

**AN SFL- AND GENRE ANALYSIS-BASED ANALYSIS OF BRAZILIAN  
EXPERT RESEARCHERS' ARTICLE INTRODUCTIONS WRITTEN IN  
ENGLISH**

Igor Antônio Lourenço da SILVA<sup>i</sup>  
Kelen C. Sant'Anna de LIMA<sup>ii</sup>  
Adriana Silvina PAGANO<sup>iii</sup>  
(UFMG)

**ABSTRACT:** *This paper reports on an ongoing research project – CORDIALL (Corpus of Discourse for the Analysis of Language and Literature), developed at Faculdade de Letras, Universidade Federal de Minas Gerais, Brazil –, which adopts a systemic-functional approach to translation studies. It describes research articles introductions written in English by Brazilian expert researchers on genetic disorders, and contrasts them with other textual productions of the same register, genre and language. It aims to map some discursive features of these experts' textual production, and to verify whether these individuals' generic competence allows them to produce texts that fulfill some basic features of RA introductions in a foreign language textual production.*

**KEY-WORDS:** *research article, thematic organization, transitivity, expert researchers.*

**RESUMO:** *Este trabalho está vinculado ao projeto CORDIALL (Corpus Discursivo para Análises Lingüísticas e Literárias), desenvolvido na Faculdade de Letras da UFMG e que adota uma abordagem sistêmico-funcional dentro dos estudos da tradução. Descrevem-se introduções de artigos acadêmicos que foram produzidos em inglês por pesquisadores brasileiros, considerados expertos em doenças genéticas, e contrastam-se essas introduções com outras pertencentes ao mesmo registro, gênero e língua. O objetivo consiste em mapear alguns aspectos discursivos da produção textual desses pesquisadores e verificar se a competência genérica desses indivíduos lhes facultam produzir textos que preenchem alguns requisitos básicos das introduções de artigos acadêmicos mesmo quando da produção textual em uma língua estrangeira.*

**PALAVRAS-CHAVE:** *artigo acadêmico, organização temática, transitividade, pesquisadores expertos.*

## 1. Introduction

This paper reports on an ongoing research project – CORDIAL (Corpus of Discourse for the Analysis of Language and Literature), developed at Faculdade de Letras, Universidade Federal de Minas Gerais, Brazil –, which adopts a systemic-functional approach to translation studies. Research article (RA) introductions which were produced in English (as a foreign language) by Brazilian expert researchers on genetic disorders are described and contrasted with other textual productions of the same register and genre written in English by international researchers. By doing so, we aim to map some discursive features of these professionals' textual production, and verify whether these individuals' generic competence allows them to produce texts that fulfill some basic features of RA articles.

Building on Bhatia's (2004) concept of generic competence – the ability to respond to recurrent and novel rhetoric situations, exploring generic conventions of specific cultures and disciplinary practices –, we hypothesize that expert researchers' generic competence allows for the production of texts that fulfill basic features of RA articles, even when these text productions are carried out in a foreign language, since the texts analyzed have already been accepted by and published in internationally indexed journals. In order to verify such a hypothesis, we draw on Swales's genre analysis, more specifically, his concepts of moves, steps, and categories for lexical and grammatical analysis such as reference and tense combination) along with a Systemic Functional Linguistics approach (SFL, henceforth), focusing on interpersonal meanings, thematic organization, transitivity and cohesion. The interface between these two theoretical grounds has not been much developed, but we assume that a multiple perspective on the same object may allow for a better understanding of expert researchers's textualization of lexicogrammar, as well as their organization of discourse, two fundamental stages for the mapping of a genre.

## 2. Literature

In spite of the fact that the concept of genre is not a consensus, particularly in literary studies, linguistics, and in rhetoric, this paper considers the definition presented by Swales (1996), for whom a genre comprises five characteristics. First, a genre is a class of communicative events. Second, the principal criterion feature that turns a collection of communicative events into a genre is some shared set of communicative purposes. Third, exemplars or instances of genres vary in their

prototypicality. Fourth, the rationale behind a genre establishes constraints on allowable contributions in terms of their content, positioning and form. And fifth, a discourse community's nomenclature for genres is an important source of insight. Furthermore, Bhatia (2004) may be insightful for the definition of genre within the scope of this paper. For the author, the most important feature of genre is that it is identifiable, sufficiently standardized, and contextualized in a social and institutional framework; it is based on a set of accessible conventions shared by most of the members of a professional, academic or institutional organization.

There is no doubt that English has, according to Swales (1996: 99), become the world's predominant language of research and scholarship. Abstracts and the very research articles are supposed to be written in English so as to be published in internationally indexed journals, which urges researchers to publish in English, either texts they translate themselves, texts they have someone to translate, or text they produce originally in English. However the English text is written, the text must fulfill some patterns and regularities of lexico-grammar and organization which are considered to be particular of that specific genre in order to be accepted. In other words, researchers are required to master both their lexical-grammatical choices and their moves and affiliations throughout the text in order to *negotiate* meaning, and their insertion in a given knowledge domain and in the context of a journal. According to Bhatia (2004), such a requirement takes place because genres are reflections of disciplinary practices, and the acquisition of generic competence implicates the acquisition of expertise in the activities that produce or consume knowledge of the disciplinary, working and professional cultures.

Still according to Bhatia (2004), studies of genre can be mapped onto three stages of development, namely textualization of lexico-grammar, organization of discourse and contextualization of discourse. The first refers to the investigations of microstructures of a text and their functional values within the framework of a set of restricted aspects of a given discourse. The second refers to the engagement with the structuring of discourse and focuses on patterns or regularities of organization in discourse. The third stage seeks to investigate disciplinary and institutional contexts, as well as professional and social contexts. While the first two stages can be carried through by investigating elements within the text, the contextualization of discourse involves investigating elements beyond the text, and thus only a brief

look at this phase will be provided here, as this is outside the scope of our study.

Organization of discourse is mapped in this paper through the analysis of some elements of RA introductions described by Swales (1990/6), that is, their moves and steps. On the other hand, textualization of lexicogrammar is mapped by observing both the reference and tense combination presented by Swales (1990/6) along with text analysis based on Systemic Functional Linguistics (SFL) categories, namely interpersonal meanings, thematic organization, transitivity and cohesion. Even though Bhatia does not point out SFL as a ground for the analysis of textualization, this theory is used for two reasons drawing on Pagano & Vasconcellos' account of the potential of SFL (2005): (i) the text is regarded as a social phenomenon and this is determined by other social phenomena (situation context), and (ii) the functional components of the linguistic system semantic stratum (ideational, interpersonal and textual metafunctions) establish organized nets of meaning potentials from which the language user makes his/her *choices*.

## 2.1. Organization of discourse

Along the production of a RA comes the necessity of getting started, that is, of writing the introduction. This, according to Swales (1996) is considered troublesome, since writing the opening paragraphs implicates several decisions (e.g., about including and excluding background knowledge, about the directness of the approach). Bearing in mind the organization of discourse, the author suggests a model for RA introductions, named CARS (Create a Research Space), in which he provides steps for the production of an article introduction. The model is represented in Figure 1 and it will be applied on the selected articles for their analysis.

**Move 1 Establishing a territory**

Step 1 Claiming centrality

and/or

Step 2 Making topic generalization(s)

and/or

Step 3 Reviewing items of previous research

**Move 2 Establishing a niche**

Step 1A Counter-claiming

or

Step 1B Indicating a gap

or

Step 1C Question-raising

or

Step 1D Continuing a tradition

**Move 3 Occupying the niche**

Step 1A Outlining purposes

or

Step 1B Announcing present research

Step 2 Announcing principal findings

Step 3 Indicating research article structure

Figure 1- CARS model for article introductions, based on Swales (1996: 141).

Source: SWALES, J.M. 1996. *Genre Analysis: English in academic and research settings*.

## 2.2. Textualization of lexico-grammar

### 2.2.1. Reference and tense combination

According to Swales (1996), there is a certain prototypicality of reference (or intertextuality for Bhatia, 2004) and tense usage for RA in each different disciplinary field. For the author, reference may be (1) integral, when the author being cited is the subject or agent of the cited sentence, or (2) non-integral, when the author is present but not as subject or agent of the cited sentence; and reference may also be done either (a) in a reporting structure, when there is a reporting verb, or (b) in a non-reporting structure, when there is no such a verb. The combination of this 1-2 and a-b tends to determine the tense employed, as shown in Figure 2.

	<b>Integral (1)</b>	<b>Non-integral (2)</b>
<b>Reporting (a)</b>	<i>Past</i> Brie (1988) showed that...	<i>Present Perfect</i> It has been shown that... (Brie, 1988)
<b>Non-reporting (b)</b>		<i>Present (or modal)</i> The moon may be made of cheese (Brie, 1988)

Figure 2 – Reference and tense, based on Swales (1996: 153).

Source: SWALES, J.M. 1996. *Genre Analysis: English in academic and research settings*.

Although figure 2 is a “rule of thumb” as Swales (1996) points out and consequently it may not always be followed by a RA writer, this paper will examine how expert researchers may deal with this matter of reference and tense, and contrast their textual production with native’s current production. Moreover, the way by which references are made and indicated will also be analyzed, that is, use of footnotes, endnotes, parentheses, and so on.

### 2.2.2. SFL

Studies of genre have considered aspects of lexico-grammar, as it can be seen in Bhatia (2004). In this paper, the textualization of lexico-grammar in expert non-translator’s translation products will be analyzed by using some categories of SFL, for its focus is on meaning and function.

SFL (Halliday & Matthiessen 2004) allows for studies focused on function, on meaning potential, and thereafter on choices. Language can be approached from the perspective of three metafunctions: (i) interpersonal; (ii) ideational; and (iii) textual metafunctions. The first one corresponds to the interactive moves in a dialog; the second one refers to an individual’s construction of experience of the world; and the third is the responsible for the organization and connection of discourse.

The ideational metafunction is analyzed in this paper by means of realization of the system of transitivity, which comprises six types of Processes: (1) material, (2) relational, (3) mental, (4) verbal, (5) existential, and (6) behavioral. Concerning relational Processes, there are two types: (i) attributive, which are irreversible, and (ii) identifying, which are reversible; there are three categories of (a) intensive, (b) possessive, and (c) circumstantial. With reference to the interpersonal metafunction, this paper aims to check how text production can be analysed in terms of polarity (the opposition between positive and negative), modality (several types of indeterminacy, which fall in between positive and negative), wh- elements, mood adjunct, comment adjunct, and interpersonal metaphor. And regarding the textual metafunction, both cohesion and thematic organization are analyzed bearing in mind the genre’s and system’s constraints. Based on SFL, Baker (1994) for example, points out some problems with regard to thematization presented by some novice translators in translating academic articles from Portuguese into English. In this case, Baker claims that the problems are due to genre constrictions. As regards thematic organization, besides Halliday’s categories, Thompson’s (2002) thematising structures (i.e., thematic equatives, predicated themes,

thematized comment, and preposed theme) are taken into account. Moreover, in relation to cohesion, Halliday & Hasan's (1976) *Cohesion in English* is also considered. On such a base, the aim is to identify instances of reference (anaphoric and cataphoric), ellipsis, substitution, and lexical cohesion (synonymy, hyponymy, repetition, collocation).

### 3. Methodology

The Neonatal Screening Program from the State of Minas Gerais of Faculdade de Medicina/UFMG – NUPAD/FM-UFMG – kindly provided some RA to be analyzed in this study. Some researchers that work at the Program can be regarded as expert researchers. All of the RA were produced in English by three different researchers from NUPAD. Three RA were selected randomly to make up the corpus for this analysis, together with three texts produced in English by three international researchers. All texts were selected because they had as core subject *genetic disorders*, in order to work with the same register field. Hereafter, T1, T2, and T3 stand for the group of written texts in English by the Brazilian researchers, or group A; and T4, T5, and T6 stand for the group of texts originally written in English by international researchers, or group B. Group B operates here as a reference corpus for comparison. Obviously the sample is not representative but it can provide some insights for the analysis of group A. The selected RA introductions were analyzed as follows.

1. Regarding genre, the moves and steps were analyzed for each text. Reference and tense combinations were also analyzed.
2. For the lexico-grammar analysis, each text was segmented into sentences and clauses, but only paratactic clauses were analyzed. For each of those clauses, Processes were categorized; Themes were categorized in simple or multiple, marked or unmarked, and whether they were a thematized structure; cohesion was analyzed focusing on the ties on the subject of the disease (genetic disorders) or people with such a disease, finding instances of reference (anaphoric/cataphoric), ellipsis, substitution, and lexical cohesion (synonymy, hyponymy, repetition, collocation).



Move 1	Step 2	Sickle cell disease is the most common genetic disorder in Brazil, and sickle cell anemia, homozygous (HbSS) is the most frequent genotype followed by hemoglobinopathy SC (HbSC) [ <u>1</u> , <u>2</u> and <u>3</u> ].
Move 2	Step 1D	A recent cross-sectional study in northeastern Brazil showed that median weights and heights for age of 76 children and adolescents with sickle cell disease were less than the National Center for Health Statistics (NCHS) reference population standards [4]. The international literature also demonstrates impaired growth in children suffering from this disorder. These children have normal weight and height at birth. In a prospective study, Jamaican children with sickle cell anemia (SS) disease had detectable deficits in both weight and height at <2 years of age. The average deficit approached 1 standard deviation (SD) <control mean for age by 9 years. As for HbSC patients, growth pattern was not different from normal controls [5]. A study recently published on the same Jamaican cohort [6] provided height and weight reference curves for children with HbSS from birth to 22 years of age. It showed that both girls and boys at 18 years of age had significant deficits in weight, more pronounced in boys. At that age, female patients had caught up with NCHS height standards, but male patients still had significant linear growth delay. Two cross-sectional U.S. studies also concluded that patients <18 years of age demonstrated weight and height deficit in both sexes [7 and 8]. SS and S $\beta^0$ thalassemia patients had significantly more pronounced deficits than the SC and S $\beta^+$ thalassemia groups [8].
Move 3	Step 1A	The objective of this study was to assess the growth of children with HbSS and HbSC during a 1-year prospective follow-up and correlate it with hematologic data.

Figure 3 – Moves and steps in an article introduction written by a Brazilian expert researcher.

#### 4. Analysis

##### 4.1. Moves and steps

Both the group of written texts in English by Brazilian researchers and the group of texts produced in English by international researchers presented the moves in the order suggested by Swales (1996) that is, establishing a territory, establishing a niche, and occupying the niche, respectively. Differences can only be found among the steps used within each move. Figure 3 exemplifies the moves and the steps in an introduction written by a Brazilian expert.

As it can be observed in Figure 3, the article written by a Brazilian researcher, like the other articles, fulfills all the genre requirements regarding moves, that is, there is the establishment of a territory followed by the establishment of a niche, and then there is the occupation of the niche. The territory is established through topic

generalizations (Swales's step 2): there are generalizations about the most common genetic diseases and most frequent phenotypes in Brazil. Then, the niche is established by means of continuing a tradition (Swales's step 1D). Such a step could also be Step 3 of move 1 (reviewing items of previous research), as other researches are described briefly; however, a closer look at the article shows the author's argument is to continue a tradition, that is, many other researches are necessary so as to confirm, or contrast, or contradict available data. Finally, the niche is occupied by the outlining of the purposes of the article, made explicitly by "The objective of this study...".

#### 4.2. Reference and Tense

Concerning reference, the following setting can be observed: in the corpus of written texts in English by Brazilian researchers, one presents references in endnotes, as in sentence (1), from T1, and two present references in parentheses with the author's last name followed by the date of publication after the reference passage, as in sentence (2), from T2; in the group of texts written originally in English, the three article introduction present references in endnotes.

(1) SS and S $\beta^0$  thalassemia patients had significantly more pronounced deficits than the SC and S $\beta^+$  thalassemia groups [8].

(2) Consequently, neonatal screening programs have constantly targeted PKU (McCabe et al., 2002; Carreiro-Lewandowski, 2002).

As regards reference and tense combination among the corpus of written texts in English by Brazilian researchers, the following configurations can be presented:

T1	Integral	Non-integral
<b>Reporting</b>	<i>Past</i> 1	<i>Present</i> <i>Perfect</i> 0
<b>Non-reporting</b>		<i>Present (or modal)</i> 2

Figure 4 – Reference and tense in T1.

T2	Integral	Non-integral
<b>Reporting</b>	<i>Past</i> 0	<i>Present</i> <i>Perfect</i> 0
<b>Non-reporting</b>		<i>Present (or modal)</i> 7  <i>Present</i> <i>Perfect</i> 2

Figure 5 – Reference and tense in T2.

T3	Integral	Non-integral
<b>Reporting</b>	<i>Past</i> 0	<i>Present</i> <i>Perfect</i> 0
<b>Non-reporting</b>		<i>Present (or modal)</i> 6

Figure 6 – Reference and tense in T3.

Reference and tense combinations in the corpus of texts written in English by international researchers are as follows.

T4	Integral	Non-integral
<b>Reporting</b>	<i>Past</i> 2	<i>Present</i> <i>Perfect</i> 0
<b>Non-reporting</b>		<i>Present (or modal)</i> 1

Figure 7 – Reference and tense in T4.

T5	Integral	Non-integral
<b>Reporting</b>	<i>Past</i> 0	<i>Present Perfect</i>
<b>Non-reporting</b>		<i>Present (or modal)</i> 2  Present Perfect 3

Figure 8 – Reference and tense in T5.

T6	Integral	Non-integral
<b>Reporting</b>	<i>Past</i> 0	<i>Present Perfect</i> 0
<b>Non-reporting</b>		<i>Present (or modal)</i> 10  Present Perfect 4

Figure 9 – Reference and tense in T6.

An analysis of Figures 4-9 shows that the texts written in English by Brazilian researchers tend to follow Swales's (1996) "rule of thumb" much more than the texts written in English by international researchers. However, it is important to point out that the instances that diverge from such a rule do not seem to be inadequate. The present or modal tenses in the combination non-reporting/non-integral seem to stand for statements in which states of being are accepted as general truths. Seen from an SFL perspective those clauses usually have relational Processes as it can be observed in sentences (3), (4), and (5), from T2, T3, and T6, respectively

- (3) These differences *are* attributable to the history of the populations (Hofman et al., 1991; Perez et al., 1993, 1999; Guldberg et al., 1998; yang et al., 2001; Zschocke, 2003).

(4) PKU *is* an autosomal recessive disorder, with a mean frequency of 1:10.000 in Caucasian populations (Acosta et al., 2001; Scriver and Kaufman, 2001; Christ, 2003; Perez et al., 2004).

(5) The search for predictive markers *has* direct clinical implications, specifically the institution of early interventional strategies, including close monitoring and preventative therapies [7].

On the other hand, when Processes other than relational are used, the present perfect tense tends to be chosen, as in sentences (6), and (7), from T2 and T6, respectively.

(6) Currently, more than 490 PAH mutations *have been described* (PAH Mutation Analysis Consortium; <http://www.pahdb.mcgill.ca>, Scriver et al., 2003).

(7) Recently, the draft sequence of the human genome has generated an extraordinary resource for using single nucleotide polymorphisms (SNPs) as physical markers in either a candidate gene or eventually the entire genome to identify complex traits and modifiers of monogenic diseases. [9]

Finally, as regards the use of simple past in the combination reporting/integral supports Swales's prototypicality, as in the examples (8) and (9) from T1 and T4, respectively. Note, however, that a nominal phrase referring to a given study is provided instead of the author's name.

(8) Early reports on the outcome of pregnancy in homozygous sickle cell (SS) disease depicted an almost universal adverse outcome for mother and child.

(9) A recent cross-sectional study in northeastern Brazil showed that median weights and heights for age of 76 children and adolescents with sickle cell disease were less than the National Center for Health Statistics (NCHS) reference population standards [4].

### 4.3. SFL

#### 4.3.1. Transitivity: Processes

The analysis of the Processes is provided in Table 1.

Process / Text	Mental	Verbal	Behavioral	Existential	Material	Relational					
						Circumstantial		Possessive		Intensive	
						Identify ing	Attribut ive	Identify ing	Attribut ive	Identify ing	Attribut ive
<b>T1</b>	-	6,7	-	-	-	-	-	33,3	-	46,7	13,3
<b>T2</b>	9,5	4,8	-	-	42,9	-	-	-	-	14,3	28,5
<b>T3</b>	16,1	-	-	-	19,3	-	9,7	6,5	3,2	22,6	22,6
<b>T4</b>	18,2	-	-	-	18,2	-	-	9,1	27,2	18,2	9,1
<b>T5</b>	-	37,5	-	-	25	-	12,5	12,5	-	12,5	-
<b>T6</b>	8	-	-	4	64	-	4	8	-	8	4

Table 1 – Percentage of Processes.

According to Table 1, it is not possible to make generalizations concerning each group of introductions or to contrast the two groups. Nevertheless, some insights can be grasped from such numbers. First, it seems that behavioral and existential Processes tend not to be used. The absence of behavioral Processes can be easily explained for usually there is no conscious being participant behaving (even though a patient's behavior could be described) in the introductions. The occurrence of only one instance of existential Processes, as in sentence (8), from T6,

may be accounted for based on Halliday & Matthiessen (2004: 257), according to whom only 3 or 4 percent of clauses are existential.

(8) Currently, controversy *exists* among geneticists as to the best means for identifying the genetic determinants of multigenic or complex traits like stroke in SS disease.

Second, it is possible to state that relational Processes are present in all article introductions in a considerable rate, that is, if circumstantial, possessive and intensive relational Processes are computed, the lowest percentage would be 24%, from which *to be* and *to have* are the most prominent predicators. Such a rate could be explained by the necessity of describing the diseases, as in sentence (9), from T1, and the patients, as in sentence (10), from T1, identifying them or providing them with some particular attribute.

(9) Sickle cell disease *is* the most common genetic disorder in Brazil (...).

(10) In a prospective study, Jamaican children with sickle cell anemia (SS) disease *had* detectable deficits in both weight and height at <years of height.

After relational Processes, the next most common type of Process are material Processes. The use of material Processes is mostly due to the necessity of providing information about procedures, steps and actions concerning the researches or the treatment of patients, or even concerning substances or genes actions, as in sentences (11), (12) and (13), from T2, T2 and T4, respectively.

(11) PAH *catalyses* the hydroxylation of phenylalanine to tyrosine.

(12) Depending on the mutation, some patients *are* more efficiently *treated* with (...).

(13) Such multicenter studies, although carefully conducted and controlled, inevitably *introduce* confounding factors

such as variability in management between centers.

Material Processes are followed by mental Processes, and then by verbal Processes. Here it is important to stress that mental Processes are usually related to the authors' or other researchers' perspective throughout the texts, those authors being themselves or any other researchers the subject (even implicitly) of the active clause or the agent of a passive clause, as in sentences (14) and (15), from T4 and T3, respectively.

(14) We *believe* that the best opportunity for describing true pregnancy outcome is a representative sample of patients in a cohort study from birth.

(15) Phenylketonuria (PKU; OMIM 261600) *has always been considered to be* the classical example of a simple genetic disease for which an effective therapy is possible.

As regards verbal Processes, half the articles present this type of Process. This occurrence could be explained by the fact that references are usually not integral/reporting but non-integral/non-reporting among the article introductions in the corpora, according to Swales (1996), as it can be seen in sentences (16) and (17), from T5.

(16) Here we *describe* an analysis of UGT1A1 promoter polymorphism, bilirubin levels (...).

(17) The coinheritance of Gilbert syndrome with disorders that increase the turnover of red blood cells, or their precursors, has been reported to elevate bilirubin levels in (...).

#### 4.3.2. Thematic Organization

Table 2 is an overview of thematic organization in the analyzed texts.



Theme /	Marked							Unmarked		
	Simple							Simple	Multiple	
Text	clause	circumstance	process	Thematizing structures				Participant	Interpersonal+ ideational	Textual + Ideational
				Predicated	Equative	Preposed	Comment			
<b>T1</b>	-	13,3	-	-	-	-	-	80	-	6,7
<b>T2</b>	4,8	14,3	-	-	-	-	-	76,1	-	4,8
<b>T3</b>	-	9,7	-	-	-	-	-	45,2	16,1	29
<b>T4</b>	-	-	-	-	-	-	-	81,8	-	18,2
<b>T5</b>	-	12,5	-	-	-	-	-	75	-	12,5
<b>T6</b>	4	28	-	-	-	-	4	52	-	12

Table 2 – Thematic organization: percentage

As Table 2 shows, the texts written in English by Brazilian researchers tend to display the same thematic organization as the texts written in English by international researchers. In other words, unmarked Themes prevail over marked Themes. Among unmarked Themes, simple Themes that comprise Participants, as in sentences (18) and (19), from T6, outnumber multiple Themes. Multiple unmarked Themes carry textual and ideational elements in almost all instantiations, as in sentences (20) and (21), from T1, being T3 the only instantiation of interpersonal and ideational elements because of five wh-question clauses, as in (22) and (23).

(18) *Stroke is a major complication of sickle cell (SS) disease (...).*

(19) *Associated risk factors include peripheral leukocytosis (...).*

(20) *(...) and sickle cell anemia, homozygous (HbSS) is the most frequent genotype followed by hemoglobinopathy SC (HbSC).*

(21) (...) *but male patients* still had significant linear growth delay.

(22) *How far are we* from a less burdensome but truly effective PKU treatment?

(23) *What is the real impact of our basic knowledge of PKU* on clinical management and therapy?

Marked themes are simple Themes and almost all of them refer to an extent circumstance (specifically of time or of place), as in sentences (24) and (25), from T6. There is only one example of thematizing structure (specifically a thematized comment), as in sentence (26), from T6, and there are a few examples of simple, marked Theme that refer to a clause complex, as in sentence (27), from T6.

(24) *Recently*, the draft sequence of human genome has generated an extraordinary resource for using single nucleotide polymorphisms (SNPs) as physical markers (...).

(25) *In vitro*, sickle erythrocytes adhere to cytokine-stimulated (...).

(26) *It is estimated* that the lifetime risk for stroke is between 8% and 10%.

(27) *While traditional linkage studies are still attractive*, some have generated controversy (...).

#### 4.3.3. Interpersonal meanings

Table 3 presents categories of interpersonal meanings in the texts.

Feature / Text	Inter-personal metaphor	Wh-question	Mood adjunct	Comment adjunct	Modal-ization	Polar-ization (negative)
T1	-	-	-	-	-	-
T2	-	-	-	-	-	-
T3	-	5	5	-	2	1
T4	1	-	2	-	-	-
T5	-	-	-	-	-	-
T6	-	-	-	-	-	-

Table 3 – Interpersonal meanings.

According to Table 3, interpersonal meanings are not outstanding, at least considering the categories analyzed (i.e., interpersonal metaphor, wh-question, mood adjunct, modalization, and negative polarization). The authors seem to be straightforward in their statements and to prefer to utter affirmative sentences instead of negative. There is no doubt to what is being uttered, especially because the experiments introduce unquestionable data, and the only opening for doubt or contestation is present in percentages lower than 100%, as in sentences (28) and (29), respectively from T2 and T3.

(28) The mean coverage is *approximately 94%* of the births in the state.

(29) Mutations in those genes responsible for BH<sub>4</sub> biosynthesis or regeneration lead to *about 2%* of the HPAs.

Even so, two texts (T3, a text produced by Brazilian researcher; and T4, a text written in English by international authors) present some interpersonal features that are worth discussing. First of all, they are the only texts that discuss previous findings. In other words, instead of presenting data as general and irrefutable truths exclusively to support their argument, the authors discuss those data and their consequence for the patients' treatment and for biases related to the diseases. Perhaps, that is the reason why T3 and T4 allow the use of some interpersonal

elements that are not present among the other four texts. Moreover, T3 presents particularly the less academic-like title: “The time has come: a new scene for PKU treatment”, pointing out since the very beginning to the discussion the author carries out throughout the text. One example of each interpersonal feature is provided in sentences (30)-(34): interpersonal metaphor, wh- question, mood adjunct, modalization, and negative polarization, respectively.

(30) *We believe* that the best opportunity for describing true pregnancy outcome is a representative sample of patients in a cohort study from birth. (T4)

(31) *When are* we going to be able to adapt therapeutics to patient individualities? (T3)

(32) (...) because the patients have been ascertained *usually* through presentation at antenatal clinics (...)

(33) In this context, some questions *can* be put forward.

(34) (...) and phenotype-genotype correlations in PKU are *not* easy to establish. (T3)

Besides the interpersonal features displayed in Figure 11, interpersonal elements could also be found in lexical items such as: “unexpected complexity” and “truly effective”, “extremely restrictive”, “rarely predictable”, in T3, and “carefully conducted”, “potentially serious biases”, in T4.

#### 4.3.4. Cohesion

The analysis of the ties linked to the disease dealt with in the article introductions can be summarized in figure 13.

Type / Text	Lexical Cohesion				Reference		Ellipsis	Substitution
	Synonymy	Hyponymy	Repetition	Collocation	Anaphora	Cataphora		
T1	2	1	2	2				
T2	3		2	3				
T3	3		4	6				
T4			2	3				
T5	2	1		2				
T6		1	6	4				

Table 4 – Types of cohesion: occurrences.

According to Table 4, ellipsis and substitutions are not used among group A and group B, neither is anaphora or cataphora. It is interesting then to observe that only lexical cohesion is used to make reference to the diseases. As regards synonymy, in sentences (36) and (38), sickle cell anemia and sickle cell anemia (SS) disease (underlined) stand for the same disorder are illustrating. In terms of hyponymy, the name of the disease may be regarded as a hyponym for the superordinate disorder or disease, as in sentences (35) and (37). In terms of repetition, items are used through simple repetition, as in sentence (37), in which the item disorder is used again after the sentence (35). And in terms of collocation, the item disease, disorder or the very name of the disease is usually employed along with the node patient (or the like: e.g., children, adolescent), as in sentences (35) and (38). Due space constricts the examples are only from T1, as it has been already provided in section 4.1., and the sentences follow the order they appear in the texts, and thus the reader may have a better overall view of the text and its ties.

(35) *Sickle cell disease* is the most common genetic *disorder* in Brazil (...)

(36) and *sickle cell anemia*, homozygous (*HbSS*) is the most frequent genotype (...)

(37) The international literature also demonstrates impaired growth in children suffering from this *disorder*.

(38) *Jamaican children with sickle cell anemia (SS) disease* had detectable deficits in both weight and height at <2 years of age.

In general, there is no considerable difference between group A and group B. However, the presence of more synonymy in group A may be a result of the tendency for Portuguese writers to avoid repetition in texts.

## 5. Final remarks

Building on interpersonal features, thematic organization, transitivity and cohesion and Swales' and Bhatia's studies of genres, mainly RA introductions, the present paper sought to provide a first description of the texts written in English by Brazilian researchers.

By means of the analysis, as it could be already expected, since the written texts in English by Brazilian researchers had previously been accepted and published in internationally indexed journals, the written texts in English by Brazilian researchers seem to fulfill generally some basic features of the genre article introduction, namely the moves and steps, and the reference and tense combination. Furthermore, the texts generally cohere and present a lexico-grammatical display/organization very similar to those texts written by researchers whose mother language is English. Such fulfillment and similarity may be explained by the fact that the authors as members (both producers and consumers) of the academic community share some rhetoric and discursive knowledge that allows them to write adequately even in a second language (provided they have enough competence on that language).

## REFERENCES

- ALVES, F. 2005. Ritmo cognitivo, meta-função e experiência: parâmetros de análise processual no desempenho de tradutores novatos e experientes. In: ALVES, F.; MAGALHÃES, C.; PAGANO, A. (org.). *Competência em tradução: cognição e discurso*. Belo Horizonte: Ed. UFMG.
- BAKER, M. 1994. *In other words: a coursebook on translation*. London: Routledge.
- BHATIA, V. 2004. *Worlds of written discourse*. London and New York: Continuum.
- HALLIDAY, M.A.K.; HASAN, R. 1976. *Cohesion in English*. London and New York: Longman.
- HALLIDAY, M.A.K.; MATTHIESSEN, C.M.I.M. 2004. *An introduction to Functional Grammar*. 3<sup>rd</sup> ed. London: Arnold.
- JAKOBSEN, A. L. 2002. Translation drafting by professional translators and by translation students. In: HANSEN, G. (Ed.). *Empirical translation studies: process and product*. Copenhagen: Samfundslitteratur.
- PAGANO, A.; VASCONCELLOS. 2005. Explorando interfaces: estudos da tradução, linguística sistêmico-funcional e linguística de corpus. In: ALVES, F.; MAGALHÃES, C.; PAGANO, A. *Competência em tradução: cognição e discurso*. Belo Horizonte: UFMG, 2005.
- SWALES, J.M. 1996. *Genre Analysis: English in academic and research settings*. Cambridge: Cambridge.
- THOMPSON, G. 2002. *Introducing Functional Grammar*. London: Arnold.

---

<sup>i</sup> Igor Antônio Lourenço da Silva is an MA and is currently studying with a grant from CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico).

<sup>ii</sup> Kelen Cristina Sant'Anna de Lima is a researcher from Nupad-FM/UFMG (Núcleo de Ações e Pesquisa em Apoio Diagnóstico da Faculdade de Medicina da Universidade Federal de Minas Gerais)

<sup>iii</sup> Pagano is Associate Professor at Faculdade de Letras, UFMG.