

Towards Generating Instructional Texts: an exploration of their rhetorical and argumentative structure

Farida Aouladomar, Patrick Saint-Dizier

IRIT-CNRS

118, route de Narbonne

31062 Toulouse Cedex France

aouladom@irit.fr, stdizier@irit.fr

Abstract

Instructional texts consist of sequences of instructions designed in order to reach an objective. The author or the generator of instructional texts must follow a number of principles to guarantee that the text is of any use. Similarly, a user must follow step by step the instructions in order to reach the results expected. In this paper, we explore facets of instructional texts: general prototypical structures, rhetorical structure and natural argumentation. Our study is based on an extensive corpus study with the aim of generating such texts.

1 General typology of instructional texts

Instructional texts, also equivalently called procedural texts, consist of a sequence of instructions, designed with some accuracy in order to reach an objective (e.g. assemble a computer). Procedural texts explain how to realize a certain goal by means of actions which are at least partially temporally organized.

Procedural texts often exhibit a quite complex rational and 'irrational' structure, outlining different ways to realize something, with arguments, advices, conditions, hypothesis, preferences, evaluations, user stimulations, etc. They also often contain a number of recommendations, warnings, and comments of various sorts.

Another feature is that procedural texts tend to minimize the distance between language and action. Plans to realize a goal are made as immediate and explicit as necessary, the objective being to reduce the inferences that the user will have to make before acting, and therefore potential errors or misunderstandings. Texts are thus oriented towards action, they combine instructions with icons, images, graphics, summaries, etc.

In our perspective, procedural texts range from apparently simple cooking recipes to large maintenance manuals (whose paper versions are measured in tons e.g. for aircraft maintenance). They also include documents as diverse as teaching texts, medical notices, social behavior recommendations, directions for use, assembly notices, do-it-yourself notices, itinerary guides, advice texts, savoir-faire guides etc. Procedural texts obey more or less to a number of structural criteria, which may depend on the author's writing abilities

and on traditions associated with a given domain. There is obviously a big contrast between maintenance manuals and people magazines which both contains such texts. Procedural texts can be regulatory, procedural, programmatic, prescriptive, injunctive, or may introduce advices (for social or psychological behavior) (Adam, 2001).

The study we present in this paper has three main goals:

- First, to be able to accurately respond in natural language to *How?* (procedural) questions (Aouladomar 2005), on the Web or in texts, outlining the structure of a procedure that answers such a question,
- Next, to be able to select the best text (assuming a single text is selected as a response) w.r.t. the user profile when there are several responses (a frequent phenomenon on the Web),
- Finally, and for the long term, to be able to merge procedural texts or fragments of procedural texts to construct an optimal text, in terms of level of detail, illustration, etc. Integrating texts is obviously a long-term challenge.

This paper basically relates the structure of instructional texts as they are in French. English translations of examples are just glosses, they are given when space constraints permit. We believe that besides language realization variants, most of the characteristics we present here are language neutral. This study is based on an extensive corpus study, within a language production perspective. This approach allows us to integrate logical, linguistic (e.g. (Moschler, 1985), (Anscombre et al. 1981)) and philosophical views of argumentation.

In this paper we first introduce some elements of a general typology of instructional texts, outlining the number of components, rational and irrational, that compose them. We then give the most important structural elements, and the rhetorical structure that characterizes the relations between elements. Finally, we focus on argumentation, a major dimension of instructional texts and we briefly present the forms of arguments which are the most important, and under what constraints they can be generated.

2 State of the art

Procedural texts have been studied in psycholinguistic, linguistic and didactic circles. We briefly survey various approaches here, outlining elements of interest for our objectives.

2.1 General typology

Under the heading of procedural texts, there is a quite large diversity of texts. (Adam, 2001) notices the variability of judgments in procedural text categorization. Texts can, for example, be grouped into families according to their main objectives and style. We have, for example:

- regulatory texts (Mortara et al., 1988) that characterize expected behaviours,
- procedural texts (Longacre, 1982) defined as rather linear sets of instructions,
- 'programmatory' texts which include recipes, musical scores and architectural plan. (Greimas, 1983) identifies how knowledge from an expert is transferred via these texts to users who are expected to follow strictly the instructions which are given.
- instructional-prescriptive texts (Werlich, 1975), where a quite detailed analysis of temporal and event structures is carried out,
- injunctive texts, where (Adam, 1987) show the form and style used in short notices that relate e.g. fire instructions, security measures, etc.,
- advice texts (Lüger, 1995), which include advice texts of various sorts, such as those found in large public magazines.
- recipe texts (Qamar, 1996), which is a domain quite well-studied, for example in language generation.

Despite their goals, all these forms share common structures: specification of goals, description of lists of pre-requisites to reach the goal, and description of sequences of instructions. They also share common stylistic forms, e.g. preferences for imperative forms, and a number of typographic elements such as enumerations.

From the analysis, mainly psychological or cognitive, of the different forms of procedural texts mentioned above, we categorize procedural texts into three main categories, which will be considered in our project:

- Procedures, e.g.: recipes, maintenance and construction manuals, some medical texts, didactic texts, etc.
- Injunctions, e.g.: orders, regulations, game rules, security measures, etc.
- Advices, e.g.: beauty advices, ways to fill in forms, to behave in certain environments, how to manage a meeting, etc.

Two works will be used as the starting point of the development of the discursive structure of procedural texts that we have elaborated. (Bieger et al., 1984, 1985) propose a taxonomy of the contents of instructions in 9 points: inventory (objects and concepts used), description (of objects and concepts), operational (information that suggest the agent how to realize an action), spatial (spatial data about the actions), contextual, covariance (of actions, which evolve in conjunction), temporal, qualificative (manners, limits of an information), emphatic (redirects attention to another action).

One of the main works in Computational Linguistics is due to (Kosseim, 1996). She isolated 9 main structures or operations, called *semantic elements* from corpus analysis:

1. sequential operations: a necessary action that the agent must realize,
2. object attribute: description meant to help understand the action to realize,
3. material conditions: environment in which an action must be carried out,
4. effects: consequences of the realization of a group of operations on the world,
5. influences: explain why and how an operation must be realized,
6. co-temporal operations: expresses synchronization of operations,
7. options: optional operations,
8. preventions: describes actions to be avoided,
9. possible operations: possible operations to do in the future.

In a different range of ideas, and applied to Japanese, (Takechi et al., 2003), show that it is possible to isolate effective features used to categorize lists of Web pages as being procedural or not. They show that different techniques than in standard text categorization need to be developed, for example in terms of word cooccurrence analysis. List of procedural expressions in the computer domain have been extracted with a quite high accuracy.

Let us finally note a few papers NLG-oriented centered on the generation of arguments such as (Zuckerman et al., 2000), which focus on generic types of arguments (ad absurdum, from cases, etc.) and (Ellahad, 1995).

3 Procedural texts and argumentation

3.1 General considerations

Procedural texts are specific forms of discourse, satisfying constraints of economy of means, accuracy, etc. They are in general based on a specific discursive logic, made up of presuppositions, causes and consequences, goals, inductions, warnings, anaphoric networks, etc., and more psychological elements (e.g. *to stimulate a user*). The goal is to optimize a logical sequencing of instructions and make the user feel safe and confident with respect to the goal(s) he wants to achieve (e.g. *clean an oil filter, learn how to organize a customer meeting*).

Procedural texts, from this point of view, can be analyzed not only just as sequences of mere instructions, but as efficient, one-way (i.e. no contradiction, no negotiation) argumentative discourses, designed to help a user to reach a goal, making the best decisions (see e.g. (Amgoud et al., 2001, 2005)). This type of discourse contains a number of facets, which all are associated in a way to argumentation. Procedural discourse is informative, narrative, explicative, descriptive, injunctive and sometimes figurative. Given a certain goal, it is also of much interest to compare or contrast the

means used by different authors, possibly for different audiences.

Producing explanations is a rather synthetic activity whose goal is to use the elements introduced by knowledge explicitation mechanisms to induce generalizations, subsumptions, deductions, relations between objects or activities and the goals to reach. Explanation, a form of argumentation, does provide a motivation and an internal coherence to procedural text. This is particularly visible in the lexical choices made and in the choice of some constructions, including typographic; procedural discourse is basically interactive: it communicates, teaches, justifies, explains, warns, forbids, stimulates, evaluates.

3.2 The art of writing efficient procedural texts

The author of procedural texts must consider three dimensions (Frederiksen et al., 1992): (1) cognitive: notions referred to must be mastered and understood by the target users, (2) epistemic: take into account, possibly to deny them, the beliefs of those users, and (3) linguistic: use an appropriate language, adjust accuracy, technical level, complexity of sentences and paragraphs, visual and typographic structure of the text. The tonality of the text must be adjusted, depending on style and the domain: didactic, polemic, ironic, with a moral dimension, etc.

The producer of procedural texts starts from a number of assumptions or presuppositions about potential users, about their knowledge, abilities and skills, but also about their beliefs, preferences, opinions, ability to generalize and adapt (to adapt instructions to their own situation, which is never exactly the one described in the procedure), perception of generic situations, and ability to follow discursive processes. The producer of procedural texts has then, from this basis, to re-enforce or weaken presuppositions, to specify some extra knowledge and know-how, possibly beliefs or opinions. He has to convince the reader that his text will certainly lead to the success of the target goal, modulo the restrictions he includes.

Procedural texts are in general highly structured and modular. They exhibit a particularly rich micro-rhetorical structure integrated into the syntactic-semantic structures of instructions. Procedural texts are a difficult exercise to realize. For example they must make linear, because of language constraints, actions which may have a more complex temporal or causal structure. Connectors and referents contribute to implement this linearity. Texts are also expected to be locally and globally coherent, with no contradictions, and no space for hesitation or negotiation.

In most types of procedural texts, in particular social behavior, communication, etc. procedural discourse has two dimensions: an explicative component, constructed around rational and objective elements, and a seduction component whose goal is (1) to encourage the user, (2) to help him revise his opinions, (3) to enrich the goals and the purposes, by outlining certain properties or qualities or consequences of a certain action or prevention.

Another important feature, which is rather implicit, is the way instructions or groups of instructions are organized and follow each other, and both the logic (objective aspect) and

the connotations (subjective aspects) that underlie this organization (sequential, parallel, concurrent, conditional, etc.).

The complexity of instructional texts leads us to proceed very gradually in the study of their structure before being able to produce even simple such texts. At an intermediate level, responding in natural language to *How?* questions, using underspecified templates is an interesting step which can be evaluated.

4 A Discursive analysis of procedural texts

Here is, represented by means of a grammar, the main elements of the structure of a procedural text, see (Aouladomar, 2005a) for more details. The structures reported below correspond essentially to (1) the organization of the informational contents: how tasks are planned, according to goals and sub-goals, and (2) to the argumentative strategies used (planning, progression of tasks, warnings, advices, evaluations, etc.). General principles of argumentative discourse are given e.g. in (Eggs, 1994).

In what follows, parentheses express optionality, + iteration, the comma is just a separator with no temporal connotation a priori, / is an or and the operator < indicates a preferred precedence. Each symbol corresponds to an XML-tag, allowing us to annotate instructional texts.,

The top node is termed **objective**:
objective → **title**, **(summary)**, **(warning)+**, **(pre-requisites)**, **(picture)+** < **instruction sequences**.

summary → **title+** Summary describes the global organization of the procedure, it may be useful when procedures are complex (summary can be a set of hyper-links, often pointing to titles),

warning → **text**, **(picture)+**, **(pre-requisites)**.

pre-requisites → **list of objects**, **instruction sequences**. Pre-requisites describe all kinds of equipments needed to realize the action (e.g. the different constituents of a recipe) and preparatory actions. It may also include presuppositions on the user profile and abilities.

picture describes a sequence of charts and/or schemas of various sorts. They often interact with instructions by e.g. making them more clear.

Instruction sequences is structured as follows:
instruction sequences → **instseq** < **discursive connector** < **instruction sequences** / **instseq**.

instseq is then of one of four main types below:
instseq → **(goal)**, **imperative linear sequence** / **(goal)**, **optional sequence** / **(goal)**, **alternative sequence** / **(goal)**, **imperative co-temporal sequence**.

Goal may contain, besides the target itself motivations, manners, references etc.

Each type of instruction sequence is defined as follows:
Imperative linear sequence → **instruction** < **(temporal mark)**, **imperative linear sequence/ instruction**. (e.g. *inspect carefully if the filter is clean and then open the valve*)

Optional sequence → **conditional expression, imperative linear sequence**. (e.g. *if you prefer a stronger flavor, add curry powder and cream.*)

Alternative sequence → (**conditional expression**), (**argument**) **imperative linear sequence**, (**alternative-opposition mark**) < **instseq** / (**conditional expression, instseq**)⁺. (e.g. *if you can locate the COM1 port, then ... otherwise, or if you wish to be more cautious or cannot locate it, dismount ...*).

Imperative co-temporal sequence → **imperative linear sequence** < **co-temporal mark**, < **imperative co-temporal sequence** / **instruction**.

A co-temporal sequence relates instructions which must be realized at the same time.

Finally, Instruction is the lower level and has the following structure, with recursion on objective:

instruction → (**iterative expression**), **action**, (**reference**)⁺, (**goal**)⁺, (**manner**)⁺, (**motivation**), (**limit**), (**picture**)⁺, (**warning**) / **objective**.

Instructions can be complex since they may contain their own goals, warnings and pictures. If an instruction is complex it is analyzed as an objective.

An example of an annotated text is given at the end of this document.

4.1 Rhetorical structures

Rhetorical structures play several roles in our approach. They first give a semantics to the discursive structure syntax given above. They also contribute to enhancing the production of well-designed responses (Kosseim, 1995). They are also useful, as shall be seen below, to allow for the integration of procedural texts dealing with similar objectives or goals, but this is an extremely difficult task. Finally, they are used to answer questions with a higher accuracy by clearly identifying e.g. instruments (for the instrumental how), risks (via the warnings) and equipment needed (via the prerequisites).

The RST (Mann et al., 1988) is a descriptive theory that specifies 23 possible relations showing how two portions of a text are linked. Previous work on procedural texts (Kosseim, 1995) (Vander Linden, 1993) (Rösner et al., 1992) used limited RST relations and suggested additional relations that fit procedural texts, which we use for our own analysis (limit, alternative, concurrence). We identified 17 relations from our corpora analysis, among which we introduced five new relations: reference, prevention, pre-requisite, option and co-occurrence. Here are the relations we use, with their definitions, possibly slightly altered from their original use:

- **Sequence**: is a multinucleic relations where Nucleuses are linked up by a succession relation (*disconnect the grinder, install the Standard Abrasives Quick Connect, 2-in. holder pad into the chuck and tighten the nut*).
- **Result**: specifies that an action cannot start before a desired result from a previous action is reached (*once all the screw and the plugs positioned, assemble the prefabricated sections*).

- **Purpose**: occurs between a goal and the action meant to reach it (*clean the inside surfaces of the engine block to improve oil return*).
- **Evaluation**: is a little different from the purpose and the result relations because it is possible to evaluate whether the action was made correctly or not (*keep stirring in order to get an unctuous cream*).
- **Limit**: links up an action with a satellite representing a breakpoint (*reduce the sauce by stirring it until the liquid disappear*).
- **Alternative**: links 2 alternative actions, the choice can depend on the subject will or on the situation itself (*if it is screwed in place, remove the screws with a screwdriver, or If the panel hangs on hooks, pull the panel out and swing it up to remove it from the hooks*).
- **Means**: we consider that the manner of doing an action and the instruments needed to realize the actions are linked with the nucleus (the action itself) by a means relation (*you must reduce the air grinder's speed by using the regulator*).
- **Reference**: holds between an action and a segment which provides its procedure localization, in the text or in related texts via hyperlinks (*remove the reductor (see page 18)*).
- **Prerequisites**: occurs between an action or an objective and a list of instruments or a set of actions without which the action or the objective cannot be realized (*changing a car wheel : to change a wheel is not difficult, with the proviso of having in one's car the good tools : wheel brace, jack, clean rag, torch (if dark), warning triangle*).
- **Option**: is considered when an action depends on the realization of a conditional situation. Notice that this relation can also link two sequential actions, where one is compulsory and the other depends on the subject will or on the situation (*steam the fish for 10 minutes and pass it 5 minutes in the oven if you want it to turn golden*).
- **Prevention**: is usually a relation between an action and its warnings. Satellites include expressions such as: *be careful not to ...*, and *'don't'* expressions (*cut the wood planks, don't draw any line!*).
- **Condition**: appears when the action results from the occurrence of a conditioning situation (*if you can't do it by yourself, ask the joiner to cut the frames*).
- **Co-occurrence**: is the word we use for Vander Linden's concurrence relation: the nucleuses are linked by a co-temporal relation (*simultaneously to baking the meat, prepare the vegetables*).
- **Concurrence**: occurs between two rival co-temporal actions (*to choose the best computer, run the program A on Mac, at the same time run the program B on PC. If Mac detects the component before the PC, then use Mac, otherwise use PC*).
- **Motivation**: occurs when the information given in the satellite intends to increase the readers desire to perform the action. Enablements are also part of this category

(you've almost come to the end, now you only have to wait till the flowerings).

The following chart summarizes, for the rhetorical relations we use, the elements in our grammar which are involved.

Rhetorical relations	kernel-satellite or multi-kernel pairs
Sequence	Instruction-imperative linear sequence Instruction sequence - instseq
Result	Goal-imperative linear sequence Goal-optional sequence Goal-alternative sequence Goal-imperative co-temporal sequence Instruction-imperative linear sequence Goal-action
Purpose	Imperative linear sequence - goal Optional sequence - goal Alternative sequence- goal Imperative co-temporal sequence - goal Action - goal-
Evaluation	Goal-imperative linear sequence Goal-optional sequence Goal-alternative sequence Goal-imperative co-temporal sequence Goal-action
Limit	Action - limit
Alternative	Imperative linear sequenc - instruction sequence Instruction sequence - instseq
Means	Action - manner
Reference	Action - reference
Prerequisites	Title-prerequisites
Option	Optional expression-imperative linear sequence Instruction sequence - instseq
Prevention	Title-warning Action - warning
Condition	Imperative linear sequence - optional expression Imperative linear sequence - conditional expression Instruction sequence - instseq
Concurrence	Imperative linear sequence - imperative co-temporal sequence
Co-occurrence	Imperative linear sequence - imperative co-temporal sequence
Motivation	Action - motivation

5 Argumentation in instructional texts

Argumentation is found in the expression of procedural text objectives, in the expression of disjunction, alternatives, warnings, and within instructions.

Let us review here the 4 major forms of arguments we found frequently in corpora. We outline here the main conceptual and syntactic structures that characterize each of these forms. Verb classes referred to are in general those specified in WordNet (Fellbaum, 1998):

- **'objective or goal' arguments:** are the most usual ones. They usually introduce a set of instructions or more locally an instruction. Their target is the "goal" symbol

of the grammar. They basically introduce causality between a goal and the set of instructions that realize it. The abstract schemas are the following: (1) purpose connector-infinitive verb, (2) causal connector-deverbal and (3) titles.

- purpose connectors : pour, afin de, etc. (to, in order to) (e.g. *to remove the bearings, for lubrication of the universal joint shafts*).
- titles : infinitive verbs or deverbals (e.g. *engine dismount*).

- **prevention arguments:** embedded either in a 'positive' or a 'negative' formulation. Their role is basically to explain and to justify. Negative formulation is easy to identify: there are prototypical expressions that introduce the arguments. Negative formulation follows the abstract schemas : (1) negative causal connector-infinitive risk verbs; (2) causal connector-modal +VP(negative polarity, infinitive); (3) negative causal mark-risk verb class VP; (4) causal connector-VP(with negation); (5) causal connector-prevention verb.

- negative causal connectors: sous peine de, sinon, car sinon, sans quoi, etc. (otherwise, under the risk of) (e.g. *sous peine d'attaquer la teinte du bois*).
- risk verb class: risquer, causer, nuire, commettre etc. (e.g. *pour ne pas commettre d'erreur*).
- prevention verbs: éviter, prévenir, etc. (e.g. *afin d'éviter que la carte se déchausse lorsqu'on la visse au châssis*, gloss: *in order to prevent the card from skipping off its rack*).
- causal connector and negation: de façon à ne pas, pour ne pas, pour que ... ne ...pas etc. (in order not to) (e.g. *pour ne pas le rendre brillant*, gloss: *in order not to make it too bright*).
- modal VP: pouvoir, pouvoir-être (e.g. *car il peut être utilisé prématurément par la défaillance d'un autre*, gloss: *because it may be prematurely worn due to the failure of another component*).

Positive formulation marks are the same as for the first category of arguments described above. We have the following abstract schemas: (1) purpose mark-infinitive verb; (2) causal subordination mark-subordinate proposition, (3) causal mark-proposition:

- purpose marks: afin de, pour (so as to, for).
- causal marks: car, c'est pourquoi etc. (e.g. *car ceux-ci sont les plus délicats*).
- causal subordination marks: afin que, pour que, etc. (so that, for).
- the verbs encountered are usually of conservative type : conserver, maintenir, etc.

To discriminate arguments using purpose marks from those of the first class, we can use a reformulation criterion. Positive prevention arguments can be reformulated to a negative form using negative causal connectors or verbal inferences (e.g. *afin que la semence adhère bien au sol* → *car sinon la semence n'adhèrera pas au sol*

(gloss: in order for the crop to adhere to the ground / otherwise the crop will not adhere to the ground)).

- **performing arguments:** These arguments are less imperative than the others, they are rather advices, evaluations. The corresponding abstract schemas are: (1) causal connector-performing NP; (2) causal connector-performing verb; (3) causal connector-modal-performing verb; (4) performing proposition.
 - performing verbs: e.g. *permettre, améliorer, etc. (allow, improve)*.
 - performing NP: e.g. *Pour une meilleure finition; pour des raisons de performances.*
 - performing proposition: e.g. *Have small bills. It's easier to tip and to pay your fare that way.*
- **threatening arguments:** These arguments have a strong impact on the user's intention to realize the instruction provided, the instruction is made compulsory by using this kind of argument. This is the injunctive form. It follows the following schema: (1) otherwise connector-consequence proposition; (2) otherwise negative expression-consequence proposition
 - otherwise connectors: *sinon*.
 - otherwise negative expression: *si ... ne ...pas... (e.g. si vous ne le faites pas, nous le périmons automatiquement après trois semaines en ligne, if you do not do it, we will revoke it immediately)*.

Besides these four main types of arguments, we found some forms of stimulation-evaluation (what you only have to do now...), and evaluation.

6 Injunctive forms

Let us now say a few words about interesting syntactic and morphological characteristics. First, we found no sign of author positioning: there is no use of personal pronoun like 'I' or 'We'. However, the author's enunciation is made visible in French by the use of imperative and infinitive verbal forms. The most important form is certainly the injunctive discourse. It characterizes certain modalities of discourse: orders, preventions, warnings, avoidances, advices. These all have a strong volitive and deontic dimension.

Injunctive discourse shows how the author of an instructional text imposes his point of view to the user. The goal is that the user knows how to execute it in a way as explicit and less ambiguous as possible. The user is assumed to have the required competences to realize it.

Instructional texts are an example of a logic of action. Injunction is particularly frequent in cooking receipes, security notices, etc. Its strength is measured via the illocutionary force of the statement. In general we observed that infinitive or imperative modes are used in French. Some examples of injunction forms are given below, from which we could construct dedicated NLG templates:

- infinitive: *Mettre la poudre dans le verre* (put the powder in the glass).

- imperative: *Enlevez la bague supérieure du bol d'articulation à l'aide d'un burin* (gloss: *remove upper bushing from socket using a chisel*).
- modal verbs: *Vous devez enduire la face intérieure du pivot de pâte d'étanchéité SILICOMET* (gloss: *you must coat internal face of pivot with SILICOMET sealing compound*).
- preference expression : *"il est conseillé de ...", "nous vous recommandons de ...", "il est préférable de ..."* (it is advised to, we recommend that).
- negative infinitive syntax form: *Ne pas utiliser de façon prolongée sans avis médical* do not use on the long term without medical advice.

In everyday life, we encounter many injunctions posted in public areas, in French these injunctions follow in general these regular structural schemas:

- deverbal-infinitive (e.g. *défense d'afficher* (gloss: *stick no bills*)).
- courtesy formula-negative infinitive (e.g. *prière de ne pas fumer* (gloss: *no smoking (please)*)).

7 Conclusion

In this paper, we briefly shown the variety of structures, rational and somewhat irrational that organize instructional texts. We conducted this research with the main goal of generating responses in a cooperative way to *How?* and *Why?* questions. The long-term goals are to select the best text w.r.t. a user profile, and then to be able to integrate texts on the same topic to get a better text.

This preliminary step is now stabilized, and we designed an annotation tool, based on the grammar and related marks, to implement and evaluate our results. However, to get an more accurate view of the diversity of argumentation in this type of text, we need to also consider more subtle language forms such as: modalisators, tonality, opinion marks, evaluation marks, illocutionary force measures in injunctions, etc.

References

- [1] Adam, J.M., *Types de Textes ou genres de Discours ? Comment Classer les Textes qui Disent De et Comment Faire*, Langages, 141, pp. 10-27, 2001.
- [2] Adam, J.M., *Types de Sequences Textuelles Elementaires*, Pratiques n56, Metz, 1987.
- [3] Amgoud, L., Parsons, S., Maudet, N., *Arguments, Dialogue, and Negotiation*, in: 14th European Conference on Artificial Intelligence, Berlin, 2001.
- [4] Amgoud, L., Bonnefon, J.F., Prade, H., *An Argumentation-based Approach to Multiple Criteria Decision*, in 8th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty, ECSQARU'2005, Barcelona, 2005.
- [5] Anscombre, J.C., Ducrot, O., *Interrogation et Argumentation*, in Langue française, no 52, L'interrogation, 5 - 22, 1981.

- [6] Aouladomar, F., *A Semantic Analysis of Instructional Texts*, IWCS05, Tilburg, 2005.
- [7] Aouladomar, F., *Towards Answering Procedural Questions*, Workshop KRAQ05, IJCAI05, Edinburgh, 2005 (under submission).
- [8] Benamara, F., Saint-Dizier, P., *Advanced Relaxation for Cooperative Question Answering*, in: New Directions in Question Answering, in Mark T. Maybury, (ed), AAAI/MIT Press, 2004.
- [9] Bieger, G.R., Glock, M.D., *The Information Content of Picture-text Instructions*, Journal of Experimental Education, 53, 68-76, 1984-85.
- [10] Delin, J., Hartley, A., Paris, C., Scott, D., Vander Linden, K., *Expressing Procedural Relationships in Multilingual Instructions*, Proceedings of the Seventh International Workshop on Natural Language Generation, pp. 61-70, Maine, USA, 1994.
- [11] Ellahad, M., *Using Argumentation in Text generation*, Journal of Pragmatics, vol. 24 (1995).
- [12] Eggs, E., *Grammaire du Discours Argumentatif. Le Topique, le Générique, le Figuré* Editions Kimé, Paris, 1994.
- [13] Fellbaum, C., *WordNet An Electronic Lexical Database*, The MIT Press, 1998.
- [14] Greimas, A., *La Soupe au Pistou ou la Conservation d'un Objet de Valeur*, in Du sens II, Seuil, Paris, 1983.
- [15] Kosseim, L., Lapalme, G., *Choosing Rhetorical Structures to Plan Instructional Texts*, Computational Intelligence, Blackwell, Boston, 2000.
- [16] Longacre, R., *Discourse Typology in Relation to Language Typology*, Sture Allen éd., Text Processing, Proceeding of Nobel Symposium 51, Stockholm, Almqvist and Wiksell, 457-486, 1982.
- [17] Luc, C., Mojahid, M., Virbel, J., Garcia-Debanç, C., Pery-Woodley, M-P., *A Linguistic Approach to Some Parameters of Layout: A study of enumerations*, In R. Power and D. Scott (Eds.), *Using Layout for the Generation, Understanding or Retrieval of Documents*, AAAI 1999 Fall Symposium, pp. 20-29, 1999.
- [18] Lüger, H.H., *Pressesprache*, Tübingen, Niemeyer, 1995.
- [19] Mann, W., Thompson, S., *Rhetorical Structure Theory: Towards a Functional Theory of Text Organisation*, TEXT 8 (3) pp. 243-281, 1988.
- [20] de Mattos Pimenta Parente, M-A., Steffen Holderbaum, C., Virbel J., Nespoulous, J-L., *Text Questionability as a predictor of story recall*, Thirteen Annual Meeting of the Society for Text Understanding, Madrid, 2003.
- [21] Mortara Garavelli, B., *Tipologia dei Testi*, in G. Hodus et al.: *lexicon der romanistischen Linguistik*, vol. IV, Tübingen, Niemeyer, 1988.
- [22] Moschler, J., *Argumentation et Conversation, Eléments pour une Analyse Pragmatique du Discours*, Hatier - Crédif, 1985.
- [23] Muller, P., Tannier, X., *Annotating and Measuring Temporal Relations in Texts*, In Proceedings of Coling 2004, volume I, pages 50-56, Geneve, 2004.
- [24] Qamar, H., *Quand Dire c'est: Ecrire-Comment-faire. Un Autre Type de Texte: le RECETTAL*, these soutenue l'Universite Lumiere, Lyon II, 1996.
- [25] De Rijke, M., *Question Answering: What's Next?*, the Sixth International Workshop on Computational Semantics, Tilburg, 2005.
- [26] Rösner, D., Stede, M., *Customizing RST for the Automatic Production of Technical Manuals*, in R. Dale, E. Hovy, D. Rosner and O. Stock eds., *Aspects of Automated Natural Language Generation*, Lecture Notes in Artificial Intelligence, pp. 199-214, Springer-Verlag, 1992.
- [27] Saint-Dizier, P., *Verb Semantic Classes Based on 'Alternations' and WordNet-like criteria*, in : *Predicative Forms in Natural language and lexical Knowledge Bases*, Eds: Saint-Dizier,P., Eds: Kluwer Academic, Cambridge, USA, 1998.
- [28] Saint-Dizier, P., *PrepNet: a Framework for Describing Prepositions: Preliminary Investigation Results*, the Sixth International Workshop on Computational Semantics, Tilburg, 2005.
- [29] Schwitter, R., Rinaldi, F., Clematide, S., *The Importance Of How-Questions in Technical Domains*, TALN, Workshop Question-Reponse, Fes, Maroc, 2004.
- [30] Takechi, M., Tokunaga, T., Matsumoto, Y., Tanaka, H., *Feature Selection in Categorizing Procedural Expressions*, The Sixth International Workshop on Information Retrieval with Asian Languages (IRAL2003), pp.49-56, 2003.
- [31] Vander Linden, K., *Speaking of Actions Choosing Rhetorical Status and Grammatical Form in Instructional Text Generation Thesis*, University of Colorado, 1993.
- [32] Werlich, E., *Typologie der Texte*, Heidelberg, Quelle and Meyer, 1975.
- [33] Yin, L., *Topic Analysis and Answering Procedural Questions*, Information Technology Research Institute Technical Report Series, ITRI-04-14, University of Brighton, UK, 2004.
- [34] Zuckerman, I., McConachy, R., Korb, K., *Using Argumentation Strategies in Automated Argument Generation*, INLG2000.

```

< alt_seq >
  < imper_seq >
    < instr > Don't take any antibiotic
      for an appendicitis. < instr >
  < imper_seq >
  < opp_mark > But < \ opp_mark >
  < instseq >
    < alt_seq >
      < cond_exp > If you suffer from an
        appendicitis with appendix rupture
      < \ cond_exp >
      < imper_seq >
        < instr > you have to take Augmentin
          150mg/kg daily at 3 intervals
          during 2 days. < instr >
      < imper_seq >
      < instseq >
        < alt_seq >
          < cond_exp > If in addition,
            you're allergic,
          < \ cond_exp >
          < imper_seq >
            < instr > you have to take
              instead Flagyl 0.5g every
              8 hours and Gentamicine
              5mg/kg daily during 2 days.
            < instr >
          < imper_seq >
        < \ alt_seq >
      < \ instseq >
    < \ alt_seq >
  < \ instseq >
< \ alt_seq >

```

Figure 1: An example of an annotated instructional text