Rhetorical Organization Model for business-based technical documents

Modelo de Organización Retórica para documentos técnicos organizacionales

Bell Manrique Losada
Universidad de Medellín
Cra 87 No. 30-65, Medellín
bmanrique@udem.edu.co

Carlos Mario Zapata
Universidad Nacional de Colombia
Cra. 80 No. 65-223, Medellín
cmzapata@unal.edu.co

René Venegas
Universidad Católica Valparaíso
Av. El Bosque 1290, Viña del Mar
rene.venegas@ucv.cl

Abstract: Rhetorical analysis is concerned with discourse organization focused on the communicative purpose of texts. In this paper we propose a first approach to a rhetorical organization model related to business-based technical documents. By using discourse rhetorical analysis and genre analysis, we define such a model in terms of functional and structural features, based on macro-moves, moves, and functions comprising a kind of documents called *standard operating procedure*. So, we propose a functional-structural framework of the discourse, influencing and restricting the discourse contents and style of each genre. We are using this model as an analytical-conceptual basis for mapping business-based technical documents and generating a controlled language output.

Keywords: Rhetorical analysis, specialized discourse, technical documents, functional and structural model.

Resumen: El análisis retórico se relaciona con la forma en que se construyen los discursos, dando prioridad al propósito comunicativo de los textos. En este artículo se propone una primera propuesta de modelo de organización retórica relacionada con documentos técnicos organizacionales. Por medio del análisis retórico del discurso y el análisis de género, se define el modelo en términos de aspectos estructurales y funcionales, basado en las macro-movidas, movidas y funciones, que se encuentran en el tipo de documento llamado *standard operating procedure*. Así, se propone un marco funcional-estructural del discurso, influenciando y restringiendo los contenidos y estilo de cada género. Este modelo se utiliza como base analítico-conceptual para proyectar desde documentos técnicos de negocios la generación de un lenguaje controlado de salida.

Palabras clave: Análisis retórico, discurso especializado, documentos técnicos, modelo estructural y funcional.

1 Introduction

Discourse analysis is a vast field comprising interpretive and explanatory sub-disciplines. Commonly, such analysis is concerned with the relationships between form and function of discourse. Discourse analysis includes a particular process to treat specialized text (Biber, 2006; Nickerson, 1999). Specialized texts—generated from a specialized organizational discourse—are produced by specialists who have mastered the cognitive and conceptual organization of matter.

We follow the analysis of such specialized discourses from the genre point of view (Nickerson, 1999). Our approach for analyzing discourses is based on the rhetorical analysis approaches (van Nus, 1999; Swales, 1990; 2004).

Rhetorical analysis is concerned with the discourse construction focused on the communicative purpose of texts. We follow discourse rhetorical analysis (DRA) and genre analysis for defining a first approach to a rhetorical organization model in terms of functional and structural features comprised in a

standard operating procedure (SOP). SOP is as a kind of business-based technical document.

The structural aspects are based on macro-moves and moves. Structural moves describe texts in function of their formal features either concerning global organization or sentence-level features. The approach is focused on structural move analysis developed by Swales (1990).

Formal features are the rhetorical purposes for expressing the communicative intentions of the document authors. The structure and functions influence and restrict the contents and style of a text (Askehave & Swales, 2001) and allow for the identification of linguistics features. The proposed model is being used as the core of a mapping framework of business-based technical documents which can generate a controlled language output based on the model proposed.

The remainder sections of this paper are organized as follows: in Section 2 we describe the theoretical framework and some background in the field of discourse and rhetorical analysis. In Section 3 we propose our approach to rhetorical organization model. Finally, in Section 4 we discuss conclusions and outline future work.

2   Theoretical Framework and related work

In this work we are focused on perspectives of discourse analysis, based on the usage of language for constructing, interpreting, and exploiting technical documents.

2.1   Discourse Analysis

Language in social use is studied by researchers in several areas, and discourse analysis is one of them. Commonly, the relationships between form and function in language (Gee, 2004) are the focus of such researchers. Discourse analysis is interpretive and explanatory and is based on a systematic methodology. The study of discourse is a vast field comprising many overlapping sub-disciplines, such as sociolinguistics and semiotics, specialized discourse analysis, and discourse genre studies.

Specialized text involves a particular process in the framework of discourse analysis (Biber, 2006; Nickerson 1999). An approach to discourse analysis is based on the genre. According to Swales (1990; 2004) and Yates (1989) genres are defined as variations of a language which operate by means of linguistic features present in a text. Likewise, they are linguistically linked to their communicative purposes, participants involved, production contexts, usage contexts, and discourse organization modes, among others (Parodi, Ibañez, and Venegas, 2010).

The genre theory (Nickerson, 1999; van Nus, 1999) focuses on written practices of members from specific communities and also on the design of information and business records. Specialized texts—generated from a specialized organizational discourse—are produced by specialists who have mastered the cognitive and conceptual organization of a matter. According to Cabré (1999), the specialized discourse is derived from variables related to the subject and perspective of a topic, and the intent and level of expertise of the producer. According to this genre theory, we can analyze specialized discourses, because they are written by members of specific communities.

Genres are described in terms of texts structures and their contexts (van Dijk, 2008). Genre analysis is an attempt to relate the text structure to the macro-social context. Commonly, genre analysis is characterized by recognizable purposes and schematic structures, which can be as numerous as the social practices which people get involved in. Genre studies have been developed from the seminal work by Bakhtin (1986), the new rhetoric by Swales (1981; 1990), and the research within Systemic functional linguistics by Christie (1999) and Eggins (1994). According to the previous foundations and the approach of Meurer (2002), we follow the genre analysis based on the definition of genres as reasonably stable types of formal or informal text, which can be recognized due to their rhetorical structure and function.

2.2   Rhetorical Analysis

Rhetorical analysis (RA) is concerned with the discourse construction, prioritizing on the communicative purpose of each genre (Azaustre and Casas, 1997). Rhetoric supports discourse from its intentional and instrumental point of view. Thus, the rhetorical organization of discourse is the purpose-driven way in which textual structures are employed in order to achieve a desired effect (Connor, 1996). Structure constitutes a framework where

diverse discourses are articulated in a particular way and their relations are textually constituted.

From RA, the genre analysis is discussed in terms of rhetorical moves, which refer to the functional parts or sections of a genre. This approach for studying a particular genre comprises the analysis of a text and its description in terms of rhetorical structure (moves). The particular configuration of the text surface is defined in terms of levels of text organization, also known as rhetorical discourse organization. In this way, the macro units identified by genre analysis can be characterized as moves and steps, referring to the passages of the text which are larger than the largest grammatical units—*e.g.* clauses, complex clauses, and sentences—and possess some unity grounded in a common function/meaning.

2.3 Corpus Linguistics

Corpus Linguistics is a subfield of Linguistics for studying the linguistic properties of an extended passage, text, or corpus of texts. Such a definition includes Semiotics and Genre Analysis perspectives, and the computational analysis of text corpora.

According to Parodi (2008), corpus linguistics comprises a set of methodological principles for studying any language domain. Sets of linguistic features—operated by genres—can be identified from representative corpus. Currently, corpus-based approaches have been widely applied to many fields such as discourse analysis, language teaching, and stylistic analysis (Kennedy, 2000).

Biber, Conrad, and Reppen (1998) summarize features of corpus-based studies as an empirical approach, on the basis of a natural texts sample, computer-assisted analysis, and combination of quantitative and qualitative analysis. According to Sinclair (1991), working with a specialized language corpus constitutes a strict selection of texts for identifying common patterns. From a general sample of texts, many words and phrases are scarce, but they are more frequent in a strict selection of text belonging to the same king. Thus, a corpus should be representative of one or more aspects of a language.

3 Rhetorical Organization Model

In this paper, we propose a first approach of a rhetorical organization model (ROM), based on structural and functional features. The methodological procedures applied for defining the ROM in this study are divided into stages, as follows:

3.1 Corpus definition

The corpus definition starts by collecting possible technical documents on the macro-genre circulating on the web. We explore broadly 4 types of technical documents: Job Description Documents, Corporate Policy Documents, Function Manuals, and Standard Operating Procedures. Based on the analysis of collected documents, we select as a referent for the case study the 'Standard Operating Procedure' corpus. The population selection for the corpus is based on the following criteria: Named as 'Standard Operating Procedure' (SOP) or 'SOP Manual'; written in English; published online and open-accessed in the Internet; having an author affiliated to a company/corporation; and text-based, with a low percentage of images.

We use a methodology based on corpus linguistics (Simpson & Swales 2001; Tognini-Bonelli, 2001) for analyzing the documents from corpus, according to the proposal developed by Parodi (2005). We aimed to perform a descriptive, procedure-focused analysis of linguistic texts from business-based technical documents. We define a hierarchy with the macro-genre, based on the genre analysis theory and the processes of searching, reviewing and analyzing documents from the Web. The qualitative characterization following the genre point of view (Swales, 2004) has the manual macro-genre and the procedures manual genre. The term *manual* applies to academic manuals, instructional and teaching textbooks, and technical procedure manuals (Parodi, 2008). In this work, we use the meaning of procedure manual referring to technical procedure manuals. Also, we can identify several sub-types, using the idea of genre, as we present in Figure 1 and Figure 2.

Specifically, we use SOP as a linguistic genre. A SOP is a constitutive document of a quality system describing a set of recurring operations. A SOP describes procedures defined as segment of business processes and

describes how the policies are effectively implemented.

The motivation for defining SOP as the document under our research is based on such kind of technical document are used in organizations for specifying technical, administrative, and operational activities. SOPs are part of the 'manual' genre as well as typical manuals—*e.g.*, procedure, quality, and user manuals. A manual is a set of written instructions describing how procedures are defined, developed, and managed by organization members. We identify the SOPs as documents containing procedural information, which is the most important information type included in procedure manuals (Karreman and Steehouder, 2003), as a source for extracting domain knowledge and business information.
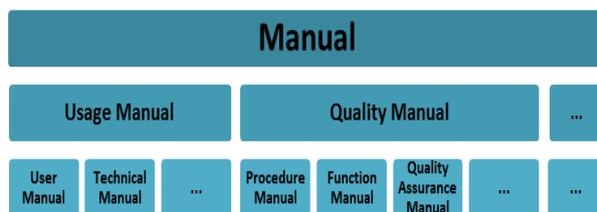

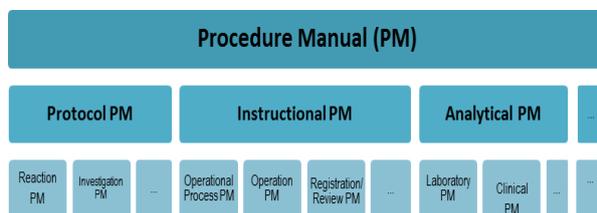Figure 1: Hierarchy of 'manual' *macro-genre*

Figure 2: Classification of the 'procedure manual' *genre*

We establish the criteria for the web searching and the formation of the corpus, and then perform a sampling procedure. The documents collected in the corpus correspond to 100% of the population—forming a significant sample, from a proportional estimate sample size, to a confidence level of 95% precision and an approximate value of 5% the measured parameter. Assuming an evenly distributed population, we selected a sample of 32 documents, corresponding to 64% of the total population. This is the minimum percentage statistically randomized, calculated with the Z-test of proportions. Then, we conducted the rhetorical analysis based on this sample.

3.2 Analysis of the digital corpus.

The analysis carried out on the corpus is related to tokenization, keyword, and stop-word identification, characterization, and creation of word frequency lists, among others. Also, a concordance analysis is carried out by looking for identified lexical forms and every occurrence of a given word together with some context. In this way, we identify a list of all of the occurrences of given search terms in a corpus, presented within the context in which they occur. The corpus analysis considered 9252 word types and 167905 Tokens.

3.3 ROM Design.

3.3.1 Defining a reference model.

Concerning discourse mapping, the first step is searching for the theoretical models in the state of the art, to be used as a reference to the rhetorical analysis of the genre under study. Some approaches we found are concerned to government and business documents (Trosborg, 2000; Renkema, 2003; McCarthy and Handford 2004; Warren, 2004) and commercial documents (Yates, 1989; Jameson, 2008; Freedman and Medway 1995). Directly oriented to the genres related with SOP, we found no references. For this reason and following Burdiles (2011), we use an inductive method for defining the preliminary model, as follows.

First, we randomly select four sample documents from the corpus. Then, we develop an incremental construction of a preliminary model from a by-hand review of the structure and superstructure, in order to identify the common organization units in the sample (moves). Incrementally—and according to the genre analysis—we define rhetorical moves as functional sections of a genre (Swales, 1990; 2004). Based on Parodi (2008), we adopt the macro-move concept which refers to a higher abstraction of rhetorical purpose than a move. Thus, each macro-move serves a communicative purpose and all macro-moves shape the overall organization of the text.

By analyzing the functional organization of a document, we identify purposes of a higher level hierarchy (macro-purposes), which comprise a set of more specific moves, which in turn will be composed of more detailed steps. Consequently, we define a preliminary model as a set of functional and structural features, resulted from the identification of moves with

recurrent presence (likely mandatory level). The mandatory moves are selected based on a set of categories we define for evaluating each move in each document from the sub-corpus. Such categories are related to the percentage of occurrence by move as follows: 0 (does not appear in the document—occurrence zero); 1–30% (low chance); 31–70% (average chance); 71–100% (mandatory). In the reference model we consider the moves as placed in the high percentages and comprise three macro-moves containing 19 moves—showing more specific functional units—as we show in Figure 1 and describe as follows.

Macro-move I
Presenting SOP
- Move 1: Identifying SOP
- Move 2: Organizing SOP
- Move 3: Introduction.
- Move 4: Presenting Foreword
- Move 5: Documenting Conventions.
- Move 6: Appointing regulations or regulatory requirements
- Move 7: Giving acknowledgments.
- Move 8: Defining Intended Audience and Reading Suggestions.
- Move 9: Establishing Purpose

Macro-move II
Developing procedures
- Move 10: Defining procedure purpose.
- Move 11: Defining roles and responsibilities.
- Move 12: Identifying prerequisites.
- Move 13: Listing definitions.
- Move 14: Listing resources.
- Move 15: Establishing methods
- Move 17: Including references

Macro-move III
Ending SOP
- Move 18: Adding Supplementary information.
- Move 19: Including references

Figure 3: Preliminary Rhetorical Organization Model

Macro-move I: Presenting SOP. This is a preliminary statement presenting an introduction of the document, describing the document purpose, conventions, revision schedule, approval authority, and document organization, among others.
- Move 1: Identifying SOP. Identifies the organization authoring the SOP, including: author, company, location, filiation, name, and verbal or nonverbal identification.
- Move 2: Organizing SOP. Includes aspects of the document body, related to content organization, lists of tables, and lists of figures, among others. This move allows for the reader to locate the document content.

This section should present the entire hierarchical organization—divisions and major subdivisions—of the document, preferably with a respective list.
- Move 3: Introduction. Justifies and presents the document. This move describes a general view of the related context.
- Move 4: Presenting foreword. Presents a general review of the document and describes what is included in each procedure. Also, it can describe those who participated in writing the SOP, how it was organized, how to read it, the re-view process that took place, and warnings about its use and distribution.
- Move 5: Documenting conventions. Gives the reader the current context of the document: date of approval, version number, author, and revision number.
- Move 6: Appointing regulations or regulatory requirements. Name standards, contractual requirements, policy, or regulations associated with the procedures included in the SOP. It can include lists of references.
- Move 7: Giving acknowledgments. Presents the compendium of helpers, people, or individuals acknowledged for their contributions in writing the SOP.
- Move 8: Defining intended audience and reading suggestions. Defines the primary audience for the SOP. It can include management team, operational team, and staff of the organization.
- Move 9: Establishing purpose. Describe the general goal of the procedures included inside the SOP, within the organization framework. This goal is oriented to contextualization and purpose description.

Macro-move II: Developing procedures. It presents the procedures associated with each organizational process in detail. Throughout, a series of specific purposes, responsibilities and functions, procedural descriptions, and rules for implementation are defined.
- Move 10: Defining procedure purpose. Defines the purpose of each procedure.
- Move 11: Defining roles and responsibilities. Defines the purpose of each procedure.
- Move 12: Identifying prerequisites. Identifies the requisites previous to the execution of the procedure. It may include

rules, cautions, warnings, and recommendations for achieving them.

- Move 13: Listing definitions. Includes a list of definitions, concepts, and terms of acronyms used in the context of SOP.
- Move 14: Listing resources. Lists the equipment, resources, and material required for the execution of the procedure.
- Move 15: Establishing methods. Establishes the methods used to characterize or guide the procedure.
- Move 16: Specifying procedure. Provides step-by-step instructions to ensure details of procedures.
- Move 17: Including references. Lists bibliographical references included in the procedure.

Macro-move III: Ending SOP. It is related to the moves I and II. It is optional, but it is intended to supplement the development macro-move.

- Move 18: Adding supplementary information. Includes attachments for supporting the development macro-move.
- Move 19: Including references. It lists bibliographical references

The backbone of this genre lies in macro-move II: development. It acts as a unit repeated as many times as necessary for covering the detailed development of the procedures. It occurs as a feature of the SOP genre. The preamble macro-move does not operate in the same way as the macro-move II, as its inclusion only occurs once throughout the text. The latter is optional, but complements the previous one in some cases, depending on the information needed to support the procedures. Finally, we develop a peer review for analyzing the preliminary model, and generate a reference model based on the comments of the experts.

3.3.2 Rhetorical analysis method

The rhetorical analysis method is based on the reference model and comprises a set of by-hand activities or meditated by computational tools

like as AntConc 3.3.5w®[1], TermoStatWeb[2], ElGrial[3], and NLTK-Demos[4].

The aforementioned activities are the following:

i. Analysis and identification of rhetorical units

This activity is focused on the registration of the analysis grid including the following information: identifier, summary, SOP category, macro-move identification and name, move identification and name, linguistic features identified, and an example. In addition, a process for identifying lexical and grammatical tracks from each rhetorical unit was developed.

ii. Definition of mandatory of rhetorical units

Each rhetorical unit from the SOP is considered optional/mandatory according to the dependency with the purpose of the macro-structure and the author communicative intention. Such a recurrent presence generates the mandatory level of the move.

iii. Corpus preprocessing

The collected texts in the corpus are converted into .txt format by using certain transformation actions. The non-relevant information for text processing—e.g. HTML tags, webpage names, ads, etc.—is manually removed from SOPs. Tools like HTML Text and Multireplacer3 can be also used.

iv. Identification of linguistics features by rhetorical unit

We perform an observation of several variables: word frequency according to mutual information; parameters of probable co-occurrence among words; lexical items recurrent in given stretches of text; keywords analyzed; and frequency of the words, among others. We perform the identification of prototypical lexical-grammatical features from the corpus sample by using several

---

[1] AntConc 3.3.5w®. A freeware concordance program (Anthony, 2009)

[2] TermoStatWeb™. A term extractor that uses linguistic and statistical methods based on the potential terms' structures and relative frequencies from corpus (Drouin, 2003)

[3] ElGrial®. A computational interface for tagging and parsing texts from corpora collected and stored online Available in: www.elgrial.cl

[4] NLTK-Demos. Tool with different word tokenizers, stemming text, part of speech tagging, and chunk extraction modules Available in: http://text-processing.com/demo/

computational tools. Some of such features are essential to current writing in specialized organizational fields and are presented in the Table 1, according to the proposal of Biber *et al.* (2004) and Venegas (2005).

v.   Morphosyntactic tagging

This analysis level aims to identify morphosyntactic categories and then features, surrounding each rhetorical unit. We describe the features for such categories based on the feature object. The features are described by the following categories belonging to the feature object: noun, adjective, verb, subordination, coordination, syntactic clauses, verbal mood, verbal periphrasis, and person. The feature description is presented with the move including them.

Category	Feature Object	Feature Description	Move
Morphological	Noun	Extensive use of abstract nominalizations as a common feature in technical documents (e.g. *Management, establishment*)	2, 5, 10, 11
Morphological	Noun	Most common are uncountable nous (e.g. *information, process, data*)	12, 13
Morphological	Adjective	Most common demonstratives: *this*, *that*.	1, 2, 10
Morphological	Adjective	Scarce usage of posesives like *its* or *his/her*	-
Morphological	Adjective	Undefined. They are rarely used (e.g. *any*)	12, 13
Morphological	Adjective	The qualifying most common are: *required, apropriate, successful, complete*. Their use is mainly subjective, rather than descriptive	11, 12, 13
Morphological	Verb	The copulative verb are the most used mode in SOP. Among them *be, is, are*.	5, 11, 12, 13
Morphological	Verb	Can be conjugated in specific mode, commonly to describe steps (e.g. *may*).	4, 13
Morphological	Verb	Mainly conjugated in indicative mode. (e.g. *use, review, report, describe*)	5, 10, 11, 13
Syntactical	Subordination	Common place and time subordinate (using *where, when, after…*)	11, 13
Syntactical	Subordination	Adjective or noun subordinate clause are used, usually begins with a relative pronoun (e.g. *that, how*) or using *as, before*, or *until*	11, 12, 13
Syntactical	Coordination	The most frequent are c*opulative coordinate clauses* to express adding one stamentent to another. They are introduced by conjunctions (e.g. *and, both, as well as*)	4, 5, 11, 12, 13
Syntactical	Coordination	The only explanatory coordinate clause found was *that is*.	4, 12, 13
Syntactical	Syntactic groups	Usage of nominal phrases (adjective+ noun) is mainly common in the roles and responsibilities section, and in the SOP steps (*e.g. appropriate project*)	11, 13
Syntactical	Verbal mood	Main moods: indicative and imperative, used to express capacity or obligation.	All of the moves
Syntactical	Verbal mood	Mood in conditional steps denoting positiblity/priority (*e.g. is possible, is important, etc.*)	4, 12, 13
	Verbal periphrasis	Modal periphrasis, indicating obligation (*e.g. may+verb, must+verb*) and its negative form or indicating potencial mood (*e.g. Can+verb…*)	4, 12, 13
	Verbal periphrasis	Aspectual periphrasis occurs as: initiation (e.g. *take*); progression (*including performing*); habits (*used to, used for*); duration (*is/are +gerund*); completion (*end of + infinitive*)	13
	Person	Prevails first person of plural in the sequence of procedures. In section 'roles an responsibilities' third person of singular is used. Low usage of pronouns has been detected.	All of the moves

Note: English is not a way for the subjunctive; there semantics of the subjunctive, but English is expressed by the present or the past. (See the complete proposal in Manrique *et al.* 2013).

Table 1: Prototypical lexical-grammatical features

vi.   Identification of occurrence frequencies for relevant features

We analyze relevant features and identify their occurrence frequencies, by using computational tools supporting the corpus-linguistic method.

Feature Name	Feature Description
Language Function	It is informative or referential focused on the message, external reality, or referent. The writer makes no judgments about the information or processes.
Descriptive sequence	Clear language, accurate, and direct. In most of cases clarity in the exposition of ideas can be detected. Descriptive style and usage of names, locations, and qualifying objects, people, or processes. Also, it may have actions in a temporal order looking for a specific purpose.
Terminology	Contains a large number of terms belonging to the body of knowledge and a lot of information shared between producer and reader. Such information is produced by individuals who possess specific knowledge of a subject.

Table 2: Prototypical Functional and Lexical Features

3.4   Evaluation of structural and functional features

We address this activity to determine what rhetorical moves are nuclear or core moves for writing a SOP, as a part of the methodology based on corpus linguistics. We follow the evaluation activities as we describe below:

i.   Selecting the experts for the evaluation of the reference model (comprising 19 moves as we presented in Figure 3).

ii. Designing a template for the evaluation, based on the reference model and considering the following components: rhetorical unit of reference, example extracted from a SOP, evaluation, specification of the optionality of the move, and a section for adding comments.
iii. Designing an instruction guide for filling in the evaluation template
iv. Sending the request for information to the experts for evaluation purposes
v. Analyzing and filtering the forms filled in

We receive the evaluation template filled in by four experts, and we developed an inter-rater reliability analysis, according to the comments, evaluated parameters, and valuation of each parameter. Such an analysis generates a new version of the model for the subsequent analysis, comprising only the moves considered mandatory by the experts. For this analysis, we consider as inclusion criteria only the moves which were evaluated with 3 or 4 positive responses. Additionally, we carried out the adjustments and changes by move in the cases where the responses disagreed or were unsatisfied. In Table 2 we show the results of the evaluation of our ROM proposal, comprising all of moves with a valuation of 3 or 4—according of the fulfilled evaluation template—an agreement among the experts as a result of the mandatory valuation assigned by them. Finally, the ROM proposal comprises only 15 moves out of 19 from the reference model.

Code	Rhetorical Unit Type	Name of Rhetorical Unit
1	Macromove	Presenting SOP
1.1	Move	Identifying SOP
1.2	Move	Organizing SOP
1.3	Move	Introducing the SOP
1.5	Move	Documenting conventions
1.6	Move	Appointing regulations or regulatory requirements
1.8	Move	Defining intended audience and reading suggestions
1.9	Move	Establishing purpose
2	Macromove	Developing procedures
2.1	Move	Defining procedure purpose
2.2	Move	Defining roles and responsibilities
2.3	Move	Identifying prerequisites
2.4	Move	Listing definitions
2.5	Move	Listing resources
2.6	Move	Establishing methods
2.7	Move	Specifying procedure
3	Macromove	Ending SOP

Code	Rhetorical Unit Type	Name of Rhetorical Unit
3.2	Move	Referencing sources

Table 3: ROM approach

4 Conclusions and future work

We proposed a first approach to a rhetorical organization model in terms of functional and structural features comprised in standard operating procedure (SOP). We followed the method of discourse rhetorical analysis (DRA) and genre analysis for defining such a model.

We stated the discourse analysis approach as efficient as others applied for the analysis of documents, such as data mining, knowledge engineering techniques, and formal methods. This approach is being used as the core of a mapping framework of business-based technical documents which can generate a controlled language output based on the model proposed.

Our proposal lies in the differentiation in the preliminary document analysis by considering each kind of business-based technical document as a genre. Based on such a consideration, we look for the functional, structural, and subsequently linguistic patterns—morphological, lexical, syntactic, and semantic—actually obey and are aligned with the writer communicative purpose. In this way, inside the information and knowledge generated from the document analysis—following our approach for mapping—the original function of the text will remain according to the communicative intention of the author.

We paid special attention to the distinction among the discourse analysis methods and NLP for business-based technical documents. Although scientific communities have approached the processing of such documents usually work in isolation, NLP applications achieve a combined use of techniques from various approaches. Our experience in this field makes us think that it is possible to use one mixed approach including methods and theories of discourse analysis for disciplines such as requirements elicitation, knowledge engineering, and software engineering.

We are developing work oriented to the proposal of heuristic rules for transforming the specific features by move into a controlled language. In addition, a number of open problems must be approached to allow for the

development of a rhetorical organization model. Such problems suggest a variety of research directions as follows:

- Discourse processing in the organizational adaptation of information systems, looking for emphasizing on the role of professional discursive practices in shaping the process of organizational adaptation of information systems.
- Rhetorical characterization of management discourses. By using our proposal of rhetorical analysis for business-based technical documents, we consider a challenging application for processing management of written discourses.

5 Acknowledgments

References

Askehave, I. and Swales, J. M. 2001. Genre identification and communicative purpose: a problem and a possible solution. *Applied Linguistics*, Vol. 22, No. 2, pp. 195-212.

Azaustre, A. and Casas J. 1997. *Manual de retórica española*. Barcelona: Ariel.

Bakhtin, M. M. 1986. The Problem of Speech Genres. *Speech Genres and Other Late Essays, Emerson & Holquist* (eds). University of Texas Press, Austin.

Biber, D. 2006. *University Language: A Corpus-based Study of Spoken and Written Registers*. Amsterdam: John Benjamins.

Biber, D. Conrad, S. and Reppen, R. 1998. *Corpus linguistics: Investigating language structure and use*. University Press, Cambridge.

Burdiles, G. A. 2011. *Descripción de la organización retórica del género caso clínico de la medicina a partir del corpus CCM-2009*. Ph.D. Thesis in Applied Linguistics. Pontificia Universidad Católica de Valparaíso, Chile.

Cabré, M. T. 1999. *La Terminología: Representación y Comunicación. Elementos para una teoría de base comunicativa y otros artículos*. Barcelona: Univ. Pompeu Fabra.

Christie, F. 1999. Genre theory and ESL teaching: A systemic functional perspective. *TESOL Quarterly*, Vol. 33(4), pp.759-763.

Connor, U. 1996. *Contrastive Rhetoric: Cross-Cultural Aspects of Second-Language Writing*. Cambridge University Press, Cambridge.

Eggins, S. 1994. *An introduction to systemic functional linguistics*. Pinter, London.

Freedman, A. and Medway, P. 1994. Locating genre studies: Antecedents and prospects. *Genre and the New Rhetoric,* Freedman & Medway (Eds.), pp. 79-101.

Gee, J. P. 200). Discourse analysis: What makes it critical? In R. Rogers (Ed.), *An introduction to critical discourse analysis in education. Mahwah*, NJ: Laurence Erlbaum, pp. 19-50.

Jameson, D. A. 2008. The Making of Positive Meaning in Thai Annual Reports during the 1997 Economic Crisis. *A Discourse Analysis of the "Message from the Chairman"*.

Karreman, J. and Steehouder, M. 2003. Effects of declarative information in instructions for use. *IEEE International Professional Communication*, 29-33.

Karreman, J. and Steehouder, M. 2003. Effects of declarative information in instructions for use. *IEEE International Professional Communication*, Vol. 1, pp. 29-33.

Kennedy, G. 2000. An introduction to corpus linguistics. *Foreign Language Teaching and Research Press*. NY Press.

Manrique, B., Zapata, C.M., and Burgos, D. 2013. Exploring MWEs for Knowledge Acquisition from Corporate Technical Documents. In *9th Workshop on Multiword Expressions MWE 2013*.

McCarthy, M. and Handford, M. 2004. Invisible to us: A preliminary corpus-based study of spoken business English. *Discourse in the professions*. pp. 167-202. Benjamins, Amsterdam.

Meurer, J. L. 2002. Genre as diversity, and rhetorical mode as unity in language use. *Revista Ilha do Desterro A Journal of English Language*, *Literatures in English and Cultural Studies*, No. 43, Florianópolis, Brasil.

Nickerson, C. 1999. The usefulness of genre theory in the investigation of organizational communication across cultures. *Document Design*, Vol.1, No. 3, pp. 203–215.

Parodi, G. 2005. *Discurso especializado e instituciones formadoras*. Valparaíso: Ediciones Universitarias de Valparaíso.

Parodi, G. 2008. Lingüística de corpus: una introducción al ámbito. *Revista de Lingüística Teórica y Aplicada*, Vol.46 (1), pp. 93-119.

Parodi, G., Ibáñez, R. and Venegas, R. 2010. Discourse genres in PUCV-2006 Corpus of Academic and Professional Spanish: Criteria, definitions and examples. En G. Parodi (Ed.), *Discourse genres in Spanish: Academic and professional connections* (pp.39-68). Amsterdam: Benjamins.

Renkema J. 2003. En torno a la solución del problema de Orwell en la comunicación gubernamental. *Revista Signos*, 36(53), 103-120.

Simpson, R. and Swales, J. 2001. Corpus Linguistics in North America. *Selections from the 1999 Symposium Ann Arbor*: University of Michigan Press.Sinclair, J. 1991. *Corpus, Concordance, Collocation*. Oxford University Press, Oxford.

Swales, J. M. 1990. *Genre Analysis: English in Academic and Research Settings*. University Press, Cambridge.

Swales, J. M. 2004. *Research Genres: Explorations and Applications*. Univ. Press, Cambridge.

Tognini-Bonelli, E. 2001. *Corpus linguistics at work: Studies in corpus linguistics*, v.6. Amsterdam: John Benjamins.

Trosborg, A. (Ed.). 2000. *Analyzing professional genres*. Amsterdam: Benjamins.

van Dijk, T. A. 2008. *Discourse and Context. A Sociocognitive Approach.* Cambridge University Press, Cambridge.

van Nus, M. 1999. Business genres and their corporate con-text. *Journal of Research and Problem Solving in Organizational Communi-cation*: Document Design, 1(3), 187-197.

Warren, M. 2004. So what have you been working on recently? Compiling a specialized corpus of spoken business English. *Discourse in the professions. Perspectives from corpus linguistics*, pp. 115-140. Benjamins, Amsterdam.

Yates, J. 1998. *Control through communication: The rise of system in American management*. Johns Hopkins University Press, Baltimore.