



Acta Scientiarum. Language and Culture  
ISSN: 1983-4675  
ISSN: 1983-4683  
actalan@uem.br  
Universidade Estadual de Maringá  
Brasil

## Applying rhetorical analysis to processing technical documents

---

**Manrique-Losada, Bell; Zapata-Jaramillo, Carlos Mario; Venegas-Velásquez, René**

Applying rhetorical analysis to processing technical documents

Acta Scientiarum. Language and Culture, vol. 41, no. 1, 2019

Universidade Estadual de Maringá, Brasil

**Available in:** <https://www.redalyc.org/articulo.oa?id=307460649013>

**DOI:** <https://doi.org/10.4025/actascilangcult.v41i1.43382>

## Applying rhetorical analysis to processing technical documents

*Bell Manrique-Losada*  
*Universidad de Medellín, Colombia*  
 bmanrique@udem.edu.co

DOI: <https://doi.org/10.4025/actascilangcult.v41i1.43382>  
 Redalyc: <https://www.redalyc.org/articulo.oa?id=307460649013>

*Carlos Mario Zapata-Jaramillo*  
*Universidad Nacional de Colombia, Colombia*

*René Venegas-Velásquez*  
*Pontificia Universidad Católica de Valparaíso, Chile*

Received: 20 June 2018  
 Accepted: 13 February 2019

**ABSTRACT:**

Rhetorical discourse analysis (RDA) emphasizes the communicative purpose and organization structures of texts. This paper proposes a rhetorical organization model for business-based technical documents. Using RDA and genre analysis, we define such a model based on functional and structural features in terms of macro-moves, moves, and functions that comprise a type of document called standard operating procedure (SOP). An SOP is a technical document that describes procedures for segments of business processes or policies to be implemented inside organizations (commonly called a procedure manual). We propose a functional-structural framework for analyzing SOP discourses by influencing and restricting their content based on the style of the 'manual' genre. We identify an SOP as a document with procedural information that serves as a relevant source for extracting domain knowledge and business information. We apply this model as an analytical-conceptual procedure for mapping business documents and generating a controlled language output.

**KEYWORDS:** rhetorical discourse analysis, standard operating procedure, functional and structural model.

**RESUMO:**

A análise do discurso retórica (ADR) enfatiza o propósito comunicativo e as estruturas organizacionais dos textos. Este artigo propõe um modelo de organização retórica para documentos técnicos da área de negócios. Utilizando a ADR e a análise de gêneros, definimos tal modelo com base em traços funcionais e estruturais em termos de movimentos-macro, movimentos e funções que abrangem um tipo de documento chamado de procedimento operacional padrão (POP). Um POP é um documento técnico que descreve procedimentos para segmentos de processos comerciais ou políticas a serem implementadas dentro de organizações (usualmente designadas de manual de procedimento). Propomos um quadro teórico estrutural-funcional para analisar os discursos, influenciando e restringindo seu conteúdo com base no estilo do gênero do 'manual'. Identificamos um POP como um documento com informações de procedimento que serve como um fonte relevante para extrair conhecimento da área e informação comercial. Aplicamos este modelo como um procedimento conceitual-analítico para mapear documentos da área de negócios e gerar como resultado uma linguagem controlada.

**PALAVRAS-CHAVE:** análise do discurso retórica, procedimento operacional padrão, modelo estrutural e funcional.

**INTRODUCTION**

Discourse analysis is concerned with the relationship between the forms and functions of discourse, and it is a vast field encompassing various interpretive and explanatory sub-disciplines. In discourse analysis, creating a specialized text involves a distinct process whereby the text is generated from a specialized organizational discourse and produced by specialists who have mastered the cognitive and conceptual organization of the subject matter (Biber, 2006; Nickerson, 1999).

We approach the analysis of such specialized discourses from the genre perspective (Nickerson, 1999). Our approach for discourse analysis is based on the rhetorical analysis techniques proposed by Van Nus (1999) and Swales (1990; 2004).

Rhetorical analysis is concerned with discourse construction such that it emphasizes the communicative purpose of texts. We applied rhetorical discourse analysis (RDA) and genre analysis to define a preliminary approach in developing a rhetorical organization model with functional and structural features that comprise a business-based technical document called 'standard operating procedure' (SOP).

Based on the structural move analysis and steps developed by Swales (1990), we define macro-moves and moves as structural aspects of SOPs that describe features of text, concerning global organization or sentence-level features respectively.

The functional features of SOP encompass the rhetorical purposes that express communicative intentions of the documents authors. The structure and functions, therefore, influence and restrict the content and style of a text and enable the identification of linguistic features (Askehave & Swales, 2001). The proposed model is currently being used as the core of a mapping framework for business-based technical documents that can generate a controlled language output.

The remaining sections of this paper are organized as follows. Section 2 describes the theoretical framework and some related work in the field of discourse and rhetorical analysis, and Section 3 presents our proposed rhetorical organization model. Finally, Section 4 presents our conclusions and outlines prospects for future work.

## THEORETICAL FRAMEWORK AND RELATED WORK

This study focuses on the following perspectives of discourse analysis through the use of language for constructing, interpreting, and exploiting technical documents.

### Discourse analysis.

Discourse analysis is one of the several fields devoted to the study of the social uses of language, particularly the linguistic relationship between forms and functions (Gee, 2000). Based on a systematic methodology, discourse analysis is interpretive and explanatory. The study of discourse is vast and encompasses several overlapping sub-disciplines, such as sociolinguistics and semiotics, genre studies, and specialized discourse analysis.

Specialized text analysis is a special field of discourse analysis (Nickerson, 1999; Biber, 2006). One approach to discourse analysis in specialized texts is through genre analysis. According to Yates (1998) and Swales (1990; 2004) genres are defined as variations of a language that operate throughout the linguistic features present in a text. Genres are linguistically linked to specialized communicative purposes, participants, production/usage contexts, and modes of discourse organization among other elements (Parodi, Ibañez, & Venegas, 2010).

Genre theory (Nickerson, 1999; Van Nus, 1999) focuses on the written practices of members in particular communities and/or the design of information and business records. Specialized texts are generated from a specialized organizational discourse and produced by subject-area experts who have mastered the conceptual and cognitive organization of a topic. According to Cabré (1999), specialized discourse derives from variables related to the subject and perspective of a topic and the producer's intent and level of expertise. Following such a Genre theory, we can analyze specialized discourses because members of specialized communities are writing them.

The text structure and its context describe a genre (Van Dijk, 2008). Genre analysis is an effort to relate the text structure to the macro-social context. It is often characterized by identifiable purposes and schematic structures, which can be as numerous as the social practices people are involved in. Genre studies have been developed from Bakhtin's (1986) seminal work, the socio-rhetorical theoretical framework by Swales (1981; 1990), and the approaches for systemic functional linguistics proposed by Christie (1999) and Eggins (1994). Based on these foundations and the approach developed by Meurer (2002), we applied genre analysis based on the characterization of genres as reasonably stable types of text (formal or informal), which can be recognized based on their rhetorical structure and function.

## Rhetorical analysis

Rhetorical analysis is concerned with the construction of discourse, giving priority to the communicative purpose of each genre (Azaustre & Casas, 1997). Rhetoric aims at discourse from its intentional (purpose-driven) and instrumental (means of fulfilling the purpose) perspectives. Thus, rhetorical discourse organization is an approach where textual structures are employed to achieve a desired effect (Connor, 1996). These structures provide a framework for articulating diverse discourses in a particular manner and textually constituting their relations.

Based on the framework of rhetorical analysis, genre analysis involves the analysis and description of a text in terms of rhetorical moves or rhetorical structures, which denote the functional parts or sections of a genre. The particular conformation of the text surface is defined by text organization levels, which is known as rhetorical discourse organization. In such a way, the structural units identified by genre analysis can be characterized as moves subdivided into steps, referring to those passages of the text that are larger than the largest grammatical units, e.g., clauses, and sentences, and possess some unity grounded in a common function and meaning.

## Corpus linguistics

Corpus linguistics analyzes the linguistic properties of an extended passage, text, or corpus of texts, and includes genre analysis and semiotics perspectives in the computational analysis of text corpora.

According to Parodi (2008), corpus linguistics encompasses a set of methodological principles for studying any language domain. Collections of linguistic features, operated by genres, can be identified from a representative corpus.

Corpus-based approaches have been widely applied in a range of fields such as discourse analysis, language teaching, and stylistic analysis (Kennedy, 2000). Biber, Conrad, and Reppen (1998) summarized features of corpus-based studies as empirical approaches involving a natural texts sample, computer-assisted analysis, and a combination of quantitative and qualitative analysis.

Working with a specialized language corpus involves a strict selection of texts for identifying common patterns (Sinclair, 1991). Certain words and phrases may be rare within a general sample of texts, but also appear very frequently in certain specific texts in a strict selection. Thus, a corpus should be representative of one or more aspects of a language.

## Proposal of rhetorical organization model

Based on the above described rhetorical analysis approaches, we use structural and functional features to propose a preliminary rhetorical organization model (ROM). ROM has been defined by employing methodological procedures, which are divided into the following phases:

## Corpus definition

Defining the corpus begins with the collection of possible technical documents on the genre that are circulating online and contain access to full text. We broadly explored four types of technical documents, namely job description documents, functions manuals, corporate policy documents, and SOPs. We then selected the corpus for ‘SOP’ as a referent for this study, after analyzing the collected documents. The population selection criteria for the corpus included the following: ‘SOP’ or ‘SOP manuals’ that are written in English, published online and have open-access on the Internet, having an author affiliated to a company and/or organization, and are text-based with a low percentage of images.

We analyzed the documents following a corpus linguistics methodology (Simpson & Swales, 2001; Tognini-Bonelli, 2001) according to the approach developed by Parodi (2005). We performed a descriptive linguistic analysis of business-based technical documents that centered on procedures, and subsequently developed the following processes:

- (i) searching, reviewing, and analyzing of documents available on the Internet;
- (ii) defining a hierarchy of the macro-genre, based on genre analysis theory; and
- (iii) developing a genre-based qualitative characterization, from the genre perspective, for the ‘manual’ macro-genre and the ‘procedures manual’ genre (Swales, 2004).

The term ‘manual’ can be applied to academic manuals, instructional and teaching textbooks, and technical procedure manuals (Parodi, 2008). In this study, we use procedures manual to refer to the latter category as a basis for identifying several genre sub-types, as presented in Figure 1 and Figure 2.

Explicitly, we define SOP as a linguistic genre such that an SOP is a constitutive document of a quality system describing a set of recurring operations. An SOP describes procedures for a segment of business processes and the effective implementation of a set of policies.

A manual is a set of written instructions describing how procedures are defined, developed, and managed by an organization’s members. SOPs belong to the ‘manual’ genre along with more typical manuals such as procedure, quality, and user manuals. We used SOP as the main document type for our research because organizations use such technical documents as a means of specifying their technical, administrative, and operational activities. We identified an SOP as a type of procedural document that presents the most important or fundamental procedural information (Karreman & Steehouder, 2003), as a source for extracting domain knowledge and business information.



FIGURE 1  
Hierarchy of the ‘manual’ macro-genre.

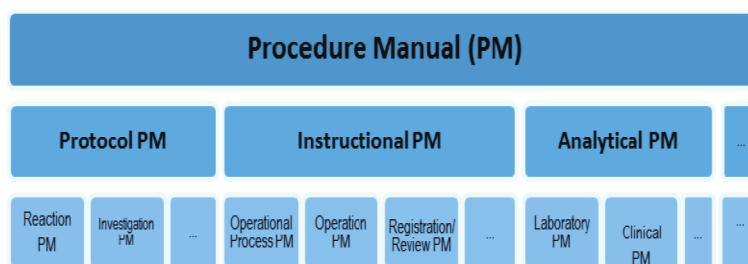


FIGURE 2  
Classification of the 'procedure manual' genre.

After establishing the criteria for Internet search and corpus formation, we implemented a sampling procedure. The documents collected for the corpus correspond to 100% of the population; thus, forming a significant sample from a proportional estimated sample size with a confidence level of 95% precision and an approximate value of 5% of the measured parameter. Assuming that the population was evenly distributed, we selected a sample of 32 documents, corresponding to 64% of the total population, which was the minimum statistically randomized percentage calculated with the Z-test of proportions.

## Digital corpus analysis

The corpus analysis involved tokenization, keyword and stop-word identification, characterization, and the creation of word frequency lists among other tasks. In addition, a concordance analysis was conducted by searching for identified lexical forms and every occurrence of a given word, together with its context. Thus, we generated an occurrences list of the given search terms in the corpus within the context where they occur. The corpus analysis considered 9252 word types and 167,905 tokens.

## ROM design

The first step in discourse mapping entailed performing a search for theoretical models in the relevant literature, to be used as a reference for the rhetorical analysis of the SOP genre.

Some antecedents in the administrative field include previous attempts to provide guidelines on how to create well-written SOPs, including how to organize or format such documents (Stup, 2001; Grusenmeyer, 2003; North Carolina State University [NCSU], 2014).

Such approaches are concerned with understandable instructions that will enable SOP writers to write in a clear and simple manner. Additionally, authors such as Wieringa, Moore, & Barnes (1998) and Price (2001) proposed including a discussion of procedure, grammar, and writing along with principles and practices.

We identified approaches concerned with documents related to government and business (Trosborg, 2000; Renkema, 2003; McCarthy & Handford, 2004; Warren, 2004), as well as commercial documents (Freedman & Medway, 1994; Yates, 1998; Jameson, 2008); yet we found no references to genres related to SOPs. For this reason, we used an inductive method to define the preliminary model, according to Burdiles (2016). Such a method includes the following steps:

- (i) Random selection of four sample documents from the corpus
- (ii) Incremental construction of a preliminary model based on manual review of the structure and superstructure to identify the common organization units in the sample (moves)
- (iii) Definition of rhetorical moves as the functional sections of the genre (Swales, 1990; 2004), whereby we adopted the macro-move concept, which entails a higher abstraction of rhetorical purpose (Parodi, 2008).



Thus, each macro-move serves a communicative purpose and all macro-moves shape the overall organization of the text, and

(iv) purpose identification of a higher level hierarchy of macro-purposes that comprise a set of more specific moves and detailed steps.

We defined our preliminary model as a set of functional and structural features that resulted from the identification of highly recurrent moves (the mandatory level). The mandatory moves were selected based on a set of defined categories for evaluating every move in each document from the sub-corpus. Such categories were related to the percentage of move occurrence as follows: 0% (does not appear in the document), 1–30% (low chance), 31–70% (average chance), and 71–100% (mandatory). In the reference model, we consider as moves the placed in the high percentages. As shown in Figure 3, the resulting model comprised three macro-moves, containing 19 moves—showing more specific functional units, which are described as follows:

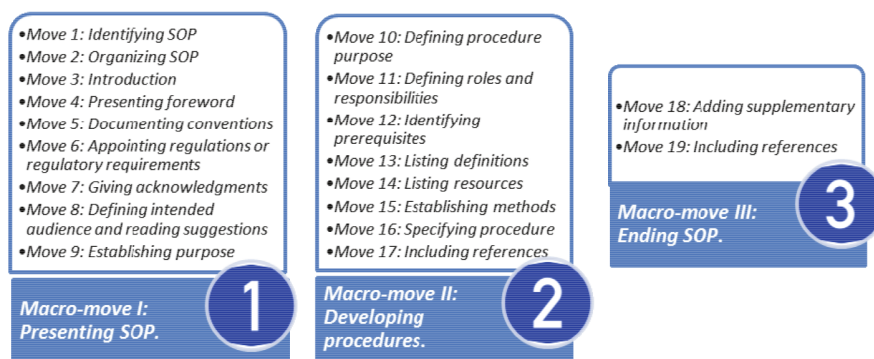


FIGURE 3.  
Preliminary ROM.

Macro-move I: Presenting the SOP. This macro-move entails the presentation of a preliminary statement that introduces the document and defines its purpose, conventions, revision schedule, approval authority, and organization, among other elements.

- Move 1: Identifying the SOP. This move includes identifying the organization authoring the SOP, including the author(s), company, location, filiation, name, and verbal/nonverbal identification.

- Move 2: Organizing the SOP. This move includes elements of the document's body related to content organization and lists of tables and figures among others. This move allows the reader to locate the document content and presents the entire hierarchical organization of the document.

- Move 3: Introduction. Justifies and presents the document. This move provides a general description of the document's context and establishes its purpose.

- Move 4: Presenting foreword. Presents a general overview of the document and describes what is included in each procedure. Additionally, it can describe who participated in writing the SOP, how it was organized, how to read it, the review process that was undertaken, and warnings about its use/distribution.

- Move 5: Documenting conventions. Establishes the document's context, namely the date of its approval, version number, author, and revision number.

- Move 6: Appointing regulations or regulatory requirements. Reviews the standards, contractual requirements, and policies or regulations associated with the procedures in the SOP.

- Move 7: Giving acknowledgments. Presents the compendium of writers and others involved in authoring the SOP and acknowledges their contributions.

- Move 8: Defining intended audience and reading suggestions. Defines the primary audience for the SOP, which can include management teams, operational teams, and organization staff.

- Move 9: Establishing purpose. Describes the general goal of the procedures within the organizational framework. This goal is oriented toward contextualization and purpose description.

Macro-move II: Developing procedures. This macro-move entails providing a detailed description of the procedures associated with each organizational process, thus defining a series of specific purposes, responsibilities, functions, procedural description, and rules for implementation.

- Move 10: Defining procedure purpose. Defines the purpose of each procedure.
- Move 11: Defining roles and responsibilities. Defines roles and responsibilities involved in each procedure.
- Move 12: Identifying prerequisites. Identifies the prerequisite steps in the procedure's execution, and may include rules, cautions, warnings, and recommendations for achieving them.
- Move 13: Listing definitions. Includes a list of definitions, concepts, and terms or acronyms used in the context of SOP.
- Move 14: Listing resources. Lists the equipment, resources, and materials required for the procedure's execution.
- Move 15: Establishing methods. Establishes the methods used to characterize or guide the procedure.
- Move 16: Specifying procedure. Provides step-by-step instructions to elucidate the details of procedures.
- Move 17: Including references. Lists bibliographical references that support the procedures.

Macro-move III: Ending the SOP. This is an optional macro-move related to macro-moves I and II.

- Move 18: Adding supplementary information. This includes attachments supporting the development of the macro-moves.
- Move 19: Including references. Lists a set of bibliographical references.

Macro-move II, 'Developing procedures', forms the backbone of the SOP genre because it serves as a unit that can be repeated a number of times to cover the details of all of the necessary procedures development. In contrast, macro-move I only occurs once throughout the text. Macro-move III is optional, but often complements macro-move II depending on the information needed to support the procedures.

Following the construction of the preliminary model, we solicited detailed and systematic peer review analyses and generated a reference model based on the feedback received from the experts.

## Rhetorical analysis

The rhetorical analysis was based on the reference model and comprised a set of manual activities as well as those supported by computational tools like AntConc 3.3.5w<sup>[1]</sup>, TermoStatWeb<sup>[2]</sup>, ElGrial<sup>[3]</sup>, and NLTK-Demo<sup>[4]</sup>. The activities proceeded as follows:

i. Analysis and identification of rhetorical units. Rhetorical units refer to the organization units as macro-moves, moves and steps. This activity focused on the identification and registration of the identifier, summary, SOP category, macro-move identifier and name, move identifier and name, identified linguistic features, and an example. Additionally, we developed a process for identifying lexical and grammatical tracks from each rhetorical unit.

ii. Definition of mandatory rhetorical units. Each rhetorical unit from the SOP was considered optional or mandatory based on its relationship with the purpose of the macro-structure and the author's communicative intention, whereby a recurrent presence indicated the mandatory level of the move.

iii. Corpus preprocessing. The collected texts were converted into a .txt format through certain transformation actions, and non-relevant information for text processing such as HTML tags, webpage names, and ads, was manually removed using tools like HTML Text and Multireplacer3.

iv. Identification of linguistic features by rhetorical unit. We identified several variables, including word frequency that was based on mutual information, parameters of probable co-occurrence among words, recurrent lexical items in given stretches of text, keywords, and word frequency among others. We further used computational tools to identify prototypical lexical-grammatical features from the corpus sample



(Table 1). According to Biber, Conrad, and Cortes (2004) and Venegas (2010), some of these features are essential to current writing in specialized organizational fields.

v. Morpho-syntactic tagging. This analysis aimed to identify the morpho-syntactic categories and features that surrounded each rhetorical unit. We described the features for such categories based on feature object, i.e., noun, adjective, syntactic clause, verb, subordination, coordination, verbal mood, verbal periphrasis, and person. The feature description is presented with the included move in Table 1 and functional features in Table 2.

TABLE 1.  
Prototypical lexical- grammatical features.

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
		Most common are uncountable nouns (e.g., information, process, data)	12, 13
		The most common demonstratives: this, that	1, 2, 10
		Scarce usage of possessives like <i>its</i> or <i>his/her</i>	-
		Undefined. They are rarely used (e.g., any)	12, 13
	Adjective	The most common are: required, appropriate, successful, complete. Their use is mainly subjective rather than descriptive	11, 12, 13
		The copulative verbs are the most used mode in SCP. Some examples include <i>be, is, are</i>	5, 11, 12, 13
		Can be conjugated in specific mode, commonly to describe steps (e.g., may)	4, 13
		Mainly conjugated in indicative mode (e.g., use, review, report, describe)	5, 10, 11, 13
		Common place and tense subordinate (using <i>where, when, after...</i> )	11, 13
Syntactical	Subordination	Adjectives or noun subordinate clauses are used, usually begins with a relative pronoun (e.g., <i>that, how</i> ) or using <i>as, before, or until</i>	11, 12, 13
		The most frequent are copulative coordinate clauses to express adding one statement to another. They are introduced by conjunctions (e.g., <i>and, both, as well as</i> )	4, 5, 11, 12, 13
		The only explanatory coordinate clause found was that is	4, 12, 13
		Usage of nominal phrases (adjective + noun) is mainly common in the rules and responsibilities section and the SCP steps (e.g., appropriate project)	11, 13
		Main moods: indicative and imperative, used to express capacity or obligation	All of the moves
	Verbal mood	Mood in conditional steps denoting possibility/priority (e.g., <i>is possible, is important, etc.</i> )	4, 12, 13
		Modal periphrasis, indicating obligation (e.g., <i>may + verb, must + verb</i> ) and its negative form or indicating potential mood (e.g., <i>can + verb</i> )	4, 12, 13
		Aspectual periphrasis occurs as follows: initiation (e.g., <i>take</i> ), progression (including performing), habits (used to, used for), duration ( <i>is/are + gerund</i> ), completion ( <i>end of + infinitive</i> )	13
		Prevalence of the first person placed in the sequence of procedures. The third person singular is used in the "rules and responsibilities" section. Low usage of pronouns has been detected	All of the moves
		Verbal periphrasis	

**Notes:** English is mostly/usually expressed by the present or the past and for the subjunctive mood has different usages and frequencies. \*See the complete proposal in Manrique et al. (2013).

vi. Identification of occurrence frequencies for relevant features. We used computational tools related to corpus linguistics to analyze relevant features and identify their frequency of occurrence.

TABLE 2  
Prototypical functional and lexical features

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

## Evaluation of structural and functional features

We evaluated the structural and functional features in order to determine the rhetorical moves, which are fundamental for SOP writing. We followed the evaluation activities as described below:

- Selection of experts to evaluate the reference model (comprising 19 moves as presented in Figure 3);
- Designing a template for evaluation based on the reference model and considering the following factors: rhetorical unit of reference, example extracted from an SOP, evaluation, specification of the move's optionality, and a section to add comments;
- Designing an instruction guide to fill the evaluation template;
- Sending the request to experts for evaluation purposes;
- Analyzing and filtering the completed forms.

We received the completed evaluation templates from the experts and developed an inter-rater reliability analysis according to the comments, evaluated parameters, and valuation of each parameter. For the subsequent analysis, we generated a new version of the model that comprised only the moves considered mandatory by the experts. For this analysis, the inclusion criteria consisted only of the moves that were evaluated with three or four positive responses and adjustments and/or changes of moves were made in cases where the responses disagreed or were unsatisfactory Table 3 shows the evaluation results of our ROM proposal. The final ROM proposal includes only 15 of the original 19 moves proposed in the reference model.

TABLE 3  
ROM approach.

Code	Rhetorical Unit Type	Name of Rhetorical Unit
1	Macro-move	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	Macro-move	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	Macro-move	Ending SOP
3.2	Move	Referencing sources

## CONCLUSIONS AND FUTURE WORK

This study proposed a first approach to a ROM in terms of functional and structural features comprised in SOP. In defining such a model, we used the methods of RDA and genre analysis.

We identified the discourse analysis approach that is equally efficient as other techniques applied for the analysis of documents, such as data mining, knowledge engineering methods, and formal methods. Discourse analysis 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 proposed model.

Our proposal has a distinct method of preliminary document analysis that considers each kind of business-based technical document as a genre. Based on this consideration, we aim to identify the functional, structural, and subsequently linguistic patterns, which are morphological, lexical, syntactic, and semantic, that is closely aligned with the writer's communicative purpose. Our approach toward mapping the information and knowledge generated from the document analysis state that the original function of the text will reflect the communicative intention of the author.

We paid particular attention to distinctions among discourse analysis methods and natural language processing techniques (NLP) for business-based technical documents. Although scientific communities, which have approached the processing of such documents, generally work in isolation, NLP applications combine techniques from various approaches. Our experience in this field enables us to believe it is possible to use one mixed approach that incorporates discourse analysis methods and theories with disciplines such as requirements elicitation, knowledge engineering, and software engineering.

We are currently in the process of developing proposals of heuristic rules for transforming the specific features of moves into a controlled language. In addition, a number of other problems must be addressed to further the development of a ROM. Such problems suggest a variety of research directions as follows:

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

## ACKNOWLEDGEMENTS

This work has been partly funded by the Vicerrectoría de Investigación from both the Universidad de Medellín and the Universidad Nacional de Colombia, under the project “*Especificación de un lenguaje controlado de dominio específico: fundamentos lingüísticos y bases de transformación desde documentos técnicos corporativos en lenguaje natural*”. It was further supported in part by the COLCIENCIAS, under the doctoral program “*Generación del bicentenario, 2010*.”

## REFERENCES

- Askehave, I., & Swales, J. M. (2001). Genre identification and communicative purpose: a problem and a possible solution. *Applied Linguistics*, 22(2), 195-212.
- Azaustre, A., & Casas J. (1997). *Manual de retórica española*. Barcelona, ES: Ariel.
- Bakhtin, M. (1986). The problem of speech genres. In M. Bakhtin, *Speech genres and other late essays*, (p. 209-218). Austin, TX: University Press.
- Biber, D. (2006). *University language: a corpus-based study of spoken and written registers*. Amsterdam, ND: John Benjamins.
- Biber, D., Conrad, S., & Cortes, V. (2004). If you look at ...: lexical bundles in university teaching and textbooks. *Applied Linguistics*, 25(3), 371-405.
- Biber, D., Conrad, S., & Reppen, R. (1998). *Corpus linguistics: investigating language structure and use*. Cambridge, MA: University Press.
- Burdiles, G. A. (2016). Género Caso Clínico: Organización retórica de su macromovida Relato del Caso en publicaciones médicas chilenas. *Revista Signos*, 49(91), 192-216.
- Cabré, M. T. (1999). La terminología: representación y comunicación. Elementos para una teoría de base comunicativa y otros artículos. Barcelona, ES: Univ. Pompeu Fabra.
- Christie, F. (1999). Genre theory and ESL teaching: A systemic functional perspective. *TESOL Quarterly*, 33(4), 759-763.
- Connor, U. (1996). *Contrastive rhetoric: cross-cultural aspects of second-language writing*. Cambridge, MA: Cambridge University Press.
- Eggins, S. (1994). *An introduction to systemic functional linguistics*. London, UK: Pinter.
- Freedman, A., & Medway, P. (1994). Locating genre studies: Antecedents and prospects. In A. Freedman, & P. Medway (Ed.), *Genre and the new rhetoric*, (p. 79-101). Londres: UK: Wiley.
- Gee, J. P. (2000). Discourse analysis: what makes it critical? In R. Rogers (Ed.), *An introduction to critical discourse analysis in education* (p. 19-50). Mahwah, NJ: Laurence Erlbaum.
- Grusenmeyer, D. (2003). Developing effective standard operating procedures. *Dairy Business Management*, 1(1), 1-7.
- Jameson, D. A. (2008). The making of positive meaning. In *Proceedings Thai annual reports during the 1997 economic crisis* (p. 80-93). Pattaya, TH.
- Karreman, J., & Steehouder, M. (2003). Effects of declarative information in instructions for use. *IEEE International Professional Communication*, 1(1), 29-33.

- Kennedy, G. (2000). *An introduction to corpus linguistics*. New York, US: NY Press.
- Manrique, B., Zapata, C. M., & Burgos, D. (2013). Exploring MWEs for knowledge acquisition from corporate technical documents. In *Proceedings of the 9th Workshop on Multiword Expressions MWE 2013* (p. 82-86). Atlanta, GE.
- McCarthy, M., & Handford, M. (2004). Invisible to us: a preliminary corpus-based study of spoken business English. In U. Connor, & T. A. Upton (Ed.), *Discourse in the professions* (p. 167-202). Amsterdam, ND: Benjamins.
- 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*, 1(43), 61-82.
- Nickerson, C. (1999). The usefulness of genre theory in the investigation of organizational communication across cultures. *Document Design*, 1(3), 203-215.
- North Carolina State University [NCSU]. (2014). How to write standard operating procedure (SOP). Retrieved from: <https://www.ces.ncsu.edu/wp-content/uploads/2014/03/how-to-write-an-SOP.pdf>
- Parodi, G. (2005). *Discurso especializado e instituciones formadoras*. Valparaíso, CH: Ediciones Universitarias de Valparaíso.
- Parodi, G. (2008). Lingüística de corpus: una introducción al ámbito. *Revista de Lingüística Teórica y Aplicada*, 46(1), 93-119.
- Parodi, G., Ibáñez, R., & Venegas, R. (2010). Discourse genres in PUCV-2006 corpus of academic and professional spanish: criteria, definitions and examples. In G. Parodi (Ed.), *Discourse genres in spanish: academic and professional connections* (p. 39-68). Amsterdam, ND: Benjamins.
- Price, B. (2001). Set monitoring protocols for SOPs. *Dairy Herd Management*, 38(3), 1-8.
- Renkema, J. (2003). En torno a la solución del problema de Orwell en la comunicación gubernamental. *Revista Signos*, 36(53), 103-120.
- Simpson, R., & Swales, J. (2001). *Corpus linguistics in North America. Selections from the 1999 symposium Ann Arbor*. Arbor, MI: University of Michigan Press.
- Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford, UK: Oxford University Press.
- Stup, R. (2001). Standard operating procedures: a writing guide. Retrieved from <http://extension.psu.edu/animals/dairy/hr/tools-for-dairy-employee-supervisors/ud011>
- Swales, J. M. (1981). *Aspects of article introductions*. Birmingham, UK: University of Aston.
- Swales, J. M. (1990). *Genre analysis: english in academic and research settings*. Cambridge, MA: University Press.
- Swales, J. M. (2004). *Research genres: explorations and applications*. Cambridge, MA: University Press.
- Tognini-Bonelli, E. (2001). *Corpus linguistics at work: studies in corpus linguistics* (Vol. 6). Amsterdam, ND: John Benjamins.
- Trosborg, A. (2000). *Analyzing professional genres*. Amsterdam, ND: Benjamins.
- Van Dijk, T. A. (2008). *Discourse and context. A sociocognitive approach*. Cambridge, MA: Cambridge University Press.
- Van Nus, M. (1999). Business genres and their corporate con-text. *Journal of Research and Problem Solving in Organizational Communication: Document Design*, 1(3), 187-197.
- Venegas, R. (2010). Multi-dimensional analysis of the PUCV-2006 corpus of Academic Spanish. In G. Parodi (Ed.), *Discourse genres in Spanish: academic and professional connections* (p. 112-134). Amsterdam, ND: Benjamins.
- Warren, M. (2004). So, what have you been working on recently? Compiling a specialized corpus of spoken business English. In U. Connor, & T. A. Upton (Ed.), *Discourse in the professions. Perspectives from corpus linguistics* (p. 115-140). Amsterdam, ND: Benjamins.
- Wieringa, D., Moore, C., & Barnes V. (1998). *Procedure writing: principles and practices*. Columbus, Ohio: Battelle Press.
- Yates, J. (1998). *Control through communication: the rise of system in American management*. Baltimore, MA: Johns Hopkins University Press.



## NOTES

- [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 at: [www.elgrial.cl](http://www.elgrial.cl)
- [4] NLTK-Demos. A tool containing different word tokenizers, stemming text, part of speech tagging, and chunk extraction modules available at: <http://text-processing.com/demo/>