

# When *must* should be chosen

Ralf Klabunde

Ruhr-Universität Bochum  
Department of Linguistics  
D- 44780 Bochum  
klabunde@linguistics.rub.de

## Abstract

This paper describes ongoing work on the choice of modal expressions in the generation of recommendations about courses of study within the B.A.-program of the Ruhr-Universität Bochum. We focus our work on the German modal verbs *müssen* (must), *können* (can), the subjunctive form *sollte* (should), and *dürfen* (may). The idea is to combine insights from formal semantics into the meaning of modal verbs with requirements for their choice in NLG-systems in order to achieve a linguistically satisfying model of their choice. The overall model is implemented in the CAN-system that plans courses of study and generates recommendations about them.

## 1 Introduction

From a typological perspective, modality concerns the speaker's estimation of the possible and necessary occurrence, respectively, of some state of affairs [van der Auwera and Plungian, 1998]. In many languages, modality is linguistically realized by a whole bundle of different means, with modal verbs being the most prominent ones.

In this paper, we will describe the realization of uses of the four German modal verbs *müssen* (must, have to, need to), *können* (can, to be able to), *dürfen* (may, could, to be allowed to do) and the subjunctive form of *sollen*, i.e. *sollte* (should). As the English translations indicate, there is no simple correspondence between the meaning of these German modal verbs and their counterparts in other languages. In particular, the pragmatic conditions for their use differ widely so that a straightforward adoption of the conditions we are giving in this paper to other languages will probably not be possible.

In general, the problem of the choice of modal verbs is the fact that the use of modals is, among other things, determined by information that is usually only implicitly given in a discourse. This information must be made available during macroplanning for the entire text in a way that enables the choice of the contextually appropriate modal verbs.

In what follows, we will first outline the semantics and pragmatics of modal verbs in general, and the four German modals plus *sollen* in particular. Based on this, we will de-

lineate our choice model for modals and its application to the aforementioned four modal verbs. Finally, we will show how this model is integrated into the CAN-System that we are currently developing.

## 2 The meanings of modal verbs

The meaning of modals is deeply rooted in information that is only implicitly given in a discourse, but plays a crucial role in the full understanding of the underlying speaker's intention. This context-dependence results in different readings. A sentence like *Peter must attend the lecture on semantics* can be interpreted in numerous ways: It could express that according to the knowledge of the speaker it is necessary that Peter attends this lecture, or that Peter's course of study requires his attendance in the lecture on semantics.

Since Kratzer's (1977) seminal work on a uniform and formal treatment of the semantics of modal verbs, the context-dependence of the meaning of modals is traced back to so-called *conversational backgrounds* (also called *modal domains* or *modal bases*) that provide that information against which the modalized sentence has to be interpreted. According to her widely accepted analysis in formal semantics, the meaning of modal verbs depends on the class of permitted conversational backgrounds and the so-called *modal force*, i.e. whether the modal verbs express possibility or necessity. Possibility and necessity are defined as consistency and entailment with respect to the respective conversational background: a proposition is possible if it is consistent with the relevant background, and it is necessary if it is entailed by the conversational background.

Accordingly, modal verbs differ in their semantics with respect to these two parameters: their inherent modal force, and the classes of modal domains associated with them.<sup>1</sup>

---

<sup>1</sup> In more recent works, Kratzer's (1977) original work has been extended to a doubly relative analysis of modal verbs: Their meaning does not only depend on one modal domain, but additionally on a second one, the "ordering source". An ordering source is a set of propositions that describes the plausibility of the states given in the primary modal domain. Since this refinement of Kratzer's work does not concern our treatment of the choice of

The classifications of relevant conversational backgrounds that have been postulated for an adequate analysis of the meaning of modal expressions differ widely in the literature. They range from two types – a basic distinction between epistemic and “root” modals – to more than twenty in von Stechow’s (2004) classification.

Kratzer uses a modal logic to describe the meaning of modal verbs. However, she leaves open what conversational backgrounds really are (besides their characterization as sets of propositions), although this concept is the linchpin of a satisfying analysis of the meaning of modals.

Also discourse-oriented approaches to the meaning of modals as, e.g., Roberts (1989), Frank (1996), or Stone (1999) do not say much about the structure and content of conversational backgrounds. These approaches analyze specific uses of modals as anaphoric expressions: in some discourse segments, modals pick up modal domains introduced by previous modal expressions, and must be interpreted against this already introduced hypothetical scenario. Stone (1999), for example, gives the following sentence pair:

(1) *A wolf might walk into the house. It would eat you.*

In order to interpret the second modalized sentence correctly, the anaphoric connection between the content of that sentence and the possible scenario of the first one must be made explicit.

## 2.1 The meaning of *müssen*, *dürfen*, *sollen*, *sollte*, and *können*

We will outline what semanticists have to say about the meaning of these modal verbs, and characterize those pragmatic conditions for their use that are of importance for this paper.

*Müssen* and *sollen* both express necessity as modal force. The semantic difference between these two verbs concerns the admissible conversational backgrounds: While *müssen* can be used with any conversational background, the use of *sollen* is restricted to specific ones, viz. those that express a specified ideal like laws or social norms. For example, while the sentence

(2) *Peter muss den Kurs besuchen*  
(Peter must attend the lecture)

can, in principle, be interpreted with respect to any kind of conversational background, some of these backgrounds are not possible for the sentence *Peter soll den Kurs besuchen* (Peter shall attend the lecture). The former sentence can be interpreted as:

1. In view of what is known by the speaker/the listener/..., it is necessary that Peter attends the lecture (epistemic interpretation)

---

modals in this paper, we will not go into the details. However, we will pick up that point in the depiction of our future work.

2. In view of what the conditions of study determines, it is necessary that Peter attends the lecture (deontic interpretation)
  3. in view of someone’s will, it is necessary that Peter attends the lecture (volitional interpretation)
  4. in view of achieving certain goals it is necessary that Peter attends the lecture (teleological interpretation)
- and so on.

*Sollen*, however, cannot presuppose conversational backgrounds that are evoked by the subject (cf, von Stechow, 2004; Zifonun et al., 1997). For example, a deontic interpretation of *Peter soll den Kurs besuchen* is not possible with the respective obligations as conversational backgrounds determined by Peter himself.

In addition to the modal verbs, mood is also a linguistic means to express modality. Often, mood weakens the modal force so that *sollte* does not express the necessity of the embedded proposition anymore, but a weaker modal force that comes close to possibility.

*Können* and *dürfen* seem to be the counterparts of *müssen* and *sollen*. They express possibility as modal force. *Können* is not confined to specific conversational backgrounds, while *dürfen* can – similar to *sollen* – only be used with extra-subjective ideal backgrounds, preferably deontic ones. The sentence

(3) *Peter darf den Kurs besuchen*  
(Peter may attend the lecture)

means: his attendance of the lecture is compatible with his duties (imposed by someone else), the conditions of study, the will of a third person, some relevant facts, and so on.

Although this semantic analysis shows convincingly the relevance of the modal force and the conversational backgrounds for a systematic treatment of the meaning of modals, it cannot be the whole story of the meaning of modals. If conversational backgrounds and modal forces would sufficiently explain the meaning of modals, every modal verb with a restricted set of admissible conversational backgrounds could be replaced by the modals *können* and *müssen*, respectively, without changing the meaning of the respective sentence. Although this might be true with respect to the truth conditions of the single sentences, replacing the specific modals by more general ones nevertheless results in sentences with different readings. For example, replacing the sentence

(4) *Du darfst den Semantik-Kurs besuchen und du solltest den Syntax-Kurs besuchen* (You may attend the semantics lecture and you should visit the lecture on syntax)

by the sentence

(4') *Du kannst den Semantik-Kurs besuchen und du müsstest den Syntax-Kurs besuchen* (You can attend

the semantics lecture and you ought to attend the lecture on syntax)

leads to different prominent meanings: while the first sentence clearly expresses a permission and a recommendation, the second one is more neutral with respect to the relevant pragmatic meaning. Additionally, the second clause expresses a weak deontically motivated entailment.

Hence, the modal verbs are associated with specific pragmatic conditions for their use. We will refer to those pragmatics aspects only that are crucial for our treatment of the modals. We start from the assumption that the uses of *müssen* and *können* do not require specific pragmatic constraints. *Sollte* does not express necessity (as the indicative *sollen* does) but possibility, and shall be used if the goal of the system is to express the proposition as a strong advice. *Dürfen* shall be used if the system expresses a permission that is based on the courses the user has completed successfully so far, or if the user has the permission to study at all.

## 2 Modality in natural language generation systems

In general, there is not much work on the computational treatment of modals, be it in NL Understanding or NL Generation.

There seems to be a clear gap between the relevance of modality in texts and dialogues produced by human speakers and its consideration in NLG-systems. While modal expressions appear in all kinds of text genres produced by humans, be it newspaper articles or dialogue contributions, their realization in NLG-systems did not receive much attention so far.<sup>2</sup> This disregarding of modality might be due to the nature of the text types that typically describe entities as, e.g., animals (Peba-II; Milosavljevic 1997) or air-qualities (TEMSIS, Busemann and Horacek 1997). All of these generation systems present data or knowledge units that are “true” in a certain sense. None of these systems have to cope with an estimation whether some events will happen with a specific probability, given the information state of the system and the user. However, estimations of such a kind are the basis for the use of modals.

Our work is loosely inspired by the treatment of modals in Crouch and Pulman (1993), but differs from their work in several crucial respects. Crouch and Pulman investigate the use of temporal and modal expressions in an NL-interface to a planner that produces partial plans in an engineering domain. Instead of using a modal logic with its underlying possible world semantics, the authors determine possible and necessary propositions against the range of plans available to solve a task. The necessity relation holds if the proposition (more specifically, the temporal orders and rela-

tions in the corresponding formula) can be inferred from the plan description. In contrast to this, a formula is possible if its negation can be inferred from the plan description.

## 3 The choice of modal verbs in CAN

We are currently developing the CAN-system (Conceptualization for Modal Expressions) that generates recommendations about courses of study within the B.A. program of the Ruhr-Universität Bochum. A user provides the system with his/her current term number and the lectures he/she has finished so far, and the system generates a recommendation which lectures he should/must/shall/may/can take.

A planner provides the partially ordered plan which describes the progression of the overall study from the current semester to the final degree. The propositions are checked in a piecemeal fashion in order to determine whether they necessarily or possibly hold with respect to the conversational background. This procedure determines the modal force, and a subclass of the relevant modal verbs with it. Afterwards the pragmatic conditions will be checked to select the appropriate modal verb, and that verb is inserted into a syntactic template for the corresponding sentence.

We confine our analysis to epistemic and deontic meanings of the aforementioned modal verbs, because these readings are the most prominent ones in our domain. Deontic meanings are determined by the structure and content of the plan to be verbalized. Epistemic meanings are determined by the discourse state. Both knowledge sources will be handled by the same procedure which makes an extension to other types of modal meanings manageable. In other words, we adopt the bipartite distinction between necessary and possible formulae, although this might not completely correspond to the fine-grained modal forces expressed in natural languages. Furthermore, we adopt the modeling of possibility as consistency with the conversational background, and necessity as entailment from that background.

If one of the modal relations is true, we know what the modal force is, and are able to determine that subclass of modal verbs that expresses the proven modal force. Otherwise the underlying proposition will linguistically not be modalized.

The deontic background is determined by the plan. According to the partially ordered plan, some lectures can only be attended if other lectures had already been completed successfully, but this order does not hold between all lectures. The obligation to attend certain lectures during some semester can directly be read from the plan. The modal force will be derived from the plans.

An epistemic background is given by what the system knows of the user’s current course of his study and the information provided by the system so far. For this, we use a discourse record that keeps track of the user’s input and what he has been told.

### 3.1 Planning for the choice of modals

The task of the planner is to place the obligatory lectures at the right nodes in the plan, and to determine possible nodes of the optional lectures. For example, while all students of Linguistics have to attend a lecture on pragmatics in the

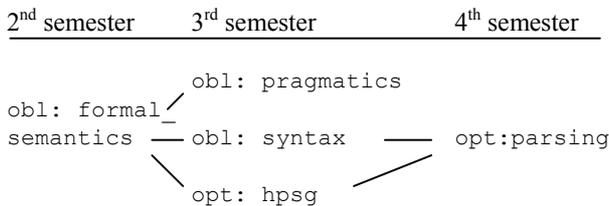
---

<sup>2</sup> This statement might be too strong in case of the use of systemic grammars. Systemic grammars are able to convert complex modality-related speech-functions into corresponding modal expressions. However, systemic grammars do not consider content planning tasks that we are describing in this paper.

third semester, they can attend an additional lecture with a topic of their own choice, but the student must have satisfied its prerequisites already.

An “action” of the plan is realized as a pair that represents the status of the lecture with respect to the semester (obligatory, optional), and the title of the lecture.

Suppose a fully specified plan contains the following order of single lectures, starting with the second semester:



If each “action” of this part of the plan shall linguistically be realized, only some of the four modal verbs are applicable.

The use of the modals depends on the status of the lectures as being obligatory or optional in a certain semester, and the existence of certain preconditions of the planning operators. The distinguishing feature of obligatory lectures is that there are no alternatives for them; they are necessary courses in a certain semester. Hence, the lectures on semantics and pragmatics will be mentioned in sentences with *müssen* as modal verb. Optional lectures appear to be more interesting because in principle they can be attended in any semester, as long as the prerequisites for their attendance have been achieved by the student. The prerequisites are modelled as preconditions of corresponding planning operators. Linguistically, optionality is expressed by modals which express possibility as modal force, and the use of these modals is constrained by additional pragmatic conditions for their use. Whether the optional lectures on HSPG and parsing, respectively, will be mentioned in sentences with *können*, *dürfen*, or *sollte* depends on the decision to recommend or permit these lectures, which will be derived from conditions we will mention in the following section.

To sum up, the choice of the appropriate modal verbs depends on:

- the status of the lecture, whether it is optional or obligatory in one semester,
- the user’s possible achievement of prerequisites of optional lectures,
- the placement of optional lectures in the overall plan,
- pragmatic functions like expressing a permission or a recommendation.

The first three factors belong to the conversational background, while the last one seems to be an additional constraint for the use of some of the modal verbs. Dealing with these aspects of the conversational background can be characterized in the following way:

Let P be the plan with the obligatory lectures placed in the right positions, and T(P) the set of all totally specified plans based on P that include possible placements of the optional lectures.

Every lecture in P marked as obligatory is necessary with respect to the study process.

Additionally, for every proposition p we check whether  $T(P) \rightarrow p$  holds or  $T(P) \cup \{p\}$  is true. If the former formula is true, the modal verb *müssen* can be realized. If the latter formula is true, we have to check whether the pragmatic conditions for the use of one of the modals verbs that express possibility are given in the discourse record. The corresponding decision results in the choice of the verb *dürfen*, *können*, and *sollte*, respectively.

### 3.1 An example: the choice of *dürfen* and *sollte*

The choice of *dürfen* and *sollte* may illustrate the choice mechanism that takes into account the plan structure and the discourse state, respectively.

Both *dürfen* and *sollte* express possibility as modal force so that they compete with *können* as the unmarked modal verb for expressing possibility. Since *dürfen* and *sollte* can only be used if specific pragmatic conditions are satisfied (and with restricted sets of conversational backgrounds), *können* will always be chosen if the modal force is possibility and the other two modals cannot be used.

Expressing a permission by using *dürfen* is based on the following conditions:

- (a) if the user is willing to begin his studies and CAN does not know whether he is registered at the university at all, CAN asks for a corresponding confirmation. In case of a positive response the permission to study the courses of the first semester can be given, and *dürfen* will be used.
- (b) If CAN does not know whether some courses as preconditions of a course are satisfied it asks the user whether he attended these required courses. The user has to give this information, and the system integrates the courses into the plan. Since the preconditions for the respective course are now satisfied, the permission to attend this course will be signalled by the use of *dürfen*.

In general, our impression is that permissions – the pragmatic function of *dürfen* when a deontic conversational background becomes relevant – can best be derived in dialogue situations. An NLG-system that permits to attend a certain course must be set into the position to express the permission at all. In our scenario, this is the case if CAN requires additional information and asks for it. In this respect permissions are different from recommendations expressed by *sollte*, which can be determined without reference to a dialogue situation.

Recommendations declare a certain state of affairs as advantageous with respect to some goal. In our scenario, attending a course in a specific semester is advantageous for the completion of the course of study, if its attendance optimizes the number of semesters that are necessary for that completion.

The system checks whether the course can successfully be integrated into the overall plan and computes the corresponding consequences: If attending this course during the current semester does not increase the number of semesters required to achieve the completion within the normal dura-

tion of study, but a different course does, that is also offered for the current semester, then it is advantageous to attend this course, and the modal *sollte* is chosen. In other words, the system compares the consequences of attending courses with respect to the length of study.

#### 4 Future work

The work reported in this paper describes the current state of the CAN-system. Our long-term goal is a wide coverage of modal expressions including all German modal verbs as well as other modals as, e.g., modal nouns, adverbs, and particles. The evaluation of the generated modal expressions is an additional research topic, as well as the interplay of modality with mood and temporality.

Currently, our discourse record is a simplifying collection of propositions, but we will need a more sophisticated version of this knowledge store if modal subordination [Roberts, 1989; Stone, 1999] becomes a topic of our research. Modal subordination becomes relevant in sentence pairs of the following kind:

- (5) *Du könntest den Semantik-Kurs besuchen. Dann musst du im folgenden Semester den Pragmatik-Kurs besuchen.* (You could attend the semantics lecture. In this case you must attend the lecture on pragmatics in the next semester.)

In order to choose *müssen* for the second sentence, the hypothetical event introduced by the first sentence must be considered as background so that ordering sources will become an additional factor for a discourse-sensitive choice of modals.

Finally, the decision to modalize a proposition is often grounded in specific dialogue settings like the existence of social power relations between the participants, or their respective beliefs of the beliefs states of the addressees. We assume that a dialogue situation explains the use of modals in a more natural way than the current model, as our current treatment of permissions already indicates. Therefore, in the long run CAN shall be able to react dynamically to user requests.

#### References

- [Busemann and Horacek, 1997] S. Busemann and H. Horacek. Generating air quality reports from environmental data. *Proceedings of the DFKI Workshop on NLG*, pages 15-21, 1997.
- [Crouch and Pulman, 1993] R.S. Crouch and S.G. Pulman. Time and modality in a natural language interface to a planning system. *Artificial Intelligence*, 63: 265-304, 1993.
- [Frank, 1996] A. Frank. *Context dependence in modal constructions*. Doctoral Dissertation, University of Stuttgart, 1996.
- [Kratzer, 1977] A. Kratzer. What „must“ and „can“ must and can mean. *Linguistics and Philosophy*, 1:337-355, 1977.
- [Milosavljevic, 1997] M. Milosavljevic. Content selection in comparison generation. *Proceedings of the 6th European Workshop on Natural Language Generation*, pages 72-81. Duisburg, Germany, 1997.
- [Roberts, 1989] C. Roberts. Modal subordination and pronominal anaphora in discourse. *Linguistics and Philosophy*, 12: 683-721, 1989.
- [Stone, 1999] M. Stone. *Reference to possible worlds*. Technical report, Rutgers. Available at: <http://www.cs.rutgers.edu/~mdstone/>
- [van der Auwera and Plungian, 1998] Modality's semantic map. *Linguistic Typology*, 2: 79-124, 1998.
- [von Stechow, 2004] A. von Stechow. *Schritte zur Satzsemantik*. Ms., University of Tübingen. Available at: [http://vivaldi.sfs.nphil.uni-tuebingen.de/~arnim10/Lehre/Japan2004/Schritte1\\_17.pdf](http://vivaldi.sfs.nphil.uni-tuebingen.de/~arnim10/Lehre/Japan2004/Schritte1_17.pdf)
- [Zifonun et al., 1997] G. Zifonun, L. Hoffmann and B. Strecker (eds.) *Grammatik der Deutschen Sprache*, Vol. III; Chapter F5: Bedeutung der Modalverben. Berlin: Walter de Gruyter, 1997.