Choosing Rhetorical Structures to Plan Instructional Texts

Leila Kosseim
Guy Lapalme
RALI, Département d'informatique et de recherche opérationnelle
Université de Montréal, CP 6128, Succ. Centre Ville
Montréal (Québec) Canada, H3C 3J7
e-mail:{kosseim, lapalme}@iro.umontreal.ca

Abstract

This paper discusses a fundamental problem in natural language generation: how to organize the content of a text in a coherent and natural way. In this research, we set out to determine the semantic content and the rhetorical structure of texts and to develop heuristics to perform this process automatically within a text generation framework. The study was performed on a specific language and textual genre: French instructional texts. From a corpus analysis of these texts, we determined 9 *senses* typically communicated in instructional texts and 7 *rhetorical relations* used to present these senses. From this analysis, we then developed a set of *presentation heuristics* that determine how the senses to be communicated should be organized rhetorically in order to create a coherent and natural text. The heuristics are based on 5 types of constraints: conceptual, semantic, rhetorical, pragmatic and intentional constraints.

In order to verify the heuristics, we developed the SPIN natural language generation system. SPIN performs all steps of text generation but focuses on the determination of the content and the rhetorical structure of the text.

keywords: natural language generation, text planning, Rhetorical Structure Theory, instructional texts

Introduction

This paper discusses a fundamental problem in natural language generation: how to organize the content of a text in a coherent and natural way. When human writers compose a text, they must select from a wide array of rhetorical means how they will

[0]to appear in *Computational Intelligence* vol. 16, no. 3.

present the information to be communicated. For one type of information, many "points of view" are available, but only a few will actually create the most appropriate expression in the communicative context. Consider, for example, the following sentences:

(1) a. Appuyer sur la touche de réglage de canal (CHANNEL SET) *pour régler le canal de réception*.
Press the channel selector (CHANNEL SET) *to set the reception channel*.

b. Régler le canal de réception *en appuyant sur la touche de réglage de canal (*CHANNEL SET*)*.
Set the reception channel *by pressing the channel selector (*CHANNEL SET*)*.

Semantically, these two sentences present the same information: two operations are involved: setting the channel and pressing the selector, and the first operation is at a higher level of abstraction than the second. However, the first sentence presents the hierarchical relation between the operations as a purpose; while the second sentence presents the relation as a means. In other words, the first sentence answers the question *why should I press the selector?* by presenting the hierarchical relation bottom-up; while the second sentence answers the question *how should I set the channel?* by presenting the relation with a top-down point of view.

In order to generate a coherent and natural text, a text generation system needs to know what rhetorical relations are available to present a particular type of information, and needs to have a principled set of rules to select the most appropriate relation in the communicative context.

Through a corpus analysis of French instructional texts, we identified 9 senses, which make up most of the content of instructional texts, and 7 rhetorical relations used to present these senses. We then developed a set of content heuristics to select the most appropriate senses to be communicated and a set of presentation heuristics to select the most appropriate rhetorical relations to use. Both sets of heuristics are based on conceptual, semantic, rhetorical, and pragmatic constraints, and also take into account the knowledge and intentions of the reader. Finally, we implemented these heuristics in an automatic text generator called SPIN[1]. SPIN performs all steps of the text generation process from the conceptual determination to the lexical and typographical choices; however, the expertise of SPIN is in the semantic and the rhetorical selection.

In this paper, we will only discuss the presentation heuristics, that is, once we have decided what to say, how do we decide how to present it? Section 1 explains the advantages of separating the semantic and rhetorical levels. Section 2 presents the corpus

[1] SPIN stands for "Système de Planification d'INstructions".

used in our analysis. Section 3 discusses the senses identified through the corpus analysis. Section 4 shows the presentation heuristics. Finally, sections 5 and 6 present an overview of the SPIN system and an evaluation of its results.

1 Separating the Semantic and the Rhetorical Levels

Our model of the planning of instructional texts is a 3 step process:

1. We first determine the sequence of operations to execute in order to reach a particular goal; this leads to the conceptual representation of the instructions.

2. We then choose which information will be communicated in the text and which will be left implicit; this leads to the semantic representation of the text and is where the content heuristics come into play.

3. Finally, we choose the rhetorical relations to use; this leads to the text's rhetorical structure and is where the presentation heuristics come into play.

We believe that by separating the semantic and the rhetorical levels, we can achieve a larger rhetorical diversity than if the two questions were not distinguished. The important point is not that the processes be distinct, but that there be room for two distinct levels: the semantic and the rhetorical. For example, although in Moore and Paris (1993)'s text planner both questions are treated within a single component, both semantic and rhetorical operators co-exist. However, in a constructive Rhetorical Structure Theory Mann and Thompson (1988) or a schema based approach to text planning the content determination and rhetorical structure determination are performed simultaneously.

Delin, Scott, and Hartley (1993) have demonstrated that in multilingual instructions, the same information can be conveyed using different rhetorical structures depending on the language of communication. More generally, many researchers have argued that within a unilingual environment the mapping between the semantic and the rhetorical levels is many-to-many (Moore and Pollack 1992; Korelsky and Kittredge 1993). It thus becomes necessary to have linguistically motivated guidelines on how to organize a text's content; that is, how to map a semantic representation onto the most appropriate rhetorical structure.

In French instructional texts, sentences like the following appear quite often:

(2) a. Brancher le cordon d'alimentation du magnétoscope dans une prise secteur 120V et appuyer sur l'interrupteur POWER. *Le voyant* POWER *s'allume et l'horloge*

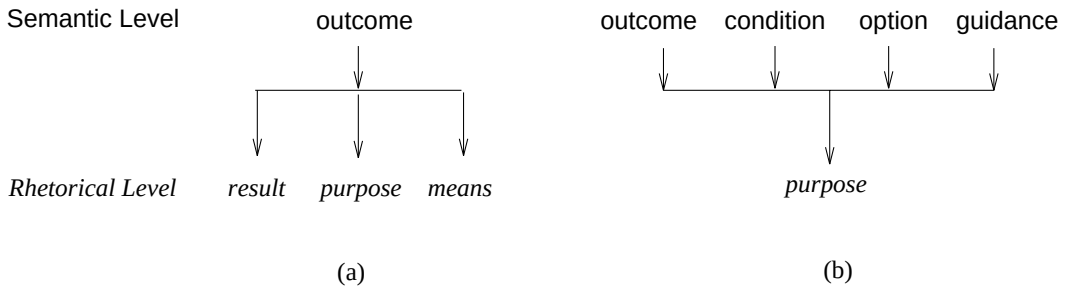


Figure 1: Many-to-many mapping between the semantic and the rhetorical levels

commence à clignoter.
Plug the electrical cord of the video-tape recorder in a 120V outlet and press on the POWER button. *The* POWER *light is turned on and the clock starts to blink.*

b. Revisser l'écrou capuchon sur la lyre *pour ne pas le perdre.*
Screw the nut-cap on the lamp-shade holder *so that you do not lose the it.*

c. *Vous pouvez voir le niveau de volume* en observant la barre rouge sur la gamme de 15 barres affichées sur l'écran.
You can see the volume by observing the red bar on the 15-bar scale displayed on the screen.

In these three examples, the same semantic information is conveyed by the expression in italics: it expresses the OUTCOME[2] of some action. However, these OUTCOMES are communicated through different rhetorical relations. In the case of (2a), a *result* is used; in (2b) a *purpose* is used; and in (2c) it is the nucleus of a *means* relation. This mapping is shown in Figure 1a.

Inversely, one rhetorical relation can be used to convey different semantic information. For example, the relation of *purpose* in (2b) communicates the OUTCOME of an action; while in (3a) it communicates a CONDITION on an action; in (3b), it communicates the OPTIONAL NATURE of an action; and finally in (3c) it communicates a GUIDANCE[3] on how to perform an action. This is illustrated in Figure 1b.

(3) a. *Pour [vérifier] un commutateur ordinaire [. . .]*, touchez la vis de la borne de cuivre avec la pince du vérificateur.
To [check] an ordinary switch [. . .], touch the screw of the copper terminal with the clip of the tester.

[2]In this paper, senses are written using SMALL CAPS while rhetorical relations are written using *italics*.

[3]A GUIDANCE is an instance of what Di Eugenio and Webber (1996) call pragmatic overloading.

b. Tirer la roue et le pneu; *pour vous aider*, poussez fermement le flanc du pneu avec votre pied.
Pull the wheel and the tire; *to ease the task*, firmly press on the side of the tire with your foot.

c. Tourner cette touche à droite et à gauche *pour minimiser les parasites*.
Turn this knob clockwise and counter-clockwise *to minimize interference*.

These examples illustrate the many to many mappings between the semantic and the rhetorical levels. But once the rhetorical level is reached, the question of choice is certainly not over. A particular rhetorical relation can be presented through different grammatical forms. For example, a *purpose* relation can be presented by an infinitive clause, as in examples (3b) and (3c) or a preposition followed by a goal metonymy as in (3a) or even a subordinate with a subjunctive verb (e.g. *pour que la syntonisation soit précise* / so that the tuning be precise).

Before describing the presentation heuristics, we will first describe the corpus study.

2 The Corpus

Instructional texts exhibit certain characteristics that make them interesting for natural language processing, and particularly for natural language generation. Instructional texts are widely available and understandable by many readers and are usually well structured. Compared to other narrative texts, they can be represented more objectively. This representation is generally viewed by means of relations between states and operations. More specifically, it is commonly and productively viewed as a hierarchy of operations (Dixon 1987; Dixon, Faries, and Gabrys 1988; Donin, Bracewell, Frederiksen, and Dillinger 1992) that can be built using an automatic planner. Typically, instructional texts also include non-procedural information like details of an instrument, but in this paper we focus on the problem of generating procedural information that can be modeled by a hierarchy of operations.

We analyzed a corpus of French instructional texts in order to capture natural linguistic phenomena without over-simplifying them. Although our study was done on French texts, we believe that our results are also applicable to English in the sublanguage of instructional texts. As Kittredge (1982) noted:

> *One is therefore drawn to conclude that English and French technical texts show the strongest parallels because the text purpose is more similar here than in descriptive texts. Weather reports, recipes and aviation manuals, which show the strongest parallels, all have very well-defined text purpose.* (p. 135)

Type	Domain	Length (words)	Lexical specialization	Typo-graphical cues	Required judgment	Required instruments
	emergency respiratory care	50	low	strong	average	low
	cooking recipes	370	low	strong	low	average
	assembling a dresser	270	low	strong	average	average
Execu-	using a video tape recorder	55	low	strong	low	low
tion	using a television	70	low	strong	low	low
	fixing electrical appliances	100	low	strong	average	low
	car maintenance	100	low	strong	low	strong
	wine-making	250	strong	average	strong	strong
Compre-	glass painting	550	average	low	strong	strong
hension	glass blowing	120	strong	low	average	strong
	restoration of antique furniture	820	average	low	average	average
	Mazda car manual	120	average	strong	low	low
	techniques of rotin	255	average	average	average	average
Hybrid	techniques of photography	200	average	average	average	average
	organizing a hard drive	200	average	average	low	low

Table 1: Characteristics of the corpus

This can also be confirmed in this paper where we have given a literal translation (in this type) after each example from our corpus. These translations are almost always acceptable English instructions. On the other hand, other research has shown preferences in rhetorical relations according to the language studied (Delin, Scott, and Hartley 1996).

To describe the natural process of generating instructions, we have taken a corpus of texts and analyzed it both at the semantic and the rhetorical levels. When selecting the corpus, we took a few precautions: we rejected what seemed to us the "obviously flawed" texts and we only considered original French texts or translations that did not seem "biased" by their original language. We collected instructional texts from 15 different sources from every-day appliances and how-to books. The corpus is made up of 79 procedures (≈13,300 words) that have different communicative goals, domains and intended readers. When one procedure was divided into sub-procedures, we counted it as one a procedure for reaching a single goal (e.g. oil changing in a car). If the procedure had a less precise goal (e.g. using a remote control) then each procedure was counted separately. We analyzed only the procedural parts of these texts. Table 1 lists the texts of the corpus along with some textual, conceptual, lexical and typographical characteristics. References to the original texts are given in (Kosseim 1995). The corpus was divided into 3 classes according to the communicative goal of the texts.

Execution texts are characterized by their simplicity and are aimed at an immediate execution. These texts are typically short, have a low level of lexical specialization, have strong typographical cues (e.g. bullets, alert icons, ...) and are characterized by their conceptual simplicity. They require little judgment and non-specialized instruments.

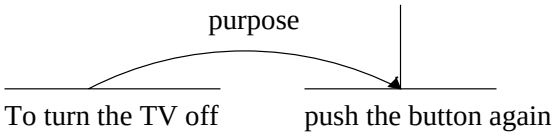


Figure 2: An RST relation

Comprehension texts are characterized by their complexity and are aimed at a possible execution; their goal is to explain, not to tell. These texts have longer procedures, use specialized terminology, have fewer typographical cues and generally require more judgment, and many and specialized instruments.

Hybrid texts exhibit characteristics of both execution and comprehension texts.

The corpus is rather varied with respect to the instructional domain and the communicative goals. The analysis of a less diversified corpus would certainly have permitted of more precise results, but their applicability to other instructional domains would not have been possible. Following Mellish (1988), our aim is to define general rules for the generation of instructions in any domain, then particular rules for adapting the text to a specific domain and reader.

Instructional texts are characterized by stereotypical relations between parts of the text. Like (Rösner and Stede 1992; Vander Linden 1993), we used Rhetorical Structure Theory (RST) described by Mann and Thompson (1988) as a descriptive tool for the rhetorical analysis of our corpus. RST was developed to identify relations between adjacent portions of text and is defined in terms of:

relations between a main part, the *nucleus* and an auxiliary part, the *satellite*.

schemata that specify the structural composition of a text.

Figure 2 shows the schema corresponding to a *purpose* relation. Here, the satellite *To turn the TV off* is linked to the nucleus *push the button again*. In a RST schema, the horizontal lines span the linguistic expression, the vertical line shows the nucleus and the arc between them indicates a relation between them.

3 The Semantic Content

By viewing the conceptual representation of an instructional text as a hierarchy of plans, we realized that not all the procedural information available in the conceptual

representation is given in the text. Inversely, the text often contains information that was not represented in a hierarchy of plans. Therefore, the goal of the semantic analysis of our corpus was to find out:

1. What type of information (or *sense*) is typically communicated in instructional texts.

2. What constraints influence the inclusion (or exclusion) of this information in instructional texts.

Through the corpus analysis, we identified 9 *senses* that identify the meaning of textual expressions[4]. They are called *senses* rather than *semantic elements* or *semantic relations* because although most refer to elements of meaning that are mapped onto the satellite of an RST relation (Mann and Thompson 1988); some refer to relations whose nucleus and satellite will be determined at the rhetorical level when the most appropriate relation is chosen. To define the senses, we have been inspired by the semantic relations of *type generation*, *instance generation* and *enablement* of Goldman (1970) which have been used by Delin, Scott, and Hartley (1996) as the basis for the generation of multilingual instructional texts. These relations, however, do not cover the whole range of senses found in instructional texts; thus, we have developed more specific entities particular to this discourse genre. The senses have been identified only through functional criteria and, thus, are independent of their syntactic form. Figure 3 shows the criteria we used to identify the senses.

The procedural nature of the information: This first criteria has been chosen as the most discriminating factor because the basic purpose of an instructional text is to present a procedure to be followed. Most expressions will convey procedural information, but some expressions are used to reinforce procedural information with background information. As these expressions were not the main focus of our study, any such expression is classified as an ATTRIBUTE.

The state/operation dichotomy: As one describes a procedure, one invariably speaks of states and operations. This distinction is a delicate one at the deep semantic level. Because a state is brought about by an operation, one can argue that anything at the semantic level can be characterized as an operation. However, in AI planning (our level of conceptual knowledge), the distinction between states and operations is made because an operation is defined as a function between 2 states. Furthermore, although a state is brought about by an operation, in the context of instructional texts, it may not always be possible to determine what operation

[4]The granularity of the analysis is consistent with a typical RST analysis.

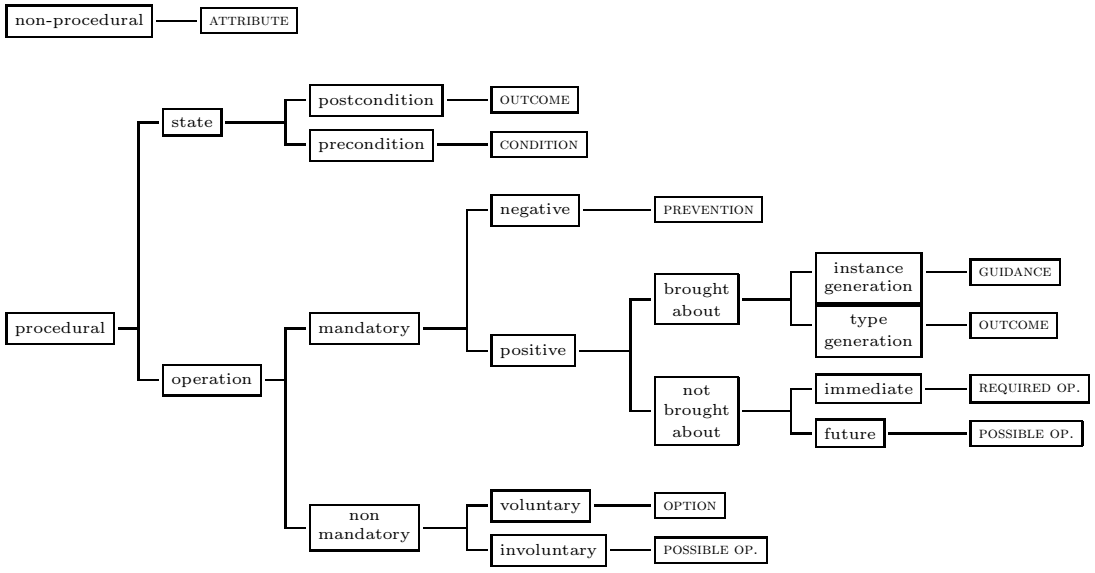


Figure 3: Functional criteria used to identify the senses

has brought about a certain state; in addition, even if this operation could be determined, it may be of no interest to the procedure to be described. For example, in the expression *if you have a front-wheel drive*, it is difficult to determine what operation has brought about the type of drive (the owner bought it? the manufacturer built it? ...). Even if this operation could be determined, it would be of no interest to the writing of the procedure. In our study we classified any procedural sense executed by the agent exclusively for the procedure at hand as an operation; anything else is considered to be a state.

The pre- and post-condition nature: States can be further classified depending on their relation with the operation they relate to. Pre-conditions to operations are called CONDITIONS, while post-conditions are called OUTCOMES.

As the primary goal of an instructional text is to indicate to the reader the steps to take to achieve a goal, most of the senses pertain to operations. These are classified along the following criteria:

The mandatory and voluntary nature: Some operations are mandatory to the successful achievement of the procedure; while other operations are not. If an operation is not mandatory, the decision to execute it can be made explicitly by the agent or can be performed involuntarily (e.g. by accident). In the first case, the

sense that is conveyed is called an OPTION, in the second case, it is a POSSIBLE OPERATION.

The polarity: Most operations indicate what to do, but some indicate what *not* to do. They warn the reader not to do something. Mandatory negative operations are called PREVENTIONS.

The type of relation between operations: As operations are the most frequent expressions in instructional texts, one can further characterize them by the relation that holds between them. Common relations in instructional texts are Goldman (1970)'s instance and type generation. If an operation O_1 always generates an operation O_2, independently of how O_1 is performed, then there exists a type generation between the two operations and O_2 is said to be an OUTCOME of O_1. If, on the other hand, O_1 generates O_2 only if O_1 is performed in a very specific manner, then there exists an instance generation between the two operations and O_1 is said to be a GUIDANCE of O_2 (O_1 guides the correct execution of O_2).

The time and relative order of execution: Finally, an operation can be executed at the same time as the rest of the procedure or at a much later time. In the latter case, the sense conveyed is a POSSIBLE OPERATION; in the first case (more frequent) a further distinction is made between concurrent and individual execution of operations.

These criteria allowed us to identify 9 different senses:

ATTRIBUTES OF OBJECTS do not participate in the procedural part of the instructions but rather give some background information to the reader. For example:

(4) *Une lampe à une seule douille comporte habituellement une lyre qui tient à une barre de retenue par des manchons.*
A single-socket lamp usually has a shade holder attached to a retaining bar by a few fittings.

REQUIRED OPERATIONS describe the steps to follow in order to execute the instructions. For example:

(5) *Placez une disquette MS-DOS dans le lecteur [A].*
Place an MS-DOS disquette in drive [A].

CONDITIONS describe a situation that must be true before an event can take place. For example:

(6) *L'enfant a moins de 2 ans ?* Placez-le sur votre avant-bras ou sur votre cuisse.
The child is less than 2 years old ? Put him on your forearm or on your thigh.

OUTCOMES show what happens after an event has taken place. For example:

(7) Appuyer sur la touche PLAY. *La lecture normale apparaît.*
Touch the PLAY button. *Normal reading appears.*

OPTIONS are operations that may be performed but are not necessary. For example:

(8) *Pour monter un dispositif de sécurité pour enfants*, [...] immobilisez le dispositif de sécurité avec la ceinture ventrale.
To mount the child safety system, [...] secure the safety mechanism with the belt.

GUIDANCES show how or why an operation should be executed while guiding its execution. A guidance is an instance of what Di Eugenio and Webber (1996) call pragmatic overloading. For example:

(9) Tourner cette touche à droite et à gauche *pour minimiser les parasites*.
Turn this knob clockwise and counter-clockwise *to minimize interference.*

CO-TEMPORAL OPERATIONS indicate two or more operations that should be executed simultaneously. For example:

(10) *Effleurer la surface du verre, tout en le faisant tourner.*
Touch lightly the glass surface, while making it turn.

PREVENTIONS indicate operations that should not be performed. They are realized grammatically by *preventative expressions* (Vander Linden and Di Eugenio 1996).

(11) Utilisez une clé en croix pour desserrer les écrous de la roue [...] *N'enlevez pas complètement les écrous.*
Use a cross key to loosen the nut of the wheel [...] *Do not completely remove the nuts.*

POSSIBLE OPERATIONS indicate operations that might be executed by mistake or without any awareness from the user; they are given to explain a behavior that might occur but is not in the normal sequence of operation. Contrarily to PREVENTIONS, they do not necessarily result in a undesirable side-effect.

(12) *Si vous tentez d'accéder au disque "C >"*, vous n'y parviendrez pas.
If you try to access the "C >" disk, you will not succeed.

The decision tree of figure 3 is generally sufficient to uniquely identify the sense conveyed by an expression. However, in some cases the context or knowledge of the reader is not sufficient to verify if a criterion is satisfied or not. Consider, for example, the following sentence:

(13) *Pour pallier à cet inconvénient*, peindre ces surfaces transparentes avec les couleurs à l'huile et attendre qu'elles sèchent.
To avoid this inconvenience, paint these transparent surfaces with the oil-based colors and wait until they dry out.

Here, it is not clear from the context if *to avoid this inconvenience* is mandatory (e.g. *you must avoid*) or optional (e.g. *if you wish to avoid*). In the first case, we are dealing with an OUTCOME, in the second case an OPTION is specified. In situations of ambiguity, we choose the strongest interpretation in context, in this case, an OUTCOME.

We do not claim that these senses cover every type of information encountered in instructional texts. Table 2 shows that out of 1471 expressions analyzed, 12 could not be classified by our criteria. We do not claim either that this sense classification is the only or best way to describe the informational content of instructional texts. The semantic criteria could possibly be further specified, other criteria could be taken into account and other combinations of criteria could be used to create new senses. We do claim though, that this classification has a coverage large enough for our corpus (≈99%), and the granularity is appropriate for the rest of the analysis. This set of senses should be seen as an open set, subject to extensions and modifications to fit the corpus and granularity of the analysis.

Table 2 shows the number of occurrences of the senses in our corpus. From this table, we can see that about half the content of the texts (52%)[5] are made up of REQUIRED OPERATIONS; this is not surprising as the main goal of an instructional text is to indicate to the reader what actions to take to achieve some goal. Table 2 also shows the frequency of the senses by types of texts (execution, comprehension and hybrid). We can see that senses are influenced by the type of text. For example, execution texts contain far fewer ATTRIBUTES but more REQUIRED OPERATIONS than comprehension texts. These results are not surprising, considering that an execution text tells what to do; while a comprehension text explains. More details about the semantic analysis of our corpus can be found in Kosseim (1995).

[5]In this paper, all percentages have been rounded off, so their sum can differ from 100.

	Entire corpus		Execution Texts	Hybrid Texts	Comprehension Texts
sense	**Number of occurrences**	**%**	**%**	**%**	**%**
ATTRIBUTE	158	11	3	17	95
REQUIRED OPERATION	762	52	65	40	29
CONDITION	164	11	11	12	9
OUTCOME	136	9	7	13	9
GUIDANCE	124	8	9	8	8
CO-TEMPORAL OPERATION	45	3	1	4	7
OPTION	34	2	2	3	3
PREVENTION	21	1	1	2	2
POSSIBLE OPERATION	15	1	1	1	2
OTHER	12	1	0	0	5
Total	1471	≈ 100	100	100	100

Table 2: Frequencies of senses

Once a sense has been chosen to be communicated, it may be expressed from a given point of view, but the mapping between these two levels is not direct as we showed in section 1. The next section describes the presentation heuristics for choosing the best rhetorical relations in a given context.

4 The Rhetorical Structure

As instructional texts exhibit a rather stereotypical structure, the set of rhetorical devices used in this genre is rather limited compared to the whole spectrum of RST relations. However, the advantage of RST, is that the list of relations initially defined do not form a closed set; it is susceptible to extensions and modifications to fit a particular textual genre.

Following this idea, Vander Linden (1993) and Rösner and Stede (1992) have identified the rhetorical relations typically used in instructional texts. For Vander Linden, the most important RST relations are *temporal sequence*, *precondition* (which we call *c-condition*), *purpose*, *result* and *concurrence*. Rösner and Stede identified other relations: *until* and *alternative* for which we found very few occurrences (2%) so we did not take them into account. We also decided to combine their *step-sequence* with the usual *sequence* as the distinction is more typographical than rhetorical.

We now give more details on the rhetorical relations we kept in our study. This list of relations was necessary in our analysis, but cannot be considered sufficient. In our analysis, these relations made up about 98% of the expressions. But, similarly to RST's relations, this list should be seen as an open set.

sequence is a multinuclear relation (that is, a relation where no component is a subor-

dinate of another) where there exists a succession relation between the situations conveyed by the nuclei.

c-condition presents an unrealized situation that must be realized in order to make it possible or sensible to carry out the nucleus (Rösner and Stede (1992)). This relation combines features of RST's relations of *circumstance* and *condition*. Rösner and Stede call this new relation a *precondition* but in our study, this term is already used in its AI planning definition. As Rösner and Stede point out, this new relation is more adapted to instructional texts, where some expressions do not really fit RST's definition of *circumstance* or *condition*.

elaboration present additional detail about the nucleus, for example, the relations of set:member, abstract:instance, ...

purpose presents a situation to be realized through another situation presented in the nucleus. Vander Linden (1993) does not distinguish this from the *means*. For him, the propositions:

(14) a. Do O_1 by doing O_2.

b. To do O_1, do O_2.

present the hierarchical relation between O_1 and O_2 through the same point of view. Following Mann and Thompson (1988), we prefer to distinguish these relations. In (14a), a *means* is used because by doing activity presented in the satellite, the reader is better equipped to do activity presented in the nucleus; while in (14b), a *purpose* is used as it presents a situation to be realized through the satellite.

means presents a situation which allows the reader to do the activity presented in the nucleus.

result presents a situation (in the nucleus) that may have caused the situation presented in the satellite. The situation in the nucleus is more central to the writer's purpose than the situation in the satellite; this is what distinguishes this relation from a *cause*. The relation of *result* includes RST's volitional and non-volitional results. The distinction between those two types of results is not central in instructional texts, and is typically not made.

action concurrency is multinuclear like the sequence, but there exists a simultaneous relation between the nuclei.

Table 3 shows the results of the rhetorical analysis; while the mapping between senses and rhetorical relations in our corpus is shown in Table 4. As we can see, almost

Rhetorical relation	Number of occurrences	%
sequence	770	52
c-condition	172	12
elaboration	170	12
purpose	118	8
result	98	7
means	97	7
concurrency	45	3
Total	1471	≈ 100

Table 3: Frequency of rhetorical relations

all REQUIRED OPERATIONS are presented within a *sequence* (98% of the time), but an OUTCOME can be presented through 3 different rhetorical relations: a *purpose* (29%), a *result* (68%) or a *means* (4%). This study enabled us to determine heuristics for choosing appropriate rhetorical relations to express the senses. These heuristics are based on several factors given in the next section.

4.1 The Communicative Context

We stated in the introduction that choosing the most appropriate rhetorical relation to present a given sense depends on the communicative context. But what exactly do we mean by that? During the corpus analysis, we took into account five types of constraints that together define the communicative context that influence the choice of a rhetorical relation. We have found these types of constraints necessary for the rhetorical selection in general[6]. This set of constraints was sufficient for our study; however, as this study deals with an unclear and somehow subjective subject manner, we believe that a finer-grained or more linguistically-motivated analysis may find these criteria insufficient. Although our corpus did not show evidence of this, grammatical or lexical realizations may play a role in rhetorical choice.

Conceptual constraints: As many studies in psychology (Dixon 1987; Dixon, Faries, and Gabrys 1988; Donin, Bracewell, Frederiksen, and Dillinger 1992) suggest, the content and structure of the conceptual representation of a procedure should be taken into account in order to construct instructions that are easy to interpret. An important conceptual criterion is the notion of *basic level operations* which can informally be defined as the operations that are specific enough to be descriptive, yet general enough to be meaningful. More specifically, according to Rosch (1978), basic-level operations seem to be operations that are most easily remem-

[6]But not all types of constraints are necessary for every selection (see section 4.2.2).

Name of heuristic	Sense	Rhetorical Relation						
		sequence	*c-condition*	*elaboration*	*purpose*	*result*	*means*	*concurrency*
AT	ATTRIBUTE			100%				
RO	REQUIRED OPERATION	98%	1%		1%			
CD	CONDITION	2%	90%		4%	4%		
OU	OUTCOME				28%	68%	4%	
GD	GUIDANCE				31%		69%	
CO	CO-TEMPORAL OPERATION							100%
OP	OPTION		21%		79%			
PR	PREVENTION	86%					14%	
PO	POSSIBLE OPERATION		73%				27%	

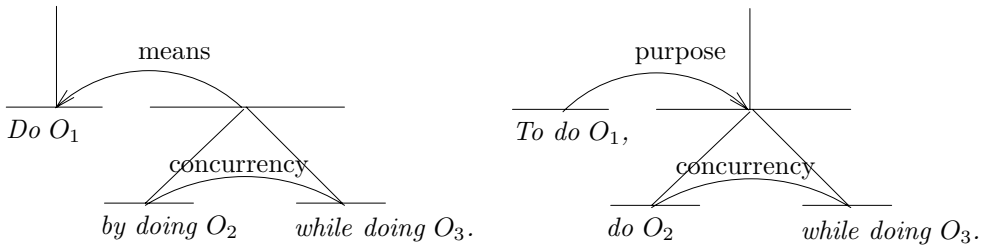
Table 4: Global mapping between senses and rhetorical relations in the corpus

bered. Along the same lines, Pollack (1986) defines the notion of domain-basic act-types. For Pollack, *typing a character on a keyboard* cannot be considered a basic-level operation in the computer domain because it is too specific; just like *typing a string of characters* or typing anything at all. In this domain, a basic-level operation would be *issuing a command*. Pollack stipulates that within a discourse domain there exists a set of basic-level operations, and these may be agent-specific. We use the notions of basic-level operation of Rosch and Pollack, to explain the communication of certain operations in instructional texts. Indeed, basic-level operations are included in the text because writers take for granted that readers have an easily accessible mental representation of them and because they ease the reader's recall since the readers can easily construct a memory representation of them.

Semantic constraints: The most important factor in determining what rhetorical relation to use is what sense we wish to convey. For a particular sense, only a subset of the rhetorical relations are acceptable (see Table 4). For example, a CONDITION cannot be conveyed through a *means* or a *concurrency*.

Rhetorical constraints: The choices of rhetorical relations are co-dependent. That is, the selection of a rhetorical relation to present some sense $S1$ will impose constraints on the presentation of another sense $S2$. This is because the senses are generally mapped into a portion of a rhetorical relation (the satellite or the nucleus). In order to have a well-formed relation, and therefore a coherent text, we must ensure that all portions of the relation will be filled by some sense. Once a relation $R1$ is chosen to convey a sense $S1$, part of $R1$ will be filled by $S1$, and another sense, $S2$, will have to fill the remaining part[7].

[7]In the case of multi-nucleic relations like *sequence* at least 2 nuclei must be filled.



a. Unlikely structure

b. Preferred structure

Figure 4: Preferred rhetorical structure

When looking for a filler for the other portion of the rhetorical relation, it should be kept in mind that in instructional texts, some rhetorical relations seem to co-occur while some other combinations are unlikely. For example, if two CONDITIONS are to be presented to constrain the same operation, we can use a *c-condition* for the "easiest" one to verify and a *result* to convey the other. For example:

(15) *S'ils [les écrous] portent la marque "L", ils ont le filetage à gauche* et vous devez les dévisser ....
If they [the nuts] have an "L" mark, they have a left-hand thread, and you must unscrew them ...

However, it is unlikely to find a *concurrency* related to its nucleus by a *means*. This unlikely form[8] is shown in:

(16) * Do O_1 by doing O_2 while doing O_3.

and in Figure 4a. To convey the same information, a *purpose* related to a *concurrency* is preferred, as in:

(17) To do O_1, do O_2 while doing O_3.

and in Figure 4b. Compared to (16), (17) states more clearly the relation between O_1, O_2 and O_3. In our study, we only observed a preference of certain combinations of relations based on their number of occurrences. Although the question is very interesting, we did not perform a linguistically-motivated analysis on why certain combinations of relations are more felicitous than others.

[8]By notation, awkward examples are preceded by an asterisk.

Pragmatic constraints: These constraints take into account specifications of the nature of the procedure (i.e. pragmatic characteristics of the operations and states of the procedure) to select a rhetorical relation. This includes the optionality and degree of desirability of an operation (if an optional line of operations is generally desirable, it will be conveyed differently than one rarely chosen), the level of danger of a negative operation and the internal/external status of states. Consider, for example, the sentences:

(18) a. Turn off the car engine; *this will turn off the radio.*

b. * *To turn off the radio*, turn off the car engine.

c. * Turn off the radio *by turning off the car engine.*

While example (18a) seem natural, examples (18b) and (18c) seem awkward. This is because the *purpose* and the *means* relations imply that the method prescribed in the instruction is the normal procedure to achieve the goal. In the preceding examples, turning off the car engine is not the normal method to turn off a car radio. In this case, the nature of the relation between these two operations would rule out the selection of a *purpose* or a *means* relation, and leave only the relation of *result* as a possible choice.

Intentional constraints: What a "generic" reader believes about the operations and states of the procedure and his pursued goals greatly influences how information is conveyed in the text. This is why a model of the imagined reader's knowledge and intentions must be taken into account to generate appropriate relations. For example, in:

(19) a. *Si l'on souhaite une ligne plus large*, alors s'attarder sur le verre de façon à laisser s'écouler plus de couleur.
If you wish a thicker line, stay on the glass longer so that more paint can flow.

b. *Pour une ligne plus large*, alors s'attarder sur le verre de façon à laisser s'écouler plus de couleur.
To have a thicker line, stay on the glass longer so that more paint can flow.

If the two expressions in italics present the sense of OPTION, a relation of *c-condition* (19a) is preferred for novice readers as the optional aspect is explicit. A relation of *purpose* (19b) does not convey the optionality as explicitly.

A sixth type of constraint that may be taken into account is the communicative goal of the text. That is, how does the communicative goal influence how a given sense is

Sense	**Rhetorical Relation**														
	% sequence			*% c-condition*			*% elaboration*			*% purpose*			*% result*			*% means*			*% concurrency*		
	ex	hy	co	ex	hy	co	ex	hy	co	ex	hy	co	ex	hy	co	ex	hy	co	ex	hy	co
ATTRIBUTE							100	100	100												
REQUIRED OP.	98	98	100	1	1	0				1	1	0									
CONDITION	3	0	0	90	92	86				3	8	0	3	0	14						
OUTCOME										20	28	27	77	70	64	3	2	9			
GUIDANCE										37	27	10				61	73	90			
CO-TEMPORAL OP.																			100	100	100
OPTION				19	27	14				81	73	86									
PREVENTION	100	88	50													0	12	50			
POSSIBLE OP.				100	100	20										0	0	80			

Table 5: Mapping between senses and rhetorical relations in execution, hybrid and comprehension texts. For example, an OPTION is mapped 19% of the time to a *c-condition* in an execution text, 27% of the time in an hybrid text and 14% of the time in a comprehension text.

presented? Hartley and Paris (1996) have found that the communicative goal influences the linguistic realization of instructions in French; but does it also influence the rhetorical choice? Table 5 shows how the senses in our corpus are presented in the 3 types of texts analyzed: execution, hybrid and comprehension texts. The figures in this table have been compiled by the first author using the same categorization policies in cases of ambiguity as in the general rhetorical analysis. The table shows that most senses are not influenced by the communicative goal of the text. For example, REQUIRED OPERATIONS are presented by a sequence 98% of the time in execution texts and hybrid texts and all the time in comprehension texts. However, the senses of GUIDANCE, PREVENTION and POSSIBLE OPERATION seem influenced by the textual type; but in the case of PREVENTION and POSSIBLE OPERATION, the number of occurrences in our corpus is too low to verify or contradict this hypothesis. OPTIONS and CONDITIONS seem to behave in a rather unpredictable manner in hybrid texts as compared to other types of texts. OPTIONS, for example, seem to use many more *c-conditions* rather than *purposes* in hybrid texts; a phenomenon that we cannot explain satisfactorily with the current results. The conclusion we drew from these figures, was that a larger and more formal corpus analysis is necessary to determine the influence of the communicative goal on the selection of rhetorical relations. This includes the analysis of many more texts (currently 8, 4 and 4), the use of many more annotators (currently 1) and the decomposition of the term *communicative goal* into more measurable factors. In light of this conclusion, we did not take into account the influence of the communicative goal in the current study.

In the next section, we will discuss the presentation heuristics based on the criteria we identified in this section.

4.2 The Presentation Heuristics

This section presents the criteria for selecting the most appropriate rhetorical relation to communicate a given sense. The following rules are given in order of preference, i.e. as soon as the communicative context satisfies a given criterion, the corresponding rhetorical relation is chosen. These heuristics are implemented in SPIN, presented in the next section. As the justification of the heuristics for each sense is rather long, this section only shows the justification for REQUIRED OPERATIONS and enumerates the heuristics for the other senses. The justification of the heuristics for the rest of the senses are given in the Appendix.

4.2.1 Required Operations

An REQUIRED OPERATION can be presented by 3 rhetorical relations:

- a *sequence* of actions (98 % of the time):

  (20) *Mettre le magnétoscope sous tension* et *placer le sélecteur* TV/VCR *sur* "VCR".
  Turn the recorder on and *place the* TV/VCR *selector on* VCR.

- a *purpose* (1 %):

  (21) Vérifiez le commutateur de la douille ou réinstallez la douille *pour vérifier le commutateur du socle.*
  Check the socket switch or install the socket again *to check the switch on the base.*

  In the above example, the two operations should be performed in sequence, but the second is rhetorically demoted to the satellite of a *purpose* relation. Example (21) could be re-formulated as:

  (22) Vérifiez le commutateur de la douille ou réinstallez la douille *et ensuite vérifier le commutateur du socle.*
  Check the socket switch or install the socket again *and then check the switch on the base.*

- a *c-condition* (1 %), as in examples (23a) and (23b),

(23) a. Retrait de la lyre:

[...] sur certaines lampes, il faut enlever la douille *avant de soulever la lyre.*

Dégagement de la douille:

[...]

Removal of the lamp-shade holder:

[...] on certain lamps, the socket must be removed *before lifting the lamp-shade holder.*

Removal of the socket

[...]

b. Appuyer sur la touche OTR pour specifier l'heure d'enregistrement.
Lorsque la touche est enfoncée une fois, 'PM 10:35 (30 min)' sera affiché.
Lorsque la touche est enfoncée deux fois, 'PM 11:05 (1h)' sera affiché.
Press the OTR button to specify the recording time.
When the button is pressed once, 'PM 10:35 (30 min)' will be displayed.
When the button is pressed twice, 'PM 11:05 (1h)' will be displayed.

The presentation heuristics for REQUIRED OPERATIONS are:

RO-1 [9] A *c-condition* is used in 3 distinct cases:

RO-1a:

semantic constraint: Two REQUIRED OPERATIONS RO_1 and RO_2 are to be communicated.

conceptual constraint: RO_1 should be executed before RO_2.

rhetorical constraint: none

pragmatic constraint: none

intentional constraint: The reader believes that RO_1 should be executed after RO_2.

This is the case in (23a) where *the removal of the lamp-shade holder* and *the removal of the socket* are, most of the time, to be executed in that order (as indicated by the titles of the procedures) but on some lamps the same operations must be executed in the reverse order. If the order of execution is reversed, a *c-condition* is chosen for one of the operation (either RO_1 or RO_2) instead of a *sequence* because it puts the emphasis on the temporal order of

[9] For easier reference, each heuristic is identified by a distinct name like RO-1.

operations. We have not analyzed which of the two operations is typically demoted to the satellite position; in the implementation of the heuristics, we always use the preposition *avant* (before) and demote the operation that must be executed second.

RO-1b:

semantic constraint: A REQUIRED OPERATION *RO* is to be communicated.

conceptual constraint: The agent *of RO* is not necessarily the agent of rest of the procedure.

rhetorical constraint: A *result* or a *sequence* will be related to *RO* in the text to ensure it a nucleus.

pragmatic constraint: none

intentional constraint: none

In this case, *RO* becomes the satellite of a *c-condition* because this relation does not mention explicitly or implicitly, who must perform the operation. In addition, we say that a *result* or a *sequence* must be related to *RO* in the text in order to ensure a nucleus for the *c-condition*; as in:

(24) a. Lorsque *RO* est exécuté, ceci se produira.
When *RO* is completed, this will happen.

b. Lorsque *RO* est exécuté, faites RO_2.
When *RO* is completed, do RO_2.

RO-1c:

semantic constraint: A REQUIRED OPERATION *RO* is to be communicated along with its OUTCOMES OU_i.

conceptual constraint: *RO* is divided into a set of sub-operations RO_i who have postconditions OU_i.

rhetorical constraint: The OUTCOMES OU_i will be communicated in the text by *results* to ensure a nucleus for RO_i.

pragmatic constraint: none

intentional constraint: none

In this case, even if the reader knows how to perform *RO*, its sub-operations RO_i must be included in the text in order to have a nucleus for the *results* OU_i. These sub-operations RO_i become satellites of temporal *c-conditions*[10]. This is the case in (23b), where the sub-operations *press the button once* and *press*

[10]A *means* could also be used, but in this case, the presentation heuristics give preference to a *c-condition* (see the appendix).

the button twice should normally not appear because they are too primitive to be communicated to the user as part of the operations to perform. They are nevertheless communicated in the text for rhetorical reasons, not mainly for semantic ones although we could argue that this redundant information could be useful.

RO-2 A *purpose* is used if:

semantic constraint: Two REQUIRED OPERATIONS RO_1 and RO_2 are to be communicated.

conceptual constraint: RO_1 is a basic-level operation; while RO_2 is a precondition to RO_1.

rhetorical constraint: RO_2 will be presented within a *sequence*.

pragmatic constraint: RO_2 is a modifiable condition.

intentional constraint: none

In that case, RO_1 will become the satellite of a *purpose* relation. As defined by Vander Linden (1993), a condition is modifiable if the agent must check it or perform an operation so that it becomes true. This is the case in (21) where, although, there is no hierarchical relation between the two operations *check the socket switch or install the socket again* and *check the switch on the base*, they are related rhetorically by a purpose, but semantically by a sequence.

RO-3 In all other cases, a REQUIRED OPERATION RO is presented as a nucleus of a *sequence*; as in (20).

4.2.2 Summary of the Presentation Heuristics

Tables 6 though 10 provide a summary of the presentation heuristics. The reader interested in their justification is invited to consult the Appendix. CO-TEMPORAL OPERATIONS are always presented through a *concurrency* relation; while ATTRIBUTES OF OBJECTS are always presented through an *elaboration* relation. The heuristics of section 4.2.1 and tables 6–10 have been implemented in the SPIN system described in the next section.

5 The SPIN system

In order to verify our analysis, we have implemented the SPIN Natural Language Generation system. SPIN performs all levels of text generation going from a conceptual

Common semantic constraint: A REQUIRED OPERATION RO is to be communicated.
RO-1a *c-condition* semantic: Another REQUIRED OPERATION RO_2 is to be communicated. conceptual: RO should be executed before RO_2. intentional: The reader believes that RO should be executed after RO_2.
RO-1b *c-condition* conceptual: The agent $of RO$ is not necessarily the agent of rest of the procedure. rhetorical: A *result* or a *sequence* will be related to RO in the text to ensure it a nucleus.
RO-1c *c-condition* semantic: The OUTCOMES of RO (OU_i) are to be communicated. conceptual: RO is divided into a set of sub-operations RO_i who have postconditions OU_i. rhetorical: OU_i will be communicated by *results* to ensure a nucleus for RO_i.
RO-2 *purpose* semantic: Another REQUIRED OPERATION RO_2 is to be communicated. conceptual: RO is a basic-level operation; while RO_2 is a precondition to RO. rhetorical: RO_2 will be presented within a *sequence*. pragmatic: RO_2 is a modifiable condition.
RO-3 *sequence* otherwise

Table 6: Presentation heuristics for REQUIRED OPERATIONS – see section 4.2.1

Common semantic constraint: A CONDITION CD is to be communicated.
CD-1a *c-condition* semantic: An equivalent CONDITION CD_2 is to be communicated. pragmatic: CD_2 is more difficult to evaluate than CD. CD will be presented by a *c-condition*.
CD-1b *result* semantic: An equivalent CONDITION CD_2 is to be communicated. pragmatic: CD_2 is easier to evaluate than CD. CD will be presented by a *result*.
CD-2 *sequence* semantic: CD is not related to the operation it constrains by an enablement. pragmatic: CD is a modifiable condition.
CD-3 *purpose* or *c-condition* pragmatic: CD pertains to the type of device.
CD-4 *c-condition* otherwise

Table 7: Presentation heuristics for CONDITIONS – see appendix A.1

Common semantic constraint: An outcome *OU* is to be communicated.
OU-1 *purpose* pragmatic: The reader knows or can guess that *OU* is desirable and the method described in the instruction is the normal way to realize the OUTCOME.
OU-2 *result* otherwise

Table 8: Presentation heuristics for OUTCOMES – see Appendix A.2

Common semantic constraint: A REQUIRED OPERATION *RO* and a set of GUIDANCES *G* are to be communicated. Common conceptual constraint: *RO* is the parent operation of all members of *G*. Common rhetorical constraint: *RO* will be presented by a *sequence*.
GD-1a *means* pragmatic: At least one member of *G* specifies the use of an instrument.
GD-1b *means* conceptual: *G* is a single operation and *RO* is a basic-level operation.
GD-2a *purpose* conceptual: *G* is a pair of operations and both members of *G* should be executed co-temporally.
GD-2b *purpose* conceptual: *G* is a single operation and *G* is a basic-level operation.
GD-2c *purpose* otherwise

Table 9: Presentation heuristics for GUIDANCES – see Appendix A.3

Common semantic constraint: An OPTION *OP* is to be communicated.
OP-1 *purpose* rhetorical: A *sequence* will be related to *OP* in the text to ensure it a nucleus. intentional: *OP* will probably be followed by the agent.
OP-2 *c-condition* otherwise

Table 10: Presentation heuristics for OPTIONS – see Appendix A.4

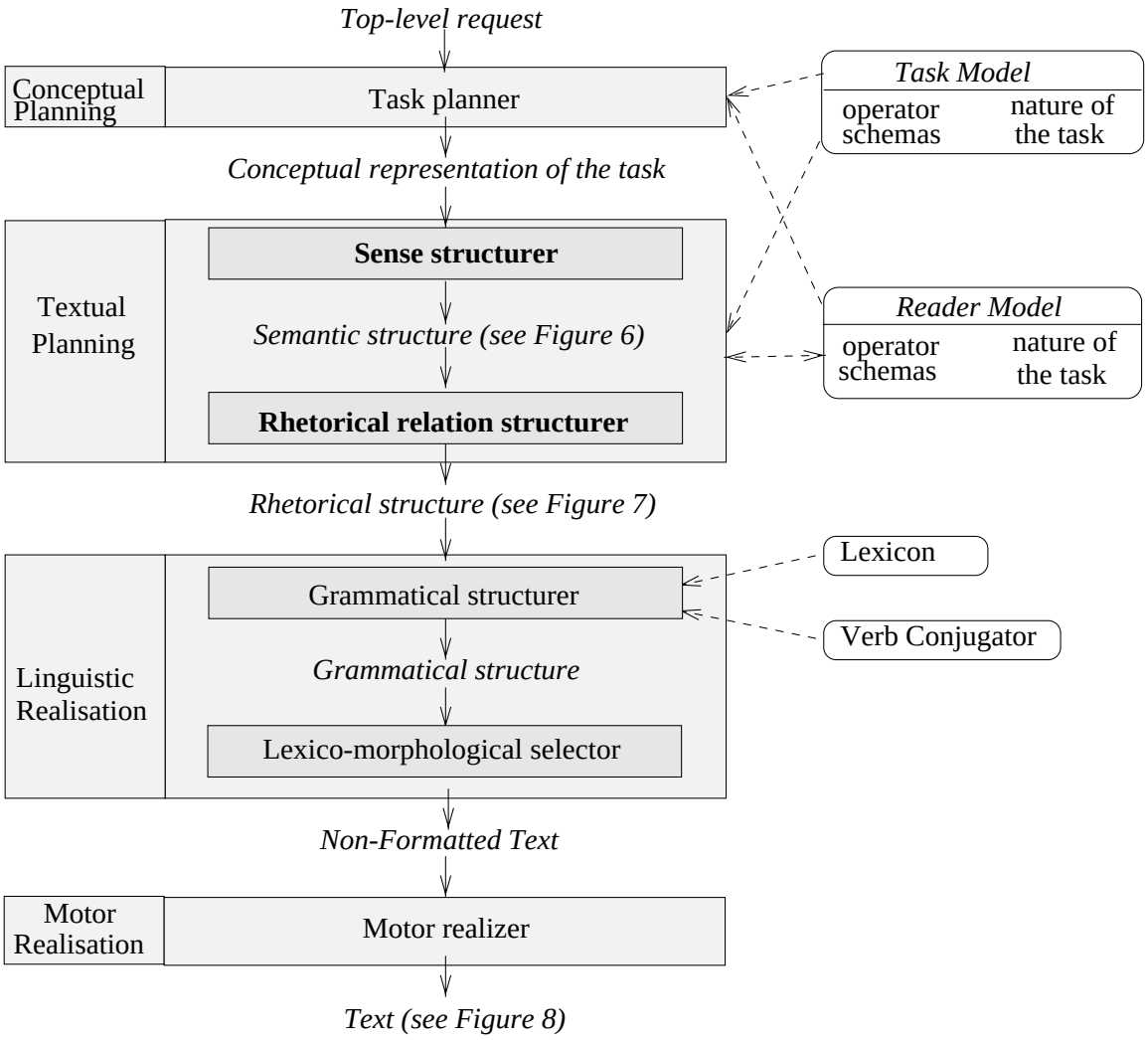


Figure 5: Architecture of SPIN

representation to a French formatted text. The emphasis of SPIN was put on the text planning stage: the sense and the rhetorical relation structurers. The other modules are rather straightforward, but by building a complete system, we have achieved three goals:

1. We have shown the feasibility of the overall generation system.

2. We made it harder to "shove off" problems in the semantic and rhetorical stages to another module.

3. We ensured that a hierarchy of plans and a conceptual representation of the domain can indeed give all the necessary information to the semantic structurer and that the output of the rhetorical structurer is sufficient to obtain a readable text.

operator:	`record_by_OTR(objet:X)`
precondition:	`inside(object:cassette, location:recorder)`
body:	`set(object:tape speed, to:SP)`
	`select(object:channel 4)`
	`specify(object:duration time of 1 h 30 min)`
	`push(object:timer button)`
success postcondition:	add:`programmed(objet:X)`
failure postcondition:	add:`not(programmed(objet:X))`

Figure 6: Example of operator schema for the `vcr` example

SPIN follows the linear architecture shown in Figure 5. This type of architecture was chosen for its implementation simplicity; although, it prevents us from having lower-level decisions influence upper-level ones iteratively by means of a feedback loop. However, SPIN is written in Prolog so that, if a choice made at one level cannot be realized by lower-level components, Prolog's built-in backtracking mechanism allows the previous component to try another possibility until the specification is realizable by the lower-level components.

Let us now describe each knowledge base and each module of SPIN. For easy of reference, we will use a running example, called the `vcr` example. It instructs on how to use the "One Touch Recording" (OTR) feature of a video tape recorder.

Two important knowledge bases are used: the task model and the reader model.

The task model contains a library of operator schemas and a description of the nature of the task. The operator schemas describe operations in terms of preconditions, sub-operations and postconditions (Sacerdoti 1977). Figure 6 shows an example of an operator schema. The description of the nature of the task lists the actual properties of the task. These properties includes, the fact that an operation is dangerous, optional or irreversible, the fact that the procedure is to be followed at the time of reading, etc. The task model is used to verify the pragmatic constraints of the presentation heuristics. For example, to verify the pragmatic constraint of **RO-2** (*RO_2 is a modifiable condition*), SPIN verifies if the predicate `modifiable(RO2)` exists in the task model.

The reader model is parallel to the task model, but contains the beliefs of the reader rather than the task's actual properties. This model indicates what the reader believes about the task: how he believes operations should be performed and what properties of the nature of the task he knows. A particular knowledge of the reader is the notion of basic-level operations which is domain and agent-specific. The model of the reader specifically lists the operations that are basic-level for the

reader. The reader model is updated dynamically when a piece of information is selected to be conveyed in the text. This model is used to verify the conceptual constraints of the presentation heuristics (for the notion of basic-level operations) and intentional constraints. For example, to verify the **RO-1a**'s intentional constraint (*The reader believes that RO should be executed after* RO_2), SPIN verifies if the predicate `believes_order([RO2, RO1])` exists in the reader model.

The modules that make up SPIN are:

The task planner is used to plan the procedure to achieve a particular goal. To do this, it constructs a *conceptual representation of the task* using a nonlinear AI planner and a library of schemas of operators (Sacerdoti 1977). This technique is typically used in the generation of instructional texts (Mellish 1988; Dale 1992; Paris and Vander Linden 1996) because the structure of the resulting plans is hierarchical, similarly to the structure of the text. The result of this step is a tree-like structure of operator schemas necessary to achieve the top-level request. For example, given the top-level request, `record_by_OTR(object:X)` and a library of operator schemas like the one shown in Figure 6, the task planner builds a hierarchical conceptual representation of the task linking operations and sub-operations in a tree structured representation. Each level of the tree is a plan at a given level of abstraction.

The sense structurer first selects which senses will be expressed and which will be left implicit. SPIN takes into account a reader model and a task model. The set of relevant senses computed by the procedure are then reordered in the following way:

1. Preparation operations are presented first, e.g. *Wash the strawberries.*
2. Autonomous operations are presented second. Autonomous operations are those operations that the agent just activates but are then completed by themselves, e.g. Melt the chocolate. *Let cool.*
3. The remaining operations are then presented while keeping adjacent the operation that depend upon a single parent; thus decreasing the number of focus changes.

Figure 7 shows the list of ordered senses for the `vcr` example (translated in English). This structure was generated by the sense structurer from the conceptual tree-representation of the task built by the task planner and from the task and reader models. This structure uses a unique predicate to represent a sense to be conveyed. Each predicate has at least 1 argument, the information content to be

```
[title  (program(obj:a 1h 30 recording, qual:using the one-touch recording button)),
 req_op (set(obj:speed selector, dest:SP)),
option  (better picture quality, set(obj:speed selector, dest:SP)),
req_op  (select(obj:channel 4)),
guidance(select(obj:channel 4), push(dest:channel button)),
req_op  (push(dest:channel button)),
req_op  (push(dest:OTR button, qual:3 times)),
req_op  (push(obj:OTR button, qual:1 time)),
outcome(PM 10:35 (30min), push(obj:OTR button, qual:1 time)),
req_op  (push(obj:OTR button, qual:2 times)),
outcome(PM 11:05 (1 h), push(obj:OTR button, qual:2 times)),
req_op  (push(obj:OTR button, qual:3 times)),
outcome(PM 11:35 (1 h 30min), push(obj:OTR button, qual:3 times)),
req_op  (push(dest:TIMER button, qual:within 9 seconds))]
```

Figure 7: Semantic structure of the `vcr` text given by the ordered list of senses

conveyed, and in the case of senses related to operations, a second argument indicates the information content of the operation the sense is related to. For example, Figure 7 states that after the title (not discussed in this paper), the text will convey a required operation (`req_op`) whose informational content[11] is `set(obj:speed selector, dest:SP)` followed by an option whose informational content is `better picture quality` related to the operation `set(obj:speed selector, dest:SP)`.

The rhetorical relation structurer takes over the semantic structure and applies the presentation heuristics described in section 4 to select the most appropriate rhetorical relation to communicate a particular sense. To verify the different types of constraints, SPIN consults different knowledge sources. Conceptual constraints are verified by consulting the task plan built by the task planner and the knowledge of the reader. Semantic constraints use the semantic structure built by the previous process. Rhetorical constraints are satisfied by verifying the rhetorical selection made so far, and by imposing constraints on the future choices of rhetorical relations. Pragmatic and intentional constraints are verified by consulting the task and reader models. The result of this process is an RST-like text representation. Once the list of rhetorical relations is selected, it is reordered to take into account some cooccurrence constraints and to position the satellites in relation with their nucleus. Vander Linden (1993) has thoroughly studied this aspect in English and in our corpus, the relative position of the elements of a rhetoric relation seem to fol-

[11] Note that the informational content is explicitly given in the task planner's operator schema.

low the same rules. This step finally adds appropriate punctuation signs between and among relations using simple heuristics.

Figure 8 illustrates the final rhetorical structure of the `vcr` text where the rhetorical relations are of the form `rhetorical_relation(Satellite, Nucleus)`. To better illustrate this Figure, let us go through the transformation of the first REQUIRED OPERATION and the first OPTION of Figure 7. According to the heuristics of section 4.2 to present the REQUIRED OPERATION the rhetorical relation structurer has a choice between a *c-condition*, a *purpose* or a *sequence*. Heuristic **RO-1a** is not satisfied because according to the reader model, the reader does not believe in a reverse order of execution between the current operation and another. Heuristic **RO-1b** is not satisfied either, because the agent of the operation is the agent of the rest of the procedure. Heuristic **RO-1c** is not satisfied because no outcome of the operation needs to be communicated. Finally, heuristic **RO-2** is not satisfied because no precondition to the operation needs to be communicated. In this case, the default heuristic **RO-3** is applied and a *sequence* is selected.

To present the OPTION, heuristic **OP-1** is satisfied because the preceding REQUIRED OPERATION will be presented by a *sequence*, and the model of the reader indicates that the option will likely be followed. A relation of *purpose* is therefore selected.

After a re-ordering of the clauses, the first 2 senses of Figure 7:

```
req_op(set(obj:speed selector, dest:SP))
option(better picture quality, set(obj:speed selector, dest:SP))
```

will therefore be transformed into the first 2 rhetorical relations of Figure 8:

```
purpose(better picture quality, set(obj:speed selector, dest:SP))
sequence(set(obj:speed selector, dest:SP))
```

The grammatical structurer selects the appropriate grammatical structures to realize the satellites and the nuclei of the rhetorical relations. To realize this step, we have adapted Vander Linden (1993)'s analysis of English instructions to our French corpus, and in addition, we took into account the textual type as a criterion for choosing grammatical structures. Indeed, our corpus analysis revealed a strong correlation between the textual genre (execution, hybrid or comprehension texts) and the grammatical form of rhetorical relations; thus confirming Hartley and Paris (1996)'s results.

The lexico-morphological selector takes over the grammatical structure and realizes it into words. The words are chosen (in the current implementation, a one-to-one correspondence exists between concepts and words) and declined. This

```
[   title,
 purpose(program(obj:a 1h 30 recording, qual:using the one-touch recording button)),
    paragraph, item,
 purpose(better picture quality, set(obj:speed selector, dest:SP)),
    comma,
 sequence(set(obj:speed selector, dest:SP)), period, item,
 sequence(select(obj:channel 4)),
 means(push(dest:channel button), select(obj:channel 4)),
    period, item,
 sequence(push(dest:OTR button, qual:3 times)),
    period, nl, space,
 c_condition(push(obj:OTR button, qual:1 time)), push(obj:OTR button, qual:3 times)),
    comma,
 result(PM 10:35 (30min), push(obj:OTR button, qual:1 time))),
    period, nl, space,
 c_condition(push(obj:OTR button, qual:2 times), push(obj:OTR button, qual:3 times)),
 result(PM 11:05 (1 h), push(obj:OTR button, qual:2 times)),
    period, nl, space,
 c_condition(push(obj:OTR button, qual:3 times), push(dest:OTR button, qual:3 times)),
 result(PM 11:35 (1 h 30min), push(dest:OTR button, qual:3 times)),
    period, nl, period, item,
sequence(push(dest:TIMER button, qual:within 9 seconds)),
    period]
```

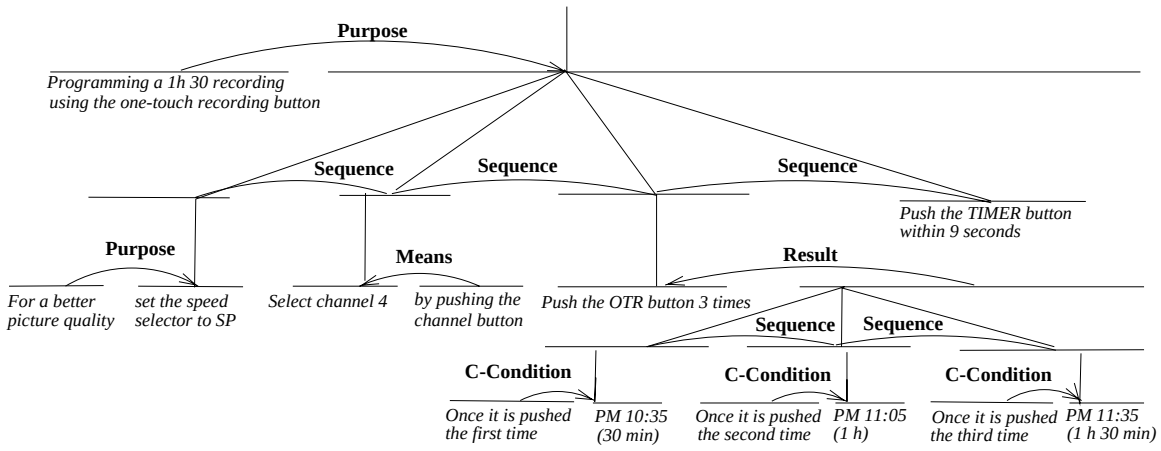


Figure 8: Rhetorical structure of the `vcr` text and its RST representation

Programmation d'un enregistrement d'1h 30 min par une touche incrémentielle

- Pour une meilleure qualité d'image, réglez le sélecteur de vitesse de bande sur SP.
- Sélectionnez le canal 4 en appuyant sur la touche de canal.
- Appuyez sur la touche OTR 3 fois.
  Lorsqu'elle est enfoncée 1 fois, PM 10:35 (30min).
  Lorsqu'elle est enfoncée 2 fois, PM 11:05 (1 h).
  Lorsqu'elle est enfoncée 3 fois, PM 11:35 (1 h 30min).
- Appuyez sur la touche TIMER dans un délai de 9 secondes.

Programming a 1h 30 recording using the one-touch recording button

- For a better picture quality, set the speed selector on SP.
- Select channel 4 by pushing the channel button.
- Push the OTR button 3 times.
  Once it is pushed the first time, PM 10:35 (30min).
  Once it is pushed the second time, PM 11:05 (1h).
  Once it is pushed the third time, PM 11:35 (1 h 30min).
- Push the TIMER button within 9 seconds.

Figure 9: Final `vcr` text generated from SPIN and its English translation

component is also responsible for producing anaphoric expressions based on Tutin (1992)'s work. For example, the anaphora-generation rules enabled SPIN to generate:

> Appuyez sur la touche OTR 3 fois.
> Lorsqu'*elle* est enfoncée 1 fois, PM 10:35 (30min).
> Press the OTR button 3 times.
> Once *it* is pushed the first time, PM 10:35 (30min).

where the pronoun *elle* refers to *la touche OTR* introduced in the text in the previous sentence. SPIN can also produce personal pronouns, partial repetitions and ellipsis of noun phrases.

The motor realizer finally realizes the typographical layout of the text. Elisions, capitalizations, etc are performed. SPIN can produce paragraph-structured procedures and itemized lists. Figure 9 shows an example of final text generated by SPIN.

6 Evaluation of the Heuristics

Today, in natural language generation, there exists no accepted methodology to validate a generation theory. The problem is amplified by two phenomena: the lack of a canonical form for the input and sometimes the output of a generation system and the difficulty of defining the set of "correct" output (Walker 1989; Bates, Hovy, and Senneff 1994).

One type of evaluation consists of comparing each decision taken by a generation system with the corresponding decision taken by a human writer. If an identical match is found, the system is considered correct, otherwise, it is considered at fault. We do not believe that such a *word-to-word/grammatical form-to-grammatical form/...*, is an appropriate evaluation method. If a human writer has taken a different choice (be it lexical, grammatical, rhetorical, or other) from the system to be evaluated, it does not necessarily imply that either one is wrong. In fact, two human writers, instructed to write a text on the same topic and in the same communicative context, will most probably come up with different texts. Natural language provides a wide array of means to communicate the same information; while only a few will create the most appropriate expression in the communicative context, in rare case can one choice be considered the best. In our view, using this technique (like (Vander Linden and Martin 1995)) is simply too hard on the generation system.

A valid and more objective evaluation method consists of giving subjects a set of "natural" and generated texts and asking them to evaluate them according to specific criteria (e.g. informational content, text coherence, comprehensibility, etc). This method is particularly interesting for instructional texts as the subjects can be asked to actually perform the prescribed procedure. Criteria like their execution errors and reading time can be measured to evaluate the quality of the texts and compare them to "natural" ones. This method, although interesting, needs an involved experimental setup and experience in interpreting psychological performances that we do not have. It would also imply separating the text output quality from the cognitive ability of the individual readers. Given these difficulties, we did not pursue this approach.

In order to validate our heuristics, we have thus followed two other more popular evaluation methods: a comparison with other instruction-generation systems and a global qualitative comparison of the generated texts with their "natural" counterparts.

6.1 Comparison with previous work

An important research in the generation of instructional texts is the DRAFTER project (Paris, Vander Linden, Fischer, Hartley, Pemberton, Power, and Scott 1995; Paris and Vander Linden 1996; Paris and Vander Linden 1996). The emphasis of the DRAFTER project was the analysis and generation of grammatical expressions to convey specific

procedural relations in a variety of languages. The DRAFTER system first plans the content of the text using Moore and Paris (1993)'s text planner. This planner allows the same conceptual information to be presented through different rhetorical relations; but in addition DRAFTER determines its choice according to the language of communication. In DRAFTER, the choice of the semantic information to be conveyed and the rhetorical relations to be used is performed at the same time through discourse strategies. However, although both steps are performed simultaneously, the discourse strategies distinguishes the semantic and rhetorical operators, similarly to SPIN.

The DRAFTER project can thus be seen as complementary to ours: the emphasis of SPIN is to determine the most appropriate semantic and rhetorical structures of the text from its conceptual representation, while the emphasis of DRAFTER is to determine the text's most appropriate grammatical structure from its procedural relations. Although the two projects do not deal with the same levels, they share the same underlying hypothesis that there exists a many-to-many correspondence between the different levels of text representation. In addition, to map two levels of text representation in a text generation framework, both studies have fundamentally used the same approach: a corpus analysis to develop linguistically-motivated heuristics. It is therefore conceivable to combine DRAFTER and SPIN to create an instructional text generator having both the strengths of SPIN's deep generation and DRAFTER's French grammatical selection. A similar experiment has been performed with SPIN and the FLAUBERT generator (Danlos and Lapalme 1999). The existence of the DRAFTER project emphasizes the appropriateness of our approach; as to the specific results, a comparison with DRAFTER'S text planner is necessary. However, to our knowledge, the discourse strategies used by DRAFTER have not explicitly presented in the literature.

6.2 Comparison with natural texts

In the second evaluation, we took 3 texts (Figure 10 shows one of them) outside our corpus of analysis and generated them by SPIN, then we compared the two versions on a qualitative basis. If we compare the content and the rhetorical structure of the texts in Figure 10 we can see for each aspect:

semantic content From the semantic point of view, the text generated by SPIN is the same as the original text except for one sense. The REQUIRED OPERATION *vous devez les dévisser* found in sentence 3 of the original text, was not communicated by SPIN. This is because SPIN was told that the operation was not a basic-level one. In light of this, the system did not deem it useful to communicate it in an execution-oriented text.

Text generated by SPIN

Desserrage des écrous de la roue

- N'enlevez pas les écrous complètement.
- Desserrer les écrous avec une clé en croix.
- S'ils portent la marque L, les écrous ont le filetage à gauche, tourner les écrous dans le sens des aiguilles d'une montre.
- S'ils ne portent aucune marque, tourner les écrous en sens contraire.
- Si les écrous sont difficiles à dévisser, mettre un peu d'huile pénétrante, attendre quelques minutes, essayer de nouveau.

Translation of the text generated by SPIN

Loosening the wheel nuts

- Do not remove the nuts completely
- Loosen the nuts with a cross wrench.
- If the nuts have an "L" mark, they have a left-hand thread, and you must unscrew them clockwise.
- If they have no mark, turn the nuts in the opposite direction.
- If the nuts are hard to unscrew, apply some penetrating oil, wait a few minutes, try again.

Original text

Desserrage des écrous

Utilisez une clé en croix pour desserrer les écrous de la roue. S'ils portent la marque L, ils ont le filetage à gauche et vous devez les dévisser en tournant dans le sens des aiguilles d'une montre. Pour les écrous qui ne portent aucune marque, tournez en sens contraire. Si les écrous sont difficiles à dévisser, mettez un peu d'huile pénétrante, attendez quelques minutes puis essayer de nouveau. N'enlevez pas complètement les écrous.

Translation of the original text

Loosening of the nuts

Use a cross-bar to loosen the nuts of the wheel. If they have an L mark, they have a left-hand thread and you must unscrew them by turning clockwise. For nuts that have no mark, turn counter-clockwise. If the nuts are difficult to unscrew, apply a small amount of penetrating oil, wait a few minutes, then try again. Do not remove the nuts.

Figure 10: Example of a text generated by SPIN

rhetorical structure From the rhetorical point of view, SPIN did not always pick the same rhetorical relations present in the original text.

1. A *means* is used in sentence 2 of the generated text:

   (25) Desserrer les écrous avec une clé en croix.
   Loosen the nuts with a cross-bar.

   while the original text used a *purpose*:

   (26) Utilisez une clé en croix pour desserrer les écrous de la roue.
   Use a cross-bar to loosen the nuts of the wheel.

   The underlying sense in both sentences a GUIDANCE, as the use of a cross-bar *guides* the agent in determining how to loosen the nuts. The heuristic **GD-1a** was responsible for the choice of the relation of *means* in that case because the use of a particular instrument was communicated. According to our corpus analysis, SPIN's choice is the most common one made, and to us, seems quite appropriate.

2. A *c-condition* is used in sentence 4 of SPIN's version:

   (27) S'ils ne portent aucune marque, tourner les écrous en sens contraire.
   If they have no mark, turn the nuts counter-clockwise.

   while a *purpose* is used in the original text:

   (28) Pour les écrous qui ne portent aucune marque, tournez en sens contraire.
   For nuts that have no mark, turn the nuts counter-clockwise.

   In both sentences, a CONDITION is communicated. The heuristic **CD-3** was responsible for SPIN's choice. Recall that in the case of conditions that pertain to the type of a device, **CD-3** suggests the use of a *purpose* or a *c-condition*, but has no preference. SPIN thus picked randomly the relation of *c-condition*, which in our opinion sounds very natural.

Kosseim (1995) has analyzed three texts outside the training corpus and found that the output of SPIN is comparable to the ones found in instruction booklets.

7 Conclusion and Further Research

This article has presented the results of an investigation and the implementation of a system for the generation of instructional texts. We take the view that the planning of instructional texts must be a 2-stage process: selecting its semantic content then its rhetorical structure. Indeed, one sense can be presented through different rhetorical

relations, while the same rhetorical relation can carry different semantic contents. This article has emphasized the results of this second step: how to select the most appropriate rhetorical relation in French instructions. First, we have introduced the 9 senses typically found in instructional texts, and the 7 rhetorical relations used to present them. The presentation heuristics have then been specified for the 7 most common senses in order to show how the most appropriate rhetorical relation can be chosen automatically. These heuristics are based on the notion of basic-level operation and 5 types of constraints. The SPIN system, implementing the heuristics, has then been presented in order to validate the research.

The most important contribution of this study is the explicit separation of two important questions in textual planning. The *semantic level* corresponding to the "What to say?" problem and the *rhetoric level*, the "How to organize it?" problem, are considered separately while being linked in many ways. These two levels had not been considered separate before. For example, in a schema or RST based approach for textual planning, these questions are more tightly linked, reducing the rhetorical diversity especially in multilingual generation where rhetorical choices might differ between languages.

SPIN's heuristics are based on a thorough linguistic corpus study and thus take into account many interesting phenomena of the sublanguage of instructional texts.

SPIN is a complete system that goes from a high level description of a task to coherent and natural short (about 12 relations) French texts instructing on how to perform the task. The longer texts we have analyzed are linked by a sequence relation and ruled by the same global textual rules. We believe that we can reuse the selection and presentation heuristics for longer texts but we would need a strategy for separating a global task into subtexts to be generated individually.

Among the questions raised by this research, one can consider the influence of the discourse domain and of the language of communication on the selection of the content and the structure. For example, cooking recipes have a higher percentage of REQUIRED OPERATIONS and *sequences* than instructions from other domains. We have attributed this phenomenon on the communicative goal of the text (execution text versus comprehension texts), but even among execution texts, cooking recipes have a particularly stereotypical content and structure. The discourse domain may in fact have a lot to do with this.

The language of communication and cultural context may also play a role in the choice of textual content and structure in instructions. These factors may very well influence the semantic content of instructional texts as different cultures have a different level of knowledge on a particular domain and have different sets of values. Warnings and

safety issues, that are required to be communicated in North-American instructions, may seem out of place in another country where the culture and the legislation are different. The number of PREVENTIONS, for example, will therefore be influenced by this factor. As to the structure of the texts, Delin, Scott, and Hartley (1993) have already noted that in multi-lingual instructions, the same content may be presented by different rhetorical means according to the language of communication. These very interesting issues remain to be investigated.

Acknowledgment

This work was supported by a scholarship and grants from the Natural Sciences and Engineering Research Council of Canada (NSERC). The authors wish to thank the anonymous referees, whose comments greatly improved the paper.

References

Bates, L., E. Hovy, and S. Senneff (1994, June). (invited panelists). Panel discussion: Evaluating Generation Systems, 7^{th} International Workshop on Natural Language Generation. unpublished.

Dale, R. (1992). *Generating Referring Expressions: Constructing Descriptions in a Domain of Objects and Processes*. ACL–MIT Press Series in Natural Language Processing. The MIT Press.

Danlos, L. and G. Lapalme (1999, April). An experiment in combining two text generators. In *Proceedings of the AISB'99 Workshop on Reference Architectures and Data Standards for NLP*, Edinburgh, pp. 1–7.

Delin, J., A. Hartley, C. Paris, D. Scott, and K. Vander Linden (1994, June). Expressing Procedural Relationships in Multilingual Instructions. In *Proceedings of the 7^{th} International Workshop on Natural Language Generation*, Kennebunkport, Maine, pp. 61–70.

Delin, J., D. Scott, and A. Hartley (1996, August). Language-specific mappings from semantic to syntax. In *Proceedings of the sixteenth International Conference on Computational Linguistics (COLING-96)*, Copenhagen, Denmark, pp. 292–297.

Delin, J., D. Scott, and T. Hartley (1993, June). Knowledge, Intention, Rhetoric: Levels of Variation in Multilingual Instructions. In O. Rambow (Ed.), *Proceedings of the ACL Workshop on Intentionality and Structure in Discourse Relations*, Ohio State University, pp. 7–10.

Di Eugenio, B. (1993). *Understanding Natural Language Instructions: A Computational Approach to Purpose Clauses*. Ph. D. thesis, University of Pennsylvania.

Di Eugenio, B. and B. L. Webber (1996). Pragmatic Overloading in Natural Language Instructions. *International Journal of Expert Systems, Special Issue on Knowledge representation and Reasonning in Natural Language Processing 9*(1).

Dixon, P. (1987, February). The Processing of Organizational and Component Step Information in Written Directions. *Journal of Memory and Language 26*(1), 24–35.

Dixon, P., J. Faries, and G. Gabrys (1988). The Role of Explicit Action Statements in Understanding and Using Written Directions. *Journal of Memory and Language*.

Donin, J., R. Bracewell, C. Frederiksen, and M. Dillinger (1992, April). Students' Strategies for Writing Instructions, Organizing Conceptual Information in Text. *Written Communication 9*(2), 209–236.

Goldman, A. (1970). *A Theory of Human Action*. Princeton, NJ: Princeton University Press.

Hartley, A. and C. Paris (1996, June). Two sources of control over the generation of software instructions. In *Proceedings of the* 34^{th} *Annual Meeting of the Association for Computational Linguistics (ACL-96)*, University of California, Santa Cruz, pp. 192–199.

Kittredge, R. (1982). Variation and homogeneity of sublanguages. Chapter 4, pp. 107–137. Berlin: Walter de Gruyter.

Korelsky, T. and R. Kittredge (1993, June). Towards stratification of RST. In O. Rambow (Ed.), *Proceedings of the ACL Workshop on Intentionality and Structure in Discourse Relations*, Ohio State University, pp. 52–55.

Kosseim, L. (1995). *Planification de textes d'instruction: Sélection du contenu et de la structure rhétorique*. Ph. D. thesis, Université de Montréal.

Mann, W. and S. Thompson (1988). Rhetorical Structure Theory: towards a functional theory of text organization. *TEXT 8*(3), 243–281.

Mellish, C. (1988). Natural Language Generation from Plans. In M. Zock and G. Sabah (Eds.), *Advances in Natural Language Generation*, Communications in Artificial Intelligence Series, Chapter 7, pp. 131–145. London: Pinter Publishers.

Moore, J. and C. Paris (1993, December). Planning Text for Advisory Dialogues: Capturing Intentional, and Rhetorical Information. *Computational Linguistics 19*(4), 651–694.

Moore, J. and M. Pollack (1992, December). A Problem for RST: The Need for Multi-Level Discourse Analysis. *Computational Linguistics 18*(4), 537–544.

Paris, C. and K. Vander Linden (1996, August). Building Knowledge Bases for the Generation of Software Documentation. In *Proceedings of the Sixteenth International Conference on Computational Linguistics (COLING-96)*, Volume 2, Copenhagen, Denmark, pp. 734–739.

Paris, C. and K. Vander Linden (1996, July). Drafter: An interactive support tool for writing multilingual instructions. *IEEE Computer 29*(7), 49–56. (Special Issue on Interactive Natural Language Processing).

Paris, C., K. Vander Linden, M. Fischer, A. Hartley, L. Pemberton, R. Power, and D. Scott (1995). A Support Tool for Writing Multilingual Instructions. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI-95)*, Volume 2, Montréal, Canada.

Pollack, M. (1986). *Inferring Domain Plans in Question-Answering*. Ph. D. thesis, University of Pensylvania.

Rosch, E. (1978). Principles of Categorization. In E. Rosch and B. Lloyd (Eds.), *Cognition and Categorization*, pp. 27–48. Hillsdale, NJ: Lawrence Erlbaum.

Rösner, D. and M. Stede (1992, April). Customizing RST for the Automatic Production of Technical Manuals. In R. Dale, E. Hovy, D. Rösner, and O. Stock (Eds.), *Aspects of Automated Natural Language Generation*, Lecture Notes in Artificial Intelligence, pp. 199–214. Trento, Italy: Springler–Verlag.

Sacerdoti, E. D. (1977). *A Structure for Plans and Behavior*. New York: Elsevier.

Tutin, A. (1992). *Étude des anaphores grammaticales et lexicales pour la génération automatique de textes de procédures*. Ph. D. thesis, Département de linguistique et philologie, Université de Montréal.

Vander Linden, K. (1993, June). Generating effective instructions. In *15th Annual Conference of the Cognitive Science Society*, Boulder, Colorado, pp. 1023–1028.

Vander Linden, K. (1993). *Speaking of Actions: Choosing Rhetorical Status and Grammatical From in Instructional Text Generation*. Ph. D. thesis, University of Colorado.

Vander Linden, K. and B. Di Eugenio (1996, June). Learning micro-planning rules for preventative expressions. In *Proceedings of the Eighth International Natural Language Generation Workshop*, Sussex, UK, pp. 11–20.

Vander Linden, K. and J. H. Martin (1995, March). Expressing local rhetorical relations in instructional text: A case-study of the purpose relation. *Computational Linguistics 21*(1), 29–57.

Walker, M. (1989, June). Evaluating Discourse Processing Algorithms. In *Proceedings of the 27th Annual Meeting of the Association for Computational Linguistics (ACL-89)*, Vancouver.

Appendix: Justification of the Presentation Heuristics

A.1 Justification of Table 7 : Conditions

A CONDITION can be presented by 4 rhetorical relations:

1. A *c-condition* (90 % of the time):

   (29) *Si des bulles apparaissent autour de la valve*, c'est qu'il y a une fuite à cet endroit. Faites-la remplacer par un mécanicien qualifié.
   If bubbles appear around the valve, there is a leak. Have it repaired by a qualified mechanic.

   Here, 2 equivalent CONDITIONS are presented: the appearance of bubbles, and the existence of a leak. As we will see later, the first condition is presented through a *c-condition*; while the second is presented through a *result* (see example (31)).

2. A *purpose* (4 %):

   (30) *Pour [vérifier] un commutateur ordinaire* [...], touchez la vis de la borne de cuivre avec la pince du vérificateur.
   To [check] an ordinary switch [...], touch the screw of the copper terminal with the tester.

3. A *result* (4 %):

   (31) Si des bulles apparaissent autour de la valve, *c'est qu'il y a une fuite à cet endroit*. Faites-la remplacer par un mécanicien qualifié.
   If bubbles appear around the valve, *there is a leak*. Have it repaired by a qualified mechanic

4. a *sequence* (2 %):

   (32) Introduire la cassette (*vérifier que la languette de la vidéocassette n'a pas été enlevée.*)
   Insert the cassette (*check that the tab of the video cassette has not been removed.*)

CD-1 A *result* and a *c-condition* are used if:

semantic constraint: Two equivalent CONDITIONS are to be communicated.

conceptual constraint: none

rhetorical constraint: none

pragmatic constraint: The condition that is difficult to evaluate presents the main condition justifying the next line of actions. This state is always presented by a *result* and the state that is easier to verify uses a *c-condition*.

intentional constraint: none

This is the case in (29) and (31).

CD-2 *sequence* is used if:

semantic constraint: A CONDITION CD is to be communicated, and CD is not related to the operation it constrains by an enablement (Di Eugenio 1993).

conceptual constraint: none

rhetorical constraint: none

pragmatic constraint: CD is a modifiable condition.

intentional constraint: none

This is the case in (32). This phenomenon seems to apply both in French and in English. In fact, Vander Linden (1993) notes that conditions that specify a modifiable state are presented by a *sequence*. He calls this phenomenon a *rhetorical promotion*. Indeed, what is considered a condition on an operation at the semantic level is promoted to a full-fledged agent-action at the rhetorical level. Instead of being presented as the satellite of a relation, it becomes a nucleus.

CD-3 A *purpose* or a *c-condition* with ellipsis of the verb is used when:

semantic constraint: A CONDITION CD is to be communicated.

conceptual constraint: none

rhetorical constraint: none

pragmatic constraint: CD pertains to the type of device; for example, a particular model or feature.

intentional constraint: none

In this case, the CONDITION is presented by either *purpose* or a *c-condition* (see e.g. (30)).

CD-4 In all other cases, CONDITIONS are presented by a *c-condition*, as in (29).

A.2 Justification of Table 8: Outcomes

An OUTCOME can be presented by 3 relations:

1. a *result* (68 %):

   (33) a. Brancher le cordon d'alimentation du magnétoscope dans une prise secteur 120V et appuyer sur l'interrupteur "POWER". *Le voyant* "POWER" *s'allume et l'horloge commence à clignoter.*
   Plug the electrical cord of the recorder in a 120V outlet and press the POWER button. *The power light is turned on and the clock starts to blink.*

   b. Engagez le levier de vitesse rapidement dans chacune des positions [...], *cela fait circuler le liquide de transmission.*
   Put the gear-shift lever rapidly in each position [...], *this makes the transmission liquid circulate.*

2. a *purpose* (28 %):

   (34) *Pour protéger les bornes contre la tension*, nouez les extrémités séparées du cordon.
   To protect the terminals from contacting each other, tie the extremities of the wire away from each other.

3. a *means* (4 %):

   (35) Vous pouvez voir le niveau de volume *en observant la barre* ROUGE *sur la gamme de 15 barres affichées sur l'écran.*
   You can see the volume level *by observing the bar on the 15-bar scale displayed on the screen.*

   In this case, the OUTCOME (*to see the volume level)* is presented in a *sequence* of actions and the operation that brings it about (*to observe*) is related to it by a relation of *means*. Here, the OUTCOME is promoted to the nucleus position of relation; while *to observe* is demoted to the satellite position.

Delin, Hartley, Paris, Scott, and Vander Linden (1994) note that in English, an OUTCOME can be presented by a *sequence*, but this phenomenon has not been found in our French corpus.

The choice between a *purpose*, a *result* and a *means* does not seem to depend on cooccurrences of rhetorical relations but on whether the outcomes are desirable or not.

Kosseim (1995) describes many examples that can be categorized by the following heuristics:

OU-1 A *purpose* is used if:

semantic constraint: An OUTCOME *OU* is to be communicated.

conceptual constraint: none

rhetorical constraint: none

pragmatic constraint: The reader knows or can guess that *OU* is desirable and the method described in the instruction is the normal way to realize the OUTCOME (as in (34)).

intentional constraint: none

OU-2 In all other cases, OUTCOMES are presented by a *result*. This is the case in (33a) where the OUTCOME specifies the reaction of a device, i.e. a non-desirable side-effect and in (33b) where the reader cannot guess that the OUTCOME is desirable.

A *means* is also used if the reader cannot guess that the OUTCOME is desirable and if the operation that brings about the OUTCOME is the normal method to achieve it. However, in this case, a *result* is always available. As our corpus only includes 4 % of *means*, but 68 % of *result* that present OUTCOMES, we decided to always use a *result*.

A.3 Justification of Table 9: Guidances

A GUIDANCE is made up of a parent operation RO_p and its sub-operations RO_i which "guide" its execution. A GUIDANCE can be presented by:

1. A relation of *means* (69 %); in this case the nucleus is constituted of the parent-operation:

   (36) a. La paraison est centrée soit *en utilisant le marbre*, soit *en la roulant dans une forme creusée dans du bois, appelée "mailloche"*.
   The glass blob is centered either *by using the marble* or *by rolling it in a wooden concave form called "mailloche"*.

   b. *Avec un tournevis plat*, grattez la saleté accumulée sur le contact.
   With a flat-headed screwdriver, scrape off the dirt accumulated on the contact.

   c. Régler la ceinture *en la tirant par la languette*.
   Adjust the seat belt *by pulling it by the strap*.

2. A relation of *purpose* (31 %); in this case the nucleus is made up of the sub-operations:

(37) a. Si la boîte-pont est munie d'une jauge d'huile, tirez-la hors du tube de remplissage, essuyez-la, réinsérez-la complètement et retirez-la de nouveau *pour lire le niveau d'huile*.
If the oil sump has a dip stick, pull it out of the filling tube, wipe it, insert it again entirely and remove it again *to read the oil level*.

b. Tourner cette touche à droite et à gauche *pour minimiser les parasites*.
Turn this knob clockwise and counter-clockwise *to minimize interference*.

c. Rouler le pinceau en le tirant vers soi *de façon à reformer la pointe*.
Twist the paint brush while pulling it towards you *in order to form a pointed tip*.

It is interesting to note that for grammatical reasons, all operations of the same level of abstraction are presented by the same rhetorical relation; the GUIDANCES in:

(38) Pour faire RO_{gp}, faites RO_p en faisant RO_1 et RO_2.
To do RO_{gp}, do RO_p by doing RO_1 and RO_2.

involve operations from three different levels of abstraction: grand parent (RO_{gp}), parent (RO_p) and children (RO_1 and RO_2). However, the operations of the same level (RO_1 and RO_2) are both presented by the same relation (a *means*).

The heuristics involved in the choice of the rhetorical relation are:

GD-1 A *means* is used if:

GD-1a:

semantic constraint: A REQUIRED OPERATION RO is to be communicated along with a set of guiding operations (GUIDANCE) G_i.

conceptual constraint: G_i is a set of sub-operations of an operation RO.

rhetorical constraint: RO will be presented by a *sequence*.

pragmatic constraint: At least one member of G_i presents the use of an instrument.

intentional constraint: none

In this case all members of G_i are presented through a means, so that operations at the same level of abstraction are presented by the same rhetorical relation. For example, in (36a), the use of *marble* dictates the choice of a *means* for all the sub-operations.

GD-1b:

semantic constraint: A REQUIRED OPERATION RO_p is to be communicated and only one of its sub-operation G_i is to be communicated as a GUIDANCE to RO_p.

conceptual constraint: RO_p is the parent operation of G_i and RO_p is a basic-level operation.

rhetorical constraint: RO_p will be presented by a *sequence*.

pragmatic constraint: none

intentional constraint: none

In this case, the GUIDANCE is generally seen top-down (the nucleus is formed of RO_p and the satellite of G_i) by a relation of *means*, as in (36c).

GD-2 A relation of ***purpose*** is used if:

GD-2a:

semantic constraint: A REQUIRED OPERATION RO_p is to be communicated along with a set of guiding operations (GUIDANCE) G_1 and G_2.

conceptual constraint: RO_p is the parent operation of operations G_1 and G_2 which should be executed co-temporally.

rhetorical constraint: RO_p will be presented by a *sequence*.

pragmatic constraint: none

intentional constraint: none

This is the case in (37c) where the sub-operations (*twist* and *pull*) should be performed co-temporally. The only way to communicate this temporal aspect is to use a *concurrency* to relate G_1 and G_2, and a *purpose* between O_p and the pair (G_1, G_2).

GD-2b:

semantic constraint: A REQUIRED OPERATION RO_p is to be communicated along with a set of guiding operations (GUIDANCE) G_i.

conceptual constraint: RO_p is the parent operation of operations G_i.

rhetorical constraint: RO_p will be presented by a *sequence*.

pragmatic constraint: No operation of G_i indicates the use of an instrument.

intentional constraint: none

This is the case in (37a).

GD-2c: In all other cases, GUIDANCES are presented by a *purpose*. This is the case in (37b).

A.4 Justification of Table 10: Options

OPTIONS can be presented in the text by:

1. A *purpose* (79 %):

   (39) Tirer la roue et le pneu; *pour vous aider*, poussez fermement le flanc du pneu avec votre pied.
   Pull the wheel and the tire; *to ease the task*, firmly press on the side of the tire with your foot.

2. A *c-condition* (21 %):

   (40) *Si l'on souhaite une ligne plus large*, alors s'attarder sur le verre de façon à laisser s'écouler plus de couleur.
   If you wish a thicker line, stay longer on the glass in order to let more color flow.

The choice of rhetorical relation only depends on semantic and intentional constraints:

OP-1 A *purpose* is used if:

semantic constraint: An OPTION OP is to be communicated.

conceptual constraint: none

rhetorical constraint: A *sequence* will be related to OP in the text to ensure it a nucleus.

pragmatic constraint: none

intentional constraint: OP will probably be followed by the agent.

This is the case in (39).

OP-2 In all other cases, a *c-condition* is used if. This is the case in (40). The *c-condition* gives an explicit choice to the agent by using expressions like *if you wish* while *purpose* expresses the OPTION less explicitly, thus restricting the possibility of rejection.

A.5 Co-temporal operations

Similarly to ATTRIBUTES, CO-TEMPORAL OPERATIONS are always presented through a single rhetorical relation. In the case of CO-TEMPORAL OPERATIONS, a *concurrency* is always used:

(41) *Rouler le pinceau en le tirant vers soi* de façon à reformer une pointe.
Twist the paint brush while pulling it toward you in order to form a pointed tip.

One should distinguish the relation of *concurrency* and the relation of *means* which, in French, are both realized grammatically by a gerund (eg. *en soufflant*). In the case of a *concurrency*, it is possible to add the adverb *tout* before the gerund (eg. *tout en soufflant*) without modifying the meaning of the expression, while a relation of *means* cannot be realized by *tout* + gerund. In English, the distinction is more obvious as a *means* is realized by a gerund; while the *concurrency* is realized without the preposition *by* but with a preposition like *while*, *meanwhile*, ... (Vander Linden 1993).

A.6 Attributes of Objects

ATTRIBUTES OF OBJECTS are among the senses that are always presented through only one rhetorical relation. According to our corpus, the relation of *elaboration* is always used. For example:

(42) *Une lampe à une seule douille comporte habituellement une lyre qui tient à une barre de retenue par des manchons.*
A single-socket lamp usually has a shade holder attached to a retaining bar by a few fittings.

We will not discuss PREVENTION and POSSIBLE OPERATION because our corpus only contained 15 and 12 occurrences of these senses, hardly enough to develop heuristics of a more general nature.