

VIVIANE POSSAMAI

**MARCADORES TEXTUAIS DO ARTIGO CIENTÍFICO EM COMPARAÇÃO
PORTUGUÊS E INGLÊS - UM ESTUDO SOB A PERSPECTIVA DA TRADUÇÃO**

Dissertação apresentada à Banca Examinadora da Universidade Federal do Rio Grande do Sul, como exigência parcial para obtenção do título de MESTRE em Teorias do Texto e do Discurso, sob orientação da Profa. Dra. Maria da Graça Krieger.

PORTO ALEGRE

2004

BANCA EXAMINADORA

Dr. João Azenha Jr. (USP)

Dra. Erica Sofia F. Schultz (UFRGS)

Dra. Maria José B. Finatto (UFRGS)

AGRADECIMENTOS

À minha orientadora, Profa. Maria da Graça Krieger.

Às professoras do Projeto Termisul: Anna Maria B. Maciel, Cleci Bevilacqua, e Maria José Finatto.

Aos professores do PPG-Letras: Luzia Araújo, Sônia Gehring e Valdir Flores.

Aos meus pais, Deomede e Jurema.

Ao Felipe.

Às minhas amigas Sandra, Ciça, Giselle, Susana e Cristiane.

Ao meu irmão Everson, à minha cunhada Andréa e à pequena Celine.

E ao CNPq, pela bolsa concedida.

SUMÁRIO

RESUMO	7
ABSTRACT	8
LISTA DE QUADROS	9
LISTA DE TABELAS, GRÁFICOS E FIGURAS	11
SIGLÁRIO	13
INTRODUÇÃO: O ARTIGO CIENTÍFICO E TRADUÇÃO.....	15
CAPÍTULO 1: OS ESTUDOS DE GÊNERO E O TEXTO ESPECIALIZADO	20
1.1 Introdução	20
1.1.1 Gêneros, classes e tipos textuais	21
1.2 Gêneros em diferentes perspectivas.....	23
1.2.1 Aristóteles e a retórica.....	23
1.2.2 Bakhtin e a noção de gênero discursivo	25
1.2.3 Swales e a noção de comunidade discursiva.....	29
1.2.4 Aproximando Swales e Bakhtin.....	35
1.3 O texto de especialidade e gênero.....	40
CAPÍTULO 2: O ARTIGO CIENTÍFICO E OS MARCADORES TEXTUAIS	50
2.1 Introdução	50
2.1.1 A produção de artigos científicos em números	51

2.2 O estudo do artigo científico.....	53
2.2.1 Características textuais do artigo científico	58
2.3 Os marcadores textuais do artigo científico.....	59
2.3.1 Os marcadores textuais e o modelo sistêmico-funcional	63
2.3.2 O modelo sistêmico-funcional e o artigo científico	69
2.4 Algumas considerações.....	80
CAPÍTULO 3: TRADUÇÃO E RETÓRICA CONTRASTIVA	83
3.1 Introdução	83
3.2 Retórica Contrastiva	83
3.2.1 Retórica Contrastiva e Tradução.....	88
CAPÍTULO 4: DESCRIÇÃO DO <i>CORPUS</i>	92
4.1 Introdução	92
4.2 O <i>corpus</i> em português	94
4.2.2 Textos selecionados	96
4.2.3 Eventos e subáreas que constituíram as fontes	98
4.3 O <i>corpus</i> em inglês	101
4.3.1 As fontes.....	102
CAPÍTULO 5: METODOLOGIA.....	109
5.1 Introdução	109
5.2 A Linguística de <i>Corpus</i>	109
5.2.1 A LC aplicada.....	114
5.3 Tratamento do <i>corpus</i>	116
5.3.1 As ferramentas utilizadas	116
5.3.2 Tratamento do corpus em português	117
5.3.3 Tratamento do corpus em inglês	121

CAPÍTULO 6: ANÁLISE COMPARATIVA DOS MARCADORES TEXTUAIS

DO ARTIGO CIENTÍFICO EM PORTUGUÊS E INGLÊS	122
6.1 Introdução	122
6.2 Marcadores Textuais em comparação.....	123
6.2.1 Marcadores relacionados ao propósito do artigo científico	124
6.2.2 Marcadores metarreferenciais	137
6.2.3 Marcadores subjetivos.....	145
6.3 Algumas considerações.....	150
CONSIDERAÇÕES FINAIS	155
REFERÊNCIAS	160

RESUMO

Nesta dissertação, faz-se um estudo do gênero artigo científico, em especial de expressões típicas de seu desenvolvimento e organização. Essas expressões foram denominadas **marcadores textuais**, e os objetivos deste trabalho foram identificá-las e verificar seus padrões de uso em textos em português e inglês, em um estudo permeado pelos interesses e pelas perspectivas da tradução e apoiado pelos pressupostos da Retórica Contrastiva. Para esse fim, foram utilizados dois *corpora*: um em português, composto de 333 artigos, e outro em inglês, composto de 111 artigos. Os mesmos foram analisados utilizando-se a ferramenta WordSmith Tools, empregada pelos estudos em Lingüística de *Corpus*. Os marcadores selecionados a partir dos *corpora* foram classificados com base nas metafunções da linguagem propostas por Halliday. Após essa classificação, comparamos as ocorrências das unidades em inglês e português, observando padrões de uso, frequência e colocação.

ABSTRACT

The present work approaches the research article genre, especially in what concerns typical expressions of its development and textual organization. These expressions were gathered under the denomination **textual markers**, and our objective was to verify their patterns of use, collocation and frequency in two *corpora*. This study has been carried out with the aim of contributing both for theory and practice of translation, and for that end it has been supported by studies of Genre, Contrastive Rhetoric, and Functional Grammar. The two *corpora* analyzed were composed respectively of 333 research articles in Portuguese and 111 research articles in English in the area of Computer Science. For *corpora* analysis we have used the WordSmith Tools software, a tool highly employed by *Corpus Linguistics*. At the end of this study we classified the textual markers selected from corpora and compared their patterns of use, frequency and collocation in English and Portuguese.

LISTA DE QUADROS

Quadro 1.1 - Propriedades das comunidades discursivas em dois momentos da obra de Swales – 1990 e 1992	34
Quadro 4.1 Eventos da SBC - 2001-2003	95
Quadro 6.1 Ocorrência dos marcadores agrupados em função da palavra-núcleo “objetivo” nos corpora nas línguas portuguesa e inglesa	125
Quadro 6.2 Ocorrência dos marcadores agrupados em função da palavra-núcleo “forma” nos corpora nas línguas portuguesa e inglesa.....	126
Quadro 6.3 Ocorrência de expressões agrupadas pelo grupo funcional “método”	128
Quadro 6.4 Ocorrência de expressões agrupadas pelo grupo funcional “ancoragem”	130
Quadro 6.5 Ocorrência de expressões agrupadas pelo grupo funcional “quantificação”	131
Quadro 6.6 Ocorrência de expressões agrupadas pelo grupo funcional “tempo”	133
Quadro 6.7 Ocorrência de expressões agrupadas pelo grupo funcional “explicação”	134
Quadro 6.8 Ocorrência de expressões agrupadas pelo grupo funcional “restrição”	136
Quadro 6.9 Ocorrência dos marcadores agrupados em função da palavra-núcleo “seção” nos corpora nas línguas portuguesa e inglesa.....	137

Quadro 6.10 Ocorrência dos marcadores agrupados em função da palavra-núcleo	
“figura” nos corpora nas línguas portuguesa e inglesa.....	139
Quadro 6.11 Ocorrência dos marcadores agrupados em função da palavra-núcleo	
“tabela” nos corpora nas línguas portuguesa e inglesa.....	141
Quadro 6.12 Ocorrência dos marcadores agrupados em função da palavra-núcleo	
“artigo” nos corpora nas línguas portuguesa e inglesa.....	142
Quadro 6.13 Ocorrência dos marcadores agrupados em função da palavra-núcleo	
“trabalho” nos corpora nas línguas portuguesa e inglesa	143
Quadro 6.14 Ocorrência dos marcadores agrupados em função de seu caráter	
explícito de organização do texto	144
Quadro 6.15 Ocorrência dos marcadores agrupados em função da palavra-núcleo	
“importante” nos corpora nas línguas portuguesa e inglesa.....	146
Quadro 6.16 Ocorrência dos marcadores agrupados em função da palavra-núcleo	
“necessário” nos corpora nas línguas portuguesa e inglesa.....	147
Quadro 6.17 Ocorrência dos marcadores agrupados em função das palavras-	
núcleo “esperar / acreditar” nos corpora nas línguas portuguesa e	
inglesa.....	147
Quadro 6.18 Ocorrência dos marcadores agrupados em função da palavra-núcleo	
“observar” nos corpora nas línguas portuguesa e inglesa.....	148
Quadro 6.19 Ocorrência dos marcadores agrupados em função da palavra-núcleo	
“poder” nos corpora nas línguas portuguesa e inglesa	149

LISTA DE TABELAS, GRÁFICOS E FIGURAS

Tabela 2.1 Publicação de artigos brasileiros, por grandes áreas de conhecimento, em periódicos científicos internacionais indexados e em anais de eventos internacionais - 1998-2001	53
Tabela 4.1 Trabalhos apresentados e respectivos artigos publicados em eventos da SBC – 2001, 2002, 2003.....	96
Tabela 4.2 Artigos de eventos da SBC que compõem o <i>corpus</i> em língua portuguesa – 2001, 2002 e 2003.....	97
Tabela 4.3 Fontes para a composição do <i>corpus</i> em língua inglesa.....	107
Gráfico 2.1 Artigos publicados em periódicos científicos internacionais indexados – 1981, 1997 e 2002	51
Gráfico 2.2 Artigos brasileiros publicados em periódicos científicos internacionais indexados - 1997 e 2002.....	52
Figura 2.1 Categorias de revisão textual enquanto realizações dos componentes metafuncionais da linguagem de Halliday.....	72

Figura 2.2 Relação entre as categorias estabelecidas para este trabalho, as
categorias apresentadas por Gosden (1985) e os componentes
metafuncionais da linguagem. 81

SIGLÁRIO

CLEI	- Centro Latinoamericano de Estudios en Informática
CTD	- Concurso de Teses e Dissertações
CTIC	- Concurso de Trabalhos de Iniciação Científica
ENIA	- Encontro Nacional de Inteligência Artificial
IA	- Inteligência Artificial
IAS	- Intelligent Society of Autonomous Systems
IEEE	- Institute of Electrical and Electronics Engineers
IFIP	- International Federation for Information Processing
ISI	- Institute for Scientific Information
JAI	- Jornadas de Atualização em Informática
LC	- Lingüística de Corpus
NILC-SP	- Núcleo Interinstitucional de Lingüística Computacional de São Paulo
NSI	- National Science Indicators
MCT	- Ministério de Ciência e Tecnologia
RC	- Retórica Contrastiva
RMc	- Alterações relacionadas às afirmações dos redatores do texto
RMd	- Modificações textuais que se relacionam à estrutura do discurso

- RMp - Alterações relacionadas ao propósito do redator do texto
- SBC - Sociedade Brasileira de Computação
- SBC&M - Simpósio Brasileiro de Computação Musical
- SBPC - Sociedade Brasileira para o Progresso da Ciência
- SEMISH - Seminário Integrado de Software e Hardware
- WEI - *Workshop* sobre Educação em Computação
- WIE - *Workshop* de Informática na Escola
- WPerformance - *Workshop* em Desempenho de Sistemas Computacionais e de Comunicação

INTRODUÇÃO: O ARTIGO CIENTÍFICO E TRADUÇÃO

Esta dissertação de mestrado é fruto de um misto de observações e de inquietações surgidas no período em que trabalhei como tradutora no Instituto de Informática da UFRGS, de 1998 a 2002. O trabalho envolvia, sobretudo, a realização de versões e de revisões de artigos científicos para e na língua inglesa, os quais eram produzidos por membros de projetos de pesquisa: mestrandos, doutorandos e professores. Recém egressa do curso de Bacharelado em Letras da UFRGS, deparava-me com muitos trabalhos nesse mesmo tipo de texto, do qual pouco conhecia as particularidades e as convenções, o que ocasionava dúvidas sobre como seriam seus similares em inglês e, portanto, como deveriam ser traduzidos para não fugir às expectativas da comunidade na qual se inseririam, assim como da banca que julgaria a aceitação ou não do trabalho. Também percebia que, nos textos em inglês produzidos por pesquisadores brasileiros, os quais minha tarefa era revisar, os aspectos sintáticos, semânticos e pragmáticos do português eram repassados para o inglês, o que muitas vezes deixava o texto truncado e estilisticamente estranho na língua de chegada.

A prática e o volume de trabalho realizados com esses textos começaram a evidenciar convenções e características que lhes eram específicas e regulares, tanto em um nível mais genérico e já bem conhecido, que poderíamos chamar de macroestrutural (a divisão em seções, os resumos em duas línguas, a identificação dos autores, etc.), como também em outro mais específico, interno e intrínseco à constituição da trama textual. Foi nesse ponto que a

observação e a inquietação acima mencionadas começaram a delinear o objeto, os pressupostos e as motivações para esta pesquisa, que, de uma observação individual, passou a ser formalmente conduzida sob orientação durante o Programa de Pós-Graduação da UFRGS¹.

Sendo assim, interessaram-nos exatamente aquelas expressões do texto que se mostravam sistemáticas e regulares no desenvolvimento, na coesão e na organização dos artigos científicos. O objetivo foi, então, a partir de um *corpus* em português, identificar que expressões eram essas, utilizando critérios de frequência e uso dessas formas regulares.

Como esse foi um trabalho que surgiu de uma observação na prática da tradução, ele envolveu também uma segunda fase de análise; após a identificação das expressões em português, procuramos fazer o mesmo a partir de um *corpus* em inglês, e, assim, identificar, através de comparações, as diferenças de uso, forma e frequência nas duas línguas. O objetivo final foi mostrar algumas dessas diferenças, para chamar a atenção para o uso dessas expressões durante o processo tradutório. Isso também porque queremos mostrar aspectos do texto científico que são tão importantes para a sua boa constituição como são os termos técnicos, já que muito se acredita que na tradução técnica o detentor de terminologia de uma dada área é obrigatoriamente melhor tradutor do que quem não a detém. Os aspectos discursivos de um texto técnico ou científico, ou seja, esses outros fatores que fazem parte das linguagens especializadas e que estão no mesmo ambiente que o termo, são muitas vezes esquecidos por parte de quem traduz e/ou ensina tradução.

Para a realização desta investigação, lançamos mão dos estudos de gênero e de alguns pressupostos da Retórica Contrastiva. Para a análise dos textos, empregamos o ferramental da

¹ Doravante este texto será escrito na primeira pessoa no plural, uma vez que, a partir deste momento relatado, passou a ser uma pesquisa realizada não apenas por mim, mas dentro de um contexto acadêmico e de pesquisa, o qual, sabemos, não recomenda solidão.

Linguística de *Corpus* e o modelo das macrofunções da linguagem da gramática sistêmico-funcional para classificação das estruturas selecionadas.

No Capítulo 1 fazemos uma introdução aos estudos de gênero que consideramos de maior visibilidade e relevância. Realizamos um percurso de ordem cronológica que abrange os estudos de Aristóteles, Mikhail Bakhtin e Swales. Com Aristóteles, pretendemos estudar as origens dos estudos de gênero, observando que muitos dos pontos ali apresentados encontram pertinência e relevância ainda hoje, como a própria noção de retórica. Bakhtin foi incluído devido à importância de suas conclusões e de suas reflexões a respeito de gênero e de comunidade discursiva. A leitura de Swales mostrou-se fundamental devido ao seu já bastante conhecido trabalho com gêneros e, principalmente, com artigos científicos.

O intuito principal desse percurso foi situar e delimitar o objeto *lato sensu* da pesquisa: o gênero artigo científico, que é observado em detalhe no Capítulo 2. Nesse capítulo, utilizamos principalmente os dados sobre artigo científico expostos por Swales, em 1990, já que este continua se constituindo como um referencial válido e importante para estudos de gênero e, em especial, de artigos científicos. De Swales, utilizamos principalmente suas reflexões históricas e sua compilação sobre estudos já realizados. Aspectos do estudo de Swales que não são contemplados nesta dissertação são a divisão do artigo em passos e movimentos, e o estudo do artigo separado pelas partes (Introdução, Métodos, Resultados e Conclusões), uma vez que logo percebemos que os textos de nosso *corpus* não estavam assim divididos formalmente, como se espera de um número grande de textos e de ocorrências reais.

Nesse capítulo, aumentamos ainda mais a nossa lente de observação, voltando-nos para características textuais do artigo, em especial, para nosso objeto específico, os marcadores textuais. Demonstramos a classificação feita, após a identificação das expressões a partir do *corpus* em português, que se baseou em duas das macrofunções da linguagem propostas por Halliday (1985): interpessoal e textual. A classificação das estruturas

identificadas ficou assim constituída: (i) marcadores textuais relacionados ao propósito do artigo científico; (ii) marcadores textuais metarreferenciais; e (iii) marcadores subjetivos. Nesse capítulo, trouxemos apenas um exemplo, contextualizado, de cada marcador.

Como dissemos, este é um estudo que tem a perspectiva do processo tradutório como fio condutor; coube, então, fazer um contraste entre modos de realização de tais marcadores em duas línguas: o português e o inglês. Pressupostos da Retórica Contrastiva, e de sua relação com estudos de tradução deram suportes teóricos para este trabalho, e estão apresentados no Capítulo 3. Resumidamente, estudos de Retórica Contrastiva demonstram que cada cultura desenvolve normas de escrita específicas e distintas e procura estudar e descrever essas especificidades e distinções.

Os *corpora* que analisamos para evidenciar os marcadores textuais típicos dos artigos fazem parte de uma grande área, a Ciência da Computação, e se constituem de textos de diferentes subáreas, como Inteligência Artificial, Desenvolvimento de Software e Hardware, Robótica, Informática na Educação, dentre outros. O *corpus* em português constitui-se de 333 artigos publicados nos Anais dos Congressos da Sociedade Brasileira de Computação de 2001, 2002 e 2003; o *corpus* em inglês constitui-se de 111 artigos publicados em diferentes periódicos de circulação internacional. O Capítulo 4 trata da descrição dos *corpora* utilizados, tentando, o máximo possível, localizá-los em seu contexto de ocorrência, isto é, nas publicações onde se realizaram. Com isso, intencionamos dar uma idéia de como os artigos científicos ocorrem dentro do que podemos chamar de uma grande comunidade científica, que estabelece normas, diretrizes e também padrões convencionais de realização desses textos, mas que, como nos mostra a Retórica Contrastiva, não impede que cada língua as realize de maneira diferente.

Para os dois momentos de análise realizados nesta dissertação, que são, primeiramente, a identificação dos marcadores somente em português e, depois,

comparativamente, em inglês e português, valemo-nos das ferramentas empregadas pela Lingüística de *Corpus*. Ao possibilitar a abordagem de uma quantidade surpreendente de exemplos e informações sobre a língua, uma vez que conta com tecnologia para tanto, a Lingüística de *Corpus* e suas ferramentas proporcionam uma análise de dados empíricos, importantíssimos, portanto, para o tradutor, que se vê envolvido em situações em que é preciso saber “como se diz” determinada porção de linguagem em um outro contexto cultural, muitas vezes distante do seu. Nossa abordagem metodológica é relatada no Capítulo 5, onde se irá encontrar também uma descrição da Lingüística de *Corpus*, seus fundamentos e pressupostos.

A seguir, no Capítulo 6, são apresentados quadros com os resultados das análises comparativas realizadas a partir dos *corpora* descritos anteriormente. Esse capítulo está organizado de acordo com a classificação das unidades descritas no Capítulo 2, e cada seção apresenta comentários e observações a respeito dessa comparação logo abaixo de cada quadro, visando demonstrar diferenças de uso, de frequência e de colocação das unidades em português e inglês. Ao final, apresentamos também uma sistematização dos resultados, mas que não constitui o fim maior desta pesquisa, já que nosso foco está, principalmente, nas diferenças e nas particularidades.

Por fim, no Capítulo 7, apresentamos nossos comentários finais, recapitulando o que foi descrito e estudado e encerrando a investigação acerca dos artigos científicos e de seus marcadores em comparação inglês–português.

CAPÍTULO 1: OS ESTUDOS DE GÊNERO E O TEXTO ESPECIALIZADO

1.1 Introdução

Enquanto falantes e usuários de uma língua, sabemos diferenciar intuitivamente distintas realizações de textos – tomando, aqui, **texto** em uma visão muito ampla, de um todo que faz sentido (Barros, 1990). Nosso conhecimento da língua e do mundo permite-nos reconhecer e distinguir uma matéria jornalística, uma bula de remédio, um conto de fadas, um artigo científico, uma receita, uma carta, um telefonema, um discurso político e muitos outros textos que fazem parte de nosso dia-a-dia e que ajudam a organizar e a estabilizar nossas atividades comunicativas. Porém, o que cabe aqui, neste capítulo, é tratar do assunto com mais acuidade e embasamento do que permite nossa intuição, nossa necessidade de ordenamento e nossa capacidade de distinguir diferentes gêneros, uma vez que esta dissertação enfoca certas expressões textuais que se mostram típicas de um gênero em especial: o artigo científico, o qual é tratado em detalhe no Capítulo 2. O que se espera com isso é trazer algumas das reflexões a respeito do tema que, entendemos, delimitarão nosso objeto de estudo e trarão à tona os conceitos de gênero vinculados a uma idéia de esfera social e de comunidade discursiva, conforme posto por Mikhail Bakhtin (1997) e John M. Swales (1990), sempre tendo como meta final uma reflexão acerca da necessidade do reconhecimento dos gêneros e de seus constituintes por parte de quem lida com esses textos, aqui, em especial, os tradutores.

Este capítulo também traz algumas reflexões acerca do texto que materializa o gênero artigo científico: o texto de especialidade. Embora não se pretenda aqui um estudo voltado para uma área de especialidade, acreditamos que uma aproximação entre as reflexões sobre gêneros e textos de especialidade pode vir a complementar esta dissertação, abrindo caminho também para os estudos do texto de especialidade que enfocam outros elementos do texto, além dos termos e das fraseologias de uma dada área. Nesta seção, também trazemos algumas questões implicadas na relação entre os estudos da linguagem e do texto especializado, da terminologia e da tradução, já que nossas preocupações, nesta dissertação, estão voltadas para o processo tradutório.

1.1.1 Gêneros, classes e tipos textuais

O título deste capítulo evidencia que tratamos de questões relacionadas aos estudos de gênero, porém consideramos adequado fazer uma pequena introdução à distinção que se faz, no âmbito dos estudos acadêmicos, entre o que são gêneros, classes e tipos textuais. A idéia é tentar delimitar melhor a que estamos nos referindo quando falamos de **gênero**, ainda assim com consciência de que, como em toda classificação, podemos estar sendo limitados e limitadores.

Em princípio, parece-nos que essas três denominações, gênero, classe e tipo textual, são sinônimos que poderiam ser empregados indistintamente, porém cada uma delas implica diferentes pontos de vista, pois foram cunhadas em diferentes momentos e utilizadas por distintas teorias no contexto dos estudos de gênero e da lingüística textual.

A expressão **gênero textual** tem seu fundamento atribuído principalmente à conceitualização de **gênero discursivo** na obra de Bakhtin, **Estética da Criação Verbal** (1997). Alguns autores, como Marcuschi (2002), defendem a idéia de que é impossível se comunicar verbalmente a não ser por algum gênero (mesma idéia de Bakhtin), assim como é

impossível se comunicar verbalmente a não ser por algum texto, concluindo então que a comunicação verbal só é possível por algum gênero textual. É evidente, pela leitura do texto desse autor que ele segue as idéias de Bakhtin, fazendo a noção de gênero discursivo evoluir para a de gênero textual, sem distinções. Posteriormente, discutimos os pressupostos que acompanham e definem o ponto de vista da denominação introduzida por Bakhtin e vemos que não há mesmo por que haver tal distinção, pelo menos nos pontos que nos interessam nesta pesquisa.

No que concerne ainda ao uso de termo **gênero**, é preciso também lembrar que “(...) hoje, gênero é facilmente usado para referir uma categoria distintiva de discurso de qualquer tipo, falado ou escrito, com ou sem aspirações literárias” (Swales, 1990, p. 33), sendo usado inclusive em áreas como folclore, retórica, sociologia e antropologia; porém vale destacar que a noção de gêneros em literatura ainda é muito presente, e as distinções entre os mesmos (soneto, prosa, poesia, romance, dentre outros) continuam sendo de muita relevância para os estudos nessa área.

Já os termos classe e tipo textual são utilizados quase exclusivamente pelas vertentes da lingüística textual. A classe textual seria exatamente o que sabemos enquanto falantes de uma língua. Ciapuscio² (1994, p. 25) define classe textual, citando Heinemann e Viehweger:

“(...) classe textual se aplica hoje a classificações empíricas, tal como são realizadas por membros de uma comunidade lingüística, isto é, classificações cotidianas que podem ser mencionadas por meio de determinados lexemas condensadores do saber sobre determinada classe textual: por exemplo, ‘isto é um conto’, ‘isto é uma piada’ (...) etc”.

Nesse caso, parece não haver maior distinção entre os conceitos de gênero e classe.

O **tipo textual** refere-se a uma classificação científica dos textos. Segundo Marcuschi (2002, p. 22) a expressão **tipo textual** designa uma

² Todas as citações feitas nesta dissertação, cujos originais não estão em português, foram traduzidas por nós.

“(...) espécie de construção teórica definida pela natureza lingüística de sua composição (aspectos lexicais, sintáticos, tempos verbais, relações lógicas). Em geral, os tipos textuais abrangem cerca de meia dúzia de categorias conhecidas como: narração, argumentação, exposição, descrição, injunção”.

Os tipos textuais “(...) constituem seqüências lingüísticas ou seqüências de enunciados e não são textos empíricos” (id., p. 23), isto porque um texto, em geral, contém segmentos de vários tipos de texto, assim, um determinado gênero pode conter um ou mais tipos de texto em sua composição, por exemplo, um artigo científico poderá conter elementos narrativos e descritivos (apresentação da metodologia de pesquisa) e elementos argumentativos (a importância do estudo).

No decorrer deste capítulo, entendemos que fica bastante claro, pelas definições e pelas reflexões de alguns autores consagrados, o motivo pelo qual o conceito de gênero se aplica muito bem ao nosso estudo. Estamos lidando com artigos científicos, textos que já são aceitos e reconhecidos como constituintes da comunicação de uma grande comunidade, a comunidade científica internacional, e que já têm um formato bastante estável, portanto textos que constituem um gênero textual.

1.2 Gêneros em diferentes perspectivas

A seguir, apresentamos alguns estudos de gênero que consideramos relevantes para este trabalho.

1.2.1 Aristóteles e a retórica

É intrínseco ao ser humano classificar, agrupar por semelhanças, denominar ou “separar em gavetas” o conhecimento. Isso parece dar ao homem uma certeza de que é possível organizar o caos em que vive. Ciapuscio (1994) comenta que essa forma do sistema cognitivo de operar humano teve, desde a Antigüidade, seu reflexo nas artes e nas

humanidades. No que se refere à distinção de diferentes tipos de textos ou discursos, a obra **Arte Retórica e a Arte Poética**, de Aristóteles, é apontada como os primórdios do aparecimento do conceito de gênero, ainda no século III a.C.

Aristóteles chamava atenção para os propósitos dos discursos e para as audiências particulares de cada um deles, o que acarretaria a produção de diferentes gêneros (Bezerra, 2001). O filósofo distinguiu três gêneros de discursos oratórios: o judiciário, o deliberativo e o demonstrativo (Aristóteles, 1959, p. 32), estreitamente dirigidos a como proceder de acordo com assunto, audiência e finalidade e se obter o resultado desejado. Observa-se, desde então, a importância do tema, da função e da audiência (que poderíamos ousar chamar de **outro**, como em Bakhtin) para a distinção entre os gêneros.

Essa maneira de proceder visando a um fim foi abarcada pela retórica, sendo esta, para Aristóteles, compreendida como a arte de persuadir, “(...) a faculdade de ver teoricamente o que, em cada caso, pode ser capaz de gerar persuasão. A Retórica parece ser capaz de, por assim dizer, no concernente a uma dada questão, descobrir o que é próprio para persuadir” (Aristóteles, 1959, p. 24).

Os estudos literários também se valeram do conceito de gênero disposto em **A poética**, de Aristóteles, em que ele fez a tradicional distinção entre gênero épico, dramático e lírico. Até hoje, a literatura retoma essa distinção e suas subcategorias, como o soneto, a ode, a epopéia, etc.

Mas as observações acima se prestam, neste trabalho, mais para fazer um retorno às origens e uma observação histórica do surgimento do termo **gênero**. Também nos interessa o caráter argumentativo da retórica, conforme definiu Todorov (1971, p. 81):

“Em seus primórdios, a Retórica é, acima de tudo, uma técnica que deve permitir, a quem a possua, atingir, dentro de uma situação discursiva, o objetivo desejado; ela tem, portanto, um caráter pragmático: convencer o interlocutor da justeza de sua causa”.

Uma aproximação que podemos fazer das noções de gênero e retórica, em Aristóteles, com o nosso trabalho, é que o caráter de persuasão da retórica é uma das características que constitui a razão de ser do artigo científico, um gênero que veio a se constituir como tal muito depois de Aristóteles. O texto e a retórica de um artigo visam, em geral, validar uma pesquisa e torná-la conhecida, sem contar ainda o fato de que, antes de (e para) atingir seu público-alvo, o texto deverá ser submetido a uma banca examinadora, que irá determinar a aceitação ou não do trabalho, e, portanto, julgar se o mesmo merece a divulgação desejada pelos autores.

Depois de Aristóteles, a reflexão sobre gêneros começa a adquirir um caráter um tanto diferenciado a partir do ensaio **O Problema dos Gêneros do Discurso**, escrito por Bakhtin em 1953,³ o qual traz, para esta pesquisa, também as noções de comunidade discursiva e gênero discursivo, e de que os enunciados estão ligados a uma esfera social determinada. Tais aspectos nos interessam por ressaltarem ainda mais a importância de se conhecer o gênero na realização de um texto, e, principalmente, no caso de um texto que vai operar em outra língua, seja ele oriundo de uma versão (tradução do português para o inglês) ou produzido em uma língua não materna.

1.2.2 Bakhtin e a noção de gênero discursivo

Existem teóricos, como a pesquisadora Roxane Rojo (Rojo, 2004), que acreditam que o ensaio **O problema dos gêneros do discurso**,

“é uma formalização explícita (e de divulgação) – incompleta e nem tão bem acabada – das idéias que estavam disponíveis no círculo (de Moscou), desde o início. Algumas abordagens anteriores a esse texto dizem mais sobre o que são e como funcionam os gêneros e de maneira mais aprofundada e complexa”.

³ Para este trabalho, utilizamos a edição brasileira de 1997, cf. referência bibliográfica.

É claro que pode haver estudos e interpretações dos temas desenvolvidos por Bakhtin mais aprofundados e mais rigorosos, mas, neste trabalho, utilizamos o texto **Os gêneros do discurso** por considerá-lo bastante esclarecedor da idéia desenvolvida por esse autor e por sugerir aportes condizentes com a pesquisa que realizamos. Precisamos também esclarecer que existem estudos que propõem fazer distinção entre **gêneros do discurso** e **gêneros textuais**, porém seguimos uma proposta semelhante à de Marcuschi (2002), citada na seção 1.2, utilizando a discussão sobre gêneros do discurso, em Bakhtin, e transpondo-a para os gêneros textuais, sem maiores distinções.

O livro **A Estética da Criação Verbal**, no qual consta o ensaio O Problema dos Gêneros do Discurso, acima mencionado, é muito citado e referido na literatura. Segundo o que escreveu Todorov em sua introdução, esse texto data do último período “sintético” da vida de Bakhtin e seria o início ou o plano de um livro consagrado aos gêneros do discurso, uma espécie de síntese das reflexões lingüísticas de Bakhtin nos anos 20. Segundo Ciapuscio (1994, p. 14), Bakhtin

“realizou um dos maiores aportes ao estudo da temática ao introduzir o conceito de gênero discursivo enquanto conjunto de enunciados relativamente estável ligado a uma esfera social determinada”.

No seu texto, Bakhtin (1997, p. 280) revela o caráter inovador de sua reflexão, mencionando que o problema geral dos gêneros do discurso **nunca** havia sido antes colocado, o que seria compreensível, devido à diversidade e à heterogeneidade dos enunciados que compõem os gêneros, assim como das esferas da atividade humana. Historicamente parece ter sido muito difícil estabelecer um traço comum a todos os gêneros, e Bakhtin vem trazer a idéia de que esse traço seria a natureza lingüística dos enunciados, uma característica única compartilhada por todos eles. Os enunciados (orais ou escritos) seriam a realização da língua, sendo usados em toda e qualquer esfera das inúmeras atividades humanas.

O enunciado reflete as condições específicas e as finalidades de cada uma dessas esferas, não só por seu conteúdo (temático) e por seu estilo verbal, ou seja, pela seleção operada nos recursos da língua – recursos lexicais, fraseológicos e gramaticais –, mas também, e sobretudo, por sua construção composicional.” (id. p. 279)

Segundo Marcuschi (2002), a construção composicional seria a forma de dizer o conteúdo temático, sua organização geral, que não é inventada a cada vez que nos comunicamos, mas que está disponível para uso social. Podemos dizer que esses traços estruturais, juntamente com o conteúdo e o estilo, permitem ao usuário reconhecer um gênero, nomeá-lo e empregá-lo como algo que é parte do uso, da linguagem típica de sua cultura. Quando menciona as “formas relativamente estáveis” dos gêneros, Marcuschi, baseado em Bakhtin, explica que os gêneros constituem ações verbais que se tornam convencionalizadas “(...) em virtude da recorrência das situações em que são investidas como ações retóricas típicas” (Marcuschi, 2002, p. 32), ou seja, suas construções composicionais estabilizam-se pelo uso constante nas práticas verbais de um determinado grupo cultural.

Essa constatação é muito valiosa para o nosso trabalho, uma vez que confirma teoricamente o pressuposto que queremos demonstrar na prática: que existem expressões textuais e discursivas no artigo científico que são frequentes e sistemáticas e, mais ainda, como estamos tratando de tradução, que uma vez inseridas em grupos culturais distintos (seja pela língua, seja pela esfera de atividade) tenderão a ter seu uso e seu emprego diferenciados, como demonstra a análise comparativa dos marcadores textuais apresentada no Capítulo 6.

Cada esfera de atividade, então, desenvolve tipos relativamente estáveis de enunciados que passam a ser comumente associados a elas. São esses tipos diferentes e estáveis de enunciados que Bakhtin chama de gêneros de discurso (Bakhtin 1997, p. 301), e essa idéia da relativa estabilidade dos gêneros é, até hoje, retomada pelos teóricos e analistas de gênero.

Outro ponto observado por Bakhtin e que nos interessa nessa pesquisa concerne à estilística. Ele considera que o estilo está ligado ao enunciado e a suas formas típicas (os

gêneros). Uma vez que o enunciado é individual, ele pode refletir a individualidade, sendo que alguns gêneros são mais propícios para tal do que outros. Porém é o estilo geral, ou seja, o lingüístico, que está indissociavelmente ligado ao gênero,

“o estilo lingüístico ou funcional nada mais é senão o estilo de um gênero peculiar a uma dada esfera da atividade e da comunicação humana. Cada esfera conhece seus gêneros, apropriados à sua especificidade, aos quais correspondem determinados estilos” (Bakhtin, 1997, p. 284).

Voltando novamente essa discussão especificamente para nossa pesquisa, podemos dizer que as expressões que estudamos fazem parte da especificidade do gênero discursivo **artigo científico** e do estilo que essa esfera de atuação (em nosso caso, em primeiro lugar a grande comunidade científica internacional) condiciona ou a qual os autores estão condicionados. Bakhtin (id. p. 284) afirma que

“o estilo é indissociavelmente vinculado a unidades temáticas determinadas e, o que é particularmente importante, a unidades composicionais: tipo de estruturação e conclusão de um todo, tipo de relação entre o locutor e os outros parceiros da comunicação verbal.”

Essa afirmação contempla alguns casos de unidades que estamos estudando nesse trabalho. Mas é importante ressaltar que a noção de gêneros do discurso nesse texto de Bakhtin se presta a uma reflexão muito mais ampla do que a subcategorização dos gêneros e as questões de estilo. Há, no texto **Os Gêneros do Discurso**, toda uma teoria da linguagem, pois, baseado nas noções de enunciado e gênero, Bakhtin discorre e reflete sobre as unidades da língua como sistema: as palavras e as orações. Para ele, o estudo do enunciado, que é a unidade real de comunicação, deve permitir compreender melhor a natureza das unidades da língua, que são as palavras e as orações.

As reflexões inovadoras de Bakhtin trouxeram uma luz diferente à reflexão sobre gêneros, principalmente nas noções de esfera de atividade e relativa estabilidade dos gêneros,

e, claro, da consideração de um **outro** como fundamental na atividade de comunicação (atitude responsiva). Diz ele, no rastro de Aristóteles, “Quando se subestima a relação do locutor com **o outro** e com seus enunciados (existentes ou presumidos), não se pode compreender nem o gênero nem o estilo de um discurso” (p. 324).

1.2.3 Swales e a noção de comunidade discursiva

O nome de John M. Swales é um dos de maior destaque na tradição norte-americana quando se trata de estudos sobre gêneros. A publicação do livro **Genre Analysis. English in Academic and Research Settings** já tem mais de 10 anos, mas seu trabalho sobre gêneros, e especificamente sobre o artigo científico, parece permanecer como uma base firme e confiável para o desenvolvimento de estudos mais contemporâneos, não só de raízes norte-americanas, ainda que com algumas ressalvas quanto ao caráter generalizador de seu modelo.

A partir das noções de gênero em quatro áreas distintas – o folclore, a literatura, a lingüística e a retórica –, esse autor faz uma nova definição de **gênero** que ele deseja, pouco ambiciosamente, que seja aplicada apenas aos propósitos que tem em mente, que são, principalmente, o ensino e o aprendizado de língua estrangeira.

Das quatro áreas que avalia para formular seu conceito de gênero, Swales resume seis aspectos comuns e presentes em todas elas. Menciona que em todas há (i) uma classificação duvidosa dos gêneros e de um prescritivismo prematuro e pouco cuidadoso; (ii) uma noção de que os gêneros são importantes na integração de passado e presente; (iii) um reconhecimento de que gêneros estão situados dentro de comunidades discursivas, em que crenças e práticas de denominação dos participantes têm relevância; (iv) uma ênfase no propósito comunicativo e na ação social; (v) um interesse em uma estrutura genérica (e no qual o artigo está baseado); (vi) uma compreensão da dupla capacidade gerativa de gêneros – de estabelecer objetivos retóricos e de incentivar sua realização.

Swales faz, então, uma caracterização dos gêneros e, metodologicamente, divide seus comentários em seis tópicos, os quais descrevemos, resumindo os principais pontos, a seguir.

(i) Um gênero é uma classe de eventos comunicativos.

O autor toma um evento comunicativo como aquele no qual a língua tem um papel significativo e indispensável. Os eventos comunicativos de uma determinada classe podem ter uma frequência maior ou menor, sendo que os de menor frequência tem que ter uma certa importância, para uma determinada cultura, para existirem como um gênero. Um evento comunicativo para Swales compreende não somente o próprio discurso e seus participantes, “mas também o papel do discurso e o ambiente de sua produção e recepção, incluindo suas associações culturais e históricas” (Swales, 1990, p. 46).

(ii) O principal critério que transforma um grupo de eventos comunicativos em um gênero particular é a existência de propósitos comunicativos em comum.

Essa é uma visão que se aproxima daquela dos estudos de gênero chamada New Rhetoric, fundamentada principalmente na obra **Genre as a Social Action** (1984), de Carolyn Miller. A ênfase desses estudos está em saber muito mais o papel social dos gêneros, do que a sua forma, sua linguagem típica e sua organização interna. Swales coloca que o primeiro e determinante fator de pertinência a um gênero está mais no compartilhamento do mesmo propósito do que em similaridades de forma ou algum outro critério. Porém ele mesmo ressalta que pode não ser tão simples assim identificar os propósitos de um gênero, e o pesquisador deve ter cuidado ao simplificar

as coisas, como ao dizer que artigos de pesquisa são simples relatos de experiências. Também não é incomum um gênero ter mais de um propósito comunicativo.

(iii) Os exemplares de gêneros variam em sua prototipicidade.

Nesse ponto, Swales menciona que outros aspectos determinariam um gênero além do propósito comunicativo, que outras características seriam necessárias para determinar um gênero. Diz, então, que há duas maneiras de ver essa questão: a da abordagem da definição e a da semelhança por família.

Swales utiliza um dos pontos de vista discutidos para a problemática da definição (que é questão muito cara aos estudos de terminologia e de lexicologia e às áreas afins aplicadas) para empregá-lo na definição dos diferentes tipos de gêneros. A proposta citada é a que ficou mais conhecida pelos estudos de Katz and Fodor (1963), em que um conjunto de propriedades, necessárias e suficientes para diferenciar e identificar todos os membros e somente os membros de uma determinada categoria os definem. Ele coloca, porém, que, na prática, poderia ser muito difícil listar tais características para determinadas categorias, como a das frutas ou dos móveis, e, assim, seria para as categorias de gêneros e que, talvez, o que permitiria pertencer a uma mesma categoria não seria uma lista de características definidoras, mas inter-relações de um tipo mais maleável, as que Wittgenstein chamou de famílias. No entanto, de acordo com Swales “(...) pode-se dizer que uma teoria de semelhanças de família possa fazer qualquer coisa se parecer a qualquer coisa” (1990. p. 49), porém a noção de família para Wittgenstein é de que as famílias têm uma coerência por outras razões além das características físicas, como os laços de sangue e as experiências compartilhadas. Essas discussões deram origem a outras, que incluem também a noção de protótipo, para que pudessem ser reconhecidas as instâncias de categorias.

Para finalizar, Swales conclui que

“(...) o propósito comunicativo foi nomeado como a propriedade *privilegiada* de um gênero. Outras propriedades, como forma, estrutura e audiência, operam para identificar até que ponto um exemplar é *prototípico* de um determinado gênero” (1990, p. 52)

(iv) A lógica subjacente a um gênero estabelece restrições a possíveis contribuições em termos de conteúdo, posicionamento e forma.

Swales afirma que “(...) membros estabelecidos de comunidades discursivas empregam gêneros para perceber comunicativamente as metas de suas comunidades” (id. p. 52). A idéia é a de que o reconhecimento dos propósitos fornece as bases, enquanto essas permitem convenções que restringem a participação de membros aprendizes ou não-membros, que as conhecem menos.

(v) A nomenclatura usada para os gêneros por uma comunidade discursiva é importante fonte de insight

Neste item, o autor chama atenção para os nomes dados aos gêneros, ressaltando que podem ser fontes para conclusões importantes. Diz que, em geral, os membros das comunidades discursivas denominam gêneros e classes de eventos comunicativos que reconhecem como tendo formas retóricas recorrentes. Algumas dessas refletem o propósito comunicativo do gênero, outras podem ser marcas de certas instituições, ao invés de serem apenas descritivas. Às vezes podem existir gêneros sem eventos comunicativos vinculados a eles, assim como alguns eventos comunicativos que não são denominados por gênero algum.

Após ter feito essa categorização, Swales propôs uma definição de gênero que ele mesmo diz talvez não ser totalmente adequada, mas que se beneficiou de discussões

anteriores a respeito do termo e de outras áreas afins, e que ela representa um avanço nas próprias formulações anteriores que havia feito:

“Um gênero compreende uma classe de eventos comunicativos, cujos membros compartilham um conjunto de propósitos comunicativos. Esses propósitos são reconhecidos pelos membros da comunidade discursiva que trabalha com eles e, portanto constituem a lógica subjacente aos gêneros. Essa lógica molda a estrutura esquemática do discurso e influencia e restringe a escolha do conteúdo e o estilo. Além do propósito, os exemplares de um gênero exibem vários padrões de similaridade em termos de estrutura, estilo, conteúdo e público alvo” (Swales 1990, p. 58).

Mas para entender esse conceito de gênero de Swales, é fundamental que se compreenda também o que ele define por comunidade discursiva. A visão que ele segue é a mesma estabelecida por Herzberg:

“O uso do termo ‘comunidade discursiva’ testemunha a idéia geral que vêm se expandindo de que o discurso opera dentro de convenções definidas por comunidades, sejam disciplinas acadêmicas ou grupos sociais” (Herzberg, 1986, apud Swales, 1990, p. 21).

Swales menciona também que, embora se considere a idéia de comunidade discursiva muito importante, ainda há muitos pontos que não estão bem claros, sugerindo que ela levanta mais questões do que consegue responder. Por esse motivo, ele estabelece um conjunto de critérios para eliminar aspectos “confusos e controversos” presentes na definição de comunidade discursiva. Além da distinção entre comunidades de fala (*speech communities*) e comunidades discursivas (*discourse communities*), ele também apresenta seis características para a identificação de um grupo de indivíduos como uma comunidade discursiva.

Em um artigo posterior a publicação do livro, Swales (1992) suspeita que foi muito seduzido pelo conceito de comunidade discursiva e afirma que a verdadeira comunidade discursiva pode ser mais rara e esotérica do que pensava. Assim, faz então uma redefinição dos seis critérios, afirmando que os anteriores não resistiram ao teste do tempo, pois ainda

manifestam um caráter reducionista, utópico e estático. Apresentamos essas características em comparação no Quadro 1, apontando quais foram as mudanças dessa “redefinição” em cada uma delas.

Neste Quadro, percebe-se que o conceito foi ampliado, abrangendo mais elementos e ficando mais flexível, no sentido de que assim há uma tentativa de “(...) representar um mundo mais complexo e um tanto obscuro”. (Swales, 1992).

Quadro 1.1 - Propriedades das comunidades discursivas em dois momentos da obra de Swales – 1990 e 1992

SWALES (1990)	SWALES (1992)
Uma comunidade discursiva tem um conjunto de objetivos públicos comuns;	A comunidade discursiva não apenas aceita os objetivos, mas também os formula ou estabelece. Esses objetivos podem ser consensuais, mas também podem ser distintos e relacionarem-se.
Uma comunidade discursiva tem mecanismos de intercomunicação entre seus membros;	Quanto aos mecanismos de intercomunicação entre os membros de uma comunidade discursiva, não houve alterações. Segundo Swales (1990), esses mecanismos variam de acordo com a comunidade.
Uma comunidade discursiva utiliza seus mecanismos de participação principalmente para fornecer informações e <i>feedback</i> ;	Nesse ponto, Swales acrescenta que uma comunidade discursiva usa mecanismos de participação para uma série de propósitos, não apenas para informação e <i>feedback</i> .

Uma comunidade discursiva utiliza e, portanto, tem um ou mais gêneros para alcançar comunicativamente seus objetivos;	Ao invés de um ou mais gêneros para alcançar seus objetivos, uma comunidade discursiva utiliza uma seleção crescente de gêneros no alcance dos mesmos.
Além dos gêneros, uma comunidade discursiva adquire um vocabulário específico;	Uma comunidade discursiva adquire e continua sempre buscando uma terminologia específica.
Uma comunidade discursiva conta com uma estrutura de membros que detém um certo nível de conteúdo relevante e perícia discursiva.	Swales acrescenta que uma comunidade discursiva tem uma estrutura hierárquica explícita ou implícita que orienta os processos de admissão e de progresso dentro dela.

1.2.4 Aproximando Swales e Bakhtin

Após olharmos com mais atenção para essa discussão sobre comunidade discursiva, interessa-nos fazer uma aproximação entre o que Bakhtin chamou de “esferas de atividade” e o que Swales chamou de “comunidade discursiva”. Se tomarmos os seguintes extratos representativos das afirmações de ambos:

“Todas as esferas da atividade humana, por mais variadas que sejam, estão sempre relacionadas com a utilização da língua. O enunciado reflete as condições de cada uma dessas esferas, não só por seu conteúdo (temático) e por seu estilo verbal mas também, e sobretudo, por sua construção composicional. Esses três elementos fundem-se indissolavelmente no *todo* do enunciado, e todos eles são marcados pela especificidade de uma esfera de comunicação” (Bakhtin, 1997, p. 279)

“O discurso opera dentro de convenções definidas por comunidades, sejam disciplinas acadêmicas ou grupos sociais. O uso da língua em um grupo é uma forma de comportamento social, o discurso é um meio de manter e estender o conhecimento do grupo e de introduzir novos membros no grupo, o discurso é epistêmico ou constitutivo do conhecimento do grupo” (Kerzberg, 1986, p. 1 apud Swales, 1990, p. 21),

veremos que eles têm pontos que podem ser discutidos tomando-se e aproximando as duas teorias no que nos interessa para a validação desta pesquisa. Vamos lembrar também que o próprio Swales (1992) já admitiu que se sente “completa e dolorosamente” culpado por não ter dado mais atenção a Bakhtin, de quem conhecia apenas de segunda mão o trabalho sobre dialogismo.

As esferas de atividade em Bakhtin são todos os eventos nos quais o homem participa e, segundo ele, todas estão relacionadas ao uso da língua. As comunidades discursivas, de Swales, são mais restritas, tanto que ele exclui de seu foco duas áreas da atividade verbal: a conversa casual e a narrativa simples, considerando-as pré-gêneros. Contrapondo as idéias de Bakhtin e de Swales, percebemos que se distanciam nesse ponto, pois, para Bakhtin, toda utilização da língua efetua-se em forma de enunciados, e cada esfera de utilização desses enunciados elabora tipos relativamente estáveis deles, ocasionando, então, que toda utilização da língua se dará por meio de algum gênero; para Swales, um gênero é uma classe de eventos comunicativos, cujos membros compartilham o mesmo conjunto de propósitos comunicativos. Na definição de Swales, esses propósitos estariam condicionados por comunidades mais concretas, em um sentido mais relacionado a uma associação formal, com mecanismos de comunicação, membros associados e regras de aceitação (ver Quadro 1). Mas, como ele mesmo diz, sua teoria é desenvolvida com base nos seus objetivos, que, em geral, são o ensino e a aprendizagem dos gêneros, diferentemente de Bakhtin, que faz uma reflexão mais ampla e filosófica sobre a linguagem.

Embora estejamos falando de diferenças, parece haver por trás dessas reflexões um mesmo fio condutor, ou seja, uma tentativa de descrever o acontecimento de um mesmo evento da linguagem: o gênero e suas implicações. O que parece resultar dessa aproximação (Swales-Bakhtin) é que a reflexão de Bakhtin acaba abrigando a de Swales. Ambos tratam de uma mesma coisa, ou seja, de como o discurso, ou o enunciado, se estabelece dentro de esferas ou comunidades, que o determinam e o fazem ser o que é e como é, resultando, assim, em diferentes formas de comunicação. O ponto é que Swales está com outro tipo de discussão em mente e acaba por focar, digamos assim, apenas esferas de comunicação (usando o termo como em Bakhtin) bastante marcadas e caracterizadas pelos seis aspectos já mencionados, e, embora ele não mencione, parece estar muito mais calcado no texto escrito. Já Bakhtin dá conta de classificar em gêneros todos os tipos enunciados, sejam eles orais, sejam escritos.

Podemos dizer, ainda, dessa aproximação que, para Swales, o gênero se dá por afinidades de propósitos comunicativos. Estes forneceriam as bases para a constituição de um gênero, enquanto, para Bakhtin, o critério de agrupamento é a própria atividade discursiva, os propósitos comunicativos parecem estar implícitos na idéia de pertinência de um enunciado a um gênero: “A variedade dos gêneros do discurso pressupõe a variedade dos escopos intencionais daquele que fala ou escreve” (Bakhtin, 1997 p. 291).

Assim, o que tentamos esclarecer é que sempre um gênero acontece dentro de uma esfera, ou de uma comunidade, que irá determinar escolhas e estilos, e que irá, no evento da comunicação, pressupor conhecimentos e buscar, digamos, uma cumplicidade dos participantes. Os gêneros também não são eventos fixos, isolados e que se excluem mutuamente, sendo que podemos agrupá-los a partir do estabelecimento de diferentes critérios, mesmo que alguns de seus elementos, considerados inerentes à sua constituição formal, não estejam presentes.

As duas propostas que acompanhamos até aqui seguem, mais ou menos, a mesma direção. Com isso, queremos dizer que os autores partem mais do ponto de vista da produção de uma determinada forma de enunciado, ou de um gênero, do que do ponto de vista da recepção. Não queremos dizer, de forma alguma, que o receptor é desconsiderado, muito pelo contrário, em Bakhtin é muito bem salientado que o **outro** é sempre levado em conta, assim como, em Swales, a audiência também é levada em conta. Mas cabe trazer para esse mesmo âmbito de discussão o conceito de **comunidade interpretativa**, desenvolvido por Stanley Fish (1980) e mencionado por Swales (1990) como uma das fontes para o conceito de comunidade discursiva.

Na perspectiva desenvolvida por Fish (1980, p. 318), “(...) todo o objeto só se constitui enquanto tal a partir de determinada interpretação”. Porém essa interpretação está condicionada social e culturalmente, afastando assim a idéia de que poderíamos ter tantas interpretações quantos fossem os sujeitos do universo, o que seria um caos inimaginável. Fish propõe que todos nós aprendemos a adquirir um conhecimento relativo às comunidades em que nos inserimos ou as quais podemos pertencer e que esse conhecimento é, às vezes, tão óbvio que não nos damos conta de que o adquirimos. Segundo Fish (1980, p. 318)

“(...) a comunicação ocorre dentro de situações e estar em uma situação já é estar de posse de (ou ser possuído por) uma estrutura de suposições, de práticas entendidas como relevantes em relação aos propósitos e metas que já estão colocados, e é de dentro desses propósitos e metas que qualquer enunciado é imediatamente ouvido”.

Toda essa idéia de condicionamento cultural e social, ou de pertinência a um grupo ou a uma situação, é comum e está presente nas três abordagens – Bakhtin, Swales e Fish – que apresentamos. Essa idéia é fundamental para o embasamento e para a justificativa do trabalho com expressões típicas do artigo científico que apresentamos aqui, pois estamos defendendo que os textos de nosso *corpus* foram produzidos dentro de e para uma comunidade, não apenas a de cada língua, mas também a da grande comunidade científica internacional. Por

isso, consideramos importante que correspondam ao que essas comunidades esperam deles, tanto em português como em inglês, enquanto realizações de um gênero específico.

Antes de passarmos para a seção 1.3, a seguir, e no que concerne à importância do estudo de um gênero para a tradução, gostaríamos de acrescentar o comentário de Hurtado Albir (2001, p. 505) a esse respeito:

“Consideramos que a descrição e a classificação dos gêneros é fundamental para o desenvolvimento dos estudos descritivos em Tradutologia, já que permitem conhecer melhor as modalidades e tipos de tradução, mas também para a didática da tradução, por suas aplicações pedagógicas. Daí a importância de seguir pesquisando para identificar mais gêneros em outros âmbitos específicos, para conhecer melhor os padrões de gêneros e sub-gêneros nos já identificados (técnicos, jurídicos, audiovisuais, etc.) e para contrastar seu funcionamento em línguas diferentes”.

Com esta dissertação, queremos contribuir com estes estudos de gênero voltados à tradução. Em nosso recorte, escolhemos o artigo científico, entendendo que, embora o mesmo já seja identificado como um gênero estável, ainda é necessário, para a tradução, identificar, no nível textual, as estruturas que o compõem e o fazem acontecer como um gênero, ainda que em culturas de línguas distintas. Algumas dessas estruturas, denominadas aqui marcadores textuais, foram o foco deste estudo, e a motivação principal foi identificar semelhanças e diferenças das mesmas nas culturas das línguas inglesa e portuguesa, visando ao auxílio na realização de traduções para que se encaixem nos padrões esperados das culturas em questão.

A seção a seguir traz, à luz da reflexão sobre gêneros feita até aqui, uma abordagem do texto de especialidade, pois entendemos que, sendo este o texto que compõe o artigo científico, merece também ser abordado neste trabalho, ainda que sucintamente e sem maiores pretensões. Como, recentemente, a Terminologia, que estuda elementos do texto de especialidade, como termos e fraseologias especializadas, também se voltou para o ambiente textual, trouxemos, para o âmbito desta pesquisa, alguns pontos em que, sob nossa ótica, pode

haver colaboração entre o estudo que desenvolvemos até aqui e a apreensão do todo que é o texto de especialidade, visando complementar as bases teóricas deste estudo.

1.3 O texto de especialidade e gênero

Entendemos que denominações como gênero, e texto especializado provêm de áreas de estudos que se desenvolveram independentemente e trilharam vias distintas para atingir seus objetivos, porque cada uma encerra uma gama muito grande de considerações. Resumidamente, podemos dizer que a principal diferença, no que concerne à abordagem de uma e outra área é que, no estudo do texto especializado, entra em questão a área em que o texto se realiza, e, no estudo de gênero, não. No entanto, percebemos que algumas questões discutidas para o estudo do texto de especialidade contribuem com o estudo de um gênero, e o contrário também pode acontecer. Por essa razão, o que pretendemos nesta seção é aproximá-los, no que têm de comum e no que podem se complementar.

O estudo do texto especializado situa-se no âmbito da lingüística das linguagens especializadas. Essa denominação, linguagens especializadas, pode dar margem a um entendimento de que as mesmas são subsistemas da linguagem, independentes do sistema da língua geral, como consideram alguns autores (Rondeau, 1983). Cabré (1999), no entanto, considera abusivo o uso do termo **linguagem** e propõe a denominação “comunicação especializada”, considerando, assim, que, dentro do sistema geral de língua, subsistem registros funcionais, empregados para realizar determinadas comunicações. Compartilhamos esse ponto de vista, porém preferimos o uso da expressão **linguagem de especialidade**, já consagrada com o sentido que dela queremos apreender, isto é, não como um sub-sistema, mas como o mesmo sistema de língua geral em uma realização de comunicação específica.

Conforme esclarece Maciel (2001, p. 87)

“(...) todos os usuários da língua, leigos ou especialistas, têm à disposição o mesmo repertório fonológico, morfológico ou sintático. Cada um faz a escolha que melhor serve a seus propósitos nas circunstâncias em que se realiza a comunicação, ainda que utilize o mesmo manancial lingüístico”.

Sendo assim, conforme pontua Azenha (1999, p. 66), ainda é reduzido,

“(...) no Brasil, o número de estudos descritivos do léxico e da sintaxe das linguagens técnicas que permitam alguma objetividade na determinação da frequência e da quantificação de termos e estruturas, critérios a partir dos quais se poderia estabelecer, em tese, uma diferenciação das linguagens técnicas entre si, e dessas para a linguagem comum.”

Com essa afirmação entendemos, não o contrário do que afirmamos acima, mas que, essas as escolhas que o usuário das linguagens técnicas ou científicas faz acabam revelando tendências e similaridades, determinadas, inclusive, pelo que defendemos na primeira parte desse capítulo, as esferas de atividade e as comunidades, que por sua vez determinam os gêneros. Essas similaridades é que constituem as linguagens técnicas e científicas e que as fazem constituírem gêneros. Acreditamos que este trabalho está dando um pequeno passo em direção a esses estudos, mencionados por Azenha, de frequência e quantificação de termos e estruturas dessas linguagens.

Sobre o caráter especializado de uma linguagem, podemos dizer que o mesmo não será conferido pelo tema focado, mas pela maneira como este será tratado na comunicação. Como observa Cabré (1999, p. 154), um conteúdo tradicionalmente científico, por exemplo, apresentado de maneira banal não se constitui em matéria de comunicação especializada, e, de maneira inversa, qualquer conteúdo, quando descrito de maneira precisa, com referência a um esquema científico ou a uma estrutura normativa, torna-se objeto de conhecimento especializado e, portanto, de comunicação em linguagem de especialidade. Chamamos

atenção aqui para o fato de estarem envolvidos, na linguagem de especialidade, elementos relacionados ao modo de dizer de certo conteúdo, sendo assim, abre-se a perspectiva para outras realizações do texto de especialidade que não somente de conteúdo, ou de área, mas de tipo de texto ou de gênero textual.

Hoffmann (2004), cujo estudo é pioneiro em aproximar a linguagem de especialidade da língua geral, esclarece que

“(...) a especificidade das linguagens especializadas, em relação à linguagem comum e a outras sublinguagens, se expressa mais claramente pelo léxico, quer dizer, pelo vocabulário especializado ou pela terminologia, mas também pelo uso de determinadas categorias gramaticais, de construções sintáticas e de outras estruturas textuais”,

e isto aponta para o que mencionamos no parágrafo anterior a respeito do que faz um tema se realizar de maneira especializada ou não. De certa forma, certos gêneros, dadas suas características, colaboram e se tornam até mais propícios para a realização de uma comunicação especializada.

Definido o que entendemos por linguagem de especialidade, chegamos ao texto especializado, que é a linguagem de especialidade na situação real de comunicação (Maciel, 2001). De acordo com Hoffmann (2004), o texto especializado “(...) é instrumento e ao mesmo tempo resultado da atividade comunicativa exercida em relação a uma atividade especializada” Esse texto se realiza de maneiras específicas, determinadas pelas suas funções e pelo seu âmbito (Möhn e Pelka, 1984 apud Hoffmann, 1994), ou, acrescentaríamos, pelos critérios textuais e discursivos, e pelos critérios temáticos. Portanto, um texto especializado não se define apenas pelo critério **tema** (área científica ou técnica), mas por “(...) particularidades de sua macroestrutura (articulação), por relações de coerência entre seus elementos e pela utilização de unidades sintáticas, lexicais, morfológicas e gráfico-fonéticas”. As preferências para o uso dessas particularidades se condicionam às circunstâncias da interação verbal dos atores do evento comunicativo, dependendo, portanto, das peculiaridades

do destinador e do destinatário (Maciel, 2001, p. 93). Como se vê, muito do que discutimos sobre gênero se sobrepõe a esses comentários.

Se tomamos as definições de gênero vistas até aqui,

- a) a de Bakhtin, para quem cada esfera de atividade desenvolve tipos relativamente estáveis de enunciados – os gêneros –,
- b) a de Swales, para quem um gênero é uma classe de eventos comunicativos cujos membros compartilham um conjunto de propósitos comunicativos

e a definição de texto especializado de Hoffmann (2004), segundo o qual

- c) “(...) o texto especializado é instrumento e, ao mesmo tempo, resultado da atividade comunicativa exercida em relação a uma atividade especializada sócio-produtiva”,

perceberemos que elas têm pontos em comum, principalmente no que concerne ao gênero ser determinado por uma atividade específica, envolvendo membros em sociedade e que compartilham atividades comunicativas.

Dessa forma, por exemplo, a escolha que fazemos dos itens que irão servir aos nossos propósitos nas circunstâncias em que se realiza a comunicação é determinada pela nossa “esfera de atividade” e pela comunidade da qual estamos falando. Obviamente, surgem também as questões de estilo, mas, como propõe Bakhtin, também o estilo estará condicionado ao meio em que se realiza o texto.

O que precisamos deixar claro é que o enfoque do texto por parte dessas duas áreas – estudos de gênero e estudos de texto especializado – são diferentes. Quando falamos em um gênero estamos falando do geral, de um agrupamento de eventos comunicativos (dentre esses

o texto) que se assemelham. Por outro lado, um texto especializado é o particular, é a matéria que compõe os textos que pertencem a um gênero.

Neste trabalho, ao tratarmos do **gênero artigo científico** estaremos falando inevitavelmente de sua materialização, o **texto especializado**, uma vez que os textos de artigos científicos, em geral, são resultado de um exercício de observação ou de uma atividade científica ou técnica em uma dada área. Não estaremos, porém, falando das particularidades das áreas envolvidas em nossos *corpora*, que, certamente, se fossem observadas, mostrariam ter implicações na realização dos gêneros.

Já mencionamos anteriormente que o reconhecimento dos gêneros, por parte dos tradutores, é essencial para que se conheçam melhor os padrões de gêneros e subgêneros, no intuito de entender seu funcionamento nas diferentes línguas e produzir textos condizentes com esses padrões nas línguas envolvidas. O estudo do texto de especialidade, assim como da terminologia que o compõe, também pode trazer resultados importantes para o estudo da tradução de textos especializados, principalmente quando ambos, texto especializado e terminologia, forem abordados através de uma perspectiva ampla, de inclusão e descritiva.

Cabré (1999) sustenta que o tradutor de textos especializados deve conhecer alguns parâmetros para que sua tradução consiga atingir os índices mínimos de qualidade, ou seja, além de ser verídica, do ponto de vista do conteúdo, e correta, do ponto de vista gramatical, deve ser adequada e soar natural. Esses parâmetros estão relacionados com três grupos de condições que definem operativamente a língua e o texto de especialidade: um primeiro grupo de caráter cognitivo, um segundo grupo de caráter gramatical e um terceiro de caráter pragmático, que descrevemos brevemente a seguir, com o intuito de localizar também onde esta pesquisa se situa.

De acordo com as condições de caráter cognitivo, um tema só é especializado se transmite um conhecimento estruturado dentro de uma área de especialidade, isto é, se é

codificado com referência a um esquema preestabelecido por cada uma das matérias ou escolas científicas. Como consequência, o fator que faz um tema ser geral ou especializado não é o conteúdo em si, mas a maneira como é conceitualizado um determinado modo de significação.

Gramaticalmente, um texto especializado apresenta especificidades em dois planos: o léxico e o textual. No plano do léxico, é uma característica marcante o uso de uma terminologia específica; do ponto de vista textual, o texto especializado caracteriza-se, entre outros, pelo caráter restritivo de suas estruturas e pela sistematicidade na apresentação da informação. Nesse ponto, um estudo de gênero pode colaborar com a identificação dessas estruturas sistemáticas, como é o caso do presente trabalho, que demonstrará algumas dessas sistematicidades no texto do artigo científico, em particular, as unidades marcadoras do desenvolvimento do texto; ainda que não entre na questão da especificidade da área, pois consideramos que esse seria um outro tipo de investigação, talvez futura.

Por fim, o terceiro grupo, o do caráter pragmático. Cabré (1999) considera que o texto especializado se caracteriza pelos elementos que intervêm em seu processo de produção e recepção. Assim, se podemos dizer que o emissor de um texto de especialidade sempre é um especialista que interiorizou certos conhecimentos, os destinatários poderão ser: outros especialistas, aprendizes e o grupo daqueles que recebe a informação em forma de discurso de divulgação, simplesmente para aumentar sua gama de conhecimentos sobre a matéria. Essas afirmações devem ser, no entanto, um pouco relativizadas, pois não há um grau absoluto nessas categorias de nível de especialidade.

Retomando, entendemos que o tradutor, como propõe Cabré, ao trabalhar com um texto de especialidade, precisará estar atento a essas três características do texto: a conceitualização e suas implicações nos diferentes sistemas lingüísticos (da língua de partida e da língua de chegada), o léxico e a gramática especialmente próprios de cada gênero textual

e os níveis de produção e de recepção de um texto. Acreditamos que, ao colocarmos em evidência algumas expressões típicas do gênero artigo científico, estaremos contribuindo para uma melhor compreensão do texto que compõe esse gênero e chamando atenção para seu estudo em tradução, o que nos parece contribuir muito para a realização de traduções adequadas e naturais. Acreditamos também que, para que isso aconteça, é necessário, cada vez mais, partir das observações teóricas e classificatórias, para as questões práticas e de ocorrências em textos reais; assim, tradutores e aprendizes de tradução poderão contar mais efetivamente com pesquisas e material de apoio em suas traduções de textos especializados.

No que tange aos estudos da Terminologia, conforme descrito por Hoffmann (2004), no início dos estudos da lingüística das linguagens especializadas, a atenção concentrava-se quase exclusivamente sobre o vocabulário especializado e sobre a terminologia. Da década de 80 para cá, o texto de especialidade como um todo já figura como o objeto central dos estudos em linguagens de especialidade. Processo parecido tem acontecido com a Terminologia, que, a partir dos primeiros postulados estabelecidos por Wüster, nos anos 30 – mais prescritivos e exclusivamente voltados à uma apreensão normalizadora dos termos técnico-científicos – tem buscado, no texto, complementos e suporte para o estudo de seu objeto, que também se ampliou, incluindo, além do termo, a fraseologia e o texto das definições (Krieger, 2001a). Resumidamente, podemos dizer então, que a terminologia, isto é, o elenco de termos de diversos domínios, continua sendo parte integrante do texto especializado, ao lado de outros elementos lingüísticos; e que a Terminologia, isto é, o campo de estudos teóricos e metodológicos que tem por objetos o termo, a fraseologia e a definição, se abre cada vez mais para uma abordagem do texto de especialidade, em seu desenvolvimento e no de seus produtos. Conforme esclarece Krieger (2001a, p. 36):

“(...) elementos da Lingüística Textual, da Semiótica Narrativa greimasiana, ao lado do enfoque enunciativo de discursos e de componentes pragmáticos da comunicação especializada, têm se constituído em aportes que permitem identificar melhor a feição e o

funcionamento das linguagens especializadas. Essa relação, entretanto, não se restringe à dimensão temática, vinculando-se também a componentes da ordem da textualidade e da discursividade de que cada discurso especializado se vale para se instituir.”

Acreditamos que a Terminologia só pode ser de utilidade para a tradução se realmente tratar seus objetos como pertencentes a um ambiente comunicativo e que se relacionam com outros elementos do texto, não estando isolados e fechados em seus conceitos. A tradução precisa de uma Terminologia que considere o funcionamento da linguagem com toda a sua complexidade e que, assim, possa descrever, em melhores condições, “(...) a especificidade dos contextos discursivos nos quais os termos se manifestam em toda a sua diversidade de configurações” (Krieger 2001a, p. 37). Conforme esclarece ainda Krieger (id., p. 37), “(...) para poder dar conta dessa diversidade, os estudos de Terminologia não podem ficar restritos aos limites da frase, nem à formalidade da análise de seus componentes morfossintáticos”, é por isso que acreditamos que, uma vez preocupada em demonstrar marcadores textuais típicos de artigos científicos, esta dissertação pode também contribuir para a maneira como a Terminologia olha para os seus objetos, uma vez que nosso enfoque são exatamente os elementos do texto e não os de conteúdo.

Um pouco em oposição à informação de que a “Terminologia teórica, à medida que evolui, está cada vez mais reconhecendo a necessidade de que o texto especializado integre seus objetos de estudo” (Krieger e Finatto 2004, p.181), entendemos que o texto de especialidade não se tornaria **objeto** de estudo, mas uma condição *sine qua non* para um olhar sobre termos e fraseologias. Nesse ponto, estamos em conformidade com a idéia de que o texto “(...) deixa de ser um mero conjunto de ‘termos técnicos’ relacionados, tornando-se um todo de estudo e análise também para a Terminologia” (id, p. 181), porém, ressaltamos, não como objeto, mas como potencial ponto de partida para observações acerca das unidades lexicais terminológicas. Assim, por exemplo, um estudo mais dedicado poderia observar que algumas das expressões que descrevemos neste trabalho contribuem para delimitar sentidos e

restringir interpretações nos textos dos artigos, tais como **no sentido de que, no que diz respeito a, de acordo com x**. Um olhar sobre o texto parece deixar claro que o rigor e a univocidade que a ciência necessita não são dados exclusivamente pelos termos, como parece ser o entendimento geral, mas pelo todo do texto, pela maneira como os termos são colocados em ação e também por outros fatores externos.

Por essas razões e nessa perspectiva, a visão que temos é que não seria necessário se falar de uma Terminologia textual, pois não há maneira de se fazer ou pensar Terminologia independentemente do texto, ou do ponto de vista do texto.

A Terminologia tende a beneficiar-se e a render mais proveitosos frutos se continuar, cada vez mais, a incluir em seu olhar outros elementos da trama textual no estudo de seus objetos, como os adjetivos, os conectores, a diagramação do texto, a pontuação, a retórica, dentre outros. Os produtos para os quais se presta a Terminologia aplicada, ou seja, principalmente os dicionários e os glossários, ficam mais enriquecidos se observarem os entornos dos termos. A descrição sistemática do texto desenvolvida no âmbito da Terminologia pode auxiliar a percepção de peculiaridades textuais de determinadas áreas de conhecimento, o que pode, por exemplo, ser muito útil para o tradutor (Bevilacqua, 1998). Como já dissemos anteriormente, esta pode ser uma perspectiva futura da pesquisa que realizamos aqui: avaliar quais (e como) os marcadores textuais típicos do artigo científico são empregados por diferentes domínios de conhecimento, como também exemplificamos no capítulo a seguir.

Este primeiro capítulo tratou de abordar conceitos relacionados aos estudos de gênero que delimitam e qualificam um estudo de gênero como necessário ao fazer e pensar tradutórios, uma vez que essas atividades envolvem sempre uma relação entre culturas, comunidades e contextos de ocorrência. O entendimento de que um gênero apresenta características típicas e recorrentes, produzidas e reconhecidas por parte dos membros das

comunidades que lidam com o mesmo, faz com que nossa preocupação em identificar algumas dessas unidades prototípicas seja validada. Também, a rápida incursão feita aqui a respeito do texto de especialidade e dos estudos da Terminologia visou mostrar um aspecto do texto do artigo científico, o de ter em sua constituição o caráter de uma área de especialidade. O objetivo maior de fazermos essa observação foi aproximar a discussão a respeito do gênero artigo científico da ótica do estudo do texto especializado, uma vez que são dois enfoques que se sobrepõem em determinados momentos e que, em geral, são vistos apenas sob os pontos de vista específicos de cada uma dessas áreas de estudo.

O capítulo a seguir restringe ainda mais o foco de estudo. Uma vez estabelecido aqui o que entendemos por gênero, em uma visão ampla, no Capítulo 2, nosso olhar se volta sobre um gênero em específico: o artigo científico.

CAPÍTULO 2: O ARTIGO CIENTÍFICO E OS MARCADORES TEXTUAIS

*O ato de “fazer ciência” pressupõe o ato de “falar ciência”,
“ler ciência”, adentrar um mundo que tem um código
e precisa ser dominado, se quisermos nos apropriar
do conhecimento.
Marlise Fontes Borges.*

2.1 Introdução

Nesta dissertação, nos voltamos para o estudo de marcadores textuais típicos do texto de artigos científicos. Já tratamos, no capítulo anterior, das condições de formação e de constituição dos gêneros e, agora, nos detemos em pontos específicos de nosso objeto, o artigo científico. Com base no que discutimos anteriormente, consideramos, *a priori*, que esse é um gênero já consagrado, que mantém um forte grau de convenções, lingüísticas e pragmáticas, que o restringem e moldam. As expressões que estudamos são parte dessas convenções; elas estruturam, organizam e dão o caráter retórico desse gênero textual, de modo que, ao final do processo de escritura, com o uso dessas expressões, se tenha em mãos um texto que se enquadra na categoria **artigo científico**.

Neste capítulo, após a apresentação de alguns dados quantitativos a respeito da produção de artigos científicos, nos detemos em estudos e considerações já realizados sobre os mesmos, principalmente nos descritos por John Swales (1990). Também enfocamos

aspectos da constituição dos artigos científicos e, em detalhe, um desses aspectos constituintes, os marcadores textuais. Para sistematização e apresentação dos mesmos, trazemos, à luz deste trabalho, alguns pressupostos e fundamentos do modelo sistêmico-funcional apresentado por Halliday (1985).

2.1.1 A produção de artigos científicos em números

Os artigos científicos, em sua maioria, são publicados em periódicos, anais de eventos, e *sites* na *internet* que, em geral, são organizados por associações profissionais e acadêmicas. O número dessas publicações tem aumentado cada vez mais. De acordo com um levantamento apresentado pelo Ministério de Ciência e Tecnologia (MCT) com base nos dados do Institute for Scientific Information (ISI) e dos National Science Indicators (NSI), no ano de 2002, foram publicados 730.229 artigos em periódicos científicos internacionais indexados, um crescimento de mais de 7,53%, se comparados aos 677.798 de cinco anos anteriores, e de 70%, se comparados aos 429.263 do ano de 1981, quando começaram a ser registrados esses dados (Gráfico 2.1).

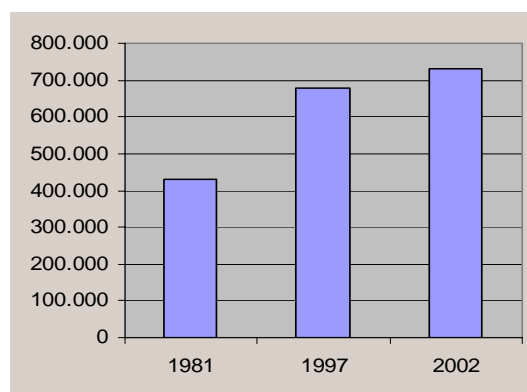


Gráfico 2.1 - Artigos publicados em periódicos científicos internacionais indexados – 1981, 1997 e 2002

De acordo com os dados, o Brasil é o sétimo país com maior crescimento no número de artigos publicados em periódicos científicos internacionais, com 11.285 artigos publicados no ano de 2002, um crescimento de mais 67%, se comparados aos 6.749 de cinco anos anteriores (Gráfico 2.2).

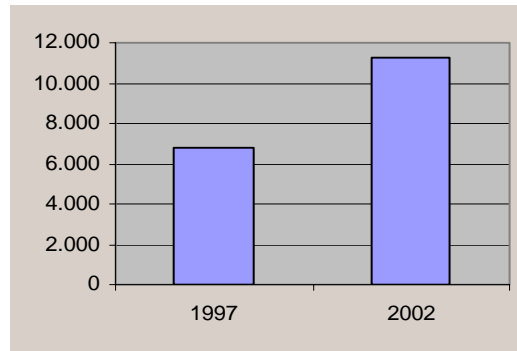


Gráfico 2.2 - Artigos brasileiros publicados em periódicos científicos internacionais indexados - 1997 e 2002

Como se percebe, os números de publicações vêm crescendo. Observamos também que a grande área escolhida para a constituição de nossos *corpora* (Engenharias e Ciências da Computação) foi a que mais apresentou número de artigos brasileiros completos em anais de eventos internacionais, como mostra a Tabela 2.1, com dados do MCT. Esses dados demonstram que há uma quantidade marcante de eventos realizados na área, o que sugere que, dado o caráter de constante e rápida evolução, esses eventos acabam sendo o local de divulgação e referencial do estado-da-arte na pesquisa e no desenvolvimento das chamadas tecnologias da informação. Com esses dados, queremos, além de justificar um olhar mais próximo a esses textos, também chamar atenção para o fato de que toda essa produção significa trabalho e mercado para tradutores, intérpretes, revisores e professores de português e de outras línguas, uma vez que os eventos se realizam também no Exterior.

Tabela 2.1 - Publicação de artigos brasileiros, por grandes áreas de conhecimento, em periódicos científicos internacionais indexados e em anais de eventos internacionais - 1998-2001

	Ciências Agrárias	Ciências Biológicas	Ciências Exatas e da Terra	Engenharias e Ciências da Computação	Humanidades	Ciências da Saúde
Artigos completos (periódicos)	9.993	31.298	37.798	15.966	5.346	19.971
Artigos completos (anais)	4.647	3.533	9.928	26.723	5.243	3.915
Total	14.649	34.831	47.726	42.689	9.589	23.886

2.2 O estudo do artigo científico

Conforme mencionamos na introdução deste capítulo, o nome de Swales é tradicionalmente citado quando se trata de estudos de gênero e, em particular, de um deles: o **artigo científico**. Seu foco e sua motivação são bastante pedagógicos, como ele mesmo estabelece: “O principal objetivo (...) é oferecer uma abordagem para o ensino do inglês científico e acadêmico” (Swales, 1990, p. 1). Embora o nosso objetivo aqui esteja voltado para a produção de textos a partir do processo conhecido como versão (em nosso caso no sentido português-inglês)⁴, utilizaremos a reflexão de Swales para embasar a caracterização desse gênero e apontar algumas das questões que estão envolvidas na sua produção.

⁴ No decorrer deste estudo, não faremos maiores distinções entre os processos de tradução e versão, principalmente no que tange às teorias sobre tradução, uma vez que estas parecem não fazer tal distinção.

Esse gênero surge, ainda que em uma forma embrionária, contemporaneamente com o estabelecimento do primeiro periódico científico em 1665, o **The Philosophical Transactions of the Royal Society**. De acordo com Ard (1983, apud Swales, 1990), o artigo científico desenvolveu-se a partir de cartas informativas trocadas entre pesquisadores, enviadas para o **Transactions**. Quando esse jornal começou a proporcionar um local para discussões, passou a assumir um novo papel, e os textos acabaram refletindo as situações retóricas novas e recorrentes, diferentes das escritas na forma de carta, e, dando origem, assim, aos artigos científicos.

Alguns cientistas que muito colaboraram para dar um formato aos primeiros artigos científicos foram Robert Boyle e seus colegas. “De acordo com Shapin [1984], através da experiência com o fato real, Boyle e seus colegas procuraram transformar reivindicações e especulações em um tipo de conhecimento que fosse mais amplamente aceito” (Swales, 1990, p.111) Para operar tal transformação, este estudioso desenvolveu algumas estratégias que mostram como ele e seus companheiros tiveram que trabalhar duro para tentar estabelecer uma retórica e desenvolver um estilo convincente para o artigo de pesquisa (id., p.112).

As estratégias, dentre outras, consistiam em usar testemunhas para provar que as experiências haviam realmente sido feitas e mostrar ilustrações do aparato em questão na ocasião de sua publicação. Estas deveriam ser realistas, exatas e detalhadas, como “testemunhas virtuais”. Boyle também era muito prolixo em seus textos, levando, assim, o leitor a acreditar que estava lendo um relato preciso da experiência realizada diante da descrição detalhada dos experimentos que funcionavam e os que falhavam, utilizando muitas expressões do tipo **talvez, parece que, não é improvável**

Um trabalho desenvolvido por Bazerman (1983) demonstrou que, no início, a maioria dos artigos eram relatórios sobre eventos naturais, como terremotos, ou sobre observações feitas a partir de microscópios, telescópios ou dissecação anatômica. Na evolução do artigo

científico, Bazerman (apud Swales 1990, p. 113) relata que a: “(...) relação do cientista com a natureza se alterou gradualmente de uma visão onde a natureza das coisas era facilmente revelada pela observação ou manipulação direta para uma onde a natureza era complexa, obscura e difícil de ser apreendida”. Isso levou a um cuidado em se descrever como os experimentos eram feitos, em explicar os métodos e em detalhar, com precisão, os resultados.

Certamente, há diferenças entre o modo de escrever um artigo científico no século XVIII e o que encontramos atualmente. O trabalho de Bazerman investigou artigos do periódico *Physical Review*, dos anos de 1893 a 1980, traçando um perfil de sua evolução e atual forma. Foram observados pontos como a extensão dos artigos, a inclusão e o tipo de referências feitas, as características sintáticas e lexicais, como construção das frases (o que indicou uma mudança de descrição para explicação), os tipos de sujeitos das principais orações (passaram de mais concretos, como **aparato** e **substância**, para mais abstratos, relacionados a processos ou à qualidade, como **ionização** e **correlação**.) Houve também uma diminuição dos verbos de relato e aumento dos verbos ativos, sugerindo que “(...) a descoberta ou teoria foi cada vez mais sendo levada para a posição gramatical central”, dentre outras mudanças (Bazerman, 1984 apud Swales 1990, p. 115). Bazerman também observou aspectos não verbais, como a diminuição da apresentação de desenhos e tabelas, o aumento no número e na complexidade de gráficos e equações e a diferente organização do artigo, com o uso de seções.

A retórica e a estrutura do artigo científico, hoje, mostram-se mais estabelecidas, e o questionamento do seu formato já não é mais motivo para discussão. Basicamente, os artigos são compostos de Introdução, Métodos, Resultados, Discussão e Conclusão, modelo esse proposto por Swales. Acrescentaríamos, porém, que, mesmo que muito dificilmente os artigos científicos sejam todos, na prática, assim formalmente divididos, eles contemplam em seu desenvolvimento aspectos semanticamente relacionados a essa divisão. Algumas áreas, como

a medicina, parecem seguir a risca essa divisão, já os textos de nossos *corpora*, em Ciência da Computação, não apresentam essa divisão como regra geral, apesar de, como dissemos, contemplarem aspectos relacionados à discussão, à metodologia no contexto geral de seu texto.

Podemos considerar, portanto, que os aspectos da composição dos artigos estudados por Bazerman continuam sendo prototicamente intrínsecos aos mesmos. Todos esses e outros podem ser observados em outras áreas, em artigos de outros periódicos e, ainda, sincrônica ou diacronicamente, mas seria prescritivo demais dizer que teriam o mesmo comportamento dos artigos estudados por Bazerman. Aqueles fatos levantados correspondem à análise feita por aquele autor, para os seus propósitos.

Em nosso caso, esta pesquisa não se volta para um estudo diacrônico, como o de Bazerman, mas também se preocupou em mostrar aspectos do artigo científico, em especial um deles, as expressões que são típicas de suas composição e estruturação textual, as quais foram aqui denominadas **marcadores textuais** e que mais adiante serão descritos.

O ato e o momento da redação de artigos científicos também já mereceram importantes e reveladoras considerações. Um estudo realizado por Knorr-Cetina (1981) demonstra que entre a primeira e a segunda versão de um mesmo artigo pode haver mudanças consideráveis, isso serve também para demonstrar que entre o trabalho no laboratório ou no campo e o artigo final há uma grande distância, o que nos alerta para o fato de que, ao estarmos diante de um artigo, não devemos ter a ilusão de que ele representa a própria pesquisa, isso porque

“(...) um artigo científico é um gênero um tanto diferente do relatório de laboratório e tem suas convenções próprias e separadas, seus próprios processo de raciocínio literário e seus próprios padrões de argumentação, dentro do qual um paradigma moldador muito forte é o do tipo de texto de solução de problemas” (Hoey, 1979 apud Swales, 1990, p. 119).

Na elaboração de um artigo, entram muitos fatores condicionantes que já fazem parte de todo contexto da comunidade científica e acadêmica à qual ele pertence, incluindo colegas, adversários, editoras, instituições e todo o entorno que envolve uma pesquisa. O reflexo dessas influências percebe-se tanto nos padrões de texto como nas expressões utilizadas. Nesse estudo de comparação entre uma primeira e uma segunda versão dos textos, Knorr-Cetina aponta, por exemplo, que muitas afirmações ‘perigosas’ são por vezes eliminadas; assim, **deveria** é substituído por **poderia**, **é** passa a ser **sugere-se**, **boa solubilidade** transforma-se em **melhor**, e assim por diante.

Sobre esse assunto, Swales menciona que não é que as notas tomadas no laboratório não pudessem ser publicadas de maneira linear, nem que a primeira versão de um artigo seria totalmente impublicável, mas o que acontece é que, na construção do artigo, há um processo de “(...) crítica técnica e controle social operando tanto no ambiente particular de pesquisa como em um outro mundo imaginado de ‘o que os outros cientistas irão pensar’” (Swales, 1990, p. 120). “O artigo publicado é um híbrido com multi-níveis, *co-produzido* pelos autores e por membros da audiência para o qual é direcionado” (Knorr-Cetina 1981, p. 106). Podemos perceber, assim, que entre a pesquisa em si e a escritura de um artigo existem muitos fatores operando, o que torna essa tarefa difícil e complicada até mesmo para membros estabelecidos de comunidades científicas.

Ao pensarmos o papel do tradutor, não podemos ignorar que todas as questões que apontamos acima – e dificuldades – estarão envolvidas também em seu trabalho, pois a ele caberá a tarefa de fazer o artigo operar, guardando sua estrutura de gênero, na língua de chegada de maneira satisfatória, tanto quanto opera na língua de partida. Segundo Hurtado Albir (2001, p. 491)

“O tradutor deve saber decodificar as convenções próprias do gênero a que pertence o texto original e saber utilizar as próprias do gênero na língua e cultura de chegada, quando essa for a finalidade da tradução”.

As dificuldades e responsabilidades enfrentadas pelo tradutor é a perspectiva que nossa pesquisa privilegia, pois pretende colaborar tanto no pensar o ensino e a prática da tradução voltados para os gêneros e suas regularidades e particularidades. Com isso, queremos também trazer uma pequena contribuição prática ao trabalho do tradutor, principalmente ao aprendiz, apresentando algumas das expressões que são típicas da estruturação de artigos científicos em português, em um primeiro momento, e, depois, fazendo uma comparação de usos e de emprego das mesmas em português-inglês.

Essas expressões, ou marcadores, foram evidenciadas neste trabalho a partir da análise de um *corpus* de 333 textos de ocorrências reais, em português, evitando-se partir de generalizações feitas em outros estudos ou gramáticas. O processo de análise que resultou na identificação desses marcadores é descrito no Capítulo 6. Agora, trazemos apenas alguns dos resultados, que serão apresentados na seção 2.3.1 a seguir.

2.2.1 Características textuais do artigo científico

Não é por ser um texto condicionado por muitos padrões que o artigo científico deixa de ser rico e de ter um caráter com muitas particularidades e diversidades. A nossa preocupação, nesta dissertação, foi bastante específica: estamos olhando para os marcadores textuais típicos do artigo científico, considerando a especificidade do gênero e não da área em que foram produzidos. Outros olhares poderiam ser lançados em relação às características de estruturação de texto, enfocando-se outros aspectos, como, as diferenças entre as áreas, por exemplo, se olhássemos os marcadores textuais típicos de textos de informática, encontraríamos “para a implementação de ...”; de medicina, “x membros em y pacientes foram estudados”; estatística, “se representarmos os elementos da amostra ordenada com a

seguinte notação: X1:n , X2:n; direito, “é de observar-se que, “O Capítulo x da lei de proteção a concorrência, Lei número 8884/94, trata de”, e assim em outros contextos também.

Também outras características dos artigos poderiam ser observadas, sendo de igual ou maior importância, para o tradutor, das que trouxemos aqui. Podemos citar, apenas como exemplos, o estudo das divisões em seções e o tipo textual envolvido em cada uma (narração, descrição, etc.), o estudo específico de cada uma das seções (resumo, introdução, conclusão), o uso de citações (em que momentos, e por que motivo), os tempos verbais e as pessoas empregadas, a indefinição e o apagamento do sujeito e muitos outros.

Seja qual for o estudo voltado para características prototípicas de um gênero que se fizer, certamente ele será interessante do ponto de vista da tradução e do tradutor, que precisa entender o gênero para fazer suas escolhas nos processos de decodificação, redação e adaptação envolvidos na tarefa.

A seguir, delimitamos ainda mais nosso objeto de estudo, apresentando um olhar sobre o recorte que fizemos a respeito do estudo do gênero artigo científico, isto é, o estudo dos marcadores textuais no artigo científico. Para esse fim, fazemos também a apresentação do modelo sistêmico-funcional, que fundamentou a nossa abordagem e a classificação dos marcadores.

2.3 Os marcadores textuais do artigo científico

Até o presente momento, ainda não foi esclarecida, exatamente, a denominação que utilizamos para as “expressões” ou “unidades” que despertaram o interesse para este trabalho. De fato, para se chegar a uma denominação, foi preciso fazer muitas idas e vindas na literatura. Tínhamos a hipótese, o *corpus* e o objeto, porém não tínhamos a denominação, ou um conceito teórico já definido. Houve sempre o desejo de trabalhar com unidades que tivessem um caráter fraseológico, ou de fixação recorrente, que fossem típicas do texto do

artigo científico, e o campo das fraseologias é outro que conta com muitos estudos, classificações e denominações diferentes.

Ficou definido que não trabalharíamos com unidades fraseológicas de caráter temático especializado, isto é, com fraseologias terminológicas, aquelas que são compostas por um termo, pois entraríamos em outro caminho, que não fazia parte de nossas inquietações no momento, as quais estão mais voltadas para questões de gênero do que de especificidades lingüísticas de uma área de conhecimento.

O elenco de unidades que levantamos a partir do *corpus* mostrou perfis sintáticos e semânticos bastante variados, com caráter às vezes de fraseologias, como em **é importante ressaltar que, a seção # apresenta, como mostra a figura #**; às vezes de colocações, como em **através do uso, a grande maioria, a partir da análise**, e às vezes de conjunções, como em **de forma que, de modo que**

Apesar desses perfis variados, e, considerando que essas expressões desempenhavam um papel no texto, predominantemente coesivo, optamos por uma denominação que fosse o mais abrangente, mas que guardasse a idéia de funcionalidade na constituição do artigo científico. Daí por que as expressões foram, então, genericamente denominadas **marcadores textuais**, em consonância com as proposições de Aluísio (1997), Berger et al. (2002) e Sáez, (2003). Ainda, poderíamos expandir o conceito para o de **marcadores multipalavras** ou **marcadores de segundo nível**, termos utilizados por Siepmann (2001), segundo o qual, um marcador de segundo nível seria, “(...) em sua definição mais simples, uma expressão de mais de uma palavra que realiza uma função pragmática ou de estruturação do texto” (Siepmann, 2001, p. 97). De acordo com esse autor, que estuda a tradução de marcadores discursivos restritores, “(...) existe uma falta visível de investigações a respeito desse tipo de unidade fraseológica”, com algumas exceções (id, p. 97). Como podemos ver, é uma definição

bastante ampla, e dá conta da variada natureza do elenco de unidades que selecionamos neste trabalho.

Sobre a variedade de denominações, Saéz (2003) fez um levantamento de termos utilizados na literatura e que denominam uma mesma idéia. Segundo esse autor, apenas em espanhol, seriam utilizados termos como *conectores*, *conectores extraoracionales*, *conectores argumentativos*, *conectores discursivos*, *conectores pragmáticos*, *conectores enunciativos*, *conectivos*, *partículas discursivas*, *enlaces textuales*, *relacionantes supraoracionales*, *elementos de cohesión*, *operadores discursivos*, *ordenadores del discurso* e *muletillas* (Martin Zorraquino Y Portolé Lázaro 1999, p. 4057 apud Sáez, 2003). Em inglês, outros termos também configuram essas estruturas, tais como: o mais genérico *metatextual elements* (Valero-Garcés, 1996), *contextualising frames* (Gosden, 1992) e *hedging devices*. Berger et al. (2002) utilizam o termo *discourse markers* considerando que estes não formam uma classe homogênea do ponto de vista sintático, mas sim do ponto de vista funcional, sendo escolhas do plano paradigmático.

Em nosso trabalho, também notamos essa característica, por esse motivo procuramos agrupar os marcadores por função e não por categoria sintática. É preciso, ainda, dizer que podem existir diferenças entre essa variedade de denominações que apresentamos acima, de acordo com os propósitos de cada autor e de cada estudo, mas, mesmo assim, essas expressões mantêm um ponto em comum: são elementos que garantem a coesão de um texto e também dão informações sobre sua estrutura retórica e, conseqüentemente, de seu gênero.

O termo genérico **metatextual elements**, por exemplo, ou elementos metatextuais, já foi utilizado em pesquisas baseadas nas macrofunções da linguagem de Halliday (Vande Kopple, 1985), que dividiram essas expressões, em inglês, em elementos com função textual de conexão (primeiro, próximo), de glosas de códigos (x significa y), marcadores de ação (em resumo) e narradores (de acordo com); e em elementos com função interpessoal de

modalidade (obviamente, claramente), de atitude (surpreendentemente); e de comentários (pode-se discordar com o fato de).

Outros estudos desenvolveram mais ainda essa classificação (Crismore and Farnworth, 1990 apud Valero-Garcés, 1996), incluindo outros elementos, como metadiscorso tático, por exemplo, de previsão (nesse trabalho proponho), de mudança de tópico (por outro lado); e metadiscorso léxico, por exemplo, limitador (nesse exemplo em particular), e enfático (é obvio que).

Como percebemos, esses marcadores são de categorias lingüísticas que poderiam pertencer a várias classificações e subclassificações. No entanto, mesmo sem nos aprofundarmos em uma classificação mais específica dessas estruturas, queremos ressaltar, novamente, que elas se constituem em elementos típicos do artigo científico. Conforme explicado por Azenha (1999, p. 85), já que não podemos falar de uma sintaxe das linguagens técnicas, podemos falar de

“diferenças no uso e na freqüência de certas estruturas em certos tipos de textos e que, nesses textos, tais estruturas, associadas à escolha lexical, à divisão do texto e ao emprego de recursos icônicos, dão sustentação aos objetivos traçados para a produção do texto técnico”.

Assim, acreditamos que a definição que adotamos é simples, mas adequada a tais estruturas: “(...) um marcador (...) é uma expressão de mais de uma palavra que realiza uma função pragmática ou de estruturação do texto” (Siepmann, 2001). Dessa forma, além de contemplarmos o caráter funcional das estruturas, contemplamos também o seu caráter, pode-se dizer, fraseológico, ou de colocação e recorrência.

Entendemos também que a denominação empregada não é o ponto mais importante, uma vez que a intenção e a razão de ser desta pesquisa está mais voltada a uma observação de comportamento desses elementos do que a uma conceitualização teórica ou ao enquadramento dos mesmos em categorias fixas. Mesmo assim, o elenco de marcadores que selecionamos a

partir do *corpus* foi classificado primeiro empiricamente, encontrando posteriormente fundamento nos pressupostos da gramática sistêmico-funcional proposta por Halliday (1985), que descrevemos a seguir, assim como a grande maioria dos estudos de marcadores e similares acima descritos, que também desenvolvem seus estudos com base nesse modelo.

2.3.1 *Os marcadores textuais e o modelo sistêmico-funcional*

Nas primeiras observações feitas sobre os dados levantados com o ferramental e a metodologia da Linguística de *Corpus* (ver Capítulo 5), conseguimos já estabelecer algumas diferenças entre a natureza funcional das expressões selecionadas. Distinguimos três tipos:

- (i) unidades relacionadas ao propósito do artigo científico, qual seja, apresentar um trabalho, avaliação, pesquisa ou hipótese; aqui estariam contempladas as expressões que dizem respeito ao objeto/problema de pesquisa (com o objetivo de), como é tratado/abordado (através do uso de, pode ser entendido como), quando acontece (durante a realização de, ao longo do processo), dentre outros que serão demonstrados em mais detalhe no Capítulo 6;
- (ii) unidades metatextuais, que fazem referência à própria estruturação e à organização discursiva do texto, bem como de outros elementos, como gráficos, figuras e tabelas (A seção # apresenta, como pode ser visto na tabela, este artigo apresenta, o artigo está organizado da seguinte forma);
- (iii) unidades modalizadoras, que dizem respeito ao caráter persuasivo do artigo científico e evidenciam a presença do sujeito na realização do texto (é importante ressaltar que, espera-se que, pode ser visto como).

Após essa classificação prévia, encontramos suporte teórico para a classificação no modelo sistêmico-funcional desenvolvido por Halliday (1985) acerca das três macrofunções da linguagem (textual, interpessoal e ideacional).

Nesta seção, optamos por fazer uma introdução aos conceitos do modelo sistêmico-funcional, deixando claro que tal abordagem é direcionada ao propósito de justificar a classificação das unidades selecionadas e não se propõe como uma revisão ampla dessa teoria e de seus conceitos.

Halliday (1985), na introdução de **An Introduction do Functional Grammar**, explicita por que assim chamou o seu livro, esclarecendo seu entendimento de **functional** e de **grammar**. Esclarece que é uma gramática **funcional** em três sentidos:

- (i) no sentido que é designada para dar conta de como a língua é utilizada. “Tudo que foi dito ou escrito se desenvolveu em algum contexto e, além disso, são os usos da língua que, por milhares de gerações, vêm moldando o sistema. A linguagem evoluiu para satisfazer necessidades humanas; e a maneira como é organizada é funcional com respeito a essas necessidades” (id, xiii);
- (ii) os componentes de significado na língua são componentes funcionais, “todas as línguas são organizadas em torno de dois principais tipos de significados, o ‘ideacional’ e o ‘interpessoal’”. Na teoria sistêmico-funcional, esses componentes são chamados de metafunções e são as manifestações no sistema lingüístico de dois propósitos mais amplos que subjazem todos usos da língua (a) entender o ambiente (ideacional) e, (b) agir em outros a partir dele (interpessoal). Combinados com esses dois componentes está um terceiro componente metafuncional, o textual, que dá relevância aos outros dois;
- (iii) cada elemento na linguagem é explicado por referências à sua função na totalidade do sistema lingüístico. Nesse terceiro sentido, “(...) uma gramática

funcional é aquela que constrói todas as unidades de uma língua – suas sentenças, frases e etc – como configurações orgânicas de funções. Em outras palavras, cada parte é interpretada como funcional com relação ao todo”.

Nesse contexto, o enfoque de uma análise está na investigação do uso da língua em condições reais de ocorrência (Halliday, 1985, xiii). Para tal, estuda-se e descreve-se a língua a partir de produções textuais autênticas, sejam elas orais e/ou escritas, relacionando as possibilidades lingüísticas formais de escolha, dentro de um sistema de significados, à função ou à necessidade social (intenção e sentido) a ser desempenhada (Egins, 1994, p.1 e Halliday, 1985, p. xiv-xv apud Santos, 1996) dentro de um contexto sociocultural observável, porque a língua é o que é porque desempenha funções para os falantes (Halliday, 1978, p.19).

Se pensarmos nessas afirmações no âmbito do que viemos defendendo até aqui, veremos que a os pressupostos de Halliday entram em acordo com outros que já apresentamos, tais como: (i) os contextos social e cultural são determinantes no uso que se faz da linguagem, retomando aqui a idéia de comunidade discursiva e esferas de atuação (ver Capítulo 1); (ii) a investigação do uso da língua em condições reais de ocorrência, o que demonstramos na escolha da metodologia e dos pressupostos da Lingüística de *Corpus* (ver Capítulo 5); e (iii) as escolhas lingüísticas dão-se em função da necessidade social dentro de um contexto sociocultural; também é o que tentamos deixar claro em nossa visão de gêneros descrita no Capítulo 1.

Essa feição de trabalhar com o uso real da língua, e as noções de funcionalidade da linguagem, isto é, “linguagem que desempenha um papel em um contexto” social é que justificam o fato de o modelo sistêmico-funcional ter sido já bastante aplicado pelos estudos de ensino de línguas para fins específicos. (Santos, 1996, p. 1). Quando se trata de inserção em culturas e contextos sociais distintos, entramos diretamente em questões importantes

dentro do ensino de línguas e também da essencialidade da tradução – enquanto produto e processo –, pois estamos lidando com pontos que interessam ao querermos ser coerentes em culturas e contextos distintos, ou diferentes dos nossos de origem. Estes também são motivos que nos levaram a escolher o modelo sistêmico-funcional para embasar o tratamento que demos às unidades que levantamos no *corpus*.

Em uma perspectiva sociosemiótica como essa proposta por Halliday, de acordo com Barbisan (1995, p. 52), o texto enquanto produto é “(...) instância de significado social num contexto particular de situação. É um produto de seu contexto, produto de um contínuo processo de escolhas no significado, escolhas representadas nas redes que constituem o sistema lingüístico”. Esse **contexto de situação** está no texto por meio de uma relação sistemática entre **contexto social** e **organização funcional da linguagem**.

Para a interpretação do **contexto social**, Halliday (1985) propõe três conceitos: campo, teor e modo. Esses conceitos são expressos por traços lingüísticos particulares, por exemplo: (i) campo, faz-se notar no vocabulário; (ii) teor, na escolha da função da fala; (iii) modo, na escolha dos temas.

O **contexto de situação** constrói-se na mente das pessoas que irão receber o texto, como um modelo montado a partir do que se percebe nos traços de **campo, teor e modo**. Aqui, podemos retomar o conceito de **comunidade interpretativa**, apresentado no Capítulo 1, que defende que a recepção do texto estará condicionada às predições que as pessoas fazem a partir do contexto de onde recebem o texto. Acreditamos que o modo de perceber esse contexto (o campo, o teor e o modo) é condicionado por todo o contexto social, como esclarece Fish (1980, p. 331):

“... todos os objetos são construídos e não encontrados prontos, e são construídos pelas estratégias interpretativas que colocamos em ação. Isso, porém, não fica comprometido pela subjetividade, pois os meios pelos quais estas estratégias acontecem são sociais e convencionais. Isto é, o ‘eu’ que realiza esse trabalho interpretativo é um ‘eu’ comunitário e não um indivíduo isolado. Ninguém levanta de manhã e

reinventa a poesia ou elabora um novo sistema educacional (...). Não fazemos essas coisas porque não poderíamos fazê-las, porque as operações mentais que podemos realizar são limitadas por instituições nas quais *já* estamos inseridos.”

O texto para Halliday é feito de sentidos, “uma unidade semântica” e, como tal, deve ser considerado como produto, isto é, resultado de um contínuo processo de escolhas, e como processo, a própria escolha semântica na rede de significados potenciais (Barbisan, 1995). Nas próprias palavras de Halliday (1978, p.122) um texto é “(...) a realização, em termos lingüísticos, das escolhas léxico-gramaticais feitas por um indivíduo de acordo com a função, com o tipo de interação e com o ambiente social dentro de um sistema de potenciais significados”.

De acordo com Halliday (1985, p. 53), existem “(...) três tipos de significado **incorporados** à linguagem humana como um todo, formando a base da organização semântica de todas as linguagens naturais” e, portanto, significados que estão presentes nos textos e nas sentenças. Esses significados são chamados, na abordagem sistêmico-funcional, de ‘metafunções’ ideacional, interpessoal e textual”. Essas metafunções “(...) englobam todas as possibilidades de escolhas semântico-lingüísticas possíveis para um falante em situações de interação com os outros ou com o mundo” (Santos, 1996, p. 9). Elas estão relacionadas a: representação do mundo real, modo de interação entre falante e ouvinte (ou redator e leitor) e relações lógicas, ou seja, a textura, que envolve a coesão do texto.

A **metafunção ideacional** está relacionada à natureza do evento social, isto é, à “(...) expressão do conteúdo de acordo com a experiência do falante inserido em uma comunidade de fala” (Halliday, 1978, p. 37); expressa relações entre o autor e a matéria (Aluísio, 1997); é a metafunção que se “(...) costuma chamar de função cognitiva ou referencial da linguagem (Koch, 1983, p.37). A função ideacional de uma sentença “(...) é a que representa o que, no sentido mais amplo, chamamos ‘processos’: ações, eventos, processos de consciência e relações” (Santos, 1996, p. 9). Tal função encontra-se nos traços do texto que representam o

mundo real como apreendido por nossa experiência. Está relacionada ao **campo**, que se faz notar no vocabulário, é “(...) o que está acontecendo, a natureza da ação social” (Barbisan 1995, p. 52).

A **metafunção interpessoal** está relacionada ao papel desempenhado e ao *status* ocupado pelos participantes, ou seja, à “expressão das relações sociais e pessoais” dos papéis desempenhados nas interações; “(...) é o significado como forma de ação: o falante ou escritor fazendo alguma coisa para o ouvinte ou leitor através da linguagem” (Halliday 1985, p. 52); ela expressa “relações entre o autor e a audiência” (Aluísio, 1997), a “posição que o locutor assume perante o ouvinte, no processo da enunciação. A metafunção interpessoal concretiza-se no sistema de **teor**, refletindo o engajamento e o envolvimento do falante com a asserção e com os participantes da interação.

A **metafunção textual** relaciona-se ao “cumprimento das exigências de operacionalidade de uma língua” (Halliday 1978, p. 42 apud Santos, 1996, p. 10), ou seja, ao papel da linguagem, com quais escolhas lingüísticas a mensagem está sendo construída e organizada dentro de um sistema determinado pelo contexto e pela intenção do autor. Relacionado à metafunção textual está o **modo**, que aparece na escolha dos temas, em traços léxico-gramaticais específicos à língua oral ou à escrita, no uso de conjunções ou outro tipo particular de coesão (Barbisan, 1995, p. 52)

Halliday (1985, p. xiii) comenta que a metafunção textual dá relevância às outras duas e é somente em razão de podermos selecionar a forma desejada da mensagem que podemos utilizar a linguagem eficazmente, tanto para entender o ambiente (função ideacional) como para interagir (função interpessoal) com quem está ao nosso redor.

Com base nesses breves esclarecimentos acerca dos conceitos do modelo sistêmico-funcional e das metafunções da linguagem, apresentamos, a seguir, algumas de suas

aplicações na abordagem de artigos científicos, também já nos dirigindo para como o modelo sistêmico-funcional pôde ser aplicado a este trabalho.

2.3.2 *O modelo sistêmico-funcional e o artigo científico*

Há um estudo clássico, já citado neste trabalho, realizado por Knorr-Cetina (Swales 1990, p. 119), em que foram observadas as atividades de pesquisa em laboratório e o processo de escrita dos artigos a respeito dessas atividades por parte de um grupo de bioquímicos. Nesse trabalho, Knorr-Cetina fez uma análise detalhada da primeira e da última versão da introdução de um artigo e classificou as principais estratégias empregadas nas alterações que os autores faziam nos textos em três tipos principais (Gosden, 1995, p. 42):

- (i) o apagamento de certas afirmações, tanto argumentos óbvios que, em geral, reforçavam algum ponto, como afirmações consideradas “fracas” ou “perigosas”;
- (ii) o reordenamento de afirmações, levando a uma nova estrutura de parágrafo;
- (iii) as mudanças na modalidade de algumas afirmações, como, por exemplo, do **necessário** ao **possível** e, em geral, do que fora afirmado **com certeza** para uma afirmação **mais sutil**.

Esse estudo da revisão dos artigos tornou possível entender melhor seu processo de escritura e evidenciou fatos inusitados também, como, por exemplo, que o artigo escrito evidencia elementos que não são necessariamente mencionados durante a pesquisa como os nomes de marcas de aparelhos ou as origens de alguma técnica. O trabalho de Knorr-Cetina permitiu conhecer mais sobre a natureza do artigo científico, enquanto produto e processo. Seu trabalho originou um outro que veio a ser relevante na realização de nossa pesquisa com

artigos científicos: o de Hugh Gosden (1995), pesquisador que tem estudado o artigo científico no Tokyo Institute of Technology, sob a ótica da gramática sistêmico-funcional.

Em um de seus trabalhos e a partir das observações feitas por Knorr-Cetina, Gosden (1995) expande esse quadro de estratégias de revisão e reelaboração dos artigos com base em um pequeno *corpus*. Ele percebe, por exemplo, que a estratégia b (reordenamento de afirmações) teve menos ocorrências. No entanto, inclusões significativas de detalhes técnicos aconteceram em maior número. Além disso, muitas das inclusões feitas nas versões finais não se relacionavam apenas à matéria em questão, mas tinham funções importantes na organização textual. O pesquisador apresenta, então, cinco categorias de revisão textual:

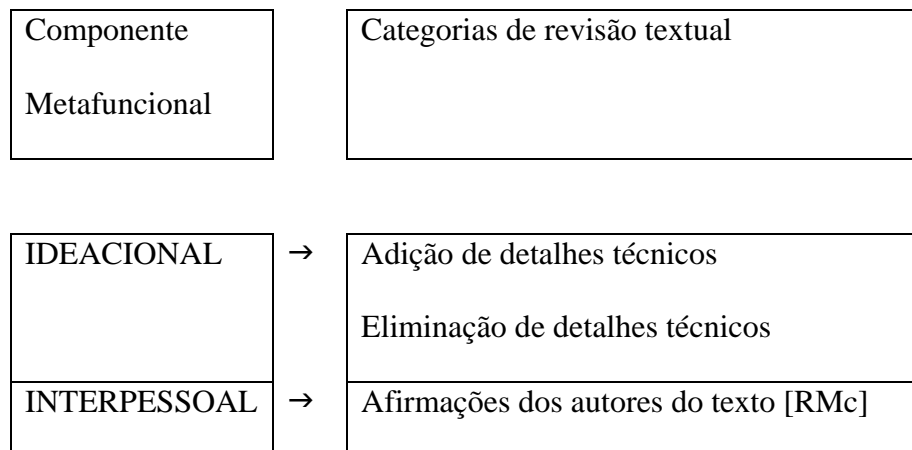
- (i) a *eliminação* de detalhes ou afirmações técnicas;
- (ii) a *adição* de detalhes ou afirmações técnicas;
- (iii) a *reorganização* das afirmações;
- (iv) as *modificações textuais*, refletindo o que Swales (1990) chamou de *rhetorical machining* do discurso científico;
- (v) o *polimento* da linguagem, limpeza de pequenos erros.

A categoria **d** é a mais desenvolvida no trabalho de Hugh Gosden e também a que mais nos interessa, pois nela está incluído o tipo de unidades que levantamos no presente estudo de nosso *corpus*, que fazem parte dos mecanismos retóricos do texto do artigo científico. As categorias elaboradas por Gosden e a subdivisão da categoria **d**, apresentada a seguir, também vêm sendo utilizadas em outros estudos, como o Projeto Amadeus, coordenado pela Professora Sandra Maria Aluísio, do Núcleo Interinstitucional de Linguística Computacional (NILC-SP), que visa ao desenvolvimento de recursos e ferramentas para auxiliar falantes não nativos do inglês a escreverem artigos científicos.

A categoria **d** do trabalho de Gosden foi subdividida em três orientações básicas:

- (di) modificações textuais que se relacionam à estrutura do discurso (RMd). Essa categoria inclui o uso de temas marcados como **contextualizadores**, tais como: “in addition”, “here”, “furthermore”, assim como marcadores da organização do discurso como “x will be discussed in the next session”, “as shown in Figure 1”.
- (dii) alterações relacionadas às afirmações dos redatores do texto (RMc) e que, portanto, refletem a consciência de um retorno por parte da comunidade acadêmica. Essa categoria inclui a adição do uso de modalizadores: “it can be suggested that”, “it seems reasonable to conclude that”, “x may be interpreted as”.
- (diii) alterações relacionadas ao propósito do redator do texto (RMp) e à expressão das razões **para** e resultadas **das** ações tomadas e das conclusões alcançadas, como: “therefore”, “consequently”, “accordingly”, “in order to”, “because”.

Gosden faz uma relação entre essas categorias de revisão textual e as metafunções da linguagem propostas por Halliday. O autor salienta que as categorias lingüísticas raramente são passíveis de ser classificadas e que Halliday aponta a necessidade de se falar em **tendências** e não em regras. Assim, a Figura 2.1 mostra tendências dessas relações.



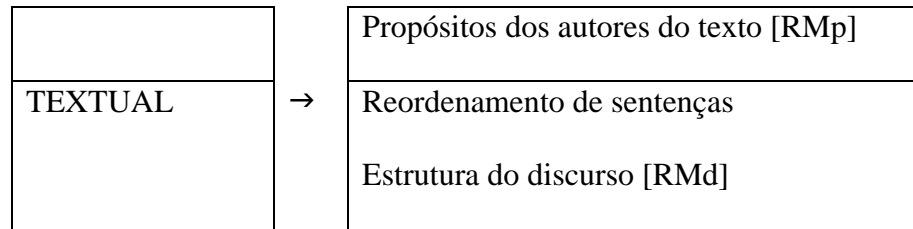


Figura 2.1 - Categorias de revisão textual enquanto realizações dos componentes metafuncionais da linguagem de Halliday.

Fazendo uma leitura da Figura 2.1, temos que:

- a reorganização de detalhes técnicos e mecanismos retóricos da estrutura do discurso (RMd) tendem a refletir a manipulação *textual* e as relações coesivas;
- as afirmações dos autores (RMc) e a expressão do propósito (RMp), como no esforço de convencer a audiência sobre a validade da pesquisa, tendem a refletir os significados **interpessoais**, como percebemos nos sistemas de modalidade;
- a adição e a eliminação de detalhes técnicos relacionam-se à matéria em questão, portanto a aspectos da metafunção **ideacional**.

Em nosso trabalho, as unidades selecionadas e que haviam sido previamente separadas de acordo com aquelas três funções que descrevemos anteriormente foram, então, formalmente enquadradas de acordo com os três grupos formulados por Gosden, com base nas metafunções da linguagem: (i) marcadores relacionados ao propósito do artigo científico, equivalentes ao RMp; (ii) marcadores metarreferenciais, equivalentes ao RMd; e (iii) marcadores subjetivos, relacionados ao RMc. Dentro de cada um dos três grupos, em nossa classificação, os marcadores foram agrupados de acordo com uma palavra-núcleo ou por semelhanças de função no texto. Os quadros com as relações completas dos marcadores

textuais que selecionamos são apresentadas no Capítulo 6, e a metodologia empregada para obtenção dos marcadores a partir do *corpus* é relatada no Capítulo 5. Nesta seção, trazemos apenas alguns exemplos de cada tipo, extraídos do *corpus*, para melhor visualização e entendimento dessas categorias estabelecidas em conformidade com o modelo sistêmico-funcional.

2.3.2.1 Marcadores relacionados ao propósito do artigo científico.

Esses marcadores relacionam-se a elementos do propósito do artigo científico, cujo texto visa, de maneira geral, a apresentar objetivos e objetos do estudo, descrever a fundamentação teórica, ações e métodos e apresentar as conclusões. Foram agrupados em torno de duas palavras-núcleo **objetivo** e **forma** – seis grupos funcionais **método**, **ancoragem**, **quantificação**, **tempo**, **explicação** e **restrição**.

A seguir, apresentamos contextos de uso de um marcador de cada um dos oito subgrupos.

Objetivo – as expressões agrupadas por essa palavra-chave tratam de apresentar objetivos ou propósitos do artigo, da pesquisa, do uso de um método, dentre outros, como podemos perceber na expressão a seguir:

com o objetivo de

“Com o objetivo de comparar os resultados com os das redes MLP, foram usadas redes de Elman com topolog.”

Método – essas expressões têm o caráter de mostrar como determinada ação foi feita, tomada.

Ex.:

através do uso

“A disjunção é simulada através do uso de várias regras com o mesmo conseqüente.”

Quantificação – essas expressões trazem elementos quantificadores. São características dos artigos que trazem dados, números, comparações de resultados, etc., e são esses marcadores que realizam tal função. Ex.:

a maior parte

“A interface do MAESTRO é a parte da aplicação com a qual o usuário tem contato direto. A maior parte do processamento do MAESTRO é feita no lado do servidor.”

Forma – agrupamos essas expressões pelas palavras-chave **forma, modo e maneira**, já que tinham uma frequência bastante grande de ocorrências. Algumas desempenham um papel conclusivo, como mostra o exemplo abaixo:

de forma que

“O ambiente virtual deve, ainda, fornecer ao aprendiz recursos de navegação não preestabelecidos, de forma que ele possua a liberdade de construir o seu próprio conhecimento através da agregação de valor.”

Ancoragem – essas expressões têm um sentido de ancorar um ponto de partida para uma afirmação, por exemplo:

a partir da análise

“As contribuições foram identificadas a partir da análise da interação dos alunos com a ferramenta.”

Tempo – esses marcadores dão uma idéia de localização no tempo de algum aspecto da pesquisa ou experimento; algumas também realizam o papel de ordenamento e organização textual. Ex.:

até o presente momento

“Os resultados colhidos até o presente momento mostram o crescimento da maturidade em Engenharia de Software, como pretendido.”

Explicativos – esses marcadores realizam a função de explicitar algum ponto da pesquisa, dos resultados ou das conclusões. Também podem introduzir um exemplo ou acrescentar alguma informação no sentido de trazer uma explicitação ao texto.

se por um lado ... por outro

“Se por um lado descrições declarativas tendem a ter um nível de abstração mais alto, facilitando a modelagem, por outro, à medida que as descrições crescem ...”

Restritivos – essas expressões foram chamadas de restritivas pois entendemos que conduzem a um foco específico, ligado a um referencial, restringindo desvios de enfoque e outras possibilidades de interpretação.

no que diz respeito a

“No que diz respeito à linguagem de programação, optou-se pelas Java Server Pages [Brown 2001], uma vez que, em comparação com outras linguagens ...”

2.3.2.2 Marcadores metarreferenciais

Fazem referência ao próprio texto e a outros elementos gráficos, como figuras e tabelas, que, por sua vez, não deixam de pertencer à categoria anterior. Foram agrupados em torno de sete palavras-núcleo: **seção, figura, tabela, artigo, trabalho, exemplo** e um grupo

específico de estruturas que se referem diretamente à organização do texto. A seguir apresentamos um exemplo de cada um dos grupos.

Seção – as expressões agrupadas em torno da palavra-chave *seção* fazem referência a alguma das sub-divisões do artigo, como mostra o exemplo:

a seção # apresenta

“A Seção 2 apresenta a conceituação de redes Bayesianas, seu mecanismo básico de inferência e o modelo noisy-or.”

Figura – as expressões agrupadas em torno da palavra *figura* fazem referência à introdução ou exemplificação de algum ponto do artigo por meio de uma figura.

x pode ser visto na figura

“O modelo de sujeito individual (Si) em interação com uma determinada ferramenta computacional pode ser visto na figura 2.”

Tabela – as expressões agrupadas em torno da palavra *tabela* fazem referência à introdução de dados apresentados em uma tabela.

x é apresentado na tabela #

“O resultado desse processo é apresentado na Tabela 1. Para simplificação, alguns nomes de disciplinas foram abreviados.”

Artigo – as expressões que contêm a palavra *artigo* fazem referência ao próprio texto que está sendo produzido

o objetivo deste artigo é

“O objetivo deste artigo é apresentar um sistema automatizado de transcrição de textos escritos no sistema Braille para textos escritos no sistema óptico.”

Trabalho – semelhante à anterior. Nessas expressões, **trabalho** é utilizado como sinônimo de **artigo**.

Este trabalho apresenta

“Este trabalho apresenta a Ferramenta Fuzzy para acompanhamento do Desempenho dos Alunos nos Cursos à Distância.”

Organização – essas expressões explicitamente organizam o artigo, apontando como irá se desenvolver, organizar e se apresentar macroestruturalmente. Ex.:

este artigo está organizado da seguinte forma

“Este artigo está organizado da seguinte forma: na seção 2 faz-se uma breve introdução à Programação Genética. Na seção 3, a abordagem multiobjetivo proposta é descrita. Os experimentos são analisados na seção 4. As conclusões são apresentadas na seção 5.”

2.3.2.3 Marcadores subjetivos

Foram assim denominados porque carregam consigo a visão do sujeito que realiza o artigo, ou que relata a pesquisa. Contam com a presença de muitos modalizadores e foram agrupados em torno de cinco palavras-núcleo: **importante**, **necessário**, **esperar**, **observar**, **poder**.

Importante – as variações das expressões que contêm a palavra-chave importante contam com outros verbos enfáticos, como **destacar** e **salientar**.

é importante ressaltar que

“É importante ressaltar que o programa desenvolvido pode ser integrado em outras plataformas ou ambientes colaborativos.”

Necessário – apenas duas expressões com a palavra necessário mostraram-se mais frequentes no *corpus*: **é necessário que...** e **é necessário que se**. Essa diferença é apenas do opções sintáticas, como mostram os exemplos:

é necessário que se

“Portanto, na implementação computacional de tal conceito, é necessário que se torne possível a utilização, pelo professor, desses quesitos.”

É necessário que

“Para que a colaboração se realize, é necessário que a infra-estrutura do ambiente de colaboração seja projetada de modo a tornar o grupo coeso.”

Esperar – as expressões agrupadas em torno dos verbos esperar e acreditar fazem referência a expectativas dos autores dos artigos com relação aos resultados e às metas a que se propuseram. São expressões tipicamente encontradas na introdução e na conclusão dos artigos.

espera-se que

“Espera-se que este trabalho possa contribuir para a elaboração de outros sistemas de ensino musical e ensine música com concepções e parâmetros adequados ao desenvolvimento musical.”

Observar – o verbo observar, palavra-chave dessas expressões, implica a presença de um sujeito que realizou o experimento, a análise, a reflexão e observou resultados, pontos críticos

ou importantes. Essas são expressões que poderiam igualmente ter sido incluídas na categoria de marcadores relacionados ao propósito do artigo, uma vez que a observação é característica do pensar e fazer científicos.

observou-se que

“Observou-se que a maioria dos casos não cobertos por este algoritmo poderiam exigir a introdução de informação semântica.”

Poder – nessa categoria foram incluídas as expressões com um caráter modalizador, dado pelo verbo poder. Também outras expressões, representando possibilidade foram incluídas nessa categoria.

pode ser definido como

“Um componente pode ser definido como uma parte física e substituível de um sistema que a ele se conecta e que provê a ele um conjunto de interfaces [1].”

pode-se dizer que

“Assim, pode-se dizer que o software DWB, instrumento que estava mediando a relação dos alunos com o conteúdo possibilitou-lhes a aprendizagem de um novo aspecto de um mesmo assunto.”

Apresentamos então, resumidamente, como ficou constituída a classificação dos marcadores textuais em nosso trabalho.

Por fim, mas não menos importante, cabe ainda fazer uma ressalva à denominação marcadores textuais, pois entendemos que os componentes da terceira categoria apresentada, dos marcadores subjetivos, têm, na realidade, um caráter discursivo, isto é, não estão no nível de organização textual, mas expressam posições dos produtores do texto, através das modalizações aqui identificadas. Embora estes últimos marcadores tenham essa natureza

distinta dos que cumprem um papel coesivo no texto, achamos importante relacioná-los, tendo em vista que esses aspectos também afetam o conteúdo dos artigos, e o tradutor deve observá-los atentamente em seu trabalho. Tanto é assim que os resultados, decorrentes da aplicação informatizada que fizemos, identificam essas estruturas modalizadoras com alta frequência.

2.4 Algumas considerações

Gostaríamos de ressaltar que a classificação, em si, que apresentamos, dessas unidades não é o foco central deste trabalho. O que consideramos importante é o levantamento dos marcadores e a sua comparação de uso em duas línguas, que demonstramos no Capítulo 6. Ao observarmos classificações não empíricas de outros trabalhos, por exemplo Gosden (1992), vemos que acontece uma “classificação pela classificação”, sem servir a um propósito mais prático. Ainda, como o próprio Halliday propôs, é muito difícil estabelecer cortes a categorias lingüísticas. Em nosso caso, percebe-se que algumas expressões classificadas em uma determinada categoria poderiam também ser incluídas em outras, pela suas várias funcionalidade no texto. A classificação que fizemos, como toda classificação, pode ser problematizada, mas é uma possibilidade que permite olhar os marcadores selecionados de maneira mais funcional e sistemática.

Resumimos, na Figura 2.2, a seguir, a relação que foi feita entre a subdivisão que efetuamos e aquela efetuada por Gosden. Assim, as unidades pertencentes à primeira categoria apresentada, (i), estariam no nível do RMp, que se relaciona aos propósitos do artigo científico; as pertencentes à segunda categoria apresentada, (ii), estariam no nível do RMc, que se relaciona às afirmações dos autores (caráter subjetivo); e as pertencentes à terceira categoria apresentada, “iii”, no nível do RMd, que se relaciona à organização textual do artigo científico.

Encerramos este capítulo que procurou delinear o que é o artigo científico, alguns eventos que lhe são peculiares e alguns estudos já realizados sobre suas características e sua natureza. Também trouxemos a classificação feita, o modelo teórico utilizado e alguns exemplos dos marcadores extraídos do *corpus*.

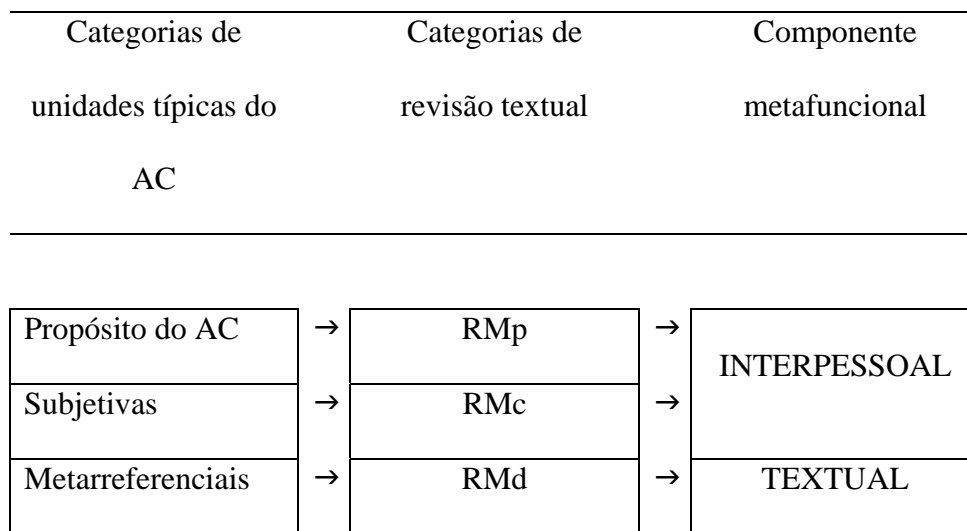


Figura 2.2 - Relação entre as categorias estabelecidas para este trabalho, as categorias apresentadas por Gosden (1985) e os componentes metafuncionais da linguagem.

Reforçamos que defendemos, neste trabalho, que as escolhas dos autores por esses recursos, ou marcadores textuais, são largamente determinadas pela estrutura interna do gênero artigo científico, que, por sua vez, é moldada a partir das expectativas e experiências da comunidade científica a qual o gênero pertence. Assim, voltamos também para uma área que procura estudar as realizações dos textos em diferentes culturas, a Retórica Contrastiva, pois, embora possa se dizer que o artigo científico tem um padrão universal, da comunidade científica internacional, entendemos que as diferentes línguas encontram maneiras distintas de realizar os elementos universais dos artigos, utilizando os próprios recursos lexicais e gramaticais.

Entendemos assim que os meios retóricos, dos quais os marcadores fazem parte, podem ser muito semelhantes em diferentes culturas de escrita (*writing cultures*), mas suas freqüências e usos preferenciais diferem (Mauranen, 1993, p. 5). Esta é a motivação para a realização deste trabalho, são essas diferenças que nos interessam como objetivo principal e final deste trabalho, pois, uma vez que culturas diferentes produzem textos diferentes, é fundamental que o tradutor e suas ferramentas de trabalho reconheçam, cada vez mais, onde estão e quais são as diferenças da língua de partida para a língua de chegada, na qual será produzido o texto, na busca por um trabalho qualificado e confiável.

O Capítulo 3 traz alguns fundamentos da Retórica Contrastiva, no que interessam a este trabalho e ao estudo da tradução.

CAPÍTULO 3: TRADUÇÃO E RETÓRICA CONTRASTIVA

3.1 Introdução

O pressuposto que originou esta pesquisa surgiu da prática da tradução. Porém não foi com pouca satisfação que encontramos estudos e teorias com objetivos muito semelhantes aos nossos e que proporcionaram suporte teórico e deram legitimidade ao trabalho que fizemos. A Retórica Contrastiva (RC) foi uma dessas teorias e, por isso, tratamos de explicitar, neste capítulo, seus principais pressupostos e motivações e como ela contribuiu para a área de tradução.

3.2 Retórica Contrastiva

Como já afirmamos anteriormente e procuramos demonstrar no Capítulo 2, as unidades que levantamos a partir dos *corpora* não dependem da área a que os textos selecionados pertencem, mas, sim, de certas convenções e regularidades próprias da retórica, e dos modos de dizer nas duas línguas em questão. O termo **retórica** está sendo empregado aqui em um sentido amplo, do “bem escrever” para a obtenção do que se deseja. No caso do artigo científico, podemos dizer que alguns dos objetivos de uma boa retórica seriam dar credibilidade às proposições e convencer o leitor do que está sendo apresentado, além de apresentar e organizar o texto de forma coerente e coesa.

Os meios retóricos de que um autor se vale para fazer um texto acabam sendo relacionados e restringidos por valores, sistemas de crenças e práticas comuns à comunidade discursiva que irá receber e/ou avaliar o texto. Em textos acadêmicos, as comunidades discursivas que mais restringem os autores são aquelas para as quais os textos estão sendo escritos, e a da cultura de escrita na qual o autor foi socializado quando aprendeu a escrever, que é, em geral, sua cultura nacional (Mauranen, 1993, p. 5). Essas duas culturas entram em questão e conflito em um trabalho de tradução, como não poderia deixar de ser. E é para com esse processo de adaptação (ou não) às formas retóricas da comunidade discursiva do texto de chegada que queremos contribuir, uma vez que assumimos, desde já, que existem formas retóricas semelhantes disponíveis nas duas línguas, mas que suas frequências e usos preferidos se diferenciam (id, p. 5). Estudos que examinam o fenômeno de que estudantes de uma segunda língua não escrevem da maneira esperada na língua-alvo têm sido relatados (Saéz, 2003; Valero-Garces, 1996; Mauraneen, 1993) desde Kaplan (1966). Isto não quer dizer, necessariamente, que os alunos estariam escrevendo de forma errada, mas apenas diferente. A RC tenta dar conta do estudo de fatores como esses, e queremos trazê-la também para mais perto de uma perspectiva da tradução.

Essa corrente de estudo surgiu nos Estados Unidos, no final dos anos 60, pela voz do lingüista Robert B. Kaplan. Sua motivação nasce de uma necessidade e de um interesse pedagógicos, pois era grande o número de estudantes estrangeiros que chegavam aos EUA naquele momento, e Kaplan, então diretor de um programa de ensino de inglês como segunda língua, percebia diferenças significativas na organização dos textos escritos por estudantes falantes de inglês e de outras nacionalidades e culturas.

Os estudos em RC foram, então, uma tentativa de entender essas diferenças e de propor metodologias de ensino apropriadas para diminuí-las (Sáez, 2003). De acordo com Connor (1996, p. 5 apud Sáez, 2003), “(...) essa é uma área de pesquisa que identifica

problemas de composição encontrados por quem escreve em uma segunda língua e, referindo-se a estratégias retóricas da língua fonte, tenta explicá-las”.

Três são os princípios básicos em que a RC se baseia (Sáez, 2003, p. 5): (i) a linguagem e a escrita são fenômenos culturais; (ii) cada língua tem suas convenções retóricas e (iii) essas convenções retóricas e lingüísticas da língua materna interferem na escrita da língua de chegada.

Kaplan observou, em seu trabalho, por exemplo, que os textos se agrupavam em quatro blocos lingüísticos distintos, além do inglês: línguas árabes, semíticas orientais, românicas e russo. Com essas considerações, seu estudo estabeleceu como conclusões que: (i) os textos escritos em árabe e outras línguas semíticas se baseavam em uma série completa de construções paralelas, tanto positivas como negativas; (ii) os textos orientais se caracterizavam por uma abordagem indireta ao tema da redação, como um espiral; (iii) os textos das línguas românicas tinham como característica dispor de muito mais liberdade para digressão e introdução de assuntos diferentes do tema principal; e (iv) os textos de escritores russos eram compostos por uma série de construções paralelas e por um número indefinido de estruturas subordinadas irrelevantes para a idéia central do texto (Sáez, 2001).

Esse trabalho recebeu muitas críticas, foi acusado de etnocêntrico, de ter privilegiado a escrita dos falantes que tem o inglês como língua materna, de privilegiar o produto em detrimento do processo, de considerar a influência da primeira língua somente como uma transferência negativa para a segunda (Connor, 1996, p. 16). Não cabe fazer uma avaliação completa desse estudo aqui, o que se pode dizer é que algumas críticas têm razão de ser quando se pensa em escritura e ensino de segunda língua, mas, como estamos fazendo uma abordagem voltada à tradução, podemos olhar as críticas de maneira diferente. Por exemplo, interferências e transferências da língua de partida para a língua de chegada podem não ser relevantes em casos de textos em que elas não prejudiquem o julgamento, a aceitação e até

mesmo a interpretação do trabalho, mas, no caso do processo tradutório, estas são questões caras, que precisam ser levantadas e para as quais o tradutor precisa estar alerta no momento em que toma suas decisões.

Cabe dizer também que, nos anos 60, a investigação de Kaplan era ainda mal projetada, fruto da curiosidade e da necessidade de encontrar respostas para um então novo problema (Sáez, 2001). Mais tarde, porém, nos anos 80, novos estudos desenvolveram-se, como os de Connor e Kaplan (1987) e de Purves (1988). Estes utilizaram apenas textos escritos em primeira língua e passaram também a considerar a audiência leitora como peça fundamental na produção do texto.

Pelo que viemos desenvolvendo até aqui, acreditamos não ser difícil inferir que esta é uma área que pode contribuir para a tradução, tanto para a área teórica (tradutologia) como para a prática, uma vez que a tradução é um campo em constante, intrínseco e saudável conflito entre as línguas e suas culturas. Essa aproximação dos estudos da RC com a tradução é apresentada na seção seguir.

Antes disso, queremos fazer um esclarecimento quanto ao uso que estamos fazendo do termo “tradução”. É fundamental destacar que temos em mente, principalmente, o processo de tradução inversa, ou versão da língua portuguesa para a língua inglesa. Em um contexto de ensino de tradução, seria importante fazer esta diferenciação – entre tradução e versão. Hurtado Albir (1999, p. 23) comenta que “(...) um dos erros que se cometeu foi colocar ‘no mesmo saco’ a tradução direta e a tradução inversa, sem levar em conta que exigem habilidades distintas do indivíduo e que, portanto, merecem espaços didáticos e objetivos diferentes na formação de tradutores”. Concordamos com essa observação e acreditamos que este trabalho pode ser de algum auxílio para estudantes de tradução inversa, porém, não estamos fazendo distinção entre teorias sobre esses dois modos de traduzir, até porque

raramente se vê essa distinção na literatura. Mais ainda, aqui não nos preocupamos com o modo de traduzir de uma perspectiva qualitativa, mas descritiva, a partir do uso de *corpora*.

Por fim, citamos alguns objetivos apontados por Beeby (1996, p. 57 apud Albir, 2001) para a aprendizagem e a metodologia da tradução inversa, os quais devem ser buscados pelo tradutor ou pelo professor de tradução. Destacamos aqueles nos quais acreditamos que esta pesquisa e os estudos de RC possam contribuir:

1. saber redigir textos gramaticalmente corretos e pragmaticamente adequados na língua de chegada;
2. conhecer as diferenças tipográficas entre as duas línguas;
3. conhecer as diferenças lexicais entre as duas línguas, em certos campos e registros;
4. conhecer as diferenças sintáticas entre as duas línguas;
5. familiarizar-se com as tipologias textuais;
6. conhecer as diferenças discursivas entre as duas línguas: coerência e coesão textuais;
7. ampliar os conhecimentos das diferenças pragmáticas e semióticas entre as duas culturas;

Tal como se pode observar, há itens desses objetivos com os quais o levantamento dos marcadores e a sua comparação nas duas línguas, que apresentamos no Capítulo 6, podem contribuir.

A seguir, apresentamos algumas relações já estabelecidas entre a RC e os estudos de tradução, visando também justificar a escolha por mais este aporte teórico.

3.2.1 *Retórica Contrastiva e Tradução*

A hipótese central da RC, de que cada cultura desenvolve normas de escrita específicas e distintas, encerra uma enorme e potencial importância para a tradução, como vimos tentando mostrar desde a seção anterior. Acreditamos ainda que, dentro de cada cultura, as comunidades e os gêneros têm formas preferenciais de estabelecer seus textos e sua comunicação.

Existem trabalhos em RC voltados à tradução, como os de Enkvist (1991 e 1995); Nord (1991); Baker (1992) e Neubert-Shreve (1992). Segundo Fernández Polo (1999), que defendeu a tese **Traducción y Retórica Contrastiva: A Propósito de la Traducción de Textos de Divulgación Científica del Inglés al Español**, esses estudos têm uma afinidade maior com os pressupostos das teorias de tradução mais clássicas, de caráter mais prescritivo, no sentido que tratam de “(...) recomendar um comportamento tradutório específico como o único e mais adequado” (Fernández Polo, 1999, p. 72). No entanto, as teorias da tradução tiveram seu caráter um pouco modificado posteriormente, tanto no que concerne a seus pressupostos como na concepção de seu objeto de estudo e dos objetivos que persegue. Nesse ponto, a relevância da RC modifica-se, pois, enquanto nos estudos de Enkvist, e outros citados anteriormente, a RC traz respostas, aconselhando que a atitude correta do tradutor face a existência de diferentes convenções retóricas, é, indefectivelmente, a adaptação do texto às normas retóricas de língua de chegada; nos estudos mais modernos de tradução “a Retórica Contrastiva se concebe não tanto como uma fonte de recomendações e respostas, mas mais como fonte geradora de perguntas” (id., p. 72).

No trabalho de Fernández Polo (1999), formulam-se algumas dessas perguntas, que reproduzimos logo abaixo, e que também contribuirão para justificar nossa pesquisa, cuja principal motivação é, justamente, auxiliar na produção de um texto, a partir da tradução para o inglês, que não surpreenda por estranhezas de nível retórico, oriundas da interferência da

língua de partida, o português, mas que surpreenda, positivamente, pela qualidade com que é escrito, deixando que a ciência que ali acontece receba todas as atenções. Nesse ponto, estamos beirando uma discussão que também é importante para outras teorias de tradução: a questão da invisibilidade ou visibilidade do tradutor. Sobre esse assunto, limitamo-nos ao que mais interessa ao nosso objetivo: a invisibilidade textual. Benedetti (2003, p. 27) esclarece muito bem nossa visão e por que, na verdade, decidimos realizar um trabalho como este que apresentamos nesta dissertação:

“Na definição da invisibilidade textual (...) detecto em geral um paradoxo. É considerado típico da invisibilidade um tradutor produzir textos que não pareçam traduzidos, que sejam redigidos com tal fluência na língua de chegada que o leitor não tenha a impressão de estar lendo um autor estrangeiro, mas sim um texto escrito diretamente na sua língua. Assim, invisibilidade do tradutor equivaleria a não se enxergar o tradutor. Há uma imprecisão aí. Porque o que não se enxerga, na realidade, é o processo de tradução, é a própria tradução, e essa invisibilidade de tradução é na realidade consequência da maior interferência do tradutor sobre o texto, logo, de sua maior visibilidade como produtor textual (...): quanto mais fluente o texto, mais influente o tradutor”.

Das perguntas formuladas por Fernández Polo (1999) a respeito da contribuição dos estudos de RC para a tradução, destacamos duas que mais se relacionam a este trabalho. A primeira é “Qual deve ser o grau de adaptação, ou não, da tradução ou corpus de traduções às normas retóricas da língua de chegada?” e a segunda é “Que fator ou fatores influenciam a adaptação, ou não, às normas retóricas da língua de chegada? Consideramos que a decisão de adaptar, ou não, um texto à retórica da língua de chegada é do tradutor, e ela se dá em função de fatores contextuais, como tipo de texto, *status* do autor do original na língua de chegada e finalidade da tradução. Entendemos que, no caso da tradução dos artigos científicos do português para o inglês, é importante que não haja muitas influências da língua materna no texto traduzido, uma vez que, além desses textos serem geralmente submetidos a bancas

examinadoras, se deseja que a pesquisa ali apresentada chame a atenção, e, para tal, o texto deve ser claro e coeso, na língua em que se encontra.

Ainda sobre o que deve ajudar a nortear as decisões e o trabalho do tradutor de artigos científicos, tomamos posição semelhante a citada por Azenha (1999) quando faz referência a autores que consideram que o fio condutor do condicionamento das traduções é a função comunicativa do texto. Para esses, uma vez que a função comunicativa do texto de chegada estiver definida ela será “a bússula que deverá nortear toda estratégia de produção da tradução”. (id., p. 37) O tradutor, ao identificar essa função poderá, então, decidir se a mesma será preservada ou não. Será a partir dessa decisão que se estruturará toda a estratégia do tradutor, em que uma outra “rede de relações, desta vez na cultura de chegada, entre os elementos empregados no plano lingüístico e os novos valores culturais e características situacionais” será estabelecida (ibid., p. 40). Acreditamos que as expressões selecionadas para este estudo sejam partes da organização textual que refletem as convenções e valores culturalmente marcados, na expressão do texto, das línguas em questão. Por isto o nosso interesse em mostrar algumas das diferenças envolvidas nesses marcadores: sabemos que elas existem e que fazem diferença na tradução enquanto produto final, para os receptores.

Por fim, não podemos esquecer de fazer uma relação entre as considerações que fizemos sobre retórica e as que fizemos, no Capítulo 1, sobre gênero, pois percebemos que fatores relacionados a retórica e ao gênero são ambos intrínsecos e constituintes de um texto. Assim, podemos dizer que diferentes gêneros contam com elementos retóricos diferentes, porque seus propósitos e suas audiências são distintos. Anna Mauranen (1993), que desenvolveu um estudo sobre metatexto em textos de economia em finlandês e inglês, aponta que alguns elementos da retórica do discurso acadêmico podem ser dirigidos pelo gênero e outros pela cultura. Em textos de matemática, por exemplo, os elementos dirigidos pelo gênero seriam algo como: seja p igual a 3, e elementos dirigidos pela cultura seriam, por

exemplo, a tendência de os autores situarem a tese principal do texto mais no início ou mais no final do trabalho. Como vimos no Capítulo 2, as expressões que selecionamos para estudo são elementos dirigidos principalmente pelo gênero, pois se relacionam a três fatores envolvidos na escritura do artigo científico: o de organização textual (que chamamos de metarreferencial), o de seu propósito (apresentação de pesquisas, hipóteses, resultados) e o de desejo de validação e legitimação desses propósitos por parte de seus autores (fator subjetivo); porém defendemos que a maneira como são empregadas essas expressões é determinada pelos recursos que cada língua tem para constituir um gênero.

Salientamos que as expressões extraídas do *corpus* não são exclusivamente empregadas em artigos científicos, mas expressões da língua geral, que aparecem em outros textos também, e que se mostraram comuns a esse gênero. Por esse motivo, apoiamo-nos em um *corpus* de 333 artigos em português e 111 artigos em inglês, para comprovar que as expressões têm uma ocorrência significativa em textos desse gênero. O Capítulo 4 descreve os *corpora* analisados e o Capítulo 5, da metodologia, explicita como foram tratados.

CAPÍTULO 4: DESCRIÇÃO DO *CORPUS*

4.1 Introdução

Este capítulo irá descrever a constituição dos *corpora* em inglês e português, tentando, ao máximo, localizá-los em seu contexto de ocorrência, isto é, no ambiente onde se realizaram. Os textos foram selecionados a partir de diferentes fontes, as quais serão devidamente citadas e documentadas, porém, um critério teve que ser mantido: os textos deveriam estar em conformidade com o gênero em questão neste trabalho: o artigo científico, já definido no Capítulo 1.

De acordo com Swales (1990, p. 93), um artigo científico “(...) pode ser ou foi publicado em um periódico científico ou, menos tipicamente, em uma coletânea de artigos em formato de livro”. Os artigos selecionados preenchem um ou outro desses requisitos, como se procura deixar claro na descrição que apresentamos.

A primeira seção deste capítulo trata de descrever o *corpus* em português, que se constitui de textos apresentados nos Congressos da Sociedade Brasileira de Computação (SBC). O número de artigos de cada evento que constitui o Congresso da SBC também foi determinado.

A segunda parte do capítulo descreve o *corpus* em inglês, relatando brevemente as fontes e descrevendo-as, assim como os objetivos de cada periódico, além do número de artigos selecionados.

Esclarecemos que, embora tenhamos escolhido a grande área da Ciência da Computação para compor nosso *corpus* de análise, não fizemos uma avaliação específica da área, como viemos repetindo ao longo do texto, nos capítulos anteriores. Temos consciência de que a área em que um artigo se insere influencia no resultado do texto e nas unidades selecionadas para constituí-lo. Afirmar o contrário seria negar até mesmo conceitos que nos apóiam, como o de comunidade discursiva e esfera de comunicação. Sabemos, porém, que nesse momento era necessário confirmarmos nossos pressupostos iniciais, trabalhando apenas em uma mesma grande área. No entanto, podemos dizer também que, por serem os textos de nossos *corpora* realizados em diferentes subáreas da computação, das mais técnicas (*hardware*, robótica, *software*) às mais humanas (inteligência artificial, informática na educação) eles contemplam uma grande gama de artigos e de “subcomunidades”, tornando, assim, possível generalizá-los como representativos do gênero.

Mauranen (1993, p.5) também dá suporte a essa idéia, mencionando que

“(...) assim como faz sentido falarmos sobre a ciência em geral, também faz sentido falar sobre o gênero da ciência e da comunidade acadêmica como objetos culturais. Assim, podemos falar, por exemplo, do ‘artigo científico’ como um gênero no mundo da ciência, apesar do fato de disciplinas particulares diferirem de alguma maneira nas suas realizações convencionais. Se insistíssemos que a comunidade de pesquisa de cada disciplina tem seus próprios gêneros, perderíamos uma importante generalização no que diz respeito à atividade científica”.

A esse respeito, Bhatia (1993, p. 18) acrescenta que Swales, em sua análise de 48 introduções de artigos de pesquisa em diversas áreas, “(...) dá evidências do fato de que a introdução de um artigo científico em ciências é tão bom exemplo desse gênero como a de um artigo da psicologia ou da sociologia”.

4.2 O *corpus* em português

A SBC foi fundada em 1978. É uma instituição acadêmica que fomenta e desenvolve pesquisa científica na área da Computação. A SBC faz parte da Sociedade Brasileira para o Progresso da Ciência (SBPC) e da International Federation for Information Processing (IFIP). A instituição também é sócia do Centro Latinoamericano de Estudios en Informática (CLEI) e afiliada ao Institute of Electrical and Electronics Engineers (IEEE). A SBC é regida por um estatuto e administrada por uma diretoria e conta com um conselho, com funções deliberativas e normativas. A instituição acadêmica, com sede em Porto Alegre, no Rio Grande do Sul, tem sido fundamental na criação de conhecimento e tecnologia brasileiros. Os associados estão divididos nas categorias de fundadores, efetivos, honorários, estudantes e institucionais. Para disseminar e criar ambientes de pesquisa e inteligência nacionais, a entidade promove, todos os anos, diversos eventos, dentre eles, o Congresso Anual da SBC.

O Congresso da Sociedade Brasileira de Computação é considerado “(...) a grande reunião anual da comunidade científica de computação em nosso País”, como os próprios coordenadores o definem. Acontece, geralmente, em universidades brasileiras, dependendo do apoio e do esforço das mesmas. O Congresso é constituído por diversos eventos de distintas áreas de interesse dentro da Ciência da Computação, que contam com uma organização interna própria, sendo uns independentes dos outros. Cada um desses eventos estabelece as atividades que serão oferecidas e as oportunidades para apresentações de trabalhos, cartazes, palestras, *workshops* e cursos. O edital de chamada para participação e apresentação de trabalhos costuma acontecer com alguns meses de antecedência. Os trabalhos são submetidos à avaliação por parte de comitês de pesquisadores e professores reunidos para tal fim.

Os eventos que podem fazer parte do Congresso anual da SBC são: Seminário Integrado de *Software* e *Hardware* (Semish), *Workshop* sobre Educação em Computação (WEI), *Workshop* de Informática na Escola (WIE), Concurso de Teses e Dissertações (CTD),

Concurso de Trabalhos de Iniciação Científica (CTIC), Simpósio Brasileiro de Computação Musical (SBC&M), Encontro Nacional de Inteligência Artificial (ENIA), Jornada de Atualização em Informática (JAI), WPerformance (Workshop em Desempenho de Sistemas Computacionais e de Comunicação). Desses, alguns são bi-anuais, como mostra o quadro a seguir:

Quadro 4.1 - Eventos da SBC - 2001-2003

	2001	2002	2003
SEMISH	✓	✓	✓
WEI	✓	✓	✓
WIE	✓	✓	✓
CTD	✓	✓	✓
CTIC	✓	✓	✓
SBC&M	✓		✓
ENIA	✓		✓
JAI		✓	✓
WPerformance		✓	✓

Os temas escolhidos para os diferentes anos foram:

- 2001 - As tecnologias da informação e a questão social;
- 2002 - Convergências tecnológicas
- 2003 - Ciência, tecnologia e inovação: atalhos para o futuro,

A Tabela 4.1 traz um quadro-resumo do número de trabalhos selecionados para serem apresentados e terem seus artigos publicados nos anais do evento.

**Tabela 4.1 - Trabalhos apresentados e respectivos artigos publicados em eventos da SBC
– 2001, 2002, 2003**

	2001	2002	2003	TOTAL
SEMISH	10	15	15	40
WEI	16	14	16	46
WIE	36	67	52	155
CTD (1)	6	6	10	22
CTD (2)	12	9	10	31
CTIC	3	3	3	9
SBC&M	25		38	63
ENIA	63		56	119
JAI		8	9	17
WPerformance		15	11	26
TOTAL	171	137	220	528

(1) Doutorado. (2) Mestrado

4.2.2 Textos selecionados

Nem todos os 528 artigos desses três congressos mencionados foram selecionados. Os textos das JAI foram excluídos por se tratarem de capítulos mais longos de livros, fugindo, portanto, do gênero de texto que desejamos para esta pesquisa, e os textos do SBC&M também, por serem, em sua maioria, em inglês. Os textos de outros eventos que se encontravam redigidos em língua inglesa e os que não permitiram a alteração do formato do arquivo de “pdf” para o formato “txt” também não fizeram parte do *corpus*.

Assim, nosso *corpus* de pesquisa ficou composto como mostra a Tabela 4.2

Tabela 4.2 - Artigos de eventos da SBC que compõem o *corpus* em língua portuguesa – 2001, 2002 e 2003

	2001	2002	2003	TOTAL
SEMISH	7	10	8	25
WEI	14	11	11	36
WIE	34	65	60	159
CTD	-	-	-	-
CTIC	-	-	-	-
SBC&M	-	-	-	-
ENIA	44	-	50	94
JAI	-	-	-	-
WPerformance	-	11	8	19
TOTAL	99	97	137	333

O *corpus* em português é constituído, então, por 333 artigos, totalizando 931.719 palavras. A decisão de trabalhar com um *corpus* de quase um milhão de palavras justifica-se pelo fato de se estar buscando unidades recorrentes e típicas desse tipo de texto, independentemente da subárea em que se está trabalhando. Concluímos que elas se sobressairiam quanto maior fosse o número desses textos. De acordo com o que apresenta a literatura (Berber-Sardinha, 2004), a extensão de um *corpus* e sua representatividade comporta três dimensões:

- (i) o número de palavras – quanto maior o número de palavras, maior será a chance do *corpus* de conter palavras de baixa frequência, que formam a maioria das palavras da língua;

- (ii) o número de textos – quanto maior o número de textos, maior garantia de que esse gênero, registro ou tipo textual esteja mais adequadamente representado;
- (iii) o número de gêneros, registros ou tipos textuais – um número maior de vários tipos de textos permite uma maior abrangência do espectro genérico da língua.

Esse número de palavras que atingiu o nosso *corpus*, 1 milhão, é considerado por Leech (Leech, 1991 apud Berber-Sardinha, 2004) a taxa usual, sugerindo ser esse o patamar mínimo. Berber-Sardinha (2004) sugere uma classificação baseada na observação dos *corpora* que já foram utilizados em estudos de Linguística de *Corpus*, considerando um *corpus* de 250 mil a 1 milhão de palavras como **médio**.

4.2.3 *Eventos e subáreas que constituíram as fontes*

Embora os artigos estejam sob um grande guarda-chuva denominado “ciência da computação”, a amostra de artigos científicos que trazemos aqui é um tanto multidisciplinar, uma vez que incorpora diferentes subáreas de interesse dentro da ciência da computação, como podemos observar na descrição a seguir, em que fizemos uma breve caracterização de cada um dos eventos nos quais os textos de nosso *corpus* foram apresentados.

O Semish é considerado o evento mais antigo da comunidade acadêmica de computação no país. Acontece anualmente e visa apresentar à comunidade a pesquisa científica desenvolvida em termos de *hardware* e *software*. No ano de 2001, o Semish favoreceu trabalhos que incluíssem alguma aplicação social, seguindo a temática proposta pelo Congresso. Em 2002, os artigos aceitos deveriam privilegiar evidências da convergência tecnológica e o impacto na ciência da computação e em diferentes áreas do conhecimento, como Bioinformática, Educação e novas aplicações, e a apresentação dos trabalhos foi organizada de acordo com esses temas. Em 2003, foi privilegiada a apresentação de trabalhos

em que a integração entre *hardware* e *software* tivesse produzido artefatos de inovação tecnológica, ou seja, para a construção e o oferecimento de novos produtos e serviços.

O WEI é um evento anual que tem por objetivo debater os diferentes temas relacionados com o ensino de informática. O WEI é considerado um importante fórum de debate para professores, alunos, coordenadores de cursos e demais interessados no aprimoramento do ensino das diversas áreas da computação.

O Workshop de Informática na Escola (WIE) é um evento anual que tem por objetivo apresentar e debater os avanços e as experiências na utilização de informática na escola. É também objetivo do evento promover uma reflexão crítica sobre a aplicação da tecnologia de informática na educação no sistema educacional brasileiro. No ano de 2002, a apresentação dos trabalhos foi organizada em seis grupos temáticos: Gestão do uso escolar e parâmetros curriculares para as novas tecnologias de informática e computação; educação à distância e comunidades virtuais de aprendizagem; ambientes de apoio à aprendizagem cooperativa; realidade virtual e educação; inteligência artificial aplicada à educação; ambientes de aprendizagem nas Ciências e na Matemática.

O ENIA acontece nos anos ímpares. Tem como objetivo reunir pesquisadores, estudantes e profissionais que atuam na área de Inteligência Artificial (IA) para divulgar trabalhos de pesquisa, debater resultados e políticas de desenvolvimento e promover cursos de atualização na área. Como a Inteligência Artificial pode ser aplicada a diversas áreas, os trabalhos apresentados podem se relacionar aos seguintes tópicos de pesquisa:

- aplicações de inteligência artificial;
- aprendizagem e aquisição de conhecimento;
- automatização do raciocínio;
- fundamentos da inteligência artificial;
- IA distribuída e sistemas multi-agentes;

- lógica e inteligência artificial;
- descoberta de conhecimento e mineração de dados;
- planejamento, tempo e ação;
- processamento de linguagem natural;
- programação em lógica;
- redes neurais artificiais;
- representação do conhecimento;
- representação e raciocínio com incertezas;
- robótica;
- sistemas tutores;
- vida artificial e algoritmos genéticos;
- visão computacional.

O evento WPerformance aconteceu pela primeira vez em 2002. Esse evento surgiu para suprir uma ampla necessidade de fóruns adequados para que especialistas em tecnologia da informação, provenientes de universidades, empresas e instituições públicas e privadas, encontrem pesquisadores, educadores, gestores de TI e outros profissionais, para compartilharem suas experiências e resultados, discutindo temas relacionados ao desempenho e ao planejamento de capacidade desses sistemas.

O *corpus* em português utilizado nesta pesquisa constitui-se, como demonstramos, de textos publicados nos anais dos congressos da SBC de 2001, 2002 e 2003, abrangendo as áreas de desenvolvimento e aplicação de *software* e *hardware*, educação em computação, informática na escola, inteligência artificial (contemplando várias sub-áreas), e gestão das Tecnologias de Informação.

4.3 O *corpus* em inglês

O *corpus* em inglês é constituído por 111 artigos que, diferentemente dos textos que constituíram o *corpus* em português, foram extraídos de diferentes fontes. Como não encontramos um CD de um evento semelhante ao da SBC, com tanta diversidade, optamos por procurar artigos que fossem de áreas semelhantes às dos textos em português. Assim sendo, utilizamos recursos da internet para fazer a busca dos artigos. Uma das fontes utilizadas foi o Portal de Periódicos da Coordenação de Aperfeiçoamento e Pesquisa em Ensino Superior (CAPES), cujo acesso é disponibilizado às Instituições de Ensino Superior (IES). Através do portal, chegamos às publicações das revistas **Artificial Intelligence** e **Robotics and Autonomous Systems**. Para os textos na área de informática na educação, educação à distância e engenharia de *software*, utilizamos o *site* de busca Cite Seer e o *site* dos periódicos eletrônicos: **Canadian Journal of Distance Education** (CADE)) e o **Canadian Journal of Learning and Technology** (CJLT).

Os textos encontravam-se no formato “pdf” e tiveram que ser convertidos para o formato “txt”, o que acarretou um trabalho pesado de revisão e readaptação, pois muitos caracteres ficam prejudicados nessa conversão, bem como partes do texto são excluídas ou desordenadas. É necessário ainda mencionar que nem todos os textos permitem essa conversão e tiveram que ser, portanto, descartados do *corpus*.

Outro critério que utilizamos para a seleção dos textos que compuseram o *corpus* em inglês foram as instituições a que estavam vinculados os autores. Cuidamos para que todos os textos tivessem sido escritos a partir de universidades de países cuja língua nacional fosse o inglês. Como não era possível verificar a nacionalidade de cada autor, utilizamos esse critério para garantir que o texto havia sido escrito, no mínimo, em um contexto de língua inglesa, sendo, assim, muito provável que os textos não houvessem sido traduzidos, mas produzidos diretamente em inglês.

Em seguida, será apresentada uma descrição de cada fonte de consulta.

4.3.1 As fontes

A revista **Artificial Intelligence** está indexada no *site* de Periódicos da CAPES. É uma publicação da editora Elsevier, internacionalmente conceituada pela qualidade de suas publicações. Os artigos são publicados em papel e também estão disponíveis para acesso *online*, para quem tem acesso ao *science@direct*, conhecido como biblioteca digital.

Essa revista teve seu primeiro número publicado em 1970, e hoje é aceita como o local internacional para publicação de resultados de pesquisas nessa área. Especificamente, aceita artigos nas áreas de: raciocínio automatizado, teorias computacionais de aprendizado, busca heurística, representação do conhecimento, física qualitativa, robótica, linguagem natural, arquiteturas de *software* e *hardware* para Inteligência Artificial, dentre outras. Segundo a página da revista na internet, a mesma apresenta os resultados alcançados e as propostas para novas formas de olhar para os problemas da IA devem incluir demonstrações de sua eficácia. O público da revista é de especialistas nas áreas de IA pura e aplicada, cientistas da computação, cognitivistas, psicólogos e lingüistas.

Segundo as instruções para publicação na revista, são aceitos trabalhos básicos e aplicados que descrevam um trabalho maduro envolvendo processos computacionais de aspectos da inteligência. Sugere-se que os artigos com conteúdos muito matemáticos sejam precedidos por uma seção introdutória, menos técnica, e que seja acessível a uma ampla audiência e que os artigos exclusivamente matemáticos, sem aplicabilidade aos problemas de Inteligência Artificial, podem não ser aceitos.

A respeito das normas, segue o padrão de publicações internacionais, estabelecendo alguns pontos, como: que o artigo deve ser original, conter palavras-chave, serem submetidos em inglês; também estabelece critérios para a formatação das referências.

A revista conta com dois editores-chefe e 42 membros no conselho editorial, pertencentes a diferentes universidades, de diferentes países como Japão, Estados Unidos, Reino Unido, Alemanha, Israel, Itália, Eslovênia, Suíça, Dentre outras.

Os números selecionados para fazerem parte de nosso estudos foram:

- volume 148, (August, 2003), **Fuzzy Set and Possibility Theory-Based Methods in Artificial Intelligence;**
- volume 149 (Setembro, 2003), sem tópico definido;
- volume 151 (Novembro, 2003), **AI and Law.**
- volume 152 (Janeiro, 2004), sem tópico definido.

O *Canadian Journal of Distance Education* é o periódico da Associação Canadense de Educação à Distância. Essa associação tem um interesse bastante local, voltado para a promoção da educação à distância no Canadá, também representando internacionalmente este país em assuntos voltados para a educação à distância. Os objetivos dessa publicação são promover e encorajar a pesquisa e o interesse em educação a distância no Canadá. O periódico contém material dividido em três grandes categorias: artigos de pesquisa que enfoquem questões relacionadas à educação à distância; relatos que enfatizem a solução de problemas críticos, e notas sobre pesquisas em andamento; e artigos voltados para o debate e a discussão de questões em educação à distância que pareçam controversos.

São publicados dois números por ano na forma impressa, e, um mês após a publicação, os textos são disponibilizados no *site*. A publicação é feita pela Faculdade de Educação da Universidade de Alberta, em Edmonton.

O periódico conta com dois editores-chefe e um conselho de seis pesquisadores de universidades na Austrália, Suécia, Reino Unido, Índia, Estados Unidos e Canadá.

A respeito das normas de publicação, seguem os padrões já mencionados, como ineditismo do trabalho e uma lista de itens voltados a tamanho e à formatação dos originais, aceitos tanto em francês como em inglês. Os manuscritos são revisados por pelo menos dois consultores, sendo que o nome do autor e a instituição não são revelados aos revisores. Os critérios de revisão, segundo as normas, incluem a validade dos argumentos e das conclusões, a clareza na organização de idéias, a competência de descrição metodologia, dentre outras.

Os números selecionados para constarem em nosso *corpus* foram:

- volume 17, n. 2 e 3 (2002)
- volume 18, n. 1 e 2 (2003)

Outra fonte de artigos científicos em inglês foi o **Canadian Journal of Learning and Technology**. Esse periódico é publicado pela Associação para Mídia e Tecnologia em Educação no Canadá, destinado à comunidade de educadores, produtores de mídia, pesquisadores, bibliotecários e outros profissionais que trabalhem para facilitar e melhorar o aprendizado, em todos os setores de educação, através da aplicação apropriada de tecnologia educacional. A associação compromete-se com a pesquisa, o desenvolvimento, o fomento, o apoio, a avaliação e a disponibilização de tecnologias educacionais que reflitam uma teoria pedagógica saudável e acessível a aprendizes e professores que promovam a inovação no ensino e no aprendizado.

O periódico recebe artigos relacionados à tecnologia de ensino e aprendizado. Os tópicos de interesse incluem: teoria e tecnologia de aprendizado, cognição e tecnologia, aprendizado *online*, informática aplicada à educação, simulações e jogos e outros aspectos do uso de tecnologia no processo de aprendizado.

Os manuscritos podem ser submetidos à comissão editorial tanto em inglês como em francês, podendo ser artigos científicos, revisões de literatura, estudos de caso, dentre outros.

Os artigos selecionados para o *corpus* foram apenas os artigos científicos, os quais podem incluir pesquisas quantitativa e qualitativas. Segundo as normas de publicação, esses artigos descrevem projeto e metodologia e apresentam os resultados de um estudo sobre uma inovação ou a aplicação de uma teoria que facilite o aprendizado com o uso da tecnologia. Tipicamente, contêm a colocação de um problema ou a descrição de uma questão a ser explorada, revisão da literatura, metodologia, resultados do estudo e implicações para futuros trabalhos.

Todos os manuscritos recebidos pelo editor são julgados por um painel de revisores anônimos que avaliarão adequabilidade, contribuição, acurácia e interesse.

Há também normas para a preparação do formato dos manuscritos, incluindo tamanho de página, letra, formato em que deve aparecer o título e outras, como nos outros dois periódicos já mencionados.

Esse periódico é publicado tanto em papel como *online*. Os números selecionados para constarem no *corpus* foram:

- volume 28 (2002)
- volume 29 (2003)

A revista **Robotics and Autonomous Systems** foi acessada através do Portal de Periódicos da CAPES. É uma publicação que se assemelha à **Artificial Intelligence**, também publicada pela Editora Elsevier e está afiliada à Intelligent Society of Autonomous Systems e Inteligentes (IAS).

A revista, que teve seu primeiro número publicado em 1985, propõe-se a publicar artigos que descrevam produções fundamentais nas áreas de robótica e sistemas autônomos, com ênfase nos sistemas autônomos, em seus aspectos teóricos, computacionais e experimentais, ou módulos de tais sistemas.

Os ambientes de aplicação que interessam incluem o ambiente industrial, ao ar livre e no espaço, onde as técnicas avançadas de robótica são necessárias para que os sistemas autônomos realizem tarefas sem a intervenção humana, incluindo-se aí o uso da robótica em ambientes perigosos e hostis. Alguns dos tópicos que a revista inclui são: controle do comportamento do robô, aplicações industriais de sistemas autônomos, técnicas robustas de Inteligência Artificial e sensoriamento, como modelagem da incerteza, programação de robôs, dentre outros.

A revista conta com três editores-chefe e um conselho editorial composto de 26 membros, também pertencentes a diferentes universidades de renome em diferentes continentes. Sobre as instruções para os autores, também segue o padrão já mencionado acima, estabelecendo critérios para formatação e submissão. A revista é publicada em papel e é possível acessá-la *online* pelo site de *science@direct*.

Os números selecionados para inclusão no *corpus* foram:

- vol 42 (2003);
- vol 44 (2003), incluindo os melhores artigos de duas conferências eurobot '01 *workshop* e *ias-7*;
- vol 45 (2003).

O **CiteSeer** foi outra fonte de coleta de textos. É um *site* de buscas de artigos científicos na internet e está disponível no endereço **www.citeseer.com**.

Nesse *site*, pesquisamos o diretório de artigos em Ciência da Computação, nos itens Engenharia de Software e Informática na Educação. O *site* retorna diretamente os links para artigos já publicados de revistas indexadas na *Web*.

Com o uso desse *site* de busca, os artigos retornados são de diferentes revistas, por isso não as detalhamos. Contudo, como os artigos retornados estão em uma ordem de “mais

citados” na bibliografia de outros trabalhos, consideramos que sua qualidade é garantida, pois, se são constantemente mencionados em outros trabalhos, respondem aos propósitos de transmitir a informação com qualidade e acurácia.

A seguir, apresentamos uma tabela geral a respeito das fontes de onde foram extraídos os textos e o número de textos selecionados, que totalizaram 872.182 palavras.

Tabela 4.3 - Fontes para a composição do *corpus* em língua inglesa

FONTE	NUMERO	ANO	NUMERO DE TEXTOS
Artificial Intelligence	148	2003	2
	149	2003	10
	151	2003	5
	152	2004	3
			20
Canadian Journal of Distance	17 – v. 3	2002	4
Education	17 – v. 2	2002	3
	18 – v.1	2003	2
	18 – v. 2	2003	1
			10
Canadian Journal of Learning and	28 - v.1	2002	3
Technology	28 - v.2	2002	5
	28 - v.3	2002	8
	29 – v.1	2003	4
	29 – v.2	2003	4
			24

Robotics and Autonomous	42	2003	9
Systems	44	2003	10
	45	2003	3
			22
CiteSeer	Informática		15
	na educação		
	Software		20
	Development		
<hr/>			
TOTAL			111
<hr/>			

Como mostra a Tabela 4.3, o *corpus* em inglês constitui-se de um número três vezes menor de textos do que o *corpus* em português. Consideramos que, por utilizarmos o *corpus* em inglês não para levantamento dos marcadores, mas como elemento de comparação, o fator tamanho não chega a afetar os resultados. Precisávamos de um *corpus* em português o mais extensivo possível, dentro de nossas limitações de tempo para ajustes e coleta dos textos, pois ele seria a ponto de partida, ele nos mostraria quais seriam as expressões que se mostrariam típicas dos artigos científicos. Assim, quanto mais textos, mais as unidades comuns aos vários textos se mostrariam. O *corpus* em inglês já foi consultado a partir das unidades evidenciadas pelo *corpus* em português. Como, para a comparação, foi utilizado um percentual, acreditamos que um maior número de textos em inglês não alteraria os resultados.

No Capítulo 5, a seguir, apresentamos a metodologia utilizada para o tratamento dos *corpora*: alguns pressupostos Lingüística de *Corpus*.

CAPÍTULO 5: METODOLOGIA

5.1 Introdução

As unidades que são objeto de investigação nesta dissertação foram extraídas de um *corpus* de artigos científicos em português e outro de artigos científicos em inglês (ver Capítulo 4 para uma descrição detalhada dos *corpora*). Utilizamos pressupostos e ferramentas empregadas pela Linguística de *Corpus* (LC) para o manuseio, preparação e escolha dos textos. A LC também auxiliou a validar o tipo de trabalho que fizemos, que realiza avaliações quantitativas e qualitativas de componentes textuais, porém, entendemos que este não é um trabalho que segue todos os princípios dessa área, da qual mais utilizamos o ferramental.

5.2 A Linguística de *Corpus*

Iniciamos esta seção transcrevendo uma citação de Amparo Hurtado Albir, a respeito da importância e necessidade de estudos de gênero em tradução e que, ao final, aponta para a realização desses estudos utilizando *corpora* eletrônicos.

“Necessitamos estudos empíricos que compilem *corpus* de textos, amplos e representativos, em diferentes âmbitos e em diversas línguas; somente a estatística permitirá estabelecer regularidades confiáveis. Estes estudos permitirão, em primeiro lugar, a classificação em gêneros e sub-gêneros, assim como a identificação de padrões estruturais e regularidades no uso das convenções de cada língua. A confrontação posterior com *corpus* em outras línguas permitirá identificar as diferenças e assim abrir caminho para a confecção de que poderíamos denominar dicionários de gêneros. A compilação de

corpus formado por originais e traduções é também de interesse para a descrição da maneira como funcionam as traduções e para analisar se ao traduzir se segue ou não as convenções estruturais e lingüísticas próprias da língua de chegada. Neste ponto, a metodologia de corpora eletrônicos abre grandes possibilidades.” (Hurtado Albir, 2001, p. 505)

É baseando-se nessa citação, e reforçando que acreditamos que o uso de *corpora* eletrônicos fornece evidências e dados de muito valor para a descrição da língua, que introduzimos aqui alguns conceitos da LC. Antes, ainda, é preciso ressaltar outro ponto nessa citação que entra em concordância com o nosso ponto de vista e com nossas motivações, o fato de que o tradutor precisa conhecer as estruturas e convenções das línguas e textos em que trabalha, para poder fazer suas escolhas com mais fundamento e até liberdade. É com isso que queremos contribuir ao realizarmos esta pesquisa.

Podemos começar a tratar da LC pela citação acima, em que a autora fala em uma “metodologia de *corpora* eletrônicos”. Em se tratando de LC, a questão de ser ela uma metodologia, uma disciplina ou uma maneira de abordagem é um tópico de debate no estabelecimento do status dessa área.

Para Berber-Sardinha (2004, p. 35), pesquisador que atua na área de LC no Brasil há mais de dez anos, a LC não poderia ser considerada uma disciplina assim como a psicolingüística ou a semântica, pois seu objeto de pesquisa não é delimitado como em outras áreas. Ela não se ocupa de um assunto definido, mas de vários fenômenos comumente enfocados em outras áreas, como o léxico e a sintaxe. Pensar na LC como uma metodologia depende da definição que tomarmos de metodologia. Se pensarmos na LC como uma metodologia no sentido de ferramental a serviço de outras disciplinas, é possível que se mantenha a orientação teórica de cada área, separadamente do ferramental. Assim, teríamos, por exemplo, a sintaxe baseada em *corpus* e a sintaxe tradicional (id., 2004).

A LC, no entanto, não se resume a um conjunto de ferramentas. Assim, se entendermos **metodologia** como a aplicação de um “conjunto de pressupostos de caráter

teórico, então a LC pode ser vista como uma metodologia, pois traz consigo algo mais do que simplesmente o instrumental computacional” (Berber-Sardinha, 2004, p. 36). Nessa definição se encaixam as pesquisas acerca do comportamento do léxico, em uma doutrina de não-categorização *a priori*, como as desenvolvidas por John Sinclair na Inglaterra. Há ainda outra razão para que a LC não seja pensada apenas como uma metodologia: seus praticantes produzem conhecimento novo, revelando fatos a respeito da linguagem que nunca se pensou em procurar e que nem se mostrariam se não fosse por esse meio.

Assim, nem totalmente metodologia e nem disciplina, surge uma terceira possibilidade: uma perspectiva, ou uma abordagem (esse termo é o utilizado por Douglas Biber, em inglês, *corpus-based approach*). Seria uma maneira de se chegar à linguagem, fazendo uma alusão ao conceito de uma teoria cuja janela molda o modo como enxergamos a linguagem (Berber-Sardinha, 2004, p. 37), seria uma espécie de lente.

Quanto a este trabalho, sem querermos avançar na questão do status da área, preferimos dizer que é uma pesquisa *baseada em corpus*, que utilizou suas ferramentas para comprovação de um pressuposto estabelecido a partir de observações dos textos. Este é um trabalho que utiliza, principalmente, as ferramentas da LC, e se não fosse por elas, não seria possível lidar com tamanha quantidade de textos. É preciso acrescentar que um trabalho em LC se desenvolveria ainda, talvez, por caminhos diversos do que fizemos aqui, pois talvez fosse necessário maior rigidez quanto a, por exemplo, tamanho do *corpus*, etiquetagem de todas as palavras em um software do tipo *parser*, e uso de métodos estatísticos para contagem das frequências das expressões. Porém, a despeito desses fatores, cabe dizer que, independentemente da diversidade de denominação e métodos, os trabalhos em LC têm características compartilhadas, e se as observarmos, a seguir, veremos que este trabalho compartilha delas também (Berber-Sardinha, 2004, p.38):

- (i) são empíricos e analisam os padrões reais de uso em textos naturais;

- (ii) utilizam coletâneas grandes e criteriosas de textos naturais, conhecidas por *corpus*, como a base da análise;
- (iii) fazem uso extensivo de computadores na análise, empregando técnicas automáticas e interativas;
- (iv) dependem de técnicas quantitativas e qualitativas;

Assim, após essa breve discussão sobre o *status* da área, acreditamos que cabe esclarecer um pouco mais o que vem a ser a LC, seu desenvolvimento e seu principal componente: o *corpus*.

Podemos dizer que essa denominação Lingüística de *Corpus* é bastante recente, surge nos anos 80, com a chegada dos microcomputadores pessoais, mas os estudos com *corpus* acontecem desde longa data. Durante boa parte do século XX, houve muitos pesquisadores que se dedicaram à descrição da linguagem por meio de *corpora*, mas os *corpora* não eram eletrônicos e a ênfase desses trabalhos era, em geral, o ensino de línguas (Berber-Sardinha, 2004). Hoje, o que prepondera nos estudos de LC é a descrição da linguagem. As aplicações dos *corpora* dentro de sub-áreas da lingüística aplicada são diversas, tais como na produção de dicionários e gramáticas, na lingüística voltada ao estudo de ideologias, nos estudos tradutológicos, literários e estilísticos, na lingüística forense e no desenvolvimento de material de apoio na produção de textos específicos (Hunston, 2002).

Uma definição considerada na literatura como bastante completa do que vem a ser, em toda sua potencialidade, um *corpus*, foi dada por Sanchez (1995, p.8-9, apud Berber-Sardinha, 2004, p. 18)

“É um conjunto de dados lingüísticos (pertencentes ao uso oral ou escrito da língua, ou a ambos), sistematizados segundo determinados critérios, suficientemente extensos em amplitude e profundidade, de maneira que sejam representativos da totalidade do uso lingüístico ou de algum de seus âmbitos, dispostos de tal modo que possam ser

processados por computador com a finalidade de propiciar resultados vários e úteis para a descrição e análise”.

Tratando separadamente seu objeto e objetivo, em linhas gerais, diz-se que a LC ocupa-se da coleta e exploração de *corpora* ou de um conjunto de dados lingüísticos textuais, coletados criteriosamente, com o propósito de servirem para a pesquisa de uma língua ou de uma variedade lingüística. Como tal, dedica-se à exploração da linguagem através de evidências empíricas, extraídas por meio de computador (Berber-Sardinha 2004, p. 3). A escolha dos textos que constituem um *corpus* que será submetido à análise e a observações não é aleatória, isto é, nem todo conjunto de textos é considerado um *corpus*. Estes devem ser textos reais e naturalmente produzidos, e não criados com a propósito de figurarem no *corpus*. Atente-se, no entanto, para o fato de que, embora os textos sejam naturais, isto é, realmente produzidos em condições normais – e não induzidas – por falantes de uma comunidade, diz-se que o *corpus*, isto é, o conjunto dos textos, é artificial. Isto quer dizer que os textos são reunidos com fins específicos de pesquisa, assim, nem todo conjunto de textos poderá ser considerado um *corpus*, pois a reunião dos textos será delimitada de acordo com os objetivos da investigação, podendo ser o *corpus*, por exemplo, constituído de textos jornalísticos, ou científicos, ou então de textos especializados em geral. Pode-se falar ainda em um *corpus* de registros dialetais, ou multilingüe, composto de transcrições de fala, ou apenas de textos escritos.

É importante dizer que, estritamente falando, um *corpus* não pode fazer nada sozinho, sendo assim nada mais é do que um depósito de língua utilizada. Os softwares de acesso aos *corpora* é que reorganizam esse depósito de maneira que diferentes observações possam ser feitas. A maioria dos pacotes de software disponíveis mostra os dados contidos em um *corpus* de três maneiras: mostrando frequência, fraseologias e colocações (Berber-Sardinha, 2004). A avaliação de um *corpus* por meio desses programas pode nos mostrar como a língua é, e o

principal argumento a favor de um *corpus* é que ele é um guia mais confiável de uma língua do que a intuição de um falante nativo, que muitas vezes não consegue explicar porque “sente” que algumas coisas na sua língua são como são (Hunston, 2002).

Como mencionamos anteriormente, existem diversas áreas que ultimamente vêm se valendo da aplicação dos *corpora*, duas das mencionadas que nos interessam nesse trabalho são a tradução e o desenvolvimento de material de apoio para quem precisa escrever textos. A seção a seguir descreve um pouco mais essas duas áreas e também as aproxima da pesquisa que desenvolvemos aqui.

5.2.1 A LC aplicada

A LC pode ser aplicada em diferentes áreas dentro dos estudos da linguagem, do ensino de línguas à descrição lingüística pura. Neste trabalho, estamos tratando de questões relacionadas a convenções de textos de artigos científicos, no sentido de auxiliar a sua tradução. Para tal, entendemos que era preciso compreender e apreender essas unidades convencionais nas duas línguas em questão, português e inglês, para depois podermos fazer avaliações de seus usos, freqüências e colocações. Assim, esse é um estudo classificado como *baseado em corpus*, pois consiste em retirar exemplos e verificar hipóteses a partir dos conjuntos de textos, diferente do que seria um trabalho *dirigido por corpus*, que partiria para a pesquisa com poucos pressupostos e esperaria pelas evidências que este proporcionaria

Essa descrição do uso real da língua, tão característica dos estudos em LC, proporciona também outros tipos de dados e informações de relevância para o estudo e prática de tradução, como sugeria Hurtado Albir no início desse capítulo. Susan Hunston (2002), pesquisadora de um dos centros de excelência em LC na Inglaterra, menciona que, em questões práticas, o uso do *corpora* na tradução pode ser resumido em duas questões: (i) que software pode ser desenvolvido que possibilite ao tradutor explorar *corpora* como uma

ferramenta de auxílio na tarefa diária de traduzir?, e (ii) o que um corpus constituído de textos traduzidos indica sobre o próprio processo de tradução? Ainda, não podemos esquecer que uma outra área de aplicação dos estudos com *corpora* que afeta o trabalho do tradutor é o desenvolvimento de materiais de referência, como glossários e dicionários. É evidente que um material de referência que contenha, por exemplo, colocações e exemplos de uso real na língua auxilia, em muito, a tarefa de traduzir.

Um estudo feito com *corpus* acerca da condição da tradução, por exemplo, foi realizado por Baker (1993), que sugere que algumas características de textos traduzidos podem ser melhor percebidas com o uso de *corpora*. Essa autora observa, por exemplo, que as traduções, em geral, são mais explícitas do que os textos de partida, incluindo informações que este omite, assim, encontram-se no texto traduzido expressões gramaticais que se prestam a essa função como, por exemplo, em inglês, as orações relativas com o uso de *that*, ao invés de sem ele. Ou então, percebe-se que, no texto traduzido, as ambigüidades são reduzidas, substituindo-se pronomes por referências mais explícitas aos sujeitos.

Além desse tipo de conclusões acerca das traduções, o uso de uma grande quantidade de textos, representando o uso da língua de uma determinada cultura, pode ser de extrema utilidade para quem precisa escrever em uma língua estrangeira. Por exemplo, “para muitos especialistas, não é a terminologia técnica, mas o que se pode chamar de terminologia da retórica que lhes causa problemas” (Hunston, 2002, p. 135). De acordo com essa autora, a análise de *corpora* compostos de tipos particulares de textos pode ser utilizada para identificar palavras e sentenças vinculadas a determinadas partes do texto. Embora voltado à tradução, nosso trabalho também foi permeado por essa preocupação do “bem escrever” em uma língua estrangeira. Também ao tradutor o conhecimento de uma terminologia às vezes não é o que mais lhe causa aflição, mas, sim, como escrever esse texto em que as terminologias se inserem.

5.3 Tratamento do *corpus*

Para que conseguíssemos apreender as ocorrências das unidades que nos interessavam, e comprovar que elas seriam típicas do gênero em questão, foi preciso uma abordagem de duas vias, quantitativa e qualitativa, que não foram aplicadas em separado. O critério quantitativo é utilizado nos dois momentos de análise desta pesquisa, tanto estabelecendo o número mínimo de ocorrências que uma determinada expressão deveria ter para constituir-se como *típica*, no *corpus* em português; como após o levantamento, servindo de critério para avaliação de preferência de uso em cada uma das línguas. O critério qualitativo foi utilizado, juntamente com o critério quantitativo, para selecionar as unidades que fariam parte de nossa análise, uma vez que, ao realizarmos uma lista de frequência, tínhamos o retorno de diversas unidades, não apenas das que interessavam para a pesquisa, como explicitaremos a seguir. Esses e outros tópicos relacionados à metodologia e trabalho com o *corpus* serão descritos nas seções a seguir.

5.3.1 As ferramentas utilizadas

O *Wordsmith Tools v.3* é um programa de computador utilizado para a operacionalização da descrição lingüística com *corpus*. Como explica Berber-Sardinha (2004, p. 86):

“o programa coloca à disposição do analista uma série de recursos que, bem usados, são extremamente úteis e poderosos na análise de vários aspectos da linguagem, como a composição lexical, a temática de textos selecionados e a organização retórica e composicional de gêneros discursivos”

Esse programa é composto de ferramentas, utilitários, instrumentos e funções. Dessas, utilizamos apenas duas ferramentas: a *WordList* e a *Concord*. A *WordList* propicia a criação de listas de palavras, que podem ser disponibilizadas em ordem alfabética ou pela ordem de

freqüência. Essa ferramenta também fornece dados estatísticos como número de palavras total (*tokens*), número de palavras diferentes (*types*) e outras estatísticas. É possível, também, definir o número de palavras que desejamos que um agrupamento tenha, os chamados *clusters*. A ferramenta *Concord* “produz concordâncias ou listagens das ocorrências de um item específico (chamado palavra de busca ou nóculo, que pode ser formado por uma ou mais palavras) acompanhado do texto ao seu redor” (Berber-Sardinha, 2004, p. 105).

A seguir apresentamos como utilizamos essas ferramentas, trazendo exemplos para tornar mais acessível a compreensão de seu uso.

5.3.2 Tratamento do corpus em português

Nesta pesquisa, primeiramente, trabalhamos com o *corpus* em português. Os 333 textos foram transformados automaticamente de seu formato de arquivo original “pdf” para “txt”, utilizando-se o software Vervypdf.com Inc. Essa operação era necessária pois o software *Wordsmith Tools v. 3* não trabalha com o formato “.pdf” de arquivo.

O primeiro passo dado foi a realização de uma lista de freqüência dos agrupamentos (*clusters*) de 4 palavras, ou seja, em todo o *corpus*, quantas vezes uma determinada seqüência de 4 palavras se repetia. Para essa tarefa, foi utilizada a ferramenta *Wordlist* do software *Wordsmith Tools*. O objetivo era encontrar resultados que nos levassem à apreensão dos marcadores textuais típicos do artigo científico. A idéia foi de não utilizar o *corpus* para buscar unidades que já conhecíamos como típicas dos artigos, mas deixar que os textos demonstrassem por si as unidades que lhe eram particulares. Essa primeira ação resultou em 29.554 ocorrências, com freqüência de repetição dos *clusters* de 127 a 2. Apresentamos, a seguir, os 20 primeiros agrupamentos, a título de exemplificação:

N	Word	Freq.
1	de ciência da computação	148
2	de acordo com a	137
3	em sala de aula	135
4	com o objetivo de	118
5	para o desenvolvimento de	104
6	de acordo com o	103
7	de informática na educação	102
8	do ponto de vista	102
9	de sistemas de informação	100
10	de computação e informática	87
11	de educação a distância	86
12	a partir de um	85
13	de # a #	82
14	de acordo com as	77
15	a figura # apresenta	74
16	de um conjunto de	69
17	em ciência da computação	67
18	a seção # apresenta	64
19	para a construção de	64
20	a figura # mostra	61

Analisamos os agrupamentos que tiveram de 148 a 5 ocorrências em todo o *corpus*, o que totaliza 3449 ocorrências. Dessas, nem todas satisfaziam o interesse da pesquisa, por exemplo, os nomes das universidades que estavam no cabeçalho dos artigos. Foram, então,

selecionadas cerca de 400 ocorrências que consideramos poderiam ser parte de ou serem as próprias unidades a serem estudadas. Essa seleção foi feita qualitativamente, utilizando o conhecimento que tínhamos do gênero de texto em questão e a própria intuição de falantes de língua portuguesa (aquela intuição de que falamos no início do Capítulo 1). Utilizando os exemplos da lista acima apresentada, selecionamos as sublinhadas:

N	Word	Freq.
1	de ciência da computação	148
2	<u>de acordo com a</u>	137
3	em sala de aula	135
4	<u>com o objetivo de</u>	118
5	<u>para o desenvolvimento de</u>	104
6	<u>de acordo com o</u>	103
7	de informática na educação	102
8	<u>do ponto de vista</u>	102
9	de sistemas de informação	100
10	de computação e informática	87
11	de educação a distância	86
12	<u>a partir de um</u>	85
13	de # a #	82
14	<u>de acordo com as</u>	77
15	<u>a figura # apresenta</u>	74
16	de um conjunto de	69
17	em ciência da computação	67
18	<u>a seção # apresenta</u>	64

19	para a construção de	64
20	<u>a figura # mostra</u>	61

É preciso ter uma noção de que classificar cerca de 400 ocorrências, das quais várias muito particulares, e verificar de que maneira ocorriam nos textos (quais os verbos que as acompanhavam, se estavam em início ou fim de frase, se realizavam a mesma função, se apresentavam variações de singular e plural, masculino e feminino, se os verbos eram conjugados) não é tarefa simples. Por exemplo, para a expressão *de acordo com a*, verificamos se era sempre seguida do artigo feminino, se não, quais eram os artigos, definidos ou indefinidos, no singular ou no plural, que ocorriam com essa expressão, e, se não fosse seguida de artigo, que outro elemento a constituía. Para a expressão *a figura # apresenta*, verificávamos se ocorria em outros tempos verbais, ou com outros verbos.

Para essa tarefa, utilizamos a ferramenta Concord do WordSmith Tools e passamos em revista uma a uma das 400 ocorrências, pois sabíamos também que poderiam existir expressões com mais do que as quatro palavras definidas para iniciar a pesquisa. Com a ferramenta Concord foi possível resgatar todas as ocorrências de uma palavra ou expressão no *corpus*, verificando suas colocações e frequências.

As ocorrências foram então sendo agrupadas de acordo com semelhanças na função que desempenhavam no texto, por conterem a mesma palavra chave ou por semelhanças sintáticas. Após esse agrupamento por função ou palavra-chave, percebeu-se que podíamos sistematizar a classificação ainda mais, e aí chegamos a classificação nos três grandes grupos, que posteriormente encontraram apoio do modelo sistêmico-funcional que demonstramos no Capítulo 2. Foi com grande satisfação que descobrimos que a nossa classificação empírica entrava em conformidade com esses pressupostos postulados por Halliday (1985) sobre as macrofunções da linguagem.

Por fim, calculou-se um percentual de ocorrências nos textos. Assim, por exemplo, a expressão *a figura # apresenta*, teve um percentual de 22.2%. Esse percentual é de quantas vezes a expressão apareceu no total de 333 textos, por esse motivo, às vezes o percentual ultrapassa os 100%, significando que tal expressão ocorreu mais de uma vez em alguns textos.

5.3.3 Tratamento do corpus em inglês

Com os textos em inglês realizou-se procedimento semelhante. Os textos foram convertidos automaticamente de seu formato “.pdf” para “.txt”. Depois, fizemos a lista de frequência de agrupamentos de 4 palavras, utilizando a ferramenta *WordList*. Em inglês, tivemos um resultado de 34081 agrupamentos, que foram de uma frequência de 489 a 2, dessas, selecionamos as que tinham ocorrência até 5 no *corpus*, o que resultou um total de 2215 ocorrências.

Fez-se a mesma avaliação qualitativa, selecionando-se as unidades que interessavam para a pesquisa. Estas foram também analisadas com auxílio da ferramenta *Concord*, classificadas em uma das três categorias já estabelecidas. As frequências de uso foram calculadas para efeito de comparação com as unidades em português.

Criamos, então, quadros comparativos, que podem ser vistos no Capítulo 6, a seguir, onde foram observadas semelhanças e diferenças de uso dessas expressões em inglês e português.

CAPÍTULO 6: ANÁLISE COMPARATIVA DOS MARCADORES TEXTUAIS DO ARTIGO CIENTÍFICO EM PORTUGUÊS E INGLÊS

6.1 Introdução

Neste capítulo, apresentamos a análise comparativa que fizemos dos marcadores textuais, mostrando uma equiparação de usos e frequências dos mesmos em português e inglês. É preciso mencionar que o que sugerimos nesta comparação visa, em primeiro lugar, a uma reflexão por parte do tradutor, ou de quem irá escrever um texto em inglês, sobre como irá traduzir/utilizar esses marcadores no texto do artigo científico. Não estamos sugerindo a tradução ideal, nem a forma mais correta, mas, sim, mostrando possibilidades de escolha, através do retrato do uso dessas expressões nos *corpora* aqui selecionados. Não estamos, também, defendendo que algumas formas devem ser usadas em detrimento de outras, mas acreditamos que o critério da frequência pode ajudar o tradutor a tomar decisões, a buscar saídas e alternativas que ajudem a qualificar a sua tradução.

A análise está, num primeiro momento, descrita sob a forma de quadros sistematizadores nos quais são apresentados os marcadores identificados e sua frequência de uso em inglês e português. Como o resultado do levantamento informatizado mostrou um grande número de expressões, selecionamos, para comentário, o que parecia ser mais relevante na comparação, além do que já fica evidente quando se compara apenas o critério de frequência de uso de cada marcador.

Ao final deste capítulo, buscamos sistematizar alguns dados recorrentes, tanto no que tange a especificidades das três grandes categorias, definidas em correspondência com as macrofunções da linguagem, como no que concerne à comparação dos marcadores nas duas línguas em questão. Essas generalidades prestam-se mais a uma apreensão geral do que encontramos, pois acreditamos que o mais importante, neste trabalho, é a compilação e as comparações desses marcadores, em que é possível visualizar questões envolvidas no seu uso nas duas línguas (colocações, escolhas lexicais, fraseologismos e preferências de uso), que, pretendemos, possam ser de auxílio para o tradutor.

6.2 Marcadores Textuais em comparação

No Capítulo 2 definimos o que entendemos por marcadores textuais e apresentamos as categorias em que os classificamos, com o acréscimo de alguns exemplos. Os marcadores textuais são, neste trabalho, expressões de mais de uma palavra que realizam uma função de estruturação e organização do texto, o que lhes dá um caráter de elementos realizadores da coesão textual. Incluímos, também, expressões em que o sujeito que realiza o artigo ou a pesquisa se mostra mais evidentemente. Essas são também expressões com alta frequência de uso em artigos científicos, por isso as incluímos neste estudo. A essas expressões chamamos marcadores subjetivos.

A seguir, apresentamos o elenco de todos marcadores compilados através da análise dos *corpora*, classificados em três categorias: marcadores relacionados ao propósito do artigo científico, marcadores **metarreferenciais** e marcadores subjetivos, conforme a seguir apresentados. Complementarmente, uma breve sistematização dos resultados encontra-se na seção 6.3, deste capítulo.

6.2.1 Marcadores relacionados ao propósito do artigo científico

Conforme descrevemos no Capítulo 2, esses marcadores relacionam-se a elementos do propósito do artigo científico, cujo texto visa, de maneira geral, a apresentar objetivos e objetos do estudo, descrever a fundamentação teórica, ações e métodos e apresentar as conclusões. É claro que os marcadores dos outros dois tipos também estão diretamente ligados ao propósito do artigo científico, mas aqui nos referimos a elementos da redação do artigo que estão vinculados à realização da pesquisa apresentada. Os outros níveis estão relacionados a como validar essa pesquisa (os marcadores subjetivos) e de que maneira organizá-la no texto (os marcadores metarreferenciais).

Os marcadores desse tipo foram agrupados em torno de duas palavras-núcleo **objetivo**, e **forma** e seis grupos funcionais **método**, **ancoragem**, **quantificação**, **tempo**, **explicação**, e **restrição**.

Na organização dos quadros, alguns símbolos e notações foram utilizados, conforme especificado:

- # - indica número, como **a seção #**;
- < > - indica que os verbos entre esses sinais encontram-se em várias conjugações nos *corpora*, como em <ser> **analisado**, que poderia ser **foi analisado**, **é analisado**;
- xy - essas letras indicam um sujeito

Apresentamos, a seguir, os quadros comparativos desse grupo, indicando os percentuais de ocorrência dos marcadores. Em seqüência, são tecidos alguns comentários feitos para cada um, incluindo-se, algumas vezes, exemplos de ocorrência nos *corpora*.

Quadro 6.2 - Ocorrência dos marcadores agrupados em função da palavra-núcleo “objetivo” nos corpora nas línguas portuguesa e inglesa

OBJETIVO	
Português	Inglês
com o objetivo de (35,4%) com a finalidade (8,7%)	<i>for the purpose(s) of</i> (18%) <i>our goal is to</i> (9%) <i>with the goal of</i> (5,4%)
0%	<i>for the purposes of this paper/article</i> (4,5%)
o objetivo central deste estudo (0,3%) o principal objetivo deste estudo (0,3%)	<i>the purpose of this study</i> (4,5%)
a fim de (48%) a fim de que (8,1%)	<i>in order to</i> (283,7%) ⁵ <i>in order that</i> (1,8%)

Pelo critério de frequência, observa-se, no Quadro 6.1, que a maneira de apresentar um objetivo em português “com o objetivo de” poderia ser traduzida para o inglês pela expressão “*for the purposes of*”, pois a tradução literal “*with the goal of*”, tem uma frequência de uso menor. As expressões “*for the purposes of*” ou “*in order to*” mostram ser expressões bastante frequentes em inglês e poderiam ser opções na elaboração do texto nessa língua, em alguns casos, menos calcada no português, que utilizaria “objetivo” na expressão.

Quanto ao “a fim de que”, percebemos que a tradução literal “*in order that*” não apresenta uma frequência muito alta, ao contrário de “*in order to*”. Num determinado texto,

⁵ Os números podem ultrapassar os 100%, no caso de expressões que ocorrem mais de uma vez em cada texto.

em que se tendesse a traduzir “a fim de que” por “in order to” deveria haver, então, uma adaptação da constituição da frase. Este é o caso de uma frase como:

Exemplo 1: “Essa particularidade das regras de associação na Web deve ser aproveitada de forma automatizada, **a fim de que** o trabalho manual de análise seja facilitado.”

O trecho “a fim de que o trabalho seja facilitado”, ao ser traduzido para o inglês utilizando-se a expressão “*in order to*”, necessitaria que o verbo viesse depois dessa expressão: “*in order to make easier*”.

Observamos também que a expressão “*the purpose of this study*” é bem mais utilizada, em comparação com a sua tradução literal, que seria “o propósito deste estudo”. Conclui-se, assim, que “*study*” poderia ser utilizado no lugar de “trabalho”, já que, como veremos mais adiante, sua tradução literal “*work*” não é muito utilizada em inglês.

Quadro 6.3 - Ocorrência dos marcadores agrupados em função da palavra-núcleo “forma” nos corpora nas línguas portuguesa e inglesa

FORMA	
Português	Inglês
da mesma forma que (6,6%)	<i>in the same way as</i> (3,6%)
de forma que (24%) de modo que (25,8%) de maneira que (5,4%)	<i>so that</i> (149,5%) <i>in a way that</i> (10,8%)
de tal forma que (6,9%)	<i>in such a way that</i> (15,3%)
é uma forma de (3,6%)	<i>it is a form of</i> (5,4%) <i>it is a way of</i> (1,8%)

como uma forma de (4,2%)	<i>as a way of</i> (13,5%) <i>as a form of</i> (1,8%)
--------------------------	--

Em relação ao Quadro 6.2, podemos dizer que essas unidades não foram exclusivamente aqui agrupadas por uma mesma função explicativa predominante, mas pela presença de palavras-chave “forma/maneira/modo” tal como se vê nos exemplos abaixo, em inglês e português.

Exemplo 2: “Além disso, o design de tecnologia para a criança, que a respeite como usuário deve, **da mesma forma que** com o adulto, envolvê-la no processo de criação.”

Exemplo 3: “A partir disto ocorre o entendimento destes efeitos **de forma que** se a mesma ação for tomada nas mesmas circunstâncias será possível antecipar o que ocorrerá (OR).”

Exemplo 4: “*It also tries to make computation manifest to us in the world **in the same way as** we encounter other phenomena, both as a way of making computation more naturally with the everyday world, and as a way of enriching our experiences.*”

Exemplo 5: “*K is a knowledge base that is to be modified **so that** the formulas in R are contained in (or implied by) the result, and the formulas in C are not.*”

Chamamos a atenção para a expressão “*so that*”, que tem uma frequência muito maior que “*in a way that*”, podendo ser uma boa alternativa de tradução para a expressão “de forma que”, que seria sua tradução literal.

Na expressão “é uma forma de”, chama a atenção a preferência pelo uso de “*form*”, ao invés de “*way*”, como na expressão mostrada a seguir, no Quadro 6.3. Nesses casos, consideramos que um estudo mais detalhado das ocorrências poderia revelar dados interessantes sobre as concordâncias dessas palavras e seus contextos de uso.

Quadro 6.4 - Ocorrência de expressões agrupadas pelo grupo funcional “método”

MÉTODO	
Português	Inglês
através do uso (12%)	
através da utilização (9%)	<i>through the use</i> (18%)
através de um processo (2,7%)	<i>through a process</i> (4,5%)
através da realização (3%)	
com o uso de (26,4%)	<i>with the use</i> (11,7%)
com a utilização (13,2%)	<i>by using</i> (49,5%)
Para a realização (20,7%)	
com o auxílio (11,1%)	
com a ajuda (5,1%)	<i>with the help of</i> (4,5%)

Agrupamos esses marcadores pela sua função ligada a método ou metodologia de pesquisa, isto é, como essa foi realizada, como os objetos foram tratados.

Observamos que as expressões “através da realização” e “para a realização”, não tiveram frequência em inglês. Em uma perspectiva da tradução, seu equivalente poderia ser “*by verbo+ing*”, que tem uma frequência de 536% no *corpus* em inglês. Na análise dessa expressão, percebemos também o caráter mais direto da língua inglesa, que não utiliza um substantivo, como em “através da **realização** de um estudo” como um intermediador, o inglês utilizaria, “*by studying ...*” diretamente.

Ainda, a expressão “para a realização de” é muito frequente em português, mas não a sua tradução literal, como “*for the accomplishment of*”. São utilizados, em inglês, os verbos,

ao invés do substantivo *accomplishment*, tais como “*to carry out*”, “*to accomplish*” e “*to perform*”, como mostra o Exemplo 6:

Exemplo 6: “*In this paper, we **carry out** a study of consistency measure with different search strategies.*”

Com os substantivos *usar* e *utilizar*, empregados quase indistintamente nas expressões “através do uso/utilização” de e “com o uso/utilização de”, chamamos atenção para o uso, em inglês, apenas de “*use*”. Novamente uma expressão com “*by* verbo + *ing*”. Nesse caso, “*by using*” tem frequência muito maior em inglês.

Para as expressões “com o auxílio de” e “com a ajuda de”, o inglês parece preferir o uso do verbo, como no Exemplo 7, ao invés do que, literalmente, e menos frequentemente “*the positive and negative aspects of technology-enabled learning was analyzed **with the help of** this pilot project.*”

Exemplo 7: “*This pilot project **helped to** analyze both the positive and negative effects of technology-enabled learning.*”

Quadro 6.5 - Ocorrência de expressões agrupadas pelo grupo funcional “ancoragem”

ANCORAGEM	
Português	Inglês
a partir de um/uma (41,1%)	<i>starting from a/an</i> (4,5%)
a partir da análise (6,6%) a partir dos resultados (3,6%) a partir da definição (2,7%) a partir de/das informações (1,5%)	<i>from the/an analysis</i> (9%)
xy <ser> ponto de partida para (1,5%)	xy <be> <i>starting point</i> (3,6%)
xy <servir> como ponto de partida para (0,9%)	xy <take, serve, chose> <i>as the starting point</i> (4,5%)
<servir> de base para (3%) <utilizar> como base para (3,3%)	<i>serves as a basis for</i> (3,6%) <i>use as the/a basis for</i> (8,1%)

Os marcadores presentes no Quadro 6.4 estabelecem pontos de partida para o segmento da frase, ou raciocínio, por isso entendemos que têm uma função de ancoragem.

Observamos que uma tradução literal da expressão “a partir de um/uma”, que é uma forma bastante empregada no português, não é tão freqüente nesse *corpus*. No caso de uma tradução para o inglês, poderíamos buscar uma reformulação da frase utilizando “from the” (sem o *starting*, que tem sentido de **partindo/a partir**). No exemplo 8, em que observamos uma tradução para o inglês, feita no próprio texto em português, a forma *from a* foi traduzida por **a partir de**, sem o *starting*, que seria a tradução literal e que não consideramos a melhor opção, como mostra o **Exemplo 8**: “O Algoritmo RISE (Rule Induction **from** a Set of Exemplars - indução de regras **a partir de** um conjunto de exemplares).

Quadro 6.6 - Ocorrência de expressões agrupadas pelo grupo funcional “quantificação”

QUANTIFICAÇÃO	
Português	Inglês
a grande maioria (4,8%)	<i>The vast majority</i> (3,6%)
na maioria dos casos (3,6%)	<i>in most cases</i> (10,8%)
a maior parte (15,6%)	<i>for the most part</i> (4,5%) <i>the majority</i> (20,7%)
em grande parte (3,3%)	<i>in large part</i> (3,6%)
uma grande parte (1,5%)	<i>a large part</i> (7,2%)
na maioria das vezes (6,6%)	0%
um grande número de (8,7%)	<i>a large number of</i> (16,2%)
uma grande variedade de (4,8%)	<i>a wide variety of</i> (15,3%)
até que ponto (2,4%)	<i>the extent to which</i> (19,8%) <i>to what extent</i> (13,5%)

Esses marcadores textuais foram agrupados no Quadro 6.5 pelo caráter quantificador de seu conteúdo. É característico de artigos científicos que tragam dados, números, comparações de resultados. Contextualizar e quantificar esses dados são as funções desses marcadores.

Dentre eles, destacamos o caráter da colocação das expressões de “grande + maioria” e “*vast + majority*”, não tendo sido registrada nesse *corpus* a colocação “*great + majority*”. A expressão “a maior parte” tem uma frequência de ocorrência de 15% no *corpus* em português. Em inglês, seu uso é bem menos freqüente e, ainda assim, na expressão “*for the most part*”, pois somente “*the most part*” não apresentou nenhuma ocorrência. “*For the most part*” ocorreu sempre entre vírgulas, podendo ser traduzido por “para a maioria” e até “na maioria

das vezes” nos contextos em que ocorreu no *corpus* em inglês; “*the majority of*” parece ser uma opção mais adequada.

A tradução literal para “na maioria das vezes”, “*most of times*”, teve apenas uma ocorrência no *corpus*. Percebemos que, em inglês, expressões utilizando “*a number of*.” tiveram 170% de frequência, o que indica que podem ser utilizadas como quantificadores em diferentes momentos, como em “*a number of strategies*”, “*a number of sources*”, ao invés de se utilizar “*majority*” ou “*most*”.

Os marcadores do Quadro 6.6, a seguir, dão uma idéia de localização no tempo ou de momento de realização de um experimento, de uma observação. Algumas também realizam o papel de ordenamento e organização do texto, como “*firstly*”, “primeiramente”.

Quadro 7.6 - Ocorrência de expressões agrupadas pelo grupo funcional “tempo”

TEMPO	
Português	Inglês
a partir do momento (em) que (3%)	<i>from the time at which</i> (0,9%)
até o presente momento (2,1%)	<i>up to now</i> (2,7%) <i>up to the present time</i> (0,9%)
primeiramente (12,6%) em um/num primeiro momento (5%)	<i>firstly</i> (9,9%)
durante a realização (3%) durante a execução (5,7%) durante a fase de (2,4%)	<i>during the course of <something></i> (8,1%)
0%	<i>over a period of</i> (6,3%)
a medida que (5,1%)	<i>as <something> + verb</i>
ao longo do tempo (6,3%) ao longo do processo (1,8%) ao longo de (tempo) (8,1%)	<i>over time</i> (45%) <i>throughout the process</i> (3,6%)

Dessas expressões, chama atenção, novamente, que o inglês tende a ser mais específico, como podemos perceber na comparação com as expressões “durante a realização/execução de”. Em inglês, o *corpus* demonstrava um uso da palavra “*during*” seguida do evento, como “*during the study*”, “*during the research*”, “*during the analysis*”. Em português temos “durante a realização do estudo”, em vez do que seria, em inglês “durante o estudo”, como mostram os dois Exemplos 9 e 10.

Exemplo 9: “O fenômeno de *priming* semântico tem sido observado e estudado **durante a realização de** experimentos com a técnica de decisão léxica desde os trabalhos de MEYER & SCHVANEVELDT (1971) e NELLY.”

Exemplo 10: “*Performance curves of fifty trials by novice subjects showed a wide range of ability in the tasks, although they all improved substantially during the experiment.*”

Quadro 6.8 - Ocorrência de expressões agrupadas pelo grupo funcional “explicação”

EXPLICAÇÃO	
Português	Inglês
D	
trata-se de um/uma (13,2%)	0%
como é o caso (8,7%)	0%
que por sua vez (4,5%)	<i>which, in turn</i> (9%)
se por um lado ... por outro (5,7%)	<i>on the one hand ... on the other (hand)</i> (14,4%)
uma vez que (68,1%)	<i>since it is</i> (16,2%) <i>once</i> (74,7%)
como por exemplo (36,6%)	<i>for instance</i> ⁶ <i>for example</i>
na medida em que (17,4%)	<i>as</i>

Das informações que contém o Quadro 6.7, podemos destacar que o *corpus* não apresentou expressões literalmente equivalentes, em inglês, a “trata-se de um/uma” e “como é

⁶ Os percentuais das expressões “*for instance*”, “*for example*” e “*as*” não foram incluídos pois ficaria inviável analisar todas as suas ocorrências nos contextos aos quais estamos nos referindo na comparação apresentada.

o caso”. Uma sugestão de tradução é o uso de formas mais genéricas como “*it is a/an, they are*” ou “*as in*”, como no Exemplo 11:

Exemplo 11: “(...) *and then, as in the null evidence case, passing just the p messages.*”

Foi interessante descobrirmos que, no *corpus* em inglês, não houve ocorrência da expressão “*as for example*”, que seria literalmente traduzida por “**como** por exemplo”. Somente “*for example*” e “*for instance*” foram utilizados.

Para a expressão bastante freqüente em português, “na medida em que”, poderia se usar, em inglês, “*as ...*”. Para determinar corretamente seu uso enquanto equivalente da expressão “na medida que”, seria necessário avaliar cada uma de suas mais de cinco mil ocorrências.

Exemplo 12: “*Wenger (1991) examines the concept of social expectancy (the expectation of appropriate interactions as it develops when people use a computer interface).*”

As expressões do Quadro 6.8, a seguir, foram chamadas de restritivas pois entendemos que conduzem a um foco específico, a um referencial, restringindo desvios de enfoque e outras possibilidades.

Quadro 6.8 - Ocorrência de expressões agrupadas pelo grupo funcional “restrição”

RESTRICÇÃO	
Português	Inglês
no sentido de que (5,1%)	<i>in the sense that</i> (11,7%)
pode ser entendida/o como (2,7%)	<i>can be thought of as</i> (11,7%)
no que diz respeito a (14,1%)	<i>with respect to the</i> (45%)
no que se refere (13,2%)	<i>in regard to</i> (5,4%) <i>with regards to</i> (5,4%)
0%	<i>referred to as</i> (45%)
no que concerne (5,1%)	<i>concerning</i> (5,4%)
do ponto de vista de (17,4%)	<i>from the/a <xy> point of view</i> (18,9%) <i>from the point of view of <xy></i> (2,7%) <i>from the <xy> viewpoint</i> (4,5%) <i>from the viewpoint of <xy></i> (4,5%)
sob o ponto de vista (4,2%)	
de acordo com (124,9%)	<i>according to</i> (141,4%)
de acordo com <referência bibliográfica> (12,9%)	<i>according to <name></i> (27%) <i>in accordance with</i> (7,2%)

Além das comparações entre as freqüências de uso que podemos observar no Quadro 6.8, chamamos atenção para a seguinte expressão: “pode ser entendido/a como”, cuja tradução literal poderia ser “*can be understood as*”. Sua freqüência é de 1,8% em inglês, bem menor do que “*can be thought of as*”.

A expressão “sob o ponto de vista” é uma variação em português de “do ponto de vista”, e não houve ocorrências de sua tradução literal, para o inglês: “*under the viewpoint/under the point of view*”. Os *corpora* também demonstram que, embora todas as formas mostradas no Quadro 6.8, em inglês, tenham frequência, o uso de “*from the/a xy point of view*” é bem mais frequente.

Exemplo 13: “*From a control point of view, it is desirable to look ahead, so that the robot can respond to changes in the path edge.*”

6.2.2 Marcadores metarreferenciais

Os marcadores textuais desse grupo dizem respeito à organização formal e à ordenação do texto do artigo científico e dos elementos que fazem parte dele: seções, figuras, tabelas. Denominamo-os metarreferenciais pois fazem referência ao próprio texto.

Esses marcadores foram divididos em sete grupos em torno das palavras-núcleo **seção**, **figura**, **tabela**, **artigo**, **trabalho** e de um grupo específico de estruturas que se referem explicitamente à organização do texto.

Quadro 6.9 - Ocorrência dos marcadores agrupados em função da palavra-núcleo “seção” nos corpora nas línguas portuguesa e inglesa

SEÇÃO	
Português	Inglês
a seção # apresenta (12,3%)	<i>section # presents</i> (10,8%)
a seção # descreve (5,4%)	<i>section # describes</i> (9,9%)
a seção # conclui (1,5%)	<i>section # concludes</i> (4,5%)
a seção # mostra (1,5%)	<i>section # shows</i> (5,4%)

na seção # <ser> apresentado/a (25%)	
na seção # apresentamos (13%)	<i>in section # we present/show ... (73%)</i>
xy apresentado(s)/da(s) na seção # (8,1%) xy descrito(s)/ta(s) na seção # (6,3%) xy discutido(s)/da(s) na seção # (1,5%)	<i>xy presented in section # (8,1%) xy described in section # (15,3%) xy discussed in section # (11,7%)</i>
como/conforme (verbo) na seção # (9%)	<i>as (verb) in section # (31,5%)</i>
nesta/nessa seção (12,6%)	<i>in this section (54,9%)</i>
na próxima seção (9,6%) nas próximas seções (1,8%)	<i>in the next section (18,9%) in the next sections (1,8%) in the following sections (9%) in the following section (5,4%)</i>
conforme/como (verbo) na seção anterior (1,5%) (verbo) na seção anterior (5,7%)	<i>as (verb) in the previous section (1,8%) (verb) in the previous section (8,1%)</i>
a seção # conclui o artigo (2,1%)	<i>section # concludes the paper (4,5%)</i>
0%	<i>(see section #) (25,2%)</i>

* xy incluem o verbo ser

Em relação ao quadro 6.9, observamos que, enquanto em português há predominância do uso do verbo “apresentar”, em inglês as ocorrências estão um pouco mais diluídas em verbos mais específicos.

Observamos que o uso da voz passiva é muito freqüente em português, não se fazendo presente, no entanto, no *corpus* em inglês; ao passo que a expressão na voz ativa com a primeira pessoa do plural “*we*” em inglês é mais utilizada do que em português.

Em inglês, na expressão equivalente à “na próxima seção”, notamos a preferência de “*next*” para o substantivo no singular e “*following*” para substantivos no plural, diferença que não ocorre em português. Esta parece ser mais uma questão de colocação.

A forma utilizada entre parênteses “*see section #*” teve uma freqüência alta em inglês, e não ocorre em português, podendo ser uma boa alternativa de tradução e possibilitando que as frases no texto em inglês sejam formuladas de outra forma, isto é, apenas indicando para ver a seção por meio da expressão “*see...*” entre parênteses ao final.

Em geral, o que pudemos perceber foi que, em inglês, se utilizam, com mais freqüência, expressões que fazem referência a seções do próprio texto. Daí, uma das conclusões a que podemos chegar é que o texto do artigo científico em inglês tende a ser um bem organizado, com muitas referências à disposição e à organização dos conteúdos.

Quadro 6.10 - Ocorrência dos marcadores agrupados em função da palavra-núcleo “figura” nos corpora nas línguas portuguesa e inglesa

FIGURA	
Português	Inglês
a figura # mostra (19,8%)	<i>figure # shows</i> (14,4%)
a figura # apresenta (22,2%)	<i>fig. # shows</i> (44,1%)
a figura # ilustra (14,1%)	<i>figure # illustrates</i> (1,8%)
	<i>fig. # illustrates</i> (4,5%)
como apresentado na figura (0,9%)	<i>as shown in figure</i> (5,4%)
como mostrado na figura (1,2%)	<i>as shown in fig.</i> (15,3%)

	<i>as illustrated in figure</i> (2,7%) <i>as illustrated in fig.</i> (7,2%)
0%	<i>(see figure #)</i> (18%)
xy apresentado na figura (4,8%) xy ilustrado na figura (4,8%)	<i>xy shown in figure</i> (11,7%) <i>shown in fig.</i> (49,5%) <i>xy illustrated in figure</i> (2,7%) <i>illustrated in fig.</i> (7,2%) <i>xy given in fig.</i> (5,4%)
como mostra a figura # (3,9%) conforme mostra a figura # (2,1%) como ilustra a figura # (2,1%) conforme ilustra a figura # (2,1%)	<i>as figure # shows</i> (0,9%) <i>as figure # illustrates</i> (0)
como pode ser visto(a) na figura (3%) (visualizado, percebido)	<i>as can be seen in figure</i> (1%)
o gráfico da figura mostra/demonstra (1,8%)	0%
xy pode ser visto/a na figura (5,1%) visualizado/percebido/observado	<i>xy can be seen in figure</i> (17,1%)

Tal como mostram os dados analisados e apresentados no Quadro 6.10, na expressão “a figura # ...” não há, no *corpus*, ocorrências do verbo “*to present*” em inglês; enquanto em português, “apresentar” é o mais frequente. Seu equivalente em termos de maior frequência seria o “*show*”, cuja tradução literal estaria mais para “mostrar” do que para “apresentar”. Outro fato interessante identificado foi que, no inglês, a abreviação “Fig.” é bastante utilizada, o que não acontece no *corpus* em português.

É interessante observarmos que, com a expressão “como apresentado(da) na figura”, tivemos em português maior frequência da voz ativa “como apresenta a figura”, ao contrário do que aconteceu com a expressão em inglês, em que a voz passiva foi mais frequente.

Registramos uma frequência de uso significativamente maior da forma “*xy can be seen in figure*” em inglês (76,5%) contra os 9,6% em português; ainda, em inglês, todas as ocorrências dessa forma sintática com a palavra “figura” empregam o verbo “*to see*”, enquanto, em português, diluem-se no uso de outros verbos também.

Quadro 6.11 - Ocorrência dos marcadores agrupados em função da palavra-núcleo “tabela” nos corpora nas línguas portuguesa e inglesa

TABELA	
Português	Inglês
a tabela # apresenta (6,6%)	<i>table # shows</i> (14,4%)
a tabela # mostra (6%)	<i>table # illustrates</i> (2,7%)
a tabela # ilustra (1,8%)	
<ser> apresentado(a)(s) na tabela (4,5%)	<i>are/is shown in table #</i> (11,7%)
<ser> mostrado(a)(s) (1,8%)	<i>are/is listed in table</i> (6,3%)
como/conforme pode ser visto/observado na tabela # (1,8%)	0%
0%	<i>(see table #)</i> (18,01%)
0%	<i>as shown in table #</i> (4,5%)

Nesses marcadores, agrupados pela palavra-chave “tabela” (quadro 6.11), observamos novamente um uso em que o verbo “*to show*” ocorre como equivalente a “apresentar” em português.

A forma “como/conforme mostrado na tabela” não ocorreu de forma representativa no *corpus* em português (menos de 1%), o que não aconteceu com seu equivalente direto “*as shown in table #*”, em inglês. Já a forma “como/conforme pode ser visto/observado” teve frequência em português, ainda que baixa, enquanto, no inglês, não apresentou ocorrências.

Outra observação que nos chama a atenção é o uso, no *corpus* em inglês, de “*see table #*” entre parênteses. É uma forma bastante empregada no inglês e que não tem ocorrência como “(ver/veja tabela #)” em português. No entanto, ela deve ser lembrada na tradução para o inglês, como uma alternativa para a apresentação e a referência às tabelas.

Quadro 6.12 - Ocorrência dos marcadores agrupados em função da palavra-núcleo “artigo” nos corpora nas línguas portuguesa e inglesa

ARTIGO	
Português	Inglês
este artigo apresenta (20,1%)	<i>this paper presents</i> (7,2%)
este artigo descreve (9%)	<i>this paper describes</i> (14,4%)
neste artigo apresentamos (1,2%)	<i>in this paper we propose</i> (2,7%)
neste artigo propomos (0,9%)	<i>in this paper we describe</i> (4,5%)
neste artigo discutimos (0%)	<i>in this paper we present</i> (0,9%)
neste artigo <outro verbo> (4,5%)	<i>in this paper we <another verb></i> (11,7%)
o objetivo deste artigo é (4,8%)	<i>the purpose of this paper is</i> (2,7%) <i>the aim of this paper is</i> (1,8%)
este artigo tem como objetivo (1,8%)	0%

o trabalho (apresentado, descrito, relatado) neste artigo (0,9%)	<i>the work (described, presented) in this paper (4,5%)</i>
--	---

No Quadro 6.12 percebemos que a estrutura “neste artigo...” em início de frase, seguida da primeira pessoa do plural, como em “*In this paper we describe two aspects of a software architecture*”, tem uma frequência bem menor em português do que em inglês, indicando que o emprego do sujeito em inglês é mais convencional do que em português, onde, em geral, os artigos procuram evitá-lo. Inclusive, há um entendimento de que “não se usa” o ‘eu’ ou ‘nós’ em um texto científico.

Observamos que, em inglês, a forma “este artigo...” não mostrou a mesma regularidade de colocações (com as palavras apresentar/descrever) do português. Os verbos que seguem “*this paper ...*” são tão variados quanto os propósitos da pesquisa (*outline, gives information, reports, deals with, compares*).

Quadro 6.13 - Ocorrência dos marcadores agrupados em função da palavra-núcleo “trabalho” nos corpora nas línguas portuguesa e inglesa

TRABALHO	
Português	Inglês
este trabalho apresenta (8,7%)	0%
este trabalho descreve (4,5%)	0%
o objetivo deste trabalho é (3,3%) o objetivo principal deste trabalho é (1,2%)	0%
esse/este trabalho tem como objetivo	0%

(1,5%) esse/este trabalho tem por objetivo	
(1,2%)	
a principal contribuição deste trabalho (1,5%)	<i>The main/major contribution of the work (2,7%)</i>

Registramos, no Quadro 6.13, que a tradução literal de “trabalho”, em inglês “*work*”, não é empregada como equivalente no sentido de “trabalho escrito”, o que justifica as casas em branco. Apenas na expressão “a principal contribuição deste trabalho” observamos essa ocorrência; e, em inglês, empregou-se o artigo definido “*the*”, enquanto, em português, se utiliza o demonstrativo “deste”.

Quadro 6.14 - Ocorrência dos marcadores agrupados em função de seu caráter explícito de organização do texto

ORGANIZAÇÃO	
Português	Inglês
este artigo está organizado - da seguinte forma (2,1%) - da seguinte maneira (0,6%)	0%
o artigo está organizado - da seguinte forma (0,9%) - da seguinte maneira (0,3%)	<i>the paper is organized as follows (4,5%)</i> <i>the rest of the paper is organized as follows (4,5%)</i>
o restante deste artigo (1,5%)	<i>in the reminder of this paper (3,6%)</i>
o trabalho/artigo está organizado como	<i>the paper is organized as follows</i>

segue (0,9%)	(4,5%)
o item #/a seção # apresenta os resultados (1,8%)	<i>section summarizes/contains/describes the results</i> (2,7%)

Os dados do Quadro 6.14, acima, evidenciam que, em inglês, não houve uso do artigo definido “*this*” na expressão que introduz a forma como o artigo está organizado, em geral presente no último parágrafo da introdução.

A respeito das expressões que complementam “este artigo está organizado...”, constatamos que, em inglês, se utiliza apenas uma forma “...*organized as follows*”, enquanto, em português, temos “da seguinte forma/maneira” e “como segue”; ainda em português, a forma “como segue” é menos freqüente do que em inglês. Os dados também evidenciaram que, em inglês, há o emprego da expressão “*the rest of the paper is organized...*”, que não é utilizada em sua tradução literal equivalente em português “o restante/resto do trabalho/artigo está organizado...”

No equivalente à expressão “o restante desse artigo”, podemos perceber que, em inglês, se utiliza a preposição: “*in the reminder of this paper*”, tornando necessária a presença de um sujeito subsequente, como em “*in the reminder of this paper we describe/explain...*”.

No que concerne à descrição do que é feito na seção, na expressão “na seção # ...”, observamos que, em inglês, são utilizados diferentes verbos, enquanto, em português, a predominância é do verbo “apresentar”.

6.2.3 Marcadores subjetivos

Nesta seção, apresentamos os marcadores textuais que denominamos subjetivos. Foram assim denominados, porque trazem consigo e deixam transparecer a visão do sujeito que realiza o artigo. Embora esses marcadores tenham essa característica que os distingue dos

mais tipicamente textuais, estão presentes nos *corpora* e são indissociáveis da constituição de um artigo científico. Contam com a presença de modalizadores e foram divididos em cinco subgrupos, de acordo com as seguintes palavras-núcleo: **importante**, **necessário**, **esperar/acreditar**, **observar**, **poder**.

Quadro 6.15 - Ocorrência dos marcadores agrupados em função da palavra-núcleo “importante” nos corpora nas línguas portuguesa e inglesa

IMPORTANTE	
Português	Inglês
é importante ressaltar que (6,3%)	<i>it is important to note that</i> (8,1%)
é importante observar que (3%)	<i>it is important to recognize that</i> (3,6%)
é importante destacar que (2,4%)	<i>it is important to consider that</i> (2,7%)
é importante salientar que (2,1%)	<i>it is important to notice that</i> (1,8%)

No Quadro 6.15, salienta-se a seleção dos verbos mais utilizados em uma e outra língua. Os verbos de língua portuguesa parecem ser mais enfáticos, como “ressaltar”, “destacar” e “salientar”, apesar de todos estarem relacionados à importância do que está sendo dito. Consideramos significativo identificar essas diferenças para que o tradutor conheça as opções mais empregadas em inglês, que diferem um pouco do português.

Quadro 6.16 - Ocorrência dos marcadores agrupados em função da palavra-núcleo “necessário” nos corpora nas línguas portuguesa e inglesa

NECESSÁRIO	
Português	Inglês
é necessário que se (3,6%)	<i>it is necessary to</i> (13,5%)
é necessário que (12%)	

Em relação ao Quadro 6.16, chamamos atenção apenas para a diferença sintática dessa construção, que acarreta uma reformulação da frase no momento da tradução; no português, a frase conta com a partícula apassivadora “se”, e, no inglês, fica no infinitivo, como mostram os Exemplos 13 e 14:

Exemplo 13: “Para a aplicação das redes Bayesianas **é necessário que se** obedeça a condição de Markov, de que não existe uma relação de dependência direta entre quaisquer dois nós.”

Exemplo 14: “*For TCSP, it is necessary to decompose into a disjunction of STPOs to get a minimal graph by PC1 and PC2.*”

Quadro 6.17 - Ocorrência dos marcadores agrupados em função das palavras-núcleo “esperar / acreditar” nos corpora nas línguas portuguesa e inglesa

ESPERAR/ACREDITAR	
Português	Inglês
espera-se que (10,8%)	<i>we hope that</i> (2,7%)
esperamos que (2,4%)	<i>it is expected that</i> (1,8%)
	<i>we expect that</i> (1,8%)

acredita-se que (10,5%)	<i>we believe that</i> (17,1%)
acreditamos que (15,6%)	

A respeito dessas expressões, contidas no Quadro 6.17, chama-nos atenção o uso da primeira pessoa do plural, que varia em frequência de uma forma para a outra. Em português, com o verbo “esperar”, é mais frequente o uso do indeterminado “espera-se que”, e com o verbo “acreditar”, é mais frequente o uso da primeira pessoa do plural “acreditamos que”, evidenciando que é arriscado fazermos generalizações sobre as preferências de uso e de colocabilidade em se tratando de culturas de diferentes línguas.

Quadro 6.18 - Ocorrência dos marcadores agrupados em função da palavra-núcleo “observar” nos corpora nas línguas portuguesa e inglesa

OBSERVAR	
Português	Inglês
observa-se que (15%) observou-se que (9,6%) observamos que (6,6%)	<i>we observe that</i> (5,4 %) <i>it is observed that</i> (1,8%)
podemos observar que (3,9%)	<i>we observe that</i> (5,4%) <i>we note that</i> (12,6%)
pode-se notar que (3,9%) pode-se perceber que (2,4%) pode-se observar que (5,1%)	<i>note that</i> (135%) (imperativo) <i>notice that</i> (27%) (imperativo) <i>observe that</i> (12,6%) (imperative)
nota-se que (10,5%)	<i>notice that</i> (32,4%)

Nessas expressões do Quadro 6.18, observamos, além da preferência pelo uso da primeira pessoa do plural em inglês, que, na forma “nota-se que”, um equivalente de uso poderia ser a forma no imperativo, no inglês, “*notice that*”.

Constatamos que a forma “podemos observar que”, conjugada na primeira pessoa do plural, é mais freqüente em inglês. Já a forma indeterminada “pode-se observar/perceber/notar que” teria como equivalente de uso o imperativo na língua inglesa, sem o uso do “*can*” ou outro modal.

Quadro 6.19 - Ocorrência dos marcadores agrupados em função da palavra-núcleo “poder” nos corpora nas línguas portuguesa e inglesa

PODER	
Português	Inglês
pode ser visto(a) como (8,1%)	<i>can be seen as</i> (7,2%) <i>can be viewed as</i> (10,8%)
pode ser definido(a) como (5,4%)	<i>can be defined as</i> (3,6%)
pode ser classificado(a) como (3%)	<i>can be classified as</i> (1,8%)
pode-se dizer que (7,5%) podemos dizer que (5,7%)	<i>we will say that</i> (6,3%) <i>we could say that</i> (1,8%) <i>we shall say that</i> (1,8%) <i>we can say</i> (0,9%) / <i>cannot say</i> (1,8%) <i>we might say</i> (0,9%)
podemos afirmar que (3,6%)	0%

Nas informações contidas no Quadro 6.19, vemos que a tradução literal de “dizer”, em inglês “*say*”, ocorre com todos os modais, mas sempre com a primeira pessoa do plural,

diferentemente da língua portuguesa, que usa também o indefinido na expressão “pode-se/podemos dizer que”.

6.3 Algumas considerações

Apresentamos, neste capítulo, os resultados do levantamento dos marcadores textuais, em inglês e português, a partir de dois *corpora* – um em português, com 333 textos e outro em inglês, com 111 textos.

Em primeiro lugar, é preciso dizer que, devido à quantidade de textos e à diversidade da natureza e da constituição dos marcadores que coletamos, nos deparamos com muitas especificidades. Cada um dos marcadores em português, e a sua comparação em inglês, mereceria um estudo que pudesse avançar ainda mais, descrevendo, mais detalhadamente, as formas de realização de cada um deles. Os resultados mostram, por exemplo, que muitas questões relacionadas a estruturas lingüísticas do artigo científico e seu comportamento em português e em inglês, podem trazer ainda mais esclarecimentos para quem lida com a tradução. No entanto, isto foge dos objetivos desta dissertação e os dados analisados nos permitem sistematizar alguns resultados, que descrevemos a seguir.

Com relação às características de cada um dos três grandes grupos de marcadores, percebemos, que, no caso do português, os marcadores mostraram algumas regularidades em sua constituição. Trazemos essas considerações aqui apenas no intuito de realizar uma apreensão geral da característica dos marcadores de acordo com os grupos. Essas constatações não implicam, diretamente, em questões para o tradutor, uma vez que estas encontram-se mais no nível de realização de cada estrutura em particular, como mostraremos em seguida, e não enquanto grande grupo. Assim,

- os marcadores textuais relacionados ao propósito do artigo científico mostraram ter mais a forma de locuções, sempre com a presença de preposições, como: “a fim de”, “de forma que”, “a partir de”, “com o uso de”;
- os marcadores metareferenciais mostraram-se como formas mais fraseológicas, compostos de uma palavra núcleo (**tabela, figura, seção**) e verbo: “a seção # apresenta”, “conforme descrito na seção #”, “como ilustra a figura”, “este artigo apresenta”;
- os marcadores subjetivos mostraram ter uma forma que introduz orações subordinadas: “é importante ressaltar que”, “é necessário que”, “esperamos que”, “observa-se que”. A eles, geralmente, está relacionada a presença de modalidades, embora com pequenas diferenças no português e inglês.

Além desses dados genéricos sobre os modos de constituição das estruturas analisadas, a comparação entre as línguas nos permite dizer que:

- no grupo dos marcadores relacionados ao propósito, a língua inglesa costuma ser mais direta na formulação das expressões, assim, em português, teríamos: “através da realização de um estudo”, e, em inglês, teríamos, “*by studying*” (estudando-se), ou “*through the study*” (através do estudo);
- a língua inglesa realiza mais construções sintáticas de frases que utilizam ou favoreçam o uso do sujeito na primeira pessoa do plural *we* (nós);

Exemplo 15: “*In this paper we present a new algorithm for DTP solving, called Epilitis.*”

Exemplo 16: “Este artigo apresenta a proposta de uma estrutura baseada em agentes inteligentes para a definição de uma seqüência.”

- a língua portuguesa utiliza algumas formas passivas inexistentes no *corpus* em inglês, mas não que sejam inexistentes em inglês, apenas são menos freqüentes:

Exemplo 17: (Em português) **Na seção # são apresentados.**

(Em inglês) nenhuma ocorrência

Esse exemplo mostra um caso em que, em inglês, a forma mais freqüente favoreceria o uso do sujeito na primeira pessoa, como “*in section # we show.*”

- os modais, em inglês (*can, could, would, will, shall*) só acontecem nas expressões que têm mesmo função de modalidade (*we could say that, can be seen as*); quando, em português, utiliza-se **poder** na função de “possibilidade”, em inglês, não utiliza-se o modal (em português “podemos afirmar que”);
- a língua inglesa apresenta maior freqüência de uso de expressões que organizam explicitamente a estrutura do texto:

Exemplo 18: “*This paper is organized as follows.*” (4,5%)

Exemplo 19: “O artigo está organizado da seguinte forma/como segue.” (1,2%)

- outra evidência, talvez de menos relevância, é que o inglês utiliza, para fazer referência a elementos como figuras e tabelas, a forma entre parênteses, como em “(see Table 1)”, e em português essa forma não mostrou nenhuma freqüência. Apesar dessa evidência parecer de pouca relevância, é um recurso que, na hora da tradução pode possibilitar que o tradutor faça uma escolha por uma frase mais direta, em inglês, apenas indicando, entre parênteses, onde

determinada informação pode ser encontrada; isso ao invés de traduzir literalmente uma expressão do português como “como mostra a figura #.”

Essa comparação feita entre as línguas mostra, talvez acima de tudo, que os textos dos dois sistemas de línguas envolvidos neste estudo, apesar de se realizarem em um único gênero, concretizam-se de maneira muito particular, em determinados contextos. Isto, inclusive, evidencia a necessidade e a importância de comparar, pontualmente, suas realizações textuais.

Era previsível que não seria possível fazer um grande quadro sistemático e generalizador dos resultados da comparação, pois o estudo e a atenção que teriam que ser dados a cada marcador extrapolariam os limites deste trabalho de mestrado. Além disso, estamos colocando lado a lado textos de ocorrências reais de duas línguas, dois sistemas diferentes e que pertencem a culturas diferentes.

A prática da tradução, e até mesmo sua teoria, têm ensinado que muito dificilmente temos respostas, fórmulas e truques para resolver todos os impasses de tradução. Conforme expôs Snell-Hornby (1998, p. 3 apud Azenha, 1999, p. 19),

“os estudos de tradução têm sido prejudicados pelos modos clássicos de categorização, que operam com divisões estanques, oposições binárias, antíteses e dicotomias. Frequentemente, tais modos de categorização não passam de construtos acadêmicos que impedem as diferenciações sutis, presentes em todos os aspectos dos estudos de tradução.”

Sabemos que, no estudo e na realização de traduções, estamos em uma área em que não há regras, em que classificações não ajudam, e em que cada texto abordado apresenta um desafio diferente, com suas particularidades e contextos. O que vai prevalecer, para uma boa tradução, será, além da experiência, o nível de atenção e a disposição do tradutor de estar atento às diferenças e de buscar soluções para elas. Por esse motivo, imaginamos que o levantamento dos marcadores, baseado em corpora, pode ser um ponto de partida para que os

tradutores, principalmente os principiantes, atentem para as duas culturas de escritas em que seu trabalho está envolvido. Nesse sentido, deve-se pensar a tradução como um novo texto, que, dependendo da situação, não carregue as marcas da língua de partida, como no caso dos textos dos artigos científicos, em que isto seria indesejável, considerando, sobretudo, o nível de internacionalização desse gênero textual. De acordo com Azenha (1999, p. 13), “(...) do tradutor técnico se exige, além do domínio de um saber específico, a habilidade de coordenar ações e garantir (...) a adequação do texto traduzido às condições de recepção”, se esse for o objetivo da tradução. Este foi um dos objetivos dessa comparação, mostrar como uma expressão pode ser melhor adaptada ao evento de recepção do texto.

Podemos dizer que este capítulo apresentou a concretização da idéia que apresentamos no início desta dissertação e que, a bem da verdade, nunca teve a pretensão de chegar a uma única e reveladora conclusão. Foi um estudo exploratório e de observação, que revelou dados que continuarão sendo observados em suas ocorrências e, principalmente, aplicados no trabalho prático da tradução.

CONSIDERAÇÕES FINAIS

Acreditamos que esse trabalho tenha cumprido principalmente uma das metas a que se propôs, a de ser um meio e um resultado, que permitissem aprender, pesquisar, refletir e estar em contato com o desenvolvimento do conhecimento, não só para a teoria, mas também para a prática da tradução.

Durante a realização desta dissertação, foi possível conhecer mais de perto o texto do artigo científico, um gênero com o qual já trabalhávamos. Em razão dessa prática, sentimos a necessidade de buscar um entendimento maior ou uma percepção da noção de gênero nele envolvida. Compreender a estrutura e o funcionamento de um gênero é essencial à prática do tradutor. Com esses princípios, avançamos no estudo de um gênero em particular, o artigo científico, identificando expressões que o constituem e lhe são características e, cujo conhecimento mais sistemático facilita a prática tradutória, bem como a reflexão sobre esta prática.

Para tanto, esta dissertação traçou, em seu primeiro capítulo, um breve percurso das reflexões sobre alguns estudos de gênero e sobre o texto que o constitui. Nele, apresentamos noções fundamentais para a abordagem de nosso objeto, como os conceitos de esferas de comunicação, em Bakhtin, e comunidade discursiva, para Swales. O entendimento de que um gênero apresenta características típicas e recorrentes, produzidas e reconhecíveis por parte dos membros das comunidades que lidam com o mesmo, validaram nosso pressuposto de estudar

algumas dessas características em comparação, pois, embora o artigo científico seja um gênero internacional, da grande comunidade científica, a maneira como seus elementos textuais prototípicos se realizam no texto dependem da língua em que ocorrem e da sua cultura de escrita.

O Capítulo 2 dedicou-se exclusivamente ao nosso objeto *latu sensu*, o artigo científico. Trouxemos para o trabalho considerações sobre a evolução histórica do mesmo enquanto gênero e também mostramos alguns trabalhos importantes já feitos sobre o artigo científico. Mencionamos que muitas são as características textuais desse gênero, porém, aqui, nos interessamos pelos marcadores – aqueles elementos do texto que lhe dão coesão, o organizam e o fazem ser o que é: um gênero estável. O intuito foi chamar a atenção para elementos que constituem tipicamente o texto científico dos artigos, que não fossem nem termos nem fraseologias especializadas, mas que, como esses, também merecem ser atentamente observados pelo tradutor, uma vez que fazem com que o artigo científico aconteça enquanto tal, lhe dão coerência e, principalmente, coesão; características, essas, fundamentais na realização de um bom texto.

Os marcadores textuais que analisamos não se mostraram pertencentes apenas a uma categoria sintática, tendo sido, então, agrupados por suas características funcionais. Dessa forma, nossa compilação ficou composta por locuções, fraseologias e colocações. Percebemos, no entanto, que os marcadores se realizaram com características sintáticas semelhantes, dentro de cada categoria de nossa classificação. Assim, como mostramos no último capítulo, os marcadores relacionados ao propósito do artigo se mostraram mais na forma de locuções; os marcadores metarreferenciais, na forma de fraseologias; e os marcadores subjetivos, na forma de orações que introduzem subordinadas.

No Capítulo 2 também justificamos a escolha pelo gênero *artigo científico*, pois esses textos têm sido extremamente importantes no quadro atual da pesquisa no Brasil e no mundo,

e isso pôde ser percebido pelos dados que mostramos, em que os números indicam que a produção de artigos cresce a cada ano. Sendo assim, o trabalho com esse gênero textual constitui um mercado de trabalho significativo para tradutores e revisores, o que também justifica que se trabalhe com ele no ensino da tradução.

Em nosso ponto de vista, para que esse trabalho pudesse também, mais praticamente, auxiliar o aprendiz da tradução, fizemos um estudo comparativo entre inglês e português. A Retórica Contrastiva, apresentada no Capítulo 3, fundamentou principalmente essa segunda etapa do trabalho que foi, após termos identificado constituintes característicos dos artigos científicos em português, compararmos os mesmos em inglês e português. Os estudos de Retórica Contrastiva mostram que as formas retóricas disponíveis em línguas distintas são semelhantes, mas com frequências e preferências de uso diferenciadas; isto porque cada comunidade é motivada por sua cultura, que conta com recursos diferentes da língua para certas realizações textuais.

Isso é de extrema importância para o tradutor, uma vez que ele é quem faz essa intermediação entre o texto de partida e o texto de chegada. O caráter do texto de chegada se dará pelas escolhas que o tradutor fizer, e, quanto mais recursos ele tiver para avaliar as diferenças e o que está envolvido nelas, mais chance de fazer um trabalho competente com autonomia ele terá. Um exemplo disso, que pôde ser visto nesse trabalho, foi como, em inglês, o texto tende a ser em uma ordem mais direta, sem “intermediadores”. Essa é uma afirmação até conhecida dos profissionais do inglês, mas mostrar e reconhecer onde essa característica acontece é que é pouco mencionado. Mostramos, no Capítulo 6, algumas dessas situações, no uso dos marcadores textuais.

Chegar ao fim desse trabalho que utilizou como *corpora* textos que, unidos, totalizaram mais de quatrocentos, como descrito no Capítulo 4, nos dá uma certeza de que trabalhamos com um número talvez não o ideal, mas que oportunizou traçar um retrato no

mínimo verdadeiro dos marcadores utilizados nos artigos. Pelos dados que foi possível extrair dos *corpora*, deixamos para trás as memórias de quanto trabalho deram para serem localizados, extraídos, compilados, transformados, organizados e analisados. As ferramentas da Lingüística de *Corpus* nos deram o apoio necessário para essas tarefas, e não se mostraram nem um pouco hostis, realmente ajudando e otimizando o tempo. Apenas nos mostraram que trabalhar com textos de ocorrências reais pode dar um pouco mais de trabalho, principalmente quando se permite que os textos mostrem suas características e peculiaridades, diferente do que seria trabalhar com um protótipo, ou estruturas prototípicas, de artigos científicos. Devido a essa diversidade, muito nos agradou quando constatamos que uma classificação feita empiricamente – após horas de exame a uma lista imensa de expressões, que não pareciam ter os menores pontos em comum – encontrava suporte em um importante modelo lingüístico: o modelo sistêmico-funcional e as três metafunções da linguagem apresentadas por Halliday (1985).

O trabalho de análise particular de cada expressão mostrou-se difícil, pois definirmos e fixarmos, uma a uma, a forma das expressões selecionadas, tentando contemplar ao máximo suas variações em cada uma das línguas, e observando seus contextos, exigiu tempo e atenção, tamanha era a diversidade de constituição e função desses objetos. Classificamos, organizamos e comparamos essas expressões e, ao final, talvez tenhamos encontrado menos regularidades do que gostaríamos, e comprovamos, mais uma vez, que no mundo da linguagem e da tradução há sempre mais perguntas do que respostas. Entendemos que, para um traçado da dimensão exata de realização específica de cada um dos marcadores que estudamos, seria necessário um exame minucioso de cada um, contando ainda com mais textos, de outras áreas e até com um *corpus* de outros gêneros textuais, para contraste. . Ainda assim, foi possível perceber algumas particularidades das duas línguas que nos chamaram bastante atenção e que, certamente, ajudarão na realização dos futuros trabalhos de tradução.

Entendemos que esse tipo de estudo mais detalhado sobre marcadores textuais pode também vir a interessar abordagens voltadas para o processamento da linguagem natural e tradução automática e essa é, talvez, uma perspectiva futura de trabalho e continuação da investigação aqui iniciada.

Realizamos, nesta dissertação, uma trajetória de busca de conhecimento, guiada pelos interesses em compreender melhor determinados aspectos envolvidos na prática de tradução de artigo científico. Alcançamos resultados que, esperamos, venham a contribuir não apenas para a prática, mas também para os estudos teóricos de tradução e texto especializado. Talvez não seja muito comum encerrar um trabalho de mais de dois anos afirmando que ele nos mostra que muito ainda há por fazer, no entanto, para nós, continuá-lo é uma renovação gratificante de inquietudes e do enfrentamento de novos desafios para aprofundar o conhecimento sobre a tradução e aparelhar melhor o profissional, tal como nos sentimos agora. Esperamos que este estudo possa também vir a provocar esse mesmo efeito em outros tradutores e, especialmente, que ele auxilie os alunos de tradução na aquisição de sua competência profissional.

REFERÊNCIAS

- ALUISIO, S.M.; GANTENBEIN, R.E. Towards the application of systemic functional linguistics in writing tools. *Proceedings of the ISCA 12th International Conference*. Tempe: M.C. Woodfill; 1997. p. 181-185.
- ANDRADE, M.L.F. *Gêneros e tipos. Uma aproximação*. Disponível em: www.filologia.org.br/soletras/2/06.htm. Acessado em 3 de outubro de 2004.
- ARISTÓTELES. *Arte retórica e arte poética*. São Paulo: Difusão Européia do Livro; 1959.
- AZENHA Jr., J. *Tradução técnica e condicionantes culturais – primeiros passos para um estudo integrado*. São Paulo: Humanitas/USP; 1999.
- BAHTIA, V.K. *Analysing genre: language use in professional settings*. New York: Longman Publishing; 1993.
- BAKER, M. *In other words. A coursebook on translation*. London: Routledge; 1992.
- _____. Corpus linguistics and translation studies: implications and applications. In: BAKER, M.; FRANCIS, G.; TOGNINI-BONELLI, E. (eds.). *Text and technology. In honour of John Sinclair*. Philadelphia: J. Benjamins Pub. Co.; 1993. p. 233-252.
- BAKHTIN, M.M. *Estética da criação verbal*. 2ª ed. São Paulo: Martins Fontes; 1997.
- BARBISAN, L.B. Texto e contexto. *Organon*, 1995, v. 9, n. 23, p. 51-60.
- BARROS, D. *Teoria semiótica do texto*. São Paulo: Ática; 1990.

- BAZERMAN, C. *Reporting the experiment: the changing account of scientific doings in the philosophical transactions of the royal society*. 1983. p. 1665-1800 [mimeo].
- BEEBY LONSDALE, A. La traducción inversa. In: ALBIR, A.H. (ed.). *La enseñanza de la traducción*. Castellón: Publicacions de la Universitat Jaume I; 1996. Col. Estudis sobre la Traducció 3.
- BENEDETTI, I.C. Prefácio. In. BENEDETTI, I.; SOBRAL, A. (orgs.) *Conversas com tradutores – balanços e perspectivas da tradução*. São Paulo: Editora Parábola; 2003.
- BENEDETTI, I.; SOBRAL, A. (orgs.). *Conversas com tradutores. Balanços e perspectivas da tradução*. São Paulo: Parábola Editorial; 2003.
- BERBER-SARDINHA, T. *Linguística de corpus*. Barueri: Editora Manole; 2004.
- BERGER, D., REITTER, D., STEDE, M. (2002). *XML/XSL in the dictionary: the case of discourse markers*. Disponível em: http://www.reitter-it-media.de/compling/papers/bergeretal_xmldiscmarkers_2002.pdf. Acessado em 3 de outubro de 2004.
- BEVILACQUA, C. Unidades fraseológicas especializadas: novas perspectivas para sua tradução e tratamento. *Organon*, 1998, v. 12, n. 26, p. 119-132.
- BEVILACQUA, C.; LORENTE, M. Los verbos en las aplicaciones terminográficas. In: *Anais do II Simpósio Iberoamericano de Terminologia*, Lisboa, 14-17 de novembro de 2000. Buenos Aires: Unico Latina; 1994.
- BEZERRA, B.G. *A distribuição das informações em resenhas acadêmicas* [dissertação de mestrado]. Fortaleza: Universidade Federal do Ceará; 2001. 137 p.
- CABRÉ, M.T. *La terminologia: representación y comunicación*. Barcelona: Institut Universitari de Linguística Aplicada; 1999.
- CIAPUSCIO, G.E. *Tipos textuales*. Buenos Aires: Oficina de Publicaciones; 1994.
- _____. *Textos especializados y terminologia*. Barcelona: Institut Universitari de Linguística Aplicada; 2003.

- CONNOR, U. *Contrastive rhetoric*. New York: Cambridge University Press; 1996.
- CONNOR, U.; KAPLAN, R.B. *Writing across languages: analysis of L2 text*. Reading: Addison-Wesley; 1987.
- CONNOR, U.; McCAGG, P. Cross-cultural differences and perceived quality in written paraphrases of English expository prose. *Applied Linguistics*, 1983, v. 4, n. 3, p. 259-268.
- ENKVIST, N.E. Discourse type, text-type and cross-cultural rhetoric. In: TIRKKONEN-CONDIT, S. (ed.). *Empirical research in translation and intercultural studies*. Tübingen: Gunter Narr Verlag; 1991. p. 5-16.
- _____. Discourse organization, crosscultural rhetoric and translation. In: WARVIK, B.; TANSKANEN, S.-K.; HILTUNEN, R. (eds.). *Organization in discourse. Proceedings from the Turku Conference, Anglicana Turkuensia N°14*. Turku: University of Turku; 1995. p. 41-58.
- FÁVERO, L.L.; KOCH, I.V. *Lingüística textual: uma introdução*. São Paulo: Cortez; 1983.
- FERNÁNDEZ POLO, F.J. *Traducción y retórica contrastiva: a propósito de la traducción de textos de divulgación científica del inglés al español*. Santiago de Compostela: Universidad de Santiago de Compostela; 1999.
- FISH, S. *Is there a text in this class?* Harvard: Harvard University Press; 1980.
- FREEDMAN, A.; MEDWAY, P. New views of genre and their implications for education. In: FREEDMAN, A.; MEDWAY, P. (orgs.). *Learning and teaching genre*. Portsmouth: Heinemann; 1994.
- GOSDEN, H. Discourse functions of marked theme in scientific research articles. *English for Specific Purposes*, 1992, v. 11, p. 207-224.
- _____. Success in research article writing and revision: a social-constructionist perspective. *English for Specific Purposes*, 1995, v. 14, n. 1, p. 37-57.

- HALLIDAY, M.A.K. *Language as social semiotic – the social interpretation of language and meaning*. Baltimore: University Park Press; 1978.
- _____. *An introduction to functional grammar*. London: Edward Arnold; 1985.
- HOFFMANN, L. Conceitos básicos da lingüística das linguagens especializadas. *Cadernos de Tradução*, 2004. No prelo.
- HUNSTON, S. *Corpora in applied linguistics*. Cambridge: Cambridge University Press; 2002.
- HURTADO ALBIR, A. (org.). *Enseñar a traducir. Metodología en la formación de traductores e intérpretes*. Madrid: Edelsa; 1999.
- _____. *Traducción y traductología. Introducción a la traductología*. Madrid: Ediciones Cátedra; 2001.
- KAPLAN, R.B. Cultural thought patterns in intercultural education. *Language Learning*, 1966, v. 16, p. 1-20.
- KATZ, J.J., FODOR, J.A. The structure of a semantic theory. *Language*, 1963, v. 39, p. 170-210.
- KNORR-CETINA, K.D. *The manufacture of knowledge*. Oxford: Pergamon; 1981.
- KOCH, I.V. *A coesão textual*. 6ª ed. São Paulo: Contexto; 1993.
- KRIEGER, M.G. A face lingüística da terminologia. In: KRIEGER, M.G.; MACIEL, A.M.B. (orgs.). *Temas de terminologia*. Porto Alegre/São Paulo: Universidade Federal do Rio Grande do Sul/Humanitas/Universidade de São Paulo; 2001a.
- _____. Sobre terminologia e seus objetos. In: KRIEGER, M.G.; MACIEL, A.M.B. (orgs.). *Temas de terminologia*. Porto Alegre/São Paulo: Universidade Federal do Rio Grande do Sul/Humanitas/Universidade de São Paulo; 2001b.
- KRIEGER, M.G.; FINATTO, M.J.B. *Introdução à terminologia. Teoria e prática*. São Paulo: Contexto; 2004.

- LAVIOSA, S. Corbus-based translation studies: where does it come from? Where is it going? *TRADTERM*, 2004, v. 10, p. 29-57.
- MACIEL, A.M.B. *Para o reconhecimento da especificidade do termo jurídico* [tese de doutorado]. Porto Alegre: UFRGS; 2001. 299 p.
- _____. (org.). *Temas de terminologia*. Porto Alegre/São Paulo: Universidade Federal do Rio Grande do Sul/Humanitas/Universidade de São Paulo; 2001.
- MARCUSCHI, L.A. Gêneros textuais: definição e funcionalidade. In: DIONÍSIO, A.P.; MACHADO, A.R.; BEZERRA, M.A. (orgs.). *Gêneros textuais & ensino*. Rio de Janeiro: Editora Lucerna; 2002. p. 19-36.
- MAURANEN, A. Contrastive ESP rhetoric: metatext in finnish-english economics texts. *English for Specific Purposes*, 1993, v. 12, p. 3-22.
- MILLER, C.R. Genre as social action. *Quarterly Journal of Speech*, 1984, v. 70, p. 151-167.
- NEUBERT, A.; SHREVE, G. *Translation as text*. Kent: The Kent State University Press; 1992.
- NORD, C. *Text analysis in translation*. Amsterdam: Rodopi; 1991.
- POSTEGUILLO, S. A comparative genre analysis of computer science English: contending rhetorics. *English for Specific Purposes*, 2000, v. 21, n. 1, p. 43-69.
- PURVES, A.C. (ed.). *Writing across languages and cultures: issues in contrastive rhetoric*. Newbury Park: Sage Publications; 1988.
- ROJO, R. *Gêneros do discurso e gêneros textuais*. Disponível em: www.fae.ufmg.br/ceale/generosdiscurso.pdf. Acessado em 3 de outubro de 2004.
- RONDEAU, G. *Introduction à la terminologie*. Chicoutimi (Quebec): Gaëtan Morin; 1983.
- SÁEZ, F.T. Escritura y cultura: la retórica contrastiva. *Pragmalingüística*, 2001, n. 8-9, p. 313-329. Disponível em: www.ugr.es/~ftsaez/cr1.pdf. Acessado em 3 de outubro de 2004.

- _____. Culture in writing: discourse markers in English and Spanish student writing. In: Departamento de Didáctica de la Lengua y la Literatura (eds.). *Tadea seu liber de Amicitia*. Granada: Imprenta Generalife; 2003. p. 345-364. Disponível em: <http://www.ugr.es/~ftsaez/tadea.pdf>.
- SANTOS, V.B.M.P. Estabelecendo as diferenças entre os termos registro e gênero. *English for Specific Purposes*, 1996, v. 19, n. 1, p. 1-40.
- SIEPMANN, D. Multi-word discourse markers in translation. *Lebende Sprachen*, 2001, v. 46, n. 3, p. 97-107.
- SWALES, J.M. *Genre analysis. English in academic and research settings*. Cambridge: Cambridge University Press; 1990.
- _____. Re-thinking genre: another look at discourse community effects. In: *Re-thinking Genre Colloquium*. Ottawa: Carleton University; 1992 [mimeo].
- TODOROV, T. *Poétique de la prose*. Paris: Seuil; 1971.
- VALERO-GARCÉS, C. Contrastive ESP rhetoric: metatext in Spanish and English economics texts. *English for Specific Purposes*, 1996, v. 15, n. 4, p. 279-294.
- VANDE KOPPLE, W. Some exploratory discourse on metadiscourse. *College Composition and Communication*, 1985, v. 36, p. 82-93.