

# Interaction between syntax and semantics: The case of gerund translation

Stephan Mehl  
Univ. Duisburg  
Computational Linguistics  
D-47048 Duisburg  
Tel. +49-203-3792876  
email he234me@unidui.uni-duisburg.de

## Abstract

Standard architectures favor sequential processes for the semantic and syntactic part of natural language generation. Some semantic decisions, however, require information from the syntactic part, as is shown at the example of translating English gerunds into German. As a solution to this problem, a model is proposed in which syntactic and semantic modules work in parallel, computing preference values for different translation variants. On the basis of these preferences, one variant is selected and further elaborated. This model draws on a maximum of information, including stylistic considerations, while avoiding the computational load of backtracking processes.

## 1. The need for interaction

### 1.1 The role of lexical choice

Traditionally, the process of text generation is split into a strategic and a tactical phase. Both may be subdivided; among others, the tactical component contains two distinct phases of lexical choice and syntactic realization. This paper will challenge two features usually coming along with this division of labour: first, that there is usually not much choice involved in lexical choice; and second, that these phases are arranged in a strictly sequential way.

Fortunately, almost any content can be formulated in a number of different ways. These variants of saying the same in other words may differ lexically, syntactically, stylistically, or, to a certain degree, even semantically (e.g. by a slight change of focus that may be irrelevant in most contexts). Sometimes, there are pragmatic reasons for repeating a certain content in other words, i.e. paraphrasing it (cf. Lenke 1995). In most cases, however, the tactical component of a generation system will simply have to decide on one or the other way of formulating a given content. Since the syntactic part of generation largely depends on the syntactic properties of the words involved, it is the phase of lexical choice that carries most of the responsibility for the resulting text.

Many natural language generation systems, however, content themselves with producing one single form per content. Concepts are assigned to lexical items in a 1:1-fashion, conceptual roles corresponding to case frames of verbs. (The term "lexical item" is intended to cover both single words as well as phrases, may they be frozen or subject to morphosyntactic variation. This includes even descriptive phrases used in case of lexical gaps (cf. e.g. Sondheimer et al. 1990). The following discussion will focus on words, but can easily be extended to phrases.)

If there is some choice, i.e. a 1:n-assignment of concepts to words, it is important to distinguish whether this choice is determined semantically or by syntactic and/or stylistic considerations. The first case typically appears if the concepts in use are so abstract that they do not specify semantic features that distinguish different lexemes of a particular language. The most extreme example of this is conceptual dependency theory as embodied in Goldman's (1975) generation system BABEL. However, this is no real situation of choice in the sense that whatever alternative one chooses, the result might be acceptable. Instead, selecting a different word will yield different contents of a sentence. This does not mean that choice will be considered an entirely arbitrary matter here; since no two linguistic forms can be used equally well in all situations, there will always be good reasons for deciding on one or the other alternative. But our interpretation of lexical choice will be that of choosing between semantically (near-)equivalent words or phrases.

A number of researchers have recently suggested improvements of the classical 1:1-assignment. The following variants have been considered:

- synonyms: the choice among synonymous or near-synonymous words and phrases
- different combinations of concepts: A single word may cover a combination of concepts. An example of this approach is Horacek's (1990) model of zoom schemata. Since concepts might be combined in different ways, alternative formulations of the same content might arise.
- mapping of concepts to syntactic functions: In some cases, parts of the conceptual representation might either be rendered by lexemes or by syntactic functions (as in *Peter's car* vs. *The car owned by Peter*; cf. Horacek 1990).

While conveying the same semantic content, one aspect of these variants is that they may differ with respect to their syntactic behavior. The focus of this paper will be on different syntactic features of these synonyms as a criterion for the choice among them. Other researchers have investigated different criteria; cf. e.g. the pragmatic features discussed by Hovy 1988.

Furthermore, this paper will concentrate on the comparison of lexemes (namely, a verb and its nominalization) rather than phrases. The determination of nominalizations is taken for granted; research in the paradigm of Mel'cuk's Meaning-Text Theory has provided many details on this subject (see e.g. Iordanskaja et al. 1991). For further aspects of lexicalization, see the survey by Stede (forthcoming).

## 1.2 Syntax and semantics

Since near-synonyms may slightly deviate from the contents stipulated by the strategic component, the latter might have to sanction the choice of such a lexeme. This is a first indication that the classical sequential architecture should be replaced by some kind of interaction between the components.

The aim of this paper is to investigate the need for such an interaction between the phases of lexical choice and syntactic generation. In general, it is only possible to combine the chosen lexemes to a well-formed sentence if exactly one of them is a verb, and the others correspond to the pattern of optional and obligatory complements of this verb. It is part of the tactical component's duty to partition the contents to be verbalized into sentences in such a way that the lexemes assigned to form one sentence fulfil the above-mentioned conditions.

Because of the usual correspondence between concepts and words as well as between conceptual roles, semantic case frames and syntactic government frames, this process is no problem in most generation systems. However, difficulties arise in the following two cases:

- a) in machine translation and in multilingual generation, since government frames are language-specific,
- b) in choosing between synonyms or near-synonyms with different syntactic features.

An obvious case for b) is the nominalization of verbs and adjectives. Most generation systems work on a restricted domain in which a simplistic view of parts of speech is employed: nouns correspond to objects, verbs correspond to actions or states, adjectives correspond to properties. Of course, this need not be the case in general; the lexical system of every language allows for a multitude of derivations so that a concept denoting an action might as well be represented by a noun or an adjective instead of a verb. This is a powerful instrument that might give the text planner ample possibilities for combining contents into one sentence, thus avoiding the production of tedious chains of sentences.

Unfortunately, these possibilities are constrained by a number of language-specific idiosyncrasies: Not always is a desired derivative part of the lexical system, or if it is, it might be unusual or ambiguous. In addition to that, since most concepts are associated with role fillers, their lexical counterparts must be attached to the corresponding words as complements or modifiers, too. The resulting construction may be either ambiguous or clumsy (or both). Even worse, a derivative might only allow for a smaller number of complements than the original verb did. In those cases, the lexical choice component must avoid using derivatives. However, it can only do so on the basis of syntactic knowledge; that is, a failure of the syntactic generation component must trigger a revision of earlier decisions.

In the following section 2, a detailed example for this catalogue of criteria will be given. Section 3 will discuss possibilities of designing the interface between syntax and semantics accordingly, as well as describe the implementation of a detailed model.

Note that most of the above-mentioned criteria don't yield binary ratings that accept or completely discard a solution. Instead, there will be a tradeoff between better and worse values on different dimensions of evaluation. In section 3 below, it will be shown how the best result might evolve in the course of the generation process.

## **2. A case study: Translating gerunds**

As an illustration of the complex of problems described in section 1, an example from the field of machine translation shall be used, namely, the translation of English gerund constructions into German. German does not permit gerund constructions; instead, either subordinate clauses or nominalizations have to be used. Not always are both alternatives possible, and if they are, they may not be equally fortunate. The following examples illustrate the factors mentioned in section 1.2; they are taken from Lyons 1977 (vol. I,II) and its translation into German (vol. I, 1980; vol. II,1983; page numbers added in brackets):

### 1. availability of a nominal derivative

*Before embarking upon the discussion of this question, ... (429)*

There is no German noun expressing the act of embarking (upon a discussion), hence the only translation possible is a subordinate clause with a finite verb:

*Bevor wir uns der Diskussion dieser Frage widmen, ... (II:59)*

Note that the German subordinate clause requires a subject that is not explicitly mentioned in the English gerund clause.

## 2. ambiguity (in the respective context) and stylistic features of this noun

*For example, 'Abiogenesis is spontaneous generation' can be understood as expressing, indirectly, a proposition about 'abiogenesis' ... (417)*

As a matter of fact, there is a German noun denoting the act of expressing, namely *Ausdruck*. However, this noun is ambiguous in a similar way as the English *expression*: it can mean an act (as in *Take these flowers as an expression of our gratitude.*) as well as an object (as in *This is not a well-formed logical expression.*) Using this noun as a description of a sentence will lead to an ambiguity that can be avoided easily by translating the gerund as a subordinate clause:

*Zum Beispiel kann der Satz 'Urzeugung ist selbsttätiges Entstehen von Leben' so verstanden werden, daß mit ihm indirekt eine Proposition über das Wort 'Urzeugung' ausgedrückt werden soll ... (II:48)*

Note that not every lexical ambiguity will enforce this decision: Readings that can easily be discarded in a particular context will not hinder the interpretation of an ambiguous word. In the same vein, rare words (or readings) as well as words that do not fit the stylistic register of the text will decrease its readability (see also Mehl 1994).

## 3. possibility to combine the noun with all complements

*The fact that the term 'expression' is in existence does not, of course, constitute sufficient reason for distinguishing it from 'lexeme', on the one hand, and from 'form', on the other. (23)*

In this example, one complement of the gerund verb consists of a pronoun. The respective German noun, however, does not permit the use of a pronoun complement (*\*seiner Unterscheidung von ...*). This leaves us with a subordinate clause:

*Die Tatsache, daß der Terminus 'Ausdruck' existiert, stellt jedoch natürlich keinen genügenden Grund dafür dar, ihn von 'Lexem' einerseits und 'Form' andererseits zu unterscheiden. (I:36)*

## 4. syntactic and semantic ambiguity of this combination

It is not always the case that pronouns exclude the use of a noun. The following example

*Having made this point and given it due emphasis, ... (12)*

might be translated as

*Nach der Feststellung dieses Punktes und seiner gebührenden Betonung ...*

However, *seiner gebührenden Betonung* might as well mean *its having given emphasis*. This is a systematic ambiguity that always occurs when a constituent in the German genitive might fit different semantic roles. In such cases, the subordinate clause variant makes relations more explicit. (In this example, the translator chose to drop some of the original contents: *Nachdem wir diesen Punkt gebührend betont haben ... (I:25)*).

## 5. stylistic evaluation of both alternatives

Even if its interpretation is unambiguous, a noun phrase containing several modifiers and complements is hard to understand. In any case, constructions with multiply embedded noun phrases (which are easy to build in German) should be avoided. In the following example, a gerund verb with two complements, one of which is very complex, has actually been translated as a noun, yielding a clause in which head (*Einfügung* [*insertion*]) and second modifier (*vor 'John'* [*before 'John'*]) have been torn apart by the first modifier:

*[...] it can be made clear by inserting the phrase 'the name', or some similar descriptive expression, before 'John'.* (6)

*[...] dies kann durch die Einfügung des Ausdrucks 'der Name', oder eines ähnlichen beschreibenden Ausdrucks vor 'John' klar werden.* (I:20)

Translating the gerund by a subordinate clause would have produced a text that is by far easier to understand:

[...] *dies kann klar werden, wenn man den Ausdruck 'der Name' oder einen ähnlichen beschreibenden Ausdruck vor 'John' einfügt.*

On the other hand, subordinate clauses may become equally confusing if too many of them are lined up or embedded into each other:

*One of the problems that arises in describing precisely the relationship that holds between lexemes and expressions [...] (25)*

Instead of a series of subordinate clauses such as

*Eines der Probleme, das entsteht, wenn man die Beziehung, die zwischen Lexemen und Ausdrücken besteht, genau beschreibt, [...]*

the translator chose a nominalization in addition to dropping a subordinate verb (*hold*):

*Eines der Probleme, das gerade bei der Beschreibung der Beziehung zwischen Lexemen und Ausdrücken [...] existiert [...] (I:38)*

The next section will show how this complex of criteria can be brought to bear on the decision for a certain lexical item.

### 3. System design

The following section will use gerund translation as an example of natural language generation. For the sake of simplicity, some specifications will be given only for this example. However, the system design to be developed here is easily extendable to other natural language generation tasks.

As was shown in the previous sections, the choice between translating a gerund as a subordinate clause or as a noun phrase depends on numerous syntactic, semantic and stylistic factors that can be evaluated only in course of the entire tactical generation process. Different architectures are possible for coping with this situation:

Dimension 1: number of candidate structures to be elaborated

- 1.1 selection of one alternative only; if the chosen candidate fails, a process of backtracking causes the original decision to be revised.
- 1.2 elaboration of all candidates; at the end of the generation process, the best solution is selected.
- 1.3 elaboration of all candidates the rating of which exceeds some threshold, again with final determination of the winner (at the risk of finding none because the threshold suppressed the only possibility).

Dimension 2: arrangement of processes

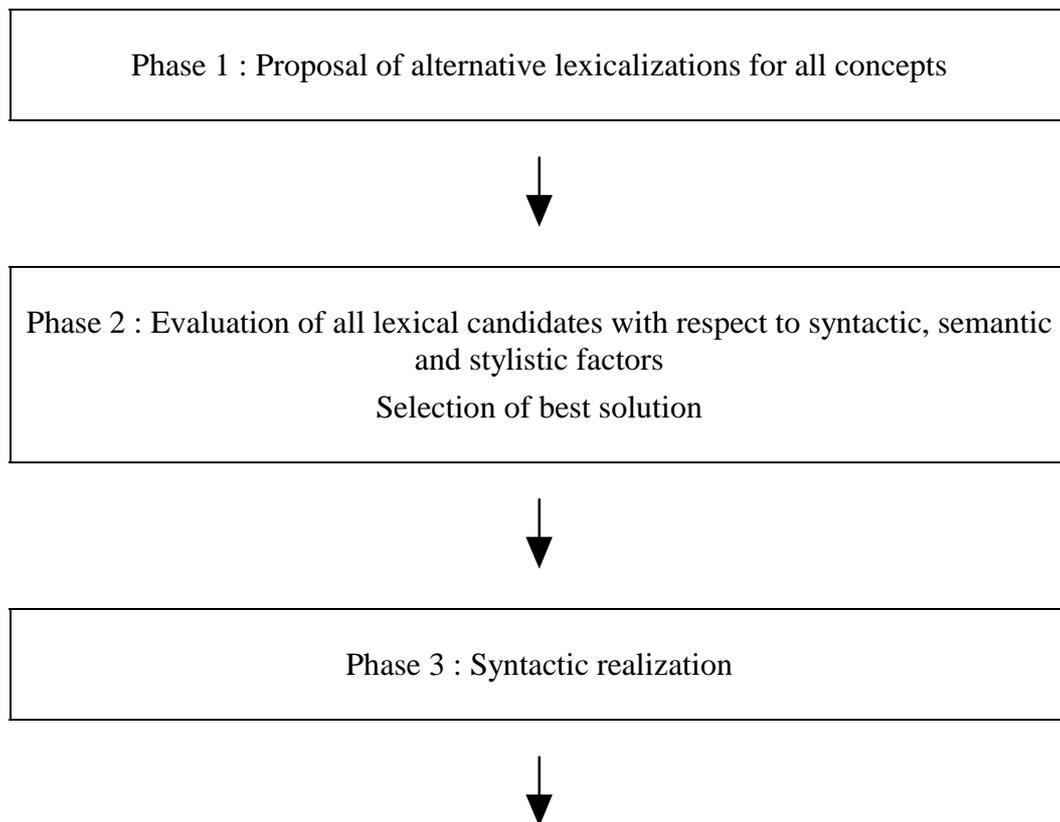
- 2.1 sequentially (e.g. lexical choice before syntax before stylistics)
- 2.2 incremental-interactively: several modules work in parallel, communicating either by immediate inquiries and messages or via a blackboard

The success of strategies 1.1 - 1.3 (in terms of computational load and of quality of the result) depends on the quality of the heuristics: If the only candidate to be elaborated has to be selected arbitrarily, it is quite probable that backtracking will be necessary, eventually leading to a repetition of some of the generation work already done. A good heuristics, on the other hand, might avoid the elaboration of numerous structures that will finally be thrown away. -

Note that heuristical evaluation is comparative; every candidate has to be weighted initially before one can be selected and further developed.

To judge the prospects of an alternative properly, factors from the semantic, syntactic as well as stylistic evaluation will have to be taken into account. This in turn favors an interactive architecture (2.2) over a sequential one. (Cf. Stede [forthcoming] who provides arguments for a model of incremental lexicalization.)

In view of these considerations, the following model of the generation process is currently being implemented:



Phase 4 : Syntactic and semantic evaluation of the structural candidates proposed  
by phase 3

Return to phase 2 if no satisfying result is achieved

The components of this model will now be explained in detail.

### 3.1 Phase 1

Phase 1 yields a semantic case frame that specifies alternative lexemes together with their semantic roles. For the purpose of illustration, only the various correlates of the gerund verb will be listed. The translation of the main clause will be regarded to be independent of the translation of the gerund (which might not be the case). Since gerunds may modify verbs as well as nouns, the following two patterns are possible:

Pattern 1: Modification of verb

Main clause

Verb

Other Complements & Adjuncts

Complement/Adjunct corresponding to English gerund

Alternatively: Verb - Noun

Complements/Adjuncts

Example:

English original:

*This interpretation of (2) can be accounted for in terms of the notion of type-token identity by generalizing from any particular occurrence of 'reference' to the class of all its tokens. (14)*

Result of phase 1 (simplified):

Main clause:

Verb            *erfassen [account]*

Agent --

Object *Interpretation*

Instrument    *Begriff [notion]*

Specification            *Typ-Exemplar-Identität*  
*[type-token identity]*

Means

Verb/Noun            *generalisieren / Generalisierung,*  
*verallgemeinern/Verallgemeinerung*

Agent                    --

Source                *Vorkommen [occurrence]*

Goal	Object ' <i>Referenz</i> ' <i>Klasse</i> [class] Specification <i>Exemplare</i> [tokens]
------	--

<u>Pattern 2: Modification of noun</u>	
Main clause	
	Verb
	Other Complements & Adjuncts
	Complement/Adjunct containing English gerund
	Complement/Adjunct corresponding to English gerund
	<u>Alternatively:</u> Verb - Noun
	Complements/Adjuncts
	Other Complements & Adjuncts

Example:  
English original:

*Any other convention would serve for the purpose of constructing metalanguage-names (11)*

Result of phase 1 (simplified):

Main clause:	
Verb	<i>diene[n]</i> [serve]
Agent	<i>Konvention</i> [convention]
Beneficient	<i>Zweck</i> [purpose]
	Specification
	Verb/Noun <i>konstruieren</i> / <i>Konstruktion</i>
	Agent                                      --
	Object <i>Namen</i> [names]

The interface between the strategical component and phase 1 as the beginning of the tactical component is deliberately left open. There might be different possible assignments between concepts and lexemes or phrases, as mentioned in section 1. The lexematic structure rendered by this phase might as well be the result of the transfer process of a transfer-based machine translation program. It will only be assumed that alternative lexemes may be weighted according to their degree of matching the conceptual input. Thus, near-synonyms can be proposed as an alternative to synonyms, but synonyms will be preferred; the ensuing evaluation process will then determine whether the preference for synonyms is outweighed by some syntactic or stylistic preference for a near-synonym.

Again for sake of simplicity, no translation variants will be considered here in which the original division of main clause and gerund clause is completely dropped (e.g. by using some fixed phrase to express the conceptual content of the source-language sentence). In general, of course, phase 1 may end up in proposing alternative case frames as well as alternative role fillers.

### 3.2 Phase 2: Evaluation of lexical candidates

As soon as phase 1 presents lexical candidates, they are evaluated according to the following criteria:

#### 3.2.1 Semantic evaluation

Besides its semantic closeness to the concept that it is supposed to express, a lexeme may be ambiguous and thus lead to false interpretations. However, an ambiguity is relevant only if it is valid in the context of the sentence to be produced. For instance, in the above-mentioned example the noun *Konstruktion* is proposed which denotes both 'construction' as a process and as the result of this process. If there were any objects mentioned in the text that could coreferentially be described as *Konstruktion* in the second sense, the use of this noun would give rise to an ambiguity. (As a matter of fact, such an ambiguity would be resolved in a moment by the modifier *metasprachliche Namen* which can only be attached to the process sense of *Konstruktion* for semantic reasons. However, the semantic component cannot detect this before phase 4, in which structures rather than lexemes are evaluated.)

Technically, the degree of ambiguity is determined by a semantic network that is tuned to a given context by way of a spreading activation mechanism (cf. Mehl 1993). The lexemes proposed by phase 1 are connected to relations in the semantic network representing the desired meaning. The degree of ambiguity is measured by summing up the activation of those semantic relations that represent other meanings of the lexeme in question.

#### 3.2.2 Syntactic evaluation

Different languages differ in their way of connecting semantic roles and syntactic valency frames. Obviously, the semantic role filled by a gerund in English must be expressed by some other syntactic construction in German. The respective head lexeme (a verb according to pattern 1, a noun according to pattern 2) determines the possible syntactic forms of a certain semantic role.

In the first place, the task of the syntactic evaluation is to find out the preferred way of expressing the semantic role in question. In the case of complements, many verbs and nouns allow several syntactic realizations (e.g. prepositional phrases, infinitives or subordinate clauses with finite verbs), but prefer one of them (cf. Ford et al. 1982). For example, *Zweck* prefers an infinitive or a prepositional phrase to a subordinate clause:

*irgendeine Konvention könnte dem Zweck dienen, Namen zu konstruieren*

*irgendeine Konvention könnte dem Zweck der Konstruktion von Namen dienen*

*? irgendeine Konvention könnte dem Zweck dienen, daß Namen konstruiert werden*

In extreme cases, only one syntactic realization might be possible, thus excluding either the noun or the verb variant of gerund translation.

Secondly, the syntactic module will assign a syntactic form to the other semantic roles of the main verb (or head noun). The best translation of a gerund largely depends on the overall structure of a sentence: Are there any other subordinate clauses? How many prepositional phrases are there? Does the preposition that is needed to translate the gerund as a prepositional noun phrase appear in any other phrase?

Consider again the examples from 3.1. Choosing the noun variant *Generalisierung*, both the Instrument and the Means role would require using the preposition *durch*:

*Diese Interpretation von (2) kann jedoch durch Generalisierung von irgendeinem bestimmten Vorkommen von 'Referenz' auf die Klasse aller seiner Exemplare auch durch den Begriff der Typ-Exemplar-Identität erfaßt werden.*

In contrast, there is no other subordinate clause that would impede a translation of the gerund as a subordinate clause.

### 3.2.3 Stylistic evaluation

The example just mentioned is not only awkward because the same preposition is used twice. Another reason is the complexity of the noun's complements, which both carry further modifications:

*von irgendeinem bestimmten Vorkommen von 'Referenz'  
auf die Klasse aller seiner Exemplare*

Therefore, a rule of thumb of improving the readability of texts might be to avoid noun phrases with more than two (possibly embedded) noun phrases as complements/adjuncts.

In 3.2.2, a rule was established not to generate two prepositional phrases with the same preposition. In addition to that, chains of noun phrases (including prepositional phrases) should be avoided in general. On the other hand, the same holds for chains of subordinate clauses. This does not mean that every sentence should contain exactly one or two noun phrase complements and one subordinate clause, but a sound mixture will be better than any kind of monoculture.

Finally, the degree of stylistic appropriateness of a lexeme might be another decision factor (cf. Stede 1993). It is calculated by comparing the respective ratings in the dictionary with the overall level of the text in question. For example, *generalisieren* as more technical a term than *verallgemeinern* fits better into Lyons's text.

### **3.3 Phase 3: Syntactic Realization**

In Phase 2, the syntactic form of the other complements and adjuncts in a sentence is determined, and it is calculated which of the possible variants of translation fits best into the sentence for semantic, syntactic or stylistic reasons. The task of phase 3 is simply to take that best-fitting variant and to state which variants again are possible for the noun phrase or subordinate clause corresponding to the gerund clause.

Should the preferred variant need a filler for an open semantic role, the semantic module is asked for help.

Since phase 2 has decided upon a translation of our example gerund as a verb, the results of phase 3 might be:

*... indem man von irgendeinem bestimmten Vorkommen von 'Referenz' auf die Klasse aller seiner Exemplare generalisiert.*

or, in passive voice:

*... indem von irgendeinem bestimmten Vorkommen ... generalisiert wird*

### **3.4 Phase 4: Structural Evaluation**

After having constructed the syntactic form of the entire sentence in phase 3, a final check should make sure that no ambiguities arise through the juxtaposition of isolated clauses. This applies mainly to the translation of gerunds as noun phrases, in which case the resulting phrase is embedded into the remaining sentence without boundary commas etc. This

embedding can lead to structural ambiguities: what was planned to be phrases on the same level might look like head-modifier structures.

Apart from structural ambiguities, a certain structure may be ambiguous with respect to its semantic interpretation. Again, this ambiguity will count only if it is relevant in a given context. Thus,

*ein Problem, das bei der Beschreibung der Beziehung entsteht, ...*

*[a problem that arises in the description of/by the relationship]*

is not ambiguous, whereas

*ein Problem, das bei der Beschreibung des Mannes entsteht, ...*

*[a problem that arises in the description of/by the man]*

might well be.

Technically, the plausibility of such an ambiguity is again checked by consulting the spreading-activation network: If there are two different semantic relations between two nouns which are both active and which can both be represented by the same syntactic form (e.g. by a genitive phrase), the structure in question is rejected.

#### **4. Summary**

It was shown that lexical choice needs to take syntactic as well as semantic and stylistic criteria into account. This leads to a number of problems: features from different levels of description have to be computed simultaneously; the ensuing preferences may be conflicting; and a solution has to be found without fully exploring all alternatives. In order to avoid ambiguities, syntactic structures are subject to a final semantic check. A model is currently implemented in which solutions to the various tradeoffs of the tasks involved are explored.

#### **Bibliography**

- Ford, Marilyn; Bresnan, Joan; Kaplan, Ronald 1982: A Competence-Based Theory of Syntactic Closure. In: J. Bresnan: The mental representation of grammatical relations. Cambridge, MA: MIT Press, p. 727-796.
- Goldman, N. 1975: Conceptual Generation. In: R. Schank (Hrsg.): Conceptual Information Processing. Amsterdam u.a.: North-Holland 1975, p. 289-371.
- Horacek, Helmut 1990: The Architecture of a Generation Component in a Complete Natural Language Dialogue System. In: R. Dale, C. Mellish, M. Zock (Hrsg.): Current Research in Natural Language Generation. London: Academic Press 1990, p. 193-228.
- Hovy, Eduard H. 1988: Generating Natural Language Under Pragmatic Constraints. Hillsdale: Erlbaum.
- Iordanskaja, Lidija/Kittredge, Richard/Polguère, Alain 1991: Lexical Selection and Paraphrase in a Meaning-Text Generation Model. In: C.Paris/W.Swartout/W.Mann (eds.): Natural Language Generation in Artificial Intelligence and Computational Linguistics. Boston: Kluwer, p. 294-312.
- Lenke, Nils 1995: Paraphrasen - Lösungen für antizipierte Leserprobleme bei der automatischen Textgenerierung. St. Augustin: infix (= doctoral dissertation, Univ. Duisburg, 1994).
- Lyons, John 1977: Semantics. Vol. I,II. Cambridge Univ. Press. German Translation: Semantik. München: Beck, Vol. I: 1980, Vol. II: 1983.

- Mehl, Stephan 1993: Dynamische semantische Netze. Zur Kontextabhängigkeit von Wortbedeutungen. St. Augustin: infix (= doctoral dissertation, Univ. Koblenz, 1992).
- Mehl, Stephan 1994: Synonyme bei der Erzeugung, Generierung und Produktion natürlicher Sprache. In: H. Trost (ed.): Proc. KONVENS '94, Wien, p. 240-249.
- Sondheimer, Norman; Cumming, Susanna; Albano, Robert 1990: How to Realize a Concept: Lexical Selection and the Conceptual Network in Text Generation. In: *Machine Translation* 5, 1990, p. 58-78.
- Stede, Manfred 1993: Lexical Choice Criteria in Language Generation. In: Proc. 6th EACL, Utrecht, 1993.
- Stede, Manfred (forthcoming): Lexicalization in Natural Language Generation: A Survey. To appear in *Artificial Intelligence Review*.