Use of natural language for knowledge acquisition: Strategies to cope with semantic and pragmatic variation

by T. Wetter
R. Nüse

The large amount of verbal data from common knowledge-elicitation methods suggests using the data directly for knowledge acquisition by means of sophisticated natural-language analyzers (NLAs). In this paper, we analyze the feasibility of such an approach theoretically and present a number of examples. In the theoretical part of the text we first provide a detailed analysis of the *entities* involved, i.e., the domains of expertise, the qualities of knowledge about domains, the properties of generic sentences and texts in natural languages, and the conclusions to be drawn from the limited expressiveness of formal representations. Then we discuss the *processes* of transforming knowledge into natural language and of transforming natural language into formal language. Since much can go wrong in both processes, we derive desired relations or *validity criteria* among the entities and *strategies* to meet the criteria. We believe that this broad theoretical framework can be used to analyze and compare existing attempts at directly using natural language for knowledge acquisition, and thus assess the present status of the field.

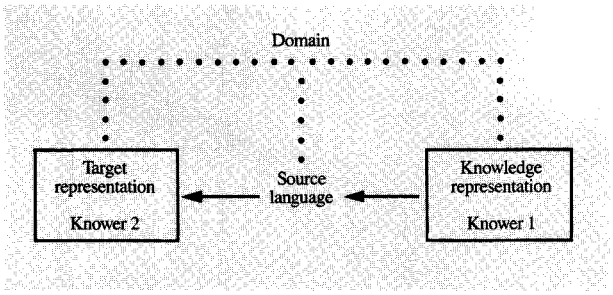

Figure 1

Basic model of natural-language knowledge acquisition (dotted lines represent *representation*, arrows represent *transformation*).

1. Introduction

There is a remarkable discrepancy between the important role of natural language in communicating and defining knowledge, and the few systematic considerations that exist for making constructive use of natural language in knowledge acquisition. On the one hand, humans use natural language for a large variety of purposes, all of which might serve the aims of knowledge acquisition. It allows for teaching; i.e., humans can learn from other humans by being told. It allows for reflecting on one's behavior; i.e., rational arguments for behavior and decisions can be put into natural language. It is the means of codifying knowledge about domains in documents—formally sound bodies of knowledge exist as textbooks. And it is the means of spontaneously externalizing things that humans learn or discover during cognitive processes—traces of problem-solving behavior are expressed as natural-language utterances.

On the other hand, so far there have been only a very small number of serious attempts to use expressions of natural language directly for knowledge acquisition, although the primary data from many knowledge-elicitation techniques (in particular, a host of variations of interviews and "think-aloud" settings) and textbook entries are in natural language.

Two related reasons may be that

1. The variety of semantics and pragmatics of utterances is faintly felt but not systematically understood.
2. Computational linguistics is only slowly approaching a state that would allow the computer immediately to "understand" natural language in the sense of transforming texts from a natural language into a formal internal representation.

In this paper, after introducing some basic structural and content-related assumptions about entities, languages, and processes, we

- Analyze the semantics of generic sentences and expository texts, and relate that semantics to properties of domains and languages.
- Analyze the conditions under which knowledge is presented and received in natural language and the problems that result from attempting to automate the process.
- Propose certain validity criteria and (partly cooperative) strategies.
- Give examples in which the strategies have been used in existing approaches.

2. Underlying assumptions and scope

- *A model of natural-language communication*

Basically, knowledge acquisition via natural language much resembles "normal" communication: Someone (e.g., an expert) knows some relevant facts about a particular domain and tries to describe those facts by using a natural language. Someone else (e.g., a system) "understands" those utterances (or written symbols) and, as a consequence, also "knows" the facts described. This process can be conceived of as a sequence of transformations of representations: Knowledge, a *mental* representation of facts, is "coded" in a (non-mental) natural-language expression representing the *same* facts, which in turn is retransformed into the "mental" representation language of the receiving system. Thus, we start with the simple basic model shown in **Figure 1**.

Hence it appears that we assume at least four entities involved in the process: knower 1, who has knowledge coded in some mental language; the source language, which conveys this knowledge to some receiver; knower 2, who is the recipient of the source-language expression (and may be, say, a software product); and finally a domain to which the aforementioned representations refer in one way or another. We also assume that natural-language expressions (and sequels of them such as texts) as such ("objectively," as it were) *represent* facts being in existence. This is not to deny that a receiver must have a certain structure to adequately grasp and understand the information "contained" in a text, that the receiver must be built in a certain way in order to be affected by a stimulus such as a text. On the contrary, we emphasize the fact that a receiver must possess large amounts of world knowledge to fulfill this task. But—unlike current "constructivist" approaches to the problem of text meanings (e.g., [1, 2])—we maintain that the structure of the text and thus the information it conveys "exists" to be processed, and that this structure is systematically related to the structure of knower 2 so that, given full knowledge of the latter, the text will affect knower 2 in a predictable way. (For detailed arguments against the philosophy underlying the constructivist approaches, see [3].)

The entities and processes of the basic model guide the following discussion. First, we focus on the entities involved and discuss different types of domains, qualities of knowledge, etc., and the conditions they impose on the whole enterprise. We then describe the processes which transform the respective representations, that is, the production and reception of a natural-language text, which we consider to be the main sources of problems concerning the transfer (and thus the acquisition) of knowledge. Since these problems go back to the idea of an "optimal" communication between knowers 1 and 2, where knower 1 says what he wants and knower 2 understands all of the sentences knower 1 produces, we subsequently discuss some possibilities for overcoming these problems by adhering to more relaxed demands. This requires us to establish criteria of validity, i.e., desirable relations between entities (see Section 4 following), and to ask in what way and to what extent transformations may meet these validity criteria. We then introduce strategies for the processing and production of source language, so as to better meet the criteria. Finally, by using examples (Section 5), we show how to apply the established framework to produce solutions from our own and other investigations to some of the problems raised.

• *Properties of the entities of the model*

In this paragraph, we only consider the *entities* mentioned in Figure 1. Here we want to distinguish

- among different target representations,
- among different types of domains,
- among different qualities of knowledge about a domain,
- among different ways in which the knowledge can represent the domain, and
- among different source languages that can be used.

A first impression of the distinctions we want to make can be derived from **Figure 2**, which is a focused view of Figure 1.

Target representations and the problem of static and dynamic languages

Mapping into a target representation has two major aspects: First, choices can be made among more or less suitable languages. Reasons for making a careful choice are given in the subsection on formal target representations. Any admissible choice remains in the realm of formal languages and inherits their limited expressiveness.[1] Second, once a choice has been made about the formal language as a whole, there remains the problem of mapping into individual expressions of that

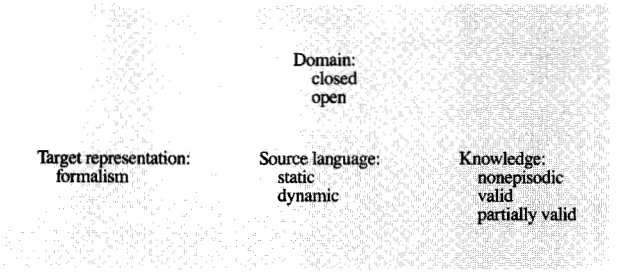

Figure 2

Distinctions among different properties of the entities.

language. The extension of this problem and strategies to solve it are outlined later in this section. Here we deal with an additional principal consequence of the necessity to use a formal language, the problem of static and dynamic languages. It follows from the distinction in [4] between static and dynamic languages, which may resemble the philosophical distinction between ideal and ordinary language.

The meaning of "meaning" is completely different in static and dynamic languages. A static language can only cover one aspect of the real world. Meaning emerges as the final result of identifying the essential concepts of this domain, introducing unique names for those concepts, and writing down known regularities. Nothing new can happen in a static language, and no discovery or development can be made. Formal languages are a subclass of the static languages; i.e., they inherit the stationarity of static languages.

In dynamic languages the meaning of an expression is extensionally characterized by its occurrences in all contexts. Every new use of an expression adds an occurrence and hence changes its meaning. Newly mentioned exceptions, for example, gradually change the scope of a default rule. Dynamic languages support development, discovery, and (this is the main argument in [4]), scientific progress. Hence, one particular formal language can be used only to precisely describe one state of cognition, not the progress from state to state.

Closed and open domains

The main distinction to be introduced subsequently concerns openness or closedness of domains, with the practically relevant intermediate form of artificial closure [5].

Typical examples of open worlds would arise in medical diagnosis, job applicant assessment, etc. Degree of illness, experience, etc. are arbitrary temporary groupings, which may undergo *gradual* change through the occurrence of new instances or counterexamples, or *dramatic* change

[1] Progress in complexity theory, logic, computer architecture, etc. may widen the class of tractable languages over time, but at any point in time there will be some definite limitation.

through the appearance of qualitatively new groupings, such as new environmental diseases.

Typical examples of closed worlds are games such as chess. The physical or world state of 32 pieces or 64 squares can be mapped uniquely to a representation of that state. Classes or sequences of such states could also be given unique names such as "final" or "gambit" (though in practice there might occur fierce discussion among enthusiasts as to whether some state is in class A or class B). This closedness is reflected in a way by the fact that people can play chess blindfolded with no other expressions than "e2–e4," "h7–h5," etc. There is nothing to add, no "untypical e2–e4," no "most h7–h5"; the world of the states denoted by "a1" through "h8" is definitely closed.

Between the open and the closed worlds there is the intermediate class of artificially closed worlds. Digital circuits are an illustrative example. Each component has a real-number description of its state in terms of voltage, current, capacity, etc. From the perspective of its functionality, however, subthreshold values are interpreted as 0, above-threshold values as 1. This transforms the open world of real states of conductive components into the closed world of on and off switches. This artificial closure (which is the basis of all quality-assurance circuit testing) may be sufficient for some problems (e.g., diagnosis of faulty components) and insufficient for others (e.g., amplifier layout to compensate for fading voltages). The borderline of sufficiency cannot even be defined precisely, since some fault attributed to a certain switch (which is the closed-world interpretation) can as well be caused by reduced amplification in a neighboring component. This clearly indicates a distinction from real closed domains, where physical variations (whether a chess piece is in the center or somewhere else in the square) do not play any role.

Source languages and domains

It appears that dynamic languages are natural instances of open domains: The meaning of an expression (which may denote a type or set) changes slightly with every new situation to which the expression is applied. That is, talking about medical diagnosis, job applicant assessment, etc. in natural language as the prototypical form of a dynamic language is a recommended combination. It should, however, be noted that natural language is also used to communicate about the functionality of a software product, which can only be a closed domain spanned by the uniquely determined, distinguishable states of a computer. In this case the precisely knowable regularities of the domain can be misconstrued by expressions which would allow undesirable, competing generic interpretations. This finishes the combination that will be followed up in this outline on use of natural language for knowledge acquisition. (As a small aside, we note that in real life static languages are also used for natural communication about closed and open domains.)

The exemplary closed case is the aforementioned blind chess game, whose communicational aspect is fully covered by the exchange of chess square coordinates. Decimal classifications of medical diagnosis, such as ICD9, are an instance of a "naturally" occurring static language for an open domain. The interesting aspect of this last case is that the problems outlined below of effectively mapping into a static language now occur between the knowledge of a knower and the source language.

- *Qualities of knowledge about the domain*

Independently of the real nature of a domain, we must deal with the knowledge about the domain that individuals (or communities) possess. We start with some basic framing assumptions, which are not necessary in principle but help to keep the size of the approach feasible.

Existence of laws The knower has available some laws about the domain, i.e., something expressing regularity, which allows him to sufficiently predict and control the domain. At this stage we make no claims about the concrete form or contents of such laws, but refer the reader to the section on formal target representations, below, which illustrates the expected variety of laws by means of the respective variety of semantics of generic sentences.

Nonepisodic nature of laws The only content-related assumption is that regularities are not known merely in terms of cases or episodes, but in terms of relations between attributes and values of cases.

Validity of knowledge That knowledge about the domain need not coincide with the nature of the domain can be argued historically by means of the Kopernikus/Galilei or the Newton/Planck–Heisenberg transitions. The planetary system or physics did not change, but human knowledge did.

More practical examples can be based on investigations about poor job performance (cf. [6]): Medical doctors as individuals know much less than could in principle be known. That is, the actual properties of a domain and the reflection of those properties in an individual instance of knowledge cannot be expected to be a one-to-one correspondence. Regularities that knower 1 "knows" need not be regularities which are really present in the domain. The relation between the purpose of the knowledge a person applies and the degree of correctness or deviation has been studied intensely in the literature on mental models.

- *Formal target representations*

In the last paragraph, we emphasized the fact that the target representation language has to be a formal language. Describing the content of natural-language expressions denoting law-like regularities by using a formal language does, however, present some difficulties, as is well known in linguistics with regard to the formal description of truth conditions of so-called generic sentences, that is, sentences of natural language expressing "laws" or rules. Among others, these well-known difficulties include the following:

- Since the truth of generic sentences such as "dogs bark" is *not* affected by the fact that there are exceptions to this statement, while the truth of a sentence such as "All dogs bark" is, it follows that the truth conditions of generic sentences cannot involve universal quantification [7–9]. In addition, generic sentences sometimes not only fail to be valid for all members of a class, but, strictly speaking, refer to only a relatively small number of members. Between these two extremes, all graduations are possible [10]:

  Human beings are mammals (is valid for all of them).

  Telephone books are thick (is valid for most of them).

  Swallows lay two eggs (is valid for almost half of them, namely for females).

  Scotsmen drink whisky (is valid for a rather small percentage, which nevertheless is high compared to other countries).

  Frenchmen eat horsemeat (is worth mentioning even if only a few Frenchmen do it).

  Finally, for sentences such as "At the end of the Ice Age, man migrated from the east toward Europe," the question is whether there is any sense in either universal or existential quantification over individuals at all.
- In other cases it may even be unclear in which way a given generic sentence is to be read as a conditional at all. For instance, a sentence such as "Dogs bark" may be interpreted as "If something is a dog, it barks," or —especially with a stress on "dogs" and as an answer to a question such as "What's that barking round the corner?"—just the opposite, as "If something barks, it is a dog" [11].
- Natural-language sentences such as "Ted votes for liberals" are indeterminate with regard to the values of certain variables. In consequence, they are compatible with more than one formal description of their truth conditions. In the above-mentioned sentence, for example, it is not specified whether Ted votes for liberals at *every* election or whether Ted does so only occasionally. Similarly, it is not marked whether Ted *always* votes for liberals, whether he votes for liberals *only*, and whether he votes for the *same* group of liberals, that is, whether the liberals he votes for remain the same set of people from one election to the next, etc. [12].

For other problems of formally representing the meaning of generic sentences see, e.g., [10].

Now, these difficulties of coding knowledge in a formal language are not only due to the limited expressiveness of the respective formal language, but may arise from each of the four entities of Figure 1 mentioned above. Thus, each may call for a different problem-solving strategy. The problems of indeterminacy in the sentence "Ted votes for liberals," for instance, are problems of *verbal presentation* only. In principle, the corresponding denoted facts can be described in a precise and formal manner; the denoting natural-language expression merely does not indicate which of these facts is present. A similar case can be made with regard to those Frenchmen eating horsemeat. This fact can be described in a formal manner, too (perhaps even by using simple existential quantification). The only problem with this sentence is that it looks like a generic sentence if one looks only at its surface structure, and that its interpretation depends on pragmatic factors such as the strangeness and the corresponding noteworthiness of the behavior described. In both cases, the problems of formalization lie only in the determination of the correct interpretation of the respective natural-language expression and are not due to any limit on the expressiveness of the formal language itself. Consequently, an adequate solution to these problems would consist not in optimizing the formal target representation language but in clarifying the natural-language expression.

By contrast, the problems of universal quantification in the case of Scotsmen drinking whisky seem to have their roots in the *organization of human knowledge*, especially knowledge of category membership, which in certain cases is stored as information about a *prototypical* example of that category. The same applies to problems of universal quantification in sentences such as "If one scrapes a match against a striking surface, it will ignite," which arise from the fact that certain (not stated but presupposed) background conditions (such as the presence of oxygen) must be met in order for the sentence to be true. These problems are problems of the organization of human knowledge, too: The respective background conditions are specifiable and formally manageable in principle (e.g., by using "circumscription" or related methods of default reasoning); they merely do not come to mind without effort. Consequently, an adequate strategy for solving this problem would be to enhance the method used to ascertain the respective knowledge.

By way of contrast, the problems with exceptions to rules such as "Dogs bark" are not due to the organization

of knowledge. Rather, they arise directly from the facts that are inherent in the respective domain itself: It happens that some dogs don't bark, so you cannot conclude from "Dogs bark" that Fido barks. In a similar vein, you cannot conclude from the fact that man migrated from the east toward Europe at the end of the Ice Age that any particular human being does or did so (we didn't, for instance). This case is a particularly interesting one, since this problem is not solved by inserting a default rule such as "if nothing else is mentioned." Rather, the predicate "migrated from the east toward Europe" does not apply to any human individual at all, but instead to the *kind* "human being," and this is the deeper reason why one cannot conclude anything about a particular human being even if it is known that man migrated from the east toward Europe. Hence, we are dealing with at least two kinds of generics, or facts that are denoted by generics, respectively: one that is valid for a set of individuals of a particular kind, and another one which is valid for the kind itself (see also [9, 13, 14]). Problems with universal quantification due to this distinction therefore call for other means of representation than simple universal quantifiers, e.g., an operator that turns predicates into kind names (cf., e.g., [10, 14]) or other operators that are sensitive to the above-mentioned distinction.

To be sure, there also remain some problems caused by the limited expressiveness of formal languages in general. This is even true when we imagine the practically irrelevant case of having available all possible extensions of some base language, e.g. 1stPL (first-order predicate logic). For practical purposes in the realm of this text, we must also account for computational tractability. This implies that we must make some choice, e.g. whether to use circumscription *or* inference rules for kind operators together with some base representation such as 1stPL.

In this paper, we do not intend to treat all of these problems, of course. Rather, we want to concentrate on those problems which are due to the transfer of knowledge via natural language, that is, problems such as the missing indication of an intended reading of a natural-language expression. This means in particular that we simply only talk about domains which allow for a (sufficiently) formal description, where some knowledge exists about them which furthermore can be recalled and may be uttered correspondingly. We assume that the knowledge thus verbalized can be taken as a starting point of language processing, and that there is an approximately adequate formal representation of this domain as the end-product of language processing. Thus, we first comment on the difficulties that may arise during the respective *processes* that must be passed through in the transfer of expert knowledge into a formal representation language via natural language. The consequences of having to deal with target languages of insufficient expressiveness are dealt with in a later section.

3. Problems of natural-language knowledge acquisition

Having given a basic perspective and explained our assumptions about natural-language knowledge acquisition, we are in a position to discuss the problems that may arise in the process. In doing so, we base our treatment in part on experience with our current system of natural-language knowledge acquisition, KALEX (see [15]), and in part on investigations of real presentations of knowledge in various domains, such as commentaries on the penal law and medical or physical textbooks, interviews or oral definitions, and explanations by experts on certain topics.

The above-mentioned basic model of natural-language knowledge acquisition includes two transformation processes that must be passed through in the process of natural-language knowledge acquisition: The expert knowledge must be transformed into natural language, and the linguistic expressions resulting from this process must be transformed into "knowledge" again, that is, into expressions of a chosen representation language, by the recipient (or the respective system). Since both processes are subject to specific dynamics and have to satisfy their own respective conditions, these two processes are the "weak points" of natural-language knowledge acquisition. On the one hand, the properties of language production are derived from the requirements of the communicative situation in which people normally speak or produce texts. These requirements differ from those prevailing in the knowledge acquisition of an expert system. Accordingly, it may easily happen that expert and system "are talking around each other": The expert says something which is useless for the system, and the expert fails to mention certain things which the system would need to know.

On the other hand, there are those requirements that mechanical, automated language processing imposes on the construction of an NLA. Such a form of language processing calls for explicit rules for mapping verbal expressions onto the respective elements of knowledge. Now, as we would like to show in the following, this kind of mapping is anything but easy to provide. Thus, even if expert and system are not talking around each other, that is, if the system is able to use what the expert says, it remains difficult for the system, on the basis of the available information, to infer what the expert *meant* from what he *said*.

The problems of the acquisition of knowledge transferred via natural language therefore arise from two sources: the *production* of the (natural-language) knowledge presentations by *man* and the *reception* of these tests by a *machine*. In the following, we illustrate these difficulties in relation to these two aspects. The

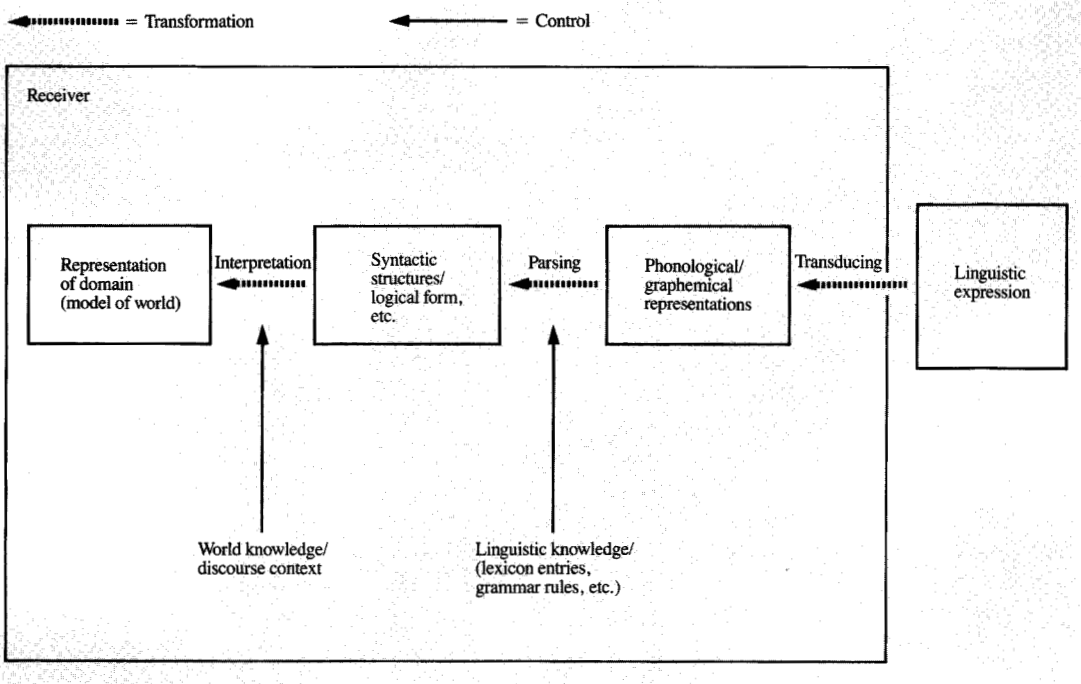

Figure 3

Language understanding as translation into representation language, cf. [16–18].

"logical" relationship between production and reception, of course, is such that the problems of production precede the problems of reception, hence the reception problems remain even after the production problems have been solved. However, since an illustration of those knowledge presentations easily "received" or processed by a system reveals what an expert should not say if he intends to transfer "useful" knowledge to the system, we comment on these problems in the reverse order.

• *Problems of language processing*

In our basic model, the following framework with regard to language comprehension has already been established: The "starting point" of language comprehension is an expression of natural language, while the end-product is a (synonymous) expression of the corresponding representation language of the system. In this sense, language comprehension can be conceived of as the transformation of one verbal expression into another without changing its meaning. Since a transformation of this type is normally called a "translation," one might also say that the understanding of a system consists in the *translation* of a natural-language expression into the corresponding expression of the system representation language. If, for instance, the representation language of a system is a version of 1stPL, the understanding of the sentence "A cat sat on a mat" would consist in the system providing the formula "$\exists x\ \exists y\ \text{cat}(x) \wedge \text{mat}(y) \wedge \text{sat on}(x, y)$." Other common representation languages are semantic networks, graphs, frames in the sense of Fillmore, or graphic representations, etc.

Figure 3, which is in a sense a kind of blow-up of our basic model, shows this conception of language comprehension. As indicated by References [16–18] in the illustration, our model is completely in line with accepted concepts of language comprehension in AI. In addition to this, similar concepts are found in all other branches of cognitive science (compare, e.g., the diagrams of [19], p. 577; [20], p. 652; [21], p. 274 f.).

If language comprehension is conceptualized in this way, one can derive those difficulties which confront natural-language understanding and thus natural-language knowledge acquisition. To begin with, there are all the well-known difficulties in establishing the intended meaning of a natural-language expression from its surface form. Remember that the target representation must be a formal,

machine-processable language. This implies in particular that any meaning variations in natural language will have to be denoted by different expressions of the representation language. This, in turn, means that different meanings unmarked at the language surface in natural language, that remain to be resolved in the process of comprehension, must be represented explicitly in the target representation language. For instance, when representing the meaning of the sentence

The shelf is the appropriate place for the diaper, if it is dry

the corresponding formula should (in contrast to the natural-language sentence) contain an expression stating (explicitly) that the personal pronoun "it" and the noun-phrase "the diaper" refer to the same entity.

But how does a hearer or system know whether the pronoun refers to the diaper or to the shelf, that is, whether it is the former or the latter which must be dry? Since this merely depends on the "plausibility" of the possible interpretations, the relevant question can be decided not by inspecting the structure of the sentence, but only by recourse to world knowledge. In other words, the structure (or linguistic form) of a natural-language expression is essentially *ambiguous* with regard to its possible interpretations. In terms of our illustration, one might say that the transformation processes are underdetermined with regard to their outcomes, or that both parsing and semantic interpretation are many-to-many functions. This is exactly why they are controlled by further knowledge such as lexicon entries and world knowledge: It is only with the aid of these stored stocks of knowledge that ambiguities can be resolved.[2]

The second difficulty arising from the model is not so well known, but nevertheless points in the same direction. It relates to the fact that the very "translation mode" of understanding as conceptualized in Figure 3 might be inappropriate for certain "pieces" of knowledge; that is, for some elements of natural-language knowledge presentations, although in a sense they constitute "knowledge," other kinds of processing than "mere" translation into a representation language are required. In the context of knowledge acquisition, for instance, this is especially true of *examples* meant to illustrate a particular point made in describing a domain or a mechanism. In this case, it would be senseless to translate the sentences describing an illustrating example into a corresponding formula of the representation language and store this formula in the knowledge base; this would rather seem to be missing the "point" the example is supposed to make. Instead, examples must be analyzed in a different way: Rather than just *storing* the facts described, the analysis must include interpreting each example as a special case of a general rule, which means (among other things) that it must recognize which features are essential for the example to be an example and which are peripheral with regard to the general case to be illustrated. Thus, a system that understands examples must first of all recognize that a certain piece of text is *meant* as an example. Second, it must be capable of *analyzing* this text in the appropriate way. This is not to say that examples (or in the general case, the recognition of changes in the appropriate mode of analysis) pose problems that must be regarded as unsolvable, but they do require a whole new class of computations and decisions. That is, just as getting the meaning of a natural-language expression right requires *knowledge* on the part of the receiving system, so does recognition of the "point" of certain expressions.

Thus, the well-known implication of both of these difficulties is that understanding natural language requires knowledge. While this might not seem to be so serious in the context of computational linguistics *per se*, it may have some unwelcome consequences in the more application-oriented field of knowledge acquisition. On the one hand, building extensive knowledge bases before doing "real" knowledge acquisition in natural language does not seem reasonable (notice, however, that even one of the best knowledge-acquisition systems using natural language one can imagine, a human child, needs four years at the very least to pick up all the knowledge it must possess in order to understand natural-language texts). On the other hand, there arises the interesting paradox that a natural-language text which is supposed to convey knowledge to a system, to some extent requires exactly this knowledge in order for the text to be understood!

In what follows, therefore, we do not want to rehearse all the well-known difficulties of natural-language understanding. Rather, we focus on those problems which have a special relevance in the context of knowledge acquisition, either because they lead to the paradoxical situation described above or because they involve linguistic constructions that are typical of knowledge presentations written (or spoken) in natural language. The first difficulty, "misunderstanding" of the intended meaning, we will call "semantically incorrect resolution"; the second difficulty, a correct semantic analysis of a sentence (in the sense of a correct translation into the representation language) that as such constitutes an incorrect "use" of the respective natural-language expression, will be called "pragmatically incorrect resolution." In linguistics, the presentation of knowledge and the expression of laws and interrelations have been studied, especially at the sentence level, with regard to the comprehension and the semantic properties of so-called "generics" (see "Formal target representations," above).

[2] Besides the problem of anaphora resolution illustrated in the example, well-known problems resulting from ambiguities include multiple parses, the determination of the scope of quantifiers and negation, time and tense, word sense ambiguities, etc.; for a short enumeration and a quick review of some of these problems see, e.g., [22–24] and others.

In addition, some (rather psycholinguistically minded) studies on the comprehension of expository *texts* have been conducted. Accordingly, we treat the sentence and text levels separately. Because both difficulties occur on both levels, we get a kind of fourfold table of the problems to be faced.

How generic sentences can be misunderstood

Semantically false resolutions Just as in the example above relating to the resolution of personal pronoun reference, in generic sentences also it is usually not clearly marked whether they refer to individual situations or law-like regularities. Bertrand Russell has already pointed out that sentences do not wear their logical form "written on their sleeves," and a semantically false resolution of generics is often due to the lack of such a marking. A well-known example of this phenomenon, which is especially relevant to the interpretation of knowledge presentations, is the generic use of the definite article. A sentence such as "The jay has a special beak" has the same surface structure as "The jay has a broken wing." The first sentence, however, is a generic one, which would have to be analyzed as a universally quantified conditional (and thus constitutes "knowledge" about jays), while the second sentence describes an event; here the definite article would have to be read as an existential quantifier (or in an otherwise nongeneric way, e.g. as a reference to an aforementioned specific jay). Thus, as in most other areas of natural language, it is true of generic sentences that there are no clear *syntactic* markers indicating that a generic reading of a sentence or linguistic form is required. Almost all linguistic constructions which are normally referred to as generic descriptions (that is, as having a generic reading) allow for a nongeneric reading as well. For instance, of the seven "generic forms" which Heyer ([9], p. 94; see also [13]) mentions—definite and indefinite NPs, each singular as well as plural, as well as mass terms, quantified NPs, and habituals—only No. 6, "quantified NPs," is unambiguous.[3] "All X are Y" permits a generic interpretation only, although in "real" uses of this form the universe of discourse is often extremely restricted by the context ("In this bucket there are three logs and two balls. All balls are blue."), such that one might hesitate to call such a sentence generic. All other items listed by Heyer are either classical cases of ambiguity or otherwise problematic (see also [25]).

At this point one may wonder why one should confine oneself to the use of syntactic markers only and not include semantic information such as "type of verb" or others to resolve the above-mentioned ambiguities. Apart from semantic criteria such as mass terms or stative verbs, verb or noun classifications might be generally used to disambiguate the linguistic construction. Bare plurals, for instance, are definitely generic if used with so-called individual-level predicates, and nongeneric if used with stage-level predicates (e.g., "Pigs are intelligent" vs. "Pigs are in my back yard" [26, 27]). Similarly, there are other kinds of verbs that permit only a generic reading of an NP. Verbs such as "to invent," for instance, cannot be applied to individuals but only to kinds, and thus give rise to a "kind-reference reading" of an attached object-NP (the same is true of predicates such as "extinct," "is a fruit," and others). And finally, characteristics of the discourse context may also clarify how a generic description is to be read. The definite article, for example, if it is to be used in a nongeneric way, is subject to special conditions of usage that may be approximately characterized as follows: The definite article refers only to previously mentioned or generally "known" entities. Thus the fact that an entity has already been mentioned could be used as a criterion to exclude a generic reading.

It should be noted, however, that such semantic characteristics are not infallible in each case. This is certainly true of the disambiguation of the definite article. Even if there is an entity that has already been mentioned to which the definite NP could refer, the NP may still be given a generic reading, as in the following example:

In this fruit bowl there are a kiwi, a mango and a passion fruit. These are all fruits from other continents. The kiwi is a fruit from New Zealand, which has only recently become known in Europe.

The same applies to verb or noun classifications. They may at the most act as heuristics which are not necessarily infallible (see also [9], p. 98 on the interpretation of the definite article). Moreover, although in linguistics some proposals concerning verb classifications have been offered (e.g., the well-known classification of [28], Ch. 4), to the best of our knowledge there is no complete application of these classification schemes; that is, we do not know of a case in which all verbs of a given language (e.g., all English verbs) have actually been classified according to such a classification scheme—which is exactly what an NLA needs in order to solve the problems listed above. Accordingly, a solution of those problems is still not in sight. This is why it is worthwhile to look for other strategies to facilitate the analysis of presentations of regularities in natural language here and now.

[3] Items 5 and 7 in Heyer's list are "mass terms" and "habituals." Although these two items may unequivocally indicate a generic reading, they do not constitute linguistic *forms*, or syntactic markers, but basically must be regarded as *semantic* classifications which already presuppose a generic interpretation of the corresponding syntactic form. Thus, the sentence "Gold is heavier than silver" permits a generic reading only; the sentence structure "X is heavier than Y," however, does not (cf. "Peter is heavier than John"). The same applies to habituals: "John smokes" may clearly be a generic sentence (irrespective of the fact that "Hans raucht" would definitely *not* be unequivocal in German); however, this is not the case in "X x-s" or "NP VPpres" (e.g., with state verbs, as in "The carpet smells" or "She loves me"). In both of these cases, there are no *syntactic* markers indicating a generic reading, but semantic ones: Only if the system has realized that the verb is not a stative one, or that a mass term has been used, can it decide whether the sentence must be interpreted generically or not.

In addition, the use of semantic criteria is still plagued by other problems. For, although the discussed types of sentences and linguistic phenomena correspond to those forms discussed as "genericity" in linguistics, it can be observed that in "real" natural-language descriptions of knowledge, totally different types of sentences are used to express "generic" contents. Paradoxically, it turns out that the only unproblematic form, that is, "All X are Y," is virtually never used in practice. Similarly, the other forms referred to above are also used relatively rarely. Instead, constructions are employed that wear their genericity "written on their sleeves" even less than the above-mentioned forms. Accordingly, there are also other types of problems to be solved than the ambiguities discussed above.

One of these problems, for instance, is that in natural-language texts, certain verbs and other typical constructions are often used to express laws and interrelations for which it is likewise not obvious that they are to be interpreted generically. Examples of such verbs and constructions are

X is defined as Y,
X counts as Y,
X belongs to Y,
X is Y,
X entails Y,
X includes, etc.

Now, NLAs directly transforming such constructions into target representation language would not only fail to interpret a sentence such as

Armed robbery falls into the category of aggravated larceny.

as a generic sentence in the sense of a universally quantified conditional, but would also decompose it incorrectly, namely as "$\exists x\ \exists y$ armed robbery(x) $\wedge$ category of aggravated larceny(y) $\wedge$ fall into(x,y)." They would not only fail to capture the implicit "if-then" construction, but would also, departing from the surface structure, interpret the content in other than the appropriate way. What is required in order to understand such a sentence correctly is the complete abstraction from the surface structure. The auxiliary construction "falls into the category" must be, as it were, "cut out." The necessity of such a "cut" is particularly obvious in cases where the surface structure of a sentence, although embedded in an auxiliary construction, actually indicates a generic assertion, such as in case of the sentence

An accident constitutes a case of negligence if the driver has failed to exercise the care which the circumstances demand.

In this case, if one does not cut out the term "is a case of," one gets a reading along the lines of "$\forall x$ fail to exercise demanded care(x) $\rightarrow$ *is a case of*(negligence(x))" or something similar, instead of the intended reading "$\forall x$ fail to exercise demanded care(x) $\rightarrow$ act negligently(x)." An inference engine provided with such a piece of knowledge would fail to conclude that someone acts negligently, but instead would infer that there is *a case of* negligence, and this may come to a bad end if, for instance, only "act negligently" is stated in the antecedent of the next rule to be employed, and the proof is interrupted.

Perhaps this can be avoided by installing a semantic interpretation mechanism for the corresponding constructions which treats them as implicit "if-then" sentences from the beginning. In addition, and similarly to what is in part already being done for modal verbs and similar constructions, a mechanism could be installed which blocks the direct translation of surface constructions such as "is a case of" into the target representation language. But the deeper problem behind the difficulty illustrated is once again the lack of an unequivocal mapping from linguistic constructions to intended interpretations. Thus, one cannot simply install one of the above-mentioned interpretation mechanisms for constructions such as "X belongs to Y," or "X includes Y," because these constructions are sometimes used in contexts where they must not be interpreted as "if-then" sentences and where they should not be "cut out" (which is especially true of "X is Y").

But even if the NLA had succeeded in interpreting the constructions in an appropriate way, there is still another difficulty which is based on the problem of abstraction from the surface structure. This problem, again, concerns the identification of the definiendum of a rule, this time with regard to the abstraction from subordinate clause boundaries. In natural language, parts of the content of the "if"-part of a rule can easily be, syntactically as well as with regard to the surface structure, incorporated into the "then"-part, as has already been the case in the example above:

An accident constitutes a case of negligence if the driver has failed to exercise that care which the circumstances demand.

Concerning this sentence, an NLA guided by surface structure markers, even if it is able to resolve the auxiliary constructions "constitutes" and "case of," still is likely to offer the following complete translation:

$\forall x\ \forall y$ fail to exercise the demanded care(x)
$\wedge$ driver(x,y)$\rightarrow$accident(x,y) $\wedge$ ac negligently(x)

The NLA has failed to realize that abstraction from the surface structure of the sentence is again required, and that part of the "then"-sentence must be included in the "if"-part of the conditional. The definiendum of the rule

clearly is the predicate "act negligently," which would have to be identified and isolated accordingly. A correct translation of the above sentence would look like this:

$$\forall x\ \forall y\ \text{accident}(x,y) \wedge \text{driver}(x,y) \wedge \text{fail to exercise demanded care}(x) \rightarrow \text{act negligently}(x)$$

In this case, there is no syntactic hint whatsoever as to what to treat as the definiendum of a rule. Thus, sometimes even the correct identification of the definiendum of a rule expressed in natural language will depend on the overall context or reference to world knowledge.

This reveals the basic underlying problem behind the individual problems listed above: In each case an adequate reading can be determined only with recourse to world knowledge. This principle is valid not only for these examples of "real" knowledge presentations, but also for those simple sentences discussed under the heading of "genericity" in linguistics. However, the overall objective of natural-language knowledge acquisition originally consisted in communicating to the system its (entire) knowledge in natural language. If, however, a system requires prior knowledge in order to understand a rule formulated in natural language, all knowledge cannot be communicated via natural language. We later return to this problem.

Pragmatically false resolutions With the example of the interpretation of verbs such as "to include," we have already almost reached the border of pragmatics. For in one sense, a system which analyzes the sentences given as examples, for example the sentence containing the term "is a case of," in the way described, does so perfectly correctly. Thus, one can certainly view this as a case where a sentence has been correctly analyzed, but something has nevertheless gone wrong. This is perfectly in line with what we propose to regard as a pragmatically false resolution.

However, there are certain "more typical" cases of this sort of pragmatically false resolution which we want to discuss in this section, although it might be argued that they are more concerned with the later use of knowledge in proofs than with the problem of grasping the meaning of a natural-language expression. For instance, it might be required that certain details found in the natural-language expression are not to be treated as asserted facts but as requirements to test the validity of the respective fact. While in a narrative text,

He drove along a public street without driving permit,

the representation of "public street" may well be treated as an asserted fact, the generic correspondent

Driving without driving permit on public streets is . . .

may require excluding that a street has been temporarily blocked (for construction, car race, etc.) for the rest to apply. Now "public" must be represented as "prove that 'public' is consistent with all you know."

As another example, consider the metapredicates "consistent" and "inconsistent," which are to serve the above-mentioned purposes in our current legal expert system LEX1 (cf. [29–31]):

- **consistent(***predicatename(x, . . .)***)** serves the deliberate reversal of the burden of proof. In-depth legal analysis of those rules, where finally **consistent** was used, reveals that it is not appropriate to prove *predicatename* but to make sure that *not predicatename* was not provable. Therefore, **consistent** suspends the superordinate proof. It does not make *predicatename* part of that proof but initiates an intermediate proof, which aims at proving *not predicatename*. From the point of view of the superordinate proof, a **fail** of the intermediate proof is taken as a **success**. This corresponds to the *negation as failure* of extensions of logic.
- **inconsistent(***predicatename,partialtheory***)** also serves the purpose of controlling the proof procedure. It takes into account that insufficient knowledge of a domain (be it fundamental or due to the actual status of cognition) does not allow the formulation of a closed consistent theory. If, however, consistent partial theories for parts of the domain are possible, and if it is known in which partial theory the proof of some predicate might succeed, proof search can be delimited to that partial theory. This would at some time enable a formulation of the whole theory without the risk of *ex falso quodlibet*. This approach corresponds to the logical tree theories.

The rule

Street traffic is the traffic on public roads or squares, but not on special terrains such as ski slopes or railway tracks,

for example, is represented in the knowledge base by using the following metapredicate (here "u" and "e" represent objects and events, respectively, in accordance with the DRS notation introduced by [32] for the Discourse Representation Theory of [33]:

$$\forall u1\ \forall e1\ \textit{'traffic area'}(u1) \wedge \textit{consistent}(\neg\ \textit{'special terrain'}(u1)) \wedge \textit{'public'}(u1) \wedge \textit{'traffic'}(e1,u1) \rightarrow \textit{'street traffic'}(e1)$$

Obviously, an NLA that was to analyze the sentence above would never think of providing the above expression

in its representation language. This is not just because it is incapable of abstracting from the surface structure, as has previously been the case. Rather, it is above all because it can only provide one *sort* of expression of the representation language. The mere possibility of translating phrases such as "but not" or "all things being equal" into a metapredicate obviously never even occurs to it.

As we have mentioned when introducing such "pragmatic" problems, it may well be possible to solve them. Besides the notorious problem of having no really certain indicator as to when a particular reading and hence a change in the mode of translation is advisable, we also see that at the very least the decision to change the mode of analysis requires a new class of computations. Natural-language knowledge acquisition hence can best be employed to save work for the knowledge engineer in formalization within a given objective; letting an NLA itself decide at which point which mode of translation is most suitable does not seem to be recommendable at present.

Understanding a text: What one must be able to do
Apart from the individual content expressed by a sentence, what is relevant for comprehending a text, above all, is the correct resolution and interpretation of the relations between individual sentences. Are the sentences of a given sequence mutually relevant, or, to put it another way, does a relationship exist between them? This question is, in essence, the touchstone for whether or not a certain sequence constitutes a "text." The minimal criterion for "textuality" is generally held to be referential coherence, that is, repeated references to the same entity. Beyond this minimum requirement there exist other important coherence relations which transform a sequence of sentences into a coherent whole: temporal and spatial relations (of the facts described) and causal relations (between the facts described), to mention only two (cf. [34–36] for a compilation of different types of relations between sentences).

In the context of natural-language knowledge acquisition, it is vital to keep in mind that these relationships between the individual sentences of a text in themselves also constitute knowledge which is to be acquired. Consider the following sequence:

If one presses the brake pedal, the brake block is pushed against the brake disc. As a result the wheel can no longer rotate.

If this sequence is analyzed into its various propositions and these in turn are stored in the knowledge base, important information is lost. However, this is not all: Under certain circumstances totally senseless pieces of knowledge may result from treating a subordinate clause as an independent unit. A complete representation of the meaning of this sequence would have to represent not only the explicitly expressed facts but also—analogous to the example of pronoun resolution discussed earlier—contain an expression denoting the causal relationship between the two parts of the text above. (See Meyer [37] for a representation language that is capable of doing exactly this. Meyer, however, constructed this language for the description of text meanings with a human user in mind; i.e., she was not concerned with the problem of how to get from the structure of the text to the meaning representation in an automated manner.)

In contrast to the analysis of "generic" sentences, computational theories of discourse and text have been a more central topic in computational linguistics (cf., e.g., [38, 39]) as well as in the context of AI [40, 41] and the design of natural-language interfaces ([42]; for a review of these approaches see [43]), although it might be argued that none of these approaches provides us with workable algorithms and existing systems "really" capable of understanding discourse ([30], p. 115; [44], p. 170). In the following, however, we again do not treat all the general objections that could be raised against existing theories, but rather illustrate some of the typical problems in this field with a few (but rather "authentic") examples from knowledge presentations such as textbooks, and discuss the feasibility of some proposed solutions in the context of natural-language knowledge acquisition. Here, too, we wish to distinguish between semantic and pragmatic problems.

Uncovering and recognizing relations between propositions On the semantic level we confine ourselves chiefly to a discussion of two of the feats necessary for understanding text: the resolution of and differentiation between various formal relations and relationships between propositions, and the identification of the correct relata (reference points) of a particular relationship.

Since sentences are symbols that represent or express something other than themselves, the relations between sentences in a text may exist on two levels: The relationship may be between the denoted content of two sentences, or it may exist between the sentences as linguistic entities *per se*. It is in this sense that, for example, van Dijk ([45], Ch. 6) distinguishes between semantic and pragmatic connectives. In the sentence sequence

The old apple tree was rotten. As a result it didn't survive the storm.

a causal relationship exists between the facts which are the denotata of these two sentences. Compare

Don't put your feet up on the table. After all, it's my table.

Table 1 Possible combinations of functional relations with content relations for the connective "because."

Reference	*Expressing a cause-effect relationship*	*Expressing an inference relationship*
Expressed content	"The tree fell down because the wind was so strong."	"I'm color-blind because I can't distinguish green and red."
Performed speech act	"Your right of access is hereby withdrawn because certain allegations have been made against you."	"I forbid your putting on any function here, because I am the householder."

Here the relationship is between the two speech acts performed. If someone owns something, he may dispose of this object as he wills. The second proposition therefore establishes the right of the speaker to forbid the hearer's putting his feet up on the table, hence it functions as a justification of the speech act "forbidding" rather than of the content expressed as such.

The relationships that exist within the expressed content of sentences are referred to here as "content relations," whereas the relationships that exist between sentences (or the illocutionary roles they have) as linguistic entities *per se* are termed "functional relations." Obviously the example presented above ("justification of a speech act") is far from being the only kind of functional relation. Other important relationships subsumed under this category would be, for instance, those of "question–answer," "assertion–example in support of assertion," or the connectedness of larger pieces of text expressed by the relation "defining the problem–solution" (see [36] in particular, but also the literature cited on "relations between sentences" in general).

Even more diverse are the possible content relations which may exist within the expressed content of a sentence sequence. At this point, however, we wish to consider only two of these: first, the already mentioned causal connection, and second, the implication or inference relation. A relationship of the latter type is expressed by the following sentence:

The atmospheric pressure is dropping because the barometer is falling.

The fact that the barometer is falling is not in itself the cause for the drop in atmospheric pressure, but a sign, a piece of evidence, that this is so. This distinction corresponds to the old distinction between "ratio essendi" and "ratio cognoscendi" (which may perhaps be translated as "reasons for being" and "reasons for knowing"; see [46]), a distinction seen as increasingly significant within AI as well (see [47], for example).

Interestingly, these very different relationships may be denoted by one and the same word, that is, the connective used. The sentence with the barometer can easily be turned around:

The barometer is falling because the atmospheric pressure is dropping.

Now the relationship between the two expressed propositions is clearly one of cause and effect rather than inference (evidence and conclusion)—this despite the fact that the sentence has the same structure and is linked by the identical connective. It is possible to go further: One may freely combine the different content relations with various functional relations, so that, for example, the connective "because," also capable of expressing the functional relationship "justification," can be shown to have (at least) the four different meanings set out in **Table 1.**

Of course, relations between sentences can also be unmarked. In addition, one and the same relationship may be marked by quite different connectives. To sum up, it is not possible to assign linguistic markers unambiguously to particular relationships, nor conversely are particular relationships always signaled by linguistic markers (see also [48]). The presence of connectives may at best lead to the assumption that particular relationships do *not* apply, thus saving—from the processing point of view—certain checks (this holds for the system developed by Cohen [49]). However, the presence of a connective is neither sufficient nor necessary for the existence of a particular relationship between propositions. Analogous to the phenomena as they manifest themselves on the sentence level, it is not possible to decide what relationship exists between two sentences or the content expressed in them solely by means of linguistic markers.

The second problem which confronts a system trying to comprehend a natural-language text further illustrates this point. In most cases, the exact relata being linked by a given connective are not at all explicit. There is no direct equivalence between linguistically defined units and the units of meaning in a content relationship. An inference relationship marked by "therefore," for instance, may relate to a part of the preceding sentence, or to an implicit conclusion based on the preceding sentence, or it may indeed relate to the "essence" of the preceding paragraph as a whole. In the following sentence sequence, for instance, the first three sentences taken together function as the relatum for the content relationship of inference expressed in the last proposition:

Paul loves Mary. He knows that she loves chocolate. He is passing a shop where chocolate is on special. **Therefore** *he enters and buys ten bars of chocolate.*

In order to correctly identify the content relationship expressed by the connective "therefore," it is necessary to have access to knowledge about the domain (human motives and the ways in which they influence behavior) to which it appertains. Here, too, syntactic structure and lexicon entries in themselves are not sufficient to allow the text to be understood ([49], p. 15, presents a similar example).

To sum up: Comprehending a text, correctly identifying and differentiating the relationships existing between the facts described (again) presupposes the existence and accessibility of a large stock of world knowledge and knowledge about possible relationships. Accordingly, most of the approaches within the framework of AI or computational linguistics to the problem of text comprehension have been based on compiling a predetermined list of possible, formally defined relationships, the presence of which may then be ascertained by running specified checks [see, above all, [35, 42], who explicitly specify *lists* of relationships; the script approach by Schank and his students can be mentioned in this connection, since scripts likewise (though in a comparatively implicit way) contain possible relationships between propositions such as "means–end"]. In her system, Cohen [49] even employs a set of inference rules specified as to content, thereby enabling the system to "recognize" the existence of an "evidence relationship" (what we have referred to as an "inference relationship") between two propositions; that is, the system "recognizes" an evidence relationship when one proposition can be derived from another according to the stored inference rules. Applying such an approach to knowledge acquisition from natural-language texts would, however, lead to the very paradox described at the beginning of this chapter: To be in a position to correctly interpret the relations which exist within a particular domain and which are being expressed by the relationships between propositions appertaining to this domain, the system would have to possess substantial prior knowledge of this domain. One can imagine a situation wherein the text through which a system is to acquire knowledge about a particular domain is comprehensible only if the system already has at its disposal the very knowledge it is meant to acquire from the text. Obviously, new approaches and strategies for dealing with this problem are needed.

Pragmatics Having described in detail the feats of semantic sophistication required of an NLA in order to comprehend a natural-language text, we wish to add some brief remarks regarding pragmatic considerations.

Pragmatic problems on the textual level arise primarily because a text is normally produced by one human being for another human being. This is the topic of the following section; the examples given below serve as an introduction or transition.

When someone wants to explain something to another person by means of a text, he will, as a rule, not only describe the relevant domain but also include a commentary about how the text has been organized, about the purpose of particular sections of the text, etc. A typical example of such a commentary is the preceding paragraph on the purpose of *this* section. In addition, particular sets of background information about the domain to be represented will also be provided in order to "block off" any unwanted chain of thought that may suggest itself spontaneously to an intelligent, involved reader. An example of this kind of background information is the following comment: "XY is called Z." Although this choice of words is rather unfortunate, it has established itself.

Clearly, what is being conveyed by these comments is not information about the domain to be represented as such, but rather information about the knowledge being expressed about this domain and/or how this knowledge is being applied. What we are dealing with here is "meta-information," or "meta-knowledge." For human beings such meta-knowledge is extremely useful; from the point of view of optimizing comprehensibility it is, in fact, expressly required of a text. An expert system, on the other hand, when confronted by such information, can do little with it. It is of no use at all to the system, structured as it is, to incorporate such propositions into its knowledge base. Analogous to the pragmatic problems of resolution encountered on the sentence level, it is quite possible for these sentences to be correctly resolved semantically and yet for their meaning representation to "loiter about" quite uselessly in the knowledge base.

A very similar fate would probably await another kind of information which, like meta-information, is required as regards both text comprehensibility and the "quality" of text processing. We are, of course, referring to examples and other illustrative devices such as analogies. Examples and analogies occur spontaneously and relatively frequently in natural texts. It is a truism which holds not only for teachers, that when presenting new material to be learned, it is best to explain it by using examples; every human being seeking to elucidate something will do this by means of examples. For example (!) note that we, too, always gave examples of the linguistic phenomena discussed, and that the sentence you are reading at the moment is itself just such an example.

An NLA, however, would not be capable of dealing with such examples in the desired way. Instead of extrapolating the general rule they are meant to exemplify,

it would translate them into its representation language and then "wonder" what to do with them. [Novice learners, interestingly enough, often make the identical mistake: They take the example itself for the information to be learned rather than grasping what it is meant to illustrate (see [50, 51])]. But even if an NLA were able to recognize the difference between information to be stored and an illustrative example, it would be no better off. It would still not be able to process them, for we do not understand *how* we understand examples—for instance, how one recognizes the extent to which a particular case is an example of a general rule, or by what means one distinguishes the relevant from the irrelevant aspects of a given example. See the psychological literature on analogical and inductive reasoning, e.g., [52, 53].

However, even assuming it were possible to process meta-information and examples adequately, they would nonetheless be of questionable use to a knowledge base. In principle, an expert system already knows everything it needs to know when it knows the relevant rule. It is of no use to the system, in contrast to a human being, to encounter the same rule again, this time embodied in an example. Whereas in the case of pragmatic difficulties encountered on the sentence level (as we saw above), the sentences contain important information and therefore *ought* perhaps to be adequately processed, in the case of meta-information and illustrations it remains doubtful whether they provide an expert system with any useful information, i.e., whether the system should, in fact, use them at all. One could thus present a good case for the view that, since such information is of no use to an expert system, it is no great loss to the system if it does not understand it correctly.

This brings us finally to the issue of whether human-produced texts and expert systems are in fact compatible—whether natural-language texts are, in fact, at all suited to building up a knowledge base. As has been mentioned, information such as comments on organization, meta-commentary, and illustrations is certainly important and useful to *human* recipients. And since human producers of texts are guided by what is useful and important to their (human) counterparts, such information appears relatively frequently in texts. The inclusion of meta-information may thus be seen as a typical characteristic natural-language textual feature arising out of human communication which, while useful and important within the human context, becomes more of a hindrance than a help in the process of knowledge acquisition. The following section is concerned with precisely this issue.

• *Production problems: Problematic features of knowledge exposition as a result of the strategy of the speaker*

In dealing with the difficulties arising from the production of knowledge presentations, we proceed in the same way as we did before with regard to the problems relating to the reception of knowledge represented by natural language—we begin with the general features of the basic process of language production and thereby derive the resulting problems from this basic process. Consequently, our starting point for the following discussion is a model of language production which largely concurs with ideas psychologists have formed of that domain. Fundamental to this model is the assumption that language production is the transformation of a mentally represented content into "external" [i.e., audible or visible (written)] symbols or linguistic expressions. This process passes through many stages, the most important being the selection of what is to be verbalized, the syntactic, lexical, and (with speech) prosodic encoding of the contents selected, and finally their articulation (conversion into sound). This concept is shown in **Figure 4**, which again is a kind of "blow-up" of our basic model, this time with regard to the sender.

For our purposes, it is the first process (i.e., the selection of what is to be verbalized) which is of particular relevance. Here, the question is which part of the entire body of the speaker's knowledge the speaker will actually choose to express. For instance someone, if asked the way to Heidelberg Castle from where he was standing, would in fact, even if he activated all his knowledge on this subject, express only a certain part of this knowledge. He would say, for instance, "Opposite St. Peter's Church you must turn right"; but he would not, for example, further explain that at this point there is a road into which one can turn, and that one should use the indicator light and turn the steering wheel to the right (cf. [55], p. 124). However, one can assume that this is knowledge that the speaker already has when thinking about how to reach the castle.

There have been several attempts to discover the particular set of rules which govern the speaker's decision about which information from his stock of knowledge he chooses to verbalize. Grice's [56] is the best known of these attempts, whereas Herrmann [54] and his Mannheim co-workers were foremost in conducting empirical studies of this problem. In all of these approaches it is more or less explicitly assumed that in a given situation the speaker will only express those things that are *relevant* in that situation and not believed to be already known to the hearer (cf. also [55] and [57], Ch. 2). What exactly is and is not relevant does, of course, differ from situation to situation. For instance, the speaker may believe something to be irrelevant because he assumes that the hearer already has the appropriate facts, or he regards certain information as redundant.

Empirical investigations of this "principle of relevance" in language production so far exist for the fields of the choice of an expression to refer to an object and the choice of a formulation for requests (cf. [54] and [55], Ch. 4). To the best of our knowledge, there has been no

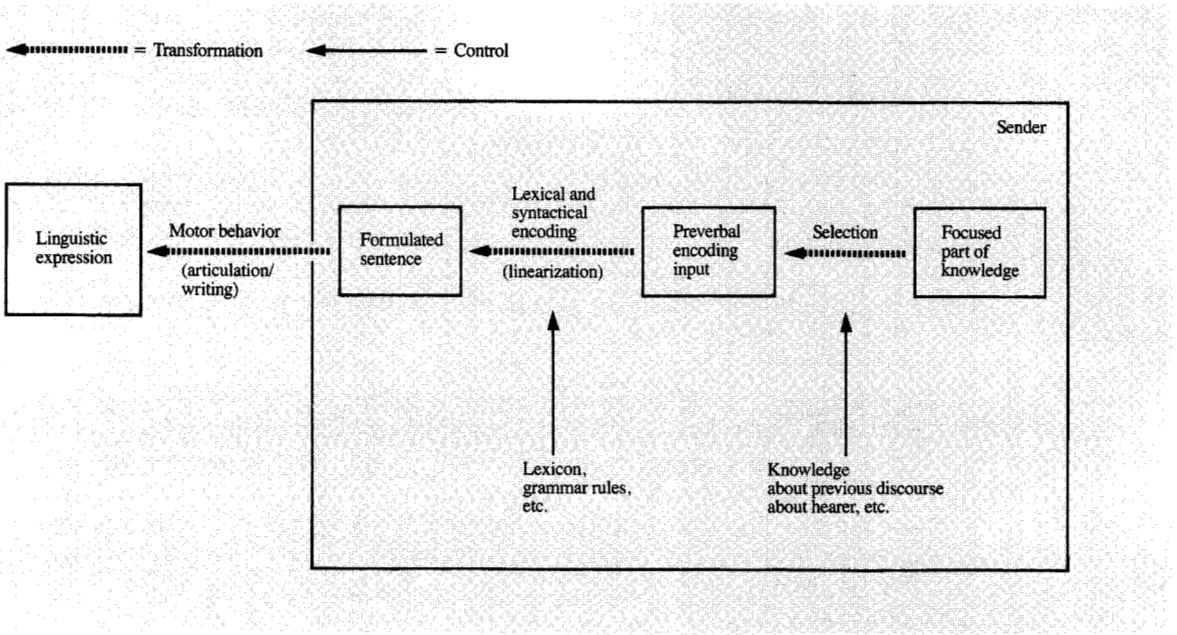

Figure 4

Basic model of language production, cf. [54, 55].

application of these results to the production of knowledge presentations. However, as far as the coordination of the respective demands of human language production and knowledge acquisition by an expert system is concerned, one can in our opinion nevertheless assume that there is a discrepancy between what appears to be relevant to a person and what is in fact relevant to the system to which he is trying to impart his knowledge. This discrepancy concerning respective principles of relevance between men and machines leads to two types of misunderstanding, which we have already mentioned in the introduction to this section on the problem of natural-language knowledge acquisition:

- The expert says something that is irrelevant for the system; that means that in certain respects he is saying *too much*.
- The expert doesn't mention certain things, which the system must, however, know; that means that in other respects he is saying *too little*.

If one now looks at "real" knowledge representations (and especially at those that have been provided in the context of natural-language knowledge acquisition) with an eye toward these two phenomena, one actually finds that the principle of relevance of language production does indeed have the expected effect. In the following, we would again like to illustrate with the aid of examples. Just as we differentiated between the levels of the sentence and the text in the discussion of the problems of reception, we would like to lay out the problems of these two levels separately. Inasmuch as both the "too much" and "too little" cases occur at both levels, we will correspondingly maintain a four-cornered discussion of the problems of production.

The sentences experts express

Our initial question was "What part of his knowledge will an expert express, and what part will he not express?" Part of the answer would be this: Above all, he would not express that which he does not regard as relevant. In relation to knowledge exposition, this means that knowledge which he assumes is well known (or at least easily inferable) will no longer be expressed. In the case of the example mentioned above of the man trying to find his way to the castle, the helper will not, for instance, say that there is a road somewhere if he has already stated that one should turn right there (and that implies that one *is able* to turn right there). The former piece of information is easily deducible from the latter.

In knowledge exposition, this selection process made by the expert is particularly applicable to one area, which is the large field of so-called (physical, psychological, etc.) "naive" (or common-sense) knowledge. Not only is naive knowledge assumed to be knowledge shared by all normal adults, it is furthermore virtually unheard of for such

information to be verbally passed on at any stage in someone's life. Naive knowledge is either innate knowledge (i.e., it is "hard-wired"), or it is simply picked up in the course of experience. An expert will then correspondingly follow the normal practice of speech and be loath to express naive knowledge explicitly.

On the sentence level, this "refusal" to articulate naive knowledge on the part of the expert has consequences with regard to "saying too much" as well as "saying too little." On the one hand, it can be seen that naive knowledge is simply omitted as being already known, easily deducible, or taken for granted. On the other hand, experts who are "forced" to articulate naive knowledge (for instance because they are to be employed in natural-language knowledge acquisition), as a result of following the principle of relevance, commonly express this knowledge in such a way as to fall into the category of "saying too much" from the standpoint of the system.

For instance, should one attempt to persuade an expert who has devoted some time to the study of naive physical knowledge, to impart this knowledge to a system in natural language, from our experience he will be likely to say something along the following lines:

If someone is in a smaller place, which itself is part of a larger place, **one can conclude** *that he is also in the larger place.*

Instead of the rule "If someone is in a smaller place, which itself is part of a larger place, he is also in the larger place," the expert here cites a rule in which he explains what **one can conclude** under which conditions. The reason for this is that it is obviously strange under normal circumstances to tell someone that anyone who finds himself in a small place also finds himself in the larger place which surrounds the smaller place. This is a piece of knowledge which one can assume every normal, socialized adult will somehow implicitly have at his disposal. Conversely, it is not so very strange to tell someone that under particular circumstances, one can draw this conclusion. This is a type of explicit knowledge which a normal adult does not yet have, but has only just discovered in connection with the formalization of naive physical knowledge.

Whereas the information about being able to conclude something may be relevant for people, for an expert system this is in a certain way more likely to be misleading. Because of its limited capabilities for resolving semantic and pragmatic ambiguities, which we described in the previous section, the system does not know what to make of this information. Because of this limited ability to resolve ambiguities, the system is not in a position to infer that the rule "If X then Y" is valid if it only knows that one is allowed to conclude "Y" from "X." Correspondingly, the system does not need to know that it can conclude something (no more than it needs to prove that something *is a case of X*; cf. "Semantically false resolutions" above). At this point the system just wants to be told the rule "If that and that is the case, then that and that is the case." And this is exactly what an expert is loath to express, because he follows the principle of relevance.

Now, this effect arises especially in the direct transmission of naive knowledge. In the exposition of other areas of knowledge, it is perhaps not so irrelevant for people to be informed of rules valid for these areas. But even here one can see how naive physical, naive psychological, or any other form of naive knowledge is really taken for granted and not explicitly mentioned. As a consequence, rules are expressed which are specifications of more general laws tailored to suit particular situations, and in this respect fulfill the condition of relevance; or rules are expressed in a form in which their content is condensed because background conditions which constitute naive knowledge are not mentioned.

If, for instance, a master mechanic were to explain the function of a brake to his apprentice, he would say those sentences which we introduced earlier as an example of the relations holding in sentence sequences:

If one presses the brake pedal, the brake block is pushed against the brake disc. As a result the wheel can no longer rotate.

At this stage we first consider the rule expressed in the sentence "If one presses the brake pedal, the brake block is pushed against the brake disc." In this rule it is apparently presupposed that the listener knows the spatial relationship of the brake pedal, brake block, and brake disc, so that the connections described can be effected in the first place. It is further assumed that the listener knows how, for instance by means of a brake pipe, the three things are linked to one another. Finally, the listener must also know that mechanical forces such as pressure can be applied to such a piece of equipment or via such a mechanism, and why it is that all this happens: namely, in order to bring a moving vehicle to a standstill.

Thus, it is not only naive physical knowledge of the sort "If one exerts pressure on a liquid which is encased in a pipe on one side of this pipe, this pressure will be transmitted to the other side of the pipe" that is presupposed in the explanation of the mechanics of the brake; against this background the interrelation stated shows itself further to be a particular, more specific version of more general interrelations for these special sorts of pipes and liquids.

Similarly, from this assumption of the "deeper" mechanisms behind a particular stated interrelation, the well-known case arises that in human inference rules there are always a number of presupposed "background

conditions" which would render the rules invalid when they are taken literally and exclusively. For example, owing to the fact that the precise chemical and physical mechanisms and the chemical and physical laws which facilitate the rule "If one scrapes a match against the striking surface, it will ignite" are taken for granted and therefore are in a sense "known," it is also clear for a human listener under which conditions the stated interrelations will no longer come into operation: for instance, if there is no oxygen, or if the match is wet, or if the striking surface is worn down.

Similarly, recognition of the difference between causal relationships and inference relationships which were touched upon in the last chapter is, in the final analysis, founded on the presupposed background conditions or "deeper" mechanisms and laws of which men are quite aware. Thus, such an assumed background knowledge is exactly what is needed to understand the relationship between sentences. We examine this problem again when we come to the text level.

To conclude: Since naive knowledge is presupposed in normal spoken communication, natural-language knowledge acquisition shows itself to be unsuited for the task of direct transmission of such knowledge. In addition, the representation of other areas, such as mechanical relationships, for instance, could possibly be detrimentally affected by this phenomenon, since assuming a deeper set of laws behind a certain systematic relation brings about typical "default rules" on the surface of the language. In our opinion, the fewest difficulties regarding the presupposition of naive knowledge and the relevance of stated rules in general arise in domains where "surface correlations" based on definitions apply. This is, for example, the case in the area of (German) law in which rules such as "negligence is the failure to exercise that care which the circumstances demand" are at the same time relevant and "free" from background conditions. (An inference engine in an expert system trying to prove this case would at most have the problem of how to determine that the person in question did indeed fail to exercise the care which the circumstances demand; this is, however, a different sort of difficulty from that occurring in a rule such as that cited for matches, which is only valid under certain circumstances.)

Properties of presupposed inference rules on the text level

In keeping with our four-way scheme of considering the problems at hand, we find that on the text level, too, there are difficulties with saying "too much" as well as "too little." However, we have already cited typical examples for the first sort of problem in the section on difficulties of pragmatically correct resolution: the quoting of examples, illustrations, the insertion of additional comments, etc. From the present point of view, all of this can be understood as information which may indeed be relevant to people but which appears to be irrelevant to an expert system. All in all, one can accordingly say that a considerable part of that which is expressed by people in natural-language texts is useless to an expert system.

Contrary to the corresponding problem on the sentence level, "saying too much" can be rectified relatively easily on the text level. If one does not attempt to use for knowledge acquisition texts which have been composed for other purposes, such as textbooks or instruction manuals, but instead lets the knowledge pass directly from an expert into the system, the expert will, with appropriate practice, certainly be able to adapt to the fact that his target is not a person and will therefore not want to know certain things. In this section, we therefore deal primarily with the problem of saying too little.

On the text level as well, this problem of saying too little pertains to the exposition of naive knowledge. Earlier, we emphasized the fact that, with regard to understanding text, what really matters is not only the comprehension of individual sentences but, above all, the recognition of relations between sentences or their content. Accordingly, on the text level, background knowledge will not only be assumed in the individual sentences (cf. the "brake" text above), but also with regard to the connections between sentences. Beyond this, there is sometimes no hint whatsoever as to any sort of connection that may exist between two sentences. In a sequence of sentences such as

John must still be at home. The light is still on.

it is not necessarily clear to a current NLA at first sight that the second sentence expresses the justification for the assertion made in the first.

> The full extent of the amount of implied knowledge which a speaker assumes the hearer to possess becomes apparent only if one tries to reconstruct the various conclusions which the hearer must have checked through before accepting that the second sentence forms the justification for the statement made in the first. Following Ryle [58] or Toulmin [59], the above sequence of sentences can be understood as an *argument*, whereby at least the assumption of an inference rule such as "If the light is still shining in someone's house, he will be at home" must be presupposed in order for the argument to be complete and at all acceptable. In full length the argument would therefore go something like this: "You accept exactly as I do that the following is valid: If the light is shining in someone's house, he will be at home; you can see that the light is still shining in John's house, so he must indeed be at home." Of the content stated in its entirety, however, in this example only a premise and the conclusion are

expressed. It is therefore not only assumed that the inference rule and above all its content are known, but, as is more important, it is assumed that in the two sentences stated we are dealing with an argument, and the listener must recognize that this is so or he will not understand. (Incidentally, as far as comprehensibility is concerned, the phenomenon discussed above of premise and conclusion being rather unsystematically arranged in the text manifests itself again: The premise appears not before but after the conclusion, and so the NLA is faced with the problem of deciding what to regard as the premise and what as the conclusion before being able to consider which inference rule is at work here and whether there is in fact a valid conclusion.)

We have, however, already covered the issue of discovering relationships, and since this section deals with production, we would prefer to comment on whether it would be better for the purposes of knowledge acquisition to express naive presupposed knowledge, as it is assumed in the inference rules mentioned, explicitly in simple sentences, or whether that knowledge should rather (with the help of further procedures) be "extracted" from the text. One could, for instance, in a way analogous to the procedure just presented, reconstruct the inference rules presupposed in a text by arranging the parts of an argument expressed on the surface (e.g., the premise and the conclusion) in the form of an "if-then" sentence (see [60] for a similar method). If inference rules were stated explicitly, one might be faced with the problem that in the first place, the inference rules so formulated would lead to the difficulties of pragmatically correct resolution which were described earlier, and second, that the naive knowledge might possibly be to some extent inaccessible to the expert. In contrast, texts appear at first sight to present the very promising possibility of ascertaining naive knowledge because it is present in a complete and "undistorted" form, albeit hidden and difficult to get at. So our question is this: Do inference rules which are presupposed as naive knowledge in the production of a text and can therefore be reconstructed from that text, have other properties (e.g., fewer background conditions) than the corresponding explicit rules? If so, should one not try to ascertain such inference rules through the examination of freely spoken, natural texts, rather than formulating these rules explicitly?

The answer: On the text level, everything is very much worse. If one considers the inference rules which are assumed in each of the arguments, one will see that the phenomena described above with regard to the sentence level, such as "situational specification" and the assumption of background conditions, present themselves much more distinctly on the text level. In the explanation of particular incidents, for example "The bottle of milk in the refrigerator is gone, because Paul took it away," the presupposed inference rule, as one would understand it at first sight, "If Paul takes the milk away, then it is gone," is in principle only valid for a certain person and for a certain object. It has actually been shown [60] how to get from such inference rules to the more general laws underneath (in the above example, for instance, this can be achieved by quantifying over people and objects so that one arrives at the rule "If someone takes something away, it will be gone") by applying techniques of machine learning. In principle, though, the highest possible level of abstraction, whereby a rule has the necessary degree of generality for it to retain an "argumentative force" while still being applicable to each of the example situations, is dependent on the particular domain under consideration. In the case of the rule "If one scrapes a match against a striking surface, it will ignite," for instance, one cannot just quantify over objects (as one cannot scrape just anything against just anything else and produce fire). In individual cases it is not at all easy to determine exactly the assumed inference rule that gives an argument its necessary general validity.

With regard to the background assumptions, the situation on the text level is more complex than on the sentence level, too. Whereas those conditions that are not regarded as noteworthy on the sentence level are, by and large, "normal conditions" whose absence really would be an exception (for instance if there were no oxygen present), in presupposed inference rules background conditions sometimes are not mentioned which actually represent further positive conditions for the validity of the rules in question. For instance, in the sentence "I fractured my arm because I fell," the inference rule "If someone falls, he will fracture his arm" is assumed. However, this inference rule not only takes for granted that there are *no* exceptions to the "normal course of events" (such as a fall on a mattress, for instance), but in addition also presupposes that there *are* exceptions to a "normal" fall (for instance, that the speaker has landed very badly). In principle the presupposed inference rule is, strictly speaking, false because only in a very few cases do people fracture an arm if they fall.

This difference between assuming the absence of normal conditions and assuming that there are further conditions which make something possible, which we have discussed here, can be demonstrated by a sort of "linguistic test": In the case of the rule relating to the matches and striking surface, one can insert the word "normally": "If one scrapes a match on a striking surface, normally it will ignite" is perfectly true. However, it would be odd to say "If someone falls, he normally will fracture his arm." Above all, the difference discussed seems to correspond to various concepts of "cause" in philosophy (cf. [46, 61]).

From all this, the following can be concluded: In natural-language texts, it is not only assumed that a connection between the sentences exists, but the inference rule used to enable these connections is not even expressed. And, as if that were not enough, these presupposed inference rules are themselves governed by such a multitude of background conditions that one would immediately consider them to be false if they were ever explicitly stated. Consequently, texts prove themselves to be even less suitable than individual sentences for the reconstruction of naive knowledge.

4. Strategies for achieving valid target representations

In the preceding section we have shown what kinds of problems a project of knowledge acquisition via natural language must face. These problems, however, all go back to the rather utopian aim of giving natural-language texts produced by human beings (e.g., textbooks) as they are (that is, containing examples and meta-comments, for instance), to a system which is expected to understand them and acquire knowledge. Thus, in reading these arguments, the reader might already have thought of more "realistic" objectives, and of strategies to achieve them. For instance, he or she might have considered the possibility that the existential reading of the definite article may seldom occur in textbooks, that is, in texts which are actually *supposed* to convey "generic facts," so one might assume that (at least a large number of) sentences in textbooks are intended to be generic. Hence, one might think of a system "understanding" a text from a textbook by employing something like a "rule of thumb," in that it takes every definite article for a generic one, thereby accepting the possibility of a few cases going wrong.

Another way of enhancing "understanding" between sender and receiver would be a kind of "tuning" on the side of the (human) sender. Human beings, one might think, are quite capable of *adapting* their behavior to the demands of the respective situation, so one might assume that they are able to suppress "unwanted" pieces of knowledge such as examples (or even restrict the syntax of the sentences they produce) when communicating with a system instead of a human counterpart.

Thus, besides the utopian demands of having the system understand every sentence of a text produced unconstrainedly, one can think, e.g., of more modest enterprises involving "rules of thumb" and "fault tolerance," or of drawing more on the competence of the human beings involved. In this chapter, we present a taxonomy of such strategies and the relaxed criteria they imply with respect to the validity of the target representations produced, as well as a discussion of the usability of systems relying on one of the strategies mentioned. In the next section, then, existing systems employing natural language for knowledge acquisition are analyzed with regard to these taxonomies, and strengths and weak points of the strategies are stressed. The discussion at the end of the paper summarizes these arguments and gives a general evaluation of each of the strategies and the conditions under which they can be employed.

- *Validity criteria of target representations*

There are many ways of demanding validity of a target representation, among them the criterion of controllability or predictability of the domain, or the requirement that the target representation be a true image of the domain, of the source-language sentences about the domain, etc. Since the latter class of criteria is the most natural one when dealing with language and knowledge as other representations, we discuss that one in some detail.

Qualities of images

True image, in the strict *mathematical* sense of an isomorphic image, can only be introduced between structures of equal cardinality. Since the image of the isomorphism is a finitely generated formal language, this precludes the existence of true images of open domains or bodies of knowledge. There can be no true image of a dynamic language either, as this strict concept of validity can only be operationalized for certain situations. Interestingly, in the case of a closed domain and a natural source language, the target representation can be a true image of the domain without being one of the source language (refer to our earlier discussion of "Source languages and domains," above).

On closer inspection, the capacity of a target representation to be an isomorphic image requires first that its algebraic structure be as rich as that of the origin, and second that all items of the origin (e.g., sentences of the natural source language) be mapped adequately on sentences in the target language. The former, of course, restricts or directs the choice concerning the target representation language.

Approximate images

Presumably, the most prominent and challenging case arises with open domains and natural source languages. Then the mediating source-language representation covers some aspect of the real domain and does not explicate a (possibly large) number of further aspects. (Theoretical arguments and practical examples for this have been given in earlier sections of this paper.) On the other hand, the formal target representation, because of its inherently limited expressiveness, can only arrive at a restricted or approximate coverage of the origin. While in the case of formal target representations the choice of a target representation language was a matter of epistemology—

How do we understand the origin?, it now has more the character of a strategic decision—*What do we want preserved when reducing from the complete expressiveness of the origin to the restricted expressiveness of the target representation?*

This reduction can have a topologic and an algebraic connotation. For the former, one needs distance measures between origin elements and methods of mapping neighboring origin elements to the same target sentence. This is not further outlined in this paper, but the whole class of fuzzy modifiers—very, extremely, etc.—are an obvious example where distance-related expressions are explicit in the syntax of a source language.

Algebraically approximate mapping has two major variants, projection and homomorphism.

Projection Projection nicely reflects the view [4] of a static language as the petrified description of one aspect of a domain, i.e., a true image in a strict sense of a restricted perspective (the one of the projection) on the field under study. Implementation of projection in the subsequent strategies may hence mean to take certain details as they are and to ignore others (e.g., formulations such as **one can conclude**).

Homomorphism Homomorphic images allow dimensions of the origin to be newly recombined, while the mapping as a whole loses some of the dimensionality or expressiveness of the origin. A nonlinguistic example will serve as an illustration: Given the body weight and height of a pupil, the weighted sum of the two values may be interpreted as one indicator of readiness to attend school. Two dimensions have been combined into one in a plausible manner. More precisely, for any pair of weights and heights, the indicator value is predictable; in this sense the weighted sum mapping is valid. However, from an indicator value we cannot derive weight nor height; the full dimensionality of the original knowledge is lost.

When returning to the linguistic context, we can use

A's car was damaged in an accident

as an example to demonstrate that "A's" may represent a "driver" or an "owner" relationship. "Driver" (with the background of traffic law in mind) and "owner" (with assets of the individual in mind) would both be projections, each of which suppresses the respective other aspect. A third alternative, the "keeper" relationship, expresses a third dimension which does not coincide with either of the above but portrays some of the character of both, namely that A in some way had the car at his disposal. In this case "keeper" is the homomorphic image combining the two possible interpretations of the original sentence into a less precise one, which, however, preserves some of the contents of the source.

"On-the-average" validity

While we have thus far treated validity as a trait of a whole system which uniformly applies or does not apply to all pairs of source expressions and their respective target representations, we now take the perspective that for a certain percentage x of source-language sentences, one of the forms of validity is achieved, and for the remaining $100 - x$ percent it is violated.

This looks like a probabilistic approach. It should be clear, however, that we do not imply the use of nondeterministic programs, which would transform a given source-language sentence into one target representation at some time and into a different one at a later time. Instead, we refer to systems that systematically get a (hopefully small) percentage of source sentences systematically and reproducibly wrong. On-the-average validity hence means that in taking independent samples from the "population" of source-language sentences, the probability of arriving at valid representations (in the sense of "qualities of images" discussed above) is a characteristic measure of the method.

An obvious application of this additional measure of validity is that a large number of source expressions may principally have a number of target representations, among which a (100%) valid distinction could be made only at a very high cost (e.g., in terms of size of common-sense knowledge bases). All but one of these target representations might, however, be either pathologic outcomes from the linguist's cabinet, or just results which are uncommon in a communicational situation. In these cases it might be a successful strategy not to care about the exceptions and to create the preferred reading. Such methods would transform validity on the average, but they would not have any fallback for exceptions; i.e., they would produce a certain number of invalid target representations without any technical means for detecting them as failures.

- *Usability criteria*

To the extent that humans are involved in the process of transforming source into target representations, the appropriateness of systems as tools becomes another important criterion. The full breadth of software ergonomics is fundamentally required here, which entails that we cannot expect precise measures except for experimental situations, but certain aspects are particularly prominent. They are related to the fact that the competence to communicate in a certain (natural) language and knowledge of linguistics, i.e., of abstract descriptions of phenomena underlying NLAs, are virtually uncorrelated. This implies that individuals, even though they can communicate fluently in a natural language, cannot automatically produce sentences or texts that correspond to any partial formalization of that specific language. Hence, systems either

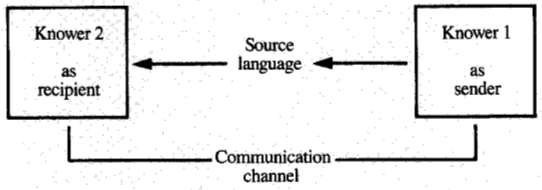

Figure 5

Starting points for strategies enhancing the communication process.

- must make the underlying linguistic structures obvious,
- must be used on the mere basis of competence in the respective natural language, i.e., must compensate for use outside their limits of formalization,
- must address users with sufficient linguistic knowledge, or
- must draw more upon contents than on linguistic knowledge.

Those systems that require human cooperation instead of just pursuing the utopian goal of "understanding" everything on their own may, e.g., be designed to teach the users about their behavior, to be fault-tolerant toward user misconceptions, to guide strictly by restricting the user to those inputs for which "nothing can go wrong," or to prompt for dialog that is dominated by domain contents rather than being linguistically intricate.

- *Concrete strategies*

We now present three different strategies for dealing with the problems outlined in the preceding section. The major distinction among the strategies is that difficulties can exist on the side of the sender of a law, on the side of a recipient, or in the communication process, and that hence we can enhance any one of the three. This is illustrated (**Figure 5**) by a refinement of Figure 1, which now focuses on the processes instead of the products.

Strategy 1: Human co-interpreter (enhancing the communication channel)

In this section we introduce two variants of a strategy that heavily involve a human *co-interpreter*. Structures implementing this strategy consist of a "naive" expert sender as knower 1, a "naive" natural-language analyzer (recipient) as knower 2, and a specifically trained human translator.

The expert is said to be naive in a situational sense; i.e., he produces his laws and explicit sentences of a natural language without making assumptions or speculations about the final technical recipient of his utterances. One might think of the expert using a dictaphone to make personal notes about a case or keying in a report for some external review board.

The natural-language analyzer is said to be naive in a pragmatic and (partly) semantic sense, on the basis of assumptions about existence and the validity of syntactic markers in the source-language sentences it receives, such as

- Syntactically declarative, object-level sentences being meant as such.
- Syntactically control-oriented sentences being meant as such and corresponding to well-defined control expressions in the internal language of the system.
- All premises, conclusions, generic or individual expressions, etc. being meant as such.

The role of the human co-interpreter in enhancing the communication channel can be twofold: He can paraphrase the expert's original source-language sentences such that they are "understood" by the NLA, or he can control and modify the results of automatic translation of authentic expert sentences translated by the NLA. The former role is subsequently called *translator*, the latter *controller* (see **Figure 6**).

To illustrate the effect of some of the validity criteria on this strategy, consider the case of a closed domain and a deviating source-language representation. The role of the human co-interpreter is to correct for the deviation, be it to paraphrase the source-language sentences or to control and correct the target representation sentences. Alternatively, consider the case of open domain and dynamic (natural) source language; here the co-interpreter can *project* in the desired direction by selecting those source sentences that bear contents of the domain and leaving out those that justify other sentences. Or he can find a concise homomorphic description, where the source allows for a variety of interpretations but does not give a cue as to which of them to prefer. This requires of the human co-interpreter

- The principal capacity to make distinctions about the pragmatics of sentences as outlined in "Pragmatically false resolutions," above.
- A sufficient understanding of the domain of expertise to classify elements from the expert's utterances according to the distinctions made in "Uncovering and recognizing relations between propositions." This includes, for example, recognizing which are causes and which are effects.
- A sufficient understanding of the formal semantics of sentences that the natural-language analyzer generates.

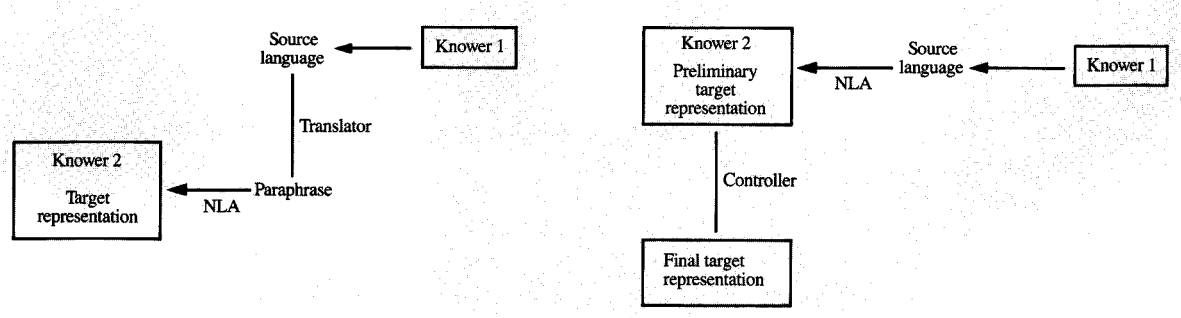

Figure 6

Two roles of a human co-interpreter.

While the former two represent reliance on the human-knowledge perspective [62, 63] to interpret the expert's verbal behavior, the latter implies mental control of its symbol-level representation.

One may ask whether a strategy which relies so heavily on the human co-interpreter is of any real use in the sense of either productivity or quality increase for the knowledge base that must be built. There are, however, good reasons to be optimistic about both. First, there is still much detail work to be done between the pragmatically appropriate translation of expert utterances and the formally correct symbol-level code. While the pragmatic translation tends to be done more easily by a human, the latter might be "easier" for a formal system. Hence, the productivity (and maybe also the quality) of the work of the "language engineer" can be increased. Furthermore, the results of the co-interpreter's translation process may be uniform, brittle, isolated, simply structured sentences, but they are still in a natural language and comprehensible to the expert. That is to say, even when all has been done except the final formalization, the expert still has a chance to check the correct interpretation of his utterances.

Strategy 2: Natural-language analyzer with common sense (enhancing the recipient)

Structures implementing this strategy include a naive expert sender and a highly sophisticated natural-language analyzer. The expert sender is naive in the sense of making no considerations about the use of his utterances. The sophistication of the natural-language analyzer may consist of additional preknowledge and/or additional processes (both indicated by "+" in "NLA+"), which are together referred to in **Figure 7** under the name *common sense* (CS).

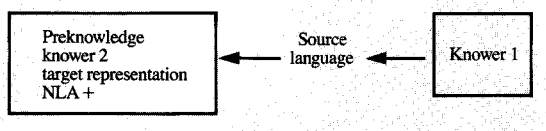

Figure 7

Basic model of Strategy 2: Natural-language analyzer with common sense.

The aim is to have the source sentences translated automatically into target sentences.[4] This requires the natural-language analyzer to make all detections of semantic and pragmatic failures. It goes without saying that here, as in the other strategies, we cannot proceed without a lexicon and a parser (providing intermediate NLA results), etc., which require expertise in their own right. We are dealing here, then, with what still remains to be done. Except for the closed domains with natural source language (where correction may be required), the major problem is indetermination or underdetermination of the intermediate NLA results. This can occur as ambiguous sentences (i.e., multiple parses), as sentences with unresolved references or other relations, or as other relations whose relata are not known. This gives all

[4] This definitely requires the target representation language to be predefined before the first sentence from the source starts being analyzed. In strategy 1 there is some choice to let the "language engineer" influence an evolving target representation language. According to [4], this is mandatory if knowledge base construction is understood as scientific discovery.

strategy 2 approaches the flavor of having to enrich the process, i.e., to introduce additional knowledge in some way. This is of course in accordance with the above analysis that production strategies are primarily *selection* strategies. Inasmuch as the sender is "naive" in the sense described, and the receiving NLA is rather sophisticated, strategy 2 approximates the "natural" situation of communication in natural language. This implies in particular that the NLA in the most optimal case would have at its disposal all that knowledge which a human sender assumes the hearer to possess and therefore does not express. For, as we have outlined, the knowledge a sender does not verbalize is exactly the knowledge which is required to understand natural language. A second implication of the resemblance between the scenario of strategy 2 and "natural" natural-language communication is that the world knowledge the NLA is supposed to possess must have been acquired in another way than via natural language. For, as we also have outlined, it is very difficult to convey common-sense knowledge via natural language. In addition, human beings do not acquire it in this manner, either. Thus, the resemblance between strategy 2 and human communication even extends to the origins of the involved knowledge: Just as common-sense knowledge is innate in humans or has been acquired through experience in the world, common-sense knowledge in an NLA+ must be either "programmed" or acquired using media other than natural language.

The subtypes of strategy 2 can be distinguished by the additional knowledge they introduce. For full functionality, the NLA+ would have to have all the knowledge used for the outlined diversity of phenomena. This almost naturally entails at the present state of cognition that any such systems can only be specializations. They can be specialized in knowing either about characteristics of the domain (subtype 1) or about the communicative situation and behavior (subtype 2).

Subtype 1: Basic knowledge about the domain For this type, the traditional school of Formal Theories of the Commonsense World [64] is central. This axiomatic, naive-physics kind of theory admits arguments about the time sequence, geometry, classical mechanics, fluid mechanics, etc. of entities introduced. Hence, possible causality, coexistence, and mutual exclusiveness of states can be concluded; i.e., when a relation such as causality is known, its relata or arguments can be uniquely determined from naive-physics (common-sense) knowledge. Where linguistic criteria do not allow distinction among the possible physical relations between states or events, common sense may discard implausible solutions (in an *a posteriori* disambiguation among a number of solutions). Or common sense may help to focus first on plausible solutions (in a more top-down-oriented process). Sentences such as *The leaves are falling because autumn is on the way* would finally be unambiguously understood as the climate influencing the vegetation and not vice versa. In this case the strategy clearly aims at satisfying a projection criterion: Knowledge about what may cause what allows this sentence to be understood in the perspective of physical causation and discards the direction of argumentation.

To give an idea of the limitations and risks of such an approach, we present an example that the reader may find pathologic in one way or another: Above, we have argued that climate influences vegetation, but in the long run, changes of vegetation, e.g. reduction of rain forest, may well change climate.

Nevertheless, different variants of knowledge about domains are briefly introduced. As additional knowledge we may assume that there exists an ontology of the natural-language concepts of a domain, which follows functional, topological, or other relations. A superconcept might, e.g., be required to perform all functions that its subconcepts perform, to satisfy all constraints that subconcepts satisfy. In other words, physical properties such as functionality and constraints can be inherited within the ontology.

Such bodies of knowledge—taxonomies or concept lattices—still bear a number of problems. They tend to be incomplete; i.e., there are common sub- or superconcepts for which a given language may have no name (Example: superconcept of snow = hail). Furthermore, whenever there are exceptions concerning some functionality or other property of some concept in the concept lattice, all the problems of default reasoning and nonmonotonicity can be expected to appear.

If we nevertheless assume that we have such a concept lattice, it can be used to distinguish whether a sentence is about terms from the concept lattice and is hence object level, or whether it is about other terms and hence presumably justificational or some other meta-comment. (Concerning the risks of such a naive procedure, cf. "Semantically false resolutions" above.)

If we encounter source-language expressions which might be specializations of intended rules (which we obviously do not know when facing the sentence but have to find out by some means), we can apply further knowledge about properties of the objects mentioned and whether they hold for superconcepts. The generalization aspect of the ignite-match example may illustrate that a common-sense knowledge base, whose concept lattice (or one of whose concept lattices) is organized according to the material properties, may arrive at the above-mentioned generalizations.

This, by the way, also corrects for an underdetermination of the source-language expression: The source rule leaves the degrees of freedom, whether or not it holds for

generalizations of its arguments. In the final intended rule this degree of freedom is removed.

Strategies that make use of basic domain knowledge are referred to as *domain CS* strategies, in contrast to the *communication CS* strategies introduced in the next section.

Subtype 2: Basic knowledge about communication
In contrast to subtype 1, where the deadlock of having to know *a priori* what one wants to acquire is omnipresent, subtype 2 ranges from general static linguistic knowledge about text coherence, stability of focus, etc., to methods of dynamically monitoring plans of the communication partner (e.g. [65]). On the one hand, since the predominant aim of knowledge representation is to convey facts, the variety of plans and elocutionary roles represented by the sentences occurring in knowledge acquisition is much less than in general communicative situations. Here knowledge acquisition lends itself to testing such strategies in that narrow range of plans. On the other hand it will turn out that approaches described so far deal only marginally with specific plan-recognition techniques. Thus, although subtype 2 seems to have some advantages at first sight, we do not discuss it further.

Strategy 3: Reduced-ambiguity source (enhancing the sender)
While in the two former strategies we allowed for a situationally naive sender and tried to enrich, project, etc. by means of knowledge and procedures in either communication channel or recipient, we now try to give the sender more information or knowledge, or to influence him to select in certain directed ways from what he knows. In those cases (**Figure 8**) we expect a *reduced source* to enter the process toward formal representation.

We introduce two possible addenda to the sender's knowledge in the two subtypes of strategy 3. Subtype 1 still leaves the expert naive about the recipient, but it creates communicational situations which enhance production of certain categories of generic sentences and suppress others. This can be understood as having influence on the level of the principles of relevance (cf. "Problems of production: Problematic features of knowledge exposition as a result of the strategy of the speaker"). Subtype 2 teaches the sender about what the technical recipient can understand and trains him to use only such formulations. Before going into the details of the two types, it should be noted that subtype 1 is in the realm of natural communication; i.e., it can be created by naturally appearing means and be brought to bear in knower 1 on a subconscious level. This may have the positive consequence that no interference between add-on knowledge and domain knowledge takes place, and hence authentic material can be captured. Subtype 2, however,

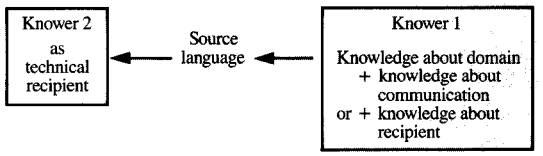

Figure 8

Basic model of Strategy 3: Enhancing the sender.

needs the conscious learning of formalization knowledge, which may be little related to domain knowledge but must be harmonized with domain knowledge. Interference will be unavoidable in subtype 2.

Subtype 1: Knowledge-elicitation methods
Verbal data are the major "source code" for both diagnosis of knowledge in cognitive psychology [66] and knowledge elicitation as a subactivity of knowledge engineering (cf. [67] or other textbooks on the acquisition aspects of knowledge engineering). The major forms of cognitively oriented diagnosis of knowledge, which bring forth verbal data, are "thinking-aloud" settings and probing. Thinking aloud is characterized primarily by influencing the subject very little in any way; authenticity and spontaneity of the pronounced source are the superordinate goal. (An example of how this can nevertheless be combined with direction concerning the useful outcome for machine use is presented in a later section.) Probing is comparatively unspecific about reducing variety in the emergent source language, except for the class of teaching back settings and structured interviews. While probing does not lend itself to knowledge acquisition, since it needs a precisely informed interviewer, thinking aloud is a good example of collecting utterances which are homogeneous in revealing sequences of individual inference steps.

In knowledge acquisition, we are aware of no systematic development of better techniques for eliciting machine-usable natural language from a source. Most approaches follow strategies 1 and 2.

Subtype 2: Preformal source language
This strategy leaves the technical recipient simple and the communication channel unmodified. It puts all of the burden of both meeting information requirements and producing a syntax that the NLA "understands" on knower 1, the domain expert. This may have a beneficial effect: Once

knower 1 knows what the NLA can understand, he is motivated to use only those formulations which match some aspect of the pragmatics supported by the target system. If, e.g., the NLA only accepts propositions, the expert will avoid imperative forms (and vice versa). However, this process of finding an acceptable formulation diverts attention from retrieving (or generating) the required contents of an item of knowledge. This may be because the technical nature of the recipient cannot be captured by subconscious "getting-used-to," but persists as a cognitive (and hence interfering) process. From other natural-language applications [68, 69] it has been observed that the user (here, knower 1) forms a model of the technical recipient which is more or less functional but also more or less wrong. This is in good accordance with observations by Norman about mental models, and is a good reason to be skeptical about the usefulness of the strategy. What makes it more critical is the fact that there remains a difference between syntactically *accepted* sentences, semantically *correct* ones, and pragmatically *adequate* ones (to the extent that such NLAs can deal with pragmatic variation at all). Feedback about correct syntax is immediate and hence efficient, whereas semantic correctness and the even more difficult pragmatic adequacy can normally not be assessed locally but must take the whole knowledge base, or at least significant partitions, into account. This is a demanding task on its own, and it can hardly be done interactively for all user entries. Inevitably the feedback concerning syntax overwhelms the feedback concerning semantics and pragmatics, which may lead to a systematic misconception of the NLA and the target system by knower 1.

One other aspect of this strategy is that knower 1 must approach the target language closely in his formulations. He must provide immediately mappable source sentences. This can be understood as the task of reducing his expressiveness to the expressiveness of the target language. If that is so, one questions whether there is any value in writing natural language instead of directly using the internal target language. The tendency of the answer is as in "Strategy 1: Human co-interpreter"; the NLA can still be of considerable help in creating well-formed expressions (e.g., correct arity of logic predicates), whereas humans do not have the respective formal information easily available [15].

5. Examples

For the purpose of analysis and clarification we have presented the three strategies as separate approaches to the use of natural language for immediate knowledge acquisition. In reality these three strategies and their variants are rarely found in pure form. Therefore, we now discuss four natural-language-based knowledge-acquisition methods, analyzing which strategies are present as elements of the methods and how they combine. We try to give some arguments about how the presence or absence of elements is related to strengths and weaknesses of a method with respect to validity and usability criteria.

Of course, all methods have in common that they start in some sense with a natural-language source, and some automatic component is involved to transform source into target.

- *Application software manuals*

Szpakowicz [70] presents an example in which software manuals are the source from which a representation of the knowledge of the respective domain, namely use of the software, is to be constructed. This is a closed domain (see "Closed and open domains," *ff.*), whose functionality is completely determined by the expressiveness of the specification of the piece of application software. The source language is natural (English). Szpakowicz does not identify special characteristics of the source language. In agreement with our assumptions, he excludes examples, i.e., episodic or incidental details, from the analysis. He observes that "software systems . . . (are an) example of . . . relatively uncomplicated domains" and that "manuals usually contain an almost complete account of the domain" if "(we may) presuppose certain minimal general knowledge of computing . . ." ([70], p. 35-2). A variant of conceptual graphs [71] is selected as target language, although the predicate logic properties of conceptual graphs do not seem to be used. In the basic expressions used in that language (such as "activity" and "object,") a purposeful choice is made in the sense of "Formal target representations" concerning semantics (and pragmatics) of the target language.

This indicates that Szpakowicz has thoroughly analyzed the scope of his approach. His observations can also be rather clearly positioned with respect to the strategies. A human co-interpreter acts as translator (cf. "Strategy 1: Human co-interpreter"), as indicated, e.g., by "The system is only an intelligent assistant to a person who builds a KB from text" ([70], p. 35-2). "If a sentence cannot be parsed . . . partial (analyses) . . . shown to the operator who . . . submits the sentence in a slightly altered form" ([70], p. 35-12) clearly describes a paraphrasing role for the involved human, which mainly aims at overcoming syntactic deficiencies and does not take the semantics and validity considerations of the former paragraph into account.

The approach also has elements of common sense. "The network must be initialized as the domain's skeletal representation . . ." ([70], p. 35-2) indicates the necessity of common sense, with the additional interest of the author "to determine what minimal knowledge is necessary . . ." not to operate "entirely controlled manually" ([70], p. 35-2).

Elements from the reduced-source strategy cannot be detected in the approach.

By the clear decision to have a human co-interpreter involved, "the parser becomes much smaller and easier to develop . . ." and "processes a majority of sentences from the tutorial part of the manual" ([70], p. 35-13).

Validity criteria are not discussed, except for the vague "(challenge) . . . to recover the meaning from the linguistic form" ([70], p. 35-2); i.e., there is no consideration (in the sense of "Qualities of images" or "Approximate images") as to whether misconceptions in the text must be corrected for, or whether the didactic intention of the writer of the manual must be separated (e.g., in the form of a projection) from domain contents.

Responsibility for such content-related decisions is with the knowledge engineer. This is in good accordance with the usability criteria we have described, which include easy access to all the forms of basic semantic knowledge required to disambiguate. However, the NLA also needs the language engineer to paraphrase sentences that cannot be parsed. This requires the language engineer to have an appropriate (mental) model of the linguistic structure of the NLA, which contradicts one of the usability criteria, unless there is either strong guidance or good teaching facilities.

- *Explanatory utterances*

In the example of understanding student curricula, as in [72], explanations provided by using situationally naive speakers are used as a natural-language source for automatic knowledge acquisition. Since there are also examples from another domain, and no considerations about the nature of the domain are presented, we conclude that the nature of the domain does not determine the approach. The source language and, partly, its production process are among the most emphasized aspects. "User-supplied natural-language explanations provide . . . information to construct . . . knowledge-base . . . . The theory of acquisition . . . (relates) individual utterance "types" to unique acquisition process . . ." ([72], p. 20-0). Utterance types are classified semantically ("definitional," "entity-tagging," etc. [72], p. 20-3), with the tacit assumption that all produced utterances relate to subject-matter contents and not to preceding or subsequent speech acts (cf. "Uncovering and recognizing relations between propositions"). The utterance "types" are classified semantically, without concern for the possible presence of syntactic markers (cf. "Semantically false resolutions"). In effect, the respective "world knowledge pertaining to the acquisition process" ([72], p. 20-5) draws upon a preprocess, which makes the necessary categorizations.

The target language is only marginally described in terms of the processes that generate it: ". . . rule-based implementation . . . at a 'conceptual' level . . ." ([72], p. 20-4); we cannot determine how the target language was selected, nor can we judge whether it is adequate in the long run.

The neglect of domain characteristics renders validity difficult. As a matter of fact, this aspect is not dealt with. A ". . . one-to-one relationship between the (utterance types) and the KB-rules" is explicitly assumed ([72], p. 20-8). As to the strategies, there is no indication of involving a human co-interpreter nor a reduced source: The "effort focuses upon . . . (a system) capable of automatically constructing and extending KBs" ([72], p. 20-0). The distinction of utterance types can be understood as communication CS (communication common sense—cf. "Subtype 2: Basic knowledge about communication"), although its syntactic basis is lacking in the system.

In summary, the approach relies heavily on the possibility of being able to distinguish utterance types. Our analysis has made it obvious that there is little chance to make such distinctions automatically by means of syntactic markers. Since, on the other hand, there is no report of either attempting to provide homogeneous source streams or of using a human co-interpreter, it is not surprising that later developments of the approach have not been reported. At best, it can be expected that, given syntactic indicators for the utterance types which apply in a majority of cases, the system achieves an acceptable "on-the-average" validity, be it in a homomorphic or an approximative sense. Of course, the system rates high on usability, in the trivial sense that a user is not required.

- *Rule acquisition*

A patent application [73] for an automated rule-acquisition system is outlined in the example of electric circuit troubleshooting. The aim is to generate the internal production-rule representations automatically from expert, word-by-word formulations and expert pointing to elements of schematic drawings.

In contrast to our analysis that this is an artificially closed domain, the authors of the application claim that their "dual medium . . . combining diagrams . . . with a restricted formal language" ([73], p. 15) method can be transferred to "other subject matter for which a set of rules may be formulated" ([73], p. 5). The likely infeasibility of this generalization becomes obvious when we look more closely at the role of the diagrams: They provide essential aspects of the semantics of the rules expressed in (quasi-)natural language (see below). This makes sense only when the objects in the diagram have a precise meaning; this seems to be the case in electrical engineering, but can be doubted even in other technical domains such as mechanical engineering. The source language used is "English-like syntax at the surface level, . . . semantics serves as task model" ([73], p. 15.); i.e., we find an instance of an only seemingly natural dynamic language, which really is static (cf. "Source

languages and domains"). The operations of mapping from knowledge or domain into the final representation [projection, homomorphism in the case of open domains (cf. "Qualities of images"), or corrections in the case of intrigued primary formulations (cf. "Source languages and domains")] take place before the process of analyzing the source language; they must be performed during its production. The selection of the target language has been done by the global argument of the seventies and early eighties that production rules represent modeling expertise. However, there have not yet been any individual or specific demonstrations that production rules lead to adequate expressiveness for describing electronic device troubleshooting (or other troubleshooting).

The approach distinctly denies co-interpreter elements: ". . . obtain . . . rules from the expert . . . in an automatic or semiautomatic manner, without ambiguity, . . . without assistance of a human 'knowledge engineer' " ([73], pp. 14–15). It obviously implements a preformed source strategy: Only Englishlike formulations of production rules are accepted, and these must be presented word by word. There is, however, some support of formulation: "If the word does not match . . . a menu . . . all possible next steps . . . presented . . ." ([73], p. 3); "to support menu-based acquisition, . . . expectation table . . . to determine which words are syntactically admissible . . ." ([73], p. 16). This draws upon domain CS about what domain *concepts* and *relations* ["*voltage*" "*of*" (a component to ground) or "*between*" (two components) ([73], pp. 16–17)] actually are in accordance with common sense of the domain. This is both terminological and physical (cf. "Strategy 2: Natural-language analyzer with common sense"), but not communicational, since pragmatic variation is not tolerated in any sense.

Considering that feedback about formulation is provided to knower 1 on the basis of *semantic* criteria,[5] the usability disadvantages of preformal source subtype 2 outlined in "Strategy 3: Reduced-ambiguity source" seem not so severe: Feedback is as good and as specific as the available common-sense knowledge about the domain providing it. The user still needs to develop his model of the system, but the model need not be based on linguistics, which is outside his domain of expertise, but on how the semantics of his domain of expertise is represented in the common sense of the system. In other words, inasmuch as production rules are the right choice as target representation, and common-sense knowledge about admissible concepts and their relations is available and appropriately represented before using the system, the approach provides both guidance and feedback to the individual expert for formulating his rules and true image validation by enforcing correction of any source expressions which mismatch the structures of the domain. Validity hence draws more upon the common sense of the system, to accept or reject expressions, than upon the concrete translation process. More than other approaches, this one demonstrates the dilemma outlined in "Problems of language processing"—that knowledge is required in the system in order to allow acquisition of knowledge. Because this premise—availability of the required domain CS—and the one of sufficient expressiveness of production rules are unrealistic except for very simple domains, this characterizes the limited scope of the patent and supports our doubts concerning extension to domains other than circuit troubleshooting.

Above we have argued that, under the highly restrictive assumption of an appropriate common-sense model, the system may well be usable because of the dominance of this domain model over linguistic intricacies. However, usability might be hindered by superimposing the word-by-word segmentation process upon the production processes of the expert, which can be assumed to be based on larger units (such as propositions).

- *Training natural language for high-end target languages*

In the environment of the LEX project [29, 31], a natural-language knowledge-acquisition tool KALEX [15] has been developed. Its hereditary domain is the open one of traffic scenes; in a small additional evaluation experiment, the recommendation of data processing equipment for small companies could also be dealt with.

In this system the domain expert user enters a (quasi-natural, see below) German "rule sentence and requests the system . . . to translate it into a DRS"[6] ([15], p. 241). The target language has been chosen to have the highest expressiveness that remains computationally tractable, i.e., to satisfy a general and not a domain-specific criterion. The notation in DRS forms efficiently supports nonlogical connotations and reflects the discourse representation theory [33].

The restricted nature of DRSs nevertheless limits the source language. We deal with a preformal source strategy ("Subtype 2: Preformal source language") in which the expert must learn what subset of his natural language the system will accept. In KALEX this of course also has a syntactic aspect: "In the case of unrecognized syntax . . . (the user) can derive from a list of related formulations, how he/she might modify" ([15], p. 241). This already indicates the user interface strategy: "The user is efficiently supported in learning" ([15], p. 247) the source-language segment he can use. This includes semantics and pragmatics. For the semantics he "always (sees) a German sentence and a (system-provided) resulting DRS in neighboring windows . . ." ([15], p. 247);

[5] This, by the way, means that the patent applicants' formulation "which words are *syntactically* admissible" is a misconception of their own approach.

[6] Special form of first-order predicate logic formula, cf. "Pragmatically false resolutions."

i.e., he receives immediate feedback about associated patterns of the two languages. At any time he can "interactive(ly) test" a newly added item (DRS) or a "temporary version" of it ([15], p. 245), together with the consolidated old knowledge base; i.e., he can make sure that the formal pragmatics[7] are adequate. KALEX does not yet match specific source-language patterns with metapredicates available to the KALEX theorem prover, such as consistent, inconsistent (cf. "Pragmatically false resolutions"). We see the whole approach as an attempt to minimize the disadvantages of preformed source strategies concerning the system feedback and resulting expert behavior. With respect to validity, KALEX can be characterized as enabling the expert to actively ratify the authentic knowledge base contents. Usability and validity criteria are highly interrelated in KALEX. Valid translation is under the permanent control of the user and not so much a property of the system. However, the capacity of exhibiting validity control requires that the user acquire a model of all aspects of the linguistics underlying KALEX (syntax, semantics, pragmatics). The continuous process of seeing natural language and DRS formulation next to each other and the possibility of interactive testing supports him in understanding the meaning of all knowledge-base entries. It is up to him to generate either projections or homomorphic reductions of the domain.

- *Partial solutions for subtasks*

In the growing field of using natural language for knowledge acquisition with its present premature status, interesting insights may also be gained from partial solutions of isolated subtasks within the presented framework.

Modeling communicative behavior

In [74] we encounter an approach in which the central part of translating natural source (about university administration) into formal representation (framelike) is still done manually. The aspect to be discussed here is the way the system operates to integrate the manually produced frames with the intermediate state of the knowledge base. The system $K_{A^c}^{n}$ derives expectations (i.e., anticipates modifications) from . . . the state of the existing knowledge base, from cues in the discourse, from previous modifications, . . . or from the state of the knowledge acquisition task" ([74], p. 14-2; the latter is not yet detailed in the further outline of the text). Typical heuristics related to the state of the knowledge base are to expect details about underspecified objects (missing slot fillers). This deals with the present state of the knowledge base. This most resembles a communication CS strategy, although it cannot be fully subsumed under either of the alternatives—basic knowledge about the domain or basic knowledge about communicative behavior. Of course domain CS plays a role in identifying missing specifications, but the expectation that full specifications can be elicited whenever partial ones exist depends more on (planned) treatment of a complex subject matter than on (spontaneous) communicative behavior.

The second type of heuristics introduces the exploitation of coarse dialog behavior and coherence patterns to generate expectations. Details or context information, for example, are expected to follow the introduction of a term.

The third type combines the first two by introducing expectations about broader dialog coherence postulates, i.e., that neighboring concepts are candidates to be dealt with in a discourse unit. This of course requires both types of common sense, the knowledge-base-related assessment of neighborhood and the dialog-related assessment of slight modification of focus.

In conclusion, the described common-sense module might be a valuable part of an NLA+. At present the module assumes that all of its input is object-level information. Justificational or inference rule formulations are not mentioned. It may well be that they are filtered out by the human supplying the input to $K_{A^c}^{n}$. If the module becomes part of a full-fledged common-sense-based system, further provisions will have to be made.

Validity measures are not explicitly discussed. The method of having the characteristics of the knowledge base and dialog determine the interpretations has the effect of projecting onto selected axes in a dynamic way: Given one state of dialog or knowledge, an input is selectively interpreted. In a different situation, other aspects of the same input may turn out to be selected.

The authors themselves speak of heuristics. This makes it clear that they are aiming at on-the-average performance and are aware of singular failures of their situation-dependent interpretations.

Focusing natural-language utterances

One more of our own results in the scope of this text deals with tuning an established knowledge-acquisition method—thinking aloud—toward the requirements for automatically transforming the results into a formal representation. In [60] the central question is how the think-aloud setting can be varied so as to increase the usefulness of the utterances for direct use in knowledge-base construction. No attempts have been made with the recorded transcripts to use the LEX natural-language analyzer [31] for automatic generation of knowledge-base entries. The reasons will become obvious from the following attempt to present [60] an in-depth analysis of an instance of an elicitation strategy.

[7] I.e., the intended effect to the knowledge base, as opposed to the pragmatic variation in the source, which has been discussed more intensely in this text, cf. "Pragmatically false resolutions." Only one aspect of pragmatic variation in the sources, namely distinguishing the justificational and the causal "because," is currently in work.

The contents of the data raised from 60 subjects in a controlled psychological experiment was "situational (knowledge) . . . goals and motives . . . temporal and causal connections" . . . "required in addition to linguistic knowledge for understanding a text" (about traffic scenes) ([60], pp. 33-4–33-5). This hard, open domain was selected to draw upon common sense that "almost everyone has" ([60], p. 33-5) such that subjects could easily be found. The different think-aloud settings were variants of sequentially presenting segmented narrative text about traffic accidents to the subjects and asking them to express the knowledge they used to understand each segment.

The material of about 7500 utterances still contains all the "noise" to be expected according to "How generic sentences can be misunderstood" (justificational, episodes, etc.). But the percentage of ultimately useful utterances could be raised from about 27% to about 80% by appropriate adjustment of the segmentation length when presenting the text and by providing some additional supply of keywords recommended to the subjects for use in their verbalizations (with consideration of the need to define temporal, causal, etc. items). It could be shown by nonincreased reaction times that the resulting "useful" verbalizations were still authentic traces of the subjects' natural considerations, not additional rationalizations.

As to the contents, most verbalizations were conclusions drawn, not "rules" used for arriving at the conclusions ([60], p. 33-9). This is apparent because most of the communicated conclusions could readily be combined with respective segments from the presented text, supplying premises for the applied rules.

This can be taken as an argument that knowledge-elicitation methods can be tuned for the special needs of providing useful verbal data for further technical use. The quality of data can be expected to be even higher in real domains of expertise ([60], p. 33-13) such as law, where the percentage of available source-language formulations is higher than in common sense (cf. "Strategy 2: Natural-language analyzer with common sense") because of the high degree of "code" in law. A further glance at the concrete contents of the utterances suggests combination with other strategies. One regular observation was that subjects did not supply the right abstractions of rules (as shown by their conclusions) but "incomplete copies of their instances" ([60], p. 33-10; see also "Properties of presupposed inference rules on the text level" in the present paper). "An object hierarchy . . . may be used . . . for determining the appropriate level of generalization" ([60], p. 33-10) addresses the usefulness of a domain CS component in partially curing this problem. Resulting rules would then be "overgeneral" ([60], p. 33-10) in the sense of "Problems of production: Problematic features of knowledge exposition as a result of the strategy of the speaker," a condition which might be approached by strategy 2, subtype 1.

Since the whole approach deals only with production on the side of knower 1 and does not address the transformation into formal representation, validity issues can only be treated with a restricted focus, namely whether the recorded utterances are a valid image of the knowledge in knower 1. There are good arguments that they *are*, in a projection sense; utterances in the direction of the keywords are enforced, while others are (indirectly) suppressed. The approach favors selection instead of postrationalization, as proved by nonincreased reaction times in the "keyword" condition. We found no indication that any specific percentage of the utterances were invalid. More typically, there still were 20% nonobject-level utterances, but they were not false.

Usability has two facets in such a procedure. On the side of knower 1, all our observations indicate that he is little hindered by adhering to the setting that we created. With respect to further use, one must, however, admit that transcription and encoding through an NLA are still labor-intensive.

6. Discussion

It has not been a goal of this paper to arrive at computationally efficient representations of natural-language source material, nor to suggest computationally optimal NLAs. As should be obvious from the text, there are enough unsolved problems in the use of natural language for knowledge acquisition that some basic clarifications should precede fast technical solutions for the wrong problems. We hope that this text helps with some of the clarifications.

The outlined problems and strategies seem not to be language-specific. From the experience of writing this text, whose first draft was in German, and from comparing our experiences with an NLA for German with those of others in English, we can conclude that most of the phenomena described in this text occur similarly at least in these two languages. The only aspect where this is not so obvious is that of syntactic markers of genericity. Their nature and consistency varied between the two languages. It might turn out that some other natural language, in contrast to the two we have studied, does have consistent syntactic markers of genericity, although [75] seems to indicate the opposite, at least with respect to tense-aspect marking. Since the lack of such markers caused a considerable part of the problems we encountered, such a language might lend itself more easily to knowledge acquisition.

The theoretical analysis, the collection of examples of the variety of linguistic phenomena, and the (more or less) operational examples of natural-language workplaces for knowledge acquisition all indicate that strategies 2 (common sense) and 3 (reduced source) tend to be

preferred. This is in some contrast to their identified disadvantages. It might be asked why strategy 1, subtype 1 (human co-interpreter, translator) is so little used. It seems to satisfy a number of advantages:

1. It uses skills where they naturally occur.
   - Knower 1, i.e., the domain expert, is responsible for adequate knowledge items.
   - The human co-interpreter is responsible for paraphrases of source expressions, which preserve the information contents but are acceptable for an NLA.
2. It has as one of its parts a natural-language formulation which is both
   - A mandatory final document of the expert, which is conceivable for him and which he can ratify and update.
   - Input to the NLA with predictable unambiguous resulting behavior of the target system.
3. It introduces a human mediator into a highly complicated process.

It should be noted that after the first enthusiastic attempts to cover all knowledge-acquisition tasks either by standalone automated tools for knowledge extraction or by standalone machine-learning algorithms, cooperative approaches have now become common in most schools of knowledge acquisition. In this context it might be worthwhile thinking about a discipline called *language engineering*, which has a role similar to that of knowledge engineering but concentrates on domains or skills, where major aspects of the knowledge are naturally available in natural language.

If for some reason cooperative strategies cannot be used, one might consider how the other strategies can be improved, and to what fields of application they might lend themselves. In general, reduced-source strategies may be recommended where the full source has already developed stereotypes of its own, i.e., where natural communication is based on highly standardized syntax and semantics. Common sense is easier to handle the closer it is to linguistic knowledge. When knowledge about concepts and their relations is already available in an NLA, and need only be enhanced by additional relations or by additional relata in an existing relation, the upgrade does not require fully new data structures and inferences, as would be the case when requiring a causal reasoner to cooperate with a separate terminological reasoner. Some domains in law are of the former type, while all natural-science-based domains obviously require causal knowledge. The outcome of strategies 2 and 3 might be improved by two coherent decisions:

- To use such target languages, which have proved useful for knowledge *acquisition* (e.g., KADS, cf. [76, 77]) and not merely for knowledge *representation*.
- To require the system to have common sense about this method, so as to allow an assessment of pragmatically adequate processing of source-language expressions.

In the case of KADS, for example, this would mean that attempts are made to recognize a source fragment as a static domain entity, an elementary inference step, or a larger inference unit (a named task), because these are the major elements of the KADS structure. Given the quantity and variety of problems of the task as a whole, all fully automatic common-sense-based strategies will have to compromise. They will have to assume standard meaning when applying domain CS and standard behavior when applying communication CS, and will produce errors in nonstandard cases. Their validity will be bound to be "on the average."

Given the mismatch between competence in language use and linguistic knowledge, all strategies will have to be complemented in one way or another by teaching facilities.

Concerning time scales for attempts to use natural language in complex human–computer interface tasks, rapid success cannot be expected. From our experience of using a predecessor of the LEX1 NLA for natural-language access to database [78], we know that it took more than a decade from the beginning of the research to the product announcement. One reason for the long time requirements is that, besides conceptual problems such as the ones outlined in this text, large amounts of data must be present before systems can start to work. One might think, e.g., of having to classify all verbs of a language in order to disambiguate NPs as to their genericity.

On the other hand, a database-access product has been announced (IBM SAA™ LanguageAccess), and a patent has been applied for and has just recently been approved [73], indicating that natural language for increasingly complex tasks at the computer is slowly migrating from research toward technology.

Acknowledgment

H. Lehmann and P. Bosch commented on an earlier draft of this text. I. Seidel and W. Schmidt continuously supported us technically in finishing this manuscript. Thanks to them all.

References

1. S. Regoczei and G. Hirst, "On 'Extracting Knowledge from Text': Modelling the Architecture of Language Users," *Proceedings of the Third European Workshop on Knowledge Acquisition for Knowledge-Based Systems (EKAW-89)*, J. Boose, B. Gaines, and J. G. Ganascia, Eds., Paris, July 1989.

2. T. Winograd and F. Flores, *Understanding Computers and Cognition. A New Foundation for Design*, Ablex, Norwood, NJ, 1986.
3. R. Nüse, N. Groeben, B. Freitag, and M. Schreier, *Über die Erfindung/en des Radikalen Konstruktivismus. Kritische Gegenargumente aus psychologischer Sicht*, Deutscher Studien Verlag, Weinheim, Germany, 1991.
4. W. Heisenberg, *Ordnung der Wirklichkeit*, Piper, Munich, Germany, 1989.
5. Th. Wetter, "Common Sense Knowledge in Expert Systems," *Expert Systems and Expert Judgement*, J. Mumpower, L. D. Phillips, O. Renn, and V. R. R. Uppuluri, Eds., *Proceedings of the NATO Advanced Research Workshop*, Porto, Portugal, August 1986; Springer, New York, 1987.
6. A. S. Elstein, L. S. Shulman, and S. A. Sprafka, *Medical Problem Solving. An Analysis of Clinical Reasoning*, Harvard University Press, Cambridge, MA, 1978.
7. G. N. Carlson, "Exceptions to Generic Generalizations," *Mathematics of Language*, A. Manaster-Ramer, Ed., Benjamin, Philadelphia, 1987.
8. G. Heyer, "Generic Descriptions, Default Reasoning, and Typicality," *Theoret. Linguist.* **12**, 33–72 (1985).
9. G. Heyer, "Semantics and Knowledge Representation in the Analysis of Generic Descriptions," *J. Semantics* **7**, 93–110 (1990).
10. L. K. Schubert and F. J. Pelletier, "Problems in the Representation of the Logical Form of Generics, Plurals, and Mass Nouns," *New Directions in Semantics*, E. LePore, Ed., Academic Press, London, 1987, pp. 385–451.
11. B. Geurts, "The Representation of Generic Knowledge," *Genericity in Natural Language*, M. Krifka, Ed., *Proceedings of the 1988 Tübingen Conference*, University of Tübingen, SNS-Bericht 88-42, Tübingen, Germany, 1988.
12. R. Declerck, "The Manifold Interpretations of Generic Sentences," *Lingua* **68**, 149–181 (1986).
13. M. Krifka, *An Outline of Genericity*, Seminar für Natürlich-sprachliche Systeme, University of Tübingen, SNS Bericht 87-25, Tübingen, Germany, 1987.
14. L. K. Schubert and F. J. Pelletier, "An Outlook on Generic Statements," *Genericity in Natural Language*, M. Krifka, Ed., *Proceedings of the 1988 Tübingen Conference*, University of Tübingen, SNS-Bericht 88-42, Tübingen, Germany, 1988.
15. G. Schmidt and Th. Wetter, "Towards Knowledge Acquisition in Natural Language Dialogue," *Proceedings of the Third European Workshop on Knowledge Acquisition for Knowledge-Based Systems (EKAW-89)*, Paris, July 1989.
16. J. G. Carbonell and P. J. Hayes, "Natural-Language Understanding," *Encyclopedia of Artificial Intelligence*, S. C. Shapiro and D. Eckroth, Eds., John Wiley & Sons, Inc., New York, 1987.
17. C. Habel, "Das Lexikon in der Forschung der Künstlichen Intelligenz," *Handbuch der Lexikologie*, Chr. Schwarze and D. Wunderlich, Eds., Athenäum, Königstein, Germany, 1985.
18. T. Winograd, *Language as a Cognitive Process, Volume 1: Syntax*, Addison-Wesley Publishing Co., Reading, MA, 1983.
19. F. Guenthner, "Linguistic Meaning in Discourse Representation Theory," *Synthese* **73**, 569–598 (1987).
20. R. Jackendoff, *Semantics and Cognition*, MIT Press, Cambridge, MA, 1983.
21. P. N. Johnson-Laird, *The Computer and the Mind*, Harvard University Press, Cambridge, MA, 1988.
22. M. Bates, "Natural-Language Interfaces," *Encyclopedia of Artificial Intelligence*, S. C. Shapiro and D. Eckroth, Eds., John Wiley & Sons, Inc., New York, 1987.
23. F. Guenthner, "From Sentences to Discourse: Some Aspects of the Computational Treatment of Language," *Natural Language at the Computer*, A. Blaser, Ed., *Lecture Notes in Computer Science*, Vol. 320, Springer, Heidelberg, Germany, 1988.
24. C. R. Rollinger, "Simulation sprachlichen Verstehens: Generelle Probleme bei der semantischen Interpretation der natürlichen Sprache," *Computational Linguistics. An International Handbook of Computer Oriented Language Research and Applications*, I. Batori, W. Lenders, and W. Putschke, Eds., de Gruyter, Berlin, Germany, 1989.
25. R. Declerck, "The Origins of Genericity," *Linguist.* **29**, 79–102 (1991).
26. G. N. Carlson, "A Unified Analysis of the English Bare Plural," *Linguist. & Philos.* **1**, 413–457 (1977).
27. M. Diesing, "Bare Plural Subjects and the Stage/Individual Contrast," *Genericity in Natural Language*, M. Krifka, Ed., *Proceedings of the 1988 Tübingen Conference*, University of Tübingen, SNS-Bericht 88-42, Tübingen, Germany, 1988.
28. Z. Vendler, *Linguistics in Philosophy*, Cornell University Press, Ithaca, NY, 1967.
29. F. Haft, R. P. Jones, and Th. Wetter, "A Natural Language Based Legal Expert System Project for Consultation and Tutoring—The LEX Project," *Proceedings of the 1st International Conference on Artificial Intelligence in Law (ICAIL)*, Boston, May 1987, ACM, 1987.
30. Hu. Lehmann, "The Lex-Project—Concepts and Results," *Natural Language at the Computer*, A. Blaser, Ed., *Lecture Notes in Computer Science*, Vol. 320, Springer, Heidelberg, Germany, 1988.
31. Br. Bläser and H. Lehmann, "Ansätze für ein Natürlichsprachliches Juristisches Konsultationssystem," *Neue Methoden im Recht*, F. Haft, Ed., attempto, Tübingen, Germany, 1990.
32. Franz Guenthner, Hubert Lehmann, and Wolfgang Schönfeld, "A Theory for the Representation of Knowledge," *IBM J. Res. Develop.* **30**, 39–56 (1986).
33. H. Kamp, "A Theory of Truth and Semantic Representation," *Formal Methods in the Study of Language*, Groenendijk et al., Eds., Mathematical Centre Tract, Amsterdam, Netherlands, 1981. Reprinted in *Truth, Representation and Information*, J. A. G. Groenendijk, T. M. U. Janssen, and M. B. J. Stokhoff, Eds., Foris, Dordrecht, Netherlands.
34. J. B. Black, "An Exposition on Understanding Expository Text," *Understanding Expository Text*, B. K. Britton and J. B. Black, Eds., LEA, Hillsdale, NJ, 1985.
35. J. R. Hobbs, "Why Is Discourse Coherent?" *Coherence in Natural Language*, F. Neubauer, Ed., Buske, Hamburg, Germany, 1983.
36. W. C. Mann and S. A. Thompson, "Relational Propositions in Discourse," *Discourse Proc.* **9**, 57–90 (1986).
37. B. J. F. Meyer, "Prose Analysis: Purposes, Procedures, and Problems," *Understanding Expository Texts*, K. B. Britton and J. B. Black, Eds., Lawrence Erlbaum, Hillsdale, NJ, 1985.
38. B. J. Grosz and C. L. Sidner, "Attention, Intention, and the Structure of Discourse," *Computational Linguist.* **12**, 175–204 (1986).
39. *Computational Models of Discourse*, M. Brady and R. C. Berwick, Eds., MIT Press, Cambridge, MA, 1983.
40. R. G. Schank and R. P. Abelson, *Scripts, Goals, Plans, and Understanding*, Lawrence Erlbaum, Hillsdale, NJ, 1977.
41. R. Wilensky, *Planning and Understanding*, Addison-Wesley Publishing Co., Reading, MA, 1983.
42. R. Reichman-Adar, "Extended Person–Machine Interface," *Artif. Intell.* **22**, 157–218 (1984).
43. R. J. H. Scha, B. C. Bruce, and L. Polanyi, "Discourse Understanding," *Encyclopedia of Artificial Intelligence*,

S. C. Shapiro and D. Eckroth, Eds., John Wiley & Sons, Inc., New York, 1987.
44. F. Guenthner and Hu. Lehmann, "Verarbeitung natürlicher Sprache: ein Überblick," *Informatik Spektrum* **9**, No. 3, 162–173 (1986).
45. T. A. v. Dijk, *Studies in the Pragmatics of Discourse*, Mouton Publishers, New York, 1981.
46. W. Stegmüller, *Erklärung, Begründung, Kausalität (Probleme und Resultate der Analytischen Philosophie und Wissenschaftstheorie Bd. 1)*, Springer, Berlin, Germany, 1983.
47. J. Pearl, "Embracing Causality in Default Reasoning," *Artif. Intell.* **35**, 259–271 (1988).
48. S. W. Draper, "What's Going On in Everyday Explanation?" *Analyzing Everyday Explanation*, Ch. Antaki, Ed., Sage, London, 1988.
49. R. Cohen, "Analyzing the Structure of Argumentative Discourse," *Computational Linguist.* **13**, 11–24 (1987).
50. F. Marton and C. G. Wenestam, "Qualitative Differences in Understanding and Retention of the Main Point in Some Texts Based on the Principle-Example Structure," *Practical Aspects of Memory*, M. M. Gruneberg, P. E. Morris, and R. N. Sykes, Eds., Academic Press, London, 1978, pp. 633–643.
51. M. T. H. Chi and M. Bassok, "Learning from Examples via Self-Explanations," *Cognition, Learning and Instruction*, L. B. Resnick, Ed., LEA, Hillsdale, NJ, 1989, pp. 251–282.
52. *Similarity and Analogical Reasoning*, S. Vosniadou and A. Ortony, Eds., Cambridge University Press, Cambridge, England, 1989.
53. J. Holland, K. J. Holyoak, R. E. Nisbett, and P. Thagard, *Induction: Processes of Inference, Learning and Discovery*, MIT Press, Cambridge, MA, 1986.
54. T. Herrmann, *Speech and Situation: A Psychological Conception of Situated Speaking*, Springer, Heidelberg, Germany, 1983.
55. W. J. M. Levelt, *Speaking: From Intention to Articulation*, MIT Press, Cambridge, MA, 1989.
56. H. P. Grice, "Logic and Conversation," *Syntax and Semantics*, Vol. 3, *Pragmatics*, P. Cole and J. Saddock, Eds., Academic Press, Inc., New York, 1975.
57. D. Sperber and D. Wilson, *Relevance: Communication and Cognition*, Harvard University Press, Cambridge, MA, 1986.
58. G. Ryle, " 'If,' 'So,' and 'Because,' " *Philosophical Analysis*, M. Black, Ed., Cornell University Press, Ithaca, NY, 1950.
59. S. Toulmin, *The Uses of Argument*, Cambridge University Press, Cambridge, England, 1958.
60. Th. Wetter and F. Schmalhofer, "Knowledge Acquisition from Text-Based Think-Aloud Protocols: Situational Specifications for a Legal Expert System," *Proceedings of the European Knowledge Acquisition Workshop (EKAW-88)*, Bonn, Germany, 1988.
61. Y. Shoham, "Nonmonotonic Reasoning and Causation," *Cogn. Sci.* **14**, 213–252 (1990).
62. Th. Wetter and B. Woodward, "Towards a Theoretical Framework for Knowledge Acquisition," *Proceedings of the AAAI Workshop for Knowledge Acquisition*, Banff, Canada, 1990.
63. W. Clancey, "Viewing Knowledge Bases as Qualitative Models," *IEEE Expert* **4**, No. 2, 1–23 (1990).
64. J. R. Hobbs and R. C. Moore, *Formal Theories of the Commonsense World*, Ablex Publishing Corporation, Norwood, NJ, 1985.
65. *Intentions in Communication*, P. R. Cohen, J. Morgan, and M. E. Pollack, Eds., MIT Press, Cambridge, MA, 1990.
66. R. H. Kluwe, C. Misiak, and R. Schmidle, "Wissenserwerb beim Umgang mit einem umfangreichen System: Lernen als Ausbildung subjektiver Ordnungsstrukturen," *Bericht über den 34. Kongreß der Deutschen Gesellschaft für Psychologie in Wien*, D. Albert, Ed., Hogrefe, Göttingen, Germany, 1984.
67. A. L. Kidd, *Knowledge Acquisition for Expert Systems, A Practical Handbook*, Plenum Press, New York, 1987.
68. J. Krause, *Mensch-Maschine-Interaktion in natürlicher Sprache*, Niemeyer, Tübingen, Germany, 1982.
69. M. Zoeppritz, "Computer Talk?" *Technical Note 85.05*, IBM Germany, Heidelberg Scientific Center, Heidelberg, Germany, 1985.
70. St. Szpakowicz, "Semi-Automatic Acquisition of Conceptual Structure from Technical Texts," *Proceedings of the Third Knowledge Acquisition for Knowledge-Based Systems Workshop*, Banff, Canada, November 1988.
71. J. Sowa, *Conceptual Structures: Information Processing in Mind and Machine*, Addison-Wesley Publishing Co., Reading, MA, 1984.
72. K. Silvestro, "An Explanation-Based Approach to Knowledge Base Acquisition," presented at the Second AAAI Knowledge Acquisition for Knowledge-Based Systems Workshop, Banff, Canada, October 1987 (to appear in *Int. J. Man-Machine Studies*).
73. J. H. Alexander, M. J. Freiling, B. Phillips, and S. L. Messick, "Rule Acquisition for Expert Systems," patent application of Nov. 22, 1985, European Patent Office, Erhardt-Str. 27, 8000 Munich 2; Application number 85114858.5.
74. L. S. Lefkowitz and V. L. Lesser, "Knowledge Acquisition as Knowledge Assimilation," presented at the Second AAAI Knowledge Acquisition for Knowledge-Based Systems Workshop, Banff, Canada, October 1987.
75. Ö. Dahl, "The Expression of the Episodic-Generic Distinction in Tense Aspect Systems," *Genericity in Natural Language*, M. Krifka, Ed., *Proceedings of the 1988 Tübingen Conference*, University of Tübingen, SNS-Bericht 88-42, Tübingen, Germany, 1988, pp. 95–106.
76. B. J. Wielinga, B. Bredeweg, and J. A. Breuker, "Knowledge Acquisition for Expert Systems," *Advanced Topics in Artificial Intelligence*, T. Nossum, Ed., *Lecture Notes in Artificial Intelligence*, Vol. 345, Springer, Berlin, Germany, 1987.
77. Th. Wetter, "First Order Logic Foundation of the KADS Conceptual Model," *Current Trends in Knowledge Acquisition*, B. Wielinga, J. Boose, B. Gaines, G. Schreiber, and M. Van Someren, Eds., IOS Press, Amsterdam, Netherlands, 1990.
78. Hu. Lehmann, N. Ott, and M. Zoeppritz, "A Multilingual Interface to Databases," *IDEE Database Eng. Bull.* **8**, No. 3 (September 1985).

Received September 28, 1990; accepted for publication November 6, 1991

Thomas Wetter *IBM Germany, Scientific Center, Wilckensstrasse 1a, D6900 Heidelberg, Germany (WETTER at GYSVMHD1).* Dr. Wetter studied mathematics and informatics at the Technical University of Aachen. He investigated topics in mathematical modeling of physiological and medical phenomena, receiving his Ph.D. in mathematics related to logic-based medical diagnosis in 1984. Dr. Wetter joined IBM as a research staff member in 1984, working in software ergonomics, expert systems, and knowledge acquisition. He teaches graduate courses on several aspects of AI at the universities of Heidelberg and Kaiserslautern. Dr. Wetter currently is vice chair of the GMDS working group expert

systems (Association for Medical Documentation, Information Science, and Statistics) and of the technical board expert systems of the committee on artificial intelligence in the German Association for Information Science. He organized the Sixth European Knowledge Acquisition Workshop (EKAW-92) in May 1992.

Ralf Nüse *University of Heidelberg and IBM Germany, Scientific Center, Wilckensstrasse 1a, D6900 Heidelberg, Germany.* Mr. Nüse studied psychology, philosophy, and psycholinguistics at the Universities of Bonn, Bochum, and Heidelberg, Federal Republic of Germany. He works in the special Language and Situation research group ("Sonderforschungsbereich") sponsored by the German Research Association at the Universities of Heidelberg and Mannheim. Since 1989, he has also had a research contract with the IBM Scientific Center, Heidelberg, where he is investigating psycholinguistic and philosophical aspects of knowledge acquisition. Mr. Nüse is the author of *Über die Erfindung/en des Radikalen Konstruktivismus* (1991; with N. Groeben, B. Freitag, and M. Schreier) and of several articles and reports on problems of language and situation.