# From Word Form Surfaces to Communication

Roland Hausser

Universität Erlangen-Nürnberg
Abteilung Computerlinguistik (CLUE)
rrh@linguistik.uni-erlangen.de

**Abstract**

The starting point of this paper is the external surface of a word form, for example the agent-external acoustic perturbations constituting a language sign in speech or the dots on paper in the case of written language. The external surfaces are modality-dependent tokens which the hearer recognizes by means of (i) pattern-matching and (ii) a mapping into modality-independent types, and which the speaker produces by an inverse mapping from modality-independent types into tokens synthesized in a modality of choice.

The types are provided by a lexicon stored in the agent's memory. They include not only the necessary[1] properties of the surface shape, but also the associated morphosyntactic properties and the meaning. The question addressed by this paper is how to design the lexical analysis of word form types as a *data structure* (abstract data type), suitable for the purpose of Database Semantics (DBS), i.e., for a computational model of natural language communication.[2]

After discussing the conditions of automatic word form recognition and production in a talking robot, we turn to the question of what format the analyzed word forms should have. The requirements are an easy coding of lexical details, a simple detection and representation of semantic relations, suitability for storage and retrieval in a database, support of a computationally straightforward matching procedure for relating the levels of language and context, and compatibility with a suitable algorithm.

## 1 Structure of Words

The nature and function of unanalyzed word form surfaces may be illustrated by a foreign language situation. For example, if our home town is in an English-speaking country we can go to a restaurant and successfully order a glass of water by saying to the waiter Please bring me a glass of water. If we travel to France, however, we will not be understood unless we use French or the waiter speaks English.[3]

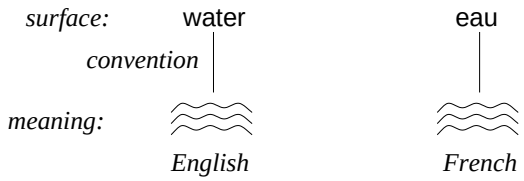
This difference between speaking in our language at home and abroad is caused by the composite structure of the words in natural language. Essential components are the *surface* and the *meaning* of a word plus a *convention* connecting the two. As an example, consider the following analyses of the English word water and its French counterpart eau:

[1] Necessary as opposed to accidental (kata sumbebêkos), as used in the philosophical tradition of Aristotle.

[2] Database Semantics describes the procedural aspects of the SLIM theory of language (FoCL'99). As an acronym, SLIM stands for the principles of **S**urface compositional, time **L**inear, **I**nternal **M**atching. As a word, SLIM stands for low (linear) mathematical complexity.

[3] "Whereas the individuals of all nonhuman species can communicate effectively with all their conspecifics, human beings can communicate effectively only with other persons who have grown up in their same linguistic community – typically, in the same geographical region." Tomasello (2003), p. 1.

## 1.1 BASIC STRUCTURE OF A WORD




The meaning is represented here in a preliminary form as three wavy lines suggesting water. While the two words happen to have the same meaning, they have different surfaces, namely water and eau. Each surface is connected to its meaning by a convention which every speaker of English or French has to learn.[4]

The surfaces of natural languages occur in different modalities, mainly spoken vs. written language.[5] In the modality of spoken language, the surfaces are sounds which are recognized by the agent's ears and produced by the agent's mouth. In the modality of written language, the surfaces are letter sequences which are recognized by the agent's eyes and produced by the agent's hands.

When the words of a first language are acquired, a child must learn (i) the meanings, (ii) the (acoustic) surfaces, and (iii) the conventions which connect the surfaces to the correct meanings. This process is embedded into child development in general, takes several years, and normally doesn't cause any special difficulties. However, when learning the words of a foreign language as an adult, the following difficulties stand out:[6]

## 1.2 TASKS OF LEARNING THE WORDS OF A FOREIGN LANGUAGE

- learning to recognize and produce the foreign surfaces in the modalities of spoken and written language, and
- learning the conventional connections between these foreign surfaces and familiar meanings.

Learning to recognize the acoustic surfaces of a foreign language is difficult and to pronounce them without accent is often nearly impossible, while learning to read and to write may come easier. There are languages like Japanese, however, for which a foreigner is considered more likely to learn to speak fairly fluently than to acquire a near-native ability to read and write.

Connecting the foreign surfaces to familiar meanings presupposes that the notions of the foreign language are identical or at least similar to those of one's own. This holds easily for basic notions like *father, mother,*[7] *child, son, daughter*, etc., as well as *sun, moon, water, fire, stone, meat, fish, bird, tree*, etc. When it comes to more culturally dependent notions, however, what is represented by a single word in one language may have to be paraphrased by complex constructions in the other.

[4]The Swiss linguist Ferdinand de Saussure (1857–1913) described the convention-based and therefore "unmotivated" relation between language-dependent surfaces like water or eau and their meaning in his *Premier Principe: l'arbitraire du signe*. Cf. Saussure 1916/1972.

[5]In addition there is the third modality of signed language for the hearing impaired and, as a form of written language, there is Braille for the blind.

[6]Once a certain number of words have been learned, there are other difficulties, such as the use of idioms.

[7]Kemmer 2003, p. 93, claims that "in some languages" the word meaning of mother would include maternal aunts. It seems doubtful, however, that the distinction between a mother's own children and those of her sisters could ever be lost. A more plausible explanation is a non-literal use. In German, for example, a child may call any female friend of the parents Tante (aunt). If needed, this use may be specified more precisely as Nenn-Tante (aunt by name or by courtesy). Similarly in Korea, where the term "older brother" may be bestowed on any older respected male.

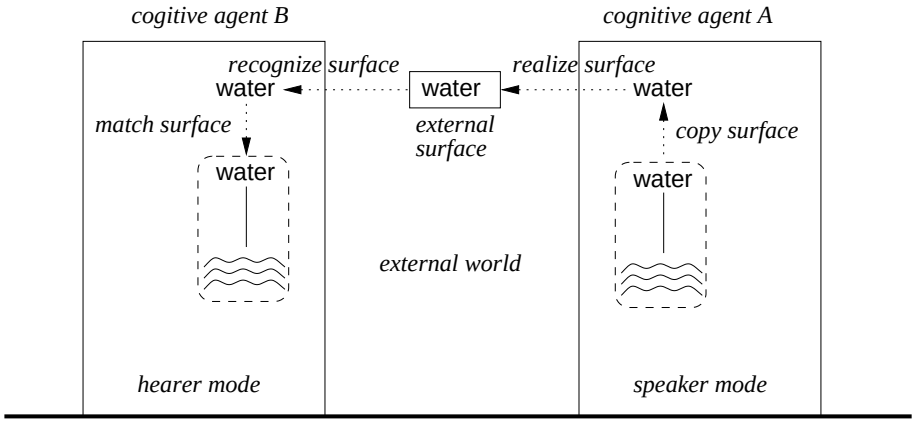
A popular example is the Eskimo language Inuit, said to have something like fifty different words for snow. In English translation, these would have to be paraphrased, for example, as soft snow, hard snow, fresh snow, old snow, white snow, grey snow, snow pissed on by a baby seal, etc.[8] An important precondition for adequate paraphrasing is a proper knowledge of the foreign notion to be described; as an example consider the first Eskimo faced with the task of introducing the notion of an electric coffee grinder into Inuit.

# 2 Modality-Dependent External Surfaces

There are essentially two basic perspectives from which communicating with natural language may be analyzed scientifically, namely *internal* and *external*.[9] The internal perspective is based on one's introspection as a native speaker, and has been illustrated in the previous section by an example of communication failure in a foreign language environment. The external perspective is that of a scientist working to reconstruct the functioning of natural language communication as an abstract theory (cf. NLC'06, Sect 1.4).

As an example of the external perspective consider two communicating agents A and B, A is in the speaker mode and produces the word form surface water, while B is in the hearer mode and recognizes this surface. Viewed from the external perspective, these two communicating agents may be shown as follows:

## 2.1 PRODUCTION AND RECOGNITION OF A WORD




The two agents are concrete individuals, e.g., humans, with bodies[10] existing in the external world (AIJ'01). Between the agents A and B, abstractly represented as boxes, there exists the external surface water, represented as the small box in the middle containing the letters of the word.

As an object in the real world, the external surface has neither a meaning attached to it nor any grammatical properties. It is simply an external object with a particular shape which may be measured and described with the methods of the natural sciences, either acoustically or optically. This shape is arbitrary in the sense that it does not matter whether the agents use

[8] See Pullum 1991, "The great Eskimo vocabulary hoax."

[9] A third approach, the sign-oriented perspective, analyzes language signs in isolation, abstracting away the communicating agents. Though popular in the current main streams of linguistics and of philosophy of language, its simplifying assumptions render the foundations too narrow for a computational model of communication.

[10] The importance of agents with a real body (instead of virtual agents) has been emphasized by Emergentism (MacWhinney 1999, 2008).

the surface water or the surface eau as long as they obey the conventions of their common natural language.

In their minds, both agents contain the word water as analyzed in 1.1, i.e., the surface, the meaning, and the convention-based connection between the surface and the meaning. As purely cognitive representations (e.g., binary code), the analyzed words are independent of any modality,[11] but include the meaning and all the grammatical properties.

The speaker produces the external surface by copying the internal word surface and realizing it externally in the modality of choice. The hearer classifies the external surface by matching it with a learned surface pattern and by matching the pattern with the internal surface of the corresponding word (lexical lookup).[12]

Even though the internal representations of surfaces and meanings are modality-free in principle, they may indirectly relate to modalities insofar as they have been derived from a modality-dependent representation during recognition or are to be realized in a modality-dependent representation during production. In this indirect sense, the internal surfaces are usually mono-modal[13] while the meanings are multi-modal. For example, the surface of the word raspberry is mono-modal in that the associated modality is purely auditive, visual, or tactile (Braille), but does not include taste. The meaning of this word, in contrast, may be viewed as relating to a multimodal conglomerate including taste.

This fundamental basis for the functioning of natural language is formulated as the First Mechanism of Communication (further Mechanisms are discussed in unpublished work):

## 2.2 THE FIRST MECHANISM OF COMMUNICATION (MOC-1)

> Natural language communication relies on modality-dependent external surfaces which have neither meaning nor any grammatical property.

MoC-1 is essential for the construction of talking robots, and differs from the assumptions of Realism (e.g., Barwise and Perry 1983) and logical semantics (e.g., Montague 1974).

The cognitive counterpart to MoC-1 is that the agent-internal representations of surfaces are modality-independent, for example, binary code. Thus, during communication there is a constant mapping of modality-independent internal representations into modality-dependent surfaces (speaker mode), and of modality-dependent surfaces into modality-independent internal representations (hearer mode).

MoC-1 is a functional complement to Surface Compositionality.[14] According to this methodological principle, the grammatical analysis of language signs may use only the concrete word forms as the building blocks of composition, such that all syntactic and semantic properties of a complex expression derive systematically from the syntactic category and the literal meaning of the lexical items – which are of a cognitive, modality-free nature.

It follows from MoC-1 that a functional reconstruction of communication with natural language cannot be limited to a grammatical analysis of the language signs, but must include a functional model of natural language communication which is defined as follows:

[11] To illustrate the agent-internal representation of the word water in 2.1, we had to resort to a suitable modality, namely vision. Though unavoidable, this is paradoxical insofar as the words in the cognition of an agent are inherently modality-free.

[12] One may recognize a surface as a real word, yet not know its meaning, for example "bedizen", "effloresce", "exigency", "minatory", or "reprobate" in English. In this case, there is an entry missing in the hearer's lexicon. This may hold even more when recognizing a surface-like shape as a word form of a foreign language.

[13] A multi-modal representation of language may be found in an opera performance in which the text sung on stage (auditive modality) is also shown more or less simultaneously in writing above the stage (visual modality). A more mundane example is watching a DVD of an English movie with English subtitles.

[14] Cf. SCG'84; FoCL'99, p. 80, 111, 256, 327, 418, 501, 502; NLC'06, p. 17–19, 29, 89, 211.

## 2.3 FUNCTIONAL MODEL OF NATURAL LANGUAGE COMMUNICATION

A functional model of natural language communication requires

1. a set of cognitive agents each with (i) a body, (ii) external interfaces for recognition and action, and (iii) a memory for the storage and processing of content,
2. a set of external language surfaces which can be recognized and produced by these agents by means of their external interfaces using pattern matching,
3. a set of agent-internal (cognitive) surface-meaning pairs stored in memory, whereby the internal surfaces correspond to the external ones, and
4. an agent-internal algorithm which constructs complex meanings from elementary ones by establishing semantic relations between them.

The importance of the agents in a language community is shown by "lost languages." Imagine the discovery of clay tablets left behind by an unknown people perished long ago. Even though the external surfaces of their language are still present in form of the glyphs on the tablets, the language as a means of communicating meaningful content is lost. The only way to revive the language at least in part is to reconstruct the *knowledge* of the original speakers – which was part of their cognition.[15]

The requirements of 2.3 are minimal, and by no means constitute a full picture of the mechanism of natural language communication. They are sufficient, however, to distinguish natural language communication from other forms of communication:

## 2.4 COMMUNICATION WITHOUT A NATURAL LANGUAGE

- endocrinic messaging by means of hormones,
- exocrinic messaging by means of pheromones, for example in ants,[16] and
- the use of samples, for example in bees communicating a source of pollen.

These kinds of communication differ from natural language because they lack a set of external surfaces with corresponding internal surface-meaning pairs established by convention.[17]

From a functional point of view, the mechanism described by MoC-1 has the following advantages for communication:

## 2.5 ADVANTAGES FOLLOWING FROM MOC-1

1. The modality of an external surface imposes no restriction on the kind of meaning which may be attached to the internal counterpart of this surface.
2. The external surfaces are much more suitable for (i) transfer and (ii) long-term storage than the associated meanings.

[15] Thereby, knowledge of the word form meanings and the algorithm for their composition alone is not sufficient for a complete, successful interpretation. What is needed in addition is knowledge of the correct context of interpretation. Cf. FoCL'99, Sect. 5.3.

[16] E.O. Wilson 1998 describes the body of an ant worker as a walking battery of exocrinic glands.

[17] Another example of not constituting a language are the macros used in programming languages: defined ad hoc by the programmer as names for pieces of code, they are (program-)internal abbreviations. As such they do not require any external surfaces agreed on by convention and are not used for inter-agent communication.

The advantage of modality independence referred to in (1) may be illustrated by a sign-theoretic comparison of symbols and icons (cf. FoCL'99, Sect. 6.4). Icons in the visual modality, for example, are limited to a visual representation of meaning, which is often difficult if not impossible for representing concepts from another modality, for example, the meaning of *sweet*. Symbols, in contrast, have no such limitation because their meaning representations are modality-free.[18]

The reason for (2) in 2.5 is that an unanalyzed external surface is extremely simple and robust compared to the often complicated meanings attached to their analyzed surfaces inside the cognitive agent. This holds especially for the written representation of language, which has long been the medium of choice for the long term agent-external storage of content.[19] Recently, however, tape recordings and video have made it possible to store also spoken and signed language for indefinite periods of time.

Naturally, the highly specialized and powerful technique characterized by MoC-1 also has an apparent disadvantage: because the surface for a given meaning can take any shape within the limits of its modality, communities of natural agents can and do evolve their own languages. This results in the difficulty of communicating in foreign language environments, which this paper began with. It is familiar to anyone who has spent some time abroad, and for many it constitutes a disincentive to leave home.

# 3 Modality Transfer in the Speaker and the Hearer Mode

MoC-1 is not only a conceptual insight into the working of natural language communication between agents viewed from the outside, but constitutes a well-defined technical challenge, i.e., the construction of machines which can recognize and produce external language surfaces. The basic task of these machines may be viewed as an automatic transfer between a modality-dependent realization of an external surface and its modality-free counterpart represented as agent-internal digital code, e.g., 7 bit ASCII. This transfer is instantiated in four basic variants, namely the speaker mode and the hearer mode, each in the modalities of vision (writing) and audition (speech).

In the hearer mode, today's systems of *speech recognition* transfer external acoustic surfaces into modality-free digital code mainly by the statistical method of Hidden Markov Models (HMM).[20] In the modality of vision, today's systems turn images of letters into modality-free digital code based on the software of *optical character recognition* (OCR).

In the speaker mode, today's systems of *speech synthesis* transfer digitally represented text into artificially realized speech, usually by concatenating pieces of recorded speech. In the modality of vision, there is the transfer from digital code to the familiar letter images on our computer screens, which may be called *optical character synthesis* (OCS).

The transfer from a modality-dependent to a modality-free representation (hearer mode) abstracts away from properties of the external surfaces such as speed, pitch, intonation, etc. in spoken language, and font, size, color, etc. in written language, which from a certain point of view may be regarded as accidental. Conversely, the transfer from a modality-free to a

[18] A modality-free representation of meaning is procedural in the sense that it is based on the recognition and action procedures of the cognitive agent. Whether such a meaning representation in an artificial agent is adequate or not is decided by agent's behavior. For example, an artificial agent's concept of *shoe* may be considered adequate if the agent picks out the same objects as a human would from a collection of different things (NLC'06, Chapt. 4).

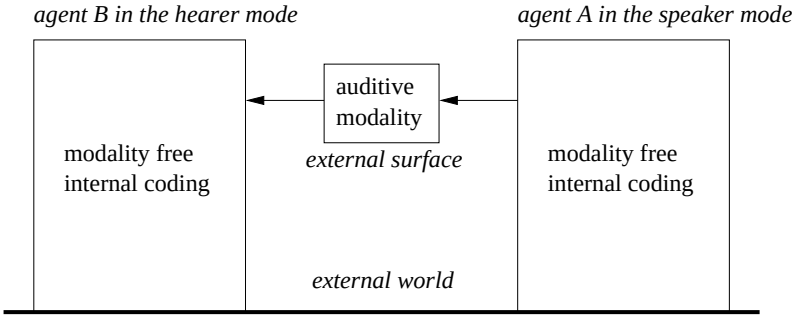
[19] On the relation between modalities and media see FoCL'99, p. 23 f., and NLC'06, p. 23 f.

[20] A historically earlier approach is called Dynamic Time Warping (DTW).

modality-dependent representation (speaker mode) must decide on which of these properties, for example, pitch, speed, dialect, etc., should be selected for the external surface.

If speaker mode and hearer mode utilize the *same* modality and are realized by *different* agents, we have interagent communication. Examples are agent A writing a letter (speaker mode, visual modality) and agent B reading the letter (hearer mode, visual modality), and accordingly in the auditive modality:

## 3.1 INTERAGENT COMMUNICATION USING SPEECH



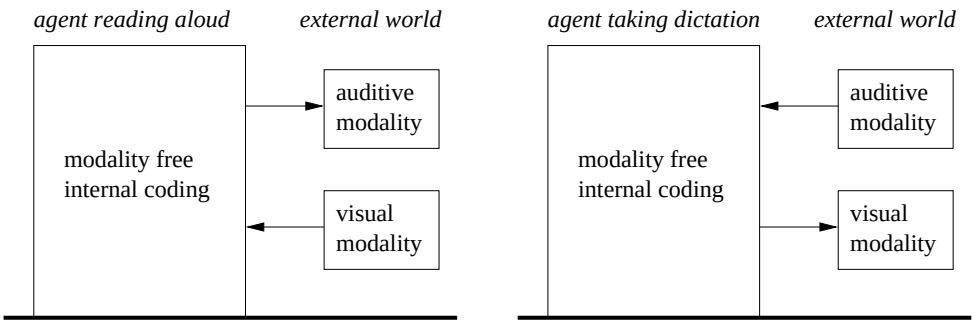

A notable difference between the auditive and the visual modality in interagent communication is that the interpretation(s) (reader, hearer mode) of a letter, for example, may be far removed in time and space from the point of production (writer, speaker mode). Another difference is that written language can be corrected, while speech usually cannot:

> Speech is irreversible. That is its fatality. What has been said cannot be unsaid, except by adding to it: to correct here is, oddly enough, to continue.
>
> R. Barthes, 1986, p. 76

Of course, as soon as spoken language is recorded, interpretation may be arbitrarily distant from production in time and space. Also, the recording may be "doctored" – which from a certain point of view is a form of correction.

If the speaker mode and the hearer mode utilize *different* modalities and are realized by the *same* agent, we have a modality conversion. In nature, modality conversion is illustrated by reading aloud (conversion from the visual to the auditive modality) and taking dictation (conversion from the auditive to the visual modality).

## 3.2 TWO KINDS OF MODALITY CONVERSION




In technology, reading aloud is modeled by combining optical character recognition and speech synthesis, which is an important application for the blind. Conversely, taking dic-

tation is modeled by a machine called “electronic secretary” which combines speech recognition with optical character synthesis.[21]

# 4 General Purpose Interfaces and the Language Channels

In recognition, language interpretation is embedded into non-language recognition. For example, in the visual modality humans can recognize non-language[22] input such as streets, houses, cars, trees, and other agents as well as language input such as shop signs and texts in newspapers, books, and letters. Similarly in the auditive modality: humans can recognize non-language input such as the sounds made by rain, wind, bird songs, or approaching cars as well as language input such as speech from other agents, the radio, or television.

In action, language production is likewise embedded into non-language action. For example, humans can use the hands for non-language output such as filling the dishwasher or drawing a picture as well as language production in the visual modality, such as handwriting a letter or typing something on the computer keyboard. Similarly in the auditive modality: humans can use their voice for non-language output such as singing without words or making other noises as well as language production, i.e., speech.

Such embedding of language recognition and synthesis into non-language recognition and action cannot be handled by today’s technologies of speech recognition and optical character recognition. Instead, they use input and output channels dedicated to language. This constitutes a substantial simplification of their respective tasks (smart solution).[23]

Despite this simplification and despite well-funded research efforts with many promises and predictions over several decades, automatic speech recognition in particular has not yet reached its goal.[24] As proof, we don’t have to embark here upon a more detailed argument, but simply point to the ever increasing number of keyboards in everyday life: if automatic speech recognition worked in any practical way, i.e., comparable to the speech recognition capability of an average human,[25] few users would prefer keyboard and screen over speech.

For building a talking robot, like R2D2 in the “Star Wars” movies, any time soon, the current attempts at automatic speech recognition present a – hopefully temporary – obstacle. Fortunately, however, it does not prevent continuing work on the computational reconstruction of natural language communication in Database Semantics (DBS), in the hearer as well as the speaker mode. The reason is an important difference between natural and artificial cognitive agents regarding what we call the *auto-channel* and the *service channel*.[26]

In a natural agent, language and non-language input and output are provided by the auto-channel. It evolves naturally during child development and comprises everything a natural

[21] Infra-red cameras are another technical means of a modality transfer, representing temperature (what is called the temperature modality) as color (visual modality).

[22] We prefer the term non-language over non-verbal because the latter leads to confusion with verb as a part of speech.

[23] For the distinction between smart and solid solutions see FoCL’99, Sect. 2.3

[24] Optical character recognition, in contrast, is quite successful at least for printed language, and widely used for making paper documents available online.

[25] A practical system of speech recognition must fulfill the following requirements simultaneously (cf. Zue et al. 1995):

1. speaker independence
2. continuous speech
3. domain independence
4. realistic vocabulary
5. robustness

[26] Cf. NLC’06, p. 14 f.

agent can see, hear, feel, taste, etc. as well as consciously do. The auto-channel also includes natural speech recognition and production.

In a standard computer, e.g., a notebook or a desktop computer, in contrast, there is no auto-channel. Instead, there are the keyboard and the screen, which allow users and scientists alike to access the hard- and software of the computer directly.

Building an artificial cognitive agent consists in large part[27] in the reconstruction of an artificial auto-channel, including the external interfaces of vision and audio for recognition and of locomotion and manipulation for action. This reconstruction is an incremental process which relies heavily on the keyboard and the screen functioning as the service channel for direct access to and manipulation of the hard- and software of the robot under development.

After completion, the artificial agent will be able to interact autonomously with the external world via its auto-channel, including communication with the user. However, in contrast to a natural agent, an artificial cognitive agent will not only have an auto-channel, but also a service channel as a remnant of the process of its construction.

The essential role of the service channel in bootstrapping the reconstruction of cognition is especially clear in the area of natural language communication. This is because it does not really matter whether a language surface gets into the computer via the auto- or the service channel. All that matters for modeling natural language understanding in computational linguistics is that the word surfaces get into the computer at all, and for this typing them at the keyboard of today's standard computers is sufficient.

Similarly for language production in the speaker mode: all that matters for the user's communication with the computer is that the surfaces derived from prior software processing are realized externally at all, and for this optical character synthesis on the screen of today's standard computers is sufficient – at least for users who can see and have learned to read.

The prospect of reconstructing the mechanism of natural language communication via the agent's service channel is good news for developing a capable automatic speech recognition[28] as part of the artificial autochannel. The reason is that the persistent stagnation in this particular application is caused by a search space too large for the statistical approaches as used today. The gigantic size of the search space results from the many possible word forms in a natural language combined with even more numerous possibilities of syntactic combinations and of variations of pronunciation between different speakers.

The best way to reduce this search space are hypotheses on possible continuations computed by a time-linear grammar algorithm as well as expectations based on prior experiences at the level of language understanding. After all, this is also the method used by humans for disambiguating speech in noisy environments. The precondition for using this method for artificial speech recognition is a theory of how communicating with natural language works.

# 5 Automatic Word Form Recognition

Assuming that the language input to an artificial cognitive agent are word form surfaces provided by the service channel as sequences of letters, the first step of any rule-based (i.e., non-statistical) reconstruction of natural language understanding is building a system of automatic word form recognition.[29] This is necessary because for the computer a word form

[27] The remaining task is the reconstruction of an autonomous control.

[28] Contextual cognition, such as non-language vision and audio, may also benefit from the service channel. By building a context component with a data structure, an algorithm, and a database schema via direct access, autonomous recognition and action may be provided with structures to map into and out of.

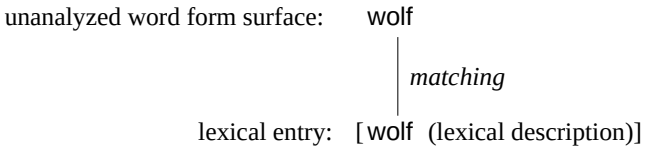
[29] We begin with the hearer mode as a means to get content into the computational reconstruction of central cognition. The availability of such content is a precondition for implementing the speaker mode.

surface like learns is merely a sequence of letters coded in seven bit ASCII, no different from snrael, for example.

Automatic word form recognition takes an unanalyzed surface, e.g., a letter sequence supplied by the service channel, as input and provides the computer's syntactic-semantic processing with the information needed for interpreting any well-formed combination of surfaces in a sentence. Its two basic tasks are (i) *categorization* and (ii) *lemmatization*. Categorization specifies the grammatical properties, which in the case of learns would be something like "verb, third person singular, present tense." Lemmatization specifies the base form, here learn, which is used to look up the core meaning common to all the word forms of the paradigm, i.e., learn, learns, learned, and learning.[30]

The recognition algorithm in its most primitive form consists in matching the surface of the unknown letter sequence with the corresponding surface in an on-line lexicon, thus providing access to the relevant lexical description.

### 5.1 MATCHING AN UNANALYZED SURFACE ONTO A KEY




There exist several techniques for matching a given surface automatically with the proper entry in an electronic lexicon.[31] For the computational linguist, the main work of building a system of automatic word form recognition for a given language is to efficiently provide the correct categorization and lemmatization for the different forms of the words, especially if there are irregularities, as in swim, swims, swam, swum, and swimming. Also, the system must be able to handle *neologisms* like insurrectionist (inmate) or migraineur.

Building such a system of automatic word form recognition for any given natural language is not particularly difficult, at least if alphabetic or syllabic writing systems are used, such as Pinyin for Chinese. Given (i) an on-line dictionary of the natural language of choice, (ii) a suitable off-the-shelf software framework, and (iii) a properly trained computational linguist, an initial system can be completed in less than six month.[32] It will provide accurate, highly detailed analyses of about 90% of the word form types in a corpus.

Increasing the recognition rate to approximate a 100% is merely a matter of additional work.[33] It consists in adding missing entries in the on-line lexicon, and improving the rules for allomorphy (handling irregular forms) and for composition (handling inflection/agglutination, derivation, and composition). To maintain a recognition rate of practically 100% over longer periods of time, the system must be serviced continually, based on a large reference corpus and regular (e.g., annual) monitor corpora.

## 6 Format of Analyzed Word Forms

Apart from the *method* of automatic word form recognition, there is the question of the *format* in which the analysis of a word form should be presented for subsequent processing. In DBS,

[30] For further information on the morphological analysis of word forms and different methods of automatic word form recognition see FoCL'99, Chapts. 13–15.

[31] See A.V. Aho & J.D. Ullman 1977, p. 336–341.

[32] This is the standard period of time for writing an MA thesis at the University of Erlangen-Nürnberg.

[33] This is in contrast to the statistical method, which is not suitable for the correction of specific errors. See FoCL'99, Sect. 15.5.

lexical word form analyses are coded as *proplets*. A proplet is a flat (non-recursive) feature structure, defined as a set of attribute–value pairs (avp).

Proplets have (i) lexical attributes, (ii) continuation attributes, and (iii) bookkeeping attributes.[34] Lexical attributes describe the sign's surface, meaning, and morphosyntactic properties. Continuation attributes establish grammatical relations, i.e., functor-argument structure and coordination, to other proplets. Bookkeeping attributes are for numbering items automatically for purposes of indexing, storage, and retrieval in a database. For better readability and improved computational efficiency the attributes are displayed in a predefined standard order. The values of proplet attributes are restricted to lists which consist of one or more atomic items.

As examples, consider the following two proplets representing corresponding words in English and French:

## 6.1 THE LEXICAL NOUN PROPLETS water AND eau

| | *English* | | *French* | | |
|---|---|---|---|---|---|
| *surface* | sur: | water | sur: | eau | |
| *core attribute* | noun: | *water* | noun: | *water* | *lexical features* |
| *category* | cat: | sn | cat: | sn | |
| *semantic property* | sem: | mass | sem: | mass | |
| *modifier(s)* | mdr: | | mdr: | | |
| *functor* | fnc: | | fnc: | | *continuation features* |
| *next conjunct* | nc: | | nc: | | |
| *previous conjunct* | pc: | | pc: | | |
| *identity* | idy: | | idy: | | *bookkeeping features* |
| *proposition number* | prn | | prn | | |

Proplets resulting from automatic word form recognition are called *isolated* proplets because only their lexical attributes have values.

The two proplets illustrate the similarity between corresponding words in English and French: the lexical features differ solely in the values water and eau of the sur attribute. These values are needed for the matching with the unanalyzed surface, as illustrated in 5.1.

Of the remaining lexical attributes, the core attribute, noun, indicates the basic grammatical use of the word form; its value, *water*, represents the meaning associated with the respective surface, and is the same for the two word forms in question. The third and fourth attributes, cat for category, and sem for semantics, specify the combinatorial and non-combinatorial morphosyntactic properties, respectively, of the word form (categorization).

The continuation attributes receive their values during syntactic-semantic composition, based on copying values between proplets. This procedure turns *isolated* (lexical) proplets into *connected* proplets, which represent the content of sentences and texts. The first two continuation attributes in 6.1 code the *functor-argument structure* of a sentence, whereby the value(s) of mdr are optional, while that of fnc is obligatory. The second two, nc and pc, are for coding an optional *coordination* of nouns in a sentence.

The values of the bookkeeping attributes are assigned by the parser. The idy attribute is incremented automatically for each new noun proplet, stipulating their non-identity. Coreference is treated as an inference which infers equivalence between certain idy values.[35] The prn attribute is common to all proplets belonging to the same proposition, and incremented automatically as soon as a new proposition is started.

[34] The idea of a proplet was introduced in Hausser 1996. The term is coined in analogy to droplet. Proplets are so-called because they are the elementary items constituting a proposition.

[35] For a detailed analysis of intra- and extrapropositional coreference see NLC'06, Chapter 10.

# 7 Essential Requirements on the Data Structure

Proplets as flat, ordered feature structures are probably the simplest format for coding whatever properties a linguist might want to attribute to a word form.[36] However, while a detailed lexical analysis is a first requirement on the data structure of proplets, it must also serve the overall purpose of modeling the functioning of natural language communication.

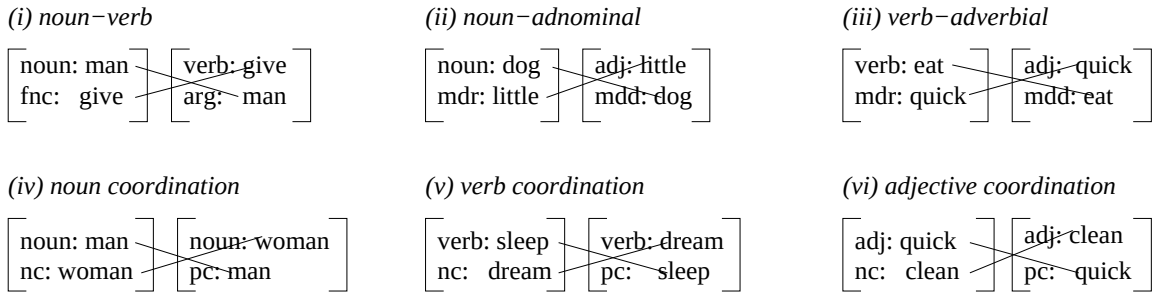
A second requirement following from this purpose is suitability for representing the semantic relations between word forms. This requirement is fulfilled by the method of bidirectional pointering between proplets, based on cross-copying:

## 7.1 SCHEMA OF BIDIRECTIONAL POINTERING


In the result, the semantic relation is characterized by the feature [attribute-2: B] in the first proplet and the feature [attribute-4: A] in the second proplet. In DBS, the bidirectional coding of a semantic relation between the two proplets is based on no more and no less than the copying of values, as indicated by the diagonal arrows.

The basic semantic relations in natural language may be illustrated schematically as follows:

## 7.2 BASIC FUNCTOR-ARGUMENT AND COORDINATION STRUCTURES




For simplicity and transparency, the representation of proplets is reduced[37] here to (i) their core feature and (ii) their relevant continuation feature. The core attributes are noun, verb, and adj (adjective), while the continuation attributes are fnc (functor), arg (argument), mdr (modifier), mdd (modified), nc (next conjunct), and pc (previous conjunct). Functor-argument structure is a semantic relation between proplets with different core attributes, while coordination structure is a semantic relation between proplets with the same core attribute.

A third requirement on the data structure of proplets is suitability for a computationally straightforward matching procedure. This procedure is needed (i) for the application of rules to their input and (ii) for the interaction between the language and the context level inside the cognitive agent. The matching between corresponding proplets at the two levels is based on compatibility of attributes and of values (cf. NLC'06, 3.2.3).

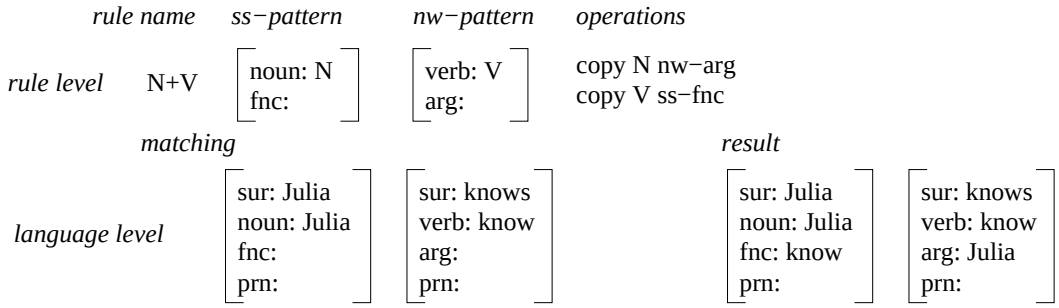
When matching a rule with language input, compatibility of values is based on the use of restricted variables, as shown in the following example (explanations in italics):

---

[36] A precursor in linguistics is lists of binary values without attributes called "feature bundle", e.g., $\begin{bmatrix}+\text{stress}\\+\text{tense}\end{bmatrix}$, used by Chomsky and Halle (1968) for purposes of morphophonology.

[37] For a more complete representation see Sect. 8 below.

## 7.3 SCHEMATIC ILLUSTRATION OF MATCHING IN A RULE APPLICATION

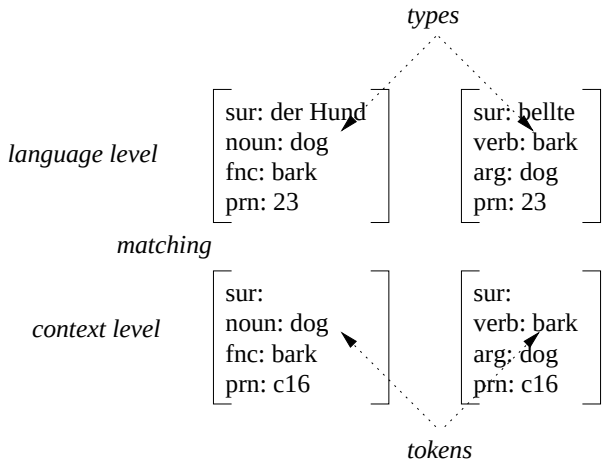



For a rule to apply, matching between the ss- and nw-patterns at the rule level and the input of proplets at the language level must be successful. The compatibility of attributes is based here on the attributes of the proplet patterns being a subset of the attributes of the corresponding language proplets. The compatibility of values is based on the value Julia being within the restriction defined for the variable N, and accordingly for know and the variable V.

During matching, the variable N is vertically bound to the value Julia and the variable V is bound to know. This enables the operations; their effect is shown in the result. The example is simplified insofar as the canceling of a valency position in the verb and the agreement conditions between subject and verb are omitted because the cat features and the associated operation are not shown (for complete detail see NLC'06, Chapt.13).

The matching between proplets at the language level and the context level uses the same attribute condition as rule matching. The value condition, however, is based on the type/token relation[38] (and not on the definition of restricted variables). Consider the following example:

## 7.4 SCHEMATIC MATCHING BETWEEN LANGUAGE AND CONTEXT LEVEL




Due to their non-recursive structure, the matching between proplet patterns and proplets and between language and context proplets is straightforward as compared to the recursive feature structures of GPSG, LFG, and HPSG; and without recursion there is no place for unification.

Just as the data structure of proplets allows to characterize word forms to any degree of lexical detail, it allows to specify the matching condition to any degree of linguistic generalization on the one hand and of any detail of restriction on the other. In the case of rule applications, these matching conditions are defined in terms of the variable definition and variable restrictions (cf. NLC'06, Sects. 11.3, 13.2) of an LA-grammar. In the case of language-context matching, the matching conditions are based primarily on the type-token relation and secondarily on inferencing (cf. NLC'06, Sect. 5.4).

[38] For a more detailed discussion see FoCL'99, Sect. 4.2, and NLC'06, Sect. 4.2.

The fourth requirement on the data structure of proplets is that they code the semantic relations between them in an *order-free* manner, so that they can be stored in a database in accordance with the needs of storage and retrieval. Storage is used in the hearer mode for sorting the proplets resulting from syntactic-semantic interpretation into the agent-internal memory, defined as a database called Word Bank. Retrieval is used in the speaker mode for selectively activating content in the Word Bank by navigating along the semantic relations between the proplets. Proplets are order-free because the semantic relations between them are coded solely in terms of attributes and their values.

# 8 Example of a Time-Linear Hearer Mode Derivation

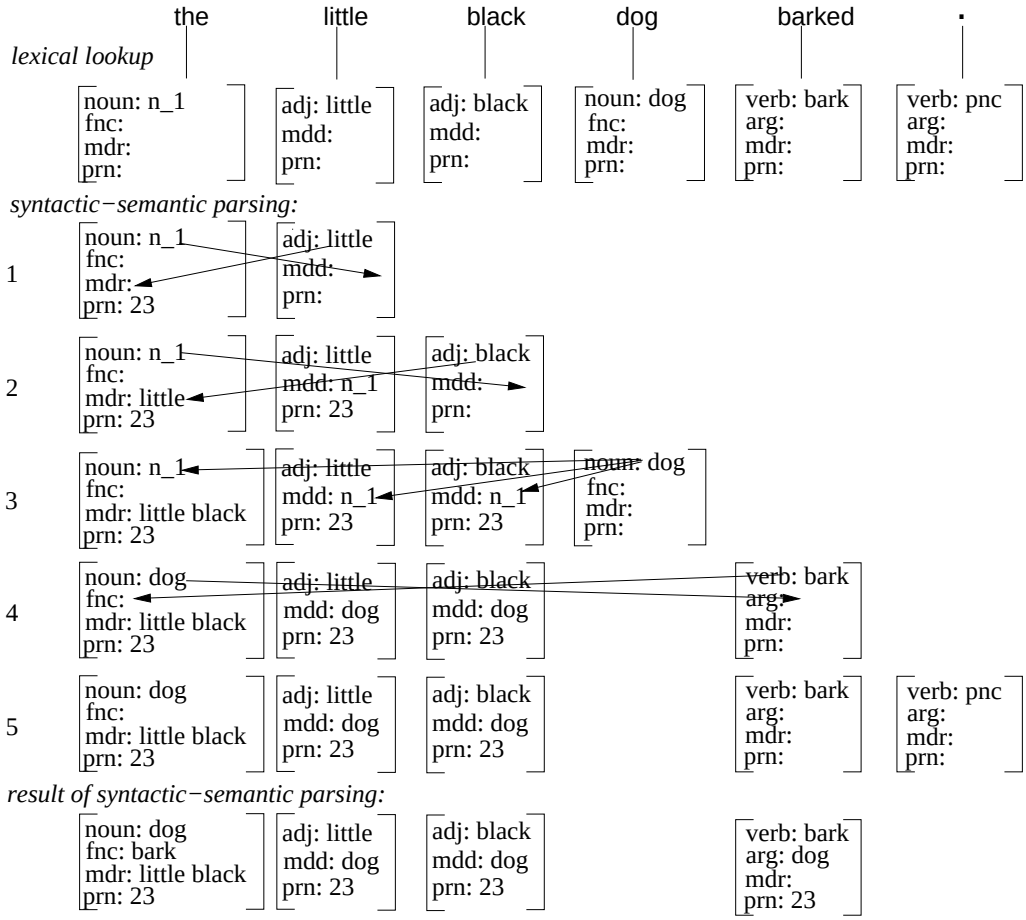
The requirements on the data structure of proplets, namely (i) coding the lexical properties of word forms, (ii) coding the semantic relations between proplets, (iii) suitability for a computationally straightforward matching procedure, and (iv) being inherently order-free, are complemented by the further requirement of (v) supporting a suitable algorithm. This is the time-linear algorithm of LA-grammar (TCS'92), which is used in the variants of LA-hear for modeling the hearer mode, LA-think for the selective activation of content in the Word Bank, and LA-speak for the realization of activated content in the speaker mode.

As an example of the interaction between the data structure and the algorithm, consider the hearer mode interpretation of the sentence

The little black dog barked (loudly).

The derivation is shown with simplified proplets:

## 8.1 TIME-LINEAR DERIVATION ESTABLISHING SEMANTIC RELATIONS



The grammatical analysis is surface compositional in that each word form of the example is analyzed as a lexical proplet (cf. lexical lookup). The derivation is time-linear, as shown by the stair-like addition of a lexical proplet in each new line. Each line represents a derivation step, based on a rule application like 7.3. The establishing of semantic relations resulting from the rule applications is indicated by diagonal arrows. There are three kinds of interaction between proplets, namely (i) cross-copying (cf. 7.2), (ii) simultaneous substitution, and (iii) absorption.

In line 1, the core values of *the* and *little* are cross-copied. The result is shown in line 2: the mdr slot of *the* now has the value little and the mdd slot of *little* has the value n_1 (i.e., a substitution value serving as the core value of *the*).

Line 2 shows an instance of cross-copying as well, this time between the core values of *the* and *black*. The result is shown in line 3: the mdr slot of *the* now has the values little black and the mdd slot of *black* has the value n_1.

Line 3 illustrates an instance of simultaneous substitution and absorption: the occurrences of the substitution value n_1 in the proplets *the, little,* and *black* are simultaneously replaced by the core value of the *dog* proplet, which is then discarded. The result is shown in line 4: the core value of what used to be the *the* proplet is now *dog*, the mdd slots of the *little* and *black* proplets have the value dog, and the lexical *dog* proplet has been deleted (absorption).

Line 4 shows another instance of cross-copying, this time between the new *dog* proplet and the *bark* proplet, as indicated by the diagonal arrows. The result is shown in line 5: the *dog* proplet now has the fnc value bark and the *bark* proplet has the arg value dog.

Line 5, finally, is an instance of absorption without prior simultaneous substitution. Though not visible due to the simplified proplets used in 8.1, the addition of the punctuation proplet triggers the following categorial operation: the [cat: v] value of the verb proplet cancels the valency position in [cat: v′ decl] of the punctuation proplet, and decl (for the sentential mood "declarative") becomes the new cat value of the verb proplet.

The result of this derivation is a representation of *content* as an order-free set of proplets, shown below with the additional adverb loudly, using the alphabetical order of the core values. As a representation of content, the language-dependent surfaces are omitted:

## **8.2** CONTENT OF The little black dog barked loudly.

| | | | | |
|---|---|---|---|---|
| sur: | sur: | sur: | sur: | sur: |
| verb: bark | adj: black | noun: dog | adj: little | adj: loud |
| cat: decl | cat: adn | cat: def sg | cat: adn | cat: adv |
| sem: past | sem: psv | sem: count | sem: psv | sem: psv |
| arg: dog | mdd: dog | fnc: bark | mdd: dog | mdd: bark |
| mdr: loud | prn: 23 | mdr: little black | prn: 23 | prn: 23 |
| prn: 23 | | prn: 23 | | |

Compared to 8.1, these proplets are shown with more detail. The adnominal adjectives *little* and *black* differ from the adverbial adjective *loud* because of their different cat values: adn for adnominal and adv for adverbial. The verb proplet *bark* has the sem value past, etc.

While the lexical analysis in 8.1 consists of six proplets, the resulting representation of content consist of only four, due to the absorption of *dog* into *the* and of the full-stop into *bark*. The semantic contributions of these function words are not lost, however. As shown in 8.2, the presence of *the* in the input is reflected by the cat value def in the noun proplet *dog* and that of the full-stop by the cat value decl in the verb proplet *bark*.[39]

[39]For corresponding LA-think and LA-speak analyses of this example see Hausser 2009.

# 9 Relating Core Attributes to Traditional Parts of Speech

The core attributes of proplets used so far in DBS are *noun, verb,* and *adj*ective.[40] These terms are among the following *parts of speech* of traditional grammar:

## 9.1 TRADITIONAL PARTS OF SPEECH IN ENGLISH

1. *verb*
   Includes finite forms like sang and non-finite forms like singing or sung of main verbs, as well as auxiliaries like was or had and modals like could and should. Some traditional grammars treat non-finite verb forms as a separate class called *participle*.

2. *noun*
   Includes common nouns like table and proper names like Julia. There is also the distinction between count nouns like book and mass nouns like wine.

3. *adjective*
   Includes determiners like a(n), the, some, all, and my as well as adnominals like little, black, and beautiful. Some traditional grammars treat determiners as a separate class.

4. *adverb*
   Includes adverbials like beautifully as well as intensifiers like very.

5. *pronoun*
   Includes nouns with an indexical meaning component such as I, me, mine, you, yours, he, him, his, she, her, hers, etc.

6. *preposition*
   Function word which combines with a noun into an adjective, such as on in (the book) on (the table).

7. *conjunction*
   Includes coordinating conjunctions (parataxis) like and and subordinating conjunctions (hypotaxis) like that (introducing subject or object sentence) or when (introducing adverbial sentence).

8. *interjection*
   Includes exclamations like ouch!, greetings like hi!, and answers like yes and no.

Despite considerable variation, most traditional grammars postulate eight parts of speech because this is the number assumed in classical Greek and Latin grammar. In daily practice, however, additional classifications are used such as *determiner, auxiliary, modal, infinitive, progressive, past participle, present tense, past tense, singular, plural, first person, second person, third person,* etc., all of which are useful for a more precise classification of word forms in English.

Of the parts of speech in 9.1, *noun, verb,* and *adjective* are perhaps the most basic, as indicated by their having counterparts in logic, namely the notions of *argument*, *functor*, and *modifier*, respectively, and in more general philosophy as *object*, *relation*, and *property*, respectively (cf. FoCL'99, Sect. 3.4). This raises the question of how the traditional parts of speech and the core attributes of Database Semantics are related.

[40] We are leaving aside here extrapropositional functor-argument structure (subclauses, cf. NLC'06, Chapt. 7). There, the core attributes n/v (noun-verb) and a/v (adjective-verb) are composed from the three basic ones.

For example, 2. *noun* and 5. *pronoun* are classified as different parts of speech[41] in 9.1, but have the same core attribute noun in DBS. Treating pronouns as nouns is motivated because a pronoun like she can serve the same grammatical function, for example as subject or object, as a name like Julia or a noun phrase like the pretty young girl.

The special property of pronouns is sign-theoretic: they are *indexicals* as compared to the other two sign kinds of natural language, namely *symbols*, e.g., girl, and *names*, e.g., Julia. (cf. FoCL'99, Sect. 6.1, NLC'06, Sect. 2.6). Given that the sign kinds *symbol* and *name* are traditionally included in the part of speech *noun*, there is no systematic reason to exclude pronouns merely because they have the third sign kind *indexical* as their core value. Consequently, in DBS all three kinds of nouns are analyzed as proplets with the same core attribute, but with core values differing in their kind of sign. The following examples also include a determiner because they are likewise analyzed as proplets with the core attribute noun:

## 9.2 ANALYZING DIFFERENT KINDS OF NOUNS AS LEXICAL PROPLETS

| *common noun* | *pronoun* | *proper name* | *determiner* |
|---|---|---|---|
| sur: books | sur: they | sur: Julia | sur: every |
| noun: book | noun: ind_3 | noun: Julia | noun: n_1 |
| cat: pn | cat: p3n | cat: nm | cat: sg |
| sem: count pl | sem: count pl | sem: sg | sem: pl exh |
| fnc: | fnc: | fnc: | fnc: |
| mdr: | mdr: | mdr: | mdr: |
| idy: | idy: | idy: | idy: |
| prn: | prn: | prn: | prn: |

The common noun proplet has a meaning defined as a concept which is represented as *book*, serving as a place holder; analyzed word forms with a concept as meaning are called symbols. The pronoun proplet has a meaning defined as a pointer which is represented as *ind_3*; analyzed word forms with a pointer as their meaning are called indexicals. The name proplet has a meaning defined as a marker which is represented as *Julia*; analyzed word forms with a marker as meaning are called proper names.

The analysis of determiners (which some traditional grammars treat as a separate part of speech) as a noun proplet facilitates the fusion of a determiner and its noun illustrated in 8.1. Some languages, e.g., German, use determiners also indexically, which may be handled by widening the restriction on the substitution value to include indexical use.

Another traditional distinction is between 3. *adjective*, 4. *adverb*, and 6. *preposition*, which are treated in Database Semantics as proplets with the same core attribute, namely adj:

## 9.3 ANALYZING DIFFERENT ADJECTIVES AS LEXICAL PROPLETS

| *adnominal* | *adverbial* | *indexical adjective* | *preposition* |
|---|---|---|---|
| sur: slow | sur: slowly | sur: here | sur: on |
| adj: slow | adj: slow | adj: ind_loc | adj: *on* n_2 |
| cat: adn | cat: adv | cat: adj | cat: adj |
| sem: psv | sem: psv | sem: | sem: |
| mdd: | mdd: | mdd: | mdd: |
| prn: | prn: | prn: | prn: |

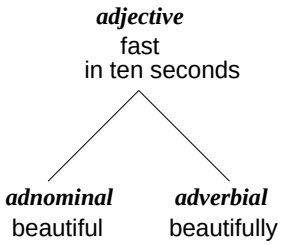
The proplet analyses of the content words slow and slowly are alike in that they have the same core value, but differ in their cat values adn (adnominal) and adv (adverbial). The obligatory continuation attribute is mdd, for "modified."

[41] This problematic distinction was postulated already by Dionysius Trax (170 BC – 90 BC).

Analyzing the adnominal and adverbial uses as proplets with the same core value may be motivated terminologically: the Latin root of *adjective* means "what is thrown in," which aptly characterizes the optional quality of modifiers in general. It may also be motivated morphologically because of the similarity between adnominal and adverbial adjectives. Consider, for example, the adnominal adjective beautiful and the adverbial adjective beautifully in English, or schöne, schöner, schönes, etc., (adnominal adjectives) and schön (adverbial adjective) in German. The two uses resemble each other also in their analytic degrees, as in more beautiful (adnominal) and more beautifully (adverbial). In synthetic degrees, as in faster, the adnominal and the adverbial form are not even distinguished in English.

Leaving aside the question whether indexicals like the pronoun they (cf. 9.2) and the adjective here (cf. 9.3) should be classified as function words or as content words,[42] we note that the latter allow adnominal use, as in the book here, as well as adverbial use, as in Goethe slept here. This is expressed in the lexical proplets by their cat value adj. The relation between adjectives marked morphologically for adnominal use only, adverbial use only, and adnominal or adverbial use may be illustrated as follows:

## 9.4 Relation between Adjective, Adnominal, and Adverbial in DBS




In lexical proplets, the distinctions in question are coded by the cat values adn, adv, and adj.

Prepositions are true function words in the sense that they get fused with associated content words, just like determiners and auxiliaries. Prepositional phrases and indexical adjectives have in common that they can be used adnominally as well as adverbially (cf. NLC'06, Chapt. 15), and therefore have the cat value adj.

Some traditional grammars also distinguish between finite verb forms, auxiliaries, and participles as different parts of speech. In Database Semantics they are treated as proplets with the same core attribute, namely verb.

## 9.5 Analyzing different verb forms as lexical proplets

| *finite main verb* | *finite auxiliary* | *non-finite main verb* |
|---|---|---|
| sur: knows | sur: is | sur: knowing |
| verb: *know* | verb: v_1 | verb: *know* |
| cat: ns3′ a′ v | cat: ns3′ be′ v | cat: a′ be |
| sem: pres | sem: be_pres | sem: prog |
| mdr: | mdr: | mdr: |
| arg: | arg: | arg: |
| prn: | prn: | prn: |

Finite main verbs and finite auxiliaries have a category ending with the constant v in common. They differ because auxiliaries, but not main verbs, have the constant be', hv', or do' as part

[42] A similar question is whether proper names should be classified as content words. We favor the general position that all words with their own reference mechanism, be it that of symbol, indexical, or name, are content words. In contrast, all words which do not have their own reference mechanism and are therefore fused with a content word (absorption), such as determiners, prepositions, conjunctions, and auxiliaries, are function words.

of their categories. Also, the core value of an auxiliary (function word) is a substitution value, while that of a main verb is a concept. Non-finite verb forms differ from finite ones by the absence of the constant v at the end of the cat value.

There remain the traditional parts of speech *7. conjunction* and *8. interjection*. Conjunctions are function words; in Database Semantics their lexical proplet analysis is adapted to that of the content words they are being fused with – as well as to their special function of connecting the main clause with the subclause in hypotactic constructions. Interjections are one-word sentences; their lexical analysis is not constrained by the compositional considerations of establishing semantic relations and still await an analysis as proplets in Database Semantics.

Even though DBS analyzes word forms as proplets with only three different core attributes, the traditional eight parts of speech may nevertheless be reconstructed in the form of proplet *patterns*. For example, using the differentiated lexical analysis of word forms in 9.2, *pronouns* as a part of speech may be characterized by a core value defined as a variable restricted to a set of pointer values, and *determiners* by a core value defined as a variable restricted to a set of substitution values. Similarly, using the lexical analysis in 9.3, *preposition* as a part of speech may be characterized by a core value defined as a variable restricted to substitution values, *adjective* by a core value defined as a concept and the cat value adj, and *adverb* by a core value defined as a concept and the cat value adv, etc.

# 10 Conclusion

The data structure of proplets is motivated by the overall purpose of DBS to model the cycle of natural language communication on a computer. Part of this cycle is the hearer mode, in which a sequence of language-dependent lexical proplets is parsed into a representation of content. This representation is language-independent insofar as (i) function words and morphological markings are interpreted as values of the cat and sem attributes and (ii) the word order of the surface is dissolved into an order-free representation of the semantic relations for purposes of storage and retrieval in a database.

The relative language-independence of content representation in DBS raises the question of whether or not the core attributes *noun, verb,* and *adj* should be regarded as universal – which corresponds to be a highly controversial issue in language typology. For example, Cayuga (spoken by native Americans in Canada) has been argued to have no distinction between verbs and nouns (Sasse 1993). Furthermore, there is the question of whether all natural languages use the same universal representation of content, discussed by Nichols (1992) in the tradition of the Humboldt–Sapir–Whorf Hypothesis (linguistic relativism).

As a computational framework, DBS is neutral on these issues. All that the abstract system of DBS requires is (i) the use of flat feature structures for the lexical analysis of word forms in terms of a surface, a meaning, continuation features, and a proposition number, (ii) a coding of semantic relations by means of attribute-value pairs, and (iii) using the data structure (a) as input to the time-linear algorithm of LA-grammar in the variants of LA-hear, LA-think, and LA-speak, (b) for storage and retrieval in a database, and (c) for matching between content at the language and the context level.

These requirements do not depend on the use of any particular attributes or of any particular values. Rather, when using the framework of DBS it is up to the linguist to select terms for attributes, values, and relations which are well-motivated for the natural language in question. Detailed computational applications to English, German, Chinese, Tagalog, Russian, and Korean have shown that DBS is well-suited to model the cycle of natural communication in a wide range of different languages.

# Bibliography


Aho, A.V. and J.D. Ullman (1977) *Principles of Compiler Design*, Reading, MA: Addison-Wesley

AIJ'01 = Hausser, R. (2001) "Database Semantics for natural language," *Artificial Intelligence*, 130.1:27–74

Barthes, R. (1986) *The Rustle of Language*, New York: Hill and Wang

Barwise, J. & J. Perry (1983) *Situations and Attitudes*, Cambridge, Mass.: MIT Press

Chomsky, N. and M. Halle (1968) *The Sound Pattern of English*, New York: Harper Row

Cole, R., et al. (eds.) (1998) *Survey of the State of the Art in Human Language Technology*, Edinburgh, Scotland: Edinburgh Univ. Press

FoCL'99 = Hausser, R. (1999) *Foundations of Computational Linguistics, Human–Computer Communication in Natural Language, 2nd ed. 2001*, Berlin Heidelberg New York: Springer

Hausser, R. (1996) "A Database Interpretation of Natural Language," *Korean Journal of Linguistics*, Vol. 21.1,2:29–55

Hausser, R. (2009) "Modeling Natural Language Communication in Database Semantics," in M. Kirchberg and S. Link (eds.), Proceedings of the APCCM 2009, Australian Computer Science Inc., CIPRIT, Vol. 96

Kemmer, S. (2003) "Human Cognition and the Elaboration of Events: Some Universal Conceptual Categories," in M. Tomasello (ed.) *The New Psychology of Language, Vol. II*, Mahaw, N.J.: Lawrence Earlbaum

MacWhinney, B. (2008) "How Mental Models Encode Embodied Linguistic Perspective," in L. Klatzky et al. (eds.) *Embodiment, Ego-Space, and Action*, New York: Psychology Press

MacWhinney, B. (ed.) (1999) *The Emergence of Language from Embodiment*, Hillsdale, NJ: Lawrence Erlbaum

Montague, R. (1974) *Formal Philosophy*, New Haven: Yale Univ. Press

Nichols, J. (1992) *Linguistic Diversity in Space and Time*, The Univ. of Chicago Press

Pullum, G. K. (1991) *The Great Eskimo Vocabulary Hoax: And Other Irreverent Essays on the Study of Language*, The Univ. of Chicago Press

NLC'06 = Hausser, R. (2006) *A Computational Model of Natural Language Commmunication: Interpretation, Inference, and Production in Database Semantics*, Berlin Heidelberg New York: Springer

Sasse, H.J. (1993) "Das Nomen – eine universelle Kategorie?" in *Sprachtypologie und Universalienforschung*, Vol. 46,3:187-221

Saussure, F. de (1916/1972) *Cours de linguistique générale*, Édition critique préparée par Tullio de Mauro, Paris: Éditions Payot

SCG'84 = Hausser, R. (1984) *Surface Compositional Grammar*, Munich: Wilhelm Fink

TCS'92 = Hausser, R. (1992) "Complexity in Left-Associative Grammar," *Theoretical Computer Science*, 106.2:283-308

Tomasello, M. (2003) "Introduction: Some Surprises to Psychologists," in M. Tomasello (ed.) *The New Psychology of Language, Vol. II*, Mahaw, N.J.: Lawrence Earlbaum

Wilson, E.O. (1998) *Consilience: The Unity of Knowledge*, New York: Alfred Knopf

Zue, V., R. Cole, & W. Ward (1995) "Speech Recognition," in R. Cole et al. (eds.) 1998