

An Exploration of the Diversity of Natural Argumentation in Instructional Texts

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 user must follow step by step the instructions in order to reach the results expected. In this short paper, we explore the different facets of natural argumentation used in such texts. Our study is based on an extensive corpus study, and within a language generation perspective.

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.

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, 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, do-it-yourself and assembly notices, itinerary guides, advice texts, savoir-faire guides, etc.

The diversity of procedural texts, their objectives and the way they are written is the source of a large variety of natural arguments. We briefly present them in this paper. 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), (Anscombe et al. 1981))

and philosophical views of argumentation. It is basically linguistic and conceptual.

2 Procedural texts and argumentation

2.1 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.

2.2 The art of writing efficient procedural texts

The author of procedural texts must consider three dimensions (Donin et al., 1992): (1) cognitive: notions referred to

must be mastered and understood by the target users, (2) episodic: 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, 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 the 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.).

3 A discursive analysis of procedural texts and argumentation

Argumentation is a process used by a person to convince an audience (Olron, 1983). Instructional texts are a form of argumentation structure since they (1) make interact the instructions producer and receiver, (2) are also a process that exert an influence on the receiver (the user must realize the instructions), (3) give justifications or elements that prove the appropriateness of the instruction, using rational elements (see our argument typology in section 4).

Plan - Goal argumental structure can be used also to describe instructional texts.

Here is, represented by means of a grammar, the main elements of the structure of procedural texts, see (Aouladomar

2005) for more details. From a methodological point of view, our approach is based on corpora analysis : first we use corpora to extract the structure of procedural texts and then we validate and enrich the grammar by using another corpora.

The structure reported below essentially corresponds to (1) the organization of the informational contents: how tasks are planned, according to goals and subgoals, and (2) the argumentative strategies deployed (via planning, progression of tasks, warnings, advices, evaluations, etc.).

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**, **argument+**, (**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.

There are many ways to organize an argumentation. Here is an example of the argumentation structure in an instruction (**goal**): *Pour nettoyer les cuirs.*, (**instruction**): *choisissez les produits spécifiques ameublement*, (**advice**): *et préférez-les incolores*, (**multiple consequences**): (**argument 1**): *ils vont jouer un rôle d'entretien et de protection*, (**argument 2**): *ajouter à la beauté des peaux* (**argument 3**): *et permettre de réparer certains dommages*.

(Gloss: To clean leathers, choose products specific to furniture. Prefer them colorless, they will care of and protect leathers, adding beauty, and will allow an easy repair of some damages).

4 Forms of arguments in instructions

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. 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 sequences or more locally an instruction implemented in the "goal" symbol of the grammar. The abstract schemas are the following: (1) purpose connectors-infinitive verbs, (2) causal connectors-deverbals and (3) titles.
 - causal 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: negative causal connectors-infinitive risk verbs ; causal connectors-modal VP-infinitive verbs; negative causal marks-risk VP; positive

causal connectors-VP negative syntactic forms, positive causal connectors-prevention verbs.

- negative 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 class verbs: *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*).
- positive causal mark and negative syntactic forms: *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 SV: *pouvoir*, *pouvoir-être* (e.g. *car il peut être usé 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: purpose marks-infinitive verb; causal subordination mark-subordinate proposition, 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*).

- **performing arguments**: These arguments are less imperative than the other ones, they are rather advices, evaluations. The corresponding abstract schemas are: causal connectors-performing NP; causal connectors-performing verbs; causal connectors-modal-performing verbs; 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: otherwise connectors-consequence proposition; 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*).

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

5 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 know 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 recipes, 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:

- 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 ..."*.
- negative infinitive syntax form: *Ne pas utiliser de façon prolongée sans avis médical*.

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)*)).

6 Conclusion

In this short paper, we briefly showed the variety of natural argumentation forms found in instructional texts. We conducted this research with the main goal of answering in a cooperative way *How?* and *Why?* questions.

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.

Let us conclude with some interesting observations and remarks that need to be developed.

- Using emphasizing arguments the importance of following instructions, can depend on the nature of procedural texts. Preliminary observations tend to show that technical procedural texts are richer in argumentation than, for instance, recipes. Arguments in procedural texts seem to depend also on the complexity or "dangerousness" of the procedure. Further investigations in those directions would be interesting to carry out.
- Arguments can depend on the user's expertise or familiarity with the domain. In that case, arguments are means, for the producer, of adapting his writing strategies according to the user

All these considerations can lead to recommendations for writing assistance tools. Results can also help to develop different strategies when generating answers to procedural questions by: (1) adding arguments in the answer of a QA system which is a way to give justifications to the generated answer ; (2) using arguments adapted to user levels.

References

- [1] Amgoud, L., Parsons, S., Maudet, N., *Arguments, Dialogue, and Negotiation*, in: 14th European Conference on Artificial Intelligence, Berlin, 2001.
- [2] 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.
- [3] Anscombre, J.-Cl. Ducrot, O., *Interrogation et Argumentation*, in *Langue française*, no 52, L'interrogation, 5 - 22, 1981.
- [4] Aouladomar, F., *A Semantic Analysis of Instructional Texts*, IWCS05, Tilburg, 2005.
- [5] Aouladomar, F., *Towards Answering Procedural Questions*, Workshop Knowledge and Reasoning for Answering Questions, IJCAI05, Edinburgh, 2005.
- [6] Donin, J., Bracewell, R. J., Frederiksen, C. H., Dillinger, M., *Students Strategies for Writing Instructions: Organizing Conceptual Information in text*, *Written Communication* 9, 209-236, 1992.
- [7] Fellbaum, C., *WordNet An Electronic Lexical Database*, The MIT Press, 1998.
- [8] Moschler, J., *Argumentation et Conversation, Eléments pour une Analyse Pragmatique du Discours*, Hatier - Crédif, 1985.