</sup>

With autonomous robots still absent in today's computational linguistics, the external framework 1.1 may be simulated, using the keyboard and the screen of standard computers as primitive recognition and action components. This, however, works only for the transfer of surfaces. It does not work for nonlanguage recognition and action, which are required for a cognitive reconstruction of reference. For example, the agent's ability to refer to agent-external items is needed for fulfilling a request like *Pick up the blue square!* or to report how many blue squares there are in the agent's current task environment.

## 2 Elementary Concepts

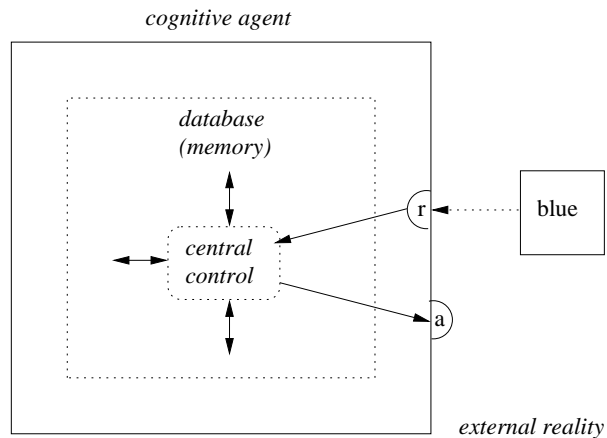
The minimum in reconstructing higher-level cognition is (i) an agent-internal memory, (ii) a central control embedded into memory, (iii) a mapping from the recognition interface to central control, and (iv) a mapping from central control to the action interface. Consider an agent recognizing the non-language object *blue square*:

<sup>1</sup>In 1.1, 4.2, and 5.2 this is expressed graphically by placing the hearer to the left and the speaker to the right. If the order is reversed, the progression of time is from right to left.

<sup>2</sup>The term *modality* is being used in several different fields of science. As employed here, modality is known as *sensory* modality (Chen 2006, Sect. 6.13.1).

<sup>3</sup>According to Wiener (1961, p. 132): "Information is information, and not matter or energy."

## 2.1 BASIC STRUCTURE OF COGNITION



The external interfaces of recognition and action constitute the agent's *peripheral cognition* (5.1). In recognition, peripheral cognition maps modality-dependent raw data into modality-independent agent-internal concepts. In action, it maps modality-independent blue-prints into modality-dependent external raw data. This provides the *grounding* of Database Semantics in the recognition and action of a natural or artificial cognitive agent.

The mappings between modality-dependent raw data and modality-independent concepts are formally based on the type-token distinction, familiar from philosophy.<sup>4</sup> The type of a concept describes the *necessary* properties, while an associated token is an instantiation with certain additional *accidental*<sup>5</sup> properties. As an example consider the recognition of colors (Hausser 1989, p. 296 ff). In physics, they are defined as intervals on the one-dimensional scales of electromagnetic wave length and frequency. Accordingly, the type and a token of the color **blue** may shown as follows:

## 2.2 TYPE AND TOKEN OF THE COLOR CALLED **blue**

*type*

$\left[ \begin{array}{l} \text{wavelength: } \alpha \\ \text{frequency: } \beta \end{array} \right]$   
 where  $\alpha \in 490\text{--}450 \text{ nm}$   
 and  $\beta \in 610\text{--}670 \text{ THz}$

*token*

$\left[ \begin{array}{l} \text{wavelength: } 470 \text{ nm} \\ \text{frequency: } 640 \text{ THz} \end{array} \right]$

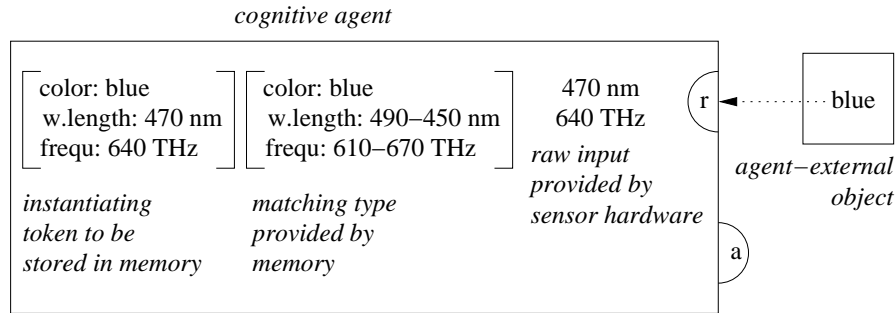
The type specifies the wavelength and the frequency of the color *blue* by means of variables which are restricted to the corresponding intervals provided by physics. The token uses constants which lie within these intervals.

<sup>4</sup>The type-token distinction was introduced by Peirce (CP 4: 537).

<sup>5</sup>The term *accidental* is used here in the philosophical tradition of Aristotle (384–322 B.C.), who distinguishes in his *Metaphysics*, Books  $\zeta$  and  $\eta$ , between the necessary and the accidental (incidental or coincidental – *kata sumbebêkos*) properties of an object in nature.

In the agent's recognition of colors, the type and the raw sensor input result in a classified token by interacting as follows:

### 2.3 TYPE AND TOKEN IN COLOR RECOGNITION



The agent's sensor measures the electro-magnetic wavelength 470 nm and frequency 640 THz in an agent-external object. These values lie within the intervals 490–450 nm and 610–6700 THz of the color blue and thus match the associated type, provided by memory. In the instantiating token, the wavelength and frequency intervals of the type are replaced by the measured values. The feature structures representing concept types and tokens may be extended as needed, for example, with an additional attribute for color intensity.

Next consider the type and the token of a two-dimensional geometric object:

### 2.4 TYPE AND TOKEN OF THE CONCEPT square

*type*

$$\left[ \begin{array}{l} \text{edge 1: } \alpha \text{ cm} \\ \text{angle 1/2: } 90^\circ \\ \text{edge 2: } \alpha \text{ cm} \\ \text{angle 2/3: } 90^\circ \\ \text{edge 3: } \alpha \text{ cm} \\ \text{angle 3/4: } 90^\circ \\ \text{edge 4: } \alpha \text{ cm} \\ \text{angle 4/1: } 90^\circ \end{array} \right]$$

where  $\alpha$  is a length

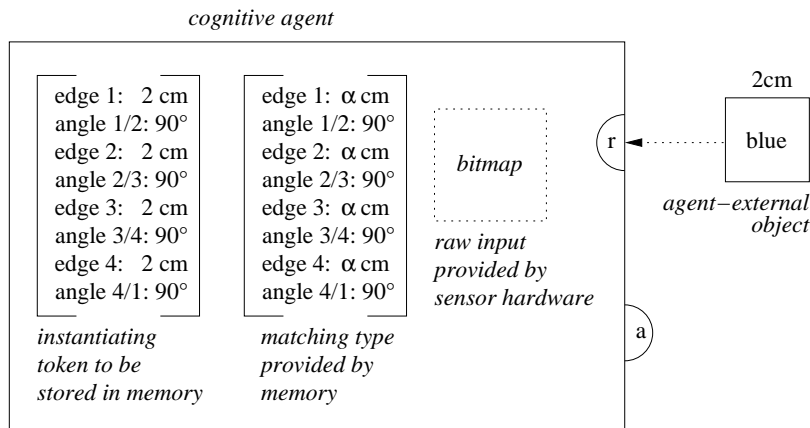
*token*

$$\left[ \begin{array}{l} \text{edge 1: } 2 \text{ cm} \\ \text{angle 1/2: } 90^\circ \\ \text{edge 2: } 2 \text{ cm} \\ \text{angle 2/3: } 90^\circ \\ \text{edge 3: } 2 \text{ cm} \\ \text{angle 3/4: } 90^\circ \\ \text{edge 4: } 2 \text{ cm} \\ \text{angle 4/1: } 90^\circ \end{array} \right]$$

Here, the type and the token share attributes which specify (i) the number of equally long edges and (ii) the angle of their intersections. The type and the token differ only in their edge lengths. Specified in the type as the variable  $\alpha$ , the edge length is 2cm in the token. The latter is accidental in that the type matches an infinite number of square tokens with different edge lengths. In an artificial agent, the type may be implemented as a pattern-matching software which recognizes tokens by approximating raw bitmap outlines Hausser (1999 3.2.1).

In analogy to 2.3, recognition of a square may be shown as follows:

## 2.5 TYPE AND TOKEN IN RECOGNIZING A SQUARE



The type matches the outline of all kinds of different squares, whereby its variables are instantiated in the resulting tokens.

Today, there exist pattern recognition programs which are already quite good at recognizing geometric objects.<sup>6</sup> They differ from our approach in that they are based almost completely on statistics. However, even if the terms of the type and the token may not be found in their theoretical descriptions, the type-token distinction is nevertheless implicit in any pattern recognition processing. Furthermore, the rule-based, incremental procedures<sup>7</sup> of pattern recognition presented in Hausser (2005) are well-suited to be combined with statistical methods.<sup>8</sup>

The elementary concepts of nonlanguage recognition are complemented by those of action. For example, the concept *take* is defined as the type of a gripping action which is instantiated as a token to be realized as raw data. The token differs from the type in that it is adapted to a specific gripping occasion. It holds in general for recognition that raw data are classified by a type and instantiated as a token, while in action a type is specialized into a token which is passed to a suitable action component for realization as raw data (Hausser 1999 3.3.5).

The interactions between the agent's external interfaces, the types, the tokens, and the memory must be hand in glove. For example, if the agent has no sensor for measuring electromagnetic wavelength/frequency, colors cannot be recognized – even if the proper types were available from memory. Conversely, without the types the raw data provided by a suitable sensor cannot be classified and instantiated as tokens. Also, without a memory the types cannot be provided for recognition and action, and the tokens cannot be stored.

<sup>6</sup>As shown by the work of Steels (1999), suitable algorithms may evolve new types automatically from similar data by abstracting from what they take to be accidental.

<sup>7</sup>They are based on an incremental, memory-based procedure of pattern recognition using geons (Biederman 1987).

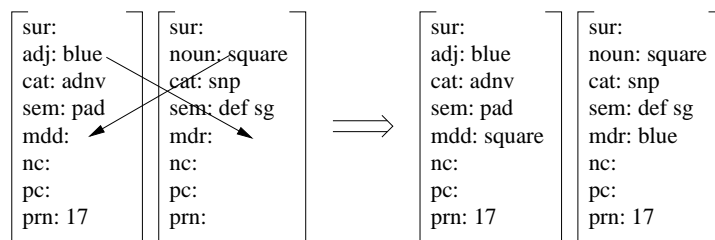
<sup>8</sup>For building a talking robot, the automatic evolution of types has to result in concepts which correspond to those of a language community. This may be achieved by presenting the artificial agent with properly selected data in combination with human guidance (guided patterns method, Hausser 2011, Sect. 6.2).

### 3 Data Structure and Database Schema

The concepts defined in 2.2 and 2.4 constitute elementary cognitive contents, but they do not provide any means for being connected, as in *blue\_square*. For this, DBS lexically embeds the concepts as *core values* into nonrecursive<sup>9</sup> feature structures with ordered attributes, called *proplets* (because they are the elementary building blocks of propositions, in analogy to droplet). A feature structure is built from features. In computer science, a feature is defined as an attribute-value pair (avp), e.g. [noun: square], with noun: as the attribute and square as the value.

The embedding of core values into proplets allows their concatenation by means of value copying. For example, the proplets *blue* and *square* may be connected into the content of *blue\_square* as follows:<sup>10</sup>

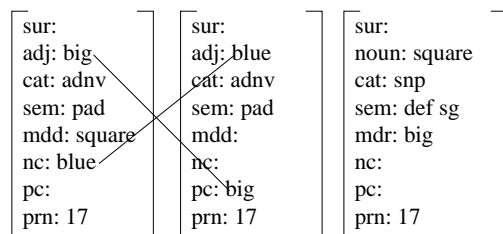
#### 3.1 CONCATENATION BY CROSS-COPYING



The nature of the semantic relation between *blue* and *square* is characterized by the attributes *mdr* (modifier) and *mdd* (modified). The relation is implemented by copying the core value of *square* into the *mdd* slot of *blue* and the core value of *blue* into the *mdr* slot *square*. In addition, the *prn* value of *blue*, here 17, is copied into the *prn* slot of the next word proplet *square*.

Next consider extending 3.1 to an intrapropositional coordination:

#### 3.2 COORDINATION IN *big blue square*



<sup>9</sup>A feature structure is nonrecursive if there is no recursive embedding of feature structures as values. Recursive feature structures are unsuitable for (i) contents with a coordination structure, (ii) the pattern matching needed for (a) modeling reference and (b) applying operation patterns to input, and (iii) storage and retrieval in a database. Unordered attributes are inefficient for computers and humans alike. Therefore recursive feature structures with unordered attributes are not used in DBS.

The relation of intrapositional coordination is coded by the **nc** (next conjunct) and **pc** (previous conjunct) attributes of the conjoined adjectives.

The diagonal lines in 3.2 are intended as optical support for the reader. Technically, however, they are redundant and may be omitted. The real method of establishing semantic relations in DBS is by addresses coded as values and implemented as pointers. This method makes the proplets forming a complex content order-free, allowing the database to store them independently of the semantic relations between them.

For example, no matter where the storage mechanism of the database puts the adnominal *big*, its modified may be found via the primary key, consisting of the **mdd** value **square** and the **prn** value **17**. Similarly, no matter where the noun *square* is stored, its modifier may be found via the **mdr** value **big** and the **prn** value **17**. And accordingly for the intrapositional coordination in 3.2.

As another example consider the content of *Julia knows John.*, represented as the following set of connected proplets:

### 3.3 CONTENT OF *Julia knows John.* AS A SET OF PROPLETS

sur: Julia noun: Julia cat: snp fnc: know prn: 625	sur: knows verb: know cat: #s3' #a' decl arg: Julia John prn: 625	sur: John noun: John cat: snp fnc: know prn: 625
--	---	--

The simplified proplets are held together by a common **prn** value, here **625**. The functor-argument is coded solely in terms of attribute values. For example, the *Julia* and *John* proplets specify their functor as **know**, while the *know* proplet specifies **Julia** and **John** as its arguments. Because of their nonempty **sur**(face) slots, the proplets are language proplets, in contradistinction to the proplets in 3.1 and 3.2, which are context proplets.

For storage and retrieval, a proplet is specified uniquely<sup>11</sup> by its **core** and **prn** values (primary key). This suggests a two-dimensional database schema, as in a classic network database (Elmasri and Navathe 2010). However, instead of using member and owner records, DBS uses member proplets and owner values.

The result is called a word bank. Its database schema consists of a column of owner values in their alphabetical order (vertical). Each owner value is preceded by an empty slot, called the now front, and a list of member proplets (horizontal); together they constitute a *token line*<sup>12</sup>

<sup>10</sup>The algorithm used for connecting (hear mode) and activating (think mode) proplets is time-linear Left-Associative Grammar (Hausser 1992). The **sur** attribute takes the language dependent surface as value. For a detailed description of the attributes and values used in proplets for describing English see NLC Appendix A3.

<sup>11</sup>Propositions containing two or more proplets with the same values, as in *Suzy loves Suzy*, require extra attention. They constitute a special case which (i) rarely occurs and (ii) is disregarded here because it may be easily handled by the software.

<sup>12</sup>The token line for any core value is found by using a trie structure (Briandais 1959; Fredkin

As an example, consider storing 3.3 as a nonlanguage content:

### 3.4 STORING THE PROPLETS OF 3.3 IN A WORD BANK

<i>member</i>	<i>proplets</i>	<i>now front</i>	<i>owner values</i>
...	noun: John		
	cat: snp		John
	fnc: know		
	prn: 625		
	noun: Julia		Julia
	cat: snp		
	fnc: know		
	prn: 625		
...	verb: know		
	cat: #n-s3' #a' decl		know
	arg: Julia John		
	prn: 625		

The proplets in a token line all have the same core value and are in the temporal order of their arrival, reflected by their *prn* values (NLC Sects. 11.2 and 11.3).

In contrast to the task of designing a practical schema for arranging the books in a private library, the sorting of proplets into a word bank is simple and mechanical. The letter sequence of a proplet's core value completely determines its token line for storage: the storage location for any new arrival is the penultimate position (now front) in the corresponding token line. When this slot is filled, the now front is reopened by moving the owner value one slot to the right (or, equivalently, pushing the member proplets one slot to the left, as in a push-down automaton).

By storing content like *sediment*, the stored data are never modified and any need for checking consistency is obviated. Changes of fact are written to the now front, like diary entries recording changes of temperature. Current data which refer to old ones use addresses as core values, implemented as pointers.

## 4 Cycle of Natural Language Communication

The transfer mechanism of content from the speaker to the hearer is based on external surfaces which have neither a meaning nor any grammatical properties (1.1). They must, however, belong to a language which the speaker and the hearer have each learned.

The learning enables the hearer to (i) recognize surfaces, (ii) use the recognized but otherwise unanalyzed surfaces for looking up lexical entries which provide the meaning and the grammatical properties, and (iii) connect them with the semantic relations of functor-argument and coordination. The learning enables the speaker

---

1960). The search for a proplet within a token line may use the *prn* value of the address in relation to the strictly linear increasing *prn* values. As pointed out by J. Handl, this may be based on binary search, in time  $O(\log(n))$  (Cormen et al. 2009), or interpolation, in time  $O(\log(\log(n)))$  (Weiss 2005), where  $n$  is the length of the token line.



to (i) navigate along the semantic relations between proplets, (ii) produce language-dependent word form surfaces from the core values of proplets traversed, and (iii) handle function word<sup>13</sup> precipitation and agreement.

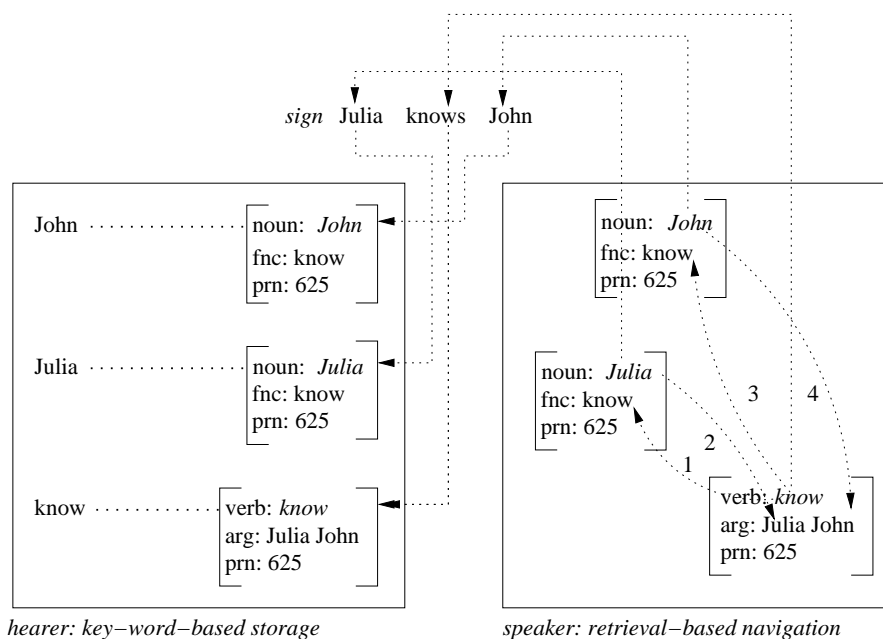
Successful communication between a speaker and a hearer is defined as follows:

#### 4.1 DEFINITION OF SUCCESSFUL COMMUNICATION

Natural language communication is successful if the content, mapped by the speaker into a sequence of external word form surfaces, is re-constructed and stored equivalently by the hearer.

Using the word bank content 3.4, the transfer of information from the speaker to the hearer, based solely (i) on unanalyzed external surfaces and (ii) the data structure and database schema of the agents' cognition, may be shown schematically as follows:

#### 4.2 NATURAL LANGUAGE TRANSFER MECHANISM



The speaker's navigation through a set of connected proplets serves as the conceptualization (*what to say*) and as the basic serialization (*how to say it*) of natural language production (McKeown 1985; Kass and Finin 1988). The hearer's interpretation consists in deriving a corresponding set of proplets, based on automatic word form recognition and syntactic-semantic parsing. The time-linear order of the

<sup>13</sup>Examples of function words in English are determiners like a(n), the, some, every, all, prepositions like in, on, above, below, auxiliaries like be, have, do, coordinating conjunctions like and, or, and subordinating conjunctions like that, who, which, when, because.

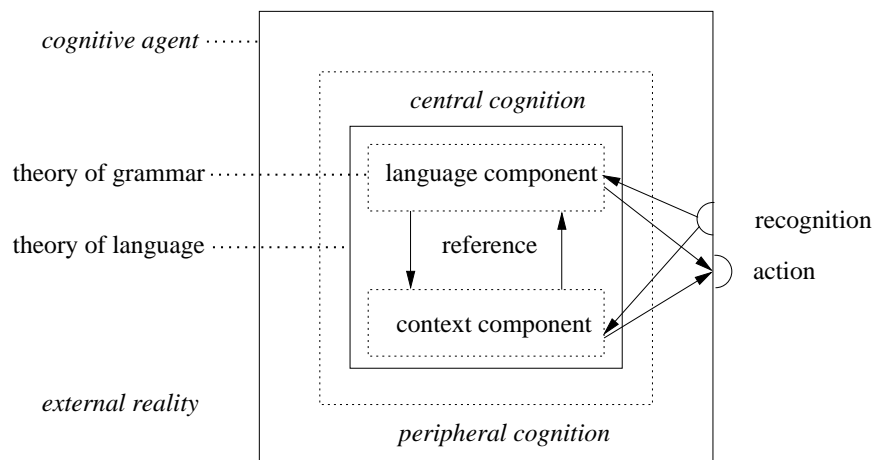
sign induced by the speaker’s navigation is eliminated in the hear mode, allowing storage of the proplets in accordance with the database schema of the content-addressable<sup>14</sup> word bank. When the agent switches into the speak mode, order is reintroduced by navigating along the semantic relations between the proplets.

## 5 Conceptual Reconstruction of Reference

In DBS, a cognitive content is defined as a set of proplets connected by address. Proplets with a non-empty *sur*(face) value (3.3) represent a language content. Proplets with an empty *sur* value (3.1) represent a context content. Otherwise, language and context proplets are alike. This holds specifically for their storage and retrieval in a word bank, which is based solely on their core value and order of arrival.

Conceptually, however, the schema 2.1 may be refined by (i) separating the levels of language and context, (ii) introducing the place of pragmatics as an interaction between the two levels, and (iii) distinguishing peripheral and central cognition.

### 5.1 CONCEPTUAL VIEW OF INTERFACES AND COMPONENTS



Externally, the agent’s interfaces for language and nonlanguage recognition are the same, as are those for language and nonlanguage action.<sup>15</sup> Internally, however, raw input data are separated by peripheral cognition into language and nonlanguage

<sup>14</sup>As a database, a word bank is content-addressable because it does not use an index (inverted file), in contrast to the widely used coordinate-addressable databases (RDBMS). See Chisvin and Duckworth (1992) for an overview.

<sup>15</sup>A differentiation into the sensory modalities (vision, audition, locomotion, manipulation) is omitted – not only for simplicity, but also because the meaning of a word or expression is independent of the modality of its external surface. For example, the meaning of the word **square** (2.4) is the same regardless of whether its surface is realized in speech, writing, or signing. A nonlanguage concept like the shape, color, taste, etc., of a blueberry may also be assumed to be independent from the modalities of its recognition (Hausser 2011, Sects. 2.2–2.4).

content (diagonal input arrows). Conversely in action, which realizes a content as raw output regardless of whether it originated at the language or at the context level (diagonal output arrows).

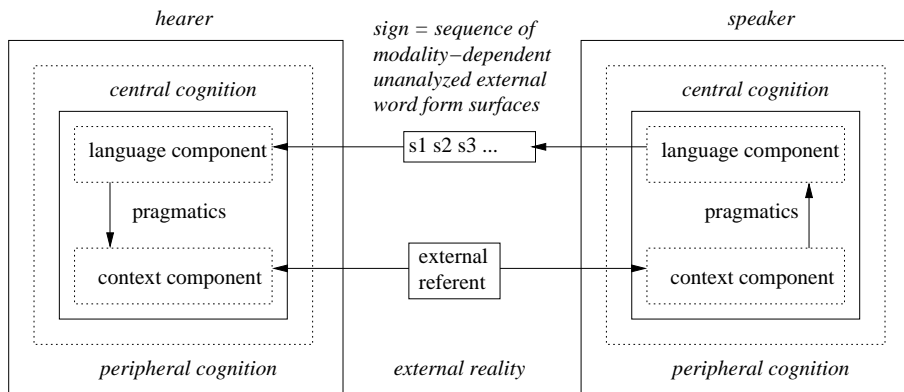
For example, as a sound pattern the surface **blue square** will have a meaningful interpretation at the language level by someone who has learned English, but be treated as an uninterpreted noise at the context level by someone who has not. Conversely, even though the action of denying entrance may be realized by telling to go away (originating at the language level) or by slamming the door (originating at the context level), both result in raw output data.

The distinction between the language and the context component provides a cognitive treatment of reference. Reference to an object in the agent’s current environment is called *immediate* reference, while reference to cognitive content existing only in the agents’ memory, for example, J.S. Bach, is called *mediated* reference. For mediated reference, the agent-based ontology of DBS (Sect. 1) is essential.

As an example of immediate reference consider a speaker and a hearer in a common task environment and looking at a blue square. If the speaker says **Take the blue square**, the noun phrase refers to the object in question. Similarly for the hearer, for whom fulfilling the request requires reference to the same object.

Postulating an external relation between a surface and its referent would be a reification fallacy. Instead we reconstruct immediate reference cognitively:

## 5.2 IMMEDIATE REFERENCE AS A PURELY COGNITIVE PROCEDURE



Immediate reference relies on the agents’ recognition interfaces for language (upper level) and nonlanguage content (lower level). Mediated reference, in contrast, relies on language recognition (upper level) and the existence of corresponding content in the agent’s memory. While immediate reference may be regarded as prototypical for the origin of language, it is a special case of mediated reference in that it has the additional requirement of context recognition.

### Terminological Remark

Computer Science uses the term “reference” differently from analytic philosophy and linguistics. A computational reference is an address in a storage location. This may coded as (i) a symbolic address (declarative) or as (ii) a pointer to a physical storage location in the

memory hardware (procedural). The term “generalized reference” is used in computational image reconstruction (computer vision).

In DBS, the term “reference” is used in the sense of philosophy and linguistics. However, the term is generalized insofar as no agent-external “representational token” is required.

Recanati (1997, 2004), Pelczar and Rainsbury (1998) and others use the term “generalized reference” for an analysis of the sign kind name which allows the surface *Mary*, for example, to refer to several different individuals. This is in contradistinction to Russell (1905) whose definite description<sup>16</sup> analysis of “proper” names requires a unique referent.

The DBS analysis of names in Sect. 8 also allows different referents (8.4). However, while the “generalized reference” of Recanati, Pelczar et al., and others is based on assimilating names to the parameters of indexicals, the DBS analysis is traditional in that it is based on an act of baptism which is generalized in that it may occur implicitly as well as explicitly. Moreover, generalized reference in DBS is not limited to names, but includes reference by means of concept (symbol) and pointer (indexical). In addition, DBS treats coreference by means of address, which occurs as a variant of symbol and pointer reference and underlies name-based reference.

## 6 Reference by Matching (Symbol)

Natural language uses three reference mechanisms, based on (i) matching, (ii) pointing, and (iii) baptism. The reference mechanism based on matching uses the type-token relation (2.2, 2.4) and is associated with the sign kind *symbol*.<sup>17</sup> For example, the terms *a blue square* and *blue squares* in the sentence sequence *John saw a blue square. ... Blue squares are rare.* are related as follows:

### 6.1 REFERENCE WITH LANGUAGE PROPLETS IN TOKEN LINES

	<i>member proplets</i>	<i>now front</i>	<i>owner values</i>
$\left[ \begin{array}{l} \text{sur:} \\ \text{adj: blue (token)} \\ \text{mdd: square} \\ \text{prn: 41} \end{array} \right]$	...	$\left[ \begin{array}{l} \text{sur: blue} \\ \text{adj: blue (type)} \\ \text{mdd: square} \\ \text{prn: 48} \end{array} \right]$	blue
...			
$\left[ \begin{array}{l} \text{sur:} \\ \text{noun: square (token)} \\ \text{sem: indef sg} \\ \text{fnc: see} \\ \text{mdr: blue} \\ \text{prn: 41} \end{array} \right]$	...	$\left[ \begin{array}{l} \text{sur: square} \\ \text{noun: square (type)} \\ \text{sem: indef pl} \\ \text{fnc: rare} \\ \text{mdr: blue} \\ \text{prn: 48} \end{array} \right]$	square

Stored in the agent’s word bank, the vertical relation between the language and the context component shown in 5.1 reappears as a horizontal relation between proplets within token lines. The proplets with the prn value 48 are the language

<sup>16</sup>For a modern variant of this approach see Burgess 2012.

<sup>17</sup>We follow the terminology of Peirce (CP 2.228, 2.229, 5.473).

proplets (non-empty *sur* values), those with the *prn* value 41 are the context proplets (empty *sur* slots). Reference by matching holds between the two *blue* proplets with the *prn* values 41 and 48 and similarly between the two *square* proplets. The distinction between the type and the token, here indicated after the core values, is usually left implicit.

The functor-argument relation<sup>18</sup> between the proplets *blue* and *square* is coded by the features [*mdd*: *square*] and [*mdr*: *blue*], respectively. The noun proplet with the feature [*sem*: *indef sg*] is an indefinite singular, that with the feature [*sem*: *indef pl*] is an indefinite plural.

Next consider the same reference relation without language. To emphasize the missing *sur* values, “ $\emptyset$ ” is used:

## 6.2 REFERENCE BY MATCHING WITHOUT LANGUAGE

	<i>member proplets</i>	<i>now front</i>	<i>owner values</i>
$\left[ \begin{array}{l} \text{sur:} \\ \text{adj: blue (token)} \\ \text{mdd: square} \\ \text{prn: 41} \end{array} \right]$	...	$\left[ \begin{array}{l} \text{sur: } \emptyset \\ \text{adj: blue (type)} \\ \text{mdd: square} \\ \text{prn: 48} \end{array} \right]$	blue
...			
$\left[ \begin{array}{l} \text{sur:} \\ \text{noun: square (token)} \\ \text{sem: indef sg} \\ \text{fnc: see} \\ \text{mdr: blue} \\ \text{prn: 41} \end{array} \right]$	...	$\left[ \begin{array}{l} \text{sur: } \emptyset \\ \text{noun: square (type)} \\ \text{sem: indef pl} \\ \text{fnc: rare} \\ \text{mdr: blue} \\ \text{prn: 48} \end{array} \right]$	square

Here the reference relation holds between two nonlanguage contents – and not between a language content (meaning<sub>1</sub>) and a nonlanguage content, as in 6.1.

Even though the reference relation is established between two individual proplet pairs in the same token lines, the combination into the complex content corresponding to **blue square** is accommodated as well: in order to match, the two *blue* proplets must not only have the same<sup>19</sup> core value, but also the same *mdd* continuation value, here *square*, and correspondingly for the *mdr* values of the two *square* proplets. Their *fnc* and *prn* values, however, are different.

Generalizing reference by matching to include referring with nonlanguage content results in the following constellations:

## 6.3 CONSTELLATIONS OF GENERALIZED REFERENCE

### 1. *Nonlanguage content referring to nonlanguage content*

Example: Agent sees something and identifies it with something seen before.

### 2. *Language content referring to nonlanguage content*

Example: Agent describes a landscape in speak mode.

<sup>18</sup>Apparently, Aristotle struggled to reconcile reference with content combination (Modrak 2001).

<sup>19</sup>Disregarding the type-token distinction.

3. *Nonlanguage content referring to language content*

Example: Agent identifies a current nonlanguage recognition with something it has read (for example, in a guide book) or heard about before.

4. *Language content referring to language content*

Example: Agent describes what it has heard or read.

Cognitive agents without language are capable of reference constellation 1 only, while agents with language use all four.

## 7 Reference by Pointing (Indexical)

The second reference mechanism of cognition is based on pointing. In natural language, it is illustrated by the indexical signs, such as the pronouns. The first step toward a computational implementation is the linguistic observation that the indexicals point at only five different parameters, namely (1) first person, (2) second person, (3) third person, (4) place, and (5) time.

In English, the pronouns *I*, *me*, *mine*, *we*, and *us* point at the parameter for first person. The pronoun *you* points at the parameter for second person. The pronouns *he*, *him*, *his*, *she*, *her*, *it*, *they* and *them* point at the parameter for third person. The indexical adjs *here* and *there* point at the parameter for place. The indexical adjs *now*, *yesterday*, and *tomorrow* point at the parameter for time.

The indexical nouns pointing at the parameters of first, second, and third person are varied by grammatical distinctions. Consider the following examples illustrating grammatical variation in 1st person pronouns of English:

### 7.1 1ST PERSON PRONOUNS DISTINCTIONS

sur: I	sur: me	sur: we	sur: us
noun: pro1	noun: pro1	noun: pro1	noun: pro1
cat: s1	cat: obl	cat: p1	cat: obl
sem: sg	sem: sg	sem: pl	sem: pl
fnc:	fnc:	fnc:	mdr:
mdr:	mdr:	mdr:	fnc:
nc:	nc:	nc:	nc:
pc:	pc:	pc:	pc:
prn:	prn:	prn:	prn:

The proplets all share the indexical pointer `pro1` as their core value. The different `cat` values `s1` (first person singular), `p1` (first person plural), and `obl` (oblique) control verb agreement, preventing, for example, ungrammatical *\*Me saw a tree* or *\*Peter saw we*. Ungrammatical *\*I sees a tree* and *\*he see a tree* are prevented by using the different `cat` values `s1` (singular 1st person) and `s3` (singular 3rd person).

Noun proplets of the sign kind indexical combine in the same way into propositions as proplets of the sign kind symbol or name. Consider the DBS analysis of English I heard you.:

## 7.2 REPRESENTING I heard you. AS A LANGUAGE CONTENT

sur: I noun: pro1 cat: s1 sem: sg fnc: hear mdr: nc: pc: prn: 71	sur: heard verb: hear cat: #n' #a' decl sem: past arg: pro1 pro2 mdr: nc: pc: prn: 71	sur: you noun: pro2 cat: sp2 sem: fnc: hear mdr: nc: pc: prn: 71
--	---	--

The question raised by this example is how the indexical pointers **pro1** and **pro2** are to be interpreted pragmatically relative to a context of use.

This leads to the second step of modeling the indexical reference mechanism. It is based on combining a propositional content with a cluster of parameter values of the agent's current STAR (Hausser 1999 Sect. 5.3). The STAR is an acronym for (i) location (**S**pace), (ii) time (**T**ime), (iii) self-identity (**A**gent), and (iv) intended addressee (**R**ecipient)..

The STAR has two functions: (a) keeping track of the agent's current situation (orientation) and (b) providing referents for indexicals occurring in contents.<sup>20</sup> A STAR is coded as a proplet with the value of the A attribute serving as the core value and as the owner. In a word bank, the temporal sequence of STARs is listed as a token line:

## 7.3 TOKEN LINE EXAMPLE<sup>21</sup> OF STARS DEFINED AS PROPLETS

<i>member proplets</i>				<i>now front</i>	<i>owner value</i>
S: kitchen T: t <sub>1</sub> - t <sub>2</sub> A: Sylvester R: Speedy 3rd: prn: 63-70	S: kitchen T: t <sub>2+1</sub> - t <sub>3</sub> A: Sylvester R: Speedy 3rd: prn: 71-78	S: living room T: t <sub>3+1</sub> - t <sub>4</sub> A: Sylvester R: Tweety 3rd: Speedy prn: 79-82	S: garden T: t <sub>4+1</sub> - t <sub>5</sub> A: Sylvester R: Hector 3rd: prn: 83-87	Sylvester	

<sup>20</sup>Integrating the interpretation of indexicals into the agent's on-board STAR orientation may be seen as an enhancement of Montague (1973), who's sign-oriented approach uses arbitrary parameter values, i.e.  $i \in I$  for a moment of time and  $j \in J$  for a possible world (space, location), superscripted at the end of a lambda expression. Introducing additional parameters for 1st, 2nd, and 3rd person, as has been suggested, has been made light of by Cresswell's (1972, p. 4) joking proposal of a "next drink parameter." The parameter approach has resurfaced as "Variablism," i.e. the view that names should be treated semantically as variables, like pronouns (Cumming 2008).

<sup>21</sup>Instead of the usual subjects John, Mary, etc., the following examples use the animation characters of Sylvester the cat, Speedy Gonzales the mouse, Tweety the bird, and Hector the dog, familiar from TV, as an aid to distinguish the individuals pointed at by indexicals.

In addition to attributes represented by the letters of the STAR, there is a fifth, called **3rd**, for third person indexicals. Though not required for the agent’s basic orientation, **3rd** is needed to provide the referent for items which are neither 1st nor 2nd person.<sup>22</sup> As indicated by the **prn** values, e.g. [**prn**: 63–70], several consecutive propositions may share the same STAR.

In natural language communication, three perspectives on content must be distinguished (Hausser 2011 Chaps. 10, 11). The STAR-0 is the agent’s perspective onto its current environment; it need not involve language. The STAR-1 is the agent’s speak mode perspective onto stored content as required for language production; if ongoing events are reported directly, the STAR-1 equals the STAR-0. The STAR-2 is the agent’s hear mode perspective onto language content as needed for the correct interpretation of indexicals. As an example of a STAR-0 perspective, consider the non-language content corresponding to *I hear you*:

#### 7.4 ANCHORING A CONTENT TO A STAR-0

sur: noun: pro1 cat: s1 sem: sg fnc: hear ... prn: 63	sur: verb: hear cat: #n' #a' decl sem: pres arg: pro1 pro2 ... prn: 63	sur: noun: pro2 cat: sp2 sem: fnc: hear ... prn: 63	S: kitchen T: 2013-09... A: Sylvester R: Speedy 3rd: prn: 63
---	--	---	---

This content differs from 7.2 because (i) it is nonlanguage (no **sur** values), (ii) the **sem** value of the verb is **pres** (present tense) rather than **past**, and (iii) a STAR is attached by having the same **prn** value as the content, here **63**.

The STAR shows the perspective of the agent Sylvester on his current environment. The **S** value specifies the location as the kitchen, the **pres** value of the verb points at the **T** value, the indexical **pro1** points at the **A** value Sylvester the cat, and **pro2** points at the **R** value Speedy the mouse.

Next Sylvester realizes the content in language by saying to Speedy *I heard you*. As time has moved, the language content 7.2 is anchored to a second, later STAR-0 with the **prn** value **71** (7.3). From these two STAR-0, the agent computes the following STAR-1 perspective for the language content 7.2:

#### 7.5 SPEAK MODE ANCHORING TO A STAR-1

sur: I noun: pro1 cat: s1 sem: sg fnc: hear ... prn: 64	sur: heard verb: hear cat: #n' #a' decl sem: past arg: pro1 pro2 ... prn: 64	sur: you noun: pro2 cat: sp2 sem: fnc: hear ... prn: 64	S: kitchen T: 2013-09... A: Sylvester R: Speedy 3rd: prn: 64
---	--	---	---

<sup>22</sup>According to King (2014), a context consists of time, location, agent, and world. In the STAR



The agent's perspective is looking from its present situation back on the stored content 7.4 and encoding it in language. The content is connected to the agent's current STAR via the common *prn* value 64. The content differs from that of 7.4 in (i) the *sem* value *past* (rather than *pres*) of the verb proplet and (ii) the language-dependent *sur* values of the content proplets.

When the language content *I heard you* is interpreted by the addressee (recipient), Speedy the mouse uses the content of the language sign and its current STAR-0 to derive the STAR-2 perspective. The result is as follows:

## 7.6 STAR-2 PERSPECTIVE IN HEAR MODE

$\left[ \begin{array}{l} \text{sur:} \\ \text{noun: pro2} \\ \text{cat: sp2} \\ \text{sem:} \\ \text{fnc: hear} \\ \text{mdr:} \\ \text{nc:} \\ \text{pc:} \\ \text{prn: 53} \end{array} \right]$	$\left[ \begin{array}{l} \text{sur:} \\ \text{verb: hear} \\ \text{cat: \#n' \#a' decl} \\ \text{sem: past} \\ \text{arg: pro2 pro1} \\ \text{mdr:} \\ \text{nc:} \\ \text{pc:} \\ \text{prn: 53} \end{array} \right]$	$\left[ \begin{array}{l} \text{sur:} \\ \text{noun: pro1} \\ \text{cat: obl} \\ \text{sem: sg} \\ \text{fnc: hear} \\ \text{mdr:} \\ \text{nc:} \\ \text{pc:} \\ \text{prn: 53} \end{array} \right]$	$\left[ \begin{array}{l} \text{S: kitchen} \\ \text{T: 2013-09...} \\ \text{A: Speedy} \\ \text{R: Sylvester} \\ \text{3rd:} \\ \text{prn: 53} \end{array} \right]$
---	--	--	---

Speedy as the interpreting agent uses his personal *prn* value and a different STAR: compared to the STAR of Sylvester, the A and R values are reversed and Sylvester's *I heard you* is reinterpreted by Speedy's STAR-2 perspective as *you heard me* (Hausser 2011 Chap. 10).

## 8 Reference by Baptism (Name)

Just like the other two reference mechanisms, the baptism-based reference of names is implemented as a cognitive operation. It consists of (i) establishing *object permanence* and (ii) *baptism* based on cross-copying between a name and its referent.

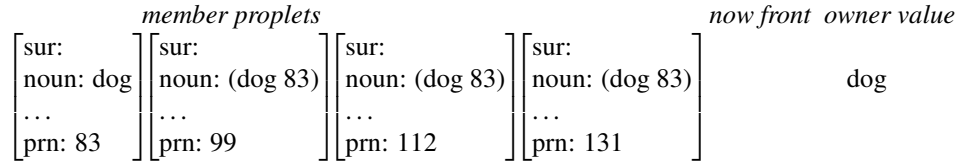
For implementing object permanence,<sup>23</sup> DBS uses identity by address. For example, when the robot observes an unfamiliar dog running through the bushes, it must understand that the different appearances are instances of the same referent. This recognition interpretation, which is based on yet another type-token relation, is coded by using an address as the core value of the non-initial proplets, pointing at the proplet representing the initial appearance of the referent:

---

of DBS, S corresponds to King's location, T to time, and A to agent. The DBS counterpart to King's world are R (intended recipient, you) and 3rd (everyone and everything that is neither A nor R). However, DBS distinguishes between the STAR parameters as the agent's on-board orientation system and basis for interpreting indexicals, on the one hand, and the context as an activated content in memory, on the other.

<sup>23</sup>The notion of object-permanence originated in cognitive psychology (Piaget 1954). It may be regarded as a non-truthconditional, non-modal counterpart to "rigid designators" (Kripke 1972).

## 8.1 OBJECT PERMANENCE BY USING ADDRESS



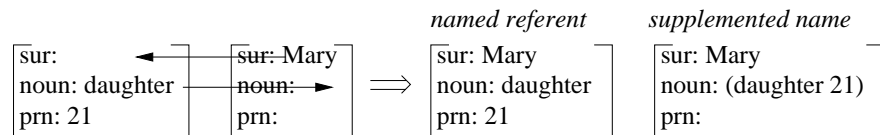
The different **prn** values indicate that each member proplet is part of a different proposition, allowing different continuation values. The extrapositional address core values (**dog 83**) of the non-initial member proplets point at the initial proplet, which is the referent and formally recognizable by its *non-address*<sup>24</sup> core value.

A token line like 8.1 may contain several initial dog referents, each referring to another individual. They are distinguished by their different **prn** values and the address numbers of the associated coreferent proplets. This is sufficient for the agent to properly discriminate between different dog referents in cognition and between their sets of coreferent proplets, all in the same token line.<sup>25</sup>

It is not sufficient, however, for language communication. This is because the **prn** values of referents are not synchronized between agents: agents may encounter the same referent at different occasions, and the number of coreferent items may differ between agents. What is needed is a name surface and an interagent consensus on which item(s) the name refers to. The consensus is simply achieved: the not yet initiated agent follows the practice observed.

The DBS implementation is based on (i) a lexical name proplet which has a **sur**(face) value but no core value and (ii) a connected referent proplet which has a core value but no **sur** value. The two proplets are supplemented by an event of generalized baptism which cross-copies the **sur** value of the name into the **sur** slot of the referent and the core value of the referent into core slot of the name.

## 8.2 BAPTISM AS CROSS-COPYING



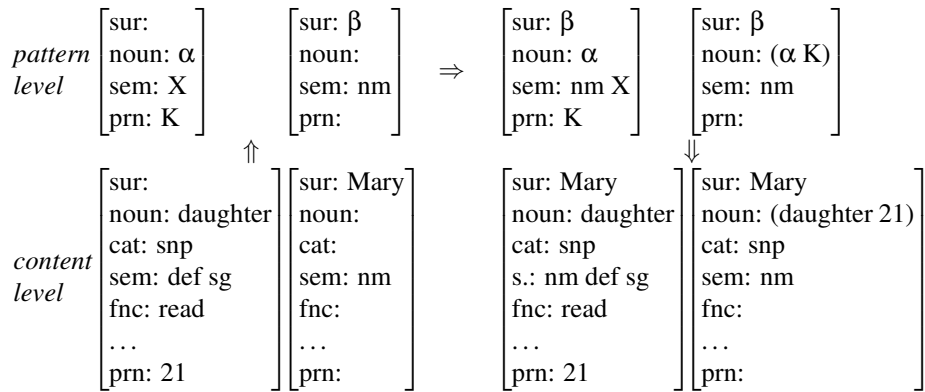
The named referent proplet is stored in the token line of the core value and used in the speak mode. The supplemented name proplet is stored in the token line of the surface and used in the hear mode.

The baptizing event is formalized as the following DBS operation:

<sup>24</sup>While continuation values are always addresses, core values may be non-address or address.

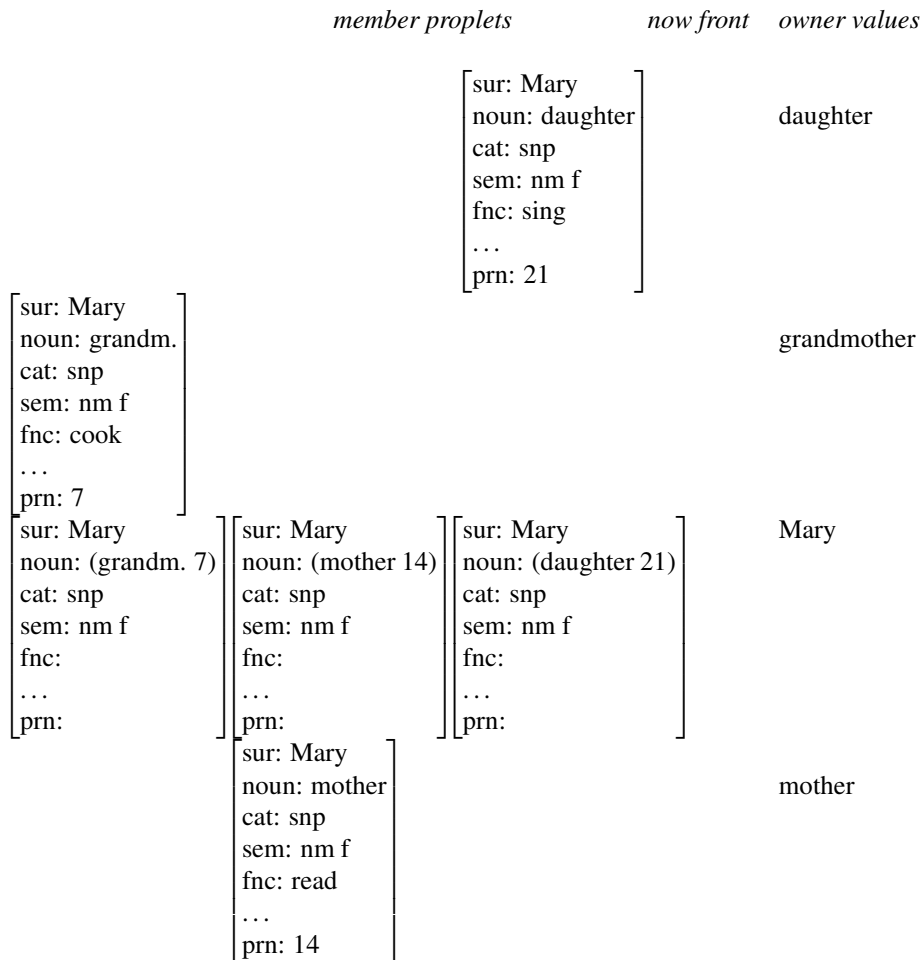
<sup>25</sup>In agents without language, it is a pre-stage of naming (part of case 1 in 6.3).

### 8.3 APPLYING THE FORMAL BAPTIZING OPERATION



The third proplet at the content level is the named referent. The fourth proplet is the supplemented name. As an example, consider the following word bank:

### 8.4 NAME REFERRING WITH MULTIPLE REFERENTS



The token lines are in the alphabetical order *daughter*, *grandmother*, *Mary*, *mother*. The member proplets show the result of three baptism operations like 8.3.<sup>26</sup> In the token line of *Mary*, each supplemented name proplet occurs only once. Supplemented names are not written into the lexicon because a core value like (daughter 21) is not a convention of the natural language at hand. Instead it is the result of a generalized baptism event witnessed by the agent. Nevertheless, the supplemented name proplets in the *Mary* token line have a lexical quality in that they have neither continuation nor prn values – like the lexical proplets resulting from automatic word form recognition.<sup>27</sup>

When the hearer interprets a sentence containing a name, the name activates the corresponding token line, here that of *Mary*. The choice between different referents, here the grandmother, the mother, and the daughter, may have one of the following results: (i) the chosen referent equals the one intended by the speaker (correct choice), (ii) does not equal the one intended by the speaker (incorrect choice), or (iii) no referent is chosen (inconclusive result). The choice between multiple potential name referents is usually not at random, however. Instead, the referent most suitable to the utterance situation will usually be the correct one. If uncertainty remains, the hearer may ask the speaker for clarification.

For the speaker, there is no ambiguity. Instead, the speaker selects the intended referent, e.g. (daughter 21). If the appropriate name was acquired in the hear mode (8.3), it is preserved in the word bank (8.6) and may be used for language production. If the speaker is in the position to select and bestow a name, it is also available for realization.

Depending on whether or not an agent’s word bank already has a token line (i) for a certain name and (ii) proplets for its referent(s), the following case distinctions may be described in the name interpretation of the hearer:

## 8.5 CASE DISTINCTIONS OF NAME REFERENCE

1. the name and the referent have been used before,
2. the name has been used before, but the referent is new,
3. the name has not been used before, but the referent exists, or
4. neither the name nor the referent have been used before by the agent.

As an example of case 1, consider agent A observing daughter *Mary* eating a cookie. Later agent A reports to agent B *Mary ate a cookie*. Assuming that

---

<sup>26</sup>For clarity, each supplemented name proplet and its referent are shown in the same column. In reality, token lines order proplets contiguously, without any empty positions, solely in accordance with the proplets’ arrival. Thus, token lines are “dense” and independent of each other.

<sup>27</sup>Our computational reconstruction of the baptism event based on cross-copying differs from other theories of naming such as the descriptive theory of proper names (Russell 1905) and the rigid designator analysis (Kripke 1972), which are based on a sign-based ontology (FoCL Sect. 20.4). The formal DBS treatment of baptism-based reference (8.2) provides a simple, efficient procedural implementation suitable for building a talking robot.

agent B knows the intended referent of *Mary*, the following proposition would be stored in the word bank of agent B (hear mode):

### 8.6 KNOWN NAME REFERS TO KNOWN REFERENT

sur: mary noun: (daughter 21) fnc: eat prn: 102	sur: verb: eat arg: (daughter 21) cookie prn: 102	sur: noun: cookie fnc: eat prn: 102
--	--	--

The proplets are held together by the common `prn` value 102. As a named referent, the first proplet is stored in the token line of the core value. The `sur` values are not preserved in storage.<sup>28</sup> Instead they will be provided by language-dependent lexicalization rules in the speak mode. The address values point at the initial referent.

The second case is illustrated by agent B knowing the grandmother and mother referents of *Mary*, but not the daughter referent. This means that the token line of *Mary* exists in the word bank of agent B, though without a supplemented name proplet pointing at the referent proplet *daughter*. The token line of *daughter* may not even contain a content proplet suitable as the referent. The two missing proplets may be added to agent's B word bank by agent A saying Did you know that *Mary's daughter is called Mary as well?*

The third case is illustrated by an unfamiliar dog running through the bushes. Its proplet representation is added to the existing token line for *dog*. This still unnamed initial referent is recognizable by its non-address core value. When the dog is called *Fido*, the name is written into the `sur` slot of initial dog proplet (named referent, 8.3) and the supplemented name proplet is stored in the *Fido* token line. If the name has not been used before, a new token line is created for it by defining a token line using the name's surface as the owner value.

The fourth case is illustrated by meeting an unknown animal in a zoo and learning that it is called *Rosie*. Thus a word bank must provide for two dimensions of continuous extension: (i) the time-linear lengthening of token lines (horizontal) and (ii) the insertion of new token lines into the column of owner values (vertical).

## 9 Reference by Address (Coreference)

Coreference by address occurs with all three sign kinds. In reference by baptism (name) it is the only mechanism for relating the supplemented name to the named referent. In reference by matching (symbol) and by pointing (indexical), in contrast, it is an additional mechanism. Let us begin with the coreferential interpretation of a symbol in the sequence a unicorn ... the unicorn:

---

<sup>28</sup>The name proplet preserves the `sur` value resulting from the baptism operation as a marker, written in default font and in lower case. The marker is used for surface realization in the speak mode.

## 9.1 COREFERENCE IN A TOKEN LINE

[noun: unicorn sem: indef sg prn: 23]	<i>member proplets</i> ...	[noun: (unicorn 23) sem: def sg prn: 26]	<i>now front</i>	<i>owner value</i>	unicorn
---	-------------------------------	--	------------------	--------------------	---------

The coreference relation is coded by using the address of the first proplet as the core value of the second. The second proplet must be a definite noun.

The other sign kind with a coreferential interpretation in addition to its own reference mechanism is the indexical. The coreferential interpretation of a pointer (e.g. 3rd person pronoun) uses the address of a preceding (antecedent) or following (postcedent) referent, instead of referring to the STAR. Reusing the example in Sect. 7, consider the extrapositional coordination **Speedy hid in the cupboard. I heard him.** uttered by Sylvester and interpreted by Hector.

The question is whether **him** refers indexically, e.g. to Tweety the bird, or coreferentially to **Speedy** the mouse. For the speaker Sylvester the answer is clear because language expressions are not ambiguous in the speak mode. For Hector as the hearer, however, the answer is unknown.

Formally, the two possible hear mode interpretations compare as follows:

## 9.2 INDEXICAL INTERPRETATION OF him

[noun: Speedy fnc: hide prn: 78]	[verb: hide arg: Speedy mdr: cupboard nc: (hear 79) prn: 78]	[noun: cupboard mdd: hide prn: 78]
[noun: pro2 fnc: hear prn: 79]	[verb: hear arg: pro2 pro3 pc: (hide 78) prn: 79]	[noun: pro3 fnc: hear prn: 79]
[S: garden T: 2013-09-23T08:20:00 A: Hector R: Sylvester 3rd: Tweety prn: 78-79]		

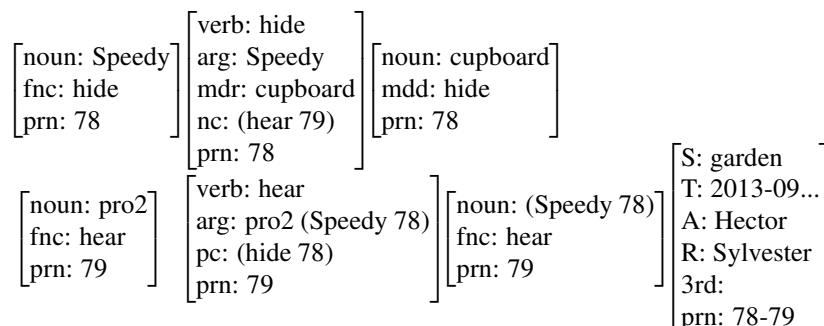
This hear mode interpretation uses Hector's STAR with a STAR-2 conversion<sup>29</sup> of the content from pro1 to pro2. Represented as pro3, him points at the Tweety value of the 3rd attribute in the STAR.<sup>30</sup>

Next consider Hector's coreferential interpretation of **him** as referring to **Speedy**:

<sup>29</sup>Hausser 2011, Chap. 10.

<sup>30</sup>The indexical interpretation of pro2 and pro3 is defined to refer to the R and 3rd values, respectively, of the STAR. For a more extensive treatment see Hausser (2011, Chaps. 10 and 11).

### 9.3 COREFERENTIAL INTERPRETATION OF him



The representation of the first proposition is the same as in 9.2. The STAR is different, however, in that its 3rd attribute has no value. Represented as the address value (Speedy 78), him is interpreted coreferentially.

#### Case Study

The indexical reference mechanism has been illustrated as a case study in a fictional dialog between Jean Paul Sartre and Simone de Beauvoir in Chapter 10 of Hausser (2011). The dialog illustrates the production and interpretation of statements, of WH and Yes/No questions with answers, and of a request with fulfillment. The indexicals are interpreted with explicit STARs for the speak and the hear mode.

## 10 De Dicto/De Re

A puzzle in ancient philosophy of language has been called the *de dicto/de re* ambiguity by medieval logicians, e.g. Thomas Aquinas (1225–1274) and William of Occam (1288-1347).<sup>31</sup> More recently, roughly the same distinction has been called *uneven/even* by Frege (1892), *opaque/transparent* by Quine (1960), and *intensional/extensional* by Montague (1973). The de dicto/de re ambiguity occurs with certain verbs and modal operators.

The de dicto reading arising with verbs constitutes an exception to two fundamental laws of medieval semantics, known as (i) Existential Generalization and (ii) Substitutivity of Identicals (Hausser 1999 Sect. 20.4). Existential Generalization fails if the meaning of a noun is equated with its real world referent (ontological assumption) and the referent doesn't exist.

This is illustrated by Montague (op. cit.) as follows:

#### 10.1 DIFFERENT ENTAILMENTS OF EXISTENCE

##### a. *De re*

John **finds** a unicorn. > There exists a unicorn.

##### b. *De dicto/de re*

John **seeks** a unicorn.  $\not>$  There exists a unicorn.

In example **b**, the entailment fails ( $\not>$ ) if John is looking for any old unicorn (de dicto), because there might not exist one. It is also possible, however, that John is looking for a particular unicorn, in which case the referent exists (de re). Example **a**, in contrast, has only a de re reading ( $>$ ).

This difference between examples **a** and **b** is caused by their different verbs. In the terminology of Montague, *find* creates an extensional context in which truth of the premise entails truth of the conclusion, while *seek* creates an intensional context in which truth of the premise does not. Technically, Montague treats the de dicto/de re alternative produced by an intensional verb as a scope ambiguity based on the position of the quantifier used to represent **a unicorn** in predicate calculus.

DBS takes a different approach: (i) reference is treated as a purely cognitive relation between two contents (6.3), (ii) there are neither quantifiers nor any of the scope ambiguities they create (Hausser 2006, Sect. 6.4), and (iii) ambiguity arises only in recognition, including the hear mode,<sup>32</sup> but not in action.

For example, if the agent is looking for a spoon (action) in an unfamiliar house, this content is represented as follows:

## 10.2 NONLANGUAGE ACTION USING DE DICTO REFERENCE

<i>member proplets</i>	<i>now front</i>	<i>owner value</i>			
...	<table style="border-left: 1px solid black; border-right: 1px solid black; border-bottom: 1px solid black; padding: 2px 5px;"> <tr> <td style="padding: 2px 5px;">noun: spoon <i>type</i></td> </tr> <tr> <td style="padding: 2px 5px;">fnc: seek</td> </tr> <tr> <td style="padding: 2px 5px;">prn: 67</td> </tr> </table>	noun: spoon <i>type</i>	fnc: seek	prn: 67	spoon
noun: spoon <i>type</i>					
fnc: seek					
prn: 67					

The search would be successful if the context were to provide a token matching the type (6.2). However, because the fnc value is the intensional verb *seek*, no assumption about the outcome is made at this point.

In a familiar house, in contrast, the agent may be looking for a particular spoon. In DBS, this de re constellation is represented as the following token line:

## 10.3 NONLANGUAGE ACTION USING DE RE REFERENCE

<i>member proplets</i>	<i>now front</i>	<i>owner value</i>			
...	<table style="border-left: 1px solid black; border-right: 1px solid black; border-bottom: 1px solid black; padding: 2px 5px;"> <tr> <td style="padding: 2px 5px;">noun: spoon <i>token</i></td> </tr> <tr> <td style="padding: 2px 5px;">fnc: have</td> </tr> <tr> <td style="padding: 2px 5px;">prn: 67</td> </tr> </table>	noun: spoon <i>token</i>	fnc: have	prn: 67	spoon
noun: spoon <i>token</i>					
fnc: have					
prn: 67					
...	<table style="border-left: 1px solid black; border-right: 1px solid black; border-bottom: 1px solid black; padding: 2px 5px;"> <tr> <td style="padding: 2px 5px;">noun: (spoon 67)</td> </tr> <tr> <td style="padding: 2px 5px;">fnc: seek</td> </tr> <tr> <td style="padding: 2px 5px;">prn: 93</td> </tr> </table>	noun: (spoon 67)	fnc: seek	prn: 93	
noun: (spoon 67)					
fnc: seek					
prn: 93					

The second *spoon* proplet, here with the prn value **93**, has an antecedent serving as the initial referent, here with the prn value **67**. The agent refers by address to the spoon being searched for (coreference, Sect. 9).

If the action is language production, e.g. uttering *I am looking for a spoon*, the speaker knows full well whether the intended reference mechanism is 10.2 or 10.3. In the hear mode, in contrast, there is a pragmatic ambiguity. Communication is successful if the hearer selects the reading meant by the speaker.

<sup>31</sup>For an overview see Dutilh Novaes Dutilh (2003).

<sup>32</sup>Even an ambiguity deliberately created in the speak mode (“diplomatic ambiguity,” Pehar 2001) arises only for the hearer.



In DBS, the failing Existential Generalization on the hearer's de dicto interpretation (10.1) is implemented as an inference. The antecedent pattern ensures that the input noun proplet (i) has a non-address core value, (ii) is indefinite, and (iii) has an intensional verb as its fnc value. The consequent pattern derives the new proposition *the unicorn might not exist*.

The valid Existential Generalization of the corresponding de re interpretation is also derived by a hear mode inference. The antecedent pattern ensures that there is an initial referent which the *unicorn* proplet is coreferent with (as in 10.3). The consequent pattern derives the new proposition *the unicorn exists* (no negation in the sem slot of the verb).

The de dicto/de re distinction in its application to a transitive verb construction may be summarized as follows. If the verb is extensional or the object noun is definite, the interpretation is de re for speaker and hearer. If the verb is intensional (in the sense of Montague) and the noun is indefinite, in contrast, the interpretation depends on whether or not the indefinite noun has an antecedent, i.e. an initial referent which it refers to by address (coreference). If the answer is yes (10.3), the interpretation is de re, otherwise de dicto (10.2). For the speaker, who knows the answer, the intended interpretation is a matter of choice. For the hearer, who doesn't know the answer, the alternative interpretations constitute a pragmatic ambiguity.

Finally consider the nature of existence, the entailment of which is at the center of the de dicto/de re distinction (10.1). In the ontology of an agent-based approach (Sect. 5), there are two kinds: (i) real existence, as in immediate reference (5.2), and (ii) assumed existence, as in mediated reference. For example, when the noun *unicorn* is used de re in a piece of fiction (mediated reference), the unicorn exists in the story, but the kind of existence is fictional in that no agent-external counterpart is entailed.

## Conclusion

The construction of a talking robot (artificial cognitive agent with language) is an interdisciplinary project which requires the cooperation of linguistics, philosophy, psychology, artificial intelligence, cybernetics, mathematical complexity theory, and robotics. It challenges the participating sciences to cooperate within a consistent, comprehensive, computationally verifiable framework.

The linguistic part of designing a talking robot is reconstructing the cycle of natural language communication, consisting of the hear, the think, and the speak mode (Sect. 4). The philosophical part includes the computational reconstruction of the different mechanisms of generalized reference, as shown here in Sects 6–9. The part of cybernetics and artificial intelligence is the design of an autonomous control which is capable of dealing with unfamiliar situations in the agent's ecological niche, with long-term survival as the standard of success.

These are theoretical tasks which require the design of a data structure, a database

schema, an algorithm, and a functional flow connecting the agent's recognition and action. They are essential for the hardware development of an artificial agent because they provide the robotic interfaces for recognition and action (1.1) with an agent-internal cognitive system to map into and out of.

A continuous cycle of upscaling and computational verification constitutes a self-correcting research strategy. Based on the identification and permanent correction of errors as well as the inclusion of more and more phenomena, systematic upscaling allows to approximate completeness of function and of data coverage

Direct access to the cognition of a talking robot via the service channel (Hausser 2006, 2.4.3) allows to view the reasoning of the artificial agent objectively, providing epistemology in philosophy and behavior modeling in psychology with the possibility of computational testing as a modern methodology. Regular software releases containing the improvements of the last upscaling cycle may be used by paying subscribers to satisfy their natural language processing needs, resulting in solid long-term funding.

### **Acknowledgment**

Thanks to an anonymous reviewer for helpful remarks.

## **Bibliography**

- Aristotle (1974). *The Complete Works of Aristotle*, edited by Jonathan Barnes, 2 vols., Princeton University Press
- Biederman, I. (1987). "Recognition-by-components: a theory of human image understanding," *Psychological Review* Vol. 94:115–147
- de la Briandais, R. (1959). "File searching using variable length keys," Proc. Western Joint Computer Conf. 15, 295–298
- Burgess, A. (2012). "Metalinguistic descriptivism for Millians," *Australasian Journal of Philosophy*, Volume 91:443–457, 2013 - Issue 3
- Chen, F. (2006). Human Factors in Speech Interface Design. *Designing Human Interface in Speech Technology*, 167-224
- Chisvin, L., and R.J. Duckworth (1992). "Content-addressable and associative memory," p. 159–235, in M.C. Yovits (ed.) *Advances in Computer Science*, 2nd ed., Academic Press
- Cormen, T.H., C.E. Leiserson, R.L. Rivest, and C. Stein (2009). *Introduction to Algorithms*, 3rd ed., Cambridge, MA: MIT Press
- Cresswell, M. (1972). "The world is everything that is the case," *Australasian Journal of Philosophy*, Vol. 50:1-13

- Cumming, S. (2008). “Variabilism,” *Philosophical Review*, 117(4): 525–554
- Dutilh Novaes, C. (2003). “A medieval reformulation of the de dicto/de re distinction” *LOGICA Yearbook 2003*, Prague, Filosofia, 2004, pp. 111–124
- Elmasri, R., and S.B. Navathe (2010). *Fundamentals of Database Systems, 6th ed.*, Redwood City, CA: Benjamin-Cummings
- Fredkin, E. (1960). “Trie memory,” *Comm. ACM*, Vol. 3.9:490–499
- Frege, G. (1892). “Über Sinn und Bedeutung,” *Zeitschrift für Philosophie und philosophische Kritik*, Vol 100: 25–50
- Handl, J. (2008). *Entwurf und Implementierung einer abstrakten Maschine für die oberflächenkompositionale inkrementelle Analyse natürlicher Sprache*, Diplom thesis, Department of Computer Science, U. of Erlangen-Nürnberg
- Handl, J. (2012). *Inkrementelle Oberflächenkompositionale Analyse und Generierung Natürlicher Sprache*, Inaugural Dissertation, CLUE, Univ. Erlangen-Nürnberg (available at <http://opus4.kobv.de/opus4-fau/frontdoor/index/index/docId/3933>)
- Handl, J., and C. Weber (2010). “A multilayered declarative approach to cope with morphotactics and allomorphy in derivational morphology”, in *Proceedings of the 7th Conference on Language Resources and Evaluation (LREC 2010)*, p. 2253–2256, Valletta, Malta
- Hausser, R. (1989). *Computation of Language, An Essay on Syntax, Semantics, and Pragmatics in Natural Man-Machine Communication*, Symbolic Computation: Artificial Intelligence, pp. 425. Soft cover reprint 2013. Springer
- Hausser, R. (1992). “Complexity in Left-Associative Grammar,” Elsevier: *Theoretical Computer Science*, Vol. 106.2:283-308
- Hausser, R. (1999). *Foundations of Computational Linguistics, Human-Computer Communication in Natural Language, 3rd ed. 2013*, Springer
- Hausser, R. (2001). “Database Semantics for natural language,” Elsevier: *Artificial Intelligence*, Vol. 130.1:27–74
- Hausser, Roland (2005) "Memory-based pattern completion in Database Semantics", *Language and Information*, Vol. 9.1:69–92
- Hausser, R. (2006). *A Computational Model of Natural Language Communication: Interpretation, Inference, and Production in Database Semantics*, Springer
- Hausser, R. (2011). *Computational Linguistics and Talking Robots; Processing Content in Database Semantics*, Springer
- Kass, R., and T. Finin (1988). “Modeling the user in natural language systems,” *Computational Linguistics*, Vol. 14:5–22
- King, J. (2014). “Speaker intentions in context,” *Noûs*, Vol. 48.2: 219–237

- Kripke, S. (1972). "Naming and necessity," in D. Davidson and G. Harman (eds), *Semantics of Natural Language*, p. 253–355, Dordrecht: D. Reidel
- McKeown, K. (1985). "Discourse strategies for generating natural-language text," *Artificial Intelligence*, Vol. 27: 1–41.
- Modrak, D. (2001) *Aristotle's Theory of Language and Meaning*, New York: Cambridge Univ. Press
- Montague, R. (1973). "The proper treatment of quantification in ordinary English," in Jaakko Hintikka, Julius Moravcsik, Patrick Suppes (eds) *Approaches to Natural Language*, Dordrecht, Reidel: 221–242
- Peirce, C.S. *Collected Papers*. C. Hartshorne and P. Weiss (eds), Cambridge, MA: Harvard Univ. Press 1931–1935
- Pelczar, M., and Rainsbury, J., 1998. "The Indexical Character of Names," *Synthese*, Vol. 114: 293–317.
- Piaget, J. (1954). *The Construction of Reality in the Child*, New York: Basic Books
- Quine, W.v.O. (1960). *Word and Object*. Cambridge, MA: MIT Press
- Recanati, F. (1997). *Direct Reference: From Language to Thought*, Oxford: Blackwell
- Recanati, F. (2004). *Literal Meaning*, Cambridge: Cambridge University Press
- Reimer, M., and E. Michaelson (2014). "Reference", *The Stanford Encyclopedia of Philosophy* (Winter 2014 Edition), Edward N. Zalta (ed.),  
URL = <<http://plato.stanford.edu/archives/win2014/entries/reference/>>
- Russell, B. (1905) "On denoting," *Mind*, Vol. 14:479–493
- Steels, L. (1999). *The Talking Heads Experiment*. Antwerp: limited pre-edition for the Laboratorium exhibition
- Weiss, M.A. (2005). *Data Structures and Problem Solving Using Java. 3rd ed.*, Upper Saddle River, NJ: Pearson Addison-Wesley
- Wiener, N. (1961). *Cybernetics, or control and communication in animal and machine, 2nd ed.*, Cambridge, MA: MIT press