

RENATA BENEDEZI

**COLOCAÇÕES SUBSTANTIVO + ADJETIVO:
PROPOSTAS PARA SUA IDENTIFICAÇÃO E TRATAMENTO LEXICOGRÁFICO
EM DICIONÁRIOS ATIVOS PORTUGUÊS-ESPAÑOL**

PORTO ALEGRE

2008

RENATA BENEDUZI

**COLOCAÇÕES SUBSTANTIVO + ADJETIVO:
PROPOSTAS PARA SUA IDENTIFICAÇÃO E TRATAMENTO LEXICOGRÁFICO
EM DICIONÁRIOS ATIVOS PORTUGUÊS-ESPANHOL**

Dissertação de Mestrado em Terminologia e Lexicografia: Relações Textuais apresentada como requisito parcial para a obtenção do título de Mestre em Estudos da Linguagem pelo Programa de Pós-Graduação em Letras da Universidade Federal do Rio Grande do Sul

Orientador: Prof. Dr. Félix V. Bugueño Miranda

PORTO ALEGRE

2008

RENATA BENEDUZI

**COLOCAÇÕES SUBSTANTIVO + ADJETIVO:
PROPOSTAS PARA SUA IDENTIFICAÇÃO E TRATAMENTO LEXICOGRÁFICO
EM DICIONÁRIOS ATIVOS PORTUGUÊS-ESPANHOL**

Dissertação de Mestrado em Terminologia e Lexicografia: Relações Textuais apresentada como requisito parcial para a obtenção do título de Mestre em Estudos da Linguagem pelo Programa de Pós-Graduação em Letras da Universidade Federal do Rio Grande do Sul

Aprovado em 29 de abril de 2008.

BANCA EXAMINADORA

Prof^a. Dr^a. Cleci Regina Bevilacqua

Instituto de Letras - UFRGS

Prof^a. Dr^a. Erica Sofia Luisa Foerthmann Schultz

Instituto de Letras - UFRGS

Prof. Dr. Werner Heidermann

Centro de Comunicação e Expressão - UFSC

*Ao Leandro,
inspiração pra tudo o que eu viver.*

AGRADECIMENTOS

A Deus, que torna possível todas as coisas.

Aos meus pais, Antônio e Vera, por terem me proporcionado as melhores oportunidades de estudo; pelo amor incondicional e constante dedicação. Ao meu irmão, Luis Fernando, pelo exemplo acadêmico e, especialmente, pelo envio da Europa de grande parte da bibliografia, a qual dificilmente eu teria acesso sem seu auxílio.

Ao Leandro, meu companheiro de todas as horas, pelo amor e dedicação nos momentos mais difíceis. Por sua presença constante e participação ativa na construção deste trabalho.

Aos meus segundos pais, Aurélio e Nara, pelo carinho e pela compreensão das ausências. Ao Alessandro e à Josi, pelo material sempre prontamente retirado na Unisinos.

Aos colegas da Universidade, especialmente à Karla, por sua disponibilidade em me auxiliar a qualquer hora em algumas traduções especialmente complicadas do inglês.

A todos os professores do Programa de Pós-Graduação que contribuíram em minha formação. Especialmente ao meu orientador, Prof. Dr. Félix Bugueño Miranda, por sua dedicação profissional e por ter me possibilitado o acesso à pesquisa científica desde o início de minha graduação, contribuindo muito para meu crescimento acadêmico. Também à Prof^a. Dr^a. Cleci Regina Bevilacqua, pela solicitude ao me disponibilizar sua bibliografia sobre o tema das colocações e pelas importantes discussões em sua disciplina de “Fraseologia”.

À Prof^a. Cecília de Leão Martins, do Núcleo de Assessoria Estatística da UFGRS, pelo seu interesse e disponibilidade em me auxiliar sempre que necessário.

Finalmente, agradeço à CAPES, pela concessão da bolsa de pesquisa, que possibilitou minha dedicação exclusiva ao desenvolvimento deste trabalho.

RESUMO

As colocações correspondem a combinações léxicas idiossincráticas que representam uma grande dificuldade para aprendizes e tradutores de uma língua estrangeira. Apesar de sua importância no estudo de idiomas, esse fenômeno léxico não apresenta sequer uma definição conceitual unívoca. Tal fato aparece refletido na impossibilidade de aplicação de uma metodologia de análise que permita delimitar tais combinações de forma objetiva e conclusiva. Essa dificuldade em seu reconhecimento resulta, ainda, em uma escassez de materiais para seu ensino-aprendizagem e, mais especificamente, de obras lexicográficas que abordem tais estruturas. Em virtude dessas constatações, o objetivo desta dissertação é, por um lado, proporcionar uma maior delimitação conceitual para o fenômeno léxico das colocações e, por outro lado, apresentar estratégias para sua representação lexicográfica. Para isso, o presente trabalho divide-se em três grandes etapas. A primeira delas consiste em, a partir de uma revisão da literatura especializada, identificar as principais propriedades que caracterizam esse fenômeno léxico. Em seguida, realiza-se um estudo das combinações formadas exclusivamente por substantivos e adjetivos, através de diferentes perspectivas de análise: quantitativa, qualitativa e contrastiva. A partir de um *corpus* textual composto por diversos jornais de língua espanhola, os candidatos a colocações são selecionados, primeiramente, por sua frequência de coaparição. Sobre essas estruturas, são identificadas as relações semânticas existentes entre substantivos e adjetivos, através de uma taxonomia realizada previamente para esta classe de palavras. Finalmente, o reconhecimento das colocações dá-se através do contraste dessas combinações entre as línguas portuguesa e espanhola. Como resultado desse processo, obtém-se, a partir das combinações do *corpus*, uma lista de colocações entre esses dois idiomas. Somente de posse dessas informações, é possível elaborar algumas propostas para o tratamento lexicográfico das colocações, levando em conta o perfil de um possível usuário e a função da obra. A partir deste estudo, percebe-se que somente através de uma análise detalhada dos candidatos a colocações, considerando diferentes perspectivas de investigação, é possível compreender essas combinações em toda sua complexidade estrutural e proporcionar melhores estratégias para sua inclusão nos dicionários.

PALAVRAS-CHAVE: Colocações. Dicionários Bilíngües. Lexicografia. Metalexigrafia.

RESUMEN

Las colocaciones corresponden a combinaciones léxicas idiosincrásicas que representan una gran dificultad para aprendices y traductores de una lengua extranjera. Aunque sea importante en el estudio de idiomas, ese fenómeno léxico no presenta ni siquiera una definición conceptual unívoca. Este hecho aparece reflejado en la imposibilidad de aplicación de una metodología de análisis que permita delimitar dichas combinaciones de forma objetiva y conclusiva. Esta dificultad en su reconocimiento también resulta en una escasez de materiales para su enseñanza-aprendizaje y más específicamente de obras lexicográficas que traten de dichas estructuras. En virtud de las constataciones anteriores, el objetivo de este trabajo es, por un lado, proporcionar una delimitación conceptual más precisa para el fenómeno léxico de las colocaciones y, por otro lado, presentar estrategias para su representación lexicográfica. Para tanto, este estudio se divide en tres grandes etapas. La primera consiste en identificar las principales propiedades que caracterizan ese fenómeno léxico a partir de una revisión de la literatura especializada. Enseguida, se realiza un estudio de las combinaciones formadas exclusivamente por sustantivos y adjetivos a través de diferentes perspectivas de análisis: cuantitativo, cualitativo y contrastivo. A partir de un *corpus* textual formado por distintos periódicos de lengua española, se seleccionan los candidatos a colocaciones, primeramente, por su frecuencia de coaparición. Basado en estas estructuras, se identifican las relaciones semánticas existentes entre sustantivos y adjetivos, a través de una taxonomía realizada previamente para esta clase de palabras. Finalmente, el reconocimiento de las colocaciones ocurre por el contraste de dichas combinaciones entre las lenguas portuguesa y española. Como resultado de ese proceso, se obtiene, a partir de las combinaciones del *corpus*, una relación de colocaciones entre esos dos idiomas. Sólo con esas informaciones se pueden elaborar algunas propuestas para el tratamiento lexicográfico de las colocaciones, teniendo en cuenta el perfil de un posible usuario y la función de la obra. A partir de este estudio, se percibe que solamente a través de un análisis detallado de los candidatos a colocaciones, que considere diferentes perspectivas de investigación, se puede comprender dichas combinaciones en toda su complejidad estructural y proporcionar mejores estrategias para su inclusión en los diccionarios.

PALABRAS CLAVE: Colocaciones. Diccionarios Bilingües. Lexicografía. Metalexigrafía.

LISTA DE ILUSTRAÇÕES

Quadro 1 - Síntese dos pressupostos teóricos	56
Figura 1 - Exemplo de aplicação da ferramenta <i>Collocate</i>	73
Figura 2 - Exemplo de aplicação da ferramenta <i>Concord</i>	74
Esquema 1 - Taxonomia dos adjetivos	104
Quadro 2 - Exemplos de adjetivos classificados	105
Gráfico 1 - Distribuição dos adjetivos segundo sua classificação.....	106
Esquema 2 - Modelo de Análise Contrastiva - Nível léxico-semântico	127
Esquema 3 - Modelo de Análise Contrastiva - Nível morfossintático	128
Esquema 4 - Adaptação do Modelo de Análise Contrastiva para as colocações	128
Esquema 5 - Níveis de conhecimento propostos pelo MARCO (2002)	152
Figura 3 - Modelo de verbete retirado do DiBU (2003, s.v. <i>tocar</i>).....	164
Figura 4 - Modelo de verbete retirado do DiBU (2003, s.v. <i>coral</i>).....	166

LISTA DE TABELAS

Tabela 1 - Aplicação do Teste Exato de Fisher às variáveis <i>cuervo</i> e <i>desnudo</i>	76
Tabela 2 - Tabela-modelo do Teste Exato de Fisher	76

LISTA DE ABREVIATURAS E SIGLAS

ac.	- acepção
cf.	- confira
COL	- colocações
etc.	- <i>et cetera</i> , e outros
i.e.	- isto é
LOC	- locuções
p.	- página
sc.	- <i>scilicet</i> , a saber
s.d.	- sem data
sic	- assim, precisamente como o texto apresenta
s.p.	- sem página
s.v.	- <i>sub voce</i> , sob o lema
v.	- ver

LISTA DE SÍMBOLOS

- ✓ - sim
- ✕ - não
- ∅ - não consta na teoria
- ◊ - indicador das colocações nos dicionários
 - marcador de inversão da estrutura
 - indicador das colocações exclusivas em língua estrangeira

SUMÁRIO

1 INTRODUÇÃO	15
1.1 Objetivos	17
1.2 Hipóteses de pesquisa	18
1.3 Estrutura geral do trabalho	18
2 PRESSUPOSTOS TEÓRICOS: UMA REVISÃO DA LITERATURA	21
2.1 Enfoque lexicológico	23
2.1.1 Abordagem semântica.....	23
2.1.1.1 <i>Bally (1951 [1909])</i>	23
2.1.1.2 <i>Porzig (1970 [1950])</i>	25
2.1.1.3 <i>Coseriu (1977 [1967])</i>	27
2.1.1.4 <i>Mel'čuk (1995, 2003)</i>	30
2.1.1.5 <i>Corpas Pastor (1996, 2001)</i>	33
2.1.1.6 <i>Síntese da abordagem semântica</i>	37
2.1.2 Abordagem estatística.....	38
2.1.2.1 <i>Firth (1957)</i>	38
2.1.2.2 <i>Halliday (1961, 1966)</i>	39
2.1.2.3 <i>Sinclair (1991)</i>	41
2.1.2.4 <i>Tagnin (2005a)</i>	42
2.1.2.5 <i>Síntese da abordagem estatística</i>	43
2.2 Enfoque lexicográfico	44
2.2.1 <i>Cowie (1981, 1994)</i>	45
2.2.2 <i>Haensch et al (1982)</i>	46
2.2.3 <i>Benson (1985, 1986, 1997)</i>	47
2.2.4 <i>Seco (1987a, 1987b)</i>	49
2.2.5 <i>Hausmann (1979, 1985, 1989, 1998a)</i>	51
2.2.6 <i>Síntese do enfoque lexicográfico</i>	54
2.3 Síntese dos pressupostos teóricos	55
2.4 Delimitação do objeto de estudo	56
3 ELABORAÇÃO DO CORPUS DE ESTUDO	59
3.1 Reconhecimento e estruturação da área	60
3.2 Constituição do corpus	63
3.3 Caracterização do corpus	67
3.4 Coleta e organização dos textos	68
4 ANÁLISE QUANTITATIVA: RELAÇÕES DE FREQUÊNCIA	71
4.1 Coleta dos dados	71
4.2 Avaliação dos resultados	75
4.3 Limitações da análise quantitativa	79
5 ANÁLISE QUALITATIVA: RELAÇÕES SEMÂNTICAS	81
5.1 Natureza do adjetivo	83
5.1.1 Enfoque formal.....	84
5.1.1.1 <i>Bello (1942)</i>	84
5.1.1.2 <i>Real Academia Española (1973)</i>	85

5.1.1.3 <i>Gili Gaya (1979)</i>	86
5.1.1.4 <i>Alcina, Blecua (1987)</i>	87
5.1.1.5 <i>Weinrich (1989)</i>	88
5.1.1.6 <i>Alarcos Llorach (1999)</i>	90
5.1.2 Enfoque semântico	91
5.1.2.1 <i>Demonte (1999)</i>	91
5.2 Necessidade de um novo tratamento dos adjetivos	92
5.3 Taxonomia dos adjetivos	94
5.3.1 Adjetivos determinativos	95
5.3.2 Adjetivos qualificativos	95
5.3.2.1 <i>Adjetivos participios</i>	95
5.3.2.2 <i>Adjetivos relacionais</i>	97
5.3.2.3 <i>Adjetivos situacionais</i>	98
5.3.2.4 <i>Adjetivos qualificativos stricto sensu</i>	99
5.3.2.4.1 <i>Adjetivos antepostos</i>	99
5.3.2.4.2 <i>Adjetivos pospostos</i>	100
5.4 Resultados da análise qualitativa	104
5.5 Limitações da análise qualitativa	107
6 ANÁLISE CONTRASTIVA: RELAÇÕES ENTRE AS LÍNGUAS	109
6.1 Apresentação de uma nova perspectiva de análise	109
6.2 Fundamentos de estruturação de um modelo de análise contrastivo	112
6.2.1 Princípio do relativismo lingüístico	113
6.2.2 Princípio do anisomorfismo lingüístico	114
6.2.3 Classificação das línguas	117
6.2.4 Transferência entre as línguas	119
6.2.4.1 <i>Análise Contrastiva</i>	120
6.2.4.2 <i>Análise de Erros</i>	122
6.2.4.3 <i>Interlíngua</i>	124
6.3 Uma proposta de análise contrastiva para a identificação das colocações	126
6.4 Aplicação do modelo de análise	129
6.4.1 Análise dos dados	130
6.4.1.1 <i>Equivalentes “literais”</i>	130
6.4.1.2 <i>Posição divergente do adjetivo</i>	131
6.4.1.3 <i>Seleção idiosincrática do colocado</i>	132
6.4.1.4 <i>Estrutura morfológica divergente</i>	134
6.4.1.5 <i>Combinações particulares a uma língua</i>	135
6.4.2 Resultados da Análise Contrastiva	137
6.5 Limitações da Análise Contrastiva	141
7 TRATAMENTO LEXICOGRÁFICO DAS COLOCAÇÕES EM DICIONÁRIOS BILÍNGÜES	144
7.1 Parâmetros teóricos para a elaboração de dicionários bilíngües	146
7.1.1 Direção e função do dicionário	146
7.1.2 Perfil do usuário	149
7.1.2.1 <i>Marco Común Europeo de Referencia para las Lenguas: aprendizaje, enseñanza, evaluación</i>	151
7.1.2.1.1 <i>Níveis de conhecimento propostos pelo MARCO (2002)</i>	151
7.1.2.2 <i>Perfil do usuário de um dicionário bilíngüe ativo português-espanhol</i>	153
7.1.3 Componentes canônicos do dicionário	155

7.1.3.1 <i>Macroestrutura</i>	156
7.1.3.2 <i>Microestrutura</i>	160
7.1.3.3 <i>Medioestrutura</i>	167
7.1.3.4 <i>Partes introdutórias</i>	168
7.2 Necessidade de um novo tratamento lexicográfico das colocações: análise de dicionários bilíngües espanhol-português/português-espanhol	169
7.3 Propostas para o tratamento lexicográfico das colocações em dicionários bilíngües ativos português-espanhol	172
7.3.1 Dicionário bilíngüe ativo português-espanhol para usuários de nível “Plataforma”	173
7.3.2 Dicionário bilíngüe ativo português-espanhol para usuários de nível “Avançado”	176
8 CONCLUSÃO	181
REFERÊNCIAS BIBLIOGRÁFICAS	186
APÊNDICE A - DOCUMENTAÇÃO DO CORPUS NO <i>MICROSOFT ACCESS</i>	200
ANEXO A - LISTA DAS FUNÇÕES LÉXICAS DE MEL’ČUK (1988)	202
ANEXO B - LISTA DE ETIQUETAS MORFOSSINTÁTICAS DO VISL	206
ANEXO C - EXEMPLO DE TEXTO ETIQUETADO PELO VISL	210

1 INTRODUÇÃO

Existem determinados fenômenos lingüísticos perante os quais, à primeira vista, parece haver certo consenso, mais ou menos explícito, na literatura especializada. No entanto, uma análise mais aprofundada desses fenômenos revela que apresentam uma maior complexidade estrutural. Como exemplo de tal fato, encontram-se as colocações, objeto de estudo deste trabalho, as quais são caracterizadas, segundo a já clássica definição de Cop (1991), como combinações léxicas situadas entre a combinatória livre e as expressões idiomáticas. Hausmann (1989) acrescenta como traços definitórios desse fenômeno seu caráter transparente (em oposição à opacidade semântica das expressões idiomáticas) e sua combinabilidade restrita (em oposição à combinatória livre, regida unicamente pelos princípios da gramaticalidade, seja ela formal e/ou semântica, e da aceitabilidade). Tais definições, recorrentes na literatura especializada sobre o tema, demonstram-se, porém, demasiadamente vagas e abrangentes, visto que não possibilitam uma delimitação conceitual precisa das combinações que constituem o fenômeno léxico das colocações.

Na tentativa de propor uma conceituação mais específica dessas combinações léxicas, surgem, então, diversos estudos sobre o tema das colocações (conforme poderá ser observado na multiplicidade de autores tratados no capítulo 2 do presente trabalho). No entanto, devido à ausência de uma conceituação unívoca que permita delimitar tais combinações, os estudos são realizados a partir de diferentes abordagens metodológicas (que partem de perspectivas semânticas, estatísticas ou lexicográficas), o que reflete em resultados práticos bastante divergentes. Com isso, sob o mesmo rótulo de “colocações”, encontramos um grupo muito heterogêneo de combinações léxicas.

A falta de clareza conceitual que existe a respeito dessas combinações, aliada à inexistência de uma metodologia conclusiva que permita delimitá-las com precisão, aparece refletida na escassez de subsídios que possibilitem seu ensino e aprendizagem. Consideramos que seja importante suprir essa ausência porque o emprego adequado das colocações por parte dos estudantes de uma língua estrangeira revela a alta fluência desses falantes no idioma que estão aprendendo. Isso se deve ao fato de a seleção léxica realizada em cada combinação obedecer a um critério arbitrário, próprio de cada língua particular. Em vista disso, as colocações correspondem a estruturas intuitivas para os falantes nativos de uma língua, de forma que eles não têm consciência de seu uso (cf. ALONSO RAMOS, 2001, p. 11). Tal fato

permite compreender porque essas combinações representam uma grande dificuldade para os falantes não-nativos de um idioma.

Entretanto, apesar da importância do reconhecimento dessas combinações léxicas por parte dos estudantes de uma língua estrangeira, apenas em língua inglesa pode ser encontrada uma grande quantidade de obras que tratem das colocações (cf. TAGNIN, 2005b, p. 200). Para a língua espanhola, praticamente não existem instrumentos disponíveis que possibilitem aos estudantes o acesso a essas combinações, tanto no que diz respeito ao seu ensino quanto ao seu tratamento lexicográfico (cf. ALONSO RAMOS, 2002a, p. 551; FARGHAL, OBIEDAT, 1995, p. 318). Na lexicografia bilíngüe espanhol-português nacional (e também na de origem hispânica), não encontramos indícios sobre o tratamento das colocações. A partir da análise de diferentes dicionários bilíngües envolvendo esses dois idiomas, pudemos identificar uma carência na inclusão dessas combinações em obras lexicográficas (cf. BENEDUZI, 2007).

Para suprir essa deficiência, consideramos pertinente a realização de uma reflexão sobre o problema das colocações com vistas a proporcionar um tratamento mais adequado dessas combinações em obras lexicográficas. Acreditamos que a inclusão das colocações deva ser realizada especialmente em dicionários bilíngües, pois essas obras, além de serem as mais consultadas pelos estudantes de uma língua estrangeira (cf. LAUFER, LEVITSKY-AVIAD, 2006, p. 136), ressaltam as divergências existentes entre os idiomas contrastados. As relações de contraste entre as línguas materna e estrangeira, possibilitadas pelos dicionários bilíngües, são fundamentais no estudo das colocações, uma vez que a principal dificuldade na produção dessas combinações em língua estrangeira deriva da seleção léxica idiossincrática de seus elementos. Assim, por exemplo, para indicar a intensidade de um erro, a língua espanhola utiliza a combinação *error mayúsculo*, enquanto em língua portuguesa, para exprimir o mesmo significado, utilizamos a expressão *erro gritante*.

A identificação dessas estruturas bastante divergentes entre duas línguas aparentemente tão próximas como o português e o espanhol motivou nosso interesse pelo estudo das colocações. Além disso, como não existe nenhum trabalho que ofereça subsídios mais definitivos sobre esse tema, inquietava-nos as razões para um fenômeno lingüístico, que consideramos tão importante na aquisição de uma língua estrangeira, permanecer com tantas lacunas. Em vista disso, surgiram nossos primeiros questionamentos:

- a) as colocações, enquanto combinações léxicas situadas entre as estruturas livres e as fixas, possuem um conjunto de características específicas que permita

identificá-las e opô-las claramente às extremidades constituídas, de um lado, pelo “discurso livre” e, de outro, pelo “discurso repetido”¹?

- b) considerando que essas combinações representam tanta dificuldade para os estudantes de uma língua estrangeira, como incluí-las em obras lexicográficas a fim de facilitar seu aprendizado?

Desenvolvemos, então, esta dissertação em função dos questionamentos apresentados, os quais evidenciam uma necessária convergência entre três diferentes disciplinas: Lingüística de *Corpus*, Lingüística Aplicada e Metalexicografia Bilíngüe. As duas primeiras são responsáveis por oferecer um suporte teórico para a identificação das colocações, enquanto a última proporciona uma reflexão sobre a melhor forma de seu tratamento lexicográfico.

1.1 Objetivos

A partir de nossos questionamentos iniciais e do pressuposto de que não há uma definição unívoca para o problema das colocações, nosso trabalho objetiva:

- a) delimitar conceitualmente o fenômeno léxico das colocações;
- b) proporcionar estratégias para uma representação lexicográfica desse fenômeno.

Para alcançar esses objetivos gerais, porém, será necessário cumprir os seguintes objetivos específicos:

- a) comprovar o poder heurístico das diferentes metodologias aplicadas na detecção e tratamento das colocações;
- b) elaborar um *corpus* de textos em língua espanhola que sirva de base de dados para aplicar as diferentes metodologias utilizadas;
- c) detectar, com base no *corpus* compilado, se é possível isolar determinados padrões lingüísticos que permitam compreender melhor a natureza do fenômeno léxico das colocações;

¹ Essa terminologia é empregada por Coseriu (1980).

- d) procurar proporcionar uma integração entre os resultados obtidos a partir de diferentes disciplinas, quais sejam, Lingüística de *Corpus*, Lingüística Aplicada e Metalexigrafia no tratamento das colocações.

1.2 Hipóteses de pesquisa

Nossas hipóteses derivam de uma intersecção dos objetivos gerais propostos para este trabalho. Tais objetivos partem do pressuposto de que não existe uma delimitação conceitual de consenso para as colocações, o que dificulta a elaboração de instrumentos lexicográficos que auxiliem em seu aprendizado. Em função dessa premissa, formulamos as seguintes hipóteses de pesquisa:

- a) somente através da aplicação de metodologias consecutivas é possível obter resultados mais precisos no reconhecimento das colocações;
- b) uma proposta de tratamento lexicográfico dessas combinações em dicionários que auxiliem o ensino/aprendizagem de uma língua estrangeira deve fundamentar-se nas particularidades lingüísticas tanto da língua materna do estudante quanto da língua que está aprendendo.

Como consequência de nossas hipóteses, consideramos que esta dissertação deva estruturar-se a partir de três âmbitos principais. O primeiro deles centra-se na obtenção de subsídios que permitam estabelecer uma definição para as colocações. O segundo âmbito contempla uma reflexão sobre como integrar diferentes perspectivas de estudo da linguagem para alcançar uma compreensão mais ampla desse fenômeno léxico. Finalmente, a última parte de nosso trabalho reserva-se à geração de propostas lexicográficas concebidas à luz dos resultados obtidos nos dois primeiros âmbitos de investigação.

1.3 Estrutura geral do trabalho

A fim de atingir os objetivos propostos e procurar comprovar nossas hipóteses de pesquisa, organizamos nossa dissertação da forma como expomos a seguir.

Iniciamos nosso trabalho com uma revisão da literatura que tratou, explicitamente ou de uma forma mais geral, sobre o tema das colocações. Procuramos, a partir das características de consenso entre os diversos autores, isolar alguns critérios que permitam o reconhecimento dessas combinações léxicas. No entanto, devido à ampla diversidade de estruturas que configuram esse fenômeno léxico, optamos por analisar somente as combinações formadas por substantivos e adjetivos². Esse percurso é realizado no segundo capítulo.

A fim de termos acesso às combinações léxicas realmente utilizadas pelos falantes, elaboramos um *corpus* formado por textos de diferentes jornais de língua espanhola, o qual serviu de base de dados para nossa pesquisa. As informações sobre coleta e caracterização de nosso *corpus* encontram-se no capítulo três.

Para proceder à identificação das colocações, realizamos uma análise quantitativa dessas combinações léxicas a partir da frequência de coaparição dos elementos em nosso *corpus*. Os resultados dessa análise aparecem contemplados no capítulo quatro.

No quinto capítulo, realizamos uma análise qualitativa das combinações obtidas a partir do estudo desenvolvido no capítulo anterior, a fim de identificar as relações semânticas estabelecidas entre os elementos que conformavam cada combinação. Com base nos trabalhos de diferentes autores, elaboramos uma classificação semântica dos adjetivos, com o intuito de identificar aqueles que apresentam uma maior tendência a formar colocações.

O capítulo seis discute a importância de considerar as características específicas de cada língua particular no estudo das colocações, culminando com uma proposta de análise contrastiva desse fenômeno léxico entre as línguas portuguesa e espanhola.

A partir dos resultados oferecidos pelas análises realizadas nos capítulos anteriores e de posse de conceitos fundamentais da prática metalexiconográfica, apresentamos, no capítulo sete, algumas propostas para um tratamento lexicográfico mais adequado do fenômeno das colocações.

Finalmente, no oitavo capítulo, encerramos nossa dissertação apresentando algumas considerações finais e retomando nossas hipóteses iniciais de pesquisa. Com isso, esperamos que o presente trabalho possa contribuir para ampliar a discussão em torno do

² Os adjetivos correspondem às palavras que se unem aos substantivos para exprimir a qualidade do objeto, do ser ou da noção designados pelos substantivos - nos casos de adjetivos qualificativos -, ou, ainda, para permitir a atualização dos substantivos na frase - nos casos de adjetivos determinativos (DUBOIS et al, 1999, s.v. *adjectif*). Para a língua portuguesa, no entanto, de acordo com a Nomenclatura Gramatical Brasileira (N.G.B.), somente são considerados adjetivos os chamados qualificativos, uma vez que os determinativos foram todos incluídos entre os pronomes (JOTA, 1976, s.v. *adjetivo*).

problema das colocações, tanto no que concerne a sua identificação quanto a seu tratamento em obras lexicográficas.

2 PRESSUPOSTOS TEÓRICOS: UMA REVISÃO DA LITERATURA

Apesar de o termo “colocação” [*collocation*] ter surgido somente com Firth (1957)³, tal noção já se vislumbrava em Saussure (2004 [1916]) através das relações sintagmáticas e associativas entre as palavras⁴. Desde então, diversos autores trataram o tema das colocações - sob diferentes denominações -, resultando em uma multiplicidade de caracterizações. Segundo Smadja (1993, p. 145), “dependendo de seus interesses e pontos de vista, os pesquisadores enfocaram diferentes aspectos das colocações”⁵. A partir disso, no presente capítulo, pretendemos oferecer um panorama não apenas cronológico dos estudos desenvolvidos sobre essas combinações, mas também organizar as diversas abordagens de forma a ressaltar suas semelhanças e diferenças⁶.

Na tentativa de oferecer uma classificação com maior sistematicidade e clareza dos trabalhos realizados sobre as colocações, apresentamos, primeiramente, uma divisão entre as duas principais perspectivas das quais partiram os estudos desenvolvidos sobre esse tema. De um lado, identificamos um enfoque puramente lexicológico (v. 2.1), representado por autores que se preocuparam unicamente com a descrição lingüística desse fenômeno, isto é, que procuraram caracterizar as colocações e identificar as relações estabelecidas entre os elementos que as compõem. Por outro lado, encontramos um enfoque lexicográfico (v. 2.2), no qual se incluem os autores que desenvolveram um estudo metalexigráfico das

³ Segundo Mitchell (1975, p. 134), o termo “colocação” não é original em Firth, mas parece ter sido inspirado em uma monografia escrita por Palmer sobre esse tema. Infelizmente, não tivemos acesso a essa publicação.

⁴ Não aprofundaremos, porém, os estudos desenvolvidos por Saussure (2004), pois esse autor enfatiza as relações paradigmáticas, quando acreditamos que as colocações situam-se no eixo sintagmático. Além disso, embora seu estudo possa ser ampliado para a combinatoria léxica, Saussure (2004) centra-se na análise apenas morfológica das palavras (cf. BEVILACQUA, 1996, p. 20).

⁵ [depending on their interests and points of view, researchers have focused on different aspects of collocations]. Esta e as demais traduções do presente trabalho foram por nós realizadas.

⁶ Koike (2000) oferece um amplo panorama dos estudos sobre as colocações através de dezoito diferentes autores; sua apresentação, porém, é organizada somente sob a base cronológica. Pazos Bretaña (2005), por sua vez, inicia com uma sistematização das colocações a partir das diferentes teorias das quais os estudos partem, elencando, assim, a escola britânica, a semântica léxica e a lexemática, o que resulta em praticamente um autor para cada escola. Após essa divisão, porém, Pazos Bretaña (2005) acrescenta, ainda, um grande grupo intitulado “lexicologia e lexicografia”, no qual inclui uma diversidade de autores. Finalmente em Maurer-Stroh (2004), encontramos uma maior ordenação dos estudos sobre as colocações, divididos nas seis maiores “linhas de pensamento colocacional” [*lines of 'collocational thinking'*]: abordagem pedagógica, fraseológica russa, fraseológica (contemporânea), lexicográfica, lingüística (baseada na frequência) e computacional. Acreditamos, porém, que essa classificação não é orientada por critérios coerentes em sua estruturação dos estudos sobre as colocações, pois, se a divisão se estabelece a partir de “linhas de pensamento”, não há razão para as abordagens fraseológicas serem separadas por critérios eminentemente cronológicos. Além disso, consideramos que a abordagem lingüística não se constitui como tal, pois os autores apresentados nessa seção entendem o fenômeno das colocações apenas como frequência de coaparição, sem que sejam estabelecidas relações lingüísticas entre os elementos.

colocações, ou seja, aqueles que se preocuparam, principalmente, com a forma como essas combinações devem ser apresentadas em um dicionário e que, a partir dessa questão inicial, procuraram também estabelecer as relações entre seus elementos constitutivos. Dessa forma, percebemos que o primeiro enfoque reflete a necessidade de compreender a complexidade do fenômeno léxico das colocações e o segundo, além dessa tarefa, procura transportar os conhecimentos adquiridos à elaboração de obras lexicográficas⁷.

É importante salientar, porém, que, além dessas diferenças de enfoque, há também uma divergência no estudo das colocações conforme sejam tratadas como um fato ontológico da linguagem ou um fato de uma língua particular. No primeiro caso, são investigadas explicações para o porquê de determinadas palavras aparecerem combinadas, independentemente da língua em que se realiza o estudo. Essa concepção ontológica aparece refletida nos trabalhos com uma abordagem mais voltada à identificação das relações semânticas entre os elementos de uma colocação. Essas relações podem ser representadas através de universais lingüísticos (como as funções léxicas⁸ propostas por Mel'čuk (1995)), sendo, portanto, aplicadas concretamente às diversas línguas. Já no segundo, as pesquisas indicam como essas combinações manifestam-se em uma língua particular, sendo mais enfatizados os aspectos formais das colocações. Essa concepção reflete-se nos estudos que pretendem classificar tais combinações a partir de seu comportamento morfossintático - como Hausmann (1989) e Benson (1997) -, assim como nas pesquisas realizadas a partir de uma abordagem estatística, como Jones, Sinclair (1974). As primeiras propõem uma categorização que só pode ser aplicada à língua específica em que se desenvolve o estudo, enquanto as segundas buscam identificar as recorrências frequentes de palavras dentro de uma língua particular. Acreditamos que esses diferentes enfoques para o mesmo fenômeno léxico também devam ser considerados para explicar as divergências entre os teóricos no momento de delimitar empiricamente seu objeto de estudo.

⁷ Salientamos que nossa classificação, em oposição às classificações naturais, realizadas a partir da própria natureza ou essência do classificado, corresponde a uma classificação artificial, pois é fundamentada em características arbitrariamente selecionadas (cf. CENTRO DI STUDI FILOSOFICI DI GALLARATE, 1957, s.v. *classificazione*; FOULQUIÉ, 1967, s.v. *clasificación*). Dessa forma, temos consciência de que outras categorizações, igualmente válidas, seriam possíveis.

⁸ Uma “função léxica”, segundo Mel'čuk (1984, p. 6) corresponde a “um significado abstrato com uma certa relação sintático-semântica tal que sua expressão lingüística depende do lexema ao qual essa função léxica se aplica” [un sens assez abstrait (...) ou un certain rôle sématique-syntaxique (...) tel que son expression linguistique dépend du lexème auquel cette FL s'applique]. Retomaremos esse conceito em 2.1.1.4.

2.1 Enfoque lexicológico

Os estudos que abordaram as colocações sob o ponto de vista lexicológico podem, ainda, ser divididos em duas abordagens principais, as quais, para o reconhecimento desse fenômeno léxico, procuraram identificar as relações semânticas (v.2.1.1) ou de frequência (v.2.1.2) estabelecidas entre os membros das combinações.

2.1.1 Abordagem semântica⁹

Os trabalhos dessa abordagem visam a identificar as relações estabelecidas, no plano do conteúdo, entre os elementos que conformam cada combinação. Alguns autores, embora não tratem especificamente do tema das colocações, constituem estudos pioneiros que chamaram a atenção para o problema da combinatória léxica. A seguir, apresentamos os trabalhos dessa abordagem em uma ordenação cronológica, na tentativa de recuperar o raciocínio que foi sendo desenvolvido ao longo dos estudos.

2.1.1.1 Bally (1951 [1909])

Bally (1951 [1909]), seguindo a distinção Saussureana entre relações associativas e sintagmáticas¹⁰, identifica que os fatos da língua são assimilados por associações e agrupamentos. Percebe, porém, que há uma gradação entre esses agrupamentos e divide-os entre os “agrupamentos passageiros” [*groupements passagers*], isto é, aqueles que são livres e se decompõem rapidamente, formando novas combinações, e as “unidades indecomponíveis”

⁹ O termo “semântica” será empregado neste trabalho para designar o estudo do significado dos signos lingüísticos (cf. GLÜCK, 2000, s.v. *Semantik*), mais especificamente o que Glück (2000, s.v. *Lexikalische Semantik*) define como a análise da significação estrutural das unidades lexicais. Segundo esse autor, à semântica léxica pertencem as seguintes áreas de pesquisa: (a) o estudo de estruturas de significação léxica interna; (b) o estudo das estruturas paradigmáticas de campos léxicos e (c) o estudo das possibilidades combinatórias sintagmáticas dos lexemas. No âmbito descrito em (c), que será focado em nosso trabalho, a semântica léxica apresenta pontos de contato com a semântica frasal.

¹⁰ Apesar de o texto original de Bally ser de 1909, enquanto que a publicação da obra de Saussure data de 1916, Bally foi discípulo de Saussure e, portanto, inspirou-se em suas idéias.

[*unités indécomposables*], que são fixas e nas quais as palavras, por serem utilizadas constantemente em conjunto, perdem sua autonomia, não podem ser separadas e adquirem seu significado¹¹ somente na combinação. Entre esses “casos extremos”, situa o que chama de “séries fraseológicas” [*séries phraséologiques*] ou “agrupamentos usuais” [*groupements usuels*], nos quais “os elementos do grupo conservam sua autonomia, deixando ver uma afinidade evidente que os aproxima, de forma que o conjunto apresenta contornos convencionados e dá a impressão de um ‘déjà vu’”¹² (BALLY, 1951, p. 70). Destaca, assim, a coesão apenas relativa entre os termos, os quais conservam sua autonomia¹³, e sua localização na norma¹⁴, como combinações pré-fabricadas¹⁵ e convencionais. Alonso Ramos (1994-95, p. 15), porém, ao centrar-se na denominação proposta por Bally (1951) de “agrupamentos usuais”, associa essas combinações a estruturas de uso freqüente. Dessa forma, relaciona-as com as combinações livres e descarta do estudo das colocações. Acreditamos, porém, que o adjetivo “usual” reflete a freqüência não de uso, mas de coaparição dos elementos, ou seja, indica que os membros da combinação costumam aparecer associados, dando, por isso, a impressão de um “déjà vu”.

Ainda para Bally (1951), dentro desse grupo intermediário de combinações, são incluídas as “séries verbais” [*séries verbales*] e as “séries de intensidade” [*séries d'intensité*]. As primeiras correspondem a expressões que parafraseiam os verbos, como *prendre une décision* [tomar uma decisão] e *avoir peur* [ter medo]. Já as “séries de intensidade” dizem respeito a combinações em que um substantivo, um adjetivo ou um verbo aparecem unidos a uma palavra cuja função é reforçar seu significado fundamental, sem acrescentar uma determinação nova, nem restringir ou precisar suas noções primeiras¹⁶. Como exemplos,

¹¹ Entendemos significado como “o conteúdo de um signo ou de uma expressão enquanto dado numa determinada língua e exclusivamente através dessa mesma língua” (COSERIU, 1980, p. 99) em oposição a sentido, que corresponde ao “conteúdo próprio de um texto, o que o texto exprime além e através da designação e do significado” (COSERIU, 1980, p. 99). Para outras noções de significado, v. Dubois et al (1999, s.v. *signifié*) e Crystal (1997, p. 100-102).

¹² [les éléments du groupe conservent leur autonomie, tout en laissant voir une affinité évidente qui les rapproche, de sorte que l'ensemble présente des contours arrêtés et donne l'impression du “déjà vu”].

¹³ Opinião compartilhada por Zuluaga (1998, p. 102), para quem “cada termo conserva o seu valor e a sua definição própria” [cada termo conserva o seu valor e a sua definición propia].

¹⁴ Entendemos “norma” como uma estrutura da língua que “contém tudo o que, no falar correspondente a uma língua funcional, é fato tradicional, comum e constante, ainda que não necessariamente funcional” (COSERIU, 1980, p. 122).

¹⁵ Utilizamos a denominação “pré-fabricadas”, neste trabalho, para caracterizar as combinações que, por serem um fato de norma, apresentam uma estrutura relativamente restrita que costuma ser recuperada freqüentemente pelos falantes. Destacamos, porém, que, por “pré-fabricadas”, não estamos nos referindo às combinações fixas, as quais também aparecem na norma e apresentam restrições entre os seus membros.

¹⁶ Essa mesma noção aparecerá em Hausmann (1989, p. 1010), quando afirma que o colocado “acrescenta uma caracterização que não modifica a identidade do caracterizado” [ajoute une caractérisation qui ne modifie pas l'identité du caractérisé].

temos *chaleur suffocante* [calor sufocante], *diamétralement opposé* [diametralmente oposto] e *refuser catégoriquement* [refutar categoricamente]. Percebe-se, dessa forma, que as combinações léxicas podem ocorrer entre diferentes classes de palavras e, além disso, que os adjetivos e advérbios, acrescentados aos substantivos e verbos, cumprem apenas uma função de intensificação, isto é, são responsáveis somente por delimitar em que grau se realiza o expressado pelo verbo ou substantivo.

Ainda que não tenha tratado diretamente das colocações, Bally (1951) é o primeiro a perceber que há um agrupamento de palavras situado entre as combinações livres e as fixas. Além disso, consideramos pertinente destacar seu estudo, uma vez que foi um dos poucos autores a se preocupar com a elaboração de critérios para a identificação das unidades fraseológicas. Através de sua proposta de que sejam utilizados índices exteriores e interiores para o reconhecimento dessas unidades - os primeiros relacionados à forma das combinações e os segundos, à maneira como os falantes as entendem -, Bally (1951) apresenta alguns parâmetros que posteriormente serão resgatados para a identificação das colocações.

2.1.1.2 Porzig (1970 [1950])

Porzig (1970 [1950]), assim como Bally (1951), situa as combinações léxicas no eixo sintagmático. Esse autor propõe que as palavras de uma língua encontram-se estruturadas em “campos verbais” [*campo verbal*] ou “significativos” [*significativo*], nos quais cada palavra é definida por sua relação com as demais. Os campos verbais em que os vocábulos podem ser agrupados contígua e conseqüentemente são chamados de “paratáticos” [*paratáticos*]¹⁷. Como exemplos, temos o campo das estações do ano, das cores ou, dentro do campo “paisagem”, as relações entre *monte*, *vale*, *encosta*, *montanha*, *colina*, etc. Dessa forma, percebemos que tais relações situam-se no eixo paradigmático. Por outro lado, há campos verbais em que a relação entre os vocábulos opera no eixo sintagmático, ou seja, as palavras selecionam apenas algumas específicas para se combinar, constituindo um campo verbal “sintático” [*sintático*].

Esse segundo campo traz a essência das relações estabelecidas entre os vocábulos que conformam uma colocação, pois, segundo Porzig (1970, p. 133), “todas as palavras têm

¹⁷ O conceito de “campos verbais paratáticos” introduzido pelo autor não apresenta relação com o conceito de “parataxe”, entendida como a justaposição de frases (cf. DUBOIS et al, 1999, s.v. *parataxe*).

um campo magnético ao seu redor, no qual podem penetrar somente palavras de um tipo muito específico”¹⁸. E acrescenta que “o círculo dessas possibilidades pode ser menor ou maior”¹⁹ (PORZIG, 1970, p. 133). Assim, enfatiza a restrição léxica entre os elementos, mas indica que esta pode ser variável de acordo com os vocábulos envolvidos. Como exemplos de restrição total, temos as relações entre o verbo *fällen* [cortar] e o substantivo *Baum* [árvore], no alemão, assim como entre *blond* [louro] e *Haar* [cabelo]. No entanto, consideramos importante ressaltar que essa seleção restrita dos vocábulos é apenas unilateral, ou seja, *louro* pode ser aplicado exclusivamente a *cabelo*, mas este pode selecionar um número maior de vocábulos para se combinar. O exemplo entre *fällen* [cortar] e *Baum* [árvore], por sua vez, não funciona para a língua portuguesa como combinação restrita, o que reafirma a idéia de Coseriu (1980, p. 99) de que o significado representa o conteúdo de um signo apenas dentro de uma determinada língua²⁰. Quanto à restrição variável, Porzig (1970, p. 132) traz como exemplo a relação entre *Schnee* [neve] e os adjetivos *weiß* [branca] e *schmutzig* [suja]. Deve-se observar, porém, que a possibilidade de se combinar diferentes vocábulos com *neve* é condicionada pela necessidade de qualificar de diferentes formas esse substantivo, exprimindo diferentes significados. Essa mesma relação poderia ocorrer com *cabelo*, utilizado como exemplo de restrição total, ao ser qualificado por diferentes adjetivos, tais como *liso*, *crespo*, *ruivo*, *preso*, etc²¹.

Corpas Pastor (1996, p. 55) destaca que Porzig (1970), ao tratar dos “campos verbais”, mistura dois fenômenos diferentes sob a mesma denominação. Por um lado, encontram-se exemplos diretamente relacionados com o problema das colocações. Porém, por outro lado, são apresentadas combinações entre unidades como *beißen* [morder] e *Zähne* [dentes], “que não costumam coaparecer no discurso e cuja relação semântica é de inclusão”²². Apesar de haver certa heterogeneidade nas combinações léxicas analisadas por Porzig (1970), não podemos negar sua contribuição para o estudo das colocações. As noções introduzidas pelo autor, ainda que não explicitamente, já propõem uma taxonomia para as combinações léxicas, pois são abordadas as relações entre verbo + substantivo (tanto sujeito quanto objeto), substantivo + adjetivo e verbo + advérbio. Além disso, ao afirmar que “se uma

¹⁸ [tienen todas las palabras un campo magnético en torno a ellas, donde sólo pueden penetrar palabras de muy determinada clase].

¹⁹ [el círculo de estas posibilidades puede ser menor o mayor].

²⁰ Em virtude desse fato, serão mantidos, no presente trabalho, os exemplos em suas línguas originais acompanhados de tradução, ainda que, muitas vezes, a tradução não consiga refletir o fenômeno exemplificado no original.

²¹ Opinião semelhante em Aisenstadt (1981).

²² [que no suelen coaparecer en el discurso, y cuya relación semántica es de inclusión].

vez, em uma situação, foi selecionada ou imposta uma palavra, então o falante já não é completamente livre na formação do discurso, mas se encontra limitado a certas possibilidades na escolha das outras palavras”²³ (PORZIG, 1970, p. 134-135), situa tais combinações léxicas no nível da língua²⁴ ou, conforme a terminologia Coseriana, na norma.

2.1.1.3 Coseriu (1977 [1967])

Coseriu (1977 [1967]), em seus estudos sobre as estruturas lexemáticas, ou seja, as estruturas do significado léxico, distinguiu entre as relações paradigmáticas e sintagmáticas e, no segundo eixo, situou as “solidariedades léxicas” [*solidariedades léxicas*]. Esse conceito, desenvolvido a partir dos estudos de Porzig (1970), refere-se à “determinação semântica de uma palavra por meio de uma classe, um arquilexema ou um lexema, no sentido de que uma classe determinada, um determinado arquilexema ou um determinado lexema funciona como traço distintivo da palavra considerada”²⁵ (COSERIU, 1977, p. 148), ou seja, uma classe²⁶, um arquilexema²⁷ ou um lexema²⁸ fazem parte do conteúdo sêmico da palavra em questão como mais um traço distintivo²⁹. Quando a relação é estabelecida por uma classe, como entre *miles* [soldado] e *senex* [velho (aplicado somente a pessoa)], ocorre uma solidariedade léxica por “afinidade” [*afinidad*]. No caso de relação com um arquilexema, como a que ocorre entre *fahren* [deslocar-se em um veículo] e o arquilexema de *Schiff* [barco], *Zug* [trem], *Wagen* [carro], etc., a solidariedade léxica é por “seleção” [*selección*]. Finalmente, quando a relação

²³ [si una vez en una situación ha sido elegida o se ha impuesto una palabra, entonces el hablante no es ya plenamente libre en la formación del discurso, sino que se encuentra limitado a ciertas posibilidades en la elección de las otras palabras].

²⁴ Essa delimitação já aparece em Saussure (2004, p. 145): “cumpre atribuir à língua e não à fala todos os tipos de sintagmas construídos sobre formas regulares”.

²⁵ [determinación semántica de una palabra por medio de una clase, un archilexema o un lexema, en el sentido de que una clase determinada o un determinado archilexema o un determinado lexema funcionan como rasgo distintivo de la palabra considerada].

²⁶ Uma classe representa um conjunto de elementos que possuem pelo menos um traço em comum (cf. GLÜCK, 2000, s.v. *Klasse*).

²⁷ Um arquilexema corresponde a um lexema cujo conteúdo é idêntico ao significado total de um campo léxico (cf. BUSSMANN, 1990, s.v. *Archilexem*).

²⁸ Um lexema corresponde a uma unidade abstrata do léxico no nível da *langue* que pode ser realizada gramaticalmente por meio de diferentes palavras (cf. BUSSMANN, 1990, s.v. *Lexem*). Segundo Trask (2006, s.v. *palavra*), o termo “lexema” é sinônimo de “item lexical”, denominação utilizada nos estudos com uma abordagem estatística das colocações (v. 2.1.2). Para outras noções desse conceito, v. Galisson, Coste (1976, s.v. *lexème*).

²⁹ Os traços distintivos, para Coseriu (1977, p. 148), correspondem a diferenças semânticas mínimas entre os lexemas. Para outras noções desse conceito, v. Dubois et al (1999, s.v. *trait*).

acontece com um lexema, como *bayo* [baio], que se aplicaria somente a *caballo* [cavalo], existe uma solidariedade léxica por “implicação” [*implicación*]. É importante salientar, porém, que, nesse último exemplo, *cavalo* não funciona somente como um lexema, mas como um arquilexema, que incluiria também *égua*, *potro*, entre outros (SALVADOR, 1989-90, p. 344).

Para Coseriu (1977, p. 151-152), em toda solidariedade léxica existe um lexema determinante, aquele cujo traço distintivo faz parte do outro lexema envolvido na solidariedade, e um lexema determinado, que constitui o receptor desse traço distintivo. Dessa forma, a relação entre os elementos na solidariedade léxica é orientada, isto é, o significado do lexema determinante implica o determinado, e não o contrário³⁰ (cf. CASTILLO CARBALLO, 1998, p. 48; CORPAS PASTOR, 1996, p. 64).

Coseriu (1977), aprofundando os trabalhos de Porzig (1970), propõe, ainda, que suas combinações sejam divididas em “solidariedades multilaterais” [*solidariedades multilaterales*] e “solidariedades unilaterais” [*solidariedades unilaterales*]. Nas primeiras ocorre uma determinação externa, pois o traço distintivo do lexema determinante é acrescentado ao conteúdo do lexema determinado, possibilitando uma oposição paradigmática, como ocorre entre *cavalorelinchar*, *cachorrollatir*, *leãolrugir*, etc., em que ao conteúdo “emitir um som [um animal]” são acrescentados diferentes animais como traços distintivos. Já nas “solidariedades unilaterais”, a determinação é interna, pois o traço distintivo do lexema determinante está incluído no determinado e não possui valor opositivo no eixo paradigmático, como ocorre entre *morder* e *dentes*. Para Salvador (1989-90, p. 341), Coseriu (1977) “superestima a suposta aproximação de Porzig ao conceito de solidariedade, até o extremo de aceitar que exista solidariedade em alguns pares indicados por esse autor, nos quais, a meu ver, a relação é meramente substancial e não formal”³¹, ou seja, existe uma relação entre os vocábulos no que se refere a seu conteúdo, mas esses vocábulos não costumam aparecer unidos no discurso³². Dessa forma, a maior parte dos exemplos apresentados por Coseriu (1977) para estabelecer a distinção entre solidariedades unilaterais e multilaterais é classificada como “solidariedades referenciais” [*solidariedades referenciales*]

³⁰ Essa mesma noção aparecerá em Hausmann (1989), ao estabelecer uma hierarquia entre a base e o colocado de uma combinação, na qual a primeira seleciona o segundo.

³¹ [sobrestima la supuesta aproximación de Porzig al concepto de solidaridad, hasta el extremo de aceptar que la hay en algunas de las parejas indicadas por éste, en las que, a mi juicio, la relación es meramente sustancial y no formal].

³² Entendemos “discurso” como “qualquer fragmento conexo de escrita ou fala” (TRASK, 2006, s.v. *discurso*) e, mais especificamente, como a língua em uso pelos falantes (DUBOIS et al, 1999, s.v. *discours*).

por Salvador (1989-90, p. 342), pois não constituem relações estabelecidas pela língua, mas sim restrições condicionadas pela realidade extralingüística.

Corpas Pastor (1996, p. 65) acredita que a noção de “solidariedade multilateral” se corresponderia com a de “colocação”; no entanto, para a autora, a noção de colocação seria mais ampla que a de solidariedade, pois esta não poderia explicar combinações como *diametralmente oposto*, na qual nenhum elemento implica o outro pelo compartilhamento de traços distintivos, ocorrendo apenas uma relação de intensificação. Acreditamos, porém, que nos casos de solidariedade em que a relação é determinada por uma classe, essas combinações tornam-se mais abrangentes que as colocações³³. Ainda segundo Copas Pastor (1996, p. 65), as colocações diferem das solidariedades no aspecto da frequência, pois, para Coseriu (1977, p. 160), “a probabilidade estatística geral das combinações não tem praticamente nada a ver com as solidariedades e não é prova de sua existência”³⁴. No entanto, segundo Alonso Ramos (1994-95, p. 23), “a frequência não tem nenhuma validade para considerar uma determinada combinação como colocação”³⁵. Sendo assim, tanto nas colocações como nas solidariedades léxicas, a frequência pode ser vista como uma consequência do uso das unidades em conjunto, mas não como sua causa.

A partir disso, percebemos que, embora os conceitos de “solidariedade léxica” e de “colocação” não estejam diretamente relacionados, muitas contribuições de Coseriu (1977) no estudo das solidariedades podem ser aplicadas às colocações. Através dos traços distintivos, esse autor propõe uma primeira aproximação à tentativa de estabelecer as relações semânticas mantidas entre os elementos que formam essas estruturas. Somente a partir dessas relações é possível a identificação de uma hierarquia entre os membros da colocação. Finalmente, Coseriu (1977) já apresenta uma classificação dos diferentes tipos de combinações ao distinguir entre as solidariedades léxicas por afinidade, seleção ou implicação, o que, segundo Pazos Breña (2005, p. 14), já contém a idéia de policolocabilidade e monocolocabilidade.

³³ Estes casos de solidariedade léxica por afinidade aproximam-se da noção de “contorno” [*contorno*], introduzida por Seco (1987), a qual trataremos em 2.2.4.

³⁴ [la probabilidad estadística general de las combinaciones no tiene prácticamente nada que ver con las solidariedades y no es prueba de su existencia].

³⁵ [la frecuencia no tienen [sic] ninguna validez para considerar una determinada combinación como colocación].

2.1.1.4 Mel'čuk (1995, 2003)³⁶

Mel'čuk (1995) estabelece uma distinção entre as combinações em que as unidades lexicais podem ser combinadas livremente e aquelas em que não há tal liberdade de escolha. As segundas, denominadas “frasemas semânticos” [*semantic phraseme*], correspondem a “um sintagma composto de duas UL [*i.e.* unidades lexicais] A e B, onde o significante não é construído livremente e o significado não é construído regularmente”³⁷ (ALONSO RAMOS, 2001, p. 8). Isso significa que existe uma “atração” entre os elementos, não podendo haver uma escolha arbitrária no léxico a ser combinado, além de essas construções não apresentarem necessariamente uma regularidade com relação a sua estrutura gramatical.

Os frasemas semânticos são divididos em três tipos que, segundo Wanner (1996, p. 14), variam de acordo com o grau de opacidade semântica:

- a) frasema completo [*full phraseme*] ou expressão idiomática [*idiom*], em que seu significado não inclui o dos elementos que o compõem, como em *to shoot the breeze* [jogar conversa fora]. Essa forma corresponde às combinações fixas para os demais autores.
- b) semi-frasema [*semi-phraseme*] ou colocação [*collocation*], em que seu significado inclui o de um de seus membros - A - e um significado novo - C -, de forma que o elemento B exprime C somente em função de A. Nesse caso, as diferentes relações estabelecidas entre os elementos das combinações, que, segundo Nesselhauf (2005, p.16), correspondem aos critérios da transparência e da comutabilidade, dão origem a quatro subtipos de colocações:

³⁶ Heid (1992, p. 524) ressalta que os trabalhos de Mel'čuk, apesar de apresentarem uma orientação lexicológica e lexicográfica, possuem, sobretudo, um enfoque semântico. Dessa forma, ainda que este autor tenha retratado o problema das colocações em obras lexicográficas, optamos por estudá-lo a partir do enfoque lexicológico, uma vez que o principal objetivo de Mel'čuk não foi o tratamento dessas combinações nos dicionários, mas sua sistematização dentro de uma teoria mais ampla do léxico - o modelo Sentido-Texto [*Sens-Texte*] - por ele desenvolvida. O modelo Sentido-Texto, segundo Mel'čuk (1978, p. 273), corresponde a um “sistema de regras formais que asseguram a tradução de todos os sentidos dados em todos os textos correspondentes e vice e versa” [*systeme de règles formelles assurant la traduction de tout sens donné dans tous les textes correspondants et vice versa*]. Enfatiza, assim, a correspondência entre um conjunto infinito de sentidos e um conjunto finito de textos (cf. MEL'ČUK, 1995, p. 171); os primeiros são entendidos como a informação comunicada em um enunciado lingüístico e os segundos correspondem à representação simbólica dessas informações. A concepção dessa nova teoria explica o uso de diferentes denominações, por parte do autor, para o tratamento da combinatória léxica.

³⁷ [un syntagme composé de deux UL A et B dont le signifiant n'est pas construit librement et dont le signifié n'est pas construit régulièrement].

- C ≠ B, ou seja, o significado correspondente de B não aparece no dicionário, mas:
 1. pode ser vago, isto é, a unidade léxica B é utilizada como um “auxiliar” para dar suporte a determinada combinação sintática (como nos casos de verbo-suporte³⁸). Exemplos desse caso seriam *to do a favor* [fazer um favor] e *to give a look* [dar uma olhada];
 2. não é vago, mas B exprime C somente em combinação com A ou com um número limitado de lexemas semelhantes, correspondendo a colocações restritas dentro de um campo semântico específico, como em *black coffee* [café preto];
- C = B, ou seja, o significado correspondente de B aparece no dicionário e:
 1. B não pode ser substituído por um sinônimo possível, o que corresponde a colocações restritas por serem consagradas pelo uso, como em *strong (*powerful) coffee* [café forte];
 2. B inclui o significado de A, exprimindo, portanto, um significado específico, aproximando-se da noção de solidariedade léxica por implicação, conforme ocorre nos exemplos *aquiline nose* [nariz aquilino] e *the horse neighs* [o cavalo relincha]³⁹.

c) quase-frasema [*quasi-phrase*], em que seu significado inclui o de suas unidades léxicas constituintes acrescido de um ajuste semântico imprevisível, o que ressalta a idiosincrasia entre as línguas. Como exemplo, temos *to start a family* [constituir uma família], em que são incluídos os significados dos dois elementos além do implícito de “ter o primeiro filho com um(a) esposo(a)”.

No grupo das colocações, posteriormente, Mel’čuk (2003) substitui as unidades lexicais A e B pelas expressões “base” e “colocado”, conceitos introduzidos por Hausmann (1979), que serão tratados em 2.2.5. Isso implica no fato de a base - elemento autônomo - ser

³⁸ Os verbos-suporte correspondem, segundo Alonso Ramos (1997), a verbos que se unem aos substantivos para formar orações sem, no entanto, acrescentar um significado novo. São considerados verbos “semanticamente vazios”, não no sentido de que estejam desprovidos de significado, mas de que se “esvaziam” (e não necessariamente de maneira completa) ao se combinarem com determinados substantivos, os quais incluem o significado do verbo e determinam sua seleção. Dubsky (1998 [1984]), de forma pioneira, utiliza a denominação “verbo de apoio” [*verbo de apoyo*] para a língua espanhola. Para o francês, a revista LANGAGES (1996) dedica seu número 121 somente aos suportes [*les supports*]. Um estudo específico para a língua portuguesa, que trata exclusivamente dos casos de colocação com verbo-suporte, encontra-se em Neves (1999). Essa autora propõe uma série de testes para identificar os casos em que esse tipo de combinação léxica constitui verdadeiras colocações, diferindo-se das combinações fixas e das livres.

³⁹ Este exemplo remonta à discussão de se a combinação realmente se dá por uma afinidade entre os vocábulos ou se sua união é motivada por fatores extralingüísticos.

selecionada livremente, mas condicionar léxica e semanticamente a escolha do colocado (cf. ALONSO RAMOS, 2001, p. 11).

O significado originado pela combinação entre a base e o colocado corresponde à noção de “função léxica” (v. nota 8), as quais devem ser suficientemente abstratas para poderem ser combinadas com muitos argumentos (cf. ALONSO RAMOS, 1989, p. 327). Por representarem, assim, a partir de um conceito genérico, a relação estabelecida entre a base e o colocado, as funções léxicas apresentam a vantagem de oferecer uma perspectiva ontológica da linguagem, podendo ser aplicadas a diferentes línguas particulares. Em virtude desse fato, as funções léxicas propostas por Mel’čuk (1995)⁴⁰, utilizadas para a identificação de relações sintagmáticas - como a intensificação ou os verbos-suporte - e paradigmáticas - como a sinonímia, antonímia ou derivados - entre os vocábulos, foram amplamente empregadas para a identificação de colocações, em obras como o DICE (2004)⁴¹, o DICO (*Dictionnaire de Combinatoire*), o Dafles (*Dictionnaire d’apprentissage du français des affaires*) e o *Lexique de cooccurrents: Bourse - conjoncture économique*, entre outros⁴². No entanto, considerando que esta não constitui a finalidade principal das funções léxicas, muitas das combinações identificadas pelos autores correspondem a estruturas livres (cf. BENSON, 1985, p. 62), como acontece no DICE (2004), organizado por Alonso Ramos. Para o substantivo *admiración* [admiração], por exemplo, o dicionário inclui combinações com os adjetivos *gran* [grande] e *unánime* [unânime], vocábulos que podem coocorrer com um número muito grande de substantivos para expressar intensidade e que algo é compartilhado por todos, respectivamente. Alonso Ramos (1994-95, p. 23-24) argumenta que a inclusão dessas combinações justifica-se pelo fato de evitar que sua ausência fosse interpretada como inaceitabilidade da construção. Porém, acreditamos que, se o dicionário é elaborado para oferecer ao usuário colocações - como seu próprio nome indica -, não há como justificar a presença de combinações livres.

Além disso, devido ao número limitado de funções léxicas (aproximadamente 60) para indicar as relações entre os vocábulos, muitas combinações não podem ser categorizadas. Assim, segundo Mel’čuk (2003, p. 31), um conjunto muito grande de colocações não pode ser previsto teoricamente e deve ser recolhido de forma empírica. Isso faz com que, além de intuitivas (cf. HAUSMANN, 1998a, p. 72; PENADÉS MARTÍNEZ, 2001, p. 66), as funções

⁴⁰ Uma relação de todas as funções léxicas pode ser encontrada no anexo A.

⁴¹ Seguindo a tendência da metalexigrafia atual, as obras dicionarísticas aparecem expressas com siglas. Na bibliografia, é feita a referência completa (cf. HARTMANN, 2001, p. 11).

⁴² Infelizmente não tivemos acesso aos originais dessas três últimas obras, mas apenas a trabalhos que fazem referência a elas, como Polguère (2005), Verlinde, Selva, Binon (2003) e Heid (1996).

léxicas sejam extremamente específicas, o que exige a utilização de um grande número de novas funções para cada situação particular. Em virtude desse fato, um grupo muito grande de combinações é limitado a categorias tão pequenas que não se pode realizar uma generalização para sua análise.

Finalmente, consideramos importante destacar que, apesar de realizar um estudo eminentemente semântico da combinatória léxica, Mel'čuk (2003, p. 26) enfatiza o aspecto da frequência de uso das colocações, ao indicar que a importância dessas combinações reside em sua onipresença. A partir da análise de 40 linhas de um texto, o autor identifica treze colocações diferentes. No entanto, entre estas, inclui combinações como *au début* [no início] e *[se mesure] à deux critères* [medir-se por dois critérios], que correspondem a exemplos de coligação⁴³ e combinação livre, respectivamente. Além disso, embora muitos autores - tais como Smadja (1993, p. 147), Bahns (1993, p. 57), entre outros - utilizem a noção de frequência de uso para tratar das colocações, essa parece não ser uma característica fundamental para a definição desse fenômeno léxico. Koike (2000, p. 84), ao estudar as combinações entre verbos e substantivos, identifica que

nem sempre os verbos com maior colocabilidade com os substantivos apresentam um alto índice de frequência de uso. A colocabilidade do verbo é determinada por seu significado léxico, seja específico ou geral, e por sua condição sintática, e não por sua frequência no texto, enquanto a frequência do verbo depende muito mais de seu maior ou menor valor funcional⁴⁴.

Essa opinião é também compartilhada por Alonso Ramos (1994-95, p. 15).

2.1.1.5 *Corpas Pastor (1996, 2001)*

Na tentativa de delimitar o escopo do fenômeno da combinatória léxica, Corpas Pastor (1996) propõe tratar todas as combinações dentro do estudo de fraseologia, entendida como “a nomenclatura utilizada para se referir ao estudo das combinações de palavras”⁴⁵ (CORPAS PASTOR, 1996, p. 16). A partir disso, a autora apresenta uma sistematização dos

⁴³ O termo “coligação” é utilizado por Tagnin (2005a, p. 30) para indicar uma “combinação de elementos lingüísticos em que o colocado é uma palavra gramatical”. Corresponde, na terminologia de Benson (1985), às “colocações gramaticais” [*grammatical collocations*].

⁴⁴ [no siempre los verbos con mayor colocabilidad con los sustantivos presentan un alto índice de frecuencia de uso. La colocabilidad del verbo va determinada por su significado léxico, sea específico o general, y por su condición sintáctica, y no por su frecuencia en el texto, mientras que la frecuencia del verbo depende más bien de su mayor o menor valor funcional].

⁴⁵ [la nomenclatura utilizada para referirse al estudio de las combinaciones de palabras].

diversos aspectos que caracterizam as combinações léxicas, os quais, em *Corpas Pastor* (2001), aplica diretamente às colocações, justificando a inclusão destas no estudo das fraseologias. Propõe, então, como características da combinatória léxica:

- a) frequência - que compreende tanto a frequência de coaparição dos elementos como a de uso da expressão. Para *Corpas Pastor* (2001, p. 43), essa é a base da corrente estatística; no entanto, acreditamos que essa abordagem (v. 2.1.2) trabalhe apenas com a noção de coocorrência freqüente, sem uma referência ao uso das combinações. Conforme explicitado na seção anterior (v. 2.1.1.4), a frequência de uso é questionada por alguns autores para caracterizar o fenômeno léxico das colocações.
- b) institucionalização - em decorrência da frequência de uso, as combinações passam a ser percebidas pelo falante como um fragmento pré-fabricado, ou seja, como uma estrutura reconhecida e pré-existente. Essa concepção será utilizada por Sinclair (1991, p. 110) para caracterizar as combinações pertencentes ao “princípio idiomático” [*the idiom principle*] da língua (v. 2.1.2.3). Ao serem entendidas como um fragmento pré-fabricado, as colocações devem ser inseridas na norma, pois, como não correspondem a construções criadas livremente pelos falantes, não pertencem ao âmbito da fala.
- c) estabilidade - que implica na restrição combinatória e na especialização semântica. A primeira pode ser variável, incluindo casos como *paliza* [surra], que se combina com *pegar* [dar] e um reduzido número de sinônimos parciais em espanhol (*dar, meter, propinar, arrear*), e como *izar* [hastear], que se combina somente com *bandera* [bandeira]. Já a especialização semântica corresponde à supressão ou à adição de significado a um dos elementos. No primeiro caso, as combinações estabelecem-se com um verbo deslexicalizado (ou suporte), como *prestar* [prestar] nas combinações com *atención* [atenção], *auxilio* [auxílio], *juramento* [juramento], entre outras, em que o verbo perde sua acepção original de “dar [algo] a [alguém] com a idéia de que lhe seja devolvido”, que corresponde a *emprestar* em português, adquirindo um significado “geral e gramaticalizado, funcional e auxiliar”⁴⁶ (*CORPAS PASTOR*, 2001, p. 45). Por outro lado, nos casos de adição de significado,

⁴⁶ [general y gramaticalizado, funcional y auxiliar].

ocorre uma “evolução semântica de clara base metafórica”⁴⁷ (CORPAS PASTOR, 2001, p. 45), como em *levantar* [suspender] ao se combinar com *castigo* [castigo], *sanción* [sanção], *prohibición* [proibição], cujo significado implica em “tirar um peso de cima” ou *desarmar* [desarmar] combinado com *teoría* [teoria], que subentende a metáfora de que “discutir é lutar”. Consideramos, porém, que, apesar de sua origem metafórica, essas combinações mantêm um significado que pode ser deduzido pela soma de seus elementos constitutivos.

- d) idiomaticidade - entendida como o resultado da especialização semântica em seu grau mais elevado, corresponde à propriedade semântica das combinações em que o significado global não é obtido pela soma de suas partes, sendo aplicada somente às expressões idiomáticas. No entanto, segundo Corpas Pastor (1996, p. 27), além desse “significado denotativo figurado” [*significado denotativo figurativo o traslaticio*], pode haver um “significado denotativo literal” [*significado denotativo literal*]. Como exemplo, Corpas Pastor (2001, p. 46) apresenta a combinação *meter un gol* [fazer um gol] que pode significar tanto “colocar a bola no gol (em uma partida de futebol)” como “enganar”. Para a autora, no primeiro caso, que apresenta um “significado denotativo literal”, estamos diante de uma colocação; já no segundo, quando ocorre um “significado denotativo figurado”, o que existe é uma locução. Dessa forma, Corpas Pastor (2001) enfatiza o caráter transparente das colocações.
- e) variação - que corresponde tanto a variantes diatópicas, diafásicas e diastráticas⁴⁸ como a modificações particulares do falante para obter determinado efeito estilístico. Como exemplo de variantes diatópicas, Corpas Pastor (2001, p. 46), a partir de Koike (2000), apresenta as colocações *dar una opinión*, de uso na Espanha, e *entregar una opinión*, variedade utilizada no Chile, ambas com o significado de “dar uma opinião”⁴⁹. Ainda que tal

⁴⁷ [evolución semántica de clara base metafórica].

⁴⁸ Segundo Coseriu (1980, p. 110), “uma língua histórica apresenta sempre variedade interna” e essa variedade é representada pelas diferenças: a) diatópicas, isto é, de espaço geográfico; b) diastráticas, isto é, entre os diferentes estratos sócio-culturais da comunidade lingüística e c) diafásicas, ou seja, entre os diversos modos de expressão de acordo com a situação comunicativa.

⁴⁹ Em uma consulta realizada através da ferramenta de buscas de sites da internet Google (www.google.com.br), identificamos que a combinação *entregar una opinión* aparece 1.520 vezes para sites provenientes do Chile; porém, as páginas desse país também registram 18.400 vezes a combinação *dar una opinión*. É importante destacar que, com esses números, não queremos indicar que o caráter colocacional dos elementos seja

divergência diatópica seja questionável, os exemplos evidenciam a preferência da base por dois colocados diferentes, conforme o país em que se utilize a combinação⁵⁰. Esse aspecto da variação das colocações, ao ser contrastado com o da estabilidade (c), demonstra que a fixação das combinações léxicas é apenas relativa.

- f) gradação - identifica que as combinações léxicas seguem uma escala na apresentação das características anteriores, formando um *continuum* no qual existem pontos sobressalentes. Essa noção de continuidade entre os diferentes tipos de combinações será também enfatizada por Cowie (1981) (v. 2.2.1).

Após a identificação das características das unidades fraseológicas como um todo, a autora define as colocações como

unidades fraseológicas formadas por duas unidades léxicas em relação sintática, que não constituem, por si só, atos de fala nem enunciados, e que, devido a sua fixação na norma, apresentam restrições de combinação estabelecidas pelo uso, geralmente de base semântica: o elemento autônomo semanticamente (a base) não apenas determina a escolha do colocado, mas, além disso, seleciona nele uma acepção especial, freqüentemente de caráter abstrato ou figurado⁵¹ (CORPAS PASTOR, 1996, p. 66).

Corpas Pastor (1996, 2001), por representar uma visão mais atual nos estudos das colocações, conjuga aspectos destacados pelos diferentes autores que trataram desse tema. Assim, ressalta a importância da frequência e da restrição combinatória variável, além de reconhecer uma hierarquia entre os elementos que compõem a colocação. Para a autora, a frequência inclui tanto a de coaparição dos elementos quanto a de uso das combinações. Reiteramos, porém, que somente a primeira demonstra ser relevante para o estudo das colocações, pois nos casos de combinações como *levantar una sanción* [suspender uma sanção], apresentada pela própria autora, cuja ocorrência é restrita a contextos específicos, a frequência de uso não será significativa. Dessa forma, tal fator não pode ser considerado determinante para a identificação das colocações.

determinado pela frequência, mas sim que a combinação *dar una opinión* também é uma colocação significativa no Chile.

⁵⁰ Peters (1992), ao analisar as diferenças diatópicas das colocações, apresenta como exemplo a combinação *producto sureño*, utilizada na Espanha e no Chile, em oposição a *producto surero*, de uso na Argentina e na Bolívia, ambas significando “produto sulino”. Além do fato de a segunda combinação não apresentar nenhuma ocorrência em consultas no *Google*, consideramos que esses exemplos não ilustram uma variação diatópica das colocações, uma vez que o substantivo *producto* não seleciona adjetivos diferentes de acordo com o país em que a combinação é utilizada. A variação acontece porque o vocábulo *sureño* possui uma forma variante, independente da combinação em que é utilizado.

⁵¹ [unidades fraseológicas formadas por dos unidades léxicas en relación sintáctica, que no constituyen, por sí mismas, actos de habla ni enunciados; y que, debido a su fijación en la norma, presentan restricciones de combinación establecidas por el uso, generalmente de base semántica: el colocado autónomo semánticamente (la base) no sólo determina la elección del colocativo, sino que, además, selecciona en éste una acepción especial, frecuentemente de carácter abstracto o figurativo].

As características propostas por Corpas Pastor (1996, 2001) são importantes para o reconhecimento das fraseologias e, mais especificamente, das colocações; no entanto, não podem ser vistas como princípios de validade universal, pois não funcionam sempre para todas as combinações. Assim, na construção *izar una bandera* [hastear uma bandeira], por exemplo, proposta pela autora para exemplificar a restrição combinatória, não podemos dizer que exista uma especialização semântica do colocado, pois o verbo mantém sua acepção original, sem acréscimo ou supressão de significado.

2.1.1.6 Síntese da abordagem semântica

A partir dos estudos desenvolvidos sob uma abordagem semântica, percebemos a coincidência dos autores em situar as combinações léxicas no eixo das relações sintagmáticas. Nesse eixo, existe uma restrição variável entre os vocábulos que podem se combinar para formar uma colocação. Esses vocábulos, ao serem fixados na norma (cf. BALLY, 1951; PORZIG, 1970; CORPAS PASTOR, 1996), constituem construções pré-fabricadas a disposição dos falantes. Por isso, correspondem a combinações de uso freqüente para Mel'čuk (2003) e Corpas Pastor (1996), mas não necessariamente para Coseriu (1977).

No que diz respeito aos elementos que compõem cada combinação, Bally (1951) ressalta sua autonomia semântica. Para Mel'čuk (1995) e Corpas Pastor (1996), porém, somente a base apresenta autonomia, idéia inspirada em Coseriu (1977), para quem os traços de um lexema determinam a coocorrência de outro. Dessa forma, a restrição na seleção léxica é determinada pelas relações semânticas estabelecidas entre os elementos da combinação. Segundo Wanner (1996, p. 16), nos estudos de Porzig (1970) e Coseriu (1977) o foco ainda está centrado no significado dos lexemas e não em sua relação colocacional. Acreditamos, porém, que, na medida em que essa relação auxilia na conformação do significado dos elementos da combinação, pois, segundo Corpas Pastor (1996, p. 66), a base seleciona uma “acepção especial” do colocado, as noções desenvolvidas por Porzig (1970) e Coseriu (1977) representam contribuições importantes para o estudo das colocações.

Esses autores, assim como Bally (1951), partindo do significado dos vocábulos que participavam das combinações, passaram a analisar as relações estabelecidas entre os elementos. Tais relações, classificadas como “séries verbais” e “séries de intensidade” por Bally (1951) ou solidariedades léxicas por “afinidade”, “seleção” e “implicação” por Coseriu

(1977), apresentam uma aplicabilidade universal e aparecem, posteriormente, formalizadas a partir das funções léxicas propostas por Mel'čuk (1995). Dessa forma, podemos considerar que, dentro desse grupo, o problema das colocações - ainda que não diretamente - é tratado como um fato ontológico da linguagem.

2.1.2 Abordagem estatística

Os trabalhos da abordagem estatística centram-se, especialmente, no aspecto probabilístico da linguagem e na frequência com que os elementos costumam aparecer em conjunto. Esses estudos adquiriram importante projeção com o avanço da informática e o advento da Lingüística de *Corpus*. A seguir, apresentamos os autores que abordaram o tema das colocações sob esta perspectiva em uma ordenação cronológica.

2.1.2.1 Firth (1957)

Firth (1957) foi o responsável pela introdução do termo “colocação” [*collocation*] em lingüística: “Proponho apresentar como um termo técnico, significado por ‘colocação’”⁵² (FIRTH, 1957, p. 194). Ainda que não defina claramente o que entende por esse conceito (cf. LYONS, 1977, p. 612; CORPAS PASTOR, 1996, p. 56), Firth (1957, p. 195) acredita que o significado de uma palavra depende de sua relação com os demais vocábulos do contexto. Nesse sentido, afirma que “um dos significados de *night* [noite] é sua colocabilidade com *dark* [sombria], e de *dark*, é claro, sua colocação com *night*”⁵³ (FIRTH, 1957, p. 196). Alonso Ramos (1994-95, p. 11), porém, utilizando os exemplos de *pelo* [cabelo] e *rubio* [louro], discorda da opinião de Firth (1957), pois acredita que o significado de *louro* não é determinado por sua colocabilidade com *cabelo*, uma vez que *louro* implica, em sua definição, uma menção a *cabelo*. Dessa forma, o significado dos vocábulos não é determinado pelo contexto, mas seleciona a coocorrência dos elementos.

⁵² [I propose to bring forward as a technical term, meaning by “collocation”].

⁵³ [One of the meanings of *night* is its collocability with *dark*, and of *dark*, of course, collocability with *night*].

Firth (1957) apresenta contribuições importantes no estudo das colocações ao utilizar as noções de frequência de coaparição e de colocabilidade restrita, pois percebe que algumas palavras ocorrem mais frequentemente em companhia de outras e que há uma possibilidade limitada de combinações. Segundo Hausmann (1985, p. 127), essa possibilidade de substituição paradigmática, posteriormente desenvolvida pelo conceito de “campo lexical” [*lexical set*] por Halliday (1961), constitui uma das maiores contribuições do Contextualismo britânico no estudo das colocações. No entanto, para Alonso Ramos (1994-95, p. 12), o mérito de Firth (1957) consiste somente em ter despertado o interesse pela coocorrência léxica, pois considera que seu estudo, baseado na frequência, apresenta uma visão muito limitada. O uso somente da frequência e a ausência de uma delimitação precisa do fenômeno das colocações fazem com que Firth (1957) inclua em seus trabalhos tanto o que chama de “colocações gerais ou usuais” [*general or usual collocations*], ou seja, combinações habituais, quanto as “colocações técnicas ou pessoais mais restritas” [*more restricted technical or personal collocations*], que corresponderiam a combinações únicas, relacionadas a um estilo particular, isto é, a um idioleto⁵⁴.

2.1.2.2 Halliday (1961, 1966)

Halliday (1961), ao desenvolver as idéias apresentadas por Firth (1957), situa as colocações no nível léxico (em oposição ao da gramática), no qual não há uma hierarquia entre os elementos. Ao contrário de Firth (1957), porém, formaliza uma definição para o que considera colocação: “associação sintagmática de itens lexicais, textualmente quantificáveis, como a probabilidade de que ocorram em n intervalos (uma distância de n itens lexicais) a partir do item x, os itens a, b, c ...”⁵⁵ (HALLIDAY, 1961, p. 276), introduzindo, assim, importantes conceitos como “item lexical”, “campo lexical”, “distância colocacional” e “probabilidade”.

⁵⁴ Segundo Bussmann (1990, s.v. *Idiolekt*), um idioleto corresponde a uma forma característica que um indivíduo tem de empregar a linguagem. Essa forma específica e pessoal de expressão manifesta-se diferenciadamente na pronúncia, no vocabulário e na sintaxe.

⁵⁵ [syntagmatic association of lexical items, quantifiable, textually, as the probability that there will occur, at n removes (a distance of n lexical items) from an item x, the items a, b, c ...].

A noção de item lexical⁵⁶, em oposição à palavra ortográfica, será importante para explicar a ocorrência de uma mesma combinação sob diferentes estruturas formais, como o próprio autor exemplificará posteriormente em:

Strong [forte], *strongly* [fortemente], *strength* [força] e *strengthened* [fortalecido] podem todas ser consideradas para o presente propósito como o mesmo item; e *a strong argument* [um forte argumento], *he argued strongly* [ele argumentou fortemente], *the strength of his argument* [a força de seu argumento] e *his argument was strengthened* [seu argumento foi fortalecido] são todos exemplos de uma e a mesma relação sintática⁵⁷ (HALLIDAY, 1966, p. 151).

Essas combinações, porém, admitem certa variação entre os elementos que compartilham o mesmo “campo léxico”. A distância colocacional, por sua vez, embora ainda não apareça precisamente delimitada, cumprirá uma função essencial no momento da detecção das combinações léxicas em *corpora* textuais. Finalmente, a probabilidade de coocorrência dos elementos, que permite delimitar estatisticamente uma colocação, parece também incluir combinações livres, pois, como afirma Halliday (1961, p. 276), “qualquer item dado, então, entra no alcance da colocação, sendo os itens com os quais ele se coloca classificados como mais ou menos prováveis”⁵⁸. Nesse sentido, surge também a crítica de Corpas Pastor (1996, p. 56) ao conceito de “colocação” adotado por Halliday (1961), para quem “o fenômeno das colocações parece reduzir-se à mera coaparição freqüente e linear das palavras léxicas no discurso”⁵⁹. Alonso Ramos (1994-95, p. 14) acredita que o fato de que dois lexemas coocorram freqüentemente não é motivo para considerá-los colocações, pois isso constitui o resultado da seqüência argumentativa que seguem os textos.

⁵⁶ Para uma discussão sobre a diferença entre “item lexical” e “palavra ortográfica”, v. Trask (2006, s.v. *palavra*). No presente trabalho, além de “item lexical”, utilizaremos como sinônimo o termo “unidade lexical”, entendida como membro de um conjunto aberto (JOTA, 1976, s.v. *conjunto aberto*). A primeira forma é empregada nos estudos de língua inglesa e a segunda, pelos autores que seguem uma tradição hispânica.

⁵⁷ [*Strong, strongly, strength* and *strengthened* can all be regarded for this present purpose as the same item; and *a strong argument, he argued strongly, the strength of his argument* and *his argument was strengthened* all as instances of one and the same syntactic relation]. Essa mesma noção será, posteriormente, desenvolvida por Mitchell (1975, p. 117), ao considerar a colocação como uma combinação entre raízes [*roots*] e não entre palavras [*words*].

⁵⁸ [any given item thus enters into a range of collocation, the items with which it is collocated being ranged from more or less probable].

⁵⁹ [el fenómeno de las colocaciones parece reducirse a la mera coaparición frecuente y lineal de las palabras léxicas en el discurso].

2.1.2.3 Sinclair (1991)⁶⁰

Sinclair (1991) insere os estudos sobre as colocações no segundo ponto da dicotomia entre o “princípio da escolha aberta” [*the open-choice principle*] e o “princípio idiomático” [*the idiom principle*] da língua. Para o primeiro, as palavras combinam-se livremente, respeitando apenas a gramaticalidade da construção. Já para o segundo, “as palavras não ocorrem ao acaso em um texto”⁶¹ (SINCLAIR, 1991, p. 110), mas estão sujeitas a diversas restrições. Essas restrições devem-se ao fato de que o falante dispõe de um número de estruturas “pré-fabricadas” que auxiliam em sua produção lingüística, tendo em vista a economia e a agilidade da comunicação.

Para esse autor, as colocações aparecem, inicialmente, definidas como

a coocorrência de dois elementos em um texto dentro de um ambiente específico. Uma colocação significativa é uma colocação habitual de elementos, tal que sua coocorrência é mais freqüente do que se poderia prever por suas respectivas freqüências e pela extensão do texto em que aparecem”⁶² (JONES, SINCLAIR, 1974, p. 19).

Posteriormente, Sinclair (1991, p. 170) ampliaria sua definição de colocação para “a ocorrência de duas ou mais palavras com um curto espaço entre cada uma delas em um texto”⁶³. A distância entre os membros que conformam uma colocação será delimitada por Jones, Sinclair (1974) como um espaço de quatro elementos à direita e à esquerda da palavra-chave. Além dessa contribuição, Jones, Sinclair (1974) incluíram o uso de amplos *corpora* textuais no estudo das colocações e introduziram conceitos importantes como o de “nó” [*node*] e “colocado” [*collocate*]. O primeiro refere-se à palavra que se está consultando no *corpus* e o segundo, aos elementos que se associam a ela; no entanto, não há uma diferença de *status* entre “nó” e “colocado”, como será postulado posteriormente por Hausmann (1989).

O enfoque estatístico proposto por Sinclair (1991), baseado unicamente na freqüência de coaparição dos elementos, trata as unidades léxicas de forma isolada, sem

⁶⁰ Ainda que este autor seja responsável pela elaboração de importantes dicionários da casa editorial *Collins*, os quais constituem obras de referência para a língua inglesa, optamos por não incluí-lo na abordagem lexicográfica, visto que o enfoque principal de Sinclair (1991) no estudo das colocações diz respeito a sua identificação em *corpora* textuais, sem uma preocupação com sua forma de apresentação em obras lexicográficas. Williams (2003, p. 41) afirma que somente após perceberem a necessidade de inclusão das combinações nos dicionários, os pesquisadores do projeto *Collins COBUILD* conferiram um tratamento mais formal a essas combinações léxicas.

⁶¹ [words do not occur at random in a text].

⁶² [the co-occurrence of two items in a text within a specified environment. “Significant” collocation is regular collocation between items, such that they co-occur more often than their respective frequencies and the length of text in which they appear would predict].

⁶³ [the occurrence of two or more words within a short space of each other in a text].

considerar as relações sintáticas e semânticas estabelecidas entre elas (CORPAS PASTOR, 1996, p. 57). Nesse sentido, Bosque (2001, p. 15) sugere que o termo “colocação” seja substituído por “coaparição” [*coaparición*] nos estudos de Sinclair (1991), uma vez que não designa um conceito propriamente lingüístico, por utilizar somente o critério da frequência. Além disso, a partir da coocorrência freqüente, Bosque (2001, p. 12) constata que muitas combinações livres são incluídas por Sinclair em seu dicionário de colocações (*Collins Cobuild English Collocations on CD ROM*), ao qual, infelizmente, não tivemos acesso.

2.1.2.4 Tagnin (2005a)

No Brasil, as pesquisas sobre o tema das colocações aparecem representadas por Tagnin (2005a) que, ao estudar o que chama de “expressões convencionais”, estabelece três níveis de convencionalidade: sintático, semântico e pragmático. As colocações são inseridas no primeiro nível e caracterizam-se pela combinabilidade entre os elementos, ou seja, são combinações de palavras que costumam “andar juntas” devido ao uso consagrado, sem que haja uma explicação para tal fato. Tagnin (2005a, p. 37-38) ressalta que há casos em que a coocorrência entre as duas palavras que conformam uma colocação é extremamente restrita, como entre *velha* e *coroca*, em que o adjetivo se combina somente com esse substantivo; porém, em outros casos, há uma maior liberdade de escolha dentro de um mesmo campo semântico, como ocorre com *impending* [imminente], que pode se combinar com *doom* [castigo], *disaster* [desastre], *horror* [horror] e *evil* [mal]⁶⁴. No entanto, a autora não aprofunda a idéia de relações semânticas entre os elementos que conformam as colocações, apresentando apenas uma organização das estruturas morfossintáticas que compõem esse fenômeno. Ao afirmar que “em geral, não há regra sintática ou semântica que justifique a coocorrência desses dois (ou mais) elementos”⁶⁵ (TAGNIN, 2005a, p. 30), utiliza como único critério para identificar as colocações sua frequência de coaparição. A autora difere, porém,

⁶⁴ Essa noção corresponde à de prosódia semântica, utilizada em Lingüística de *Corpus*, que diz respeito a “associação entre itens lexicais e conotação (negativa, positiva ou neutra) ou instância avaliativa” (BERBER SARDINHA, 2004, p. 40).

⁶⁵ Opinião semelhante apresenta Lipka (2002, p. 183): “Regras sintáticas ou semânticas não podem ser formuladas para elas [*sc.* as colocações]” [Syntactic or semantic rules cannot be formulated for them]. Para Lo Cascio (1997, p. 67), porém, “sua combinação [*sc.* das palavras] não é arbitrária, mas deve responder a critérios de coerência sintático-categorial, coerência semântica e enciclopédica” [la loro combinazione non è arbitraria ma deve rispondere a criteri di coerenza sintattico-categoriale, coerenza semantica ed enciclopedica].

dos estudos anteriores ao reconhecer uma hierarquia entre os membros da colocação, pois “a base é a palavra que [...] determina a ocorrência da outra”⁶⁶ (TAGNIN, 2005a, p. 30). Com essa constatação, ainda que afirme não haver uma regra que explique a coocorrência das palavras, a própria autora identifica que um dos elementos da colocação orienta ou seleciona a ocorrência do outro.

2.1.2.5 Síntese da abordagem estatística

Dentro da abordagem estatística, a principal característica apresentada pelas colocações é a frequência de coaparição dos elementos que as compõem, originada pela visão da língua como um sistema probabilístico. Essa frequência, para Sinclair (1991), também é motivada pela visão das colocações como combinações pré-fabricadas, em conformidade com os autores da abordagem semântica que as situavam na língua (aqui entendida em oposição à fala). No entanto, Halliday (1961) e Sinclair (1991) acreditam que ambos os elementos apresentam o mesmo *status*, diferindo de Tagnin (2005a) que, mais recentemente, já reconhece uma hierarquia entre base e colocado, o que também a aproxima das idéias da abordagem semântica.

O uso da frequência de coaparição como o principal índice para a identificação das colocações, tal como postulam os autores da abordagem estatística, é questionado em diversos estudos como Hausmann (1985), Alonso Ramos (1994-95), Bosque (2001), entre outros. Até mesmo Mitchell (1975), considerado, assim como Halliday e Sinclair, um neo-Firthiano (cf. MAURER STROH, 2004; CORPAS PASTOR, 1996), afirma que “uma colocação não é uma simples justaposição ou coocorrência”⁶⁷ (MITCHELL, 1975, p. 120). Dessa forma, consideramos que o cálculo da frequência de coocorrência dos elementos de uma combinação deve ser empregado com cautela no estudo das colocações.

Os trabalhos desenvolvidos pelos autores da abordagem estatística foram também importantes para o desenvolvimento da Lingüística de *Corpus*. Segundo Williams (2003, p. 36), “o Contextualismo e a Lingüística de *Corpus* são indissociáveis”⁶⁸, sendo esta essencial

⁶⁶ Essa noção de direcionalidade na colocação é também compartilhada por Hausmann (1985), Mel’čuk (1995), Alonso Ramos (2001), entre outros.

⁶⁷ [a collocation is not a mere juxtaposition or co-occurrence].

⁶⁸ [le contextualisme et la linguistique de *corpus* sont indissociables].

para o estudo das colocações ao proporcionar um grande conjunto de dados para serem analisados. Além disso, a incorporação de ferramentas informatizadas permitiu a identificação de colocações cujos elementos não aparecem imediatamente unidos no texto, o que, segundo Hausmann (1985, p. 128), representa uma importante contribuição do Contextualismo britânico.

Finalmente, além da possibilidade de detectar estatisticamente as recorrências léxicas, acreditamos que a principal contribuição dessa abordagem para os estudos das colocações aparece representada por Halliday (1966), ao identificar que uma mesma combinação pode ser apresentada sob diferentes estruturas formais. Segundo Lipka (2002, p. 181), essa característica atribuída às colocações, acrescida da ausência de hierarquia entre os elementos e de uma análise semântica, confere a essas combinações um caráter neutro. Devido a essa vagueza e neutralidade, o autor acredita que o conceito de colocação postulado pelos contextualistas apresenta “certas vantagens em capturar as relações sintagmáticas entre os elementos lexicais”⁶⁹ (LIPKA, 2002, p. 181).

2.2 Enfoque lexicográfico

Os autores incluídos na presente seção partem todos de um questionamento prático de como as combinações léxicas deveriam ser incluídas nos dicionários. Como as colocações correspondem a estruturas formadas por dois elementos, existem divergências se sua inclusão nessas obras deve ser realizada a partir da base ou do colocado⁷⁰. Para responder a essa questão, os lexicógrafos perceberam que era necessário realizar um estudo lexicológico prévio, isto é, era preciso delimitar as combinações a serem trabalhadas e identificar as características dos elementos que conformavam cada colocação. Em virtude dos diferentes recortes estabelecidos, surgiram diferentes propostas para o tratamento lexicográfico desse fenômeno léxico, as quais serão discutidas em 7.1.3.1. A seguir, apresentamos os trabalhos

⁶⁹ [certain advantages in capturing syntagmatic relations between lexical elements].

⁷⁰ O lexicógrafo, ao decidir por apresentar as colocações a partir da base ou do colocado, deve levar em consideração as necessidades do usuário de sua obra. Nesse sentido, é importante estabelecer a distinção entre dicionários para a produção e para a recepção de textos e identificar em qual destes contextos as colocações poderiam oferecer maiores dificuldades. Além disso, conforme o usuário necessite produzir ou decodificar um texto, poderá consultar o dicionário a partir de entradas diferentes. Aprofundaremos essa discussão no capítulo 6 deste trabalho.

que refletem sobre as relações estabelecidas entre os elementos das colocações em uma ordenação cronológica.

2.2.1 Cowie (1981, 1994)

Ao identificar que cada tipo de combinação léxica vai oferecer uma dificuldade diferente ao usuário do dicionário, Cowie (1981) propõe uma classificação dessas unidades a partir de seu grau de transparência e de sua variação léxica e gramatical, ou seja, suas possibilidades de substituição e/ou inserção de elementos. Dessa forma, são identificados quatro tipos de combinações:

- a) colocações abertas [*open collocations*], que apresentam um significado transparente e podem ser livremente combinadas, como em *drink one's tea* [beber um chá]. Esse tipo de colocações corresponde às combinações livres para os demais autores.
- b) colocações restritas [*restricted collocations*], cujo significado total é transparente. São formadas por um elemento com significado literal e outro com significado figurado. O elemento figurado determina a ocorrência do outro, havendo uma restrição arbitrária entre eles, como ocorre em *wholesome fare* [comida saudável], em que o adjetivo pode ser substituído por um grupo restrito de vocábulos como *plain* [pouco industrializada] e *simple* [leve]. Dessa forma, as colocações são definidas como a associação de lexemas realizados em diferentes estruturas gramaticais (COWIE, 1994, p. 3169), o que vai ao encontro das propostas de Halliday (1966) e Mitchell (1975), sem que haja uma diferenciação no *status* de cada elemento (cf. NESSELHAUF, 2005, p. 17). Caracterizam-se, então, por seu significado transparente e pela seleção restrita e arbitrária dos vocábulos que podem ser combinados (cf. COWIE, 1994, p. 3169).
- c) expressões idiomáticas “figuradas” [*figurative idioms*], que correspondem a uma nova categoria proposta por Cowie (1981) entre as colocações e as combinações fixas. Compreende as combinações caracterizadas por apresentarem um significado literal e outro figurado, cuja substituição de algum elemento é bastante rara. Como exemplo, temos *close ranks* que,

literalmente, significa “fechar uma fileira (em um grupo militar)”, mas também apresenta o significado de “defender outros membros do grupo”, correspondendo à combinação *cerrar fileiras*, em português. Corpas Pastor (2001) também destaca a existência de combinações com um significado literal e outro figurado, como no exemplo supracitado *meter un gol* (v. 2.1.1.5); porém, para essa autora, no primeiro caso, estamos diante de uma colocação e no segundo, de uma locução.

- d) expressões idiomáticas “puras” [*pure idioms*], cujo significado é completamente opaco e a estrutura é invariável, como em *spill the beans* [contar um segredo], em que as palavras, separadamente, significam “derramar” [*spill*] e “feijões” [*beans*].

Segundo Cowie (1981, p. 229), “figurative” e “pure idioms” costumam ser abordados como uma única categoria nos estudos de fraseologia, e essa distinção constitui a principal contribuição desse autor no tratamento da combinatória léxica. No entanto, de acordo com Nesselhauf (2005, p. 25), o problema da classificação proposta por Cowie (1981) é que, embora os critérios da transparência e da comutabilidade “sejam relacionados em certo grau, eles não coincidem de forma regular”⁷¹. Assim, algumas combinações podem ser classificadas em grupos diferentes, conforme seja privilegiado um ou outro critério, como ocorre no exemplo de *commit a crime* [cometer um crime], em que existe uma comutabilidade restrita - como nas colocações -, mas nenhum elemento apresenta um significado figurado. Além disso, Howart (1998 apud NESSELHAUF, 2005, p. 26) ressalta a dificuldade de definir quando um elemento é utilizado em seu significado literal ou figurado. Em virtude desse fato, Nesselhauf (2005, p. 27) propõe a utilização somente do critério da comutabilidade para a classificação das combinações léxicas.

2.2.2 Haensch et al (1982)

Ettinger (1982) divide as unidades léxicas pluriverbais em “colocações” [*colocaciones*] e “combinação fixa de lexemas” [*combinación fija de lexemas*] devido à maior possibilidade de combinações dos lexemas no primeiro grupo. Define, então, as colocações

⁷¹ [correlating to some degree, do not regularly coincide].

como “a tendência sintático-semântica das palavras isoladas de uma língua a adotar somente um número limitado de combinações com outras palavras entre uma grande quantidade de possíveis combinações”⁷² (ETTINGER, 1982, p. 251). Acrescenta que, por motivos práticos, além dos critérios sintático e semântico, também devem ser considerados critérios extralingüísticos no estudo das colocações. Nesse sentido, aproxima-se da noção de solidariedade referencial, o que, segundo Salvador (1989-90), não corresponderia a colocações.

Para Corpas Pastor (1996, p. 46), “o tratamento dado por esses autores às UFS [*i.e.* unidades fraseológicas] está longe de ser sistemático”⁷³. Ao estabelecer como critério para a divisão entre as colocações e as combinações fixas o seu grau de restrição combinatória e utilizar esse mesmo critério para a separação entre os diversos tipos de combinações fixas, torna desnecessária a primeira divisão. Além disso, posteriormente, Haensch (1982, p. 415-416), ao tratar das unidades léxicas que devem compor a nominata do dicionário, classifica como dois tipos diferentes as “colocações usuais” [*colocaciones usuales*], exemplificadas por *un éxito clamoroso* [um sucesso absoluto] ou *un soltero empedernido* [um solteiro inveterado], e as “unidades fraseológicas (de significado transparente)” [*unidades fraseológicas (de significado transparente)*], como *dar un paseo* [dar um passeio] ou *abrir una cuenta* [abrir uma conta]. Acreditamos, porém, que ambas as combinações pertencem ao grupo das colocações, pois tanto os primeiros exemplos apresentam um significado transparente quanto os segundos constituem combinações usuais. Além disso, a separação proposta pelo autor é fruto da utilização de dois critérios diferentes - semântico e estatístico - para delimitar as combinações léxicas.

2.2.3 Benson (1985, 1986, 1997)

Benson (1986) situa as colocações entre as combinações livres e as fixas [*idioms*] e as define como um grupo de palavras recorrente na língua que forma uma construção fixa, identificável, na qual o significado total reflete o significado de seus constituintes⁷⁴ (cf.

⁷² [la tendencia sintáctico-semántica de las palabras aisladas de una lengua a adoptar tan sólo un número limitado de combinaciones con otras palabras entre una gran cantidad de posibles combinaciones].

⁷³ [el tratamiento que estos autores dan a las UFS dista mucho de ser sistemático].

⁷⁴ [a group of words that occurs repeatedly, i. e. recurs, in a language]; [the meaning of the whole does reflect the meaning of the parts]; [fixed, identifiable, non-idiomatic phrases and constructions].

BENSON, 1985, p. 61; 1986, p. 4-5; 1997, p. xv). Apesar de sua transparência semântica, as colocações são combinações imprevisíveis, devido ao fato de serem arbitrárias. Esse caráter arbitrário fica evidenciado somente no contraste entre as línguas (BENSON, 1986, p. 11).

Benson (1997) divide as colocações entre gramaticais [*grammatical collocation*] e lexicais [*lexical collocation*]. As primeiras incluem as combinações de uma palavra dominante (substantivo, adjetivo ou verbo) e uma preposição ou estrutura gramatical. Já as colocações lexicais são formadas por substantivos, adjetivos, verbos e advérbios, sendo que entre eles não há uma palavra dominante. Sobre o tratamento conjunto das colocações lexicais e gramaticais, compartilhamos a opinião de Alonso Ramos (1994-95, p. 20):

Do nosso ponto de vista, dizer que determinado lexema verbal rege uma oração de infinitivo ou que admite a alternância de dativo não é uma informação colocacional. Certamente essa informação deve aparecer em uma descrição do léxico, mas não vemos o interesse de tratá-la como colocação no BBI (...) Acreditamos que a coocorrência sintática e a coocorrência léxica devem ser tratadas separadamente⁷⁵.

Dessa forma, no presente trabalho, trataremos apenas das colocações lexicais de Benson (1997).

Dentro desse grupo, o autor propõe uma classificação em sete diferentes combinações, o que, segundo Koike (2000, p. 12) “constitui a primeira tentativa de apresentar uma taxonomia desse tipo de colocações”⁷⁶. São elas:

- a) verbo + substantivo, na qual o verbo denota criação ou ativação, como em *compose music* [compor uma música] e *launch a missile* [lançar um míssil], respectivamente.
- b) verbo + substantivo, na qual o verbo denota erradicação ou anulação, como em *annul a marriage* [anular um casamento].
- c) adjetivo + substantivo, como *strong tea* [chá forte].
- d) substantivo + verbo, em que este denota uma ação característica do substantivo, como em *bombs explode* [uma bomba explode].
- e) substantivo + (preposição) + substantivo, em que se estabelecem relações com uma unidade ou com a coletividade, como em *an act of violence* [um ato de violência] e *a bouquet of flowers* [um buquê de flores], respectivamente.
- f) advérbio + adjetivo, como *deeply absorbed* [profundamente absorto].
- g) verbo + advérbio, como em *argue heatedly* [discutir calorosamente].

⁷⁵ [A nuestro modo de ver, decir que un lexema verbal dado rige una oración de infinitivo o que admite la alternancia de dativo no es una información colocacional. Por supuesto que esa información debe aparecer en una descripción del léxico pero no vemos el interés de tratarla como colocación en el BBI (...) Creemos que la coocurrencia sintáctica y la coocurrencia léxica deben ser tratadas separadamente].

⁷⁶ [constituye el primer intento taxonómico de este tipo de colocaciones].

Apesar de organizar uma taxonomia centrada em questões morfológicas e sintáticas, Benson (1997) procura introduzir o componente semântico em sua classificação ao separar as combinações com verbo conforme o que este denote (itens a. e b.). No entanto, consideramos que tal divisão seja de pouca utilidade, pois além dessas duas categorias, Corpas Pastor (1996, p. 71) cita outros quatro tipos de funções léxicas relacionadas à combinação de “verbo + substantivo”. Mesmo Benson (1985, p. 65) já havia percebido a existência de outras relações diferentes entre o verbo e o substantivo, ao afirmar que: “o lexicógrafo não deve hesitar em incluir [*sc.* no dicionário] importantes colocações imprevisíveis de verbo + substantivo, ainda que elas não signifiquem ‘criação’ ou ‘ativação’”⁷⁷.

Além disso, por ser baseada em aspectos morfossintáticos, a classificação proposta por Benson (1997) apresenta estruturas aplicáveis somente à língua inglesa, especialmente nas colocações gramaticais, que não serão tratadas em nosso trabalho. No caso das colocações lexicais, para a língua espanhola, Koike (2000, p. 43) acrescenta a existência de combinações formadas por verbo e adjetivo, como *resultar ileso* [sair ileso]. Em virtude dessas limitações apresentadas pelas classificações morfossintáticas, Wanner (1996, p. 20) salienta a universalidade das funções léxicas, que constituem um viés mais semântico.

2.2.4 Seco (1987a, 1987b)

Segundo Corpas Pastor (1996, p. 61), o termo “colocação” [*colocación*] foi introduzido em língua espanhola por Seco (1987a [1978], p. 16), que afirma ter se inspirado no conceito proposto por Firth (1957). Sua noção de colocação, no entanto, não se limita à coocorrência léxica, mas, devido a seu ponto de vista lexicográfico, Seco (1987a, p. 25) a define como uma explicação sobre o uso de determinado vocábulo⁷⁸. Dessa forma, inclui as

⁷⁷ [the lexicographer should never hesitate to include important unpredictable verb + noun collocations even if they do not mean “creation” or “activation”].

⁷⁸ Castillo Carballo (1998, p. 49) também concorda que, apesar de afirmar seguir a noção Firthiana de colocação, Seco (1987a) não deixa muito clara essa concepção; pelo contrário, tal interpretação se desvanece ao longo da obra.

colocações no primeiro enunciado⁷⁹ de um verbete, juntamente com a etimologia, as marcações geográficas, de uso, etc.

Posteriormente, Seco (1987b [1979]) substitui o termo “colocação” pela incorporação do “contorno” [*contorno*] na definição lexicográfica e, ainda que mantenha as combinações léxicas no primeiro enunciado, aproxima-as das definições. Em oposição ao “conteúdo” da definição, que fornece informações sobre a constituição sêmica da palavra-entrada, o “contorno” auxilia na compreensão de seu uso. Seco (1987a, p. 28) ressalta, assim, que o dicionário deve indicar

todos aqueles elementos que são “contorno” necessário da palavra definida, mas que não são componentes semânticos dela; por exemplo, nos verbos, o objeto direto, o objeto indireto, o sujeito; nos substantivos, o complemento de “posse”, (...) em adjetivos (...), é “contorno” necessário o substantivo⁸⁰.

Essas concepções são empregadas por Seco na elaboração do DEA (1999), através da utilização dos colchetes para marcar os “contornos”, como pode ser observado nos exemplos que seguem:

- a) Mover [a alguien o algo] hacia arriba, o poner[lo] en un lugar más alto (DEA, 1999, s.v. *elevantar*, ac. 1);
- b) Dirigir [un escrito o petición a una autoridad] (DEA, 1999, s.v. *elevantar*, ac. 4);
- c) Personaje principal [de una obra literaria, teatral o cinematográfica] (DEA, 1999, s.v. *protagonista*, ac. 1);
- d) [Cuerpo de metal] que está al rojo (DEA, 1999, s.v. *candente*, ac. 1).
- e) [Cosa] muy viva y del momento (DEA, 1999, s.v. *candente*, ac. 2).

Através das definições propostas para *elevantar* e *candente*, podemos perceber que a noção de “contorno” aplica-se tanto a relações estritas entre os vocábulos, como nos exemplos de (b) e (d), quanto a relações com toda uma classe, conforme os exemplos (a) e (e). Em vista disso, tal noção apenas se aproxima da de colocação, pois enquanto estas correspondem a relações restritas entre as palavras, o “contorno” pode recobrir a marcação de qualquer complemento da palavra-entrada, sendo, portanto, mais abrangente. Nesse sentido, o contorno estaria relacionado à noção de solidariedade léxica por afinidade proposta por Coseriu (1977).

⁷⁹ Seco (1987a) distingue dentro de um verbete entre o primeiro e o segundo enunciado. Aquele corresponde a informações que podem ser facilmente padronizadas e dizem respeito ao uso da palavra-entrada. Já o segundo enunciado corresponde à informação sobre o conteúdo do lema, ou seja, à definição lexicográfica propriamente dita. Em lexicografia essas noções serão recobertas pelos conceitos de “comentário de forma” [*comment on form*] e “comentário semântico” [*comment on semantics*], respectivamente (cf. HAUSMANN, WIEGAND, 1989, p. 353; HARTMANN, JAMES, 1998, s.v. *comment*).

⁸⁰ [todos aquellos elementos que son “contorno” necesario de la palabra definida, pero que no son componentes semânticos de ella; por ejemplo, en los verbos, el complemento directo, el complemento indirecto, el sujeto; en los nombres, el complemento de “posesión”, (...) en adjetivos (...), es “contorno” necesario el nombre].

Poderíamos, então, dizer que a principal diferença entre as colocações e o contorno é que aquela enfatiza uma relação léxica e semântica, enquanto este uma relação sintático-semântica.

2.2.5 Hausmann (1979, 1985, 1989, 1998a)

Hausmann (1985, p. 118) situa as colocações no nível da língua e, mais especificamente, na norma, por não serem combinações criadas livremente pelos falantes, e as diferencia das combinações livres e das fixas por apresentarem uma combinabilidade restrita entre os vocábulos, serem transparentes e não-fixas (HAUSMANN, 1989, p. 1010). Esse autor, segundo Zuluaga (2002, p. 59), foi responsável pela introdução dos termos “base” [*base*] e “colocado” [*collocatif*] nos estudos sobre as colocações, o que consideramos ter sido sua principal contribuição⁸¹. Para Hausmann (1989), o primeiro corresponde ao elemento autônomo do ponto de vista semântico e o segundo refere-se ao membro da combinação que acrescenta uma caracterização, mas não modifica a identidade da base. Segundo Hausmann (1979, p. 192), “a base completa a definição do colocado, enquanto o colocado apenas acrescenta uma qualidade à base, suficientemente definida em si mesma”⁸². Dessa forma, o autor ressalta que os dois elementos da colocação não possuem o mesmo *status*, pois a colocação é orientada da base ao colocado.

Essa distinção, apesar de ser relevante no momento da produção lingüística, em que o falante parte da base ao colocado para construir seu texto (cf. ALONSO RAMOS, 2001, p. 10; HAUSMANN, 1985, p. 119), não parece aplicar-se a todos os casos de colocações. Consideramos que a hierarquia entre os elementos pode ser facilmente identificada no exemplo proposto por Hausmann (1989), entre um substantivo e um adjetivo, ou nos casos de combinação com verbo-suporte, como em *dar um passeio*, em que evidentemente o

⁸¹ Os termos “base” e “colocado” costumam ser equiparados aos de “nó” [*node*] e “colocado” [*collocate*], propostos por Jones, Sinclair (1974); no entanto, para Hausmann (1985, p. 126), esses termos não se equivalem, pois os contextualistas não estabeleceram uma distinção conseqüente entre base e colocado. Nos estudos desenvolvidos dentro da abordagem estatística, o “nó” e o “colocado” diferem somente no fato de o primeiro ser a palavra utilizada nas consultas ao *corpus*. Por outro lado, nos trabalhos de Hausmann (1989) - e também Corpas Pastor (1996), Lo Cascio (1997), Alonso Ramos (2001), entre outros -, a diferença entre a “base” e o “colocado” fundamenta-se no fato de a primeira apresentar um *status* autônomo e, conseqüentemente, orientar a direcionalidade da combinação.

⁸² [la base complète la définition du collocatif, allors que le collocatif se contente d'ajouter une qualité à une base en elle-même suffisamment définie].

substantivo (base) é o responsável pelo “conteúdo sêmico” da expressão. Porém, em combinações como *formular uma hipótese*, acreditamos que o verbo também cumpra uma função importante para a constituição sêmica do enunciado⁸³.

Partindo, então, da concepção de que os membros de uma colocação não possuem o mesmo *status*, Hausmann (1989, p. 1010) define tais construções como combinações binárias que apresentam as seguintes estruturas:

- a. substantivo + adjetivo, como em *confirmed bachelor* [solteiro inveterado]
- b. substantivo + verbo, como *his anger falls* [sua raiva passa]
- c. verbo + substantivo, como em *to withdraw money* [sacar dinheiro]
- d. verbo + advérbio, como *it is raining heavily* [está chovendo torrencialmente]
- e. adjetivo + advérbio, como *seriously injured* [gravemente ferido]
- f. substantivo + (preposição) + substantivo, como *a gust of anger* [um ataque de raiva]

Nestas combinações, a base será sempre o substantivo, exceto na construção “verbo + advérbio”, em que será o verbo, e “adjetivo + advérbio”, em que a base corresponde ao adjetivo. É importante salientar que, por utilizar critérios morfossintáticos para a organização de sua taxonomia, as mesmas limitações oferecidas pela classificação de Benson (1997) - apresentadas em 2.2.3 -, podem ser aplicadas à de Hausmann (1989). Wanner (1996, p. 16) acrescenta que essa classificação não é suficientemente detalhada nem para uma descrição teórica do fenômeno das colocações nem como um auxílio para a produção lingüística.

Posteriormente, Hausmann (1998a, p. 66) formaliza uma definição para as colocações como “a combinação adequada de uma palavra auto-semântica com uma sinsemântica”⁸⁴. A introdução dessa nova terminologia⁸⁵, a partir de Hausmann (1995),

⁸³ Também Otani (2005, p. 3-4) acredita que “alguns tipos de seleção mútua normalmente acontecem com diferentes graus de prioridade entre um substantivo e um verbo, dependendo da situação” [some kind of mutual selection usually takes place with varying balances of priority found between a noun and a verb, depending on the situation]. Siepmann (2005, p. 418-422), por sua vez, apresenta quatro problemas inerentes à distinção entre elemento autônomo e dependente. São eles: (a) a linha que distingue um elemento autônomo de um dependente é muito tênue; (b) o caráter colocacional de uma combinação é indicado pelo contexto e não pela relação hierárquica entre seus elementos; (c) a distinção entre elemento autônomo e dependente pressupõe a descrição do segundo a partir do primeiro, o que exclui toda a gama combinatória do elemento dependente e (d) existem colocações formadas por dois elementos autônomos semanticamente.

⁸⁴ [a combinación dunha palabra autosemántica cunha sinsemántica axeitada].

⁸⁵ Tradicionalmente, as palavras “auto-semânticas” correspondem a lexemas que possuem um significado léxico próprio e contextualmente independente, podendo fazer parte de um sintagma, como os substantivos, adjetivos, verbos e advérbios (GLÜCK, 2000, s.v. *Autosemantikum*). Por outro lado, as palavras “sinsemânticas” caracterizam-se pela ausência de um significado lexical autônomo, de forma que não podem constituir um sintagma. Como exemplos desse grupo, temos os artigos, preposições e conjunções (GLÜCK, 2000, s.v.

considera que as primeiras correspondem a palavras semanticamente autônomas, que podem ser aprendidas sem recorrência a um contexto sintagmático; já as segundas correspondem a palavras semanticamente dependentes, que não podem ser definidas sem recorrência a um contexto⁸⁶. Equivalem-se, portanto, às noções de base e colocado, respectivamente. As palavras “sinsemânticas” são “imprevisíveis, geralmente diferentes de uma língua para outra, freqüentemente figuradas, preferenciais em relação a outras logicamente possíveis, mas não realizadas na língua”⁸⁷ (HAUSMANN, 1998a, p. 66). O caráter arbitrário e idiosincrático das colocações é ressaltado por Hausmann (1989) para sua inclusão nos dicionários, especialmente no que se refere à produção de textos, pois a escolha do colocado adequado será a maior dificuldade para o aprendiz de uma língua estrangeira⁸⁸.

Somente a partir do contraste entre as línguas, então, é possível perceber essa arbitrariedade que caracteriza as colocações, pois, segundo Hausmann (1989, p. 1013), “a idiosincrasia da colocação não é ressaltada definitivamente senão sob a ótica de uma outra língua que combina, para exprimir o mesmo fato, palavras diferentes”⁸⁹. Esse autor oferece exemplos comparativos entre as línguas inglesa, francesa e alemã. Porém, podemos identificar essa divergência na escolha do colocado também para a língua portuguesa através da tradução das colocações propostas por Hausmann (1989), as quais foram apresentadas anteriormente. Na combinação formada por um adjetivo e um advérbio (item e.), por exemplo, temos para o inglês a colocação *seriously injured*, cujas palavras, tomadas isoladamente, significam “seriamente” e “ferido”. Ambos os vocábulos fazem parte do léxico ativo dos falantes de língua portuguesa; no entanto, para manifestar que alguém está muito ferido, nossa língua optou por combinar esse adjetivo com o advérbio “gravemente”, resultando na colocação

Synsemantikon). Para outras noções desses conceitos, v. Hartmann, James (1998, s.v. *functional word e content word*).

⁸⁶ Hausmann (1995, p. 21) percebe, porém, que a divisão entre palavras “auto-semânticas” e “sinsemânticas” não é tão dicotômica, pois verbos e adjetivos, tradicionalmente considerados “auto-semânticos”, podem apresentar também um comportamento “sinsemântico”. Como exemplo, destaca o caso do verbo francês *serrer* [apertar] que, em combinações como *serrer le dents* [cravar os dentes] e *serrer le frein* [pisar no freio], apresenta um comportamento “sinsemântico”, correspondendo a um verbo-suporte. Para essa distinção entre os adjetivos, v. Menuzzi (1992) e Borges Neto (1992).

⁸⁷ [impredicibles, habitualmente diferentes dunha lingua a outra, frecuentemente figuradas, preferenciais en relación a outras lxicamente posibles pero non realizadas na lingua].

⁸⁸ Calderón Campos (1994, p. 81), ao analisar os erros de produção cometidos por alunos aprendizes do espanhol, afirma que “se observamos os erros, percebemos que o aluno não costuma se equivocar ao escolher o que chamamos de núcleo ou base da colocação (...); o que desconhece é com que palavra tal elemento conhecido se combina, ou seja, qual é o ‘colocado’” [si observamos los errores, advertimos que el alumno no se suele equivocar al elegir lo que hemos llamado el núcleo o la base de a colocación (...); lo que desconoce es con qué palabra se combina dicho elemento conocido, es decir, cuál es el “colocativo”].

⁸⁹ [l’idiosyncrasie de la collocation ne se relève définitivement que dans l’optique d’un autre langue qui combine, pour exprimer le même fait, des mots différents]. Opinião semelhante apresentam autores como Allerton (1984, p. 27), Wanner (1996, p. 1) e Alonso Ramos (2003, p. 47), entre outros.

gravemente ferido. A partir dessas constatações, percebemos a importância de analisar as colocações também de forma contrastiva.

2.2.6 Síntese do enfoque lexicográfico

Dentro do enfoque lexicográfico, identificamos que uma característica recorrente na descrição das colocações, além da noção de combinação de restrição variável entre os elementos, é seu significado transparente. No entanto, apesar de considerarem as colocações como combinações semanticamente transparentes, os autores indicam que o colocado pode assumir um caráter figurado (cf. COWIE, 1981; HAUSMANN, 1998a), o que resulta na imprevisibilidade desse tipo de combinação. O caráter imprevisível e arbitrário das colocações é responsável pelo alto grau de dificuldade dessas estruturas para os aprendizes de uma língua estrangeira e, devido a esse fato, os lexicógrafos perceberam a necessidade de incluí-las nos dicionários. Dessa forma, manifesta-se a importância da realização de um estudo contrastivo entre as línguas para a identificação das combinações mais problemáticas.

Devido à necessidade de inclusão das colocações nos dicionários, Hausmann (1989) postula que existe uma hierarquia entre base e colocado⁹⁰, diferindo de Cowie (1981) e Benson (1997), para quem os elementos possuem o mesmo *status*. Finalmente, Haensch et al (1982) e Benson (1997) destacam também a frequência de coaparição dos elementos das combinações⁹¹. As divergências entre as concepções apresentadas pelos autores que partem de um enfoque lexicográfico possibilitam uma associação entre os que privilegiaram uma abordagem mais estatística - ainda que não utilizando os conceitos propostos por essa visão -, como Haensch et al (1982) e Benson (1997), e os que privilegiaram as relações semânticas

⁹⁰ Zuluaga (1998, p. 23) critica a opinião de Hausmann (1979, p. 192) de que, em função dessa hierarquia, o colocado só pode ser definido em relação a sua base, pois acredita que, “se assim fosse, tais combinações deveriam sua existência a uma relação de implicação semântica entre os seus elementos constitutivos e não propriamente ao uso repetido, de forma que teríamos que deixar de considerá-las colocações” [se así fora, ditas combinacións deberían a súa existencia a unha relación de implicación semántica entre os seus elementos compoñentes e non propiamente ó uso repetido, co que teríamos que deixar de consideralas como colocacións]. Hausmann (1998b, p. 102) adverte, porém, que a definição do colocado em função de sua base está relacionada ao tratamento lexicográfico (e não teórico) das colocações, o que não significa que o colocado não possua um significado autônomo em nível teórico.

⁹¹ Ainda que Benson (1997) afirme que as colocações são combinações recorrentes na língua, remetendo à noção de coocorrência frequente, segundo Williams (2003, p. 43), seu dicionário - o *BBI Dictionary of English Word Combinations* -, por preceder os grandes *corpora* de referência, inclui as combinações de forma puramente intuitiva. Essa opinião também é compartilhada por Fontenelle (2003, p. 82).

estabelecidas entre os membros da combinação, como Cowie (1981), Seco (1987) e Hausmann (1989). Essa distinção resulta na inclusão de combinações muito diferentes em cada obra lexicográfica.

2.3 Síntese dos pressupostos teóricos

Em uma tentativa de homogeneizar a terminologia utilizada pelos diferentes autores no tratamento da combinatória léxica, estabelecemos cinco parâmetros que pudessem nos auxiliar na avaliação das diferentes concepções do conceito de “colocação”. Considerando, então, que uma colocação consiste em uma relação entre dois elementos, os parâmetros propostos visam a avaliar como cada teoria caracteriza essa relação. Dessa forma, os parâmetros arrolados pretendem identificar se a relação entre os membros de uma colocação:

1. obedece ou não a uma relação de conteúdo;
2. corresponde ou não a uma relação sintagmática;
3. obedece a uma relação de constelação⁹² ou de subordinação;
4. inscreve-se em diferentes níveis de organização de uma língua natural, à luz de uma concepção Coseriana da linguagem;
5. está relacionada ou não a um padrão de ocorrência reiterada.

A partir desses parâmetros, derivam os cinco critérios utilizados para caracterizar as relações estabelecidas entre as bases e os colocados, de acordo com as propostas dos diferentes autores estudados:

1. semântico: significado transparente das colocações;
2. sintático: restrição (variável) entre os elementos;
3. lexical: *status* entre os elementos (igual ou diferente);
4. categorial⁹³: localização das colocações no diassistema, mais especificamente na norma;

⁹² Segundo Alarcos Llorach (1972, p. 29), na relação de constelação “a dependência entre dois termos é mais livre, de modo que têm certa conexão, mas não se supõem mutuamente nem unilateralmente” [la dependencia entre dos términos es más libre, de suerte que tienen cierta conexión, pero no se suponen mutuamente ni unilateralmente].

⁹³ Designação baseada na concepção de Coseriu (1980), para quem a linguagem se divide em três níveis de estruturação: universal, histórico e individual.

5. estatístico: frequência de coaparição relevante ou não para a determinação das colocações.

Aplicando tais critérios às teorias analisadas, foi possível identificar, a partir de uma terminologia mais homogênea, a caracterização que cada autor confere ao fenômeno léxico das colocações. A fim, então, de facilitar a visualização dessas informações, apresentamos, a seguir, um quadro que sintetiza as concepções de cada autor, dispondo-os de acordo com a ordem em que aparecem neste capítulo.

Autores \ Critérios	significado transparente	restrição entre os elementos	diferente <i>status</i> entre os elementos	localização na norma	frequência de coaparição
Bally	∅	✓	✗	✓	✓
Porzig	∅	✓	∅	✓	∅
Coseriu	∅	∅	✓	∅	✗
Mel'čuk	✓	✓	✓	✓	✓
Corpas Pastor	✓	✓	✓	✓	✓
Firth	∅	✓	∅	∅	✓
Halliday	∅	✓	∅	∅	✓
Sinclair	∅	✓	✗	✓	✓
Tagnin	∅	✓	✓	∅	✓
Cowie	✓	✓	✗	∅	∅
Haensch et al	✓	✓	∅	∅	✓
Benson	✓	✓	✗	✓	✓
Seco	∅	∅	∅	∅	∅
Hausmann	✓	✓	✓	✓	∅

Legenda: ✓ = sim; ✗ = não; ∅ = não consta na teoria

Quadro 1: Síntese dos pressupostos teóricos

2.4 Delimitação do objeto de estudo

A partir da homogeneização da terminologia utilizada pelos autores, foi possível perceber que algumas características apresentadas para as colocações são recorrentes nas diversas teorias⁹⁴. São elas:

⁹⁴ Essa constatação contraria a idéia de Nesselhauf (2005, p. 11), para quem “o único denominador comum é que o termo [sc. colocação] é (ao menos na maioria das vezes) usado para se referir a algum tipo de relação sintagmática de palavras” [the only common denominator is that the term is (at least mostly) used to refer to some kind of syntagmatic relation of words].

- a) situadas na norma, ou seja, correspondem a estruturas pré-fabricadas à disposição dos falantes, as quais aparecem entre as combinações livres (situadas na fala) e as fixas (também localizadas na norma);
- b) restrição variável entre os elementos, o que, em alguns casos, possibilita a substituição de um deles por um grupo restrito de vocábulos;
- c) freqüência de coaparição das unidades lexicais;
- d) diferente *status* entre os elementos, em que um determina a coocorrência do outro;
- e) significado transparente.

Porém, apesar de haver um consenso mínimo entre os teóricos na definição do fenômeno das colocações, no momento de delimitar essas combinações, os exemplos apresentados são muito divergentes. Dessa forma, Corpas Pastor (1996) considera que a noção desenvolvida por Porzig (1970) de “campos verbais” é mais ampla que a de colocação, ao passo que a de “solidariedade léxica”, elaborada por Coseriu (1977), tem um alcance mais limitado. Ambos os autores, porém, tratam de explicar as relações semânticas estabelecidas entre os vocábulos como um fato ontológico da linguagem, que pode ser realizado em qualquer língua particular. Por outro lado, Sinclair (1991) e os demais autores que compartilham sua abordagem, ao utilizarem cálculos estatísticos para identificar as palavras que coocorrem no discurso, enfocam o problema das colocações como um fato de uma língua particular, ou seja, como uma recorrência de palavras na língua que se está analisando. Em virtude desse fato, surgem as críticas de Bosque (2001, p. 15) aos estudos desenvolvidos a partir de uma abordagem estatística, que, por utilizarem somente o critério da freqüência para o reconhecimento das colocações, não realizam um estudo propriamente lingüístico desse fenômeno léxico. Além disso, a coocorrência freqüente faz com que muitas combinações livres sejam tratadas por Sinclair, em seus dicionários, como colocações.

Essas constatações evidenciam que algumas divergências apresentadas pelos diversos estudos, mais do que atreladas a diferentes metodologias, parecem estar relacionadas ao tratamento das colocações como fato de uma língua particular (como os autores da abordagem estatística), enquanto outros as vêem como um fato ontológico da linguagem (como na abordagem semântica). Embora um enfoque não anule o outro, as diferentes perspectivas resultam em objetos de estudo diferenciados.

Em virtude, então, das diversas perspectivas adotadas para o estudo das colocações, uma multiplicidade de critérios foi elaborada para sua identificação. No entanto, foi possível verificar uma convergência parcial entre eles, através de nossa tentativa de

homogeneizar a terminologia, o que resultou em cinco características principais. Essa coincidência existente entre as diferentes abordagens nos critérios por nós arrolados, por ser apenas parcial, não permite a seleção de uma única metodologia como a que ofereça melhores subsídios para um reconhecimento efetivo das colocações. Tal fato levou-nos a realizar um estudo que recuperasse as propostas das diferentes abordagens⁹⁵. Essa opção justifica-se pelo fato de não havermos encontrado, na literatura especializada, nenhum trabalho que realizasse uma análise das colocações relacionando os diferentes enfoques. Além disso, as abordagens, ainda que divergentes, não se contradizem, o que permite sua utilização em conjunto.

Dessa forma, seguindo os pressupostos da abordagem estatística, iniciamos com a realização de uma análise quantitativa das colocações formadas somente por substantivos e adjetivos (v. capítulo 4). Essa opção de estudar apenas tais classes de palavras deriva da necessidade de obter um conjunto homogêneo de combinações que possibilite uma investigação mais aprofundada. A partir de um *corpus* textual, identificamos, então, as combinações que apresentavam uma coocorrência freqüente. No entanto, segundo Fontenelle (2003, p. 86-67) e Pazos Breña (2005, p. 98), a análise estatística das combinações constitui apenas uma primeira etapa para a identificação das colocações, a qual deve se seguir uma interpretação dos dados e sua adaptação a um modelo, ou seja, um estudo lingüístico propriamente dito. Em virtude desse fato, então, utilizamos os pressupostos oferecidos pela abordagem semântica para realizar uma análise qualitativa das colocações (v. capítulo 5). Nesse momento, procuramos identificar as relações estabelecidas entre os substantivos e os adjetivos, a fim de verificar, entre esses últimos, aqueles que apresentavam uma maior tendência a formar colocações. Finalmente, a partir dos pressupostos do enfoque lexicográfico, desenvolvemos uma análise contrastiva das colocações (v. capítulo 6), de forma a delimitar as combinações idiossincráticas.

⁹⁵ Segundo Gabrielatos (1994, p. 3), a discrepância entre as diversas teorias deve-se ao fato de que “nem todos os fatores são levados em conta por todos os analistas” [not all factors are taken into account by all analysts].

3 ELABORAÇÃO DO *CORPUS* DE ESTUDO

As propriedades que caracterizam as colocações, apresentadas no capítulo anterior, oferecem uma definição bastante vaga desse fenômeno léxico e não permitem um reconhecimento *a priori* das combinações de palavras que poderiam formar colocações. Em vista disso, a fim de estudar a maneira como adjetivos e substantivos costumam se combinar na língua, foi necessário realizar uma pesquisa lingüística baseada em *corpus*, pois somente com a utilização de uma grande quantidade de dados autênticos e de *softwares* que permitam identificar as concordâncias, é possível chegar a uma “conclusão confiável de que palavras ‘andam juntas’ com quais”⁹⁶ (SINGLETON, 2000, p. 52). Um *corpus*, segundo Pérez Hernández (2002, p. 24-25), consiste na compilação de um conjunto de textos escritos ou falados, utilizados para a realização de determinadas análises lingüísticas. Esses textos devem ser produzidos em situações reais de uso e sua inclusão no *corpus* deve ser orientada por critérios lingüísticos explícitos, de forma a assegurar que possam ser usados como uma amostra representativa de uma língua.

O conceito de representatividade de um *corpus* está diretamente relacionado a sua extensão; porém, como afirma Berber Sardinha (2004, p. 23), “não há critérios objetivos para a determinação da representatividade”, a qual vai depender das questões de pesquisa. Para Biber, Conrad, Reppen (2000, p. 248), por exemplo, um *corpus* que pretende representar a língua deve preocupar-se com a diversidade de textos que inclui, sendo estes textos de diferentes registros⁹⁷, dialetos e assuntos. Por isso, consideram que o tamanho do *corpus* é importante, mas não deve ser medido apenas em relação ao número de palavras, visto que a diversidade de textos e autores é também importante para garantir a variedade. Kennedy (1999, p. 66), no entanto, salienta que qualquer *corpus*, por maior que seja, será sempre apenas uma amostra da língua⁹⁸.

⁹⁶ [reliable conclusions about which word ‘keep company’ with which].

⁹⁷ O termo “registro” é utilizado por Biber, Conrad, Reppen (2000) na mesma acepção que, neste trabalho, caracterizamos o “gênero” (v. 3.1).

⁹⁸ Opinião semelhante apresenta Seco (2003, p. 89).

3.1 Reconhecimento e estruturação da área

Para a análise das combinações léxicas, selecionamos trabalhar com um *corpus* composto exclusivamente de textos jornalísticos, entendidos como o “material não publicitário, resultante do esforço editorial produzido por um corpo relativamente estabilizado de jornalistas e de colaboradores fixos e ocasionais” (KAUFFMANN, 2005, p. 3). Segundo Martínez Albertos (1983, p. 206-207), o texto jornalístico caracteriza-se por utilizar uma linguagem não-literária próxima da língua coloquial culta, além de ser conciso (predomínio de sintagmas nominais) e claro (uso de verbos na forma ativa com tempo indicativo). É fruto da construção de diferentes autores e seu objetivo é o de atrair a atenção do leitor desde as primeiras linhas.

Nossa escolha por trabalhar com esses textos deveu-se ao fato de os textos jornalísticos não corresponderem a um conjunto homogêneo, mas apresentarem diversidade de gêneros e de tipos textuais⁹⁹. Dessa forma, não constituem um tipo específico de textos dentro de uma tipologia textual¹⁰⁰, mas correspondem ao que Isenberg (1987, p. 101) chama de “classe de texto” [*clase de texto*] e define como

toda forma de texto, cujas características podem ser fixadas mediante a descrição de determinadas propriedades, não válidas para todos os textos, independentemente de se essas propriedades podem ser concebidas teoricamente e de que maneira dentro do marco de uma tipologia textual¹⁰¹.

Esse conceito será proposto em oposição ao de “tipo de texto” [*tipo de texto*], que corresponde a uma “designação teórica de uma forma específica de texto, descrita e definida dentro do marco de uma tipologia textual”¹⁰² (ISENBERG, 1987, p. 101).

Ainda que esse autor acredite que “uma tipologia lingüística do texto, satisfatória do ponto de vista teórico, continue sendo um *desideratum*”¹⁰³ (ISENBERG, 1987, p. 95),

⁹⁹ Um equilíbrio entre diferentes tipos e gêneros textuais é fundamental, segundo Čermak (2003, p. 21), para garantir a representatividade do *corpus*.

¹⁰⁰ Uma tipologia textual, segundo Bussmann (1990, s.v. *Texttypologie*), corresponde a uma classificação de textos no marco da Lingüística Textual. Essa classificação pode ser organizada segundo diferentes critérios, tais como: (a) função textual: texto utilitário, literário, informativo, etc.; (b) canal de comunicação ou o número e a familiaridade dos destinatários: texto escrito, oral, emissão de rádio, carta, etc.; (c) elementos estruturais ou de conteúdo como, por exemplo, o desenvolvimento temático: texto descritivo, narrativo, argumentativo, etc.; (d) combinações específicas dos critérios internos e externos, apresentados nos itens de (a) até (c), o que origina gêneros textuais *stricto sensu*, como um informe meteorológico, uma receita culinária, entre outros. Segundo o autor, até o momento, uma tipologia textual consistente e terminologicamente unitária não existe.

¹⁰¹ [toda forma de texto, cujas características pueden fijarse mediante la descripción de determinadas propiedades, no válidas para todos los textos, independentemente de si estas propiedades pueden concebirse teóricamente y de qué manera, dentro del marco de una tipología textual].

¹⁰² [designación teórica de una forma específica de texto, descrita y definida dentro del marco de una tipología textual].

ressalta que a taxonomia proposta por Werlich (1976) constitui a classificação mais homogênea e coerente. Para Werlich (1976 apud ISENBERG, 1987, p. 114), os tipos de textos são divididos em descrição, narração, exposição, argumentação e instrução, os quais correspondem aos cinco tipos textuais propostos por Adam (2001). Esse autor, ao elaborar uma classificação dos textos a partir de sua organização seqüencial prototípica, identifica os seguintes tipos textuais:

- **narração:** conjunto de proposições caracterizado pela sucessão de eventos, unidade temática, presença de uma ação e uma intriga e avaliação final (explícita ou implícita);
- **descrição:** enumeração dos atributos de um lugar, uma época, um ser animado ou inanimado, uma ação, etc.;
- **argumentação:** manifestação da opinião sobre determinado tema, na tentativa de conseguir a adesão do leitor;
- **explicação:** apesar de possuir uma base informativa, caracteriza-se pela vontade de fazer compreender um fenômeno;
- **diálogo:** caracteriza-se não somente pela presença de duas pessoas que falam sobre um tema, mas pelo fato de seus enunciados serem mutuamente determinados¹⁰⁴.

Com relação à diversidade de gêneros, entendidos, conforme Swales (2000, p. 58), como uma classe de eventos comunicativos, cujos membros compartilham os mesmos propósitos e cujos exemplares exibem diversas características semelhantes em termos de estrutura, estilo, conteúdo e público-alvo¹⁰⁵, Melo (1992, p. 16) afirma que não pode haver uma classificação universal dos gêneros justamente pelo fato de estarem relacionados ao modo de produção e à cultura. Em decorrência disso, diferentes classificações de gêneros

¹⁰³ [una tipología lingüística del texto, satisfactoria desde una perspectiva teórica, sigue siendo todavía un desideratum].

¹⁰⁴ Van Dijk (1996), ao tratar das superestruturas textuais, considera como tipos textuais a narração, a argumentação, o tratado científico e o diálogo. Posteriormente, porém, Van Dijk (1996, p. 167) apresenta uma relação de vinte diferentes tipos textuais, incluindo “conferência acadêmica”, “sermão”, “lei”, “notícia”, entre outros. Tal classificação corresponde, no presente trabalho, ao conceito de “gênero”. No entanto, o próprio autor percebe uma assimetria em sua classificação ao verificar que “uma grande quantidade deles [sc. tipos textuais] poderá ser incluída, por exemplo, na argumentação” (VAN DIJK, 1996, p. 167) [una gran cantidad de ellos se podrá incluir, por ejemplo, en la argumentación].

¹⁰⁵ [a class of communicative events, the members of which share some set of communicative purposes. (...) In addition to purpose, exemplars of a genre exhibit various patterns of similarity in terms of structure, style, content and intended audience]. Para outras noções desse conceito, v. Trask (2006, s.v. *gênero*).

foram propostas, na literatura especializada, para o texto jornalístico¹⁰⁶. Porém, segundo Glück (2000, s.v. *Textsorte*), apesar dos esforços da Lingüística e da Teoria do Texto por racionalizar os conhecimentos sobre os conjuntos de textos específicos, a elaboração de uma classificação mais abrangente dos gêneros textuais ainda não foi possível, devido à dificuldade de estabelecer traços bem diferenciados e constituídos nos diversos níveis da estrutura textual. Kauffmann (2005, p. 77-81), no entanto, trabalhando com os pressupostos da Lingüística de *Corpus*, consegue identificar treze gêneros diferentes dentro do texto jornalístico, a saber:

- **artigo**: texto opinativo e com autoria;
- **carta**: expressão do ponto de vista do leitor;
- **chamada**: texto curto na primeira página que remete o leitor ao texto completo;
- **coluna de notas**: coleção de pequenos textos em estilo mais livre;
- **comentário**: texto opinativo sobre fatos da atualidade;
- **crítica**: discussão fundamentada sobre uma manifestação artística;
- **crônica**: texto livre e pessoal, com teor literário, político, esportivo, etc.;
- **editorial**: texto opinativo que expressa o ponto de vista do jornal;
- **entrevista**: texto em forma de diálogo entre o repórter e o entrevistado;
- **nota de correção**: retificação de informações;
- **notícia**: relato objetivo de um fato;
- **reportagem**: relato mais extenso, abrangente e contextualizado, que traz informações mais detalhadas sobre as notícias, interpretando os fatos;
- **resenha**: resumo e crítica de livros.

Esse autor propõe, ainda, um grupo denominado “outros”, do qual fariam parte os gêneros não incluídos nos demais. Dessa forma, indica que a quantidade de gêneros presentes no texto jornalístico poderia ser ainda maior e confirma a dificuldade encontrada pela Lingüística Textual em oferecer uma classificação precisa para todos os diferentes gêneros textuais possíveis.

¹⁰⁶ No Brasil, os autores advindos do jornalismo que mais se dedicaram aos estudos sobre a classificação dos gêneros jornalísticos são Melo (1992) e Chaparro (2000). Bonini (2003), porém, ao analisar os gêneros jornalísticos sob um enfoque lingüístico, apresenta críticas às classificações promovidas por jornalistas, devido ao uso de conceitos de gênero já ultrapassados e uma preocupação em oferecer “rótulos”, mais do que critérios de seleção para os textos.

Martínez Albertos (1983, p. 212), por sua vez, propõe que existem diferentes estilos de escrita conforme os gêneros jornalísticos. Dessa forma, identifica que cada gênero relaciona-se com um dos quatro primeiros tipos textuais propostos por Adam (2001), quais sejam, narração, descrição, argumentação e explicação. Kauffmann (2005, p. 115) também salienta a relação entre os gêneros e os tipos textuais, mas percebe que, dentro de um mesmo gênero, podem ser utilizados diferentes tipos¹⁰⁷. No entanto, ambos os autores não incluem a forma de diálogo em suas classificações. Acreditamos, porém, que o texto dialogal também se faz presente no jornal, especialmente no gênero de entrevistas.

Consideramos, assim, que o texto jornalístico possa ser utilizado como uma amostra representativa da língua por sua diversidade tanto de gêneros quanto de tipos textuais, bem como pela variedade temática e de autoria. Além disso, devido à ampla circulação dos jornais - diários e impressos - um grande número de leitores é atingido¹⁰⁸, o que o torna um meio de comunicação importante na conformação da linguagem¹⁰⁹.

3.2 Constituição do *corpus*

Para o estudo das colocações, foram compiladas sessenta edições - durante dois meses - de diferentes jornais de língua espanhola, entre os dias 01/10/2006 e 29/11/2006¹¹⁰.

¹⁰⁷ Opinião semelhante apresenta Isenberg (1987, p. 115). Esse autor utiliza o trabalho de Werlich (1976), o qual, para o gênero jornalístico da reportagem, identificou três diferentes tipos textuais: narração, descrição e argumentação. Para Isenberg (1987, p. 116), porém, uma tipologia textual deve ser monotípica, de forma que, embora existam diferentes tipos textuais dentro de um mesmo gênero, um deles deve ser predominante.

¹⁰⁸ Para Kennedy (1999, p. 63), “individualmente, a fala tem uma maior proporção do que a linguagem escrita que a maioria de nós produz ou recebe em um dia típico. No entanto, um texto escrito (como um artigo de jornal) pode ser lido por 10 milhões de pessoas, enquanto um diálogo falado, envolvendo a compra de um par de sapatos, pode não ser ouvido nunca por pessoas que não os dois interlocutores originais”. [Individually, speech makes up a greater proportion than does writing of the language most of us receive or produce on a typical day. However, a written text (say, in a newspaper article) may be read by 10 million people, whereas a spoken dialogue involving the purchase of a pair of shoes may never be heard by any person other than the two original interlocutors].

¹⁰⁹ Salientamos que nossa dissertação não visa a elaborar um dicionário de colocações, mas apresentar critérios que permitam uma identificação mais precisa das combinações entre substantivos e adjetivos, a partir de um *corpus*, para seu posterior tratamento lexicográfico. Dessa forma, mais importante que os textos selecionados, é a caracterização das combinações léxicas recorrentes. Acreditamos, porém, que a escolha de textos com ampla diversidade, permitirá a obtenção de dados mais heterogêneos e, dessa forma, mais próximos aos utilizados em grandes *corpora* para fins lexicográficos.

¹¹⁰ Para nossa dissertação, não será utilizada a técnica da “semana construída” para a coleta dos textos, bastante utilizada em Linguística de *Corpus* para a elaboração de *corpus* de jornais. O princípio da “semana construída” é selecionar um menor número de exemplares de forma sistemática durante um longo período de tempo, a fim de evitar distorções na amostragem, ou seja, para que um determinado tema em voga no momento da coleta não se

Selecionamos os jornais de maior tiragem e circulação de dez países mais populosos de língua espanhola, quais sejam, Argentina, Bolívia, Chile, Colômbia, Equador, Espanha, México, Guatemala, Peru e Venezuela¹¹¹. Acreditamos que, utilizando os jornais de maior circulação, seja possível uma maior aproximação com a língua realmente utilizada pela maioria dos falantes, de forma a encontrar recorrências léxicas efetivamente empregadas. Tal fato demonstra-se relevante no momento de incluir as colocações em um dicionário.

Foram coletadas manualmente as versões eletrônicas dos diversos jornais, que correspondem fielmente à versão impressa; porém, para nossa pesquisa, excluímos as partes do jornal que não correspondem ao que definimos como “texto jornalístico”. São elas: tabelas, gráficos, mapas, legendas de fotos, textos-legenda (textos curtos que acompanham uma imagem), publicidade, propaganda legal (atas, avisos, etc.), roteiros (lista de estabelecimentos, atrações, etc.), horóscopo, obituário, expediente, classificados e previsão do tempo (KAUFFMANN, 2005, p. 50).

Assim, foram inseridos os textos das diferentes seções que constituem cada jornal selecionado. Para a Argentina, utilizamos o jornal *El Clarín* e extraímos os textos das matérias de capa do jornal e das seções “El País”, “El Mundo”, “Sociedad”, “La Ciudad”, “Policiales”, “Deportes” e “Espectáculos”, bem como os textos que chamamos de “opinião”, os quais incluem editoriais, cartas dos leitores e colunas com opiniões de diferentes colaboradores do jornal. Foram incluídos também os textos de “suplementos”, isto é, de cadernos sobre diferentes temáticas, que trataram sobre viagens e comportamento (edição de 01/10/2006), informática (edição de 11/10/2006), literatura, cidades e mundo rural (edição de 21/10/2006), moda e beleza (edição de 31/10/2006) e música e entretenimento (edição de 10/11/2006).

Como representante da Bolívia, utilizamos o jornal *El Diario*, cujos textos utilizados correspondem às matérias de capa e às seções “Política”, “Economía”, “Sociedad”, “Nacional”, “Cultural”, “El Deportivo”, “Internacional” e “Mundo Social”, além dos editoriais e das colunas assinadas. Foram também incluídas as revistas sobre ciência e computação (edição de 02/10/2006), mulher, saúde e beleza (edição de 12/10/2006), música e

sobreponha aos dados como um todo. Como nosso propósito é apenas o reconhecimento das combinações léxicas, quaisquer que sejam, esse princípio não se mostrou relevante.

¹¹¹ Nossos dados foram retirados do site <http://www.census.gov/cgi-bin/ipc/idbrank.pl>, onde pode ser visualizada a listagem de todos os países do mundo ordenados pelo número de habitantes. A ordem que apresentamos, porém, corresponde à utilizada para a coleta diária dos jornais, mais próxima ao ordenamento alfabético. Cabe ressaltar que, ainda que República Dominicana e Cuba apresentassem um maior número de habitantes que a Bolívia, esses países não foram incluídos em nossa pesquisa porque não dispunham de um jornal em versão eletrônica que pudesse ser utilizado para nosso *corpus*. As referências completas dos jornais encontram-se no apêndice A.

entretenimento (edição de 22/10/2006), cinema (edição de 01/11/2006), automóveis (edição de 11/11/2006) e história e educação (edição de 21/11/2006).

Para o Chile, selecionamos o jornal *El Mercurio*, composto pelas matérias de capa e das seções “Nacional”, “Claves”, “Internacional”, “Economía y Negocios”, “Deportes”, “Vida Social”, “Actividad Cultural”, “Espectáculos”, “Internet”, “Ciencia y Tecnología”, “Vida y Salud”, “Educación”, “Reportajes”, “Artes y Letras”, além de editoriais, colunas e cartas dos leitores. Apesar do grande número de seções, a maioria não apresenta matérias todos os dias. O jornal também possui revistas semanais sobre mulher, beleza e comportamento (edição de 03/10/2006), artistas, entretenimento e culinária (edição de 13/10/2006), agricultura (edição de 23/10/2006) e literatura e turismo (edição de 12/11/2006).

Com relação à Colômbia, foram incluídos os textos do jornal *El Tiempo* que correspondem às matérias de capa e às seções “Bogotá”, “Cultura”, “Deportes”, “Económicas”, “Internacional”, “Justicia”, “Nación”, “Política”, “Primer Plano”, “Televisión” e “Vida de Hoy”, bem como editoriais, colunas e cartas de leitores. Incluí ainda seções especiais sobre saúde (edição de 04/10/2006), literatura e agricultura (edição de 14/10/2006) e viagens (edição de 24/11/2006).

Para o Equador, utilizamos o jornal *El Universo* e coletamos os textos das seções “Política”, “Economía”, “Sucesos”, “Migración”, “El País”, “Internacionales”, “Deportes”, “El Gran Guayaquil”, “Vida”, “Religiosa”, “Educativas”, “En Escena” e “Un día como hoy”, além dos editoriais, colunas e cartas dos leitores e de uma seção temática sobre agropecuária (edição de 04/11/2006). As matérias de capa não foram utilizadas porque eram idênticas as apresentadas no interior do jornal. A edição de 15/10/2006 foi a de maior extensão dentro desse grupo, pois, além de ser dia de eleições presidenciais no país, contava com um caderno dominical - sobre comportamento - e um suplemento de temas variados (moda, viagens, internet, culinária, entre outros).

A Espanha contou com o jornal *El Mundo*¹¹², do qual foram coletadas as seções “España”, “Mundo”, “Ciencia”, “Economía”, “Motor”, “Deportes”, “Cultura”, “Toros”, “Comunicación”, “Última” e “Madrid”, além de editoriais, colunas e cartas dos leitores. Incluí, ainda, cadernos sobre imóveis (edição de 06/10/2006), personalidades, crônicas, economia e informática (edição de 05/11/2006), saúde (edição de 25/11/2006) e uma revista

¹¹² Ainda que o jornal *El País* tenha uma maior circulação na Espanha (435.238/dia exemplares entre junho/2005 e junho/2006), optamos por utilizar o jornal *El Mundo* (320.161/dia exemplares entre junho/2005 e junho/2006) em nosso *corpus*, pois somente este disponibilizava uma versão completa, via internet, da edição impressa. Os dados de circulação dos jornais foram obtidos através da OJD (*Oficina de Justificación de la Difusión*) e estão disponíveis em <http://www.ojd.es>.

de variedades para o público adolescente, presente nas edições de 06/10/2006, 16/10/2006, 26/10/2006 e 15/11/2006.

Para o México, foram coletadas as edições do *El Universal*. Foram incluídos textos da capa do jornal, bem como das seções “México”, “El Mundo”, “Los Estados”, “DF”, “Finanzas” “Deportes”, “Espectáculos”, “Cultura”, “Estilos”, além das colunas, editoriais e um caderno sobre automóveis da edição de 07/10/2006.

Na Guatemala, selecionamos o jornal *Prensa Libre* e incluímos os textos das seções “Nacional”, “Departamental”, “Económicas”, “Internacionales”, “Deportes”, “Cultura”, “Buena Vida”, além de colunas com textos opinativos. Na edição de 08/10/2006, o jornal também apresentava a “Revista D”, com textos dos mais variados assuntos, como novidades, história, viagens, cartas dos leitores, entre outros.

Para o Peru, utilizamos os textos de capa do jornal *El Comercio*, bem como das seções “Tema del día”, “Buenos Días”, “Política”, “Lima”, “Perú”, “Mundo”, “Economía”, “Vida y Futuro”, “Luces”, “Deporte Total”, “Sociales”, “Crónica”, “Análisis”, além de colunas e editoriais. Foi incluído, ainda, o suplemento “Día 1” (edição de 09/10/2006), que trata sobre política, economia, tecnologia, entre outros temas.

Finalmente, para a Venezuela, recolhemos os textos do jornal *El Universal*. Foram incluídas as seções “Nacional y Política”, “Economía”, “Internacional”, “Deportes”, “Caracas”, “Espectáculos”, “Qué Hay”, “Calidad de Vida”, “Universo Empresarial”, “Sociales”, bem como colunas assinadas e cartas dos leitores.

Com relação aos textos retirados de cadernos sobre uma temática especializada, acreditamos que, pelo fato de serem veiculados em um jornal de ampla divulgação, apresentam uma linguagem acessível ao leitor comum, sendo, portanto, um texto escrito pelo jornalista (um intermediador da comunicação) para o público-leigo. Assim, de acordo com Cabré (1993, p. 137), estes e os demais textos jornalísticos não constituiriam um *corpus* especializado, mas representariam a língua comum, pois, segundo a autora, um texto não pode ser considerado especializado quando:

do ponto de vista da temática (...) são matéria de conversa e intercâmbio lingüístico habitual em situações gerais da vida cotidiana, temas que definitivamente não requerem uma aprendizagem de vocabulário específico por parte dos falantes. Do ponto de vista dos interlocutores, não são sujeitos especializados todos os falantes de uma língua (...) Do ponto de vista das situações comunicativas, não são especializadas as situações comuns à maioria de membros de uma comunidade social. Do ponto de vista das funções, não são especializadas as comunicações de tipo informativo¹¹³.

¹¹³ [desde el punto de vista de la temática (...) son materia de conversación e intercambio lingüístico habitual en situaciones generales de la vida cotidiana, temas que en definitiva no requieren un aprendizaje de vocabulario

Como todas essas características podem ser aplicadas aos textos veiculados nos jornais, consideramos, então, que esse meio de comunicação não pode ser classificado como utilizador de uma linguagem especializada.

3.3 Caracterização do *corpus*

Conforme a tipologia proposta por Berber Sardinha (2004, p. 20-22), o *corpus* utilizado em nosso trabalho, que totaliza 1.139.621 palavras, poderia ser caracterizado como:

- de modo **escrito**;
- utilizando um recorte temporal **sincrônico**, visto que todos os textos correspondem a um mesmo período de tempo, e **contemporâneo**, pois os jornais utilizados representam uma amostra da língua falada atualmente;
- selecionado por **amostragem**, pois é composto por variedades textuais que pretendem ser uma amostra finita da linguagem como um todo;
- com conteúdo **geral** (em oposição a especializado), pois, como vimos, apresenta diversidade de gêneros e de tipos textuais, **monolíngüe** e com relativa **variação dialetal**¹¹⁴, pois contém somente textos originalmente escritos em língua espanhola, mas a partir de diferentes variantes segundo os países selecionados;
- com **pluralidade** de autoria e de **língua-nativa**, pois os textos são produzidos por diferentes autores falantes do espanhol;
- com finalidade de **estudo**, uma vez que constitui o *corpus* que se pretende descrever;

específico por parte de los hablantes. Desde el punto de vista de los interlocutores, no son sujetos especiales todos los hablantes de una lengua (...) Desde el punto de vista de las situaciones comunicativas, no son especiales las situaciones comunes a la mayoría de miembros de una comunidad social. Desde el punto de vista de las funciones, no son especializadas las comunicaciones de tipo informativo].

¹¹⁴ Consideramos a variação dialetal somente relativa, pois, embora nosso *corpus* seja composto por textos de diferentes países de língua espanhola, na língua culta – utilizada nos jornais – não existe uma variação considerável (cf. LIPSKI, 1994; MORENO DE ALBA, 1988), fato que também pôde ser constatado em nosso *corpus*. Segundo Lipski (1994, p. 168), “o espanhol da América apresenta numerosas características supra-regionais e uma norma de prestígio bem definida, embora não oficialmente válida para os dois continentes” [el español de América presenta numerosas características supra-regionales, y una norma de prestigio bien definida, aunque no oficialmente, válida para los dos continentes].

- com **integralidade**, pois foram incluídos textos completos (e não fragmentos), e **equilíbrio relativo**, visto que a distribuição das variedades no *corpus* reflete a distribuição dos gêneros em cada jornal.

3.4 Coleta e organização dos textos

Todos os textos utilizados em nossa pesquisa foram retirados dos *sites* dos diferentes jornais utilizados. A coleta foi realizada de forma completamente manual, ou seja, as matérias eram copiadas da internet e, posteriormente, salvas em arquivos de extensão *.txt*.

Cada edição dos jornais contou com um arquivo próprio, nomeado com as iniciais do jornal e um número, indicando a ordem em que foi recolhido. Por exemplo, a primeira edição que coletamos foi do jornal *El Clarín* e seu arquivo recebeu o nome “CLA01”. No segundo dia, utilizamos uma edição do jornal *El Diario*, cujo arquivo foi nomeado como “DIA01” e assim sucessivamente para as demais edições. Ao término da primeira coleta dos dez jornais, voltamos ao *El Clarín* e, então, salvamos as informações em um arquivo chamado “CLA02”. Essas siglas, que identificam cada edição dos jornais, serão utilizadas para acompanhar os exemplos extraídos de nosso *corpus* e aparecem detalhadas no apêndice A.

As informações sobre cada edição coletada foram armazenadas em uma base de dados no programa *Microsoft Access*. Nessa base, organizada a partir do código do arquivo de cada edição dos jornais, incluímos o nome de cada jornal, seu país de origem, a data da coleta, o número de *tokens* e de *types*¹¹⁵, o número de textos, informações sobre cadernos de temática especializada e o *site* em que cada edição encontra-se disponível (cf. apêndice A).

Devido à forma de coleta, não foi necessário que fizéssemos uma limpeza no material recolhido, pois todas as informações que não eram consideradas textos jornalísticos (conforme definido em 3.2) e que, portanto, não deveriam fazer parte do *corpus*, não foram

¹¹⁵ Os *tokens* correspondem, segundo Hartmann, James (1998, s.v. *token*), a cada ocorrência de uma palavra dentro de um *corpus*. Já os *types* representam cada uma das diferentes palavras que ocorrem em um *corpus* (HARTMANN, JAMES, 1998, s.v. *type*). No entanto, há uma diferença de concepção entre esses conceitos, conforme sejam enfocados do ponto de vista lexicográfico ou da Linguística de *Corpus* (LC). Assim, por exemplo, na frase “O carro que foi roubado era de meu irmão” existem 9 *tokens* e também 9 *types* para a LC, ao passo que para a Lexicografia há somente 8 *types*, pois as duas ocorrências do verbo *ser* (“era” e “foi”) contabilizam um único *type*. Como o número de *types* e *tokens* de nosso *corpus* é calculado a partir do *software WordSmith Tools*, o qual se baseia na LC, os dados que apresentamos correspondem a essa perspectiva.

sequer recolhidas. Apenas foram colocados entre “< >”¹¹⁶ os nomes dos redatores das matérias e as cidades correspondentes ao fato noticiado, os quais, muitas vezes, apareciam no início dos textos. Isso foi necessário para que, no momento do processamento dos dados, pudesse ser apresentada uma informação mais precisa acerca das colocações. Nosso *corpus* também contou com um sistema de anotação¹¹⁷ para a inserção de informações extralingüísticas, tais como a indicação dos títulos das diferentes seções de cada jornal. Essa marcação é importante porque facilita o acesso à informação, caso haja necessidade.

Finalmente, nosso *corpus* precisou também ser etiquetado¹¹⁸, para que pudéssemos acessar de forma mais precisa e agilizada as combinações formadas por substantivos e adjetivos, sem ter que percorrer toda a lista de 98.816 *types* presentes em nossos textos. A classificação morfossintática dos vocábulos foi feita por um processo automático de etiquetagem para o espanhol via internet, o VISL (*Visual Interactive Syntax Learning*), disponível em: <http://visl.hum.ou.dk/visl/pt/>. Esse recurso possui um índice de confiabilidade de 99% em termos de classes de palavras (cf. BICK, 2005, p. 100)¹¹⁹.

Após essa etapa, utilizamos uma rotina informatizada que permitiu isolar somente os vocábulos que correspondiam aos adjetivos, isto é, foi possível gerar uma lista contendo somente as palavras que apareciam seguidas de “ADJ” pelo etiquetador. Essa primeira relação, que continha 14.427 adjetivos, precisou, ainda, sofrer uma “limpeza” manual, para que fossem excluídas palavras que o programa considerou, de forma equivocada, adjetivos, tais como nomes, sobrenomes ou, até mesmo, erros de digitação. Como exemplo de nome etiquetado equivocadamente como adjetivo, temos “Patricio”, presente na frase *El boxeador Patricio Calero disputa hoy la primera de las cuatro finales a las que clasificó* (UNIE05)¹²⁰ [O boxeador Patricio Calero disputa hoje a primeira das quatro finais para as quais se

¹¹⁶ Essa marcação gráfica permite que os *softwares* desenvolvidos para o processamento dos dados, como o *WordSmith Tools* por nós utilizado, não considerem em suas análises as expressões que se encontram entre esses marcadores.

¹¹⁷ A anotação corresponde a comentários feitos em uma parte do texto (HARTMANN, JAMES, 1998, s.v. *annotation*). Segundo Berber Sardinha (2004, p. 145), esses comentários são inseridos manualmente por meio da observação do conteúdo de cada texto. Contêm informações não veiculadas verbalmente nos textos, mas fornecem detalhes sobre sua constituição, tais como proveniência, tipologia, autoria, etc.

¹¹⁸ A etiquetagem constitui um importante processo na preparação de corpora textuais e corresponde a uma análise das características gramaticais ou semânticas de palavras e frases (cf. HARTMANN, JAMES, 1998, s.v. *tagging*). Esse processo, segundo Berber Sardinha (2004, p. 150), permite o tratamento de grandes quantidades de textos rapidamente, além de auxiliar na desambiguação lexical e na descrição de padrões léxico-gramaticais.

¹¹⁹ Das 13 etiquetas para classes de palavras presentes no VISL, selecionamos para nosso trabalho somente os vocábulos etiquetados como ADJ, correspondente aos adjetivos. Uma lista de todas as etiquetas morfossintáticas oferecidas pelo programa, extraídas integralmente de <http://visl.sdu.dk/visl/pt/info/portsymbol.html>, pode ser vista no anexo B. Em seguida, a fim de ilustrar o funcionamento do etiquetador, apresentamos, no anexo C, um pequeno texto gerado pelo processo de etiquetagem.

¹²⁰ Esta e as demais siglas que acompanham os exemplos correspondem aos textos de nosso *corpus* dos quais extraímos as exemplificações. A referência completa encontra-se no apêndice A.

classificou], em que o programa considerou o nome “Patricio” como um qualificador do substantivo “boxeador”. Após, então, a remoção desse tipo de vocábulos, obtivemos uma lista com um total de 13.429 adjetivos, a partir dos quais iniciamos nossas buscas de combinações com os substantivos, conforme apresentamos no capítulo seguinte.

4 ANÁLISE QUANTITATIVA: RELAÇÕES DE FREQUÊNCIA

Para proceder à identificação das combinações léxicas entre adjetivos e substantivos em nosso *corpus*, utilizamos o *software WordSmith Tools*, versão 3.0¹²¹. A utilização dessa ferramenta de análise foi importante porque, além de oferecer uma série de recursos integrados que facilitaram nossa pesquisa, permitiu a identificação das possíveis colocações¹²² de forma mais agilizada.

4.1 Coleta dos dados

Para a recuperação dos adjetivos “candidatos” a colocações, foi necessário, primeiramente, a geração de uma lista de frequência das palavras de nosso *corpus*, através da ferramenta *WordList*, do *software WordSmith Tools*. A partir disso, foi possível verificar a frequência de todos os adjetivos que haviam sido identificados pelo etiquetador VISL e excluir de nossa pesquisa aqueles cuja ocorrência correspondia a apenas uma aparição. Considerando que as colocações correspondem a combinações de coocorrência frequente entre os vocábulos (cf. FIRTH, 1957; HALLIDAY, 1961; JONES, SINCLAIR, 1974; HAENSCH et al, 1982; CORPAS PASTOR, 2001; TAGNIN, 2005a), acreditamos que esse procedimento seja importante para que não sejam selecionadas estruturas utilizadas somente por um falante. Salientamos, no entanto, que existe a possibilidade de algumas combinações que realmente constituíam casos de colocações, mas que não eram expressivas em nosso *corpus*, não terem sido identificadas em nossa pesquisa.

Após, então, a exclusão dos vocábulos de frequência 1¹²³, passamos a verificar os substantivos que se combinavam com maior frequência com cada adjetivo. Para isso, geramos

¹²¹ Esse programa, desenvolvido por Mike Scott, da Universidade de Liverpool, está disponível para *download* em sua versão de demonstração no *site* <http://www.lexically.net/wordsmith>.

¹²² Chamamos de “possíveis colocações” porque o *software* oferece as recorrências léxicas mais frequentes. Essas combinações apenas serão consideradas colocações propriamente ditas se também apresentarem as características tratadas nos capítulos 5 e 6 do presente trabalho. A coocorrência frequente constitui, em nossa opinião, apenas um primeiro indicativo de que uma combinação léxica pode ser uma colocação.

¹²³ Jones, Sinclair (1974, p. 18) descartam de sua pesquisa palavras com uma frequência inferior a 10 ocorrências, devido a dificuldade de estudar todas as colocações a partir das palavras de baixa frequência em um amplo *corpus*. No entanto, acreditamos que se um vocábulo aparece poucas vezes, mas em todas as suas

os *Collocate* de cada vocábulo, também a partir do *software WordSmith Tools*. Essa ferramenta permite a visualização da quantidade de vezes que um vocábulo aparece à esquerda e à direita de cada adjetivo selecionado. Porém, é importante ressaltar que a noção de *Collocate* ou, em língua portuguesa, “colocados” empregada pelo *WordSmith Tools* relaciona-se ao conceito de “colocações” utilizado em Linguística de *Corpus*. Dentro desse enfoque, as colocações constituem “associações entre itens lexicais, ou entre o léxico e campos semânticos” (BERBER SARDINHA, 2004, p. 40). Dessa forma, são indicados como colocados de determinada palavra de busca todos os vocábulos que aparecem no mesmo contexto a uma distância previamente estipulada pelo pesquisador. Para nosso estudo, configuramos os *Collocate* de forma que apresentassem os vocábulos localizados a uma distância de quatro espaços a cada lado do adjetivo¹²⁴. Assim, foi possível a recuperação de combinações com adjetivos que não apareciam imediatamente unidos aos substantivos no texto, como no exemplo *país tropical muy pobre* (MER05) [país tropical muito pobre], em que entre o adjetivo *pobre* e o substantivo ao qual ele se refere aparecem outros dois vocábulos inseridos. A figura a seguir ilustra os resultados apresentados pelo *Collocate* a partir do adjetivo *emergente(s)* [emergente(s)].

ocorrências está relacionado a outro específico, seu estudo também será relevante para a identificação de uma relação colocacional entre eles.

¹²⁴ O programa permite visualizar combinações que variam de 2 a 8 palavras. Corpas Pastor (1996, p. 78-79) indica que não há uma distância específica entre os membros de uma combinação para que ela seja considerada uma colocação. A distância utilizada pelos diferentes estudos apresentados por essa autora variou entre uma e quinze posições entre as palavras. É importante salientar, no entanto, que as combinações que apresentam uma maior distância entre os elementos correspondem a construções formadas por verbos e substantivos, geralmente em frases com inversões e intercalações.

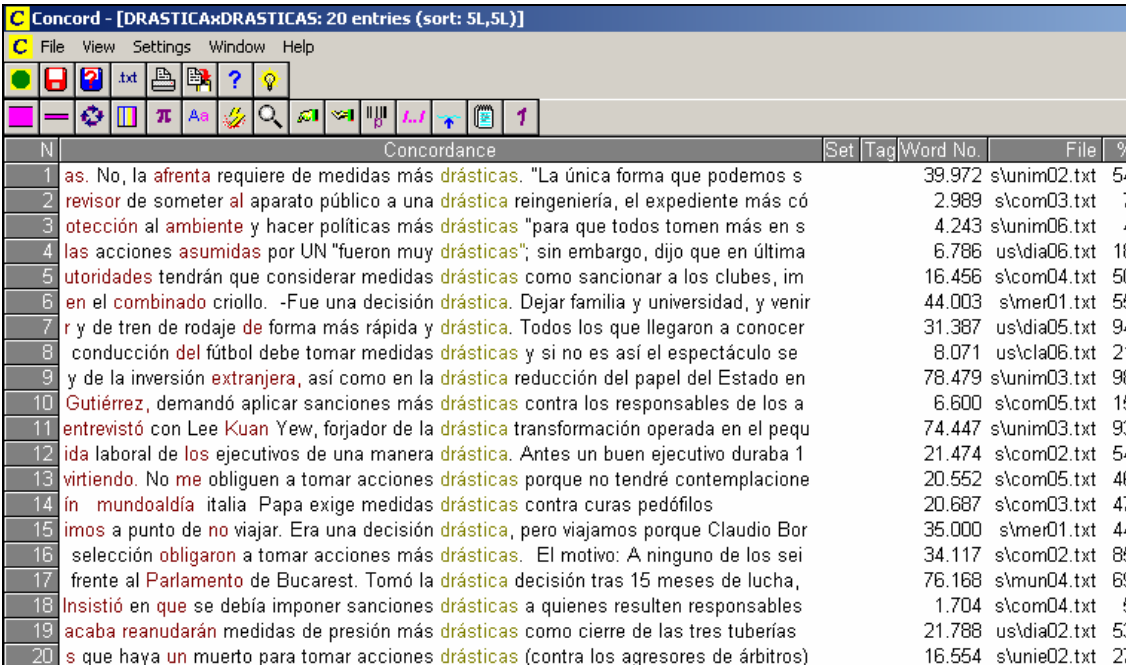
Concord - [collocates (total)]													
File View Settings Window Help													
[Icons: Green circle, Red X, Question mark, .txt, Print, Copy, Paste, Help, Lightbulb]													
[Icons: Refresh, Vertical bars, Aa, =, Eraser, Redaction, Magnifying glass]													
N	WORD	TOTAL	LEFT	RIGHT	L4	L3	L2	L1	*	R1	R2	R3	R4
1	EMERGENTES	43	0	0	0	0	0	0	43	0	0	0	0
2	DE	41	25	16	2	13	9	1	0	1	4	4	7
3	LOS	24	19	5	2	2	14	1	0	2	2	0	1
4	LA	20	7	13	3	2	2	0	0	6	1	1	5
5	EMERGENTE	19	0	0	0	0	0	0	19	0	0	0	0
6	EN	19	11	8	3	6	2	0	0	3	0	3	2
7	LAS	14	9	5	1	0	8	0	0	0	2	3	0
8	MERCADOS	12	10	2	0	0	0	10	0	0	2	0	0
9	Y	10	1	9	0	0	1	0	0	0	2	3	4
10	ECONOMÍAS	8	8	0	0	0	0	8	0	0	0	0	0
11	EL	8	2	6	0	0	1	1	0	1	1	2	2
12	QUE	8	3	5	1	2	0	0	0	0	1	1	3
13	COMO	7	2	5	0	1	0	1	0	1	1	1	2
14	UN	7	4	3	1	1	2	0	0	1	0	1	1
15	PAÍSES	6	6	0	0	0	0	6	0	0	0	0	0
16	A	5	1	4	0	1	0	0	0	0	1	2	1
17	CON	5	2	3	0	2	0	0	0	1	1	1	0
18	PAÍS	5	4	1	1	0	0	3	0	1	0	0	0
19	PARA	5	3	2	1	2	0	0	0	0	0	0	2
20	POR	5	2	3	2	0	0	0	0	1	1	0	1
21	AL	4	2	2	1	0	0	1	0	0	2	0	0
22	SE	4	1	3	1	0	0	0	0	0	1	1	1
23	?CUÁL	4	0	4	0	0	0	0	0	1	1	1	1
24	ACCIONES	3	3	0	0	2	1	0	0	0	0	0	0
25	DEL	3	2	1	1	1	0	0	0	0	1	0	0
26	MÁS	3	1	2	0	0	0	1	0	0	1	1	0
27	SUS	3	0	3	0	0	0	0	0	1	1	0	1
28	UNA	3	3	0	0	0	3	0	0	0	0	0	0
29	ASIA	2	0	2	0	0	0	0	0	0	0	2	0
30	BAJOS	2	0	2	0	0	0	0	0	0	0	2	0
31	BONOS	2	1	1	0	1	0	0	0	0	0	1	0
32	CHINA	2	0	2	0	0	0	0	0	1	1	0	0
33	CONDICIÓN	2	1	1	0	1	0	0	0	0	0	1	0
34	CONSOLIDADOS	2	0	2	0	0	0	0	0	0	0	1	1
35	CONTAR	2	1	1	1	0	0	0	0	0	1	0	0
36	CRECIMIENTO	2	2	0	2	0	0	0	0	0	0	0	0

Figura 1: Exemplo de aplicação da ferramenta *Collocate*

É importante reiterar que os *Collocate* gerados pelo *software* não correspondem exatamente à nossa definição de colocação. Muitas das combinações oferecidas, como entre o adjetivo *emergente* e artigos ou preposições, bastante frequentes na figura 1, não constituem casos de colocações. Além disso, até mesmo alguns substantivos, como *crecimiento* [crescimento] - linha 36 na figura 1 -, não apresentam uma relação colocacional com o adjetivo de busca, conforme pode ser observado no exemplo: *La falta de inversión y de*

crecimiento en varias economías emergentes reside en sus estrategias económicas (UNIM04) [A falta de investimento e de crescimento em várias economias emergentes reside em suas estratégias econômicas].

Em virtude desse fato, foi necessário gerar as concordâncias de alguns adjetivos, através da ferramenta *Concord*, do *WordSmith Tools*, a fim de verificar se todas as ocorrências dos substantivos arroladas pelo *Collocate* referiam-se exatamente ao nosso adjetivo de busca. Tal procedimento evitou que fossem considerados como candidatos a colocação substantivos que, embora aparecessem próximos aos adjetivos, não mantivessem uma relação direta com ele, como em: *Éste [sc. o ministro Tarso Genro] ha intentado permanentemente mantener el partido fiel al frente de izquierda que se formó en 1998 (...)* (UNIM03) [Este tem tentado permanentemente manter o partido fiel à frente de esquerda que se formou em 1998], em que o adjetivo *fiel* forma uma combinação com o verbo *mantener* e não com o substantivo *partido*. A seguir exemplificamos o uso da ferramenta *Concord* para o vocábulo *drástica(s)* [drástica(s)].



The screenshot shows the Concord software interface with a concordance table for the word 'drásticas'. The table has columns for 'N', 'Concordance', 'Set', 'Tag', 'Word No.', 'File', and '%'. The concordance entries are numbered 1 through 20 and show various contexts where the word 'drásticas' is used.

N	Concordance	Set	Tag	Word No.	File	%
1	as. No, la afrenta requiere de medidas más drásticas. "La única forma que podemos s			39.972	s\unim02.txt	54
2	revisor de someter al aparato público a una drástica reingeniería, el expediente más có			2.989	s\com03.txt	7
3	otección al ambiente y hacer políticas más drásticas "para que todos tomen más en s			4.243	s\unim06.txt	4
4	las acciones asumidas por UN "fueron muy drásticas"; sin embargo, dijo que en última			6.786	us\dia06.txt	18
5	utoridades tendrán que considerar medidas drásticas como sancionar a los clubes, im			16.456	s\com04.txt	50
6	en el combinado criollo. -Fue una decisión drástica. Dejar familia y universidad, y venir			44.003	s\mer01.txt	55
7	r y de tren de rodaje de forma más rápida y drástica. Todos los que llegaron a conocer			31.387	us\dia05.txt	94
8	conducción del fútbol debe tomar medidas drásticas y si no es así el espectáculo se			8.071	us\cla06.txt	21
9	y de la inversión extranjera, así como en la drástica reducción del papel del Estado en			78.479	s\unim03.txt	98
10	Gutiérrez, demandó aplicar sanciones más drásticas contra los responsables de los a			6.600	s\com05.txt	15
11	entrevistó con Lee Kuan Yew, forjador de la drástica transformación operada en el pequ			74.447	s\unim03.txt	93
12	ida laboral de los ejecutivos de una manera drástica. Antes un buen ejecutivo duraba 1			21.474	s\com02.txt	54
13	virtiendo. No me obliguen a tomar acciones drásticas porque no tendré contemplacione			20.552	s\com05.txt	46
14	fn mundoaldía italia Papa exige medidas drásticas contra curas pedófilos			20.687	s\com03.txt	47
15	imos a punto de no viajar. Era una decisión drástica, pero viajamos porque Claudio Bor			35.000	s\mer01.txt	44
16	selección obligaron a tomar acciones más drásticas. El motivo: A ninguno de los sei			34.117	s\com02.txt	85
17	frente al Parlamento de Bucarest. Tomó la drástica decisión tras 15 meses de lucha,			76.168	s\mun04.txt	69
18	Insistió en que se debía imponer sanciones drásticas a quienes resulten responsables			1.704	s\com04.txt	5
19	acaba reanudarán medidas de presión más drásticas como cierre de las tres tuberías			21.788	us\dia02.txt	53
20	s que haya un muerto para tomar acciones drásticas (contra los agresores de árbitros)			16.554	s\unie02.txt	27

Figura 2: Exemplo de aplicação da ferramenta *Concord*

Conforme pode ser observado na figura anterior, para realizar a consulta das concordâncias, agrupamos as formas de singular e plural dos vocábulos analisados, a fim de que todas as suas combinações aparecessem unidas. O uso do *Concord* possibilitou, ainda, que não fossem consideradas as combinações que ocorriam uma única vez no *corpus*, bem

como aquelas que apareciam somente na mesma edição de um jornal. Essa decisão permitiu a exclusão de combinações que representavam construções próprias de um falante, como no exemplo de *rostros siniestros* (CLA01) [rostos sinistros], que aparecia 2 vezes no jornal *El Clarín* do dia 01/10/2006, mas cuja frequência no buscador *Google* correspondia a apenas 74 ocorrências. Cabe ressaltar, porém, que, apesar de não serem expressivas em nosso *corpus*, algumas dessas combinações excluídas poderiam configurar casos de colocações.

Após a realização de todos esses procedimentos, conseguimos isolar 2.263 combinações formadas por 1.081 adjetivos¹²⁵ unidos com diferentes substantivos.

4.2 Avaliação dos resultados

A verificação da frequência de coaparição das combinações de substantivos e adjetivos é um procedimento importante para a identificação das colocações. No entanto, a maior parte das combinações que obtivemos através desse método correspondia a estruturas livres, formadas por um adjetivo de alta frequência que apresentava a propriedade de se combinar com uma grande quantidade de substantivos. Isso acontece, por exemplo, com o adjetivo *enorme* [enorme], e sua forma no plural *enormes* [enormes], que possui 265 ocorrências e pode ser combinado com um número muito grande de substantivos para denotar sua “grandeza”. Dessa forma, tal adjetivo aparece 8 vezes com *responsabilidad* [responsabilidade] e *cantidad(es)* [quantidade(s)], além de se combinar 5 vezes com *poder* [poder], *capacidad* [capacidade], *desafío(s)* [desafio(s)] e *potencial* [potencial], entre outros. Em virtude desse fato, sentimos a necessidade de utilizar um cálculo estatístico que permitisse realizar um recorte nas combinações identificadas em nosso *corpus*, a fim de que pudessemos eliminar a maior parte dos casos que configuravam combinações livres.

Jones, Sinclair (1974, p. 19), ao considerarem que há uma colocação entre dois elementos quando “eles coocorrem mais frequentemente que suas respectivas frequências e o tamanho do texto em que eles aparecem poderiam prever”¹²⁶, sugerem o uso de testes estatísticos para identificar os casos em que a coocorrência entre dois vocábulos é

¹²⁵ Para a contabilidade dos adjetivos, agrupamos suas formas de masculino, feminino, singular e plural, as quais apareciam como ocorrências diferentes em nossa lista de adjetivos original.

¹²⁶ [they co-occur more often than their respective frequencies and the length of text in which they appear would predict].

significativa. Para isso, utilizam o “Teste Exato de Fisher” [*Fisher’s Exact Test*], cálculo desenvolvido por Ronald Fisher que permite identificar se há uma associação não-randômica entre duas categorias variáveis. Esse teste calcula, então, a probabilidade de associação das características que estão em análise, ou seja, a probabilidade de tais características serem independentes ou estarem associadas. No entanto, seu uso é indicado somente quando o número total de dados é pequeno. Cabe ressaltar que Jones, Sinclair (1974), apesar de explicarem o uso do “Teste Exato de Fisher” para a identificação de colocações, não desenvolvem suas pesquisas a partir desse teste.

Com o auxílio do NAE (Núcleo de Assessoria Estatística), projeto de extensão desenvolvido no Departamento de Estatística do Instituto de Matemática da UFRGS, aplicamos o “Teste Exato de Fisher” a dois vocábulos de nosso *corpus*, quais sejam, *cuervo* [corpo] e *desnudo* [nu], o que resultou na seguinte tabela cruzada:

		<i>DESNUDO</i>		Total
		Ocorre	Não-ocorre	
<i>CUERPO</i>	Ocorre	7	505	512
	Não-ocorre	43	3.139.072	3.139.108
Total		50	3.139.577	3.139.621

Tabela 1: Aplicação do Teste Exato de Fisher às variáveis *cuervo* e *desnudo*

As variáveis dessa tabela podem ser explicadas a partir do seguinte modelo:

		<i>DESNUDO</i>		Total
		Ocorre	Não-ocorre	
<i>CUERPO</i>	Ocorre	a	b	a+b
	Não-ocorre	c	d	c+d
Total		a+c	b+d	n

Tabela 2: Tabela-modelo do Teste Exato de Fisher

Nessa tabela, a variável “a” corresponde ao número de vezes que as palavras *cuervo* e *desnudo* apareceram associadas, a variável “b” representa o número de vezes que *cuervo* foi utilizado em nosso *corpus* sem a presença de *desnudo*. Por outro lado, a variável “c” corresponde ao número de vezes que *desnudo* apareceu sem referência a *cuervo*. Finalmente, a variável “d” significa o número total de palavras que não correspondem nem a *cuervo* nem a *desnudo* em nosso *corpus*. Ao redor dessas variáveis, nas tabelas, encontram-se suas somas e, ao final, representado por “n”, aparece o número total de *tokens* de nosso *corpus*.

Para saber se a associação entre as variáveis é significativa, ou seja, se não é aleatória, aplica-se, sobre os dados da tabela, a seguinte fórmula:

$$p = \frac{(a + b)!(c + d)!(a + c)!(b + d)!}{n!a!b!c!d!}$$

Quando o valor da significância do teste é menor que 0,05 ($p < 0,05$), existe uma associação entre as duas variáveis. No caso de nosso exemplo entre *cuerpo* e *desnudo*, o valor do “Teste Exato de Fisher” apontou como resultado $p = 0,000$, o que indica que há associação entre as variáveis em análise. Porém, essa associação ocorre pelo fato de a não-ocorrência de ambas as variáveis ser muito grande em relação às outras frequências. Para realizar uma contraprova, reduzimos, então, o número de coocorrência dos vocábulos analisados de 7 para 1 e, ainda assim, o teste apresentou um valor de significância de $p = 0,007$, ou seja, mesmo que a combinação *cuerpo desnudo* tivesse ocorrido apenas 1 vez em nosso *corpus*, ainda assim seria significativa para os padrões do teste. O que define essa falsa significância, de acordo com o NAE, é a grande frequência de não-ocorrência de ambas as palavras, ou seja, como nosso *corpus* representa uma amostra grande para os padrões exigidos pelo “Teste Exato de Fisher”, o número de não-ocorrência das duas palavras é muito elevado. Isso faz com que qualquer coocorrência de dois elementos seja considerada significativa. Em virtude desses fatos, tal teste não se mostra adequado para nossas necessidades.

Outros métodos estatísticos utilizados em Lingüística de *Corpus* para a identificação de colocações, como a “razão Observado/Esperado”, a “Informação Mútua” ou o “Score T”¹²⁷, também apresentam as mesmas limitações supracitadas. Segundo Dunker (2002, p. 523), apesar da utilidade dessas medidas, “a lista de possíveis colocados é completamente ruidosa e inconclusiva”¹²⁸. Berber Sardinha (2004, p. 202), por exemplo, utilizando o teste da “razão Observado/Esperado”, identifica que os vocábulos *como* e *exemplo* possuem uma “força de associação maior do que o esperado” e, conseqüentemente, sua “coocorrência não parece ser aleatória” (BERBER SARDINHA, 2004, p. 203). Isso ocorre porque os testes estatísticos desenvolvidos para o cálculo de colocações utilizam uma definição desse fenômeno somente como a coocorrência não-aleatória de duas unidades lexicais, sem considerar a relação semântica que possa haver entre os vocábulos.

¹²⁷ Para uma aplicação prática de cada um desses métodos, v. Berber Sardinha (2004, p. 201-205). Para o modelo de “informação mútua”, v. também Church, Hanks (1990).

¹²⁸ [the list of possible collocates are quite noisy and inconclusive].

Dunker (2002, p. 523) já salienta que “a ambição de solucionar o enigma colocacional algoritmicamente sem qualquer intervenção humana é um equívoco”¹²⁹. Em virtude desse fato, Koike (2000) procura aliar o cálculo da frequência de coaparição dos elementos a um estudo semântico das combinações. Esse autor não utiliza nenhum método estatístico sofisticado para a identificação das colocações, mas calcula apenas a porcentagem de coocorrência das duas unidades léxicas. Utilizando a metodologia de Koike (2000) para nossa combinação entre *cuervo* e *desnudo*, obtemos o seguinte resultado:

Número de ocorrências de *cuervo*: 505

Número de ocorrências de *desnudo*: 43

Coocorrência de *cuervo* e *desnudo*: 7

Porcentagem de coocorrência de *cuervo* com *desnudo*: $7/505 \times 100 = 1,4\%$

Porcentagem de coocorrência de *desnudo* com *cuervo*: $43/505 \times 100 = 16,28\%$

Segundo Koike (2000, p. 19), “quando a porcentagem de coocorrência de duas unidades léxicas é superior a 20%, pode haver uma coesão sintática ou léxica”¹³⁰, ou seja, a partir desse índice as combinações passam a ser consideradas colocações. Esse valor, no entanto, é arbitrariamente estipulado pelo autor, sem que haja um embasamento teórico que justifique tal opção. Jones, Sinclair (1974, p. 32) também salientam que “embora os testes de significância em si mesmos sejam objetivos, eles se apóiam em escolhas essencialmente arbitrárias feitas por seus usuários”¹³¹. Essa arbitrariedade consiste em:

1. determinar o nível de significância, ou seja, definir a partir de qual valor no teste de significância uma combinação léxica será considerada colocação;
2. determinar sobre quais palavras deve ser aplicado o teste de significância, isto é, qual deve ser o número mínimo de coocorrência entre os vocábulos para que sejam considerados “candidatos a colocação”;
3. determinar o número mínimo de ocorrência de cada vocábulo no *corpus* para, a partir de então, identificar sua coocorrência com outros elementos.

A ausência de uma medida teoricamente fundamentada e empiricamente atestada para a identificação de colocações a partir da coocorrência frequente dos vocábulos em um

¹²⁹ [the ambition to solve the collocational riddle algorithmically without any human intervention is mistaken].

¹³⁰ [cuando el porcentaje de coocurrencia de dos unidades léxicas es superior al 20%, puede haber una cohesión sintáctica o léxica].

¹³¹ [although significance tests in themselves are objective, they rest on essentially arbitrary choices made by their users].

corpus faz com que nossa análise quantitativa seja limitada. O cálculo da frequência funciona apenas como um indício de que a combinação pode ser uma colocação, mas não constitui um fator determinante. Em razão disso, não foi possível, estatisticamente, estabelecer uma medida precisa que permitisse eliminar os casos de coocorrência que configuravam combinações livres.

Em virtude desse fato, mantivemos em nossa análise todos os casos de coocorrência com frequência superior a 1 gerados pelo *WordSmith Tools* para os adjetivos de nosso *corpus*. Porém, para que fosse possível a realização de uma análise mais refinada dos dados (conforme será apresentado nos capítulos 5 e 6), utilizamos somente as cinco combinações mais frequentes para cada adjetivo. Essa opção demonstrou-se essencial devido ao grande volume de dados a serem analisados, os quais fugiriam ao escopo deste trabalho.

4.3 Limitações da análise quantitativa

Após a identificação das principais propriedades que caracterizam as colocações, as quais foram apresentadas em 2.4, percebemos que essas combinações apareciam fixadas na norma e consistiam na coocorrência frequente de duas unidades lexicais com diferente *status* e graus de restrição entre os elementos, os quais, em conjunto, apresentavam um significado transparente. Tais características, no entanto, não permitiam uma delimitação precisa das combinações que realmente conformavam colocações. Em vista disso, sentimos a necessidade de realizar uma análise quantitativa dessas combinações, a fim de identificar, através do cálculo da frequência, os casos de colocação. É importante salientar que, dentre nossos cinco critérios utilizados para caracterizar esse fenômeno léxico, o da frequência foi o mais recorrente¹³² entre os diferentes autores estudados.

Para desenvolver, então, nossa análise, recorreremos a diferentes métodos estatísticos apresentados na literatura especializada. No entanto, percebemos que todos os cálculos disponíveis para o reconhecimento das colocações resultavam também em combinações livres, uma vez que estavam fundamentados em uma concepção de colocação somente como coocorrência frequente, sem considerar as relações semânticas estabelecidas entre os elementos. Além disso, de acordo com informações obtidas junto ao NAE, todos os

¹³² Também Pöll (1996, p. 95) identifica que a frequência de coocorrência dos lexemas é a característica mais citada na literatura especializada sobre o tema das colocações.

métodos empregados utilizavam um índice arbitrário para definir a partir de qual valor a frequência de coaparição dos elementos seria relevante para determinar se uma combinação constituía ou não uma colocação.

Em vista disso, ainda segundo o NAE, não é possível, somente a partir de cálculos estatísticos, determinar um índice de significância para identificar as colocações, pois o valor a partir do qual uma combinação será relevante deve ser determinado pelos lingüistas. Como não dispúnhamos desse valor, nem de uma teoria que permitisse fundamentar um índice de significância para as colocações, a análise quantitativa demonstrou ser limitada às nossas necessidades.

Dessa forma, os resultados obtidos a partir dessa análise, quais sejam, uma relação de diferentes combinações formadas por substantivos e adjetivos que apresentavam uma frequência de coaparição superior a 1, serviram apenas como uma primeira etapa no processo de reconhecimento das colocações. A fim, então, de refinar esses dados, realizamos uma análise qualitativa dessas combinações, isto é, procuramos identificar as relações semânticas estabelecidas entre os adjetivos e substantivos coletados, conforme apresentamos no capítulo 5.

5 ANÁLISE QUALITATIVA: RELAÇÕES SEMÂNTICAS

Devido às limitações apresentadas pela abordagem estatística na identificação das colocações entre substantivos e adjetivos (v. 4.3), neste capítulo procuraremos verificar os tipos de relações estabelecidas entre esses elementos, pois, embora todo o adjetivo mantenha uma relação semântica com os substantivos que acompanham, nem todos parecem poder estabelecer uma relação colocacional. Dessa forma, será necessário identificar os tipos de adjetivos que apresentam essa tendência, a fim de verificar se existe alguma recorrência entre as combinações de substantivo e adjetivo que costumam formar colocações. Salientamos que não pretendemos criar uma regra que permita prever todos os casos de colocações, mas apenas oferecer um conjunto mais restrito de adjetivos que apresentem uma pré-disposição a participar dessas combinações léxicas.

Na literatura especializada sobre colocações, encontramos somente o trabalho de Koike (2000) com um tratamento específico das combinações entre substantivos e adjetivos. Esse autor realizou um estudo aprofundado das relações entre essas classes de palavras com vistas à identificação de colocações. A partir das classificações dos adjetivos propostas por Navas (1962), Lapesa (1975), Bartos (1978) e Demonte (1999)¹³³, Koike (2000, p. 152-153) propõe sua própria taxonomia, dividindo os adjetivos em:

- a) **relacionais**, que servem para classificar, como *francés* [francês], *católico* [católico], *carnívoro* [carnívoro], entre outros;
- b) **qualitativos**, que expressam qualidades tanto “físicas”, como *grande* [grande] e *caliente* [quente], quanto “espirituais”, como *generoso* [generoso] e *perezoso* [preguiçoso], e
- c) **situacionais**, que modificam o modo - *posible* [possível] -, a intensão - *propio* [próprio] -, a circunstância - *actual* [atual] - e o aspecto - *frecuente* [frequente] - como o substantivo se realiza.

Segundo o autor, somente os adjetivos do último grupo (c) não são capazes de constituir colocações. No entanto, não concordamos que as combinações formadas por adjetivos relacionais, como *iglesia católica* [igreja católica], *registro civil* [registro civil], *actividad económica* [atividade econômica], entre outras apresentadas por Koike (2000, p. 176), representem casos de colocações. Para esse autor, os relacionais são os adjetivos que

¹³³ Infelizmente apenas tivemos acesso a esta última publicação.

mais participam das colocações, pois, como sua função é de classificar - e não qualificar -, apresentam uma escassa importância semântica. Acreditamos, ao contrário, que justamente pela complexidade de relações que podem estabelecer com os substantivos¹³⁴, os adjetivos relacionais não apresentam uma tendência a participar de colocações. Por outro lado, consideramos inexata a afirmação de que “os adjetivos que expressam situação (c) não formam colocações”¹³⁵, pois alguns adjetivos que indicam “modo” podem estabelecer com os substantivos relações colocacionais semelhantes às que ocorrem entre verbos e advérbios, como no exemplo de *investigación exhaustiva* (MER04) [investigação minuciosa], retirado de nosso *corpus*, e a construção *investigar exhaustivamente* [investigar minuciosamente]¹³⁶.

Dentre os adjetivos que, segundo Koike (2000), tendem a formar colocações, há uma nova subdivisão entre os que constituem colocações funcionais e os que formam colocações léxicas. Nas primeiras, encontram-se os adjetivos que atuam como um intensificador do substantivo e correspondem às funções léxicas de Mel’čuk (1988) que representam *Magn* - indica que o substantivo ocorre em “um grau elevado” [*à un degré élevé*] -, nos casos de intensificação quantitativa, e *Bon* - denota um “elogio” [*louange*] -, nos casos de intensificação qualitativa. Já nas colocações léxicas, incluem-se os adjetivos que desempenham as funções de *Ver* - indica que o substantivo é correto, “tal como deve ser” [*tel qu’il doit être*] -, *Pos* - exprime uma “avaliação positiva” em relação ao substantivo - e *Epit* - possui um significado comumente vago. Assim, de acordo com Koike (2000, p. 147), do conjunto de 54 funções léxicas apresentadas por Mel’čuk (1988), somente cinco relacionam-se às combinações entre substantivos e adjetivos, o que poderia implicar em uma maior homogeneidade nas relações semânticas entre essas classes de palavras. No entanto, o próprio autor percebe que as funções léxicas não são capazes de cobrir todas as colocações entre substantivos e adjetivos e que “é difícil fazer uma classificação semântica das colocações léxicas devido à ampla gama de significados que os adjetivos apresentam”¹³⁷ (KOIKE, 2000, p. 175). Em virtude desse fato, limita-se a agrupar esse tipo de colocações de acordo com o substantivo com o qual se combinam, ou seja, combinações com substantivos concretos - *color crudo* [cor crua] -, abstratos - *poder adquisitivo* [poder aquisitivo] - e referidos ao estado físico - *dolor punzante* [dor intensa].

¹³⁴ Retomaremos este aspecto em 5.3.2.2.

¹³⁵ [los adjetivos que expresan situación (c) no forman colocaciones].

¹³⁶ Para um estudo aprofundado das relações colocacionais entre verbos e advérbios e suas correspondências com as construções de substantivo e adjetivo, v. García-Page (1993, 2001).

¹³⁷ [resulta difícil hacer una clasificación semántica de las colocaciones léxicas debido a la amplia gama de significados que presentan los adjetivos].

Devido às limitações presentes nas propostas de Koike (2000), acreditamos, então, ser necessária a elaboração de uma nova classificação dos adjetivos. Essa taxonomia deve possibilitar uma delimitação mais precisa dos tipos de adjetivos que tendem a formar colocações, a fim de que não sejam incluídas combinações livres ou terminológicas e que todas as combinações léxicas consideradas colocações possam ser teoricamente justificadas¹³⁸. Além disso, deve ser realizada uma categorização bastante refinada dos adjetivos, a fim de apresentar com maior precisão os grupos que tendem a formar colocações. Dessa forma, pretendemos diminuir a lacuna existente entre as combinações de substantivo e adjetivo que não puderam ser classificadas pelas funções léxicas ou pela taxonomia de Koike (2000).

5.1 Natureza do adjetivo

Para a proposta de uma nova classificação dos adjetivos, que permita verificar se existe um padrão entre aqueles que costumam formar colocações, passamos a identificar as principais características dessa classe de palavras através de seu estudo em diferentes obras da literatura especializada. Além da análise de importantes gramáticas da língua espanhola, incluímos o estudo dos adjetivos em uma gramática textual do francês, pois essa perspectiva evidencia aspectos relevantes na configuração das combinações entre substantivos e adjetivos, os quais não costumam ser retratados nas gramáticas de base logicista e/ou estrutural¹³⁹. No entanto, como as relações entre essas classes de palavras realizam-se no nível da oração, essa obra pode ser facilmente aproximada das demais. Dessa forma, de acordo com a perspectiva privilegiada pelos autores, estabelecemos uma divisão entre, de um lado, um enfoque formal, correspondente aos estudos que enfatizaram os aspectos sintáticos e morfológicos que afetam os adjetivos e, por outro lado, um enfoque semântico, no qual são ressaltadas as características intrínsecas ao adjetivo e, em função destas, as relações que estabelecem com os substantivos.

¹³⁸ Nosso intuito é de que não sejam incluídas combinações como *iglesia católica* [igreja católica], considerada uma colocação para Koike (2000), mas pertencente a uma terminologia especializada, segundo Weinrich (1989, p. 276).

¹³⁹ Utilizamos uma gramática textual em língua francesa porque, como não existe esse tipo de obra em espanhol, o trabalho de Weinrich (1989) é o único disponível que aborda uma língua românica.

5.1.1 Enfoque formal

Dentro desse enfoque, encontram-se os autores que privilegiaram a constituição morfológica dos adjetivos, através do estudo dos sufixos que os compõem, bem como os que enfatizaram o posicionamento sintático dessa classe de palavras em relação aos substantivos que acompanham. Os diferentes sufixos utilizados para a formação dos adjetivos, como *-ble* (*amable* [amável]), *-ante* (*constante* [constante]), *-dor* (*creador* [criador]), entre outros, não parecem influenciar no caráter colocacional dessa classe de palavras. No entanto, o aspecto morfológico é importante para diferenciar os adjetivos *stricto sensu* daqueles que correspondem ao particípio de alguns verbos (v. 5.3.2.1), pois, conforme o caráter dos adjetivos, as relações estabelecidas com os substantivos serão diferenciadas. Com relação ao aspecto sintático, somente após a determinação do adjetivo como sendo utilizado de forma anteposta ou posposta, os autores deste enfoque procuraram identificar as relações semânticas que essa classe de palavras mantém com o substantivo. Assim, devido à relativa homogeneidade entre as obras analisadas, apresentamos, a seguir, os estudos desse enfoque ordenados cronologicamente.

5.1.1.1 Bello (1942)¹⁴⁰

Bello (1942) define os adjetivos como palavras que modificam o substantivo. Essa modificação pode acontecer através do acréscimo de uma informação à significação do substantivo ou pelo desenvolvimento de um aspecto já compreendido em sua significação. Segundo Bello (1942, p. 12, grifo do autor), “no primeiro caso o adjetivo *particulariza, específica*, no segundo *desenvolve, explica*”¹⁴¹. Os adjetivos que particularizam costumam aparecer pospostos ao substantivo, como ocorre em *animales mansos* [animais mansos]; já os explicativos, também chamados epítetos, geralmente são antepostos¹⁴², como no exemplo de

¹⁴⁰ Ainda que o texto original de Bello date de 1847, consideramos importante a incorporação dessa gramática, uma vez que continua sendo uma obra de referência nos estudos de lingüística hispânica.

¹⁴¹ [en el primer caso el adjetivo *particulariza, específica*, en el segundo *desenvuelve, explica*].

¹⁴² Para Gili Gaya (1979, p. 164) “o epíteto, que é um adjetivo de certa forma supérfluo e usado com intenção artística, é raro fora da obra literária” [el epíteto, que es un adjetivo en cierto modo superfluo y usado con intención artística, es raro fuera de la obra literaria].

mansas ovejas [mansas ovelhas], em que não se indica uma espécie particular de ovelhas, mas se destaca uma característica natural a todas elas¹⁴³. Dessa forma, Bello (1942) associa as relações semânticas estabelecidas entre substantivos e adjetivos com a posição que estes ocupam ao modificar o substantivo, partindo, portanto, de um enfoque sintático para organizar sua classificação dos adjetivos. Essa separação entre pospostos e antepostos será importante para o estudo das colocações, uma vez que os primeiros, por ressaltarem um traço inerente ao substantivo, parecem não apresentar uma pré-disposição a participar dessas combinações.

5.1.1.2 *Real Academia Española (1973)*

Para o tratamento dos adjetivos, o *Esbozo de una nueva gramática de la lengua española*, compilado pela REAL ACADEMIA ESPAÑOLA (1973, p. 190), propõe que estes sejam divididos de acordo com sua posição na oração. Conforme os adjetivos apareçam separados do substantivo por um verbo de ligação, como ocorre em *sus ojos son claros* [seus olhos são claros], desempenham uma função predicativa; já nos casos em que não ocorre tal separação, como em *ojos claros* [olhos claros], o adjetivo cumpre uma função atributiva. No entanto, independentemente da função que desempenham, a REAL ACADEMIA ESPAÑOLA (1973, p. 191-194) estabelece uma divisão dos adjetivos em três grupos, de acordo com sua terminação na formação do gênero gramatical. Assim, os estudos dessa gramática centram-se principalmente no comportamento sintático e morfológico dos adjetivos. Reiteramos, porém, que essa distinção morfológica, a partir das terminações, não parece ser relevante para o estudo das colocações.

Um viés semântico aparece incorporado somente quando se trata da formação do superlativo, pois se percebe que apenas os adjetivos que permitem uma gradação de sua intensidade costumam aparecer nessa forma, como acontece com *firme* [firme], que pode ser intensificado pelos advérbios *muy* [muito], *bastante* [bastante], *poco* [pouco], além de formar o superlativo *firmísimo* [firmíssimo]. Essa noção de intensificação, já retratada por Bally (1951), será importante para a configuração de um tipo de colocação estabelecida entre

¹⁴³ Na terminologia de Bello (1942), os adjetivos antepostos e os epítetos são tratados como sinônimos. No entanto, é importante que se estabeleça a distinção entre essas duas denominações, pois enquanto na primeira destaca-se a posição que o adjetivo ocupa em relação ao substantivo, a segunda refere-se à característica semântica do epíteto de ressaltar um traço inerente ao substantivo, assumindo uma função quase pleonástica.

adjetivos e substantivos (v. 5.3.2.4.2a). Por outro lado, a REAL ACADEMIA ESPAÑOLA (1973, p. 198-200) identifica que os adjetivos utilizados no grau comparativo costumam formar pares de oposição, como entre *interior* [interior] e *exterior* [exterior], constituindo pequenos inventários fechados. Finalmente, essa gramática destaca a relação que se estabelece entre adjetivos e advérbios, através da propriedade que os adjetivos têm de se transformar em advérbios pela incorporação do sufixo *-mente*. Acreditamos que essa relação deva ser considerada no estudo das colocações, pois permite uma aproximação das combinações formadas por substantivos e adjetivos, como *lluvia torrencial* (COM04) com aquelas formadas por verbos e advérbios, como *llover torrencialmente* (MUN02).

5.1.1.3 Gili Gaya (1979)

Esse autor centra-se especialmente no tratamento sintático do adjetivo, através do estudo de sua posição em relação ao substantivo que acompanha. Propõe, inicialmente, uma distinção entre adjetivos determinativos e qualificativos de acordo com a função que desempenham. Os primeiros, que aparecem nos exemplos *veinte vacas* [vinte vacas]¹⁴⁴ e *algunos amigos* [alguns amigos], costumam ser antepostos. Já os segundos permitem o uso da posposição e da anteposição, produzindo um valor expressivo diferenciado para cada caso (GILI GAYA, 1979, p. 216), como ocorre com *hombre estimable* [homem estimável] e *estimable hombre* [estimável homem]. Acreditamos, porém, que a maior diferença entre os adjetivos determinativos e os qualificativos esteja relacionada não à posição que ocupam em relação aos substantivos, mas à própria natureza dessa classe de palavras, visto que o primeiro grupo apresenta apenas adjetivos de caráter “gramatical”, isto é, que desempenham uma função gramatical na combinação, enquanto no segundo grupo os adjetivos possuem um significado léxico¹⁴⁵.

Gili Gaya (1979, p. 216-220) apresenta três explicações para as diferentes posições dos adjetivos qualificativos:

¹⁴⁴ Em língua espanhola, diferentemente da língua portuguesa, os numerais são incluídos no grupo dos adjetivos determinativos (cf. nota 2), não constituindo, portanto, uma categoria gramatical própria.

¹⁴⁵ Para essa distinção, v. nota 85.

- a) do ponto de vista “**lógico**”, o adjetivo posposto especifica o substantivo, enquanto que o anteposto desenvolve uma qualidade presente no substantivo¹⁴⁶;
- b) do ponto de vista “**psicológico**”, a posição do adjetivo está relacionada ao estilo, sendo que a anteposição assume um caráter mais subjetivo ou afetivo e a posposição apresenta um caráter objetivo¹⁴⁷;
- c) do ponto de vista “**sintático e rítmico**”, os adjetivos são organizados de acordo com a estrutura da frase, sua intenção e suas características rítmicas¹⁴⁸.

Por fim, esse autor apresenta a possibilidade dos adjetivos se converterem em advérbios - com ou sem a adição de sufixos¹⁴⁹ -, como em *el niño duerme tranquilamente* [o menino dorme tranqüilamente] ou *el niño duerme tranquilo* [o menino dorme tranqüilo]¹⁵⁰. Além disso, essa classe de palavras pode ter sua intensidade modificada por meio de advérbios - *extremadamente arriesgado* [extremamente arriscado] -, superlativos - *feísimo* [feíssimo] -, aumentativos - *grandón* [grandalhão] -, diminutivos - *bajito* [baixinho] - ou comparativos - *tan inteligente como* [tão inteligente como].

5.1.1.4 Alcina, Blecua (1987)

Alcina, Blecua (1987) apresentam uma categorização semântica dos adjetivos ao separá-los entre os que expressam qualidades inerentes ao substantivo, tais como tamanho, forma, cor, características morais, entre outras, e os que apresentam a relação desse substantivo com diferentes aspectos sociais, culturais, políticos, etc¹⁵¹. Esses autores, no entanto, não exploram as diferenças existentes entre esses dois tipos de adjetivos, mas

¹⁴⁶ Essa visão corresponde à de Bello (1942).

¹⁴⁷ Weinrich (1981, p. 289) não concorda com esse ponto de vista, o qual considera uma mitificação.

¹⁴⁸ Considerando que as colocações são combinações léxicas que ocorrem de forma arbitrária e sem uma explicação aparente para a escolha mútua de determinados vocábulos, os aspectos rítmico-situacionais também poderiam desempenhar um papel relevante na conformação desse fenômeno léxico. No entanto, esse tipo de estudo não será desenvolvido no presente trabalho, pois ultrapassa seu escopo.

¹⁴⁹ Em oposição ao proposto pela REAL ACADEMIA ESPAÑOLA (1973, p. 200-201), para quem a transformação do adjetivo em advérbio acontecia somente através da inclusão do sufixo *-mente*.

¹⁵⁰ A propósito da conversão de adjetivos em advérbios, v. Basilio (1991). Para essa autora, porém, no exemplo de “o menino dorme tranqüilo”, em língua portuguesa, “tranqüilo” funciona como predicativo do sujeito, não correspondendo, portanto, a um advérbio.

¹⁵¹ Demonte (1999, p. 137-138) vai tratar esses tipos de adjetivos como qualificativos e relacionais, respectivamente.

propõem outras duas classificações conforme a posição que ocupam em relação ao substantivo. Na primeira delas, os adjetivos podem assumir um caráter mais subjetivo, apresentando uma qualidade de forma vaga e imaterial - nos casos de anteposição - ou cumprirem uma função objetiva, expressando uma qualidade de forma mais concreta e precisa - quando pospostos -, como ocorre no exemplo *viejos amigos* [velhos amigos] em oposição a *amigos viejos* [amigos velhos]. Na segunda classificação, os adjetivos podem ser divididos em especificativos - *casa alta* [casa alta] - ou explicativos - *alto rascacielos* [alto arranha-céu], de acordo com a taxonomia de Bello (1942). Dessa forma, a gramática oferece uma categorização dos adjetivos a partir dos pontos de vista “psicológico” e “lógico” (na terminologia de Gili Gaya (1979)), sem que, no entanto, seja feita uma reflexão, como ocorre em Gili Gaya (1979), de que essas são apenas diferentes perspectivas para o mesmo fenômeno.

5.1.1.5 Weinrich (1989)

Para o estudo dos adjetivos, Weinrich (1989, p. 272) parte da possibilidade de utilizá-los em forma atributiva - *qualification adjectivale* [qualificação adjetival], segundo sua terminologia - ou predicativa - *prédication adjectivale* [predicação adjetival] para o autor -, conforme os exemplos *les sciences exactes* [as ciências exatas] e *les sciences sont exactes* [as ciências são exatas], respectivamente. No caso das construções atributivas, Weinrich (1989, p. 273) também elabora sua classificação a partir da posição sintática que os adjetivos ocupam em relação aos substantivos que acompanham. Ao contrário dos demais autores desse grupo, porém, sua justificativa para a anteposição ou posposição dos adjetivos não corresponde a nenhum dos diferentes pontos de vista que Gili Gaya (1979, p. 216-220) chamou de “lógico”, “psicológico” ou “sintático e rítmico”, mas está diretamente relacionada a uma caracterização sêmica das combinações.

Segundo Weinrich (1989, p. 274), os adjetivos, quando pospostos, inserem na qualificação todos os traços semânticos de sua significação; por outro lado, quando antepostos, utilizam apenas um de seus traços lexicais. Assim, no exemplo *enfant pauvre* [criança pobre], o adjetivo insere todos os seus traços lexicais, tanto o de “bens” [biens] quanto o de “carência” [manque], constituindo a significação de “uma criança carente de bens”; já no caso de *pauvre enfant* [pobre criança], somente o traço de “carência” é

atualizado, reduzindo a significação da combinação a uma noção de inferioridade que caracteriza a situação precária dessa criança. Essa redução nos traços do adjetivo anteposto, para Weinrich (1989, p. 274), faz com que, de lexemas, passem a funcionar quase como morfemas¹⁵², aproximando-se das chamadas palavras “gramaticais”¹⁵³. No entanto, acreditamos haver uma diferença semântica considerável entre os adjetivos antepostos de tipo “gramatical”, como os determinativos presentes nos exemplos *otro día* (UNIM05) [outro dia] e *propio país* (UNIV02) [próprio país], em oposição a adjetivos utilizados de forma anteposta que indicam uma preferência estilística do falante, como ocorre em *duras críticas* (MER04) [duras críticas] e *claro ejemplo* (PRL02) [claro exemplo].

Em virtude dessa heterogeneidade, acreditamos que, mais do que pelo fato de se aproximarem aos morfemas, os adjetivos utilizados de forma anteposta não apresentariam uma tendência a formar colocações nos casos em que, segundo o próprio Weinrich (1989, p. 275), correspondem a unidades breves e freqüentes, particularmente da língua cotidiana. Como esses adjetivos, utilizados na anteposição, costumam ser freqüentemente empregados na língua geral, correspondem a palavras que podem facilmente ser combinadas com qualquer substantivo, não apresentando, portanto, uma tendência a formar colocação com unidades específicas. No entanto, consideramos que os adjetivos antepostos possam participar dessas combinações nos casos em que não manifestarem as características anteriores, pois nessas circunstâncias seu comportamento assemelha-se ao dos adjetivos pospostos.

A caracterização sêmica adotada por Weinrich (1989, p. 271) também permite a constatação de que as relações estabelecidas entre adjetivos e substantivos podem ser de dois tipos diferentes: determinação [*détermination*] e contra-determinação [*contre-détermination*]. No primeiro caso, os traços lexicais entre o substantivo e o adjetivo são compatíveis e o adjetivo acrescenta um traço à significação do substantivo, como acontece em *triangle équilatéral* [triângulo equilátero], em que “equilátero” acrescenta os traços de “lados” e “igualdade”. Já na relação de contra-determinação, os traços lexicais do substantivo e do adjetivo não são compatíveis, produzindo-se uma metáfora, como no exemplo *triangle érotique* [triângulo amoroso]. Porém, de acordo com as propriedades gerais apresentadas pelas colocações, as quais foram expostas em 2.4, essas combinações caracterizam-se pela

¹⁵² Os morfemas, segundo Trask (2006, s.v. *morfema*), correspondem à “menor unidade gramatical que se pode identificar”.

¹⁵³ Esse comportamento dos adjetivos, ainda que possa ser identificado na maior parte dos vocábulos antepostos que pertencem a essa classe de palavras, não parece aplicar-se a todos os casos, conforme pode ser observado em 5.3.2.4.1. Também Weinrich (1981, p. 281, grifo nosso) observa que “a anteposição é frequentemente, prescindindo de algumas poucas exceções, característica dos morfemas” [la anteposición es frecuentemente, prescindiendo de unas pocas excepciones, característica de los morfemas].

transparência de significado, o que excluiria os casos de metáfora, próprios das chamadas “expressões idiomáticas”¹⁵⁴. Dessa forma, acreditamos que as relações estabelecidas entre os dois elementos que conformam uma colocação serão sempre de determinação.

5.1.1.6 Alarcos Llorach (1999)

Alarcos Llorach (1999, p. 83) considera que a distinção tradicional dos adjetivos entre qualificativos e determinativos mescla considerações semânticas (ou diferenciais) e as propriamente funcionais. Para esse autor, a diferença entre os adjetivos não se baseia nesses critérios que ele chama “gramaticais”, mas estão relacionadas às diferentes áreas da realidade que afetam. Dessa forma, Alarcos Llorach (1999, p. 83) acredita que qualquer adjetivo pode qualificar ou determinar o significado do substantivo. Em vista disso, segundo o autor, tanto *blanco* [branco] quanto *aquel* [aquele], por exemplo, em combinação com *pañó* [pano], possuem a mesma propriedade tanto de qualificar quanto de determinar o substantivo; o que os diferencia é que seus significados referem-se a diferentes áreas da realidade, como o campo das cores e das situações, respectivamente. Alarcos Llorach (1999), porém, não especifica todas as possíveis áreas que correspondem a cada tipo de adjetivo e, ao final, retoma a distinção entre qualificativos e determinativos a partir da posição que ocupam em relação ao substantivo que acompanham, indicando que os segundos preferem a anteposição, enquanto os primeiros podem aparecer em ambas as posições. Alarcos Llorach (1999, p. 81-82) salienta, ainda, que o emprego dos qualificativos em anteposição denota uma intenção explicativa ou descritiva do substantivo; por outro lado, sua posposição serve para especificar ou restringir o significado do substantivo. Dessa forma, de acordo com a intenção do falante, um mesmo adjetivo pode apresentar significados diferentes dependendo da posição em que aparece com relação ao substantivo, como ocorre na oposição *simple periodista* [simples jornalista] e *periodista simple* [jornalista simples], em que a primeira corresponde a “alguém que nada mais é que um jornalista”, enquanto que a segunda refere-se a um “jornalista modesto, sem arrogância”. Alguns adjetivos, no entanto, costumam apresentar uma posição fixa, já consagrada pelo uso, como ocorre em *idea fija* [idéia fixa] e *sentido común* [senso comum].

¹⁵⁴ Uma “expressão idiomática” corresponde a uma combinação fixa cujo significado total não corresponde à soma dos significados individuais dos vocábulos que a compõem (HARTMANN, JAMES, 1998, s.v. *idiom*).

5.1.2 Enfoque semântico

Em oposição aos estudos anteriores, que, de um modo geral, partiram da classificação sintática para a identificação das características semânticas dos adjetivos, Demonte (1999), o único trabalho incluído nesse enfoque, apresenta uma proposta inovadora ao iniciar sua categorização dos adjetivos a partir do significado intrínseco dessa classe de palavras. Somente após a determinação das características peculiares a cada grupo de adjetivos são identificadas as relações que estes mantêm com os substantivos com os quais se combinam.

5.1.2.1 Demonte (1999)

Demonte (1999) propõe uma nova classificação dos adjetivos ao incorporar o componente semântico em sua taxonomia. A partir de seu significado intrínseco, divide os adjetivos em:

- a) **qualificativos**, que se referem a um traço constitutivo do substantivo e podem ser parafraseados por uma oração predicativa, como em *camino tortuoso* [caminho tortuoso] - *camino que es tortuoso* [caminho que é tortuoso];
- b) **relacionais**, que indicam um conjunto de propriedades, dando origem a relações semânticas mais complexas e diversificadas, como em *diccionario médico* [dicionário médico], que corresponde a um “dicionário que contém termos da medicina” e não pode ser parafraseado por **diccionario que es médico* [*dicionário que é médico];
- c) **adverbiais**, que não constituem expressões designadoras de propriedades, mas indicam a maneira como um conceito se aplica a determinado referente, conforme acontece em *supuesto asesino* [suposto assassino], além de sua relação espaço-temporal, como nos exemplos: *cercano puente* [ponte próxima] e *antiguo acuerdo* [antigo acordo].

Cada um dos três grupos recebe, ainda, diversas subdivisões de acordo com as relações semânticas estabelecidas entre os adjetivos e os substantivos que acompanham. Esse comportamento semântico diferenciado de cada grupo de adjetivos será importante, em nosso

trabalho, para identificar aqueles que apresentam uma maior tendência a formar colocações. Em vista disso, retomaremos a classificação de Demonte (1999) em 5.3.2, sem considerar, no entanto, os adjetivos que não manifestam tal pré-disposição. Por outro lado, os grupos que costumam formar colocações deverão ser constantemente subdivididos a fim de constituírem conjuntos mais homogêneos.

5.2 Necessidade de um novo tratamento dos adjetivos

As classificações dos adjetivos apresentadas pelas gramáticas de língua espanhola não se mostram suficientes para elaborar uma taxonomia que permita identificar os principais tipos de adjetivos que tendem a formar colocações. Isso se deve, especialmente, ao fato de essas obras centrarem seus estudos nos aspectos morfológicos e sintáticos que afetam as combinações de substantivos e adjetivos. Assim, do ponto de vista morfológico, as gramáticas agrupam essa classe de palavras a partir dos sufixos que formam suas terminações (cf. BELLO, 1942; REAL ACADEMIA ESPAÑOLA, 1973), o que não parece influenciar no tratamento das colocações, visto que tais combinações léxicas incluem uma ampla variedade de adjetivos. Por outro lado, baseadas em critérios sintáticos, apresentam uma grande divisão dos adjetivos de acordo com seu uso atributivo ou predicativo (cf. REAL ACADEMIA ESPAÑOLA, 1973; WEINRICH, 1989) e, nos casos de atribuição, conforme sua posição em relação ao substantivo (cf. BELLO, 1942; GILI GAYA, 1979; ALCINA, BLECUA, 1987; ALARCOS LLORACH, 1999; WEINRICH, 1989).

Somente após o posicionamento sintático do adjetivo em relação ao substantivo, são apresentadas explicações semânticas para tal fato, as quais variam de acordo com a perspectiva de cada teórico. De um lado, para Gili Gaya (1979) e Alarcos Llorach (1999), a anteposição ou posposição do adjetivo justifica a separação entre determinativos e qualificativos, pois os primeiros costumam aparecer antepostos, enquanto os segundos podem variar sua posição. No entanto, acreditamos que as diferenças entre tais adjetivos não podem ser explicadas somente em virtude de seu posicionamento, mas sim em razão de sua natureza diversa.

Segundo Weinrich (1989, p. 279), o adjetivo anteposto assemelha-se aos morfemas pela pobreza de traços lexicais; porém, nos casos de adjetivos determinativos, essa pobreza não se deve a sua localização em relação ao substantivo, mas corresponde a um traço

constitutivo do próprio adjetivo. Os adjetivos determinativos correspondem, devido a sua natureza, a palavras “gramaticais” e, em virtude desse fato, não parecem ser relevantes para o estudo das colocações. Por outro lado, Bello (1942) e Alcina, Blecua (1987) utilizam o posicionamento dos adjetivos qualificativos antes ou depois dos substantivos para justificar sua função semântica de especificar ou explicar o significado do substantivo que acompanham. Finalmente, para Weinrich (1989), a diferença entre a anteposição e a posposição do adjetivo diz respeito ao número de traços lexicais inseridos na qualificação do substantivo.

Assim, devido a essa diferença no significado veiculado pelo adjetivo conforme a posição que ocupe, parece-nos importante que seja considerada a distinção entre a anteposição e a posposição no momento de investigar quais adjetivos tendem a formar colocações, pois, embora um mesmo adjetivo possa aparecer em ambas as posições, seu comportamento poderá ser diferenciado. Os adjetivos antepostos que atualizam somente um traço inerente ao substantivo, por exemplo, parecem não apresentar uma pré-disposição a formar colocações. Em razão desse fato, acreditamos que uma categorização sintagmática dos adjetivos seja importante para homogeneizar os grupos analisados; no entanto, para o estudo das colocações, consideramos que mais importante que o posicionamento dos adjetivos em relação aos substantivos, sejam as características intrínsecas que cada tipo diferente de adjetivo apresenta.

Porém, ainda que apareça nas diversas gramáticas, o componente semântico recebe um tratamento mais aprofundado somente em Demonte (1999), pois essa autora inicia sua classificação dos adjetivos a partir da natureza diferenciada dessa classe de palavras. Tal separação também havia sido feita (de forma menos sistematizada) por Alcina, Blecua (1987), mas esses autores não aprofundaram essa classificação, limitando-se a indicar, como os demais teóricos, que conforme a posição do adjetivo sua relação com o substantivo será de explicação ou especificação. Dessa forma, não se percebe que, mesmo ocupando uma posição semelhante, os adjetivos podem desempenhar funções diferentes de acordo com a modificação que provocam no substantivo que acompanham, pois não é a posição que configura o adjetivo, mas sua natureza intrínseca o faz preferir determinada posição. Assim, embora ambos sejam utilizados em anteposição, há uma diferença entre uma combinação com um adjetivo determinativo e outra com um qualificativo, como ocorre em *otro día* (UNIM05) [outro dia] e *auspicioso día* (UNIV03) [auspicioso dia], respectivamente. Por outro lado, ainda que ambos sejam utilizados em posposição, os adjetivos qualificativos e os relacionais modificam os substantivos de maneira diferente, pois suas características particulares não são equivalentes, como acontece em *tráfico alto* (UNIV01) [tráfego intenso] e *tráfico aéreo*

(CLA02) [tráfego aéreo], em que se acrescenta uma característica ao substantivo e se restringe sua extensão, respectivamente.

Apesar dessas diferenças, somente Demonte (1999, p. 137) estabelece uma distinção entre os adjetivos devido a seu significado “intrínseco”, ao separá-los em qualificativos, relacionais e adverbiais. Essa distinção apresenta grande utilidade para a identificação dos adjetivos que tendem a formar colocações e será a base de nossa classificação. Portanto, para o estudo das combinações entre substantivos e adjetivos, consideramos ser necessária a utilização de uma nova taxonomia dessa classe de palavras, que permita relacionar os aspectos sintáticos e morfológicos com os semânticos¹⁵⁵. Além disso, deve ser elaborada uma classificação semântica ainda mais aprofundada dos adjetivos, a fim de identificar de forma mais específica as relações que eles podem manter com os substantivos.

5.3 Taxonomia dos adjetivos

Para a elaboração de uma nova taxonomia dos adjetivos, foram estudadas suas combinações com substantivos apenas em uso atributivo. Acreditamos que tal escolha não influencie nos resultados de nossa análise, pois Koike (2000, p. 148), que estudou ambas as combinações a partir de seu *corpus*, percebeu que a maior parte das colocações representava o uso atributivo.

A fim, então, de identificar os tipos de adjetivos que apresentam uma maior tendência a formar colocações, a primeira grande divisão que estabelecemos para essa classe de palavras, baseada em critérios “semânticos”, corresponde à separação entre os determinativos e os qualificativos (cf. ALARCOS LLORACH, 1999, p. 83).

¹⁵⁵ Koike (2000, p. 149) também afirma que “é preciso levar em consideração a classificação semântica do adjetivo para explicar por que existem adjetivos qualificativos que são capazes de formar colocações substantivo-adjetivo e outros que não estão dotados dessa faculdade léxica” [es preciso tener en cuenta la clasificación semántica del adjetivo para explicar por qué hay adjetivos calificativos que son capaces de formar colocaciones sustantivo-adjetivo y otros que no están dotados de esta facultad léxica].

5.3.1 Adjetivos determinativos

Os adjetivos determinativos correspondem a um inventário fechado e apresentam um significado estritamente gramatical. Incluem os adjetivos demonstrativos, possessivos, indefinidos, numerais¹⁵⁶ e relativos. Devido a sua natureza intrínseca, as relações estabelecidas entre esses adjetivos e os substantivos que acompanham são de tipo gramatical e, portanto, acreditamos que suas combinações não tendem a formar colocações¹⁵⁷.

5.3.2 Adjetivos qualificativos

Os adjetivos qualificativos, por sua vez, representam um inventário aberto, com adjetivos de claro significado léxico, constituindo um conjunto muito amplo de vocábulos. Em vista disso, para que pudéssemos ter acesso a grupos mais homogêneos de palavras, esses adjetivos receberam uma nova separação, orientada por critérios “semânticos” e “morfológicos”. Do ponto de vista “morfológico”, os adjetivos que correspondem à forma de particípio dos verbos foram considerados como um grupo separado. Por outro lado, do ponto de vista “semântico”, os adjetivos qualificativos foram, ainda, divididos em relacionais, situacionais e qualificativos *stricto sensu*, de acordo com a classificação proposta por Demonte (1999).

5.3.2.1 Adjetivos participios

Os participios foram também incorporados ao nosso estudo, pois, segundo Weinrich (1989, p. 297), essas unidades podem funcionar como adjetivos ao apresentarem um caráter qualificativo. Dentro desse grupo, estabelecemos uma distinção entre os participios de

¹⁵⁶ Ainda que os numerais possam ser infinitos, consideramo-os como um inventário fechado por apresentarem uma ordenação lógica que permite identificar todas as suas ocorrências, como acontece nos demais grupos que correspondem a adjetivos determinativos.

¹⁵⁷ Koike (2000, p. 148-149) também exclui os adjetivos determinativos do âmbito de seus estudos por essa mesma razão.

presente e de passado, de acordo com sua indicação temporal. Assim, por exemplo, combinações como *situación alarmante* (TIE06) [situação alarmante] e *puestos vacantes* (MUN04) [postos vagos] foram classificadas como participípios de presente, enquanto as combinações de *crecimiento acelerado* (MER06) [crescimento acelerado], *objetivos trazados* (DIA04) [objetivos traçados], entre outras, foram incluídas no grupo de participípios passados. No entanto, percebemos que muitas das combinações de substantivos com participípios, especialmente passados, aproximam-se da estrutura de combinações formadas por um substantivo e um verbo, como acontece com *daño causado* [dano causado], que corresponde à estrutura “causar um dano”, conforme observamos no exemplo: “A todas as pessoas deve-se dar a oportunidade de que possam ser úteis à sociedade e reparar o dano causado lutando pela causa da liberdade” (MUN02) [A todas las personas hay que darles la oportunidad de que puedan ser útiles a la sociedad, y reparar el daño causado luchando por la causa de las libertades]. Devido ao fato de as relações mantidas entre verbos e substantivos obedecerem a princípios diferentes dos que orientam as relações entre substantivos e adjetivos, não aprofundamos nosso estudo no grupo específico dos adjetivos participípios.

Além disso, Weinrich (1989, p. 310) também salienta que alguns desses adjetivos, quando muito empregados, podem sofrer um processo de lexicalização¹⁵⁸, adquirindo o *status* de adjetivo verbal [*adjectif verbal*] e ganhando uma entrada própria no dicionário, como ocorre com “encantada” na combinação *flûte enchantée* [flauta encantada]. Quando, porém, uma combinação torna-se lexicalizada, já não pode existir uma relação colocacional entre o substantivo e o adjetivo que a conformam, pois tais vocábulos configuram uma unidade. Em vista disso, uma série de combinações lexicalizadas presentes em nosso *corpus*, tais como *hijo pródigo* (CLA01) [filho pródigo], *bebida alcohólica* (UNIM03) [bebida alcólica], entre outras, não foram consideradas em nossa análise¹⁵⁹. Por essas mesmas razões, as “unidades terminológicas sintagmáticas”, definidas por Cabré (2001 apud BEVILACQUA, 2004, p. 28) como “unidades de estrutura léxica ou sintática de caráter denominativo-conceitual, dotadas

¹⁵⁸ De acordo com Hartmann, James (1998, s.v. *lexicalisation*), esse processo consiste em conferir o *status* de lexema a uma palavra ou expressão.

¹⁵⁹ A distinção entre uma combinação lexicalizada e uma candidata a colocação não é tão fácil de determinar (cf. PÖLL, 1996, p. 116). Para identificar essa diferença, utilizamos a definição proposta por Glück (2000, s.v. *Lexikalisierung*), que entende “unidades lexicalizadas” como aquelas combinações de palavras que sofreram um processo de lexicalização, ou seja, correspondem a formações de palavras que apresentam independência semântica, mas que, devido a seu emprego freqüente, aparecem como unidades na língua. Segundo Galisson, Coste (1976, s.v. *lexicalisé*), as “unidades lexicalizadas” constituem uma seqüência indissociável que exprime uma noção única.

de capacidade de referência que denominam um nó da estrutura conceitual de um âmbito”¹⁶⁰, também foram excluídas de nossa análise. Como exemplos dessas combinações de nosso *corpus*, temos: *tarjeta amarilla* (PRL06) [cartão amarelo], *homicidios dolosos* (UNIM03) [homicídios dolosos], entre outros.

5.3.2.2 Adjetivos relacionais

Os adjetivos relacionais, que se referem a um conjunto complexo de propriedades dos substantivos, servem para categorizá-los, através de sua inclusão em um grupo específico¹⁶¹. Por envolverem uma ampla gama de propriedades, os adjetivos relacionais não permitem sua gradação e seu uso em comparações ou em construções predicativas (cf. DEMONTE, 1999, p. 139; WEINRICH, 1989, p. 272). Embora esse fato não tenha uma implicação direta no estudo das colocações, demonstra o comportamento particular que esses adjetivos adquirem, o que pode ser um indicativo do motivo pelo qual não costumam formar colocações.

Dentro do grupo dos relacionais, Demonte (1999) apresenta uma nova classificação semântica para esses adjetivos, na qual os divide entre os que correspondem a uma função gramatical, uma função adjunta e aqueles cujo significado aparece integrado ao substantivo que acompanham. Não aprofundaremos nosso estudo nessa classificação dos relacionais por acreditarmos que, justamente em razão dessas funções que cumprem, tais adjetivos não costumam formar colocações com os substantivos que acompanham. Koike (2000, p. 153), entretanto, chega a conclusões exatamente opostas às nossas, ao afirmar que “os adjetivos relacionais (...) são, em sua maioria, capazes de constituir colocações”¹⁶². Esse autor, porém, não aprofunda sua pesquisa nesse tipo de adjetivos e muitos dos exemplos apresentados - tais como *iglesia católica* [igreja católica], *barrio céntrico* [bairro central], *registro civil* [registro civil], *establecimiento comercial* [estabelecimento comercial], entre outros -, parecem constituir o que Demonte (1999, p. 162) classificou como “adjetivos cujo

¹⁶⁰ [unidades de estructura léxica o sintáctica de carácter denominativo-conceptual, dotadas de capacidad de referencia que denominan un nodo de la estructura conceptual de un ámbito].

¹⁶¹ Koike (2000, p. 151) apresenta como duas diferentes funções semânticas dos adjetivos as de classificar e relacionar. Acreditamos que isso se deva ao fato de o autor empregar a terminologia proposta por diferentes autores, pois, na realidade, ambas as funções aparecem unidas e são desempenhadas pelos relacionais.

¹⁶² [los adjetivos relacionales (...) son, en su mayoría, capaces de constituir colocaciones].

significado aparece integrado ao substantivo” [*adjetivo relacional de significado integrado en el nombre*]. Essas estruturas, ainda segundo Demonte (1999, p. 156), levaram diversos autores a aproximar tais combinações a certos compostos das línguas romance e, mais especificamente, a “postular que a relação entre o substantivo e o adjetivo relacional é mais uma relação morfológica que uma relação sintática”¹⁶³. Dessa forma, a maior parte das combinações de substantivo e adjetivo relacional apresentadas por Koike (2000) parece aproximar-se mais de unidades lexicalizadas do que conformar colocações propriamente ditas.

5.3.2.3 Adjetivos situacionais¹⁶⁴

Os adjetivos situacionais, segundo Demonte (1999, p. 204-205), não introduzem uma propriedade aos substantivos, mas podem modificar seus aspectos conceituais, temporais e espaciais. Conforme ocorra essa modificação, os adjetivos são divididos entre:

- a) **intensionais**, que indicam como um conceito se aplica a determinado referente, modificando apenas sua intensão¹⁶⁵, como ocorre em *presunto asesino* [suposto assassino], e
- b) **eventivos**, que se referem ao espaço - *cercano puente* [ponte próxima] -, ao tempo - *antiguo acuerdo* [antigo acordo] - e à maneira - *mirada severa* [olhar severo] - que afetam os objetos e os processos.

Ambos os grupos apresentam, ainda, subdivisões na classificação de Demonte (1999). Em nosso trabalho, porém, será considerada apenas a distinção entre os diferentes tipos de eventivos, conforme funcionem como marcadores espaço-temporais ou indiquem a maneira como uma situação ou objeto se comportam. Acreditamos que, para o estudo das colocações, apenas essa diferenciação seja suficiente, pois os adjetivos intensionais aproximam-se dos “adjetivos como morfemas de papéis” [*adjetifs en tant que morphèmes de rôles*] e dos “adjetivos como artigos específicos” [*adjetifs en tant qu’articles spécifiques*], na

¹⁶³ [postular que la relación entre el nombre y el adjetivo relacional es más una relación morfológica que una relación sintáctica].

¹⁶⁴ Empregamos a terminologia proposta por Koike (2000) para nos referirmos ao grupo de adjetivos correspondente aos advérbios em Demonte (1999).

¹⁶⁵ O conceito de “intensão” deve ser entendido em oposição ao de “extensão”. O primeiro é definido pelo conjunto de qualidades ou traços que caracterizam um conceito específico, enquanto o segundo corresponde ao conjunto de elementos que determinada expressão linguística designa (cf. BUSSMANN, 1990, s.v. *Intension e Extension*).

terminologia de Weinrich (1989, p. 284-288), que, por se assemelharem a morfemas e adquirirem o valor de um signo gramatical, não tendem a formar colocações. Segundo Koike (2000, p. 153), também “os adjetivos que expressam situação não formam colocações”¹⁶⁶; no entanto, acreditamos que os adjetivos eventivos que expressam a maneira como uma situação se realiza - denominados “modais” em nossa classificação -, podem vir a formar colocações, as quais se aproximam das combinações estabelecidas entre verbos e advérbios de modo, conforme o exemplo entre *cambio radical* [mudança radical] (MUN04) e *cambiar radicalmente* [mudar radicalmente] (UNIV01). Essa aproximação entre as combinações que utilizam adjetivos e advérbios ressalta a relação entre essas classes de palavras, já apontada nos trabalhos da REAL ACADEMIA ESPAÑOLA (1973) e de Gili Gaya (1979). No entanto, da mesma forma que os adjetivos participio, os modais não serão analisados neste trabalho, por consideramos que seu comportamento se aproxime ao das combinações regidas por verbos.

5.3.2.4 *Adjetivos qualificativos stricto sensu*

Para os adjetivos qualificativos *stricto sensu*, propomos, ainda, uma divisão “sintática”, baseada em sua posição anteposta ou posposta em relação ao substantivo que acompanham.

5.3.2.4.1 *Adjetivos antepostos*

A partir das diferentes propostas para justificar o uso anteposto ou posposto dos adjetivos, identificamos que os primeiros podem se comportar de três maneiras diferentes:

- a) como palavras gramaticais, nos casos dos determinativos: *este drama* (UNIV05) [este drama], *ningún control* (UNIE06) [nenhum controle], *tercer episodio* (TIE05) [terceiro episódio], etc.;

¹⁶⁶ [los adjetivos que expresan situación no forman colocaciones].

- b) como morfemas, através de uma redução de seus traços lexicais, nos casos em que possuem uma significação diferente conforme a posição que ocupam em relação ao substantivo: *viejo amigo* (COM03) [velho amigo], *gran tesoro* (UNIM03) [grande tesouro], etc.;
- c) como palavras de significação “completa”, que selecionam, no nível da norma, a posição anteposta, como em *craso error* (UNIE04) [erro crasso], *enconada polémica* (MUN06) [polêmica acirrada], etc.

De acordo com as razões já explicitadas em 5.2, tanto por seu caráter gramatical quanto pelo fato de corresponderem, segundo Weinrich (1989, p. 275), a unidades breves e freqüentes, particularmente da língua geral, acreditamos que os adjetivos dos grupos (a) e (b) não apresentam uma tendência a constituir colocações. Considerando que essas combinações léxicas são caracterizadas por uma combinabilidade restrita entre seus elementos (v. 2.4), o fato de os adjetivos antepostos poderem ser combinados com uma ampla variedade de substantivos, elimina seu caráter colocacional. Em virtude desse fato, a maior parte dos adjetivos antepostos foi excluída de nossa análise, restando somente aqueles que pertenciam ao grupo (c).

5.3.2.4.2 *Adjetivos pospostos*

Para o conjunto dos adjetivos pospostos, devido à grande quantidade de vocábulos que compreendia, foi necessário realizar uma nova divisão, baseada na característica semântica introduzida por cada adjetivo. A partir disso, os adjetivos pospostos foram classificados em intensificadores e qualificadores, conforme graduassem a intensidade de uma qualidade já presente no substantivo ou acrescentassem uma característica a ele, respectivamente.

a) Adjetivos intensificadores

Os adjetivos intensificadores, apresentados pela REAL ACADEMIA ESPAÑOLA (1973) como os utilizados para originar o grau superlativo, foram divididos de acordo com o tipo de graduação que conferiam ao substantivo, ou seja, conforme apresentassem uma intensificação quantitativa ou qualitativa. Embora os adjetivos de ambos

os grupos manifestem um aumento ou diminuição do significado expresso pelo substantivo, os primeiros representam a intensificação de algo que pode ser medido em termos absolutos e os segundos correspondem a uma intensificação mais subjetiva do substantivo, como na diferença entre *consumo abusivo* (MUN05) [consumo abusivo] e *trabajo arduo* (MER06) [trabalho árduo], respectivamente. Koike (2000, p. 155) também utiliza essa distinção para o que denomina “adjetivos funcionais” [*adjetivos funcionales*] e propõe, ainda, sua divisão entre intensificação positiva ou negativa. Acreditamos, porém, que tal separação não apresente relevância para o estudo das colocações, uma vez que ambas as categorias tendem a formar esse tipo de combinações léxicas.

b) Adjetivos qualificadores

Os adjetivos qualificadores - “qualificativos” [calificativos], segundo Demonte (1999, p. 175) -, são classificados por essa autora através de critérios léxico-sintáticos que ressaltam suas propriedades de indicar dimensão, velocidade, propriedade física, cor (e forma), idade, valoração e atitudes. Ainda que Demonte (1999, p. 172) advirta que “classificar adequadamente os adjetivos qualificativos não é uma tarefa fácil”¹⁶⁷, acreditamos que sua categorização apresenta separações arbitrárias, como no fato de não incluir os adjetivos de cor e forma entre os que indicam propriedades físicas. Além disso, muitas outras “classes semânticas” poderiam ser acrescentadas em sua taxonomia. Dessa forma, entendemos que seja necessária a proposta de uma nova categorização dos adjetivos qualificadores, tendo em vista sua funcionalidade para o estudo das combinações de adjetivos com substantivos, a fim de destacar aquelas que tendem a formar colocações.

Na tentativa de delimitar grupos mais homogêneos de adjetivos qualificadores, acreditamos ser necessário aprofundar nosso estudo da noção de “qualidade” veiculada pelo adjetivo. Isso se deve ao fato de as classificações existentes para essa classe de palavras, embora baseadas em formalismos irredutíveis, como os critérios morfológicos, sintáticos e semânticos, não revelarem uma concepção mais integral da noção de qualidade. Acreditamos que uma classificação exaustiva dos adjetivos deva também considerar as relações estabelecidas entre um atributo “consustancial” frente a um atributo “acessório” ao substantivo que acompanham.

¹⁶⁷ [clasificar adecuadamente a los adjetivos calificativos no es tarea sencilla].

Tal fato indica, então, a necessidade de se compreender a noção de qualificador desempenhada pelo adjetivo, por um lado, como a explicitação de traços efetivamente ligados de forma consubstancial a um substantivo. Nesses casos ocorre uma relação atributiva particular entre os elementos, como a que se estabelece, por exemplo, entre a noção de “volume” e a entidade “cadeira”, a qual nos permite qualificar uma cadeira como sendo grande ou pequena. Por outro lado, há relações qualificadoras entre substantivos e adjetivos que explicitam uma relação que obedece a uma dimensão não-particular, ou seja, em que o adjetivo não explicita traços consubstanciais ao substantivo, como ocorre no exemplo *informação relevante*.

Essa constatação permite postular uma classificação dos adjetivos em concretos e abstratos¹⁶⁸. De acordo com essa categorização, os adjetivos concretos são aqueles que explicitam uma relação de atribuição de algo particular em relação a um substantivo. Em termos fenomenológicos, um adjetivo concreto é aquele que apresenta uma qualidade “sensível”¹⁶⁹. Por outro lado, os adjetivos abstratos denotam uma qualidade não-particular, não-sensível.

O estudo das magnitudes “concreto” e “abstrato” como critérios classificatórios dos adjetivos não aparece desenvolvido na literatura especializada. Em vista disso, procuramos delimitar tais conceitos a partir de sua concepção filosófica. No entanto, os estudos realizados do ponto de vista dessa perspectiva relacionam as noções de “abstrato” e “concreto” com as relações de parte-todo, respectivamente (cf. FERRATER MORA, 1996, s.v. *abstracción, abstrato e concreto*; ENCYCLOPÉDIE, 1989, s.v. *abstrait e concret*), o que não se equivale a uma concepção lingüística desse fenômeno.

Do ponto de vista lingüístico, no caso particular da língua espanhola, a aplicação dessas magnitudes na classificação das entidades nominais (tanto substantivos quanto adjetivos) persistiu somente até a *Gramática de la lengua española*, compilada pela REAL ACADEMIA ESPAÑOLA (1931), mas foi abandonada pela Lingüística Hispânica com a consolidação do Estruturalismo¹⁷⁰. Em vista disso, a distinção entre “concretos” e “abstratos”

¹⁶⁸ Koike (2000, p. 162), ao analisar o comportamento dos substantivos que formam colocações com os adjetivos, identifica que “os substantivos abstratos costumam selecionar, por seu conteúdo, adjetivos com significado abstrato ou não-físico, enquanto os concretos selecionam adjetivos com significado concreto” [los sustantivos abstractos suelen seleccionar por su contenido adjetivos con significado abstracto o no físico, mientras que los concretos seleccionan adjetivos con significado concreto].

¹⁶⁹ Entendemos “sensível” como algo que pode ser percebido pelos sentidos (cf. FERRATER MORA, 1994, s.v. *sensible*).

¹⁷⁰ Também Alarcos Llorach (1999, p. 82, grifo nosso) ressalta que “o adjetivo é um tipo de palavra cujos morfemas em geral coincidem com os do substantivo e, por isso, ambos costumam ser reunidos na categoria dos

é atualmente utilizada somente para categorizar os substantivos. Entendemos que esse fato seja fundamental, visto que uma classificação de todas as entidades nominais à luz desses critérios torna-se insuficiente e não livre de controvérsia.

No momento atual da discussão, não contamos com parâmetros formais que possibilitem sustentar uma classificação dos adjetivos a partir desses mesmos critérios. No entanto, acreditamos que a distinção entre concretos e abstratos possa ser de grande auxílio para uma compreensão mais integral do que entendemos como “comportamento colocacional”, visto que uma quantidade significativa de candidatos a colocação parece obedecer a padrões ligados à distinção supracitada. A partir dessa diferença percebemos, por exemplo, que os adjetivos ditos “concretos” não parecem constituir colocações, como pode ser observado nos exemplos de *agua sucia* (UNIM03) [água suja], *color turquesa* (PRL01) [cor turquesa], entre outros.

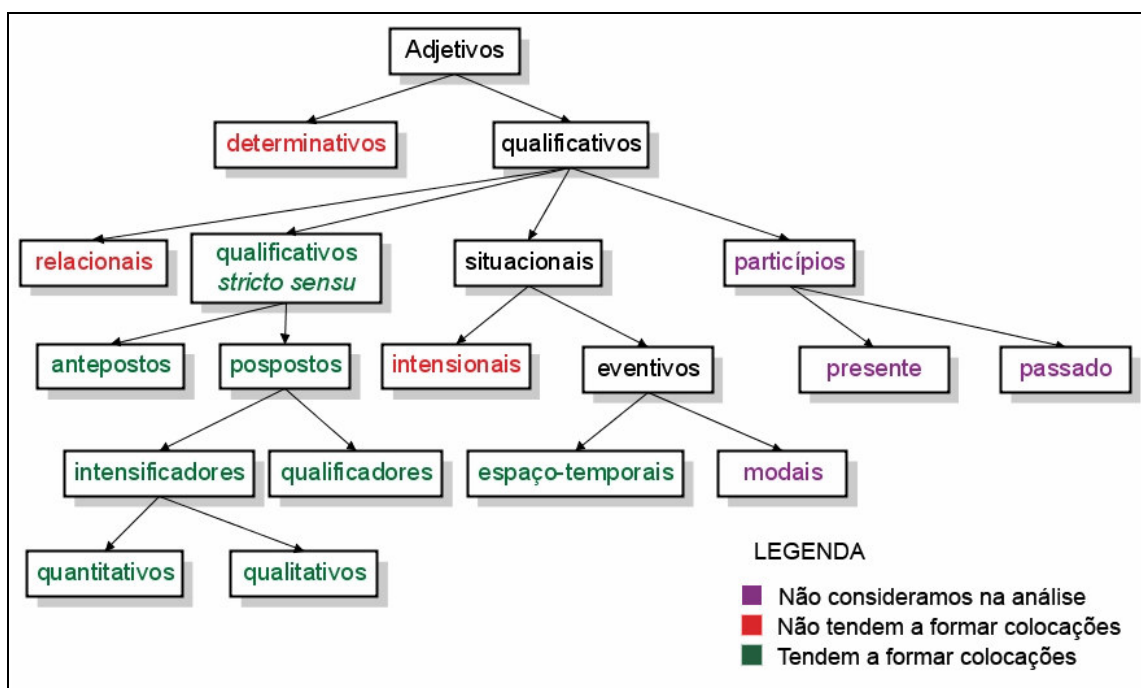
A impossibilidade, porém, de fundamentar uma classificação dos adjetivos a partir das noções de “concreto” e “abstrato” faz com que uma grande quantidade de candidatos a colocações não possa ser explicada pelos critérios taxonômicos por nós arrolados. Embora o presente trabalho almeje elucidar parâmetros que permitam gerar mecanismos para uma melhor detecção do fenômeno colocacional, existem lacunas nos estudos que concernem à Linguística Geral que devem ser previamente solucionadas para que se possa alcançar uma compreensão mais abrangente do comportamento colocacional de determinadas categorias lingüísticas.

Em virtude das limitações apresentadas, tentamos homogeneizar o grupo dos adjetivos que entendemos como “abstratos”, aplicando sobre essas unidades outras propostas de classificações existentes. Tais propostas consistiram na identificação de se os adjetivos desse grupo relacionavam-se a uma qualidade “estável” ou “transitória” (através da transformação das construções atributivas por construções predicativas com o uso dos verbos “ser” e “estar”, respectivamente) (cf. DEMONTE, 1999, p. 142) ou ainda se representavam uma avaliação (como as realizadas pelas dicotomias “bom”/“ruim”, “fácil”/“difícil”, entre outras) em oposição a pré-disposições (cf. DEMONTE, 1999, p. 179-188). No entanto, nenhuma das classificações utilizadas apresentou resultados satisfatórios para nosso estudo, pois muitos elementos não permitiam sua inclusão em grupos específicos e, além disso, os pequenos conjuntos que puderam ser isolados continuavam sendo tão heterogêneos que não permitiam a indicação de tendências.

nomes” [el adjetivo es un tipo de palabra cuyos morfemas coinciden en general con los del sustantivo, y por ello se suelen ambos reunir en la categoría de los nombres].

5.4 Resultados da análise qualitativa

Devido à impossibilidade de um maior refinamento de nossa classificação, conseguimos isolar somente grandes grupos de adjetivos em função de seu comportamento colocacional. A seguir, a fim de sintetizar a taxonomia proposta para essa classe de palavras, tendo em vista seu grau de propensão a participar de colocações, apresentamos uma sistematização de nossa classificação dos adjetivos:



Esquema 1: Taxonomia dos adjetivos

A partir da classificação dos adjetivos em diferentes categorias, segundo as relações semânticas mantidas com os substantivos, foi possível isolar 11 tipos de comportamentos para essa classe de palavras. O quadro a seguir apresenta alguns exemplos de cada um dos tipos de adjetivos de nossa classificação e permite observar as relações que esses vocábulos estabelecem com os substantivos que acompanham.

Tipos de adjetivos	Exemplos
1. Determinativos	<i>este</i> drama (UNIV05) [<i>este</i> drama] <i>ningún</i> control (UNIE06) [<i>nenhum</i> controle]
2. Relacionais	estructura <i>metálica</i> (UNIM02) [estrutura <i>metálica</i>] dolor <i>abdominal</i> (MER03) [dor <i>abdominal</i>]
3. Qualificativos antepostos	<i>flamante</i> embajador (CLA05) [<i>novo</i> embaixador] <i>despiadada</i> ofensiva (UNIE01) [<i>brutal</i> ofensiva]
4. Qualificadores	casa <i>acogedora</i> (PRL02) [casa <i>acolhedora</i>] encuentro <i>amistoso</i> (UNIV01) [jogo <i>amistoso</i>]
5. Intensificadores quantitativos	derrota <i>abrumadora</i> (MER05) [derrota <i>esmagadora</i>] frío <i>intenso</i> (PRL06) [frio <i>intenso</i>]
6. Intensificadores qualitativos	apoyo <i>irrestricto</i> (MUN02) [apoio <i>irrestrito</i>] trabajo <i>arduo</i> (UNIM06) [trabalho <i>árduo</i>]
7. Intensionais	<i>potencial</i> comprador (MER05) [<i>potencial</i> comprador] <i>supuesto</i> delito (COM06) [<i>suposto</i> delito]
8. Espaço-temporais	terreno <i>adyacente</i> (UNIV06) [terreno <i>adjacente</i>] homenaje <i>póstumo</i> (UNIM01) [homenagem <i>póstuma</i>]
9. Modais	cambio <i>radical</i> (MER01) [mudança <i>radical</i>] lluvias <i>torrenciales</i> (COM04) [chuvas <i>torrenciais</i>]
10. Particípios presente	situación <i>alarmante</i> (MUN02) [situação <i>alarmante</i>] problemas <i>acuciantes</i> (CLA06) [problemas <i>urgentes</i>]
11. Particípios passado	objetivos <i>trazados</i> (DIA04) [objetivos <i>traçados</i>] tiempo <i>perdido</i> (UNIV06) [tempo <i>perdido</i>]

Quadro 2: Exemplos de adjetivos classificados

Consideramos importante salientar que o que define a categoria a qual um adjetivo pertence não é seu significado intrínseco, mas o tipo de relação que esse vocábulo estabelece com o substantivo que acompanha. Em vista disso, não se pode dizer que um adjetivo seja, por natureza, relacional ou qualificativo, por exemplo, mas somente que desempenha essa função em sua relação com um substantivo específico. Assim, foi possível observar que um mesmo adjetivo pode ser classificado em diferentes categorias conforme o substantivo que acompanhe, como observamos nas combinações de *riesgo vital* (MER05) [risco de vida] e *importancia vital* (CLA04) [importância vital], em que o adjetivo *vital* desempenha a função de um relacional e de um intensificador, respectivamente. Por outro lado, um mesmo adjetivo pode pertencer a duas categorias diferentes, conforme a perspectiva a partir da qual seja analisado. Assim, por exemplo, do ponto de vista “sintático”, *aparente* [aparente] em combinações como *aparente contradicción* (MER05) [aparente contradição] e *aparente venganza* (DIA04) [aparente vingança] pode ser classificado no grupo dos adjetivos antepostos; porém, do ponto de vista “semântico”, tal vocábulo pertence ao conjunto formado pelos adjetivos “intensionais”. Em vista disso, o pesquisador deve determinar explicitamente os critérios mais relevantes para sua classificação. No caso de nossa taxonomia, procuramos

privilegiar as relações semânticas existentes entre os vocábulos, de modo que o adjetivo apresentado nas combinações anteriores foi incluído, por nós, no grupo dos “intensionais”.

A partir da identificação das relações semânticas estabelecidas entre cada adjetivo e os substantivos que acompanhava, foi possível selecionar aqueles que apresentavam uma maior tendência a formar colocações, quais sejam, adjetivos “qualificativos *stricto sensu*” e “espaço-temporais”. Assim, combinações com adjetivos determinativos, participípios, relacionais, intensionais e modais puderam ser excluídas de nossa análise, o que resultou em um conjunto de dados bastante reduzido. Em termos numéricos, tal redução correspondeu a 51% do total dos adjetivos que formavam combinações¹⁷¹, como podemos observar no gráfico abaixo.

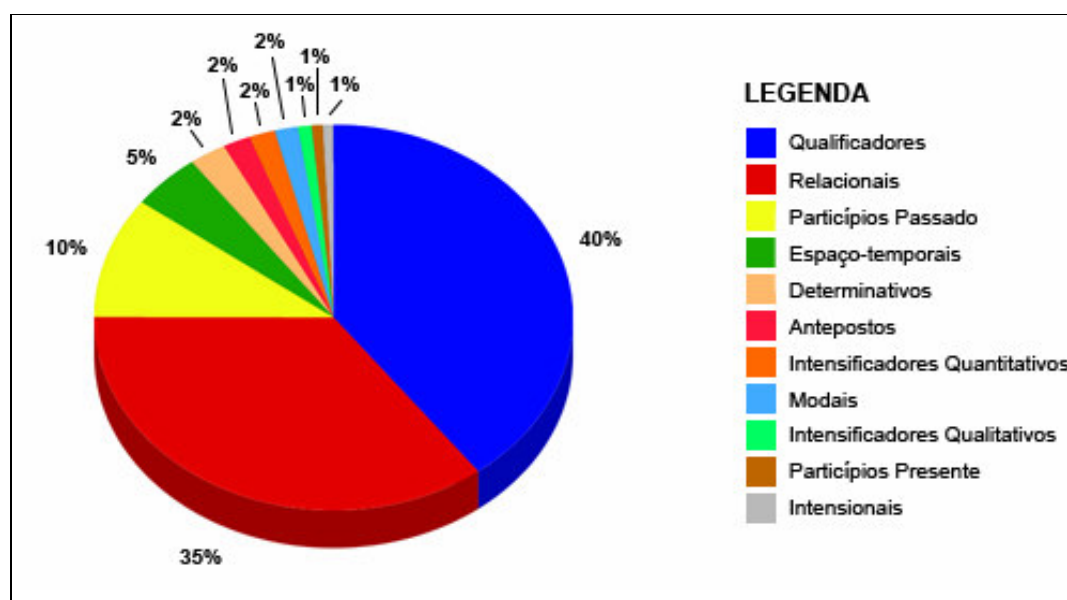


Gráfico 1: Distribuição dos adjetivos segundo sua classificação

De acordo com o gráfico, os adjetivos relacionais foram os principais responsáveis pela redução considerável no volume de dados a serem analisados. Essa redução fez com que obtivéssemos um menor número de combinações que configuravam candidatos a colocações, restringindo, assim, nosso objeto de estudo.

No entanto, apesar de nossa análise qualitativa ter possibilitado uma diminuição das combinações que poderiam constituir colocações, os resultados oferecidos não permitem uma identificação objetiva dessas estruturas léxicas, visto que, o fato de pertencer a um grupo

¹⁷¹ Quando um mesmo adjetivo apresentava diferentes tipos de relações de acordo com os substantivos que acompanhava, contabilizamos todas as suas ocorrências diferenciadas.

ou outro de nossa classificação não pode ser definitivo para considerar uma combinação como colocação ou não. Além disso, em grupos como os qualificadores, em que não foi possível estabelecer novas divisões, um número muito grande de adjetivos foi incluído, constituindo um conjunto bastante heterogêneo de vocábulos. Como exemplos dessa heterogeneidade, temos as combinações formadas por *resultado adverso* (UNIV06) [resultado adverso], *respeto mutuo* (MER05) [respeito mútuo], *vozes disonantes* (DIA01) [vozes dissonantes], entre outros. Em vista disso, até mesmo dentro de um único grupo, não é possível se falar em tendências de maneira conclusiva.

5.5 Limitações da análise qualitativa

Com base em um amplo número de adjetivos, aplicamos, no capítulo anterior, diferentes cálculos estatísticos que permitiram isolar algumas combinações de acordo com a coocorrência freqüente entre adjetivos e substantivos. A partir de tais combinações, no presente capítulo, passamos a investigar os tipos de relações semânticas mantidas entre os substantivos e os adjetivos, através de uma classificação destes últimos.

Essa análise, que chamamos de “qualitativa”, permitiu um reconhecimento mais aprofundado das combinações que, até aquele momento, mantinham uma relação apenas “estatística” entre seus membros constituintes. Em razão de nossa classificação, identificamos grupos de adjetivos que tendem a se combinar com os substantivos para formar colocações, como é o caso dos intensificadores e qualificadores, assim como aqueles que não parecem apresentar tal comportamento, como os relacionais e intensionais, por exemplo, cujas combinações originam unidades lexicalizadas ou combinações livres.

Apesar de ter possibilitado um refinamento de nossos dados, através de uma redução das combinações que se apresentam como candidatas a colocações, nossa análise qualitativa não permitiu uma delimitação precisa dos casos que configuravam esse fenômeno léxico. Tal deficiência deveu-se, por um lado, à ausência de uma taxonomia exaustiva, fundamentada lingüisticamente, que oferecesse um estudo mais aprofundado do comportamento dos adjetivos, e, por outro lado, à diversidade inerente a nosso próprio objeto de estudo, o qual, mesmo quando agrupado em categorias pequenas - como os adjetivos espaço-temporais -, não apresentava um comportamento uniforme.

Em vista disso, a execução dessa etapa de análise resultou, por um lado, em categorias bastante delimitadas, como os adjetivos intensificadores, os quais permitiram, ainda, uma divisão entre aqueles cuja intensificação é de tipo quantitativa e os que operam de forma qualitativa. No entanto, como as combinações agrupadas em ambas as categorias manifestaram uma tendência semelhante em participar de colocações, essa divisão, também presente em Koike (2000), não se mostrou fundamental. Por outro lado, o grupo formado pelos adjetivos qualificadores, devido à enorme quantidade de adjetivos incluídos, os quais mantinham as mais diversas relações com os substantivos - como de cor, de forma, de necessidade, de comparação, etc. -, não permitiu que se falasse em tendências de maneira mais específica.

Em virtude das limitações apresentadas, nossa análise das combinações a partir da classificação semântica dos adjetivos demonstrou-se uma etapa fundamental no refinamento dos dados, mas insuficiente na delimitação objetiva das colocações. Em razão disso, realizamos uma nova análise das combinações resultantes dessa etapa, através de seu estudo em contraste entre as línguas espanhola e portuguesa, o qual apresentamos no capítulo a seguir.

6 ANÁLISE CONTRASTIVA: RELAÇÕES ENTRE AS LÍNGUAS

No capítulo 4, ao realizarmos uma análise quantitativa das colocações, a partir da frequência de coaparição dos lexemas, percebemos que não era possível diferenciar, de forma objetiva, as colocações das combinações livres. Isso acontecia porque em ambas as estruturas havia uma coocorrência frequente dos elementos, mas não eram consideradas as relações semânticas estabelecidas entre eles. Em virtude desse fato, passamos, no capítulo 5, a uma análise qualitativa das colocações. Nesse momento, buscamos identificar o comportamento semântico das combinações a partir de uma classificação dos adjetivos. Dessa forma, foi possível reconhecer que alguns deles apresentavam uma maior tendência a se combinar com os substantivos para formar colocações, enquanto outros não manifestavam tal comportamento.

Essas perspectivas de análise, no entanto, embora tenham constituído etapas importantes para a identificação dessas combinações em nosso *corpus*, não permitiram isolar com precisão os casos que realmente configuravam colocações. Em virtude desse fato, consideramos necessária a realização de uma nova análise dos dados que restaram sem uma classificação específica. Para isso, buscaremos verificar o comportamento dessas combinações no contraste entre diferentes idiomas. Acreditamos que a incorporação dessa nova perspectiva de análise seja importante porque conferirá maior precisão aos resultados de nossa pesquisa.

6.1 Apresentação de uma nova perspectiva de análise

Segundo Marton (1977 apud BAHNS, 1993, p. 59), “a noção de sintagma convencional [*i.e.* colocação] encontra sua dimensão plena somente quando considerada contrastivamente, *i.e.*, quando a língua nativa do aprendiz e sua língua alvo são levadas em consideração”¹⁷². Isso ocorre porque as colocações correspondem a combinações

¹⁷² [the notion of conventional syntagm finds its full dimension only when it is considered contrastively, *i.e.* when the native language of the learner and his target language are taken into consideration].

idiossincráticas, ou seja, particulares a cada idioma¹⁷³. Assim, “visto que as colocações são combinações pré-fixadas pela tradição lingüística, deve-se esperar que as diferentes línguas apresentem diferenças na escolha dos colocados em relação a um substantivo sinônimo”¹⁷⁴ (KOIKE, 2000, p. 251). Essa divergência entre as línguas na seleção dos colocados é responsável, então, pela grande dificuldade do aprendizado das colocações para o estudante de uma língua estrangeira e, inclusive, para os tradutores.

Segundo Rundell (1999, p. 45), as colocações são estruturas intuitivas para os falantes nativos de uma língua, de forma que eles não têm consciência de seu uso, mas correspondem a combinações imprevisíveis para os aprendizes de uma língua estrangeira (L2). Isso explica porque apresentam tanta dificuldade para um falante não-nativo e revela que seu emprego adequado é um indicativo da alta fluência deste falante em L2. Para Cop (1991, p. 2776), o conhecimento das colocações representa a chave para a fluência na língua estrangeira e, quando utilizadas de forma equivocada, embora possam ser entendidas pelos falantes nativos, são motivos de riso ou, inclusive, de incômodos.

O problema das colocações fica, então, mais evidenciado no momento da produção de textos por parte de um falante não-nativo. Segundo Alonso Ramos (2001, p. 12), para produzir seu texto, o aprendiz inicia pela seleção da base e, em seguida, essa base exerce uma restrição semântica e também léxica sobre a escolha do colocado. Em decorrência do fato de que cada língua seleciona um colocado específico, o estudante encontrará maiores dificuldades em descobrir o elemento adequado para formar uma combinação não apenas aceitável, do ponto de vista semântico e sintático, mas também usual na comunidade lingüística de L2.

Calderón Campos (1994, p. 74) é ainda mais enfático em limitar a dificuldade das colocações a problemas de codificação ao afirmar que “as colocações constituem um problema de produção, mas nunca de compreensão, visto que se o estudante estrangeiro conhece o significado de cada um dos membros, não terá nenhum problema em compreender

¹⁷³ Opinião semelhante apresentam autores como Allerton (1984, p. 27), Benson (1986, p. 11), Wouden (1992, p. 449), entre outros. Uma oposição a esse argumento, no entanto, pode ser encontrada em Alonso Ramos (2002b, p. 151), para quem “o que caracteriza um colocado é sua escolha não livre e não seu caráter mais ou menos idiossincrático” [lo que caracteriza un colocativo es su elección no libre, y no su carácter más o menos idiosincrático]. Essa concepção, no entanto, que não considera o contraste entre as línguas para a delimitação das colocações, é responsável por um “inchaço” nos dicionários, através da inclusão de combinações, de certa forma, sem muita relevância para o consulente. Exemplos dessas combinações léxicas, retirados do DICE (2004), podem ser encontrados em 2.1.1.4.

¹⁷⁴ [dado que las colocaciones son combinaciones prefijadas por la tradición lingüística, cabe esperar que las distintas lenguas presenten diferencias en la selección de colocativos con respecto a un sustantivo sinónimo].

o significado completo do conjunto”¹⁷⁵. Koike (2000, p. 272), por outro lado, ao verificar que a transparência semântica das colocações nem sempre é total, acredita que essas combinações representam uma dificuldade tanto na produção quanto na compreensão de textos. Nesse sentido, Castillo Carballo (1998, p. 51-52) afirma que “existem colocações em que é necessário um conhecimento mais profundo da língua para compreender o sentido exato da combinação”¹⁷⁶.

Independentemente, porém, se a dificuldade oferecida pelas colocações dá-se, de forma mais relevante, na produção ou recepção de textos, o que constitui um consenso entre os diferentes autores é que essas combinações representam uma grande dificuldade para os aprendizes de uma língua estrangeira. Por isso, os estudos desenvolvidos a partir de um enfoque lexicográfico (v. 2.2) para esse tipo de combinações léxicas apresentam grande relevância. Com a finalidade de incluir as colocações mais problemáticas para os aprendizes em dicionários bilíngües, os autores desse enfoque verificaram a necessidade de serem realizados estudos contrastivos entre as línguas.

Em virtude dessas constatações, acreditamos que, além das análises quantitativa e qualitativa, anteriormente realizadas, é necessária a utilização de uma nova perspectiva para enfocar esse fenômeno léxico, a qual deve levar em consideração as diferenças e semelhanças existentes nas combinações entre as línguas. Dessa forma, para uma identificação mais precisa das colocações, consideramos importante a realização de uma análise contrastiva específica entre as línguas portuguesa e espanhola¹⁷⁷. Convém ressaltar, porém, que essa análise corresponde a um enfoque eminentemente fenomenológico¹⁷⁸. No entanto, em virtude das limitações apresentadas pelas análises anteriores (v. 4.3 e 5.5), o estudo contrastivo constitui uma última etapa na tentativa de isolar o fenômeno léxico das colocações. Temos consciência, porém, que, devido à vagueza na delimitação conceitual dessas combinações, fator que impede a formulação de instrumentos metodológicos mais apurados para a sua

¹⁷⁵ [las colocaciones constituyen un problema de producción, pero nunca de comprensión, puesto que si el estudiante extranjero conoce el significado de cada uno de los miembros, no tendrá ningún problema en comprender el significado completo del conjunto].

¹⁷⁶ [existen colocaciones en las que se necesita un conocimiento más profundo de la lengua para comprender el sentido exacto de la combinación].

¹⁷⁷ De acordo com Zimmer (1990, p. 128), “por meio do contraste entre línguas ficam mais claramente em evidência as discrepâncias das colocações” [Beim Sprachvergleich werden die erheblichen Diskrepanzen bei den Kollokationen offensichtlich]. Também para Herbst, Klotz (2003, p. 138) somente se pode estabelecer com total clareza que combinações são idiossincráticas e não previsíveis para um falante não-nativo quando tais combinações são contrastadas entre duas línguas.

¹⁷⁸ Por “fenomenológico”, referimo-nos a um estudo que visa à descrição de dados empíricos a partir da intuição do lingüista, com o intuito de apresentar observações sem, no entanto, se preocupar em oferecer uma explicação para os fenômenos descritos (cf. ABBAGNANO, 1998, s.v. *fenomenologia*; FERRATER MORA, 1994, s.v. *fenomenología* e *fenomenológico*).

detecção, mesmo a associação de diferentes perspectivas de análise pode não ser suficiente para uma identificação absolutamente precisa de todos os casos de colocações.

6.2 Fundamentos de estruturação de um modelo de análise contrastivo

Para a elaboração de um modelo de análise que permita estabelecer o contraste entre as diferentes combinações léxicas das línguas portuguesa e espanhola, consideramos necessário recorrer aos pressupostos da Lingüística Contrastiva. Através dos diferentes métodos de análise fornecidos por essa abordagem lingüística, quais sejam, Análise Contrastiva (v. 6.2.4.1), Análise de Erros (v. 6.2.4.2) e Interlíngua (v. 6.2.4.3) -, acreditamos que seja possível a obtenção de um modelo de análise que auxilie no processo de identificação das colocações entre o português e o espanhol.

Consideramos importante, porém, que o estudo contrastivo entre os dois idiomas leve em consideração o fato de que, segundo o “princípio do relativismo lingüístico” (v. 6.2.1), cada língua proporciona uma diferente segmentação da realidade extralingüística, o que ocasiona uma falta de equivalência absoluta entre as unidades de dois idiomas, de acordo com o “princípio do anisomorfismo lingüístico” (v. 6.2.2). Em vista disso, faz-se necessário investigar as diferentes formas de classificação das línguas (v. 6.2.3), a fim de identificar se a proximidade entre o português e o espanhol desempenha um papel relevante na conformação do fenômeno léxico das colocações.

A seguir, então, apresentamos alguns conceitos que consideramos fundamentais para a proposição de um modelo de análise que permita contrastar as combinações léxicas entre as línguas portuguesa e espanhola, com o intuito de delimitar os casos de colocações entre esses dois idiomas.

6.2.1 Princípio do relativismo lingüístico

A partir da concepção de Sapir de que “o ‘mundo real’ está ampla e inconscientemente formado pelos hábitos lingüísticos de um grupo determinado”¹⁷⁹ (Sapir apud WHORF, 1971, p. 155), Whorf (1971, p. 241), através de análises lingüísticas de um grande número de línguas, identificou que a gramática de cada língua (ou o “fundo de experiência”, em sua terminologia) não é um simples instrumento que reproduz as idéias, mas constitui seu verdadeiro formador. Assim, o processo de formação das idéias é dependente de cada gramática particular e, conseqüentemente, diferenciado de uma língua para outra. Em decorrência desse fato, Whorf (1971, p. 241) conclui que “dividimos a natureza seguindo linhas que são indicadas por nossas línguas nativas”¹⁸⁰. De acordo, então, com essa concepção da linguagem, cada língua segmenta a realidade extralingüística de uma maneira diferente¹⁸¹.

Também Humboldt (2006, p. 77) coincide com essa opinião ao afirmar que “as línguas na verdade não são meios para a representação da verdade conhecida, mas sim muito mais para a descoberta do anteriormente desconhecido. A sua diferença não reside nos sons e signos, mas na diferença de concepções de mundo em si”. Dessa forma, fica evidenciado que “percebemos o mundo através da nossa língua e os falantes de diferentes línguas percebem o mundo de maneira diferente” (HEIDERMANN, 2006, p. XLIII). Das múltiplas segmentações de uma mesma realidade extralingüística, surgem, então, as diferenças existentes entre as línguas.

¹⁷⁹ [el “mundo real” está amplia e inconscientemente conformado según los hábitos lingüísticos de un grupo determinado].

¹⁸⁰ [diseccionamos la naturaleza siguiendo líneas que nos vienen indicadas por nuestras lenguas nativas]. Também para Szende (1996, p. 111), “é a língua que cria as oposições no interior de uma realidade sem limites precisos” [c’est la langue qui crée des oppositions au sein d’une réalité sans limites précises].

¹⁸¹ Opondo-se a essa concepção do relativismo lingüístico, Pinker (2004), baseado em estudos cognitivos da linguagem, acredita que, embora haja diferenças no modo de falar entre as comunidades lingüísticas, não há como sustentar que existam também diferenças em seu modo de pensar. Segundo esse autor, Whorf (1971), ao exemplificar a forma diferente como falavam os apaches, “fez uma tradução deselegante, palavra por palavra das frases, de modo a fazer com que o sentido literal parecesse o mais estranho possível” (PINKER, 2004, p. 67). Além disso, ressalta o fato de Whorf (1971) ter realizado asserções sobre a psicologia apache baseado somente em estudos de sua gramática, o que torna sua argumentação circular: “Apaches falam de maneira diferente, portanto devem pensar de maneira diferente. Como sabemos que pensam de maneira diferente? Basta escutar a maneira como falam!” (PINKER, 2004, p. 66-67). A partir dessas constatações, Pinker (2004, p. 93) propõe que a realidade não seja segmentada de acordo com uma gramática particular a cada língua, mas que as “línguas do pensamento” parecem muito mais umas com as outras do que seus equivalentes falados, levando a crer que não difiram umas das outras, mas constituam um “mentalês universal”. As conclusões de Pinker (2004), no entanto, estão diretamente relacionadas ao funcionamento cognitivo da linguagem e não à sua estruturação enquanto sistema. Em vista disso, consideramos que essa concepção não representa um impedimento em nossa argumentação.

A partir dessas diferenças, desenvolve-se em Humboldt (2006, p. 103-105) a noção de “forma interna da língua”, segundo a qual cada língua contém um conjunto específico de particularidades lingüísticas que a caracterizam. Para Humboldt (2006, p. 105-107),

a forma característica das línguas impregna cada um de seus mais ínfimos elementos; e cada um deles é de alguma maneira por ela determinado, não importando quão imperceptível isso possa ser isoladamente. Por outro lado, é quase impossível encontrar pontos em relação aos quais se pudesse afirmar, quando considerados sozinhos, que carreguem decisivamente em si a forma característica das línguas.

Assim, ainda que se possam identificar as semelhanças existentes entre as línguas, a individualidade de cada uma prevalece e não é possível isolar seus traços específicos de forma a resumir suas particularidades, ou seja, não se pode delimitar com precisão os aspectos exatos em que duas línguas se diferenciam. Porém, embora a maior parte dos conceitos e também das características gramaticais estejam intimamente relacionados à individualidade de cada língua particular, é necessário encontrar um ponto intermediário entre as línguas, a fim de que se possa realizar um estudo contrastivo, tanto no plano gramatical quanto lexical (HUMBOLDT, 2006, p. 67-69).

Para estabelecer, então, esse contraste entre as combinações das línguas portuguesa e espanhola, é necessário considerar que a seleção léxica de cada idioma está intimamente relacionada com a visão de mundo de cada língua. Dessa forma, um vocábulo correspondente à base da colocação, para exprimir um mesmo significado, pode selecionar um colocado específico para cada uma das línguas em questão. Além disso, algumas colocações utilizadas em um idioma podem não encontrar correspondência no outro, devido a um recorte diferenciado que cada comunidade lingüística realizou de uma mesma realidade.

6.2.2 Princípio do anisomorfismo lingüístico

Segundo Neubert (1992, p. 29), “é uma obviedade antiga quando especialistas e leigos igualmente afirmam que as palavras de duas línguas nunca, ou ao menos raramente, são as mesmas”¹⁸². Essa falta de equivalência entre as unidades lexicais de línguas diferentes pode ser explicada através do conceito de “anisomorfismo lingüístico”, introduzido por Zgusta

¹⁸² [it is an obvious and age-old truism when experts and laymen alike assert that the words of two languages never or at least rarely ever are the same].

(1971, p. 294). Para esse autor, existe uma diferença na organização das designações em cada língua particular, o que seria responsável, então, pela ausência de uma coordenação entre as palavras de cada idioma¹⁸³. Dessa forma, salienta que cada unidade léxica apresenta determinadas características que lhe são específicas e que não encontram uma correspondência absoluta com as unidades léxicas de outra língua. Em virtude desse fato, Neubert (1992, p. 40), para solucionar as dificuldades provocadas pelo anisomorfismo no plano lexicográfico, propõe que as relações contrastivas entre a rede lexical de duas línguas sejam estabelecidas através de correspondências prototípicas.

Zgusta (1971, p. 295) adverte, porém, que as dificuldades originadas pelo anisomorfismo lingüístico não devem ser limitadas a questões de diferenças culturais, ou seja, a casos de *realia*¹⁸⁴, mas “deve ser esperado em todas as unidades léxicas e pode ser encontrado na maioria delas”¹⁸⁵ (ZGUSTA, 1971, p. 296). Assim, embora Zgusta (1971) aplique o conceito do anisomorfismo somente ao plano do conteúdo, essa ausência de uma equivalência entre os idiomas pode ser encontrada em diferentes níveis de organização de uma língua natural, tais como:

- a) **fonético-fonológico**, em que, enquanto a língua portuguesa apresenta vogais abertas e fechadas, como nos exemplos de *avó* e *avô*, o espanhol possui somente vogais fechadas.
- b) **morfo sintático**, em que os verbos transitivos indiretos da língua portuguesa não encontram correspondência em espanhol. Nesses casos, como no exemplo *eu gosto de maçãs*, a língua espanhola utiliza uma construção com o verbo intransitivo: *me gustan las manzanas*.
- c) **léxico**, representado, especialmente, pelos casos de falsos amigos¹⁸⁶, como *asignatura* e *firma*, em língua espanhola, que significam, respectivamente, “disciplina” e “assinatura” em língua portuguesa.

¹⁸³ Também Zimmer (1990, p. 65) acredita que “a equivalência 1-1 é sem dúvida uma ocorrência rara” [Bei der 1:1-Äquivalenz handelt es sich zweifellos um eine seltene Erscheinung].

¹⁸⁴ Segundo Dubois et al (1999, s.v. *realia*), os “*realia*” correspondem a termos que designam “uma realidade particular a determinada cultura e que são utilizados tal qual na língua [*sc.* original]” [une réalité particulière à telle ou telle culture et qui sont utilisés tels quels dans la langue]. Como exemplo de *realia* da língua espanhola, temos a palavra *chirimoya*, que corresponde a uma fruta ovalada, de sabor doce, com casca verde e polpa branca repleta de sementes pretas. Um exemplo de *realia* da língua portuguesa, por sua vez, seria o vocábulo *caipirinha*, que não encontra uma equivalência em outros idiomas.

¹⁸⁵ [must be expected in all lexical units and can be found in most of them].

¹⁸⁶ Entendemos por “falsos amigos” o par de vocábulos de duas línguas diferentes que apresentam uma estrutura fonológica idêntica ou semelhante, mas possuem uma significação diferente (cf. BENEDUZI, BUGUEÑO, 2006, p. 2). Para as diferenças entre essa denominação e outros termos relacionados, v. Bugueño (2002).

d) **pragmático**, em que, por exemplo, na língua portuguesa utilizamos a expressão *pois não* para dizer que sim, o que provoca uma interpretação equivocada por parte de falantes nativos do espanhol.

No caso específico das colocações, deve-se destacar que, além do problema do anisomorfismo, ou seja, de uma ausência de equivalência entre as línguas, existe também a dificuldade provocada pela seleção idiossincrática dos elementos. Isso ocorre porque, ainda que os dois idiomas possuam unidades léxicas equivalentes, no nível do sistema (na concepção Coseriana), cada língua pode selecionar, no nível da norma¹⁸⁷, colocados específicos para se combinarem com determinada base. Dessa forma, quando comparadas entre dois idiomas, as colocações podem representar uma divergência tanto no nível do sistema quanto no nível da norma. Em outras palavras, ao analisar as colocações em contraste, deparamo-nos, primeiramente, com uma falta de equivalência entre as línguas no nível do sistema, devido a sua natureza anisomórfica, e, em seguida, pode ocorrer também uma nova divergência, no nível da norma, decorrente do fato de cada língua particular selecionar unidades léxicas específicas para formar uma colocação. Isso faz com que a identificação de equivalências para essas combinações seja ainda mais complicada do que nos casos que configuram ocorrências somente de anisomorfismo.

Para ilustrar nosso argumento, tomemos como exemplo o uso das preposições em português e alemão nos seguintes enunciados equivalentes:

↳ “As crianças brincam na rua”

↳ “Die Kinder spielen auf der Straße”

A língua alemã, embora também possua a preposição *in* [em] - a primeira vista um equivalente total da preposição *em* do português -, utiliza em exemplos como o acima citado a preposição *auf* [sobre], caracterizando um caso de anisomorfismo léxico entre os dois idiomas. Isso acontece porque “em alemão, as preposições locais requerem um uso específico,

¹⁸⁷ Para Coseriu (1973, p. 94), a “norma” corresponde a um primeiro nível de abstração, constituído pelas estruturas “normais e tradicionais na comunidade” [normales y tradicionales en la comunidad], enquanto o “sistema” corresponde a um plano mais alto de abstração, no qual se depreende, de cada estrutura particular, “uma série de elementos essenciais e indispensáveis, de oposições funcionais” [una serie de elementos esenciales e indispensables, de oposiciones funcionales]. Dessa forma, nas palavras de Coseriu (1973, p. 98, grifo do autor), “o sistema é sistema de possibilidades, de coordenadas que indicam caminhos abertos e caminhos fechados. (...) O que na realidade impõe-se ao indivíduo, limitando sua liberdade expressiva e reduzindo as possibilidades oferecidas pelo sistema dentro do marco fixado pelas realizações tradicionais, é a norma. A norma é, de fato, um sistema de realizações obrigadas, de imposições sociais e culturais, e varia segundo a comunidade” [el sistema es un sistema de posibilidades, de coordenadas que indican caminos abiertos y caminos cerrados. (...) Lo que, en realidad, se impone al individuo, limitando su libertad expresiva y comprimiendo las posibilidades ofrecidas por el sistema dentro del marco fijado por las realizaciones tradicionales, es la norma. La norma es, en efecto, un sistema de realizaciones obligadas, de imposiciones sociales y culturales, y varía según la comunidad].

isto é, elas indicam claramente a relação local que deve ser estabelecida entre os dois elementos em questão: o que nem sempre ocorre em português” (SIMÕES, 2006, p. 33). Assim, na língua alemã, enquanto a preposição *in* expressa a idéia de “estar dentro de”, a preposição *auf* estabelece a relação de que algo “está em cima de” alguma coisa. Ainda que essa divergência entre as línguas portuguesa e alemã possa causar dificuldades nos primeiros níveis de aprendizagem, um aluno com maior proficiência em alemão já será capaz de recuperar a explicação para o fato de essa língua ter selecionado a preposição *auf*.

No entanto, nem todos os elementos das línguas possuem uma razão aparente para sua ocorrência, como no caso das colocações, em que não existe essa possibilidade de recuperar uma explicação para o fato de cada língua selecionar uma combinação, uma vez que esse processo é arbitrário. Assim, não são seguidas regras de organização internas à gramática de cada língua, como acontecia no exemplo anterior de anisomorfismo, o que pode ser uma explicação para o fato das colocações causarem tantas dificuldades para os estudantes de uma língua estrangeira.

6.2.3 Classificação das línguas

De acordo com Bussmann (1990, s.v. *Klassifikation der Sprachen*), as línguas podem ser classificadas a partir de três critérios diferentes:

- a) classificação **geográfica** ou **areal**: realizada a partir das semelhanças entre as línguas por sua proximidade geográfica ou cultural;
- b) classificação **genealógica** ou **genética**: orientada por critérios historicistas, agrupa as línguas a partir de um antepassado comum (TRASK, 2006, s.v. *relação genética*);
- c) classificação **tipológica**: baseada nas semelhanças estruturais entre as línguas, sem considerar fatores geográficos ou genealógicos. No entanto, apesar de uma classificação tipológica ser independente da história das línguas em questão, as línguas com parentesco próximo serão, naturalmente, bastante parecidas tipologicamente (TRASK, 2006, s.v. *tipologia*).

As línguas portuguesa e espanhola apresentam, portanto, semelhanças tanto do ponto de vista genético quanto tipológico. Segundo Rodrigues (s.d. apud FIALHO, 2005, s.p.), o espanhol e o português são o par de línguas latinas mais próximo. No entanto, essa

proximidade histórica e estrutural entre os dois idiomas não parece ser um facilitador do processo de aquisição da língua estrangeira. Segundo Alvarez (2002, s.p.), “a proximidade entre ambas cria o que nós chamamos de benefício no início da aprendizagem, mas que nos estágios mais avançados torna-se uma dificuldade”. Calvi (2002, p. 129), por sua vez, acredita que a semelhança entre duas línguas pode facilitar, especialmente, no processo de compreensão de textos. No entanto, essa compreensão nem sempre é satisfatória. Segundo a autora, ela “pode ser obstaculizada pelas enganadoras semelhanças, sobretudo no caso de sinónimas parciais”¹⁸⁸ (CALVI, 2002, p. 129). Tratando especificamente das colocações, que representam uma dificuldade maior para os estudantes no momento da produção de textos (v. 6.1), a semelhança entre as línguas, então, não desempenharia um papel relevante na facilitação do aprendizado. Além disso, como essas combinações correspondem a estruturas idiossincráticas, cada língua seleciona um colocado específico para se combinar com determinado vocábulo, o que independe da proximidade dos idiomas que estão em contraste, dificultando, assim, a seleção léxica adequada por parte dos aprendizes¹⁸⁹. Em vista disso, consideramos que a semelhança entre as línguas portuguesa e espanhola não será um fator relevante para auxiliar o estudante na escolha do colocado adequado e, conseqüentemente, facilitar sua produção na língua estrangeira.

Pelo contrário, a proximidade entre as línguas pode levar os aprendizes a transferir as estruturas de sua língua materna para a língua estrangeira. Assim, por exemplo, segundo Hussein (1990, p. 123),

aprendizes de inglês como língua estrangeira cometem erros ao combinar palavras simplesmente porque tendem a unir palavras que são semanticamente compatíveis, mas infelizmente a união de palavras semanticamente compatíveis nem sempre produz uma coocorrência aceitável¹⁹⁰.

Dessa forma, ressalta que o uso adequado das colocações exige uma associação específica entre a base e o colocado, de forma que o estudante não pode simplesmente utilizar as mesmas estruturas de sua língua nativa. A simples transposição das combinações de uma língua para outra é responsável pela maior parte dos erros cometidos pelos estudantes no aprendizado das colocações (BAHNS, 1993, p. 61).

¹⁸⁸ [può essere ostacolata dalle ingannevoli somiglianze, soprattutto nel caso di sinonimie parziali].

¹⁸⁹ Opinião semelhante apresentam Aguilar-Amat, Tuells (1994, p. 329), para quem “é idiossincrática [sc. uma colocação] quando efetuamos uma comparação com outras línguas, ainda que sejam muito próximas” [resulta ser idiosincrática cuando efectuamos una comparación con otras lenguas, aún cuando sean muy cercanas].

¹⁹⁰ [EFL learners commit errors in collocating words simply because they tend to join words which are semantically compatible, but unfortunately joining words which are semantically compatible does not always produce an acceptable co-occurrence].

Em vista disso, no caso específico das combinações entre o português e o espanhol, ainda que essas duas línguas apresentem vocábulos equivalentes (por sua proximidade tanto genética quanto tipológica), eles podem não ser igualmente utilizados para formar uma colocação com uma base específica. A semelhança entre os dois idiomas pode, então, causar equívocos por parte dos estudantes, que tendem a realizar uma simples transposição do elemento equivalente entre as línguas, sem uma preocupação de se essa é a combinação realmente utilizada na língua estrangeira. Em vista disso, consideramos que um estudo contrastivo das combinações entre as línguas portuguesa e espanhola seja fundamental para identificar as construções que podem apresentar uma maior dificuldade para os aprendizes, pois de acordo com Calvi (2002, p. 128), “a exigência de contrastividade é ainda mais evidente no caso de línguas afins”¹⁹¹.

6.2.4 Transferência entre as línguas

No âmbito da Lingüística Contrastiva (LC)¹⁹², através da comparação entre diferentes idiomas, Robert Lado foi quem primeiro identificou que as semelhanças e diferenças entre as línguas poderiam ser responsáveis por maiores facilidades ou dificuldades na aprendizagem. Para esse autor,

as estruturas que forem similares serão fáceis de aprender porque serão transferidas e poderão funcionar satisfatoriamente na língua estrangeira. As estruturas que forem diferentes serão difíceis porque, quando transferidas, não funcionarão satisfatoriamente na língua estrangeira e terão, portanto, de ser modificadas (LADO, 1972, p. 87).

A fim, então, de averiguar o papel desempenhado pela transferência lingüística no processo de aquisição de uma língua estrangeira, especialmente no que concerne ao problema das colocações, consideramos necessário recuperar o percurso transcorrido pela LC. Para isso, apresentamos os principais estudos desenvolvidos a partir dessa abordagem lingüística, de forma a contextualizar as diferentes perspectivas de análise e comparação de idiomas.

¹⁹¹ [l'esigenza di contrastività è ancor più marcata nel caso di lingue affini].

¹⁹² De acordo com Vandresen (1988, p. 75), a Lingüística Contrastiva corresponde a uma subárea da Lingüística Geral, cujo interesse é apontar semelhanças e diferenças estruturais entre a língua materna de um grupo de alunos e a língua estrangeira que estão aprendendo. Seu objetivo é identificar as estruturas que oferecerão maiores dificuldades na aprendizagem bem como as que, devido à similaridade com a língua materna, apresentarão facilidades.

6.2.4.1 *Análise Contrastiva*

Entre as décadas de 1950 e 1960, a partir das concepções de Lado (1957)¹⁹³, desenvolvem-se os estudos baseados na Análise Contrastiva (doravante AC) de diferentes idiomas. A AC fundamentava-se, do ponto de vista psicológico, no modelo behaviorista¹⁹⁴ e, do ponto de vista lingüístico, no Estruturalismo. A partir dos pressupostos do Behaviorismo, por um lado, a AC considerava a aquisição de uma língua estrangeira como a incorporação de um conjunto de novos hábitos lingüísticos, os quais sofriam interferência dos hábitos desenvolvidos na língua materna (L1). Por outro lado, baseada no Estruturalismo, a AC propunha uma comparação das estruturas em estudo entre L1 e L2, a fim de prever as possíveis dificuldades dos aprendizes. Assim, Vandresen (1988, p. 75) define a AC como “uma comparação do sistema lingüístico da LM [*i.e.* língua materna] com o sistema da LE [*i.e.* língua estrangeira] destacando seus pontos de semelhança e diferença e listando os erros prováveis que os alunos cometerão em sala de aula”. Gass, Selinker (1994, p. 59) acrescentam que a identificação dos erros por parte da AC objetiva determinar o que precisa e o que não precisa ser aprendido no ensino de uma L2.

Dessa forma, a AC fundamenta-se, principalmente, no pressuposto de que o processo de aquisição de L2 é diferente da aquisição de L1, o que atualmente constitui uma verdade incontestável. A diferença no processo de aquisição deve-se, segundo essa teoria, à incorporação de novos hábitos lingüísticos, os quais encontram influência na L1¹⁹⁵. Assim, para a AC, a L1, devido a sua interferência, constitui a principal fonte de erros por parte dos aprendizes e esses erros devem-se, especialmente, às diferenças entre L1 e L2. Conseqüentemente, ainda segundo essa teoria, quanto maior a diferença entre as línguas, mais serão os erros cometidos pelos aprendizes e, em virtude dessas dificuldades, a AC considera

¹⁹³ Não tivemos acesso a essa obra original, mas somente a seu exemplar traduzido para o português, correspondente a Lado (1972).

¹⁹⁴ O Behaviorismo (ou Comportamentalismo) corresponde a uma teoria psicológica que explica os fenômenos lingüísticos através da análise de seus comportamentos observáveis e de suas respostas, produzidas a partir de estímulos, às situações (DUBOIS et al, 1999, s.v. *béaviourisme*). No âmbito estritamente lingüístico, as concepções behavioristas foram introduzidas por Skinner, para quem “falar é unir palavras e estruturas de modo seqüencial em resposta a estímulos” (DURÃO, 2004, p. 22) [hablar es unir palabras y estructuras de modo secuencial en respuesta a estímulos]. De acordo com essa perspectiva, as crianças são vistas como tábulas rasas que precisam ser preenchidas com estímulos externos.

¹⁹⁵ Lado (1972, p. 87) salienta que “visto que o aprendiz tende a transferir os hábitos da estrutura de sua língua nativa para a língua estrangeira, temos aqui a maior fonte de dificuldade ou facilidade na aprendizagem de uma língua estrangeira”.

que o que precisa ser ensinado são justamente as diferenças entre as línguas. (cf. GASS, SELINKER, 1994, p. 60; RÔA, 1983, p. 116-117).

No entanto, conforme Durão (2004, p. 36-37), nem sempre as estruturas divergentes entre as línguas serão causa de dificuldades para os aprendizes de L2. Além disso, muitos dos erros cometidos pelos aprendizes não apresentam uma origem interlingual, mas correspondem a erros intralinguais, ou seja, não se relacionam com a estrutura da L1, mas decorrem das irregularidades da própria L2. Esses erros podem ser constatados pela dificuldade dos próprios falantes nativos ao aprender determinadas estruturas de sua língua¹⁹⁶. Gass, Selinker (1994, p. 64) destacam que, além da interferência, existem outros fatores que influenciam no processo de aquisição de L2: princípios inatos da linguagem, atitude, motivação, aptidão, idade e conhecimento de outras línguas. Lin (2003, p. 33) acrescenta, ainda, a esses fatores as dificuldades provocadas por estratégias equivocadas de ensino-aprendizagem.

Além dessa inconsistência, inúmeras críticas ao modelo proposto pela AC surgiram durante a década de 70. A principal delas estava relacionada à previsão de erros que não eram cometidos pelos alunos, ao passo que diversos erros que efetivamente ocorriam não eram previstos pela AC (cf. JAMES, 1998, p. 4). Embora Lado (1983, p. 19, grifo nosso) tenha advertido que os problemas resultantes da comparação entre L1 e L2 “devem ser considerados uma lista de problemas hipotéticos até que a validação final seja obtida pela sua verificação com a fala atual dos estudantes”¹⁹⁷, muitos autores tomaram as divergências entre L1 e L2 como regras para a previsão de erros.

Outra crítica constantemente dirigida à AC é que seus estudos centram-se em evitar os erros, sem se preocupar com o desenvolvimento da comunicação. Fialho (2005, s.p.) afirma que, com a metodologia proposta pela AC, “o papel central do professor era o de corrigir seu aluno e não o de promover a interação entre os membros da sala de aula”. Ainda com relação à aplicação da AC ao ensino, Lin (2003, p. 41) acredita que as previsões de erros proporcionadas por esse modelo de análise não vão além dos conhecimentos intuitivos adquiridos pelo professor através de sua experiência docente.

Finalmente, com o advento do Gerativismo, houve uma mudança de concepção no processo de aquisição da linguagem. Segundo Gass, Selinker (1994, p. 61),

¹⁹⁶ Segundo Lin (2003, p. 34), “muitos dos erros atribuídos à interferência são observados também no processo de aquisição da língua materna” [muchos de los errores que se atribuyen a la interferencia, se observan también en el proceso de adquisición de la lengua materna].

¹⁹⁷ [must be considered a list of hypothetical problems until final validation is achieved by checking it against the actual speech of students].

a língua passou a ser vista não como um conjunto de hábitos automáticos, mas como um conjunto de regras estruturadas. Essas regras são aprendidas não por imitação, mas por sua formulação ativa com base em princípios inatos, bem como pela exposição à língua que está sendo aprendida¹⁹⁸.

Dessa forma, o processo de aquisição de L2 não se baseia mais no princípio de estímulo e resposta, mas corresponde a um processo criativo e dedutivo, no qual o aprendiz, através da formulação de hipóteses, começa a se aproximar gradativamente da L2. Para Lin (2003, p. 36), “o que um aluno da segunda língua aprende não é um conjunto de habilidades ou regras novas, mas ele adapta, remodela e amplia as habilidades e conhecimentos já existentes”¹⁹⁹.

O modelo de AC, da forma como estava concebido, exigia, de acordo com Wardhaugh (1983, p. 8), o conhecimento de uma série de universais lingüísticos teoricamente fundamentados que permitissem a comparação das estruturas de cada par de línguas em análise nos diferentes níveis de organização da linguagem (sintaxe, semântica, fonologia, etc.). Esse conhecimento, no entanto, não estava disponível aos lingüistas, o que tornava impossível a utilização desse modelo (WARDHAUGH, 1983, p. 13). Em virtude desse fato, o modelo de AC recebeu uma nova versão, considerada “fraca” [*weak version*] em relação à anterior, denominada “forte” [*strong version*]. A versão “fraca” requeria somente a utilização do conhecimento lingüístico disponível para a observação e descrição das dificuldades provocadas pelas semelhanças e diferenças entre sistemas, sem o objetivo de oferecer previsões sobre possíveis dificuldades. Dessa forma, segundo Wardhaugh (1983, p. 13), demonstrava ser de maior auxílio para o estabelecimento do contraste entre as línguas.

6.2.4.2 Análise de Erros

A Análise de Erros (doravante AE) surgiu a partir do desenvolvimento da versão “fraca” do modelo de AC²⁰⁰, como uma forma de diminuir as lacunas presentes nesta última.

¹⁹⁸ [language came to be seen not as a set of automatic habits, but as a set of structured rules. These rules are learned not by imitation, but by actively formulating them on the basis of innate principles as well as on the basis of exposure to the language being learned].

¹⁹⁹ [lo que un alumno de la segunda lengua aprende no es un conjunto de habilidades o reglas nuevas, sino que adapta, remodela y amplia las habilidades y conocimientos ya existentes].

²⁰⁰ Alguns autores como Fialho (2005) e Vandresen (1988) consideram a AE como um sinônimo de AC em sua versão “fraca”. Acreditamos, no entanto, que essas teorias sejam diferentes em dois aspectos principais: em primeiro lugar, a AC (versão “fraca”) não trabalhava com dados empíricos, em oposição à AE, cuja fonte de dados corresponde aos erros produzidos pelos aprendizes. Em segundo lugar, no que concerne à comparação

A AE é definida como um tipo de análise lingüística que enfoca os erros cometidos pelos aprendizes (cf. GASS, SELINKER, 1994, p. 67), com o intuito de determinar sua incidência, natureza, causas e conseqüências (JAMES, 1998, p. 1). Diferentemente da concepção utilizada na AC, os erros não são vistos como um problema a ser corrigido, mas constituem uma parte necessária do processo de aprendizagem (cf. PIT CORDER, 1981, p. 66). Na AE, os erros podem ocorrer devido, além da interferência da L1, a generalizações intralingüísticas, ou seja, de regras já dominadas, bem como corresponder a erros não-sistemáticos devido ao esquecimento. Fialho (2005, s.p.) salienta que, “ainda assim, os erros devem ser avaliados respeitando variáveis como a idade do aluno, tensão nervosa, metodologia de ensino, motivação e interesse”.

Devido às diferentes variáveis envolvidas na produção de erros por parte do aprendiz, Pit Corder (1967 apud GASS, SELINKER, 1994, p. 67) propõe a distinção entre “erro” [*error*] e “engano” [*mistake*]. Para esse autor, um “engano” corresponde a um deslize pontual, que pode ser reconhecido e corrigido pelo próprio estudante, enquanto o “erro” caracteriza-se por ser sistemático, recorrente e não reconhecível como um erro pelo aprendiz. No entanto, ainda que seja utilizada uma definição precisa do que se entende por “erro”, uma crítica constante à AE diz respeito à subjetividade na interpretação dessas incorreções, visto que a determinação do que é ou não um erro vai variar de acordo com a concepção de cada lingüista. Somente em Durão (2004, p. 51-53), por exemplo, encontramos seis definições diferentes para o conceito de “erro”. Além dessa limitação, James (1998, p. 17) aponta a falta de inferência estatística e a ausência de poder preditivo como outras deficiências do modelo de AE.

Pit Corder (1981, p. 59-60) ressalta também que, ao analisar os erros cometidos pelos aprendizes, a AE não utiliza uma linguagem espontânea em seu meio natural de produção, pois os dados analisados partem de exercícios restritos a tópicos, funções e tempo de execução determinados. Além disso, quando estão produzindo para serem avaliados, os alunos vão preferir utilizar as estruturas de que dispõem um maior conhecimento. Dessa forma, os dados disponíveis para análise podem não corresponder às dificuldades realmente apresentadas pelos aprendizes.

Finalmente, Gass, Selinker (1994, p. 68) acreditam que o foco da AE, centrada apenas nos erros dos aprendizes, sem considerar a totalidade das produções lingüísticas, seja

estabelecida entre as línguas, a AC (versão “fraca”) realiza a comparação entre L1 e L2, enquanto que a AE não considera a L1, mas apenas os erros produzidos na Interlíngua (v. 6.2.4.3) do aluno com relação à L2.

uma das maiores deficiências desse modelo de análise²⁰¹. Dessa forma, ressaltam que, apesar da importância da AE em reconhecer que o aprendiz não é apenas um participante passivo no processo de aquisição de L2, esse método de análise falha ao obter uma descrição apenas parcial da produção em L2.

A partir do acima exposto, podemos perceber que a AE diferencia-se da AC em diversos aspectos. Segundo Durão (2004, p. 53-54), tais aspectos consistem em:

- a) teorias que subjazem a cada perspectiva de análise: Na AC, identificamos a influência do Behaviorismo e do Estruturalismo, enquanto a AE fundamenta-se no Gerativismo;
- b) identificação dos erros: Na AC, os erros são previstos através da comparação de L1 e L2, ao passo que na AE são analisados somente os erros que realmente aparecem na produção dos estudantes;
- c) concepção de erro: Na AC, os erros são vistos unicamente como interferências da L1 na aquisição da L2; já na AE, os erros correspondem a evidências do processo de aprendizado do estudante.

Apesar, no entanto, das diferenças entre esses dois métodos de análise, Yavas (1980, p. 117) adverte que a utilização da AE e da AC não se excluem mutuamente, mas, pelo contrário, são complementares.

6.2.4.3 Interlíngua

Os estudos da Interlíngua (doravante IL) surgem como uma evolução da AC e da AE (DURÃO, 2004, p. 60). O conceito de IL foi introduzido por Selinker, que a define como “um sistema lingüístico separado baseado no *output* observável que resulta de uma tentativa de produção de uma norma da L2 por parte do aprendiz”²⁰² (SELINKER, 1983, p. 176). Esse autor ressalta, assim, que as construções utilizadas pelos aprendizes não se equivalem às

²⁰¹ Opondo-se a essa visão, Lin (2003, p. 51) afirma que “a análise de erros não leva em conta somente as realizações errôneas, mas também as corretas, para conseguir uma descrição da totalidade da língua do estudante a fim de beneficiar o ensino da segunda língua” [el análisis de errores no sólo tiene en cuenta las realizaciones erróneas, sino también las correctas, para lograr una descripción de la totalidad de la lengua del estudiante con el fin de beneficiarse de la enseñanza de la segunda lengua].

²⁰² [a separate linguistic system based on the observable output which results from a learner’s attempted production of a TL norm].

produzidas por falantes nativos da L2. Dessa forma, a IL corresponde a um sistema lingüístico particular dos aprendizes, o qual está situado entre a L1 e a L2.

A concepção de IL fundamenta-se no pressuposto de que o estudante, no decorrer de sua aprendizagem, vai criando hipóteses sobre a estrutura da L2. A partir de sua interação, tanto na produção quanto na recepção, esse aprendiz começa a testar suas hipóteses, o que garante a evolução de sua IL²⁰³. Assim, a IL constitui um *continuum* dinâmico, pois é identificada em diferentes etapas do processo de aprendizagem de uma L2 com uma constante evolução²⁰⁴.

A principal característica atribuída ao modelo de IL corresponde ao conceito de “fossilização”. Para Pit Corder (1981, p. 73), quando o aprendiz atinge um estágio de conhecimento da L2 que lhe permite comunicar-se adequadamente para seus propósitos, não sente mais a necessidade de continuar ampliando seu conhecimento. Nesse momento, inicia-se o processo de “fossilização”, que corresponde a “itens lingüísticos, regras, e subsistemas que o falante de uma L1 particular tenderá a manter em sua IL relativa a uma L2 particular, independentemente da idade do aprendiz ou da quantidade de explicações e instruções que ele recebe na L2”²⁰⁵ (SELINKER, 1983, p. 177). De acordo com Alvarez (2002, s.p.), no par de línguas português e espanhol, apesar de sua proximidade, quando o nível de complexidade aumenta, a tendência dos estudantes é “cometer erros que podem se tornar fossilizáveis dentro da interlíngua criada pelo aprendiz”.

Os problemas encontrados no desenvolvimento da IL devem-se não somente à interferência da L1, pois outras línguas já conhecidas pelo aprendiz desempenham um papel importante na conformação de sua IL (PIT CORDER, 1981, p. 72). A partir dessas constatações, os estudos de IL evidenciam uma menor participação da L1 no processo de aquisição de L2, pois, segundo Pit Corder (1981, p. 89), “a interlíngua do aprendiz, seja qual for sua língua-mãe, em determinadas circunstâncias passa por uma seqüência semelhante de desenvolvimento pelo menos nos estágios iniciais”²⁰⁶. Dessa forma, o modelo de IL, assim como o de AE, diferencia-se da AC no que concerne à importância dada à L1 na aquisição da L2. O processo de identificação das dificuldades dos aprendizes é realizado na AC através da

²⁰³ Alvarez (2002, s.p.) salienta que “teorias lingüísticas de [sic] da interlíngua sugerem que a testagem de hipóteses por si só não é suficiente para explicar seu desenvolvimento”.

²⁰⁴ Essas diferentes etapas deverão ser levadas em consideração no momento de traçar o perfil do usuário de um dicionário determinado, conforme retomaremos em 7.1.2.

²⁰⁵ [linguistic items, rules, and subsystems which speakers of a particular NL will tend to keep in their IL relative to a particular TL, no matter what the age of the learner or amount of explanation and instruction he receives in the TL].

²⁰⁶ [the interlanguage of learner, whatever their mother tongue, in certain circumstances does go through a similar sequence of development at least in the earlier stages].

comparação dos sistemas de L1 e L2; na AE, parte-se da comparação dos erros produzidos pelos aprendizes com a norma da L2; já no modelo de IL, a comparação é realizada entre o *input* e o *output* do aluno, a fim de verificar as estruturas aprendidas nas quais ele apresenta maiores dificuldades.

Durão (2004, p. 70-71) destaca, ainda, outras diferenças fundamentais entre AC, AE e IL:

1. Para a AC, os erros são vistos como desvios que devem ser erradicados; na AE, sinalizam as dificuldades dos aprendizes, enquanto no modelo de IL, os erros constituem uma característica intrínseca da IL, que permite identificar o nível de conhecimento do aprendiz.
2. No modelo de AC, a L1 é a única fonte responsável pelos erros; na AE são consideradas múltiplas fontes, enquanto na IL são incorporadas essas duas concepções. A IL diferencia-se, ainda, da AE por procurar identificar o desempenho dos aprendizes como um todo, sem focar somente o erro.
3. Na AC e na AE, o importante é o produto da comunicação, já no modelo de IL também são consideradas as estratégias utilizadas pelos aprendizes.
4. Finalmente, o modelo de IL surge como uma combinação da AC e da AE. A primeira auxilia na testagem inicial de hipóteses e a segunda nas técnicas para analisar os erros produzidos pelos estudantes.

6.3 Uma proposta de análise contrastiva para a identificação das colocações

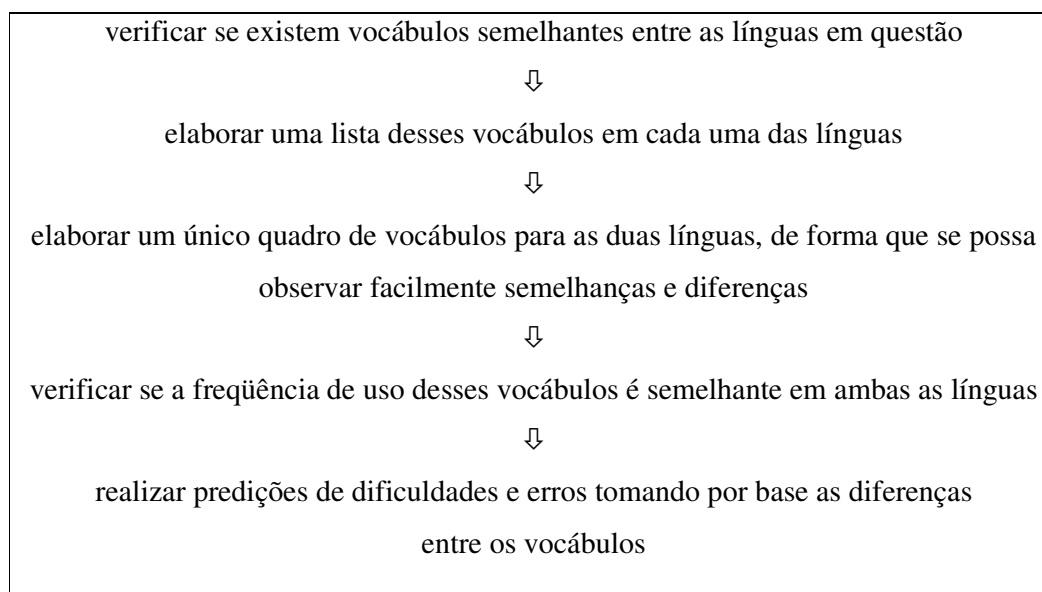
Como não partimos, em nossa pesquisa, de erros produzidos por estudantes, mas realizamos a coleta de nossos dados a partir de um *corpus* textual, a AC parece ser o modelo mais adequado para embasar nosso trabalho. No entanto, como os diferentes métodos de análise - AC, AE e IL - não são excludentes, mas complementares, utilizaremos também os pressupostos teóricos fornecidos pela AE e pela IL. A primeira será importante por sua concepção de erro, como parte do processo evolutivo de aprendizagem do estudante, e pela proposição de estratégias para suplantá-los, nos dicionários, os erros cometidos pelos aprendizes. Já o modelo de IL contribuirá com uma identificação das dificuldades dos estudantes nas diferentes etapas de seu processo de aprendizagem.

Tomando, então, como base os pressupostos acima apresentados, acreditamos ser possível a realização de um estudo contrastivo entre as línguas portuguesa e espanhola com vistas à identificação de colocações. Antes, porém, de proceder ao contraste entre as combinações léxicas desses dois idiomas, consideramos pertinente ressaltar que, apesar de a literatura especializada distinguir entre uma versão “forte” e uma versão “fraca” da AC, neste trabalho não entraremos no mérito dessa discussão. Salientamos apenas, recobrando a idéia original de Lado (1957), que a metodologia de AC aqui empregada pretende oferecer soluções de equivalência colocacional nos casos em que exista uma divergência entre a língua portuguesa (língua materna do aprendiz) e a língua espanhola (língua-alvo), ou seja, nos casos em que os mecanismos próprios da interlíngua possam levar o aprendiz à escolha do colocado equivocado. Em vista disso, procuraremos apresentar propostas para o tratamento lexicográfico das colocações que permitam suplantar tais dificuldades.

O modelo de AC utilizado em nosso trabalho baseia-se na proposta de Durão (2004), que diferencia três níveis de análise contrastiva:

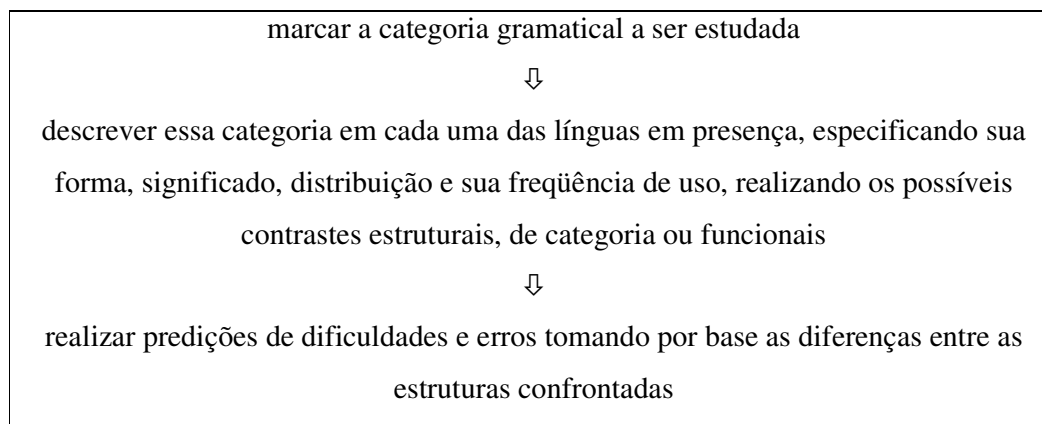
- a) Nível fonológico-fonético
- b) Nível léxico-semântico
- c) Nível morfossintático

Tendo em vista que o problema das colocações situa-se na intersecção entre os níveis (b) e (c), é necessário que realizemos uma adaptação desses modelos de análise. A proposta de AC - nível léxico-semântico, desenvolvida por Durão (2004, p. 32), organiza-se da seguinte forma:



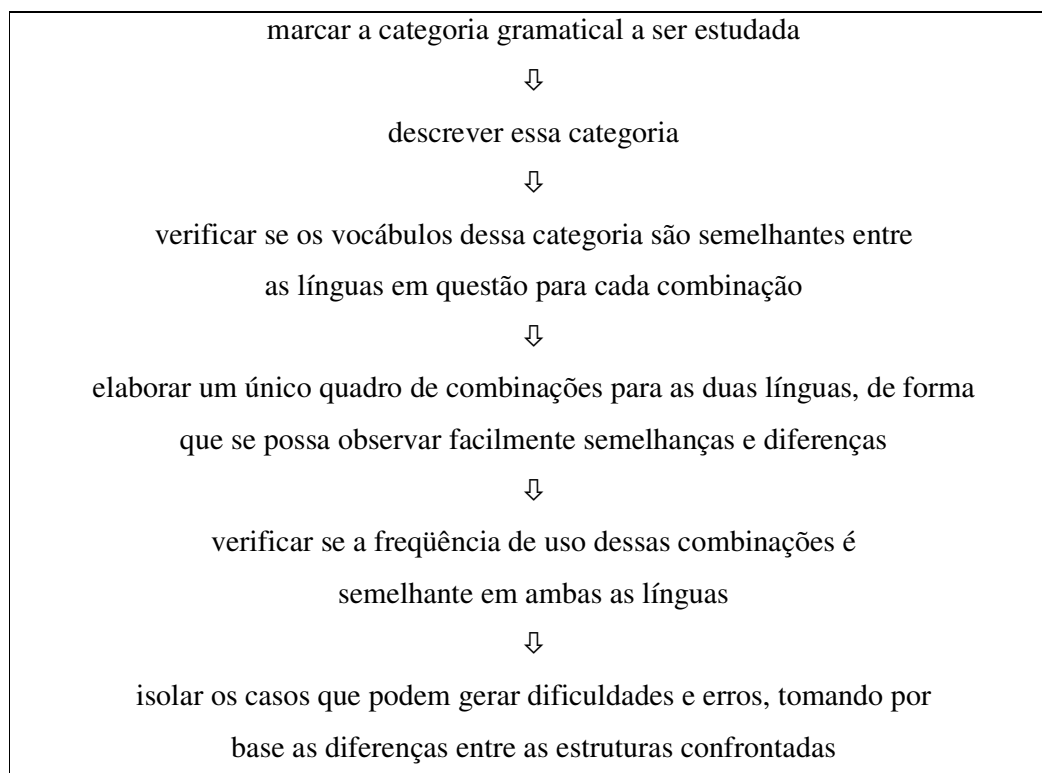
Esquema 2: Modelo de Análise Contrastiva - Nível léxico-semântico

Por outro lado, a proposta de AC - nível morfossintático, presente em Durão (2004, p. 33), apresenta a seguinte organização:



Esquema 3: Modelo de Análise Contrastiva - Nível morfossintático

Através, então, de uma adaptação desses modelos de análise, apresentamos, a seguir, o modelo de AC que utilizaremos como roteiro em nosso trabalho para o estudo específico das colocações. A partir do seguimento das etapas propostas, procuramos identificar as combinações léxicas que configuram problemas de colocações entre as línguas portuguesa e espanhola.



Esquema 4: Adaptação do Modelo de Análise Contrastiva para as colocações

6.4 Aplicação do modelo de análise

As etapas correspondentes à marcação da categoria gramatical e a sua descrição foram realizadas previamente - nos capítulos 4 e 5 deste trabalho -, ao selecionarmos estudar somente as combinações formadas por substantivos e adjetivos e identificar as relações semânticas estabelecidas entre essas classes de palavras. Como os segundos constituem os colocados e, conseqüentemente, o elemento que apresenta maior dificuldade na produção das colocações, realizamos uma descrição o mais exaustiva possível de sua categoria gramatical. Nessa oportunidade, propusemos uma classificação dos adjetivos que permitisse isolar aqueles que apresentam uma maior tendência a formar colocações. Tais adjetivos, que estabelecem com os substantivos relações “espaço-temporais” e “qualificativas *stricto sensu*” - de acordo com a terminologia por nós empregada -, serão utilizados, na presente seção, para a realização da análise contrastiva entre as línguas portuguesa e espanhola.

A partir, então, da seleção dos vocábulos em língua espanhola a serem analisados e identificação de suas combinações com substantivos, iniciamos a busca dos equivalentes em língua portuguesa para cada combinação. Devido à ausência de uma marcação das colocações em dicionários bilíngües espanhol-português/português-espanhol (v. 7.2), recorremos a dicionários monolíngües desses dois idiomas, tais como HouE (2001), AuE (1999), DRAEe (2003) e DUEe (2001). Essas obras, no entanto, também não fornecem um número muito expressivo de combinações léxicas. Em vista disso, utilizamos a ferramenta de busca de sites da internet *Google* como um auxílio no processo tradutório. Esse buscador foi escolhido por “disponibilizar conteúdo abrangente, diversificado, atualizado e por ser rápido, o que também é uma qualidade indispensável, dada a vasta quantidade de material existente na Web” (BERBER SARDINHA, 2003, p. 201).

Após a obtenção de uma lista de todas as combinações de substantivos e adjetivos em ambas as línguas, passamos a identificar os casos em que os colocados eram divergentes entre os dois idiomas. Para essa etapa, foi necessário recuperar os dados de coocorrência dos elementos fornecidos em nossas consultas ao *Google*, pois, embora muitas vezes houvesse vocábulos coincidentes entre as línguas, sua frequência em determinada combinação era diferenciada para cada idioma²⁰⁷. Assim, uma combinação poderia apresentar uma

²⁰⁷ Cabe ressaltar que o número de ocorrências fornecido pelo *Google* nem sempre coincide com o número total de exemplos disponibilizado por esse buscador (cf. BERBER SARDINHA, 2003, p. 215). Os valores de

equivalência “literal” entre as línguas portuguesa e espanhola, que, no entanto, não manifestasse uma frequência de uso semelhante entre essa duas línguas, pois havia em um dos idiomas uma combinação léxica distinta, utilizada de forma mais recorrente para exprimir o mesmo significado. A fim de ilustrar os procedimentos que nortearam nossa análise contrastiva, apresentamos a seguir alguns dos vocábulos estudados.

6.4.1 Análise dos dados

A partir da identificação dos equivalentes das combinações léxicas analisadas, foi possível classificá-las em cinco grupos diferentes, de acordo com a relação de equivalência existente entre as línguas portuguesa e espanhola. A seguir, apresentamos alguns exemplos para cada tipo de equivalência identificada, nos quais os elementos que correspondem aos colocados (adjetivos) são destacados através do uso do itálico.

6.4.1.1 *Equivalentes “literais”*

A maior parte das combinações formadas por substantivos e adjetivos que foram analisadas apresentava uma conformação léxica semelhante entre as línguas portuguesa e espanhola. Exemplos dessa equivalência, que chamamos de “literal”, constituem as combinações:

mayoría *aplastante* - maioria *esmagadora*
 crítica *feroz* - crítica *feroz*
 trabajo *arduo* - trabalho *árduo*
 quiebra *fraudulenta* - falência *fraudulenta*²⁰⁸

frequência oferecidos neste trabalho, porém, correspondem ao número total de ocorrências indicado pelo *Google*, independentemente da quantidade de exemplos apresentados.

²⁰⁸ Salientamos que, no caso de um dicionário bilíngüe voltado a falantes nativos de língua portuguesa (de acordo com nossa proposta em 7.3), as combinações léxicas deverão aparecer na ordem inversa, isto é, partindo do português ao espanhol.

A partir de uma investigação do emprego das colocações por estudantes de inglês e árabe, isoladamente, Farghal, Obiedat (1995, p. 325) verificaram que “ambos os grupos foram bem em colocações previsíveis, isto é, onde há convergência entre L₁ e L₂”²⁰⁹. Por outro lado, as colocações imprevisíveis foram as que causaram maiores dificuldades para os estudantes. Em vista disso, consideramos que as combinações que apresentam uma equivalência “literal” na língua estrangeira não necessitem ser incluídas nos dicionários.

6.4.1.2 *Posição divergente do adjetivo*

Algumas combinações, por sua vez, apesar de possuírem uma equivalência léxica “literal” entre os dois idiomas, apresentavam uma ordenação de substantivo + adjetivo em uma das línguas e de adjetivo + substantivo na outra. São exemplos dessa divergência de ordenação entre as línguas portuguesa e espanhola as seguintes combinações:

año próximo - próximo ano

craso error - erro crasso

No primeiro exemplo, encontramos 35 ocorrências em nosso *corpus* de estudo para a combinação *año próximo*. A partir de nossas consultas ao *Google*, verificamos 1,74 milhão de ocorrências para essa combinação em espanhol²¹⁰. No entanto, também a construção com o adjetivo anteposto apresenta altíssima frequência em língua espanhola, como pudemos constatar através das 2,4 milhões de ocorrências da combinação *próximo año*. No português, por outro lado, ainda que a combinação *próximo ano*, de forma semelhante à de língua espanhola, também apresente mais de 1 milhão de ocorrências, sua construção com o adjetivo posposto possui somente 21.900 ocorrências, sendo que em muitas delas o vocábulo *próximo* não funciona como adjetivo, mas desempenha a função de advérbio. Exemplos desse uso, retirados de *sites* do *Google*, são: “com o ano próximo do fim (...)”, “o aumento real do PIB (...) pode fechar o ano próximo a 2%”, entre outras. Dessa forma, embora ambas as

²⁰⁹ [both groups did well on predictable collocations, i.e., where there is convergence between L₁ and L₂].

²¹⁰ Para realizar consultas somente em língua espanhola, o buscador *Google* permite selecionar esse idioma através da opção de buscas “personalizada”. Por outro lado, para as consultas em língua portuguesa, além da seleção desse idioma, restringimos nossas buscas a “páginas do Brasil”, de forma que fossem recuperadas as ocorrências de uso somente nesse país.

construções sejam utilizadas nos dois idiomas, podemos perceber uma maior preferência da língua portuguesa pelo uso do adjetivo anteposto. Nesse caso específico, cabe salientar que, pelo fato aceitar as duas possibilidades em língua espanhola e somente uma em português, essa combinação representaria uma dificuldade apenas para o falante nativo do espanhol que pretende produzir em língua portuguesa.

O exemplo de *craso error* [erro craso], apresenta, por sua vez, uma posição muito mais fixa do adjetivo com relação ao substantivo para cada uma das línguas analisadas. Embora possam ser encontradas na *web* algumas ocorrências da combinação *error craso*, para a língua espanhola, e *crasso erro*, para a língua portuguesa, tais construções não denotam fluência para o falante nativo²¹¹.

Essas constatações evidenciam a importância de um tratamento diferenciado entre os adjetivos antepostos e pospostos, mas ressaltam que os primeiros também devem ser levados em consideração no estudo das colocações. Conforme apresentado em 5.3.2.4.1, os adjetivos antepostos podem se comportar de três maneiras diferentes:

- a) como palavras gramaticais;
- b) como morfemas, através de uma redução de seus traços lexicais;
- c) como palavras de significação “completa”, que selecionam, no nível da norma, a posição anteposta, como no exemplo de *craso error* [erro crasso].

Nesse último caso, em que cada língua seleciona uma posição diferente para a apresentação do adjetivo, consideramos importante salientar a divergência entre os idiomas, pois essa pode ser uma grande dificuldade para a obtenção da fluência em uma língua estrangeira.

6.4.1.3 Seleção idiossincrática do colocado

O principal problema, porém, no aprendizado de uma L2, acreditamos que seja causado pelas combinações que apresentam uma divergência na escolha do colocado em cada um dos idiomas. Esses casos correspondem a combinações em que, embora ambas as línguas

²¹¹ No momento de realizar consultas na internet, devemos ter sempre consciência de que esse é um canal de comunicação que pode ser “alimentado” por qualquer indivíduo, de forma que muito de seu conteúdo pode não corresponder à norma padrão dos idiomas.

possuam o mesmo par de adjetivos, cada uma seleciona um diferente para se combinar com determinado substantivo. Como exemplos dessa divergência, temos:

error mayúsculo - erro *gritante*
 discurso *encendido* - discurso *inflamado*
 victoria *resonante* - vitória *retumbante*
 éxito *rotundo* - sucesso *absoluto*

No primeiro exemplo, embora o espanhol não disponha do vocábulo *gritante*, essa formação aparece disponível no sistema dessa língua como uma forma de particípio presente do verbo *gritar*. Tal construção foi a escolhida pela língua portuguesa para manifestar a intensidade de um erro, dando origem à combinação *erro gritante*, com 1.200 ocorrências na *web*. A língua espanhola, por sua vez, optou por realizar sua combinação com o adjetivo *mayúsculo* [maiúsculo], resultando na colocação *error mayúsculo*, que aparece 6.430 vezes na *web*.

A segunda combinação demonstra a seleção de diferentes adjetivos, que compartilham o mesmo campo léxico, em cada um dos idiomas. Enquanto a língua espanhola optou pelo adjetivo *encendido* [aceso] para se combinar com o substantivo *discurso* [discurso], a língua portuguesa selecionou *inflamado*, embora também tivesse a sua disposição o adjetivo *aceso*. Esse vocábulo, por sua vez, que corresponde a um equivalente “literal” de *encendido*, costuma ser combinado com o substantivo “luz”, dando origem à construção *luces acesas*, que se equivale à combinação *luces encendidas* (MUN06) de nosso *corpus*.

Também os adjetivos *resonante* [ressonante] e *retumbante* [retumbante] fazem parte do léxico de ambos os idiomas; porém, enquanto a língua espanhola optou pela combinação *victoria resonante*, com 1.550 ocorrências na *web*, a língua portuguesa utiliza *vitória retumbante*, com 967 ocorrências. Por outro lado, se buscamos pela combinação *victoria retumbante*, em língua espanhola, obtemos somente 6 ocorrências, o mesmo número de casos encontrado para *vitória ressonante* em português. Acreditamos que as poucas ocorrências de algumas combinações devam ser minuciosamente analisadas. No caso, por exemplo, da combinação *éxito rotundo*, em língua espanhola, que corresponde a *sucesso absoluto*, em português, identificamos no *Google* 23 ocorrências para “páginas do Brasil” da construção *sucesso rotundo*, que constitui uma equivalência “literal” da construção em língua espanhola. No entanto, ao analisar cada uma dessas ocorrências, verificamos que elas estavam

relacionadas a traduções ou a temas do universo hispânico, bem como correspondiam à fala de pessoas provenientes das zonas de fronteira do Brasil com países que utilizam a língua espanhola. Em vista disso, percebemos que a combinação *sucesso rotundo* não correspondia realmente a um caso de equivalência “literal”, mas constituía um problema de interferência entre as línguas²¹². Dessa forma, consideramos importante salientar que todos os números de ocorrências apresentados neste trabalho representam indícios que foram cuidadosamente analisados e que devem, ainda, passar por outros meios de comprovação caso venham a ser utilizados em uma obra de referência.

6.4.1.4 *Estrutura morfológica divergente*

Além da escolha preferente na seleção do colocado realizada de forma divergente pelas línguas, outra dificuldade encontrada no contraste das combinações de substantivo + adjetivo entre os idiomas analisados diz respeito a construções em que uma das línguas não apresentava um adjetivo equivalente, mas utilizava um substantivo precedido de preposição, especialmente “de”, como em:

exitoso plan - plano de sucesso

Nesse exemplo, a língua espanhola utiliza, para se combinar com o substantivo *plan* [plano], o adjetivo *exitoso*, formado a partir do substantivo *éxito*, que equivale a *sucesso*, em português. Este idioma, no entanto, não dispõe de um adjetivo originado a partir desse substantivo, de modo que optou por utilizar uma construção que conte com a presença do vocábulo *sucesso*, como *plano de sucesso*. Apesar de essa combinação configurar o único caso encontrado em nosso *corpus* que ilustra tal situação, consideramos importante manter essa distinção devido, primeiramente, à limitação de nosso *corpus* de estudo, mas também porque essas estruturas deverão ter um tratamento diferenciado em sua inclusão nos dicionários, conforme trataremos em 7.3.1.

²¹² Segundo Castillo Carballo (2001, p. 135), “na maioria das vezes, os erros colocacionais são motivados pela influência da língua materna” [la mayor parte de las veces, los errores colocacionales vienen motivados por la influencia de la lengua materna].

6.4.1.5 *Combinaciones particulares a una lingua*

Por fim, consideramos importante salientar aquelas combinações que apresentam um colocado restrito na língua espanhola para o qual não foi possível encontrar um equivalente na língua portuguesa, tais como:

encuentro *maratónico* - jogo/evento [de longa duração]
 amistad *entrañable* - amizade [muito profunda]
 autor *prolífico* - autor [que produz muitas obras]

Nesses exemplos, ainda que seja possível encontrar um adjetivo em língua portuguesa que apresente um significado similar, como *longo*, *profunda* e *produtivo*, respectivamente, essas construções não configuram uma equivalência de significado com as combinações em língua espanhola. Além disso, através dessas opções, as combinações perdem o caráter mais restrito que possuem neste idioma, que as caracteriza como colocações, adquirindo o valor de combinações livres.

Além dos exemplos anteriores em que, para uma combinação na língua espanhola não havia uma construção equivalente em língua portuguesa, identificamos casos em que o espanhol apresentava duas combinações diferentes para as quais a língua portuguesa possuía somente uma opção. A única construção utilizada em português correspondia a um equivalente “literal” da combinação de maior frequência em espanhol, enquanto a outra construção existente nessa língua configurava, em nossa opinião, um caso de anisomorfismo²¹³. Como exemplos dessas combinações particulares da língua espanhola, temos:

despiadada ofensiva - [*brutal*] ofensiva
 negocio *jugoso* - negócio [*lucrativo*]
 tema *valórico* - tema [*importante*]

²¹³ Cabe ressaltar que, nos casos em que para duas estruturas do espanhol, a combinação oferecida para o português correspondia a um equivalente “literal” da construção menos frequente naquele idioma, consideramos tais combinações como uma seleção idiossincrática do colocado.

Nesses casos, as soluções oferecidas em língua portuguesa correspondem a combinações não apenas possíveis, mas, inclusive, mais utilizadas no espanhol. Para a primeira construção - *despidada ofensiva* -, por exemplo, encontramos 809 ocorrências no *Google*, enquanto *brutal ofensiva*, também em espanhol, aparece 3.610 vezes. A combinação *negocio jugoso*, por sua vez, corresponde a 1.470 casos, enquanto *negocio lucrativo* apresenta 33.100 ocorrências. Finalmente, *tema valórico* contabiliza 1.770 aparições e *tema importante*, mais de 500 mil. Acreditamos que essas combinações não representariam um problema para os falantes nativos de ambos os idiomas na produção em língua estrangeira, caso optassem pela construção mais usual, as quais possuem uma equivalência “literal” entre os idiomas. No entanto, a existência de uma combinação equivalente entre as línguas não invalida o caráter colocacional da forma utilizada somente em espanhol, visto que, nesses exemplos, ocorre uma seleção léxica na norma desse idioma que não se aplica à língua portuguesa. Apesar de dispormos dos vocábulos *suculento* e *impiedosa*, por exemplo, nosso idioma não os seleciona para se combinarem com *negócio* e *ofensiva*, respectivamente.

É importante salientar, ainda, que, nos casos em que a língua espanhola apresenta duas combinações diferentes para as quais o português conta com apenas uma, as construções no primeiro idioma podem não somente funcionar como sinônimas, manifestando o mesmo significado da combinação utilizada no português, mas também constituir uma variedade diatópica, diafásica ou diastraticamente marcada. No caso da combinação *tema(s) valórico(s)*, por exemplo, suas 6 ocorrências em nosso *corpus* correspondem a três diferentes edições do jornal *El Mercurio*, periódico originário do Chile. A partir desse dado, procedemos a uma consulta ao *Google* e identificamos que aproximadamente metade das ocorrências dessa combinação é proveniente de *sites* desse país, o que poderia indicar uma variação diatópica nessa construção. No entanto, como não dispomos de recursos suficientes para determinar essas variedades (cf. FORNARI, 2006a) e, ainda que fosse possível, fugiria aos propósitos deste trabalho, apenas constatamos essas divergências, sem oferecer uma indicação de se esses casos representam ou não uma construção variante. Porém, para um tratamento lexicográfico adequado dessas combinações, é importante que o lexicógrafo estabeleça essas diferenças com precisão antes de proceder à inclusão dessas estruturas no dicionário.

Em vista das diferentes relações de equivalência que as colocações podem estabelecer entre os idiomas, percebemos a relevância de realizar um estudo contrastivo para sua identificação, pois, embora muitas vezes haja uma correspondência dessas combinações entre duas línguas (apesar da seleção idiossincrática do elemento colocado), existem casos em que tais construções constituem uma particularidade de determinado idioma. Além disso,

pode ocorrer de uma colocação corresponder a uma combinação particular a um idioma somente no contraste com outra língua específica.

6.4.2 Resultados da Análise Contrastiva

A partir das 1.062 combinações formadas por substantivos e adjetivos pertencentes aos grupos de “qualificativos *stricto sensu*” e “espaço-temporais”²¹⁴ em língua espanhola, obtivemos uma relação de 68 combinações que apresentavam divergência com a língua portuguesa, constituindo, assim, exemplos de colocações entre os dois idiomas. A seguir, apresentamos, então, a lista de colocações obtidas, a partir de nosso *corpus*, para as línguas portuguesa e espanhola, classificando-as de acordo com a relação de equivalência entre as línguas.

a) Posição divergente do adjetivo

craso error - erro *crasso*

semana *próxima* - *próxima* semana

mes *próximo* - *próximo* mês

domingo *próximo* - *próximo* domingo

año *próximo* - *próximo* ano

b) Seleção idiossincrática do colocado

marcador *abultado* - placar *dilatado*

discurso *encendido* - discurso *inflamado*

horario *estelar* - horário *nobre*

investigación *exhaustiva* - investigação *detalhada*

prueba *fehaciente* - prova *cabal*

aterrizaje *forzoso* - aterrissagem *forçada*

²¹⁴ Somente esses grupos foram selecionados porque, conforme apresentado em 5.4, incluem os adjetivos que possuem uma maior tendência a formar colocações. No caso dos antepostos, foram considerados somente os que correspondem ao tipo (c) em nossa classificação (cf. 5.3.2.4.1).

plato *fuerte* - prato *principal*
 comida *liviana* - comida *leve*
 error *mayúsculo* - erro *gritante*
prestigioso diario - *prestigiado* jornal
 conclusión *provisional* - conclusão *provisória*
 victoria *resonante* - vitória *retumbante*
 éxito *rotundo* - sucesso *absoluto*
 dictador *sangriento* - ditador *sanguinário*
 visita *sorpresiva* - visita *surpreendente*
 persona *sospechosa* - pessoa *suspeita*
 operación *sospechosa* - operação *suspeita*
 intervención *temprana* - intervenção *precoce*
 trabajo *tesonero* - trabalho *excitante*
 información *veraz* - informação *verdadeira*

c) Estructura morfológica divergente

exitoso plan - plano *de sucesso*

d) Combinações particulares a uma língua

derrota *abrumadora* - derrota [*esmagadora*]
 crecimiento *abrumador* - crescimento [*esmagador*]
 mayoría *abrumadora* - maioria [*esmagadora*]
 victoria *ajustada* - vitória [*apertada*]
 zona *aledaña* - zona [*limítrofe*]
 sector *aledaña* - setor [*limítrofe*]
 fuente *allegada* - fonte [*próxima*]
bochornoso incidente - incidente [*vergonhoso*]
 tema *candente* - tema [atual e que provoca discussão]
 area *colindante* - área [*limítrofe*]
 zona *colindante* - zona [*limítrofe*]
 proposición *deshonesta* - proposta [*indecente*]
 cielo *despejado* - céu [*limpo*]

despiadada ofensiva - [*brutal*] ofensiva
encarnizada batalla - batalha [*sangrenta*]
encendida polémica - polêmica [muito intensa]
enconada polémica - polêmica [longa e intensa]
trámite engoroso - trâmite [*burocrático*]
amistad entrañable - amizade [muito profunda]
amigo entrañable - amigo [*do peito*]
figura estelar - figura [*de destaque*]
papel estelar - papel [*principal*]
estricto apego - apego [restrito]
flamante embajador - [*novo*] embaixador
año fructífero - ano [*produtivo*]
negocio jugoso - negócio [*lucrativo*]
momento justo - momento [*exato*]
encuentro maratónico - jogo/evento [de longa duração]
penosa agonía - [*dolorosa*] agonia
victoria pírrica - vitória [que causa grave dano ao vencedor]
autor prolífico - autor [que produz muitas obras]
silencio prudente - silêncio [que revela prudência]
parte medular - parte [*central*]
asunto medular - assunto [*central*]
tema medular - tema [*central*]
hecho sangriento - fato [muito violento]
producto sensible - produto [*perecível*]
actuación sobresaliente - atuação [*destacada*]
detección temprana - detecção [*precoce*]
diagnóstico temprano - diagnóstico [*precoce*]
piel tersa - pele [*lisa*]
tema valórico - tema [*importante*]

A partir dos exemplos arrolados, percebemos que uma mesma combinação poderia ser incluída em grupos diferentes de acordo com nossa classificação. Assim, por exemplo, a combinação *exitoso plan*, que corresponde a *plano de sucesso*, em português, apresenta uma divergência de ordem sintática e outra de ordem morfológica. Por outro lado, a

combinação, em português, equivalente a *amigo entrañable*, que significa *amigo do peito*, apesar de apresentar uma estrutura morfológica divergente, foi classificada como uma combinação particular à língua espanhola, pois consideramos que a opção por uma estrutura morfológica diferente ocorre sobre um vocábulo que não o equivalente “literal” do adjetivo em espanhol. Em vista disso, para nossa classificação das colocações a partir das relações de equivalência entre os idiomas, foram considerados, primeiramente, os aspectos de ordem semântica, seguidos dos morfológicos e, finalmente, os sintáticos.

Consideramos importante ressaltar que, nesse momento, não estamos preocupados em identificar se as colocações que classificamos nos diferentes grupos constituem uma dificuldade para o aprendiz de língua portuguesa ou espanhola. No entanto, é possível observar que algumas combinações serão mais complicadas conforme se trate da produção em um ou outro idioma. Assim, por exemplo, nos casos de combinações com divergência na posição do adjetivo, as línguas que utilizam somente a forma anteposta ou posposta representarão uma maior dificuldade para os estudantes, como ocorre com a colocação anteriormente analisada de *año próximo* para a língua espanhola e *próximo ano* para o português. Nesse caso, como a língua espanhola admite as duas posições do adjetivo, acreditamos que somente o falante nativo desse idioma encontrará problemas ao tentar utilizar a construção posposta na língua portuguesa.

As combinações particulares à língua espanhola, por sua vez, constituirão uma dificuldade para os falantes desse idioma se tentarem transportar a mesma combinação à língua portuguesa, como no exemplo de *producto sensible* [produto perecível], que resultaria na construção *?produto sensível*. Os falantes nativos do português, no entanto, partindo de combinações livres em sua língua materna, provavelmente não encontrariam dificuldades em produzir combinações igualmente livres em espanhol. Porém, o uso dessas construções revela uma menor fluência no idioma estrangeiro, ao passo que o emprego adequado das colocações manifesta um alto grau de conhecimento por parte do falante não-nativo. Em vista disso, reiteramos que cada colocação específica apresentará uma dificuldade diferente para os falantes de cada idioma, de modo que a seleção dessas combinações para seu tratamento lexicográfico deve levar em consideração o público ao qual o dicionário se destina, conforme trataremos em 7.1.2.

6.5 Limitações da Análise Contrastiva

Após a conclusão de nossas análises quantitativa e qualitativa, verificamos que suas aplicações em conjunto não permitiam uma identificação precisa das combinações léxicas que conformavam casos de colocações. A utilização de métodos estatísticos e do cálculo da frequência de coocorrência dos elementos, por um lado, mostrou-se relevante para a exclusão de combinações muito particulares e para a identificação daquelas realmente utilizadas pelos falantes. No entanto, o simples cálculo da frequência resultava na inclusão de muitas combinações livres no grupo das colocações.

Por outro lado, a utilização de uma classificação semântica dos adjetivos possibilitou excluir um grande número de combinações léxicas compostas por adjetivos que, devido a suas características intrínsecas, não apresentavam uma tendência a participar de colocações. Porém, a ausência de uma taxonomia lingüisticamente bem fundamentada, que permitisse caracterizar com precisão os diferentes tipos de adjetivos, impossibilitou que propuséssemos uma classificação suficientemente detalhada a ponto de determinar de forma objetiva os adjetivos que costumavam formar colocações.

Em virtude dessas dificuldades, foi necessário realizar uma análise contrastiva específica entre as línguas portuguesa e espanhola, de forma a identificar as combinações léxicas divergentes entre os dois idiomas. Essa análise surgiu, então, como uma última etapa na tentativa de isolar as combinações que configuram o fenômeno léxico das colocações. É importante salientar, porém, que todas as etapas do processo demonstraram ser de grande relevância para o resultado final, pois possibilitaram uma “filtragem” dos dados, de forma que cada combinação pôde ser minuciosamente analisada.

Assim, o estudo que realizamos das colocações, utilizando diferentes perspectivas de análise, possibilitou uma maior aproximação dos objetivos de nossa pesquisa. Considerando que, de acordo com a abordagem utilizada nos diferentes estudos sobre esse tema, os diversos lingüistas chegaram a resultados bastante divergentes (cf. capítulo 2) na identificação dessas combinações, foi necessário adotar um modelo de análise que permitisse satisfazer nossas necessidades práticas. Dessa forma, como nosso intuito era identificar as combinações que conformavam colocações, a fim de propor um melhor tratamento lexicográfico para tais estruturas, a utilização de uma análise contrastiva como um recurso final demonstrou ser fundamental. Somente através da comparação entre as línguas foi

possível verificar as combinações divergentes entre os dois idiomas, as quais, conseqüentemente, foram consideradas colocações.

Em vista disso, as colocações identificadas em nosso trabalho foram caracterizadas como combinações freqüentes entre um substantivo e um adjetivo, nas quais o adjetivo apresenta determinadas propriedades semânticas particulares, além de corresponder a uma seleção idiossincrática, ou seja, específica de cada idioma. É importante salientar, porém, que o fato de termos considerado colocações somente as combinações idiossincráticas e, portanto, diferentes entre as línguas portuguesa e espanhola não implica em que as combinações semelhantes entre esses idiomas não venham também a conformar casos de colocações. Exemplo disso pode ser observado na combinação *relación estrecha*, da língua espanhola, que apresenta como equivalente *relação estreita* para o português. Essa combinação, apesar de possuir uma coocorrência freqüente e ser formada por um adjetivo qualificador, o qual possui uma tendência a formar colocações, não apresenta uma escolha idiossincrática na seleção do colocado entre o português e o espanhol. No entanto, ao ser contrastada com a língua inglesa, tal combinação adquire um *status* de colocação, pois nesse idioma o colocado selecionado para se combinar com *relationship* [relação] não corresponde a *strait* [estreito], mas a *close* [perto], conformando a colocação *close relationship*, equivalente a *relação estreita* da língua portuguesa. A análise contrastiva das colocações permite, então, não uma delimitação de o que é ou não uma colocação, mas se uma combinação corresponde ou não a uma colocação com relação a uma outra língua determinada, o que é fundamental em uma abordagem lexicográfica desse fenômeno léxico, que deve enfatizar as possíveis dificuldades do aprendiz da língua estrangeira.

A utilização dos pressupostos da análise contrastiva, apesar de ter sido essencial para a identificação das colocações entre as línguas portuguesa e espanhola, apresentou também algumas deficiências. Como já apontava Wardhaugh (1983), o método contrastivo exige o conhecimento de universais lingüísticos muito bem fundamentados teoricamente, de forma a permitir a realização de um contraste eficiente entre as línguas. Em nosso estudo específico sobre as colocações, a ausência de uma teoria lingüística que fornecesse uma classificação detalhada dos adjetivos impossibilitou uma maior precisão em nossos resultados. Dessa forma, a combinação *situación similar*, por exemplo, que equivale a *situação semelhante*, em português, apresentou uma coaparição freqüente entre seus elementos, dentre os quais o adjetivo fazia parte do grupo dos qualificadores, que tende a formar colocações, além de constituir uma divergência entre as línguas portuguesa e espanhola. No entanto, apesar dessa combinação ter apresentado todas as propriedades que caracterizam nosso objeto

de estudo, consideramos que esse caso não configura um exemplo de colocação, pois a divergência na seleção do colocado não deriva de uma restrição imposta pelo substantivo, mas de uma preferência de uso de cada língua em particular. Assim, tanto em espanhol quanto em português coexistem as opções *similar* e *semejante* e *similar* e *semelhante*, respectivamente; porém, a primeira língua optou pelo uso mais difundido de *similar*, enquanto a segunda de *semelhante*.

Acreditamos que essa limitação não deriva propriamente de nossa análise contrastiva, mas das carências já identificadas anteriormente em nossos estudos quantitativo e qualitativo. Entendemos que um aprimoramento das ferramentas utilizadas nas etapas anteriores pode oferecer resultados ainda mais precisos em nossa análise. Em virtude desse fato, devemos considerar que os exemplos de colocações apresentados neste trabalho correspondem a combinações extraídas de um *corpus* em língua espanhola de pouco mais de 1 milhão de palavras, de modo que não podem ser tomados como uma solução definitiva. Para a realização de uma pesquisa com vistas à produção lexicográfica, consideramos fundamental a ampliação de todas as etapas do processo, a fim de obter colocações ainda mais confiáveis.

Finalmente, cabe ressaltar que, além dessa limitação de nossa análise, devido à impossibilidade de estabelecer o contraste entre os diferentes tipos de adjetivos para os dois idiomas estudados, os casos de combinações da língua espanhola para as quais não foram encontradas equivalências na língua portuguesa podem corresponder a uma insuficiência de nosso método de análise.

7 TRATAMENTO LEXICOGRÁFICO DAS COLOCAÇÕES EM DICIONÁRIOS BILÍNGÜES

A partir dos argumentos apresentados no capítulo anterior, consideramos que as colocações constituem uma maior dificuldade para os aprendizes de uma língua estrangeira no momento da produção de textos (v. 6.1). Em decorrência desse fato, o problema desse tipo de combinação léxica torna-se mais evidente no contraste entre dois idiomas. No caso específico das línguas portuguesa e espanhola, ainda que haja certa convergência entre esses idiomas (v. 6.2.3), existem casos em que não se pode “calcular” com facilidade a combinação na língua estrangeira e, por isso, os dicionários podem ser uma importante ferramenta no ensino-aprendizagem das colocações. No entanto, percebemos que nem todas as combinações precisam ser registradas em obras lexicográficas, mas apenas aquelas que representam divergências entre as línguas²¹⁵. Além disso, como as colocações demonstram ser mais relevantes quando estudadas contrastivamente, consideramos que sua inclusão seja fundamental em dicionários bilíngües. Esse tipo de dicionário, segundo Zagrebelsky (2002, p. 141), constitui uma fonte de consulta importantíssima no estudo de uma língua estrangeira. De acordo com Laufer, Levitzky-Aviad (2006, p. 136-137), ainda que os dicionários monolíngües apresentem mais informações sobre as palavras, os estudantes preferem utilizar os bilíngües devido a:

- (a) correlação direta entre a palavra em L2 e seu equivalente em L1 (língua que o usuário domina)²¹⁶;
- (b) ausência de necessidade de realizar novas consultas (o que ocorre nos dicionários monolíngües quando se desconhece algum dos vocábulos utilizados na definição);
- (c) desejo e necessidade do usuário de obter compreensão imediata.

O dicionário bilíngüe, então, é definido por Marelló (1996, p. 31) como

um dicionário no qual as expressões de uma língua (chamada língua fonte ou de partida) são traduzidas para uma outra (chamada língua alvo ou língua de chegada). Mas não é somente a presença de duas línguas que faz de um dicionário um bilíngüe. É a razão pela qual as duas línguas são postas em contato, isto é, a

²¹⁵ Opinião semelhante apresenta Irsula (1992, p. 166).

²¹⁶ Cabe ressaltar que essa correlação entre L2 e L1 ocorre em dicionários para a recepção, enquanto nos dicionários para produção de textos, como nos que vamos tratar a seguir, a correlação é exatamente ao contrário.

comunicação, pela tradução, entre duas comunidades que não compartilham a mesma língua²¹⁷.

Dessa forma, a autora ressalta que, mais do que o número de línguas envolvidas, o que caracteriza os dicionários bilíngües é seu papel de intermediador da comunicação entre duas comunidades lingüísticas. Também Landau (2001, p. 8), ao diferenciar os dicionários bilíngües dos monolíngües, ressalta que a função do primeiro é “oferecer ajuda para alguém que entende uma língua, mas não a outra”²¹⁸, sendo que uma das línguas corresponde a sua língua nativa. Kromann, Riiber, Rosbach (1991, p. 2717, grifo do autor) deixam mais claro o tipo de ajuda que os dicionários bilíngües podem oferecer ao indicar que sua função é a de “ser um auxílio para o tradutor *ou* na decodificação de um texto em língua estrangeira para um texto equivalente na língua nativa *ou* na codificação de um texto em língua nativa para um texto equivalente na língua estrangeira”²¹⁹. Esses autores salientam, assim, que, apesar das diferentes finalidades para as quais se pode utilizar um dicionário bilíngüe - auxílio na aprendizagem de uma L2, em viagens, em traduções especializadas, etc -, essas obras apresentam em comum o fato de serem empregadas para desempenhar duas funções específicas: auxiliar na compreensão ou na produção de textos.

Dessa forma, percebemos que o estudante de uma língua estrangeira pode consultar um dicionário bilíngüe com diferentes propósitos comunicativos relacionados tanto com o processo de produção quanto de compreensão de textos. Além disso, tais usuários podem possuir diferentes níveis de conhecimento da L2 (bem como de sua própria L1), o que também será um fator determinante do tipo de informações presentes nos dicionários e da forma como serão expostas. Em razão, então, dessas diversas variáveis envolvidas no processo de elaboração de dicionários, consideramos pertinente esclarecer alguns conceitos lexicográficos que correspondem a parâmetros teóricos na organização de dicionários bilíngües, de forma a relacioná-los com o tratamento lexicográfico das colocações.

²¹⁷ [un dictionnaire dans lequel des expressions dans une langue (dite langue source ou de départ) sont traduites dans une autre (dite langue cible ou langue d'arrivée). Mais ce n'est pas seulement la présence de deux langues qui fait d'un dictionnaire un bilingue. C'est la raison pour laquelle les deux langues sont mises en contact, c'est à dire la communication, par la traduction, entre deux communautés qui ne partagent pas la même langue].

²¹⁸ [to provide help to someone who understands one language but not the other].

²¹⁹ [to be an aid to the translator *either* in decoding a foreign-language text to an equivalent text in the native language *or* encoding a native-language text as an equivalent text in the foreign language].

7.1 Parâmetros teóricos para a elaboração de dicionários bilíngües

A elaboração de qualquer dicionário bilíngüe deve ser condicionada por três diferentes variáveis, quais sejam:

- (a) línguas envolvidas;
- (b) função do dicionário;
- (c) necessidades do usuário.

A relação estabelecida, em um dicionário bilíngüe, entre um par de línguas específico deve considerar a direção (v. 7.1.1) que será adotada por esse dicionário, isto é, a partir de qual língua serão dispostas as informações, e que função (v. 7.1.1) essa obra deverá desempenhar, ou seja, se terá como objetivo oferecer um auxílio para a produção ou para a compreensão de textos. Essas decisões deverão ser orientadas de acordo com as necessidades do público ao qual a obra é destinada e, por isso, é importante que o lexicógrafo delimite claramente o perfil do usuário (v. 7.1.2) ao qual pretende atingir. A delimitação, então, das línguas envolvidas, da função do dicionário e de seu público-alvo constituem etapas que devem ser previamente definidas, pois determinarão a organização de todos os componentes constitutivos de uma obra lexicográfica, os quais correspondem aos componentes canônicos (v. 7.1.3) e aos não-canônicos²²⁰.

7.1.1 Direção e função do dicionário

Como os dicionários bilíngües apresentam relações entre a L1 do usuário e uma L2, as informações podem ser dispostas partindo tanto da L1 para a L2 quanto da L2 para a L1. Em cada um desses casos, estamos diante de dicionários “monodirecionais”. Porém, quando uma obra lexicográfica é composta por segmentos em ambas as direções, corresponde a um dicionário “bidirecional” (cf. LANDAU, 2001, p. 9). Por “bidirecionalidade” de um dicionário bilíngüe, entende-se, então, “sua aptidão para servir a duas comunidades

²²⁰ Os componentes não-canônicos do dicionário seriam o *Back Matter* e o *Middle Matter*. O primeiro corresponde às partes funcionais que aparecem no final do dicionário, após a lista dos verbetes, e o segundo diz respeito a todas as unidades que são inseridas entre os verbetes, mas que não fazem parte da nominata do dicionário. (cf. HAUSMANN, WIEGAND, 1989, p. 331).

lingüísticas, seja como um dicionário de tradução [*i.e.* passivo], seja como um dicionário de versão [*i.e.* ativo]”²²¹ (MARELLO, 1996, p. 34). Haensch (1992, p. 399), porém, acredita que “no que se refere aos dicionários bilíngües, o ideal seria que um dicionário não tivesse que servir ‘nas duas direções’ ao mesmo tempo”²²². No entanto, ao afirmar que o dicionário não deve ser utilizado para as duas direções, esse autor está se referindo ao fato de que um mesmo dicionário - apresentando uma das direções - não pode satisfazer igualmente as necessidades dos falantes nativos de cada uma das línguas.

Dessa forma, ao aliar as necessidades dos usuários, a partir de sua língua materna, com a direção das obras lexicográficas, estabelece-se a distinção entre dicionários voltados à compreensão e à produção de textos. Os primeiros, de tipo “passivo”, objetivam “auxiliar os falantes nativos de cada uma das duas línguas a entender itens na outra”²²³ (COWIE, 1991, p. 682). Por outro lado, os dicionários para produção, também chamados “ativos”, visam a auxiliar “o usuário para traduzir de sua língua nativa para a estrangeira”²²⁴ (KROMANN, RIIBER, ROSBACH, 1991, p. 2715).

A respeito dessa dicotomia entre dicionários ativos e passivos, Landau (2001, p. 9) adverte que

muitos lexicógrafos bilíngües têm observado que é quase impossível construir um dicionário bilíngüe unidirecional para falantes de ambas as línguas. O compilador tem, ou deve ter, um grupo em mente, senão é provável que o dicionário não seja satisfatório para ninguém²²⁵.

Essa opinião é também compartilhada por Neubert (1992, p. 31), para quem “acreditar que qualquer dicionário bilíngüe pode apresentar a mesma informação em ambas as direções, isto é, para falantes de L1 bem como para falantes de L2, é outra ficção”²²⁶, ou seja, como as estratégias utilizadas para a produção e para a recepção de textos são diferentes, um mesmo dicionário não pode auxiliar de forma satisfatória a realização desses dois processos. Dessa forma, é fundamental que cada obra lexicográfica especifique a função que pretende desempenhar de acordo com as necessidades de seus usuários, pois essa decisão determinará o tipo de informações presente nos dicionários.

²²¹ [son aptitude à servir les deux communautés linguistiques soit en tant que dictionnaire de version, soit en tant que dictionnaire de thème].

²²² [por lo que se refiere a los diccionarios bilíngües, el ideal sería un diccionario que no tuviera que servir ‘en las dos direcciones’ al mismo tiempo].

²²³ [to help native speakers of each of the two languages to understand items in the other].

²²⁴ [the user for translation from the native to the foreign language].

²²⁵ [Many bilingual lexicographers have observed that is next to impossible to construct a unidirectional bilingual dictionary for speakers of both languages. The compiler has, or ought to have, one group in mind, else the dictionary is likely to be satisfactory for neither.].

²²⁶ [Believing that any bilingual dictionary can present the same information both ways, that is for L1 speakers as well for L2 speakers, is another fiction].

Segundo Marelló (1996, p. 33-34), “é corrente que a parte de tradução seja, senão mais longa, ao menos constituída de um maior número de entradas, enquanto a parte de versão possui os verbetes um pouco mais longos”²²⁷. Em vista disso, percebemos que a distinção entre dicionários de tipo ativo e passivo é responsável por uma diferente concepção da macro (v. 7.1.3.1) e da microestrutura (v. 7.1.3.2) das obras lexicográficas. Um dicionário voltado à produção de textos deve apresentar, então, uma macroestrutura mais concisa e uma microestrutura mais densa, ou seja, um menor número de entradas, mas com uma grande quantidade de informações funcionais dentro de cada verbete, de forma a facilitar o uso adequado de cada vocábulo no processo de codificação. Por outro lado, em um dicionário que vise à compreensão de textos, é fundamental uma macroestrutura mais densa, isto é, um maior número de palavras-entrada com vistas a facilitar o processo de decodificação, e uma microestrutura mais concisa.

Para Kromann, Riiber, Rosbach (1991, p. 2720), a distinção entre dicionários ativos e passivos “permite-nos eliminar sistematicamente detalhes supérfluos e maximizar as informações necessárias, estando, desse modo, de acordo com o princípio lexicográfico da economia”²²⁸. A partir disso, tais autores acreditam que, no caso das colocações, “nem todas as possíveis expressões [*i.e.* colocações] devem ser selecionadas e apresentadas no dicionário bilíngüe - somente aquelas que são imprevisíveis para o tradutor”²²⁹ (KROMANN, RIIBER, ROSBACH, 1991, p. 2722). Acreditamos que essa opção seja igualmente válida para os aprendizes de L2. Carvalho (2001, p. 83-84), por sua vez, ainda que não trate especificamente das colocações, também afirma que a inclusão de combinações léxicas nos dicionários deve ser adaptada de acordo com a função e os usuários específicos de cada obra:

Considerando-se que a grande contribuição da tipologia dos dicionários bilíngües é a melhor delimitação das necessidades do usuário, de modo a poupar espaço (além de tempo) com aquilo que é desnecessário, o exagero na inclusão de compostos transparentes com equivalências em dicionários de tradução sobrecarrega com lexemas dispensáveis exatamente a parte que precisa ser melhor [*sic*] trabalhada neste tipo de dicionário.

Dessa forma, considerando que essas combinações representam uma maior dificuldade no momento da produção de textos por parte de um falante não-nativo,

²²⁷ [il est courant que la partie version soit sinon plus longue du moins constituée d'un plus grand nombre d'adresses, alors que la partie thème a des articles en moyenne un peu plus longs].

²²⁸ [enables us to eliminate superfluous detail systematically and maximize necessary information, and is thus in accordance with lexicographical principles of economy].

²²⁹ [not all possible phrases should be selected and presented in the bilingual dictionary - only those that are unpredictable for the translator].

acreditamos que sua inclusão seja fundamental apenas em dicionários bilíngües de tipo ativo²³⁰.

7.1.2 Perfil do usuário

Conforme já mencionamos anteriormente, segundo as necessidades apresentadas pelos usuários de um dicionário - produção ou recepção de textos - e de acordo com a direção da obra, as informações oferecidas serão diferenciadas. Zgusta (1971, p. 299) chama a atenção para o fato de que o consulente possui uma maior competência de sua língua nativa e, conseqüentemente, as informações fornecidas nos dicionários deverão ser mais completas no que concerne à língua estrangeira. Neubert (1992, p. 31) esclarece essa diferença com o seguinte exemplo:

Um dicionário passivo inglês-alemão para falantes de alemão, como o que eu mesmo compilei, por exemplo, pressupõe uma competência nativa referente à compreensão, semântica e estilística de cada item oferecido em L₂ [*i.e.* na posição de equivalente]. Um usuário de inglês, por outro lado, não teria igual compreensão das palavras alemãs situadas na posição L₂²³¹.

Porém, embora a delimitação precisa de um público-alvo seja uma das primeiras decisões ao se projetar um dicionário (cf. DUBOIS, DUBOIS, 1971, p. 9), Landau (2001, p. 9) observa que “raramente um dicionário bilíngüe identifica o usuário para quem ele é planejado”²³². Tal fato deve-se, especialmente, à ausência de estudos sobre o perfil dos usuários de dicionários. Segundo Kromann, Riiber, Rosbach (1991, p. 2713), esse aspecto permanece não esclarecido na prática lexicográfica recente, de forma que lexicógrafos e editores possuem uma concepção muito vaga do público-alvo de suas obras. Santamaría Pérez (2000, p. 37) também acredita que, apesar de ser uma decisão óbvia, a adequação das informações às necessidades do usuário não é simples, pois não se pode determinar

²³⁰ Laufer, Levitzky-Aviad (2006, p. 135) consideram que, apesar de recorrerem aos dicionários bilíngües para compreender o significado de uma palavra, os aprendizes geralmente não os utilizam para a produção, pois essas obras “raramente diferenciam entre as possíveis traduções em L2 de uma palavra na L1, assim como não fornecem informação a respeito do uso de cada opção de tradução” [rarely differentiate between the possible L2 translations of the L1 word, nor do they provide information regarding the use of each translation option]. Acreditamos, porém, que o fato de os dicionários bilíngües disponíveis não desempenharem adequadamente sua função de auxiliar na produção de textos - porque não são sequer pensados para essa finalidade - não pode ser um impedimento para propostas que visem a melhoria dessas publicações.

²³¹ [A passive dictionary, English-German for speakers of German for instance, the one I compiled myself, presupposes a native competence with regard to the comprehension, semantic and stylistic, of every L₂ item offered. An English user, on the other hand, would not have an equal comprehension of the German words placed in the L₂ position].

²³² [rarely does a bilingual dictionary identify the user for whom it is intended].

exatamente o público de uma obra bilíngüe, devido à diversidade de idade, competência lingüística em L1, nível de aprendizado de L2, necessidade para compreensão ou produção, etc.

Bugueño (2007, p. 263) salienta, porém, que já se pode definir *a priori* como um traço característico do usuário o fato de que “deseja percorrer o menor caminho possível para procurar informação dentro de um dicionário”. Em vista disso, uma crescente preocupação por parte dos lexicógrafos diz respeito à apresentação das informações nos dicionários de forma mais “amigável” a seu consulente²³³. No caso específico das colocações, considerando que essas combinações correspondem a estruturas formadas por duas unidades léxicas, a principal dificuldade por parte dos lexicógrafos é criar uma estratégia que permita ao usuário encontrar a informação que necessita com o menor esforço possível. Em virtude desse fato, surgem diferentes propostas para o tratamento lexicográfico desse fenômeno léxico (v. 7.1.3.1).

Qualquer proposta para o tratamento lexicográfico das colocações, no entanto, deve reconhecer as necessidades e competências dos usuários específicos de cada dicionário, pois essas variáveis deverão orientar a organização de todos os seus componentes. O reconhecimento do público-alvo determinará, também, a quantidade e a forma de apresentação das informações oferecidas em cada obra particular, conforme exposto em 7.1.1 a propósito da dicotomia entre dicionários ativos e passivos. Segundo Haensch (1982, p. 400), se as necessidades específicas dos usuários fossem consideradas, “poderiam ser elaborados dicionários muito mais úteis”²³⁴. No entanto, na ausência de um perfil de usuário claramente definido, Bugueño (2007, p. 263) afirma ser possível trabalhar somente com uma concepção hipotética do público-alvo de um dicionário. Para traçar, então, o perfil de um possível usuário brasileiro de um dicionário bilíngüe português-espanhol, estudante do ensino superior, é necessário que conheçamos as bases curriculares que orientam as reflexões sobre o ensino da língua espanhola, a fim de identificar os conhecimentos que os estudantes devem ter em cada nível de sua aprendizagem. Para isso, recorreremos ao *Marco Común Europeo de Referencia para las Lenguas: aprendizaje, enseñanza, evaluación* [Quadro Europeu Comum de Referência para as Línguas: aprendizagem, ensino, avaliação], doravante MARCO (2002), documento responsável por formular alguns parâmetros sobre os quais se baseia o ensino do

²³³ Pérez Lagos (2001), por exemplo, destaca o didatismo do *Diccionario Ideológico de la Lengua Española VOX*, que “converte o esforço exigido (...) em um processo de aprendizagem e conhecimento da língua” (PÉREZ LAGOS, 2001, p. 192) [convierte el esfuerzo exigido (...) en un proceso de aprendizaje y conocimiento de la lengua].

²³⁴ [se podrían elaborar diccionarios mucho más útiles].

espanhol. Nosso intuito é identificar as características dos estudantes de língua espanhola e as competências que lhe são exigidas nos diferentes níveis de aprendizagem. Dessa forma, acreditamos que seja possível uma maior aproximação ao perfil de um provável usuário estrangeiro de um dicionário bilíngüe português-espanhol em seus diferentes níveis de aprendizagem, pois saberemos exatamente o conteúdo que ele deve ter acesso para atingir os níveis propostos pelo MARCO (2002).

7.1.2.1 *Marco Común Europeo de Referencia para las Lenguas: aprendizaje, enseñanza, evaluación*

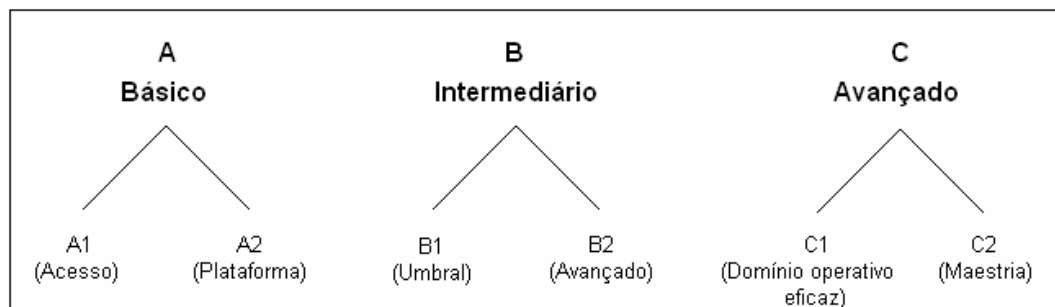
O MARCO (2002) corresponde a um projeto de política lingüística desenvolvido pelo Conselho da Europa com o intuito de unificar as diretrizes para o ensino e aprendizagem de línguas no contexto europeu. Segundo Verdía (2002, p. 4), esse documento pretende suscitar uma reflexão sobre os objetivos e a metodologia utilizada no ensino e aprendizagem de línguas estrangeiras, bem como oferecer uma base comum para o desenvolvimento de currículos, programas, provas e critérios de avaliação. Para cumprir seus objetivos, o MARCO (2002, p. 1) “descreve de forma integradora o que os estudantes de língua têm que aprender a fazer com a finalidade de utilizar uma língua para se comunicar, bem como os conhecimentos e destrezas que devem desenvolver para poder agir de forma eficaz”²³⁵. Assim, oferece uma proposta para classificação dos estudantes de língua estrangeira em diferentes níveis de conhecimento e, para cada nível, apresenta o grau de competência comunicativa que esse estudante deve ter, bem como as tarefas que deve ser capaz de realizar.

7.1.2.1.1 *Níveis de conhecimento propostos pelo MARCO (2002)*

A partir de uma ramificação da divisão clássica dos níveis de conhecimento de uma língua estrangeira em “básico”, “intermediário” e “avançado”, o MARCO (2002) propõe

²³⁵ [describe de forma integradora lo que tienen que aprender a hacer los estudiantes de lengua con el fin de utilizar una lengua para comunicarse, así como los conocimientos y destrezas que tienen que desarrollar para poder actuar de manera eficaz].

uma classificação dos estudantes em seis diferentes níveis, a saber, A1 ou “Acesso”, A2 ou “Plataforma”, B1 ou “Umbral”, B2 ou “Avançado”, C1 ou “Domínio operativo eficaz” e C2 ou “Maestria”. A figura abaixo permite identificar a correlação entre os níveis propostos pelo MARCO (2002) e a divisão clássica dos níveis de conhecimento dos estudantes.



Esquema 5: Níveis de conhecimento propostos pelo MARCO (2002)

Ainda que o MARCO (2002) não defina claramente os estudantes que pertencem a cada grupo, apresenta, através de uma escala de descritores, as capacidades (competências gerais e comunicativas) que os alunos devem apresentar para realizar atividades de produção, recepção, interação e mediação, tanto de forma oral quanto escrita, em cada um dos diferentes níveis. As características que definem cada nível apresentam uma evolução gradativa, conforme podemos observar a seguir:

- **A1** ou “**Acesso**”: nível mais baixo de uso da língua que permite aos estudantes uma interação simples sobre temas específicos do cotidiano e com um vocabulário bastante limitado.
- **A2** ou “**Plataforma**”: nível de competência que possibilita uma comunicação mais ativa, porém, repleta de limitações.
- **B1** ou “**Umbral**”: nível de competência característica de viajantes, que permite uma interação de forma flexível em situações relacionadas a problemas cotidianos.
- **B2** ou “**Avançado**”: nível de competência que permite aos estudantes utilizar a língua com fluidez e naturalidade. Caracteriza-se, especialmente, pelo desenvolvimento de uma competência metalingüística por parte dos aprendizes, capazes de identificar e corrigir equívocos.
- **C1** ou “**Domínio operativo eficaz**”: nível avançado de competência apropriado para tarefas mais complexas de trabalho e estudo. Os participantes

desse grupo possuem um amplo repertório lingüístico, que lhes permite uma comunicação fluida e espontânea.

- **C2** ou “**Maestria**”: maior nível de conhecimento lingüístico que, no entanto, não implica em uma competência de falante nativo. Os estudantes desse grupo caracterizam-se por uma grande precisão e facilidade no uso da língua, além do domínio de uma competência intercultural.

Obviamente, os estudantes mais avançados devem possuir também as competências exigidas para os níveis mais iniciais, embora essas características não sejam repetidas em todos os níveis apresentados. Além disso, conforme mencionado anteriormente, é importante destacar que os níveis propostos pelo MARCO (2002) seguem uma escala gradativa que acompanha o desenvolvimento das competências do aluno durante seu processo de aprendizagem da língua estrangeira.

Dentro, ainda, de cada nível de classificação dos estudantes, o MARCO (2002) propõe uma série de descritores que indicam os conhecimentos específicos que os alunos devem possuir, em cada nível, para a realização de diferentes tarefas. No que concerne ao uso das colocações, não há uma referência direta a esse tipo de combinações léxicas, porém, ao abordar questões relativas ao domínio do vocabulário por parte dos estudantes, o MARCO (2002, p. 109) considera que integrantes do nível B2 ou “Avançado”, apesar de uma elevada precisão léxica, cometem incorreções ao selecionar as palavras, sem que isso seja um obstáculo à comunicação. Dessa forma, reitera-se a noção de que o uso correto das colocações constitui um indicativo da alta fluência de um estudante na língua estrangeira (cf. ALONSO RAMOS, 2001, p. 10; NESSELHAUF, 2005, p. 2; RUNDELL, 1999, p. 37, entre outros).

7.1.2.2 Perfil do usuário de um dicionário bilíngüe ativo português-espanhol

Para traçar o perfil de um possível usuário de um dicionário bilíngüe português-espanhol cogitado para a produção de textos em língua espanhola, devemos conjugar três variáveis:

- a) a língua materna do usuário
- b) a língua estrangeira que está aprendendo
- c) o nível de competência que apresenta de acordo com o MARCO (2002)

Com relação aos itens (a) e (b), podemos caracterizar o usuário de um dicionário bilíngüe ativo português-espanhol como um estudante brasileiro que aprende a língua espanhola no ensino superior. No que concerne ao item (c), considerando que as colocações constituem uma dificuldade ainda para os alunos pertencentes ao nível B2 ou “Avançado” - conforme observamos em 7.1.2.1.1 -, acreditamos que todos os estudantes classificados até esse nível devam ser contemplados em um dicionário bilíngüe ativo português-espanhol. Porém, visto que as competências de um integrante do nível A1 ou “Acesso” e do nível B2 ou “Avançado” serão muito diferentes, é fundamental que haja uma distinção entre esses estudantes, através da elaboração de, pelo menos, duas obras lexicográficas distintas.

Em virtude desse fato, as colocações apresentadas nos dicionários devem ser coerentes com as necessidades de cada grupo de usuários, ou seja, devem suprir as deficiências de estudantes de níveis mais básicos e mais avançados, dependendo da obra em questão. Segundo Hanay (2003, p. 152), “não é somente o contexto de uso do usuário que ajuda a moldar um dicionário, mas também o nível de conhecimento do usuário”²³⁶. No entanto, existe uma divergência entre os autores em relação ao nível de conhecimento que o aluno deve possuir para iniciar a aprendizagem dessas combinações léxicas. Alonso Ramos (2002a, p. 553), por um lado, acredita que as colocações devam ser ensinadas desde os primeiros níveis de aprendizagem, devido à alta recorrência dessas combinações na língua. Segundo a autora, “está comprovado que os aprendizes cometem mais erros gramaticais justamente por desconhecer a colocação adequada no momento adequado”²³⁷ (ALONSO RAMOS, 2002a, p. 553). Também para Castillo Carballo (2001, p. 143) é fundamental o ensino das colocações desde os primeiros contatos com a língua estrangeira, de modo que o aluno tome consciência da coocorrência das unidades léxicas da mesma forma que aprende a gramática e os vocábulos mais essenciais. Por outro lado, Maurer-Stroh (2004, p. 88) considera que as colocações devem ser incluídas apenas em dicionários para estudantes de nível avançado e tradutores. Koike (2000, p. 271), além dos alunos de nível avançado, inclui os intermediários no grupo de estudantes para os quais as colocações devem ser ensinadas.

Acreditamos que tanto os usuários de nível básico quanto os de nível mais avançado devam ter acesso às combinações léxicas em um dicionário bilíngüe voltado à produção. O que deve diferir, porém, é a quantidade e o tipo de colocações apresentadas. Para

²³⁶ [it's not only the user's context of use that helps shape a dictionary but also the user's level of knowledge].

²³⁷ [está comprobado que los aprendices cometen más errores gramaticales, precisamente por desconocer la colocación adecuada en el momento adecuado].

o primeiro grupo, correspondente ao nível A2 ou “Plataforma” no MARCO (2002)²³⁸, é importante que seja incluído um maior número de combinações, a fim de diminuir a quantidade de hipóteses em sua interlíngua e aumentar sua competência comunicativa. Nesse caso, seriam incluídas colocações mais simples e usuais, como *error mayúsculo* [erro gritante] e *plato fuerte* [prato principal]. Por outro lado, para estudantes mais avançados, especialmente os que correspondem ao nível B2 ou “Avançado” no MARCO (2002), as colocações presentes no dicionário serviriam para um enriquecimento vocabular, através da inclusão de combinações mais específicas, tais como *victoria resonante* [vitória retumbante] e *prueba fehaciente* [prova cabal].

Dessa forma, é evidente que, conforme o nível de conhecimento do usuário, deveríamos dispor de duas obras lexicográficas distintas, cada uma destinada a satisfazer as necessidades de seu público específico. Tais necessidades devem, então, além de determinar a seleção das colocações a serem incluídas, orientar a organização de todos os componentes de cada obra lexicográfica.

7.1.3 Componentes canônicos do dicionário

A delimitação de uma função específica - para a produção - e de um grupo restrito de usuários - estudantes do nível A2 ou “Plataforma” e B2 ou “Avançado”, distintamente - para um dicionário bilíngüe será importante para orientar a organização das informações nos diversos segmentos que compõem essa obra, quais sejam, macro (v. 7.1.3.1), micro (v. 7.1.3.2), medioestrutura (v. 7.1.3.3) e partes introdutórias (v. 7.1.3.4). Segundo Bugueño, Farias (2007, s.p.), a correta definição e articulação desses componentes é um grande auxílio para elevar a qualidade dos dicionários.

²³⁸ Não incluímos os estudantes do nível A1 ou “Acesso” como possível público-alvo de um dicionário bilíngüe ativo porque as competências exigidas para esses alunos são ainda muito básicas. De acordo com o MARCO (2002, p. 26), dizem respeito à utilização de frases bastante simples, com um vocabulário limitado, destinadas a satisfazer suas necessidades imediatas.

7.1.3.1 Macroestrutura

A macroestrutura, segundo Hartmann, James (1998, s.v. *macrostructure*), corresponde à “estrutura global da lista que permite ao redator e ao usuário localizar informação em uma obra de consulta”²³⁹, ou seja, diz respeito à lista de palavras que compõem a nominata do dicionário. Landau (2001) amplia o escopo de macroestrutura, incluindo em seu conceito, além da simples lista de palavras, noções de lematização, como a diferença entre polissemia e homonímia, o agrupamento de derivados em uma mesma entrada (*run-on-entries*), o número de entradas, entre outros²⁴⁰. No presente trabalho, trataremos somente das questões que dizem respeito a uma definição das unidades que comporão a macroestrutura e aos critérios utilizados para a lematização dessas unidades. Tais aspectos fazem parte da “definição macroestrutural”²⁴¹ do dicionário, que segundo Bugeño, Farias (2007, s.p.), corresponde ao estabelecimento do número de verbetes que essa obra conterá, assim como do tipo de unidades passíveis de lematização. Para que a organização macroestrutural de um dicionário seja a mais completa e coerente possível, é necessária a realização prévia de uma definição macroestrutural quantitativa e uma qualitativa.

A definição macroestrutural quantitativa diz respeito ao cálculo do universo léxico a ser arrolado no dicionário. Tal definição não corresponde somente ao número de entradas que um dicionário apresenta, mas à pertinência destas em relação à função e ao público-alvo da obra, os quais devem estar devidamente definidos nas partes introdutórias (v. 7.1.3.4). Assim, em nosso caso específico de estudo das colocações, uma definição macroestrutural eficiente deve delimitar as combinações léxicas que serão incluídas no dicionário, tendo em vista sua função ativa e um público de nível A2 ou B2 na língua espanhola.

A definição macroestrutural qualitativa, por sua vez, permite elencar os tipos de unidades que podem fazer parte da macroestrutura, bem como determinar os critérios utilizados para a lematização de tais unidades. Com relação às colocações a serem incluídas

²³⁹ [the overall list structure which allows the compiler and the user to locate information in a reference work].

²⁴⁰ Um panorama geral sobre a evolução do conceito de macroestrutura entre diversos autores pode ser visto em Bugeño (2007).

²⁴¹ Consideramos importante ressaltar que o comportamento *sui generis* das colocações aparece também refletido em seu tratamento lexicográfico, uma vez que tais combinações podem ser abordadas tanto como um problema de organização macroestrutural quanto microestrutural. Do ponto de vista macroestrutural, as colocações são tratadas como unidades que envolvem uma relação sintagmática entre seus elementos, os quais promovem uma reflexão sobre a forma de acesso a sua estrutura no dicionário. Por outro lado, as colocações não representam uma dificuldade no reconhecimento do significado da combinação, mas na identificação do elemento colocado, constituindo, assim, um problema de designação, o que também justifica sua abordagem do ponto de vista da microestrutura.

nos dicionários, conforme explicitado em 7.1.2.2, os usuários do nível A2 deverão ter acesso a um grupo de combinações mais simples e usuais e, possivelmente, em maior número. Já para os usuários do grupo B2 serão incluídas colocações que denotam uma maior proficiência na língua estrangeira.

Com relação à forma como as informações devem ser apresentadas nos dicionários, a principal preocupação, no caso específico das colocações, diz respeito a sob qual entrada - se da base ou do colocado - essas combinações devem ser lematizadas. Em Benson (1997), as colocações são organizadas de acordo com as categorias gramaticais envolvidas, ou seja, as colocações gramaticais são listadas sob a palavra dominante e as colocações lexicais são organizadas primeiro pelo substantivo, seguido pelo adjetivo e depois pelo verbo. Essa disposição da combinatória léxica no dicionário pode ser explicada pelo fato de não se estabelecer uma hierarquia entre base e colocado (v. 2.2.3). Assim, para Benson (1985, p. 65), nas combinações de verbo + substantivo, por exemplo, as colocações devem ser apresentadas na entrada do substantivo, pois a inclusão a partir do verbo pode não fornecer ao leitor as informações de que ele necessita²⁴². Wouden (1992, p. 454) mostra-se completamente contrário a esse tipo de organização, pois, devido ao anisomorfismo das línguas, acredita que essa forma de apresentação pode dificultar no momento da consulta. Sua solução é que as colocações sejam ordenadas no dicionário a partir de regras semânticas, pois julga que essas estratégias estejam mais próximas das utilizadas por tradutores ou falantes não-nativos²⁴³. Hausmann (2002, p. 12), por sua vez, afirma que a ordem semântica repousa sobre uma interpretação fortemente subjetiva sendo, portanto, plausível, mas não previsível²⁴⁴.

Alguns trabalhos, porém, já apresentam um posicionamento mais definitivo com relação ao tratamento das colocações a partir da entrada da base ou do colocado. Por um lado, encontramos a opinião de Alonso Ramos (2001), que acredita que as colocações devam aparecer sob a entrada da base, pois considera que esta seja responsável pelo conteúdo sêmico

²⁴² Ainda que não utilize uma separação entre dicionários para a produção e para a compreensão, Benson (1985, p. 13, grifo do autor) percebe que as colocações representam uma dificuldade maior no primeiro aspecto, pois afirma que “a questão habitual é: que verbo se coloca com *desculpa*? Ou: O que se faz para uma *desculpa*?” [the usual question is: which verb collocates with *plea*? Or: What does one do to a *plea*?].

²⁴³ Wouden (1992) defende a organização interna do dicionário a partir de relações semânticas e indica o dicionário de Mel'čuk et al (1981) como um bom exemplo dessa organização. Não tivemos acesso a esse primeiro volume, mas apenas a Mel'čuk (1984, 1988).

²⁴⁴ Opinião semelhante apresenta Jarošová (2000, p. 12), para quem, devido à questão do anisomorfismo, “a divisão do espaço semântico não é idêntica entre quaisquer duas línguas” [the division of semantic space is not identical in any two languages].

da expressão²⁴⁵. A partir do argumento de Heid (1997, p. 50) de que “do ponto de vista do usuário do dicionário, o próprio fato de que haja na entrada da base uma espécie de ‘lista de colocações frequentes’ é útil, porque, normalmente, conhece-se a base e procura-se o colocado adequado”²⁴⁶, essa autora defende a inclusão das colocações na entrada da base para facilitar o processo de codificação em L2. Além disso, alguns colocados demonstram ser muito produtivos na formação de colocações, de modo que não é possível registrar todas as combinações em sua entrada. Segundo Alonso Ramos (2001, p. 17), “não se pode prever a partir do adjetivo quais são todos os substantivos que o selecionam em colocação”²⁴⁷, ou seja, a partir da entrada do colocado, não é possível identificar todas as bases com que ele se combina, uma vez que um mesmo colocado pode participar de muitas colocações (o que ocorre especialmente nos casos de verbo-suporte). Dessa forma, Alonso Ramos (2001, p. 14) discorda da incorporação das colocações na entrada do colocado, como ocorre no dicionário *Larousse* francês-espanhol/espanhol-francês analisado por ela, ainda que tal forma de apresentação pudesse facilitar a consulta do usuário, acostumado a decodificar as expressões partindo da esquerda para a direita²⁴⁸. Ruiz Gurillo (2000), ainda que não utilize os termos base e colocado, também defende que as colocações devam aparecer como subentendidas dependentes de seu “formante” mais importante²⁴⁹.

Cop (1991), por outro lado, afirma que apresentar a base na entrada do colocado auxilia a discriminar o significado deste em suas diversas acepções, enquanto incluir o colocado na entrada da base não contribui muito para a compreensão do significado desta, mas auxilia à contextualização de seus equivalentes, isto é, serve para delimitar o alcance de cada equivalente da base. Assim, a autora acredita que as colocações devam aparecer tanto sob a entrada da base quanto do colocado. No entanto, Hausmann (1989), justamente por trabalhar com as diferenças entre dicionários para produção e para recepção, identifica que, conforme a necessidade do usuário, as colocações podem ser incluídas em lugares diferentes. Dessa forma, em um dicionário voltado à recepção de textos, acredita que a apresentação

²⁴⁵ Conforme exposto em 2.2.5, no entanto, a propósito do exemplo *formular uma hipótese*, muitas vezes ambos os elementos contribuem para a conformação do significado da combinação.

²⁴⁶ [dal punto di vista dell’utente del dizionario, il fatto stesso che ci sia nell’entrata della base una specie di ‘elenco di collocazioni frequenti’ è utile, perché, normalmente, si conosce la base e viene cercato il collocatore adeguato].

²⁴⁷ [ne peut pas prévoir à partir de l’adjectif quels sont tous les noms qui le sélectionnent en collocation].

²⁴⁸ É importante salientar que a posição do colocado mais à esquerda da combinação vai depender das línguas envolvidas e da estrutura em questão. Os colocados costumam, nas línguas portuguesa e espanhola, por exemplo, vir antepostos em combinações de verbo + substantivo, ao passo que, nas combinações de substantivo + adjetivo, aparecem pospostos. Na língua inglesa, por outro lado, os adjetivos tendem a ser antepostos ao se combinarem com substantivos.

²⁴⁹ Opinião também compartilhada por Mel’čuk (2003).

dessas combinações no colocado é mais indispensável que na base, sendo que nesta “a menção da colocação é sem importância para a compreensão da palavra tratada”²⁵⁰ (HAUSMANN, 1989, p. 1010). Por outro lado, em um dicionário que vise à produção de textos, a incorporação da colocação torna-se mais relevante quando feita a partir da base, visto que o falante organiza seu texto a partir dela e, somente após sua escolha, seleciona o colocado²⁵¹. Também Heid (1996, p. 134) acredita que “um dicionário bilíngüe para a produção na língua-alvo idealmente deveria listar as colocações sob a entrada da base”²⁵². Porém, quando um dicionário servir às duas funções simultaneamente, Hausmann (1985, p. 122) concorda com Cop (1991) em que as combinações devem ser lematizadas tanto na entrada da base como do colocado. Finalmente, Koike (2000, p. 267) também afirma que “proporcionar uma lista de colocados na entrada de uma base serve para produzir textos, enquanto indicar várias bases na entrada de um colocado é útil para a compreensão do texto”²⁵³. A partir disso, reforça-se a noção de que, conforme a finalidade da obra lexicográfica, as colocações deverão ser apresentadas de forma diferenciada.

Para o tratamento lexicográfico dessas combinações em um dicionário bilíngüe português-espanhol de tipo ativo, consideramos que seja pertinente sua inclusão a partir da entrada de sua base, pois concordamos com a opinião de que esse será o elemento consultado no dicionário para a produção de um texto. Acreditamos, assim, que a presença das colocações na entrada da base vai ao encontro das necessidades do usuário, que não terá que realizar novas consultas para localizar a informação que procura. É importante ressaltar, porém, que, para que as buscas realizadas pelo consulente resultem em uma pronta solução, é necessário que os critérios utilizados para lematização estejam claramente definidos nas partes introdutórias (v. 7.1.3.4) do dicionário.

²⁵⁰ [La mention de la collocation est sans importance pour la compréhension du mot traité].

²⁵¹ Calderón Campos (1994, p. 124) concorda que o verbete do colocado desempenha uma função decodificadora, enquanto o da base auxilia na produção.

²⁵² [a bilingual dictionary for language production in the target language ideally would list collocations under the base entry].

²⁵³ [proporcionar una lista de colocativos en la entrada de una base sirve para producir textos, mientras que indicar varias bases en la entrada de un colocativo es útil para la comprensión del texto].

7.1.3.2 *Microestrutura*

A microestrutura corresponde, segundo Landau (2001, p. 99), à “organização da informação dentro de cada verbete do dicionário”²⁵⁴. Esse componente lexicográfico deve oferecer informações sobre as propriedades formais e semânticas da palavra-entrada (ou lema), constituindo o “comentário de forma” e o “comentário semântico”, respectivamente (cf. HAUSMANN, WIEGAND, 1989). No entanto, de acordo com Hartmann, James (1998, s.v. *microstructure*), “os usuários podem não ter habilidade de consulta suficiente para compreender a complexidade da microestrutura e podem precisar de orientação e/ou instrução explícitas para encontrar e extrair os detalhes requeridos”²⁵⁵. A fim, então, de facilitar as consultas por parte do usuário, é importante que os dicionários apresentem um programa constante de informações (PCI) em sua microestrutura. A partir da adoção de um PCI, “os tipos de respostas são sempre dados na mesma ordem e com a mesma formalização (...), pois o acesso às informações deve ser rápido, econômico e não-ambíguo”²⁵⁶ (DUBOIS, DUBOIS, 1971, p. 10). Dessa forma, quando o dicionário utiliza um PCI, a presença ou ausência de determinada informação é um dado relevante para o consulente, pois toda a informação tem um valor funcional em relação ao público e à função da obra.

Para identificar a relevância da presença ou ausência de informações em um verbete, porém, é necessário estabelecer a distinção entre “microestrutura abstrata” e “microestrutura concreta”, proposta por Hausmann, Wiegand (1989). A primeira “é aquela parte da estrutura completa do verbete que seleciona e ordena os (tipos de) itens da microestrutura concreta correspondente”²⁵⁷ (HAUSMANN, WIEGAND, 1989, p. 344-345), enquanto a segunda representa o conjunto de segmentos que um verbete de fato apresenta. A microestrutura concreta vai variar, então, de acordo com o lema tratado, bem como com o público-alvo e a função da obra.

Essas concepções, que orientam a disposição das informações pertinentes dentro de cada verbete, são pensadas para o tratamento lexicográfico de palavras isoladas, sem que haja um espaço destinado à inclusão da combinatória léxica. Em vista disso, a fim de

²⁵⁴ [organization of the information within each dictionary article].

²⁵⁵ [users may not have sufficient reference skills to follow the intricacies of the microstructure, and may need explicit guidance and/or instruction to find and extract the details required].

²⁵⁶ [les types de réponses sont toujours donnés dans le même ordre et avec la même formalisation (...), car l'accès aux informations doit être rapide, économique et non-ambigu].

²⁵⁷ [is that part of the complete article structure which selects and orders the (classes of) items of the corresponding concrete microstructure].

proporcionar um tratamento sistemático das combinações de palavras, Hausmann, Werner (1991 apud WELKER, 2004, p. 109) propõem uma divisão da microestrutura “básica” a partir de quatro diferentes formas de apresentação da combinatória léxica. São elas:

- a) microestrutura **integrada**: em que as combinações são apresentadas após cada acepção;
- b) microestrutura **não integrada**: em que as combinações aparecem ao final do verbete, constituindo um bloco à parte;
- c) microestrutura **semi-integrada**: inclusão das combinações ao final do verbete, mas com remissões a cada acepção particular, permitindo, assim, uma melhor identificação de suas relações;
- d) microestrutura **parcialmente integrada**: segue a organização da microestrutura integrada, mas inclui algumas combinações ao final do verbete, por não identificar a qual acepção estão relacionadas.

Para Marello (1996, p. 44), as organizações de tipo (a) e (d) dificultam a localização da informação por parte do consultante, obrigado a realizar a leitura de todo o verbete antes de chegar à informação de que necessita. Por outro lado, a adoção de uma microestrutura dos tipos (b) e (c) agiliza o acesso às combinações léxicas.

Apesar dessas diferentes propostas para a inclusão das combinações sintagmáticas em obras lexicográficas, no caso específico das colocações, de acordo com Wouden (1992, p. 449), não há uma sistematicidade em sua apresentação em dicionários bilíngües. Heid (1996, p. 117-118) acredita que, em virtude do fato de não possuírem uma seção especial para sua inclusão, as colocações costumam aparecer de forma implícita nos dicionários²⁵⁸, o que exige do usuário um amplo conhecimento metalexigráfico para sua identificação. Haensch (1982, p. 504), por sua vez, indica que as colocações podem aparecer de três formas distintas: incluídas nos exemplos, após a definição (dentro de cada acepção) ou listadas ao final do verbete²⁵⁹.

Hausmann (1998a, p. 71) acredita que, devido ao caráter transparente das colocações, não há necessidade de maiores informações que a simples lista de colocados sob a entrada da base. A inclusão dessas combinações em exemplos apenas ocupa, segundo Hausmann (1998b, p. 107), um espaço desnecessário, que poderia ser utilizado para a inserção

²⁵⁸ Um exemplo dessa organização pode ser encontrado em CCALED (2003, s.v. *smoker*), em que não há uma explicitação da colocação *heavy smoker* [fumante inveterado], mas apenas sua inclusão como exemplo.

²⁵⁹ Através de uma análise das colocações em dicionários bilíngües e monolíngües entre as línguas inglesa e espanhola, Corpas Pastor (1992) identifica catorze maneiras diferentes de inclusão das colocações, muitas delas, inclusive, dentro de uma mesma obra.

de mais colocações. No entanto, para Hanks (1998, p. 107), a utilidade dos exemplos varia de acordo com o público do dicionário:

Se nos dirigimos a principiantes, ou estudantes de nível intermediário, penso que um exemplo bem escolhido, representando algo que é típico do que está no *corpus*, é a melhor maneira de fazer entender a fraseologia apropriada. Se temos um usuário de nível avançado, então, penso que a maneira de proceder é dar uma lista de palavras de nível superior²⁶⁰.

Também para Lo Cascio (1997, p. 76), que analisa as colocações no dicionário holandês *Van Dale*, a inclusão dessas combinações nos exemplos é baseada na “estratégia que o usuário adota para procurar as informações de que necessita”²⁶¹. Heid (1997, p. 50) adverte, porém, para o fato de não haver uma correta distinção nos dicionários entre os exemplos colocacionais e os que não representam colocações²⁶².

A proposta de Hausmann (1998a, p. 71) de apresentar as colocações sob a forma de lista encontra também críticas em Haensch (1982, p. 504), para quem a inclusão de todas as colocações no final do verbete torna desnecessária a separação deste em diversas acepções. Esse autor propõe que as combinações léxicas sejam apresentadas imediatamente após cada acepção a qual estão relacionadas, conformando uma microestrutura integrada, na terminologia de Hausmann, Werner (1991 apud WELKER, 2004, p. 109). Também Heid (1996, p. 134) considera de maior proveito para o usuário se as colocações aparecem relacionadas com as diferentes acepções da palavra-entrada.

Em virtude do acima exposto, acreditamos que, no tratamento lexicográfico das colocações em um dicionário bilíngüe ativo português-espanhol, seja fundamental relacionar cada combinação com as diferentes acepções da palavra-entrada, visto que as bases das colocações mantêm seu significado original. A partir do momento que o dicionário estabelece essas relações, concordamos com Hausmann (1998a) de que seja suficiente apresentar somente a colocação em L1, acompanhada da combinação equivalente em L2²⁶³, sem a presença de qualquer outro segmento informativo como exemplos ou definições. Porém, nos casos em que para uma colocação em L2 não exista uma equivalência em L1, como na combinação anteriormente mencionada de *encuentro maratónico*, em língua espanhola, cujo significado é “jogo ou evento de longa duração”, acreditamos que a utilização de exemplos

²⁶⁰ [Se nos diriximos a principiantes, ou estudantes de nivel intermedio, penso que un exemplo ben escollido representando algo que é típico do que está no *corpus* é a maneira máis axeitada de facer entende-la fraseoloxía apropiada. Se temos un usuario dun nivel avanzado, entón, penso que a maneira de proceder é dar unha lista de palabras de rango superior].

²⁶¹ [strategia che l'utente adotta per cercare le informazioni di cui ha bisogno].

²⁶² Opinião semelhante apresenta Roberts (1996, p. 190), para quem a inclusão das colocações em exemplos não permite distingui-las das combinações livres.

²⁶³ Para um exemplo dessa aplicação, v. 7.3.2a.

seja importante para ilustrar o uso de tais combinações léxicas²⁶⁴, garantindo, assim, seu emprego adequado na produção em L2 por parte do falante nativo de língua portuguesa²⁶⁵. Varantola (2003, p. 231) salienta que exemplos sistematicamente extraídos de um *corpus* acrescentam informações essenciais a um falante não-nativo.

No entanto, consideramos que a forma de apresentação das colocações nos verbetes deva ser diferenciada, conforme a obra seja destinada a usuários de nível mais básico ou mais avançado. Para os primeiros, cujas obras contam com uma menor densidade informativa nos verbetes, acreditamos que as colocações devam aparecer em uma microestrutura integrada, de forma a facilitar seu acesso a essas combinações.

Por outro lado, os usuários que possuem um nível de conhecimento mais avançado na língua estrangeira necessitam de uma obra de consulta que lhes possibilite acessar uma maior quantidade de informações para auxiliá-los na produção textual. Em vista disso, a microestrutura apresenta uma maior densidade informativa, contando com segmentos que indicam marcações sintáticas²⁶⁶, nuances semânticas, variações fonológicas, entre outros aspectos que facilitam o processo de produção. Para ilustrar a grande quantidade de informações que um dicionário pode oferecer, apresentamos a seguir o verbete de *tocar*, extraído do DiBU (2003), que conta com mais de 10 acepções e uma série de informações complementares.

²⁶⁴ Também Farias (2008), salienta a importância dos exemplos para auxiliar na produção de colocações. Essa autora, no entanto, não oferece propostas para a apresentação dessas combinações nos exemplos.

²⁶⁵ Para um exemplo dessa aplicação, v. 7.3.2c.

²⁶⁶ Cabe ressaltar que, no caso específico das marcas sintáticas, a densidade informativa é ainda maior em uma obra dirigida a usuários de nível mais básico.

to·car 1 *tr.-intr.-prnl.* *Tocar, palpar: tocou o/no tecido e viu que era mesmo seda=tocó la tela y vio que era seda de verdad; pessoas não se contentam em olhar, precisam tocar=algunas personas no se contentan con mirar, tienen que tocar; tocava-se regularmente para ver se tinha alguma alteração na mama=se palpaba regularmente para comprobar si notaba algún cambio en el pecho.* *apalpar, sentir.* *Se dice tocar algo, em algo. – 2 tr.-intr.* *Tocar: você toca violão?=¿tocas la guitarra?; o sacristão tocou o sino para chamar os fiéis=el sacristán tocó la campana para llamar a los fieles; o pianista tocava muito bem=el pianista tocaba muy bien.* *soar, tanger.* 3 *Sonar, dar, tocar: o relógio tocou as doze badaladas=el reloj dio las doce campanadas; os sinos tocavam sempre às seis horas da tarde=las campanas sonaban siempre a las seis de la tarde.* *dar, entoar, soar.* 4 *Seguir, irse, marcharse, arrancar: toca rápido para o escritório! – disse o patrão ao motorista=vamos rápido a la oficina! – le dijo el jefe al conductor; entrei no táxi e mandei o motorista tocar=me metí en el taxi y le mandé al taxista que arrancara.* *ir, partir, seguir.* *Se dice tocar para.* 5 *Tocar, tratar: agora você tocou no/o ponto principal da questão=ahora has tratado el punto principal de la cuestión; eu o proíbo de tocar nesse assunto novamente=te prohíbo tocar ese asunto nuevamente.* *abordar, mencionar, referência, referir.* *Se dice tocar algo, em algo. – 6 intr.* *Sonar: o telefone tocou, mas ninguém uen-deu=el teléfono sonó, pero nadie respondió; tocou o despertador, é hora de acordar=ha sonado el despertador, es hora de despertarse.* *sinal, soar.* 7 *Dar la razón, felicitar: já que você gostou da minha sugestão, toque aqui=ya que te ha gustado mi sugerencia, felicítame. – 8 tr.* *Afectar, sensibilizar, impressionar: a notícia do acidente tocou-o muito=la noticia del accidente le afectó mucho; suas palavras tocaram fundo em mim=sus palabras me impresionaron profundamente.* *abalar, comover, impressionar.* 9 *Tocar, corresponder: com a morte do pai, tocou-lhe cuidar da casa e da família=con la muerte de su padre, le tocó cuidar de la casa y de la familia; esse problema toca à justiça=ese problema le corresponde a la justicia; por sorte, tocou-lhe a maior parte da fortuna deixada pelo tio=por suerte le tocó la mayor parte de la fortuna que dejó su tío.* *cabere, competir, incumbir, interessar.* *Se dice tocar a:* 10 *Tocar, llegar, golpear, chocarse: o barco tocou a margem oposta rapidamente=el barco llegó a la orilla opuesta rápidamente; o carro tocou o/no poste=el coche se chocó contra el poste.* *atingir.* *Se dice tocar algo, em algo.* 11 *BRA.* *Echar, estimular, incitar, golpear: tocou o cachorro para fora da cozinha=echó al perro de la cocina; tocou o marido para fora de casa=echó a su marido fuera de la casa; tocou os cavalos com o chicote=golpeó a los caballos con la fusta.* *expulsar, incitar.* 12 *Mover, echar: o jogador tocou a bola para o campo do adversário=el jugador echó la pelota al campo del adversario; vamos ver se tocamos esse projeto=vamos a ver si movemos ese proyecto. – 13 tr.-prnl.* *Inspirar, comover, afectar, sentir: o poema tocou-me profundamente=el poema me conmovió profundamente; foi tocada pelas palavras do padre e converteu-se=la conmovieron las palabras del cura y se convirtió; tocou-se muito com o que disse dele aquele que julgava ser o seu melhor amigo=le afectó mucho lo que dijo de él el que creía su mejor amigo.* *atingir, comover, magoar, sentir.* 14 *Tocar, rozar, probar: nem toquei na/a comida=ni toqué la comida; por definição, duas retas paralelas nunca se tocam=por definición, dos rectas paralelas nunca se tocan.* *contato, encontrar, encostar, roçar. – 15 tocar-se BRA. prnl.* *Darse cuenta, percatarse, fijarse: toquei-me do erro e corriji-o imediatamente=me di cuenta del error y lo corregí inmediatamente.* *perceber.* *Se dice tocar-se de.* *En algunos contextos, es infor.: vê se se toca, meu!=ja ver si te fijas más, tío!; □ não se tocar, no darse cuenta, no percatarse: ele não se tocava de que estava incomodando a todos=ni se daba cuenta de que estaba molestando a todo el mundo; esse cara não se toca mesmo, só dá fora=ese tipo no se da cuenta: siempre mete la pata.* *descaso, desligado, indiferença; □ ser cheio de não me toques, ser de mírame y no me toques, ser quisquilloso: a Luana é insuportável, cheia de não me toques=Luana es insoportable: es de mírame y no me toques.* *fresco, frescura.* *Como 14.*

Figura 3: Modelo de verbete retirado do DiBU (2003, s.v. *tocar*)

É importante ressaltar que essa obra não funciona como um modelo ideal de dicionário ativo, especialmente porque pretende atingir aos falantes dos dois idiomas (português e espanhol), desempenhando as funções tanto de produção quanto de recepção, mas, como as informações que apresenta são também pensadas para “facilitar o processo de codificação em espanhol” (DiBU, 2003, p. VII), consideramos que esse dicionário possa ser utilizado para exemplificar a densidade informativa de um verbete nesse tipo de obra. Em vista disso, salientamos o fato de que o que está sob análise é somente a quantidade de informações, ainda que estas, muitas vezes, não se mostrem relevantes para o usuário nem estejam dispostas da maneira mais adequada.

Dessa forma, como podemos observar no verbete anterior, a marcação sintática é oferecida através da regência verbal (*tr.*, *intr.*, *prnl.*) no “comentário de forma”; no entanto, muitas vezes, somente a informação da transitividade do verbo não demonstra ser suficiente para auxiliar o usuário a produzir na língua estrangeira. Em vista disso, o DiBU (2003) também apresenta, por meio de “informações lingüísticas e de uso” - indicadas pelo símbolo de um lápis -, os tipos de complementos que cada acepção exige. Cabe salientar, porém, que essas indicações nem sempre constituem informações discretas e discriminantes²⁶⁷, como as que aparecem já na primeira acepção: “se dice *tocar algo, em algo*”, úteis somente para um usuário de língua espanhola que precisasse produzir em língua portuguesa. Com relação à indicação das sutilezas semânticas que diferenciam cada acepção, o dicionário verificou a necessidade de apresentar alguns vocábulos que mantivessem uma relação sinonímica ou compartilhassem o mesmo campo semântico com a palavra-entrada em cada uma de suas acepções, pois a simples inclusão dos equivalentes não era suficiente para estabelecer uma diferenciação precisa. Também o aspecto morfológico aparece contemplado ao final desse verbete através da indicação de sua conjugação “como 14”, remetendo ao verbo *ficar* que aparece no *Back Matter* da obra.

É importante salientar que cada dicionário particular selecionará uma forma específica de apresentar essas e outras informações que considerar pertinentes para auxiliar no processo de produção textual. Assim, devido às decisões tomadas pelos lexicógrafos, as informações aparecerão dispostas de forma diferenciada em cada obra. No entanto, percebemos que, para suprir de modo satisfatório as dificuldades de produção dos usuários, é necessária a incorporação de uma grande quantidade de informações que sejam, como já

²⁶⁷ Por informações “discretas e discriminantes”, entendemos aquelas informações que sejam efetivamente relevantes para o consulente e lhe possibilitem um maior conhecimento sobre o uso da língua (cf. BUGUEÑO, FARIAS, 2006).

ressaltamos anteriormente, discretas e discriminantes, apesar da crescente preocupação, principalmente em razão das exigências do mercado editorial, em apresentar tais informações no menor espaço possível.

Essa grande densidade informacional, embora seja mais característica dos verbetes de verbos, como o que apresentamos, também pode ocorrer para os substantivos e adjetivos, classes de palavras que enfocamos neste trabalho, conforme podemos observar através da marcação de nuances semânticas, indicadas pelos parênteses, e de “informações lingüísticas e de uso” no verbete de *coral* presente no DiBU (2003):

co•ral 1 adj.-m. Coral: *o canto coral exige muitos ensaios=el canto coral exige muchos ensayos; o coral da escola apresentou-se com músicas belíssimas na formatura dos alunos=la coral de la escuela interpretó canciones bellísimas en la graduación de los alumnos.* ☞ **coro.** – **2 m.-adj.** (color) Coral: *Beatriz ganhou um colar de contas de coral que combinava perfeitamente com a blusa coral que comprara para sua festa de aniversário=a Beatriz le regalaron un collar de cuentas de coral que hacía juego con la blusa coral que había comprado para su fiesta de cumpleaños.* – **3 f. BIOL.** (serpiente) Coral: *a coral se escondeu no meio das pedras=la coral se escondió entre las piedras.* ☞ Se usa también **cobra-coral; falsa coral, falsa coral.** ☞ **cobra, serpente.** ☞ pl. **corais.**

Figura 7: Modelo de verbete retirado do DiBU (2003, s.v. *coral*)

Consideramos, porém, que essa breve reflexão a respeito dos verbetes de verbos seja pertinente porque as propostas de tratamento lexicográfico das colocações que apresentaremos em 7.3, embora tratem somente de combinações entre substantivos e adjetivos, também podem (e devem) ser ampliadas às demais classes de palavras para sua inclusão em dicionários.

Em virtude, então, da grande quantidade de informação que os verbetes de um dicionário bilíngüe ativo para usuários de nível avançado deverão apresentar, acreditamos que a melhor forma de apresentação dessas combinações léxicas seria a partir de uma microestrutura semi-integrada. Esse tipo de organização permite ao consulente um acesso agilizado à combinação de que necessita, no momento de sua produção textual, ao mesmo tempo que possibilita estabelecer a relação de cada combinação com as diferentes acepções da palavra-entrada, caso o usuário inicie sua busca pelas colocações a partir da descoberta do equivalente do lema.

7.1.3.3 Medioestrutura

Com relação à medioestrutura, Hartmann, James (1998, s.v. *cross-reference structure*) a definem como “a rede de referências cruzadas que permite aos redatores e aos usuários de uma obra de consulta localizar o material disperso nas diferentes partes que compõem [sc. essa obra]”²⁶⁸. Assim, o conceito de medioestrutura introduzido por Wiegand (1996 apud BUGUEÑO, 2003, p. 3) corresponde não apenas a um sistema de referência entre as diferentes partes do dicionário, mas também deve ser entendido como uma intersecção entre o dicionário e o público-alvo, de forma que o consulente encontre a informação que procura de forma fácil e rápida.

Como o conceito de medioestrutura é relativamente recente em lexicografia, não há uma grande variedade de estudos sobre essa temática, o que se reflete também na ausência de trabalhos sobre o tratamento das colocações que levem em conta os problemas de ordem medioestruturais. Assim, as únicas referências que encontramos a um sistema de remissivas entre as colocações encontram-se em Heid (1996, p. 136) que, ao analisar diferentes dicionários bilíngües, verifica que tais obras não utilizam “referências cruzadas” [cross-references] e em Cop (1991, p. 2777), pois essa autora, ao defender a inclusão das colocações tanto sob a entrada da base quanto do colocado, verifica a necessidade de utilizar um “sistema de referência cruzada” [cross-referencing system] de forma a economizar espaço. Nesse caso, é necessário que se estabeleça um sistema de remissivas (medioestrutura) bastante coerente entre bases e colocados, de modo que a obra lexicográfica não venha a repetir ou privar o usuário de importantes informações.

No tratamento lexicográfico das colocações em um dicionário bilíngüe ativo português-espanhol, porém, acreditamos ser suficiente apenas a inclusão dessas combinações sob a entrada da base. Considerando que as colocações representam uma dificuldade na produção de textos, e que os falantes partem da base para organizá-los, não nos parece necessária a presença dessas combinações na entrada do colocado e, conseqüentemente, a utilização de um sistema de remissivas entre essas duas unidades léxicas.

²⁶⁸ [The network of cross-references which allows compilers and users of a reference work to locate material spread over different component parts].

7.1.3.4 *Partes introdutórias*

As partes introdutórias do dicionário, também chamadas de *Front Matter*, correspondem, segundo Hausmann, Wiegand (1989, p. 330), a “todas as partes textuais funcionais que precedem a lista (central) de palavras”²⁶⁹. Fornari (2006b) identifica que esse componente do dicionário deve cumprir duas tarefas básicas, que são complementares entre si:

- a) Oferecer ao leitor um panorama sobre o objetivo que o dicionário pretende cumprir;
- b) Funcionar como um manual de instruções acerca da utilização do dicionário por parte do usuário.

Para que as partes introdutórias, então, cumpram suas tarefas básicas, é indispensável que *a priori* haja uma delimitação precisa do público ao qual o dicionário pretende atingir, bem como da função específica que assume (para a produção, recepção ou ambas). Essas concepções prévias, como já foi dito anteriormente, deverão orientar a elaboração de todos os componentes do dicionário e, a partir disso, as partes introdutórias devem refletir as relações estabelecidas entre os demais componentes, de forma a facilitar o acesso à informação por parte do usuário.

As partes introdutórias somente poderão ser consideradas um instrumento facilitador se as informações apresentarem uma perfeita harmonia entre abrangência e concisão. Welker (2004, p. 80) salienta que não há consenso entre os lexicógrafos a respeito da “introdução”, pois

alguns recomendam que ela seja bem sucinta porque o usuário não quer se dar ao trabalho de ler um texto longo (...), outros preconizam uma apresentação bem detalhada porque só assim o usuário saberá aproveitar todas as informações contidas no dicionário.

De qualquer forma, ressalta que, independentemente da extensão, a informação presente nas partes introdutórias deve ser clara.

Fornari (2006b), por sua vez, ao trabalhar com os conceitos de “abrangência” e “concisão”, propõe uma posição intermediária com relação ao conteúdo e à extensão das partes introdutórias. A “abrangência” relaciona-se à qualidade das informações e ao seu teor de relevância para o usuário, ou seja, as informações apresentadas devem ser relevantes para o consultante e ter a capacidade de orientá-lo a respeito da estrutura e dos conteúdos do

²⁶⁹ [all functional text parts preceding the (central) word list].

dicionário. Já o termo “concisão” está relacionado à questão da quantidade das informações, isto é, as partes introdutórias devem orientar o usuário de forma objetiva, procurando alcançar o maior poder de síntese possível. Dessa forma, as informações apresentadas nas partes introdutórias devem, de modo sucinto, orientar a consulta ao dicionário por parte de seu usuário.

A partir dessas concepções, uma proposta de tratamento lexicográfico das colocações aplicada a dicionários bilíngües deve deixar claro em suas partes introdutórias a função e o público ao qual a obra se destina. Assim, o usuário deve estar ciente de que se trata de um dicionário bilíngüe de tipo ativo, isto é, auxiliar no processo de codificação, cogitado para estudantes brasileiros de língua espanhola com um grau de conhecimento correspondente aos níveis A2 ou B2 do MARCO (2002), na língua estrangeira. Cabe lembrar que tais usuários terão acesso a duas obras lexicográficas distintas. Além disso, é importante que, nas partes introdutórias dos dicionários, sejam especificados os critérios de definição macroestrutural e de lematização (v. 7.1.3.1) utilizados na obra. O primeiro evidenciará os tipos de combinações às quais o usuário terá acesso, enquanto o segundo indicará a forma como as colocações serão incluídas. A fim de facilitar a consulta por parte do usuário, os critérios de lematização podem ser apresentados através de um exemplo de verbete, que serve para ilustrar também sua forma de organização interna. Assim, acreditamos realizar as duas tarefas básicas que, segundo Fornari (2006b), as partes introdutórias devem cumprir. Porém, em nosso estudo específico das colocações, consideramos pertinente incluir, ainda, nesse componente do dicionário, informações sobre os critérios de seleção e as fontes de coleta das combinações léxicas incluídas (cf. SANTAMARÍA PÉREZ, 2000, p. 98).

7.2 Necessidade de um novo tratamento lexicográfico das colocações: análise de dicionários bilíngües espanhol-potuguês/português-espanhol

Em virtude das diferentes propostas para o tratamento lexicográfico das colocações, realizamos uma análise da apresentação dessas combinações léxicas em seis dicionários bilíngües espanhol-potuguês/português-espanhol, quais sejam, DiBU (2003), DBE (2005), DBLA (2001), DiCP (2003), MiFF (2001) e MiPD (1996). Os resultados desse estudo, apresentados em Beneduzi (2007), permitiram-nos perceber que:

- Dos seis dicionários analisados, apenas DBLA (2001) e MiFF (2001) restringem seu público-alvo a estudantes brasileiros do espanhol de nível básico e médio, respectivamente; os demais pretendem atingir aos falantes dos dois idiomas, sem especificar o grau de conhecimento que devem ter de cada língua.
- Com relação à função que pretendem desempenhar, DiBU (2003), DBLA (2001) e DiCP (2003) – os únicos dicionários que fornecem essa informação – visam a auxiliar o usuário tanto na recepção quanto na produção de textos. A tentativa de servir para diferentes funções, então, faz com que não haja um padrão de organização para a seleção e inclusão das colocações de acordo com o propósito da obra.
- No plano microestrutural, a maior parte dos dicionários inclui as colocações ao final do verbete em negrito; no entanto, essa forma de apresentação não permite diferenciar as colocações das combinações fixas. Assim, em DiBU (2003), por exemplo, a combinação fixa *pegar la hebra* (s.v. *hebra*) [bater papo] e a colocação *dar un hervor* (s.v. *hervor*) [dar uma fervida] apresentam o mesmo *status*.
- Outra forma de apresentar as colocações é incluí-las nos exemplos; porém, os dicionários não oferecem nenhum destaque a essas combinações. Em DiBU (2003, s.v. *señal*), por exemplo, o usuário somente perceberá que esse substantivo é utilizado com o verbo-suporte “hacer” [fazer] ao identificar sua recorrência nos exemplos:

1 marca, sinal: *voy a hacer una señal en el libro para saber por dónde voy.*
 2 sinal: *un hombre nos hacía señales para que fuéramos despacio.* (DiBU, 2003, s.v. *señal*)
- No plano macroestrutural, não há um padrão para nenhum dos dicionários analisados. Embora a grande maioria das colocações seja incluída sob a entrada da base, existem alguns casos em que os dicionários indicam essas combinações sob o colocado, sem que haja uma explicação para tal fato. A colocação *surtir efecto* (s.v. *surtir*) [surtir efeito], por exemplo, aparece inserida na entrada do colocado em todos os dicionários analisados.
- No plano medioestrutural, os dicionários analisados não costumam utilizar um sistema de remissivas que permita estabelecer as relações entre bases e colocados ou, até mesmo, entre os verbetes da direção espanhol-português e os da direção português-espanhol.

- Além de uma assistemática na inclusão das colocações nos dicionários, ocorre uma arbitrariedade em sua seleção. Acreditamos que muitas combinações não eram adequadas às necessidades do público ao qual as obras se destinavam, pois, devido à semelhança entre as línguas, não representariam dificuldades tanto para o falante nativo do português que está aprendendo o espanhol quanto para o falante do espanhol que estuda o português.
- Além de uma seleção inapropriada das combinações, a falta de homogeneidade na apresentação das colocações também pode dificultar a compreensão por parte do usuário. Essa falta de homogeneidade diz respeito tanto à utilização de diferentes recursos gráficos - em uma mesma obra - para identificar as colocações quanto à falta de relação entre as duas partes dos dicionários. Com relação a essa falta de simetria entre as partes do dicionário, em DBE (2005, s.v. *sessão*), na direção português-espanhol, encontramos a informação *encerrar a sessão*, seguida da forma equivalente *levantar la sesión*; no entanto, na direção espanhol-português, encontramos as combinações *abrir/cerrar la sesión* (DBE, 2005, s.v. *sesión*) [abrir/encerrar a sessão]. O uso de *levantar la sesión* em uma direção e da combinação *cerrar la sesión* em outra, ambas significando “encerrar a sessão”, impede o leitor de ter certeza de qual construção é realmente utilizada em língua espanhola.

A partir dessas constatações, verificamos que os dicionários bilíngües espanhol-português/português-espanhol, disponíveis para o mercado editorial brasileiro, não levavam em consideração aspectos fundamentais em lexicografia para a elaboração de suas obras. A ausência de princípios orientadores básicos, tais como a delimitação de um perfil de usuário e a função do dicionário, não permitia que houvesse um embasamento teórico para justificar a forma de inclusão das colocações no que diz respeito aos diferentes componentes canônicos do dicionário. Essas deficiências resultam em uma arbitrariedade na seleção e assistemática na apresentação das colocações nos dicionários bilíngües analisados. Em virtude desse fato, consideramos que seja pertinente e, inclusive, de grande relevância, um estudo aprofundado dessas combinações léxicas que possibilite uma melhor forma de sua apresentação em obras lexicográficas. Em vista disso, procuramos oferecer uma proposta para o tratamento das colocações em um dicionário bilíngüe ativo português-espanhol, de forma a facilitar a consulta e a aprendizagem dessas combinações por parte dos estudantes brasileiros de língua espanhola.

7.3 Propostas para o tratamento lexicográfico das colocações em dicionários bilíngües ativos português-espanhol

Os resultados obtidos nas análises realizadas ao longo de nossa dissertação para a identificação das colocações devem ser aplicados, de forma prática, na geração de parâmetros que embasem uma proposta de apresentação dessas combinações léxicas em dicionários bilíngües. Para isso, é pertinente que sejam atrelados os três seguintes aspectos:

- (a) a freqüência de coaparição dos elementos,
- (b) as relações semânticas estabelecidas entre substantivos e adjetivos e
- (c) a conformação diferenciada das combinações em contraste entre as línguas portuguesa e espanhola.

Devemos salientar, porém, que esses fatores apresentarão uma relevância diferenciada de acordo com os tipos de colocações (v. 6.4.1) existentes entre as línguas. Assim, em alguns casos, será importante identificar as diferenças de freqüência entre duas combinações, enquanto em outros, o dicionário destacará a relação semântica entre os elementos.

Além desses aspectos, nossas propostas de apresentação das colocações em dicionários bilíngües devem considerar o público ao qual essa obra pretende atingir, pois, conforme explicitado em 7.1.2, essa é uma das primeiras decisões que o lexicógrafo deve tomar ao iniciar a elaboração de um dicionário. No presente trabalho, ao refletir sobre um tratamento adequado das colocações em obras lexicográficas, percebemos a necessidade de incluir tais combinações em dicionários voltados a usuários com diferentes níveis de conhecimento na língua estrangeira. Em vista disso, é fundamental que a inclusão das colocações em dicionários bilíngües ativos português-espanhol seja pensada a partir de duas obras distintas: uma voltada a um público de nível “Plataforma” ou “A2” e outra dirigida a usuários de nível “Avançado” ou “B2”. Salientamos, porém, que ambas as obras são destinadas a estudantes brasileiros de língua espanhola; o que varia, no entanto, é apenas o domínio que esses estudantes têm da língua estrangeira. Consideramos pertinente essa separação entre dois grupos de usuários diferentes porque a delimitação de um público específico orientará a concepção dos diversos componentes canônicos do dicionário, conforme veremos nos exemplos que seguem.

7.3.1 Dicionário bilíngüe ativo português-espanhol para usuários de nível “Plataforma”

Este tipo de dicionário, voltado a um público de nível mais básico, deve oferecer colocações que apresentem um caráter mais simples e usual (v. 7.1.2.2). Em vista disso, as combinações incluídas na macroestrutura dessa obra correspondem a casos de divergências na posição do adjetivo (v. 6.4.1.2), na seleção do colocado (v. 6.4.1.3) e na estrutura morfológica (v. 6.4.1.4). Ficam, assim, excluídas as combinações com uma equivalência “literal” (v. 6.4.1.1) entre os idiomas, por não representarem uma dificuldade para os estudantes, bem como as que constituem uma combinação particular a uma língua (v. 6.4.1.5), devido ao fato de serem estruturas demasiado complexas para um aprendiz do nível “Plataforma”.

No plano microestrutural, como a quantidade de informações oferecidas dentro do verbete, neste tipo de obra, é reduzida (v. 7.1.3.2), consideramos mais apropriado incluir as colocações em uma microestrutura integrada, facilitando seu acesso por esse público, ainda pouco habituado a realizar consultas a dicionários. Porém, para chamar a atenção dos usuários para as ocorrências de tais colocações, de forma que não sejam confundidas com as diferentes acepções da palavra-entrada, destacamos a presença desse fenômeno léxico no dicionário através da utilização do símbolo (◊), o qual aparecerá identificado nas partes introdutórias dessa obra como indicador das colocações.

A seguir, apresentamos as diferentes propostas de organização dos verbetes segundo o tipo de colocação implicada.

a) Divergência na posição do adjetivo

Nestes casos, ainda que o significado dos elementos que conformam a combinação seja equivalente nos dois idiomas, os adjetivos aparecem em posições diferentes com relação aos substantivos que acompanham nas línguas portuguesa e espanhola. Dessa forma, para ressaltar ao usuário essa divergência na posição dos vocábulos, utilizamos o símbolo (), que deve aparecer devidamente explicitado nas partes introdutórias como indicativo de uma inversão na ordem da estrutura. O exemplo a seguir ilustra essa marcação das colocações.

plano, na 1 *adj.* Plano, llano, liso²⁷⁰
2 *m.* Llano, llanura **3 Plano** **4 Plano,**
 mapa **5 Plan,** proyecto, propósito **◆**
 ~ **de sucesso** exitoso plan **6 FIG.**
 Plano, situación, nivel **7 Plano**
LOC ~ de saúde seguro médico

Neste verbete, a colocação *exitoso plan* [plano de sucesso] aparece integrada na acepção 5, que retrata o significado expresso pela base da combinação. No entanto, o símbolo (◆) e o uso do negrito permitem um maior destaque à colocação, de forma que o usuário perceba rapidamente sua ocorrência. Na tentativa de economizar espaço, a palavra-entrada foi substituída pelo símbolo (~), recurso amplamente empregado na prática lexicográfica e que, conseqüentemente, consideramos não causar dificuldade aos consulentes.

No final do verbete, incorporamos um segmento informativo denominado “LOC”, que significa “locuções”, no qual incluímos todas as combinações léxicas fornecidas em DiBU (2003) para as quais não é possível estabelecer uma relação com as diferentes acepções da palavra-entrada. Temos consciência, porém, que, em muitos casos, aparecem mescladas expressões idiomáticas e unidades lexicalizadas, por exemplo, mas, conforme explicitado anteriormente, mantivemos esse segmento como apresentado em DiBU (2003).

É importante salientar, ainda, que a combinação *exitoso plan* [plano de sucesso], retratada em nosso exemplo, inclui-se, simultaneamente, nos três casos de estruturas presentes em um dicionário bilíngüe ativo português-espanhol para usuários de nível “Plataforma”. Isso ocorre porque, além de apresentar uma inversão na ordem das palavras, possui uma divergência na estrutura morfológica e também na seleção do colocado, pois em língua portuguesa, apesar de dispormos do vocábulo *exitoso*, sua combinação com *plano* não se realiza no nível da norma. No entanto, o fato dessa colocação poder ser classificada de maneiras diferentes não interfere na forma como será apresentada no verbete, pois, conforme veremos a seguir, para as demais estruturas presentes neste tipo de dicionário, são utilizadas as mesmas marcações gráficas - exceto, é claro, o símbolo de inversão ().

²⁷⁰ Como a geração de equivalências foge ao escopo deste trabalho, estes e os demais equivalentes apresentados nesta seção, bem como a separação entre as diferentes acepções de cada palavra-entrada e expressões idiomáticas apresentadas, foram retirados do DiBU (2003). No entanto, a estrutura geral do verbete oferecido nessa obra sofreu algumas alterações, como a exclusão de exemplos e sinônimos e a adaptação das informações lingüísticas adicionais, as quais, muitas vezes, dificultavam a visualização de nossas propostas para o tratamento das colocações devido a sua grande extensão.

b) Divergência na seleção do colocado

Nestes casos, embora ambos os idiomas possam apresentar adjetivos equivalentes, cada língua seleciona um adjetivo particular para se combinar com o substantivo em questão. Exemplo dessa divergência pode ser visto na proposta de verbete abaixo.

horário, ria 1 *adj.* Horario, por hora 2 *m.* Horario ◊ ~ **nobre**
 horario estelar **LOC carga** ~ carga
 horaria; **fuso** ~ huso horario

Neste caso, a língua espanhola, apesar de possuir o adjetivo *noble* [nobre], não costuma combiná-lo com *horario* [horário], como acontece na língua portuguesa, pois a norma daquele idioma selecionou o adjetivo *estelar* para exprimir o mesmo significado. Porém, a combinação *horario noble*, apesar de não ser utilizada no espanhol - somente 90 ocorrências no *Google* -, é a expressão oferecida por DiBU (2003, s.v. *horário*) como equivalente da colocação *horário noble*.

c) Divergência na estrutura morfológica

Os casos de combinações com uma estrutura morfológica diferenciada entre os idiomas correspondem a ocorrências em que uma das línguas apresenta uma combinação de substantivo + adjetivo, enquanto a outra substitui o adjetivo pela construção “de” + substantivo. Neste tipo de combinações entre as línguas portuguesa e espanhola, um dicionário bilíngüe de tipo ativo, voltado a falantes nativos do português, deve registrar apenas as ocorrências que constituem casos de colocações no espanhol. Acreditamos que somente essas combinações sejam relevantes, uma vez que a dificuldade para os estudantes brasileiros de língua espanhola será justamente a produção de colocações e não de combinações livres nesse idioma. Como o único exemplo que encontramos em nosso *corpus* para essa estrutura corresponde à combinação *exitoso plan* [plano de sucesso] (v. 6.4.1.4), reproduzimos novamente a proposta de verbete apresentada em (a).

<p>plano, na 1 <i>adj.</i> Plano, llano, liso 2 m. Llano, llanura 3 Plano 4 Plano, mapa 5 Plan, proyecto, propósito ◆ ~ de sucesso exitoso plan 6 FIG. Plano, situación, nivel 7 Plano LOC ~ de saúde seguro médico</p>

Nesse exemplo, a combinação livre formada pela preposição “de” e um substantivo em língua portuguesa gera uma combinação em língua espanhola que não apresenta um vocábulo equivalente em português. Em casos como esse, embora não seja um propósito dos dicionários bilíngües, consideramos que a inclusão da combinação em língua espanhola seja importante por proporcionar um enriquecimento vocabular por parte dos aprendizes mais iniciantes. Além disso, partindo de um dicionário português-espanhol, essa seria a única forma do usuário ter acesso a esse tipo de vocábulo que não encontra uma correspondência em sua língua nativa.

Finalmente, consideramos importante salientar que, como as combinações selecionadas para serem apresentadas especificamente aos estudantes do nível “Plataforma” possuem uma conformação bastante semelhante entre as línguas portuguesa e espanhola, acreditamos que seja suficiente apenas a indicação da equivalência na língua estrangeira, sem a necessidade de utilizar uma paráfrase definitória ou exemplos. Dessa forma, o usuário pode ter um acesso mais agilizado ao significado da combinação.

7.3.2 Dicionário bilíngüe ativo português-espanhol para usuários de nível “Avançado”

Como os usuários deste dicionário possuem um nível de conhecimento mais avançado na língua estrangeira, as combinações incluídas nesta obra apresentam um vocabulário mais específico (v. 7.1.2.2), refletido em colocações com uma posição divergente do adjetivo (v. 6.4.1.2), com uma seleção idiossincrática do colocado (v. 6.4.1.3) e as que representam uma combinação particular a cada idioma (v. 6.4.1.5). Ficam excluídas, assim, colocações que possuem um equivalente “literal” (v. 6.4.1.1) e aquelas cuja divergência entre as línguas ocorre somente de acordo com a estrutura morfológica (v. 6.4.1.4).

No programa constante de informações, por sua vez, em decorrência da grande quantidade de informações que os verbetes de um dicionário bilíngüe ativo para usuários de nível “Avançado” apresenta (v. 7.1.3.2), acreditamos que as colocações devam ser tratadas em

uma microestrutura semi-integrada, a fim de facilitar o acesso a essas combinações léxicas. Dessa forma, as colocações conformariam um bloco diferenciado de combinações, indicado através da denominação “COL”, o qual apareceria separado do segmento que contém as “locuções”, uma vez que, como já dissemos anteriormente, estas não podem ser relacionadas com cada acepção da palavra-entrada.

A seguir apresentamos nossas propostas de organização dos verbetes para este tipo de dicionário de acordo com as diferentes estruturas das colocações.

a) Divergência na posição do adjetivo

Nestes casos, em que somente o adjetivo apresenta uma posição diferenciada entre as línguas, salientamos essa divergência através do símbolo (), conforme exemplificado na proposta que segue.

erro <i>m.</i> Error COL ~ crasso craso error

A utilização de um segmento *à parte* para a inclusão das colocações parece constituir um recurso desnecessário em exemplos como o acima citado, devido ao reduzido tamanho do verbete. Porém, tendo em vista que toda obra lexicográfica deve apresentar um programa constante de informações, a manutenção dessa mesma forma de inclusão das colocações é importante para garantir uma correta interpretação das informações por parte do usuário.

Além disso, consideramos importante reiterar que o DiBU (2003), obra utilizada como fonte de coleta de nossos verbetes, não corresponde a um modelo ideal de dicionário para a produção de textos. Em virtude desse fato, o emprego de uma microestrutura semi-integrada torna-se ainda mais relevante nos casos de verbetes com uma maior densidade informativa, como, por exemplo, o que apresentamos a seguir.

b) Divergência na seleção do colocado

Nas combinações em que cada língua apresenta um colocado diferente, acreditamos que seja suficiente apenas a inclusão dos equivalentes em língua estrangeira, conforme o modelo abaixo.

prova 1 *f.* Prueba, señal **2** DER. Prueba **3** Examen, prueba **4** Prueba ensayo **5** Prueba; carrera **COL 1** ~ **cabal** prueba fehaciente/cabal **LOC a toda prova** probado, a prueba de bomba; *pôr à prova* poner a prueba; *tirar a prova dos nove* comprobar, cerciorarse, estar seguro

Nesse exemplo de verbete, é possível recuperar a relação da colocação com a acepção específica a qual pertence o substantivo da combinação, nesse caso, a primeira acepção. Além disso, fica evidenciada a diferença entre as colocações e as demais combinações sintagmáticas, incluídas no segmento “LOC”. Acreditamos que a forma como os diferentes segmentos aparecem dispostos auxilia, visualmente, na identificação das informações por parte dos usuários.

Finalmente, esse verbete destaca, ainda, a possibilidade de uma colocação em língua portuguesa apresentar mais de uma forma equivalente em língua espanhola, como acontece com *prova cabal* e suas equivalências *prueba fehaciente* e *prueba cabal*. Nos casos em que isso ocorre, oferecemos, primeiramente, a combinação mais usual, seguida daquela que apresenta uma menor frequência na língua estrangeira. Ressaltamos que, evidentemente, essas decisões de ordem microestrutural devem estar devidamente indicadas nas partes introdutórias do dicionário.

c) Combinação particular a uma língua

As combinações que não apresentam uma forma equivalente na língua nativa do usuário devem aparecer somente em dicionários voltados a um estudante de nível mais avançado. Isso se deve à maior complexidade tanto do fenômeno em si quanto a sua forma de apresentação em obras lexicográficas. Para chamar, então, a atenção do consultante para essas colocações exclusivas da língua estrangeira, utilizamos o símbolo () para introduzir tais combinações. A seguir oferecemos um exemplo de verbete que ilustra tal procedimento.

amizade 1 f. Amistad **2 P. EXT. voc.** Amistad, amigo **3 pl.** Relación
COL 1 ~ [profunda] amistad
 entrañable: *El médico y el abogado, después de haber mantenido una amistad entrañable, desde hace ya varios años han sostenido una soterrada disputa* (MER) **LOC**
amizade colorida amigo, amante

Nos casos de combinações particulares a um idioma, a inexistência de uma estrutura equivalente na língua nativa do estudante impede que seja utilizada uma combinação em sua lematização. Em vista disso, optamos por apresentar, entre colchetes, algum vocábulo que denote o tipo de relação semântica que o adjetivo estabelece com o substantivo. Além disso, conforme explicitado anteriormente, nessas combinações especificamente, consideramos importante a inclusão de um exemplo que ilustre o uso da colocação na língua estrangeira, pois esse recurso confere maior segurança no emprego da combinação por parte do usuário. A colocação é destacada através de seu sublinhado e, acompanhando os exemplos, oferecemos a sigla do jornal que indica de onde cada exemplificação foi retirada. É fundamental que todas as abreviaturas apareçam devidamente explicitadas nas partes introdutórias do dicionário, bem como as fontes de onde foram coletados os exemplos. Finalmente, devido ao grau de proficiência do consulente desse tipo de obra ser mais elevado, não nos parece relevante que seja oferecida uma tradução dos exemplos.

Consideramos importante salientar que, para algumas combinações, como as que pertencem ao grupo dos adjetivos intensificadores, é possível utilizar um vocábulo que denote as relações estabelecidas entre estes adjetivos e os substantivos que acompanham de forma homogênea, como acontece com “profundo”, empregado em nosso exemplo, ou “intenso”. No entanto, tal solução representa apenas uma pequena parcela das colocações formadas por adjetivos e substantivos, nas quais o emprego de um único vocábulo é insuficiente para cobrir o significado expresso pelo adjetivo em relação ao substantivo com que se combina, como podemos observar no exemplo a seguir.

tema m. Tema **COL ~ [atual e que provoca discussão]** tema candente: *El salario mínimo es hoy un tema candente en EU* (UNIM)

Esse tipo de combinação demonstra que seria muito difícil que um dicionário conseguisse isolar um número limitado de “palavras-chave” capazes de indicar as relações estabelecidas entre substantivos e adjetivos. Essa dificuldade, já apontada pela insuficiência das quase 60 funções léxicas propostas por Mel’čuk (1988), parece ser um reflexo da ausência de uma classificação específica dos adjetivos que sirva de embasamento para a identificação de suas relações com os substantivos.

8 CONCLUSÃO

No início deste trabalho, fizemos uma revisão da literatura especializada sobre o tema das colocações, procurando obter subsídios que permitissem uma definição clara e objetiva dessas combinações. Percebemos, no entanto, que as características atribuídas pelos diferentes autores às colocações não proporcionavam uma delimitação precisa desse fenômeno léxico, visto que um consenso conceitual entre as diferentes teorias e abordagens era escasso. Em virtude desse fato, não foi possível obter subsídios que possibilitassem empregar uma metodologia para a identificação dessas combinações de forma mais objetiva. Com isso, percebemos a necessidade de aplicar, de forma consecutiva, diferentes metodologias sobre as combinações de nosso *corpus*, a fim de isolar aquelas que configuravam colocações.

Nosso próximo passo, então, foi realizar uma análise quantitativa dessas construções, a fim de identificar, através de sua frequência de coaparição, os casos que realmente constituíam colocações. No entanto, a aplicação dessa metodologia demonstrou-se insuficiente, uma vez que não dispomos de um índice estatístico que permita delimitar o valor a partir do qual a coocorrência dos elementos torna-se relevante a ponto de conformar uma colocação. Essa deficiência fazia com que muitas combinações livres fossem também consideradas colocações.

Em decorrência desse fato, procuramos refinar os dados obtidos a partir da análise estatística através de uma análise qualitativa, que consistiu no reconhecimento das relações semânticas existentes entre substantivos e adjetivos. Para isso, foi necessária a proposição de uma taxonomia dos adjetivos, que permitiu identificar que alguns tipos específicos desses vocábulos apresentavam uma maior tendência a formar colocações. Essa metodologia, porém, não pôde ser aprofundada devido à inexistência de uma ontologia dos adjetivos. Em vista disso, foi necessário recorrer a uma nova abordagem, pois um grande grupo de combinações permanecia sem uma classificação adequada.

A partir das combinações formadas por substantivos e pelos adjetivos que, de acordo com nosso estudo qualitativo, apresentavam uma maior tendência a formar colocações, realizamos a última etapa de nossa pesquisa. Tal procedimento consistiu em uma análise contrastiva dessas combinações previamente selecionadas entre as línguas portuguesa e espanhola. Para a realização desse processo, recorremos aos pressupostos da Análise Contrastiva, Análise de Erros e Interlíngua, os quais nos possibilitaram a elaboração de um

modelo de análise que permitiu o reconhecimento das combinações divergentes entre os idiomas estudados.

A realização consecutiva dessas três etapas de nossa investigação foi necessária para que pudéssemos refinar, de forma gradativa, os dados obtidos. Como resultado desse processo, obtivemos, a partir de nosso *corpus* textual, uma lista de colocações entre as línguas portuguesa e espanhola. Somente de posse dessas informações, foi possível elaborar algumas propostas para o tratamento lexicográfico das colocações, que levassem em conta o usuário e a função da obra.

Através da realização de nossas análises quantitativa, qualitativa e contrastiva, foi possível perceber a multiplicidade de aspectos que caracterizam o fenômeno léxico das colocações. Esses aspectos dizem respeito (a) à frequência de coaparição de seus elementos (ainda que sem um valor estabelecido), (b) às diversas relações semânticas existentes entre seus membros constituintes (seja de intensificação, de qualificação, de modo, etc.), assim como (c) aos diferentes tipos de divergências observadas em suas estruturas através de um estudo interlingüístico. Em virtude desse seu caráter multifacetado, as colocações foram estudadas sob perspectivas bastante diferentes, as quais oferecem uma abordagem apenas parcial desse fenômeno léxico. A ausência, então, de uma metodologia conclusiva, que possibilite um reconhecimento preciso dessas combinações, dificulta seu aprendizado e um tratamento homogêneo nas obras lexicográficas. Essas constatações vêm ao encontro das questões que nos motivaram à realização deste trabalho. Tais questões eram, em princípio, as seguintes:

- a) as colocações, enquanto combinações léxicas situadas entre as estruturas livres e as fixas, possuem um conjunto de características específicas que permita identificá-las e opô-las claramente às extremidades constituídas, de um lado, pelo “discurso livre” e, de outro, pelo “discurso repetido”?
- b) considerando que essas combinações representam tanta dificuldade para os estudantes de uma língua estrangeira, como incluí-las em obras lexicográficas a fim de facilitar seu aprendizado?

Esses questionamentos, que se constituíram como ponto de partida para nossa investigação, são respondidos, ainda que de forma não totalmente conclusiva, pelas evidências que apresentamos a seguir.

Para que possamos oferecer uma resposta ao nosso primeiro questionamento, consideramos pertinente relembrar, sinteticamente, as seguintes constatações:

- as colocações correspondem a um fato de norma, de modo que, mesmo entre duas línguas com uma alta convergência genética e tipológica no nível do sistema (como nos casos do português e do espanhol), a norma acaba determinando certas seleções léxicas particulares. Dessa forma, a “assimetria léxica” é dupla: primeiro, pelo anisomorfismo que caracteriza as línguas e, em seguida, pela seleção (no nível da norma) produzida no interior de toda realização sistêmica, ou seja, no interior de uma língua histórica;
- essas combinações mantêm uma relação de frequência de coaparição (embora não objetivamente determinada a partir de um índice estatístico) entre seus elementos;
- essas combinações mantêm determinadas relações semânticas específicas entre seus membros constituintes.

A partir dessas constatações, percebemos que o grupo formado pelas colocações apresenta uma série de características bastante diferenciadas entre si, razão pela qual essas combinações não conformam um conjunto homogêneo. Cabe lembrar que, somente para as combinações de substantivos e adjetivos (única estrutura morfossintática por nós analisada), encontramos colocações com diferentes frequências de coaparição, com diferentes relações semânticas e com diferentes tipos de equivalência entre as línguas portuguesa e espanhola. Em vista disso, a identificação dessas combinações léxicas não pode ser realizada através da aplicação de um número determinado de traços definitórios. Somente a partir de uma análise detalhada dos candidatos a colocações, considerando diferentes perspectivas de investigação, é possível identificar essas combinações em toda sua complexidade estrutural.

Com relação ao segundo questionamento, a respeito do tratamento lexicográfico das colocações, acreditamos ter sido fundamental a delimitação de um perfil de usuário e da função do dicionário, visto que tais fatores são responsáveis por orientar toda a elaboração da obra. A partir do reconhecimento das necessidades do público-alvo e da função específica que um dicionário que pretende incluir colocações deve desempenhar, a saber, auxiliar na produção de textos, entendemos ir ao encontro das dificuldades enfrentadas pelos estudantes na aprendizagem dessas combinações léxicas.

Ao lado de nossos questionamentos iniciais, apresentamos, na introdução deste trabalho, as seguintes hipóteses de pesquisa:

- a) somente através da aplicação de metodologias consecutivas é possível obter resultados mais precisos no reconhecimento das colocações;

- b) uma proposta de tratamento lexicográfico dessas combinações em dicionários que auxiliem o ensino/aprendizagem de uma língua estrangeira deve fundamentar-se nas particularidades lingüísticas tanto da língua materna do estudante quanto da língua que ele está aprendendo.

As análises realizadas ao longo de nossa dissertação confirmaram a primeira hipótese. A aplicação de cada uma das diferentes metodologias utilizadas em nosso trabalho, quais sejam, análise quantitativa, qualitativa e contrastiva, possibilitou um resultado mais apurado das combinações léxicas que configuravam colocações. A realização consecutiva desses procedimentos permitiu uma melhor filtragem dos candidatos a colocações, de modo que, ao final, foi possível oferecer um estudo contrastivo exaustivo para essas combinações.

Através das relações de contraste entre as línguas portuguesa e espanhola, obtivemos subsídios que permitiram a confirmação de nossa segunda hipótese. A utilização dos pressupostos da Análise Contrastiva, Análise de Erros e Interlíngua mostrou-se fundamental para a identificação das combinações divergentes entre os idiomas e a proposição de estratégias que permitissem diminuir o número de hipóteses na interlíngua do estudante, ampliando, assim, sua competência comunicativa desde os primeiros níveis de aprendizado. Somente através da consideração das particularidades lingüísticas dos dois idiomas, foi possível estabelecer critérios coerentes que justificassem a inclusão das colocações em obras lexicográficas, pois apenas as combinações divergentes parecem constituir uma dificuldade para os estudantes de uma língua estrangeira e, portanto, devem constar nos dicionários.

Em virtude do acima exposto, consideramos ter alcançado os objetivos apresentados no início de nosso trabalho, pois, embora não tenha sido possível a elaboração de uma definição para as colocações a partir de seus traços constitutivos, através do emprego de diferentes metodologias de análise, obtivemos uma compreensão das principais propriedades que caracterizam tais combinações, o que proporcionou uma visão mais ampla desse fenômeno léxico. Além disso, a utilização de uma teoria metalexigráfica subjacente às nossas propostas de apresentação dessas combinações nos dicionários proporcionou, embora não um tratamento definitivo para as colocações, ao menos estratégias simples e coerentes que facilitassem as consultas por parte do usuário.

Apesar, no entanto, da realização de nossos objetivos, é importante destacar que cada aspecto tratado em nossa dissertação na tentativa de identificar as colocações apresentou uma série de limitações devido à inexistência de estudos anteriores mais definitivos. Não existe, por exemplo, um índice estatístico discriminativo que permita determinar a frequência

de coaparição dos elementos necessária para considerá-los colocações. Também não identificamos uma ontologia dos adjetivos que possibilitasse uma classificação mais refinada e teoricamente fundamentada dessa classe de palavras. Além disso, apesar dos avançados estudos da Lingüística Aplicada no ensino de línguas estrangeiras, não tivemos acesso a um valor indicativo da quantidade de vocabulário que um estudante deve possuir para ser classificado nos diferentes estágios de aprendizagem.

Essas deficiências, que podem servir de motivação para a realização de estudos inovadores no âmbito lingüístico, impossibilitaram um posicionamento mais assertivo de nosso trabalho, visto que muitas das decisões que precisamos tomar para resolver problemas de ordem prática não puderam ser suficientemente fundamentadas em uma teoria. Como exemplos, temos a seleção das combinações freqüentes a serem analisadas e os tipos de colocações a serem incluídas em cada uma das obras lexicográficas propostas (para estudantes de nível A2 e B2). No entanto, apesar de nossos resultados serem apenas provisórios e passíveis de um refinamento, acreditamos ter proporcionado um avanço nas reflexões sobre o tema das colocações, especialmente pelo fato de termos chamado a atenção para uma série de questões que ainda necessitam de um maior detalhamento.

Finalmente, consideramos importante questionar se o fato de as colocações não poderem ser identificadas objetivamente através de diferentes métodos de análise deve-se essencialmente às deficiências apresentadas para cada metodologia ou a uma complexidade estrutural inerente ao objeto de estudo. Todas as teorias e suas aplicações metodológicas, tratadas neste trabalho, fundamentam-se no princípio estruturante das línguas; porém, existem aspectos em uma língua natural, como poderiam ser as colocações, que parecem obedecer a um princípio não-estruturante, comum a todas as línguas. Isso ocorre porque a ausência de um conhecimento sobre o comportamento de tais estruturas não permite sua representação em um modelo suficientemente descritivo. Tal fato não significa que a organização lingüística possa ser caótica, mas demonstra que uma língua natural, por vezes, apresenta aspectos muito mais complexos que não permitem sua redução a um modelo formal.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABBAGNANO, N. *Dicionário de filosofia*. São Paulo: Martins Fontes, 1998.
- ADAM, J-M. *Les textes: types et prototypes*. Paris: Nathan, 2001.
- AGUILAR-AMAT, A. C.; TUELLS, A. Colocaciones en un corpus: detección y aplicaciones. In: MARTÍN VIDE, C. (Ed.). CONGRESO DE LENGUAJES NATURALES Y LENGUAJES FORMALES, 10., 1994, Barcelona. *Anales...* Barcelona: PPU, 1994. p. 327-334.
- AISENSTADT, E. Restricted collocations in English lexicology and lexicography. *ITL Review of Applied Linguistics*. Louvain, v. 53, n. 1, p. 53-61, 1981.
- ALARCOS LLORACH, E. *Gramática Estructural*. Madrid: Gredos, 1972.
- _____. *Gramática de la lengua española*. Madrid: Espasa, 1999.
- ALCINA F. J.; BLECUA, J. M. *Gramática Española*. Barcelona: Ariel, 1987.
- ALLERTON, D. J. Three (or four) levels of word cooccurrence restriction. *Lingua*. Amsterdam, v. 63, p. 17-40, 1984.
- ALONSO RAMOS, M. Aproximación a un nuevo modelo lexicográfico: El *Dictionnaire explicatif et combinatoire du français contemporain*. *Recherches lexico-sémantiques* de Igor Mel'čuk. *Verba*. Santiago de Compostela, v. 16, 421-450, 1989.
- _____. Hacia una definición del concepto de colocación: de J.R. Firth a I.A. Mel'čuk. *Revista de Lexicografía*. A Coruña, v. 1, p. 9-28, 1994-95.
- _____. Coocurrencia léxica y descripción lexicográfica del verbo dar: hacia un tratamiento de los verbos soportes. *Zeitschrift für romanische Philologie*. Tübingen, v. 113, n.3, p. 380-417, 1997.
- _____. Construction d'une base de données des collocations bilingue français-espagnol. *Langages*. Paris, n. 143, p. 5-27, 2001.
- _____. Un vacío en la enseñanza del léxico del español como lengua extranjera: las colocaciones léxicas. In: EURALEX INTERNATIONAL CONGRESS, 10., 2002, Copenhagen. *Proceedings...* Copenhagen: Center for Sprogteknologi, 2002a. p. 551-561.
- _____. Elaboración del *Diccionario de colocaciones del español* y sus aplicaciones. In: SYMPOSIUM INTERNACIONAL DE LEXICOGRAFIA, 1., 2002, Barcelona. *Actes...* Barcelona: PPU, 2002b. p. 149-162.
- _____. La nature des collocatifs: leur statut en tant qu'unités lexicales. In: GROSSMAN, F; TUTIN, A. (Eds.). *Les Collocations: analyse et traitement*. Amsterdam: De Werelt, 2003. p. 45-60.

ALVAREZ, M. L. O. A transferência, a interferência e a interlíngua no ensino de línguas próximas. In: CONGRESO BRASILEÑO DE HISPANISTAS, 2., 2002, São Paulo. *Anales...* São Paulo, 2002. Disponível em: http://www.proceedings.scielo.br/scielo.php?pid=MSC0000000012002000100039&script=sci_arttext. Acesso em: 17 jan. 2008.

AuE. FERREIRA, A. B. H. Dicionário Aurélio Eletrônico. Século XXI. Versão 3.0. Rio de Janeiro; Nova Fronteira, 1999. 1 CD-ROM.

BAHNS, J. Lexical collocations: a contrastive view. *ELT Journal*. London, v. 47, n. 1, p. 56-63, 1993.

BALLY, C. *Traité de stylistique française*. Paris: Klincksieck, 1951.

BASILIO, M. Conversão adjetivo/advérbio em português: um estudo de classes de palavras. *ABRALIN*. Curitiba, n. 11, p. 143-151, 1991.

BELLO, A. *Gramática de la lengua castellana*. Buenos Aires: G.L.E.M, 1942.

BENEDUZI, R. Tratamento lexicográfico das colocações em dicionários bilíngües espanhol-português/português-espanhol. 2007. Trabalho apresentado ao VIII Congresso Brasileiro de Lingüística Aplicada, Brasília, 2007.

_____, R.; BUGUEÑO, F.V., Apresentação de um novo dicionário de falsos amigos espanhol-português. In: Fórum Internacional de Ensino de Línguas Estrangeiras, IV. *Anais...* Pelotas: [s.n.], 2006. 1 CD-ROM.

BENSON, M. Collocations and idioms. In: ILSON, R (Ed.). *Dictionaries, lexicography and language learning*. Oxford: Pergamon, 1985. p. 61-68.

_____. Lexical Combinability. *Papers in Linguistics*. Edmonton, v. 18, n. 1, p. 3-15, 1986.

_____. *The BBI dictionary of English word combinations*. Amsterdam: John Benjamins, 1997.

BERBER SARDINHA, T. Que tipo de *corpus* é a web? *Revista da ANPOLL*. São Paulo, n. 15, p. 191-220, jul./dez. 2003.

_____. *Lingüística de Corpus*. São Paulo: Manole, 2004.

BEVILACQUA, C. R. *A fraseologia jurídico-ambiental*. 1996. Dissertação (Mestrado em Letras)-Universidade Federal do Rio Grande do Sul, Porto Alegre, 1996.

_____. *Unidades fraseológicas especializadas eventivas: descripción y reglas de formación en el ámbito de la energía solar*. 2004. Tese (Doutorado em Lingüística Aplicada)-Universidad Pompeu Fabra, Barcelona, 2004.

BIBER, D.; CONRAD, S.; REPPEN, R. *Corpus Linguistics: investigating language structure and use*. Cambridge: Cambridge University Press, 2000.

BICK, E. Gramática constritiva na análise automática da sintaxe portuguesa. In: BERBER SARDINHA, T. (Org.). *A língua portuguesa no computador*. Campinas, SP: Mercado de Letras; São Paulo: Fapesp, 2005. p. 91-112.

BONINI, A. Os gêneros do jornal: o que aponta a literatura da área de comunicação no Brasil? *Linguagem em (Dis)curso*. Tubarão, v. 4, n. 1, p. 205-231, 2003.

BORGES NETO, J. Adjetivos sincategoremáticos: um problema para a semântica. *Cadernos de estudos lingüísticos*. Campinas, v. 22, p. 109-126, jan./jun. 1992.

BOSQUE, I. Sobre el concepto de 'colocación' y sus límites. *Lingüística Española Actual*. Madrid, v. XXIII, n. 1, p. 9-40, 2001.

BUGUEÑO, F. V. Falsos amigos, falsos cognatos, heterossemânticos: uma simples escolha de designações? *Organon*. Porto Alegre, v. 16, n. 32/33, p. 183-192, 2002.

_____. Problemas medioestruturais em um dicionário de falsos amigos. In: COLÓQUIO NACIONAL LETRAS EM DIÁLOGO E EM CONTEXTO: RUMOS E DESAFIOS. 2002, Porto Alegre. *Anais...* Porto Alegre: Programa de Pós-Graduação em Letras da UFRGS, 2003, p. 1-16. 1 CD-ROM.

_____. O que é macroestrutura em um dicionário de língua? In: ISQUERDO, A. N.; ALVES, I. M. (Orgs.). *As Ciências do Léxico*, v. III. Campo Grande: UFMS; São Paulo: Humanitas, 2007, p. 261-272.

_____.; FARIAS, V. S. Informações discretas e discriminantes no artigo léxico. *Cadernos de Tradução*. Florianópolis, v. 18, p. 115-135, 2006.

_____.; _____. Desenho da macroestrutura de um dicionário escolar de língua portuguesa. In: COLÓQUIO INTERNACIONAL DE LEXICOGRAFIA PEDAGÓGICA, 1., 2007, Florianópolis. *Atas...* Florianópolis: [s.n.], 2007. No prelo.

BUSSMANN, H. *Lexikon der Sprachwissenschaft*. Stuttgart: Kröner, 1990.

CABRÉ, M.T. *La terminología*. Teoría, metodología, aplicaciones. Barcelona: Empúries, 1993.

CALDERÓN CAMPOS, M. *Sobre la elaboración de diccionarios monolingües de producción*. Las definiciones, los ejemplos y las colocaciones léxicas. Granada: Universidad de Granada, 1994.

CALVI, M. V. Lessicografia bilingue italo-spagnola e didattica del lessico. In: FERRARIO, E.; PULCINI, V. (Eds.). *La lessicografia bilingue tra presente e avvenire*. Vercelli: Mercurio, 2002. p. 123-137.

CARVALHO, O. L. S. *Lexicografia bilíngüe português/alemão: Teoria e aplicação à categoria das proposições*. Brasília: Thesaurus, 2001.

CASTILLO CARBALLO, M. A. El término 'colocación' en lingüística actual. *Lingüística Española Actual*. Madrid, v. XX, n. 1, p. 41-54, 1998.

_____. Colocaciones léxicas y variación lingüística. *Lingüística Española Actual*. Madrid, v. XXIII, n. 1, p. 133-143, 2001.

CCALED. Collins COBUILD Advanced Learner's English Dictionary. Glasgow: HarperCollins, 2003. 1 CD-ROM.

CENTRO DI STUDI FILOSOFICI DI GALLARATE. Enciclopedia filosofica. Venezia: Istituto per la Collaborazione Culturale, 1957.

ČERMAK, F. Source materials for dictionaries. In: STERKENBURG, P. (Ed.). *A Practical Guide to Lexicography*. Amsterdam: John Benjamins, 2003, p. 18-25.

CHAPARRO, M. C. *Sotaques d'aquém e d'além mar*. Percursos e gêneros do jornalismo português e brasileiro. Santarém: Jortejo, 2000.

CHURCH, K. W.; HANKS, P. Word Association Norms, Mutual Information, and Lexicography. *Computational Linguistics*. Arlington, v. 16, n. 1, p. 22-29, mar. 1990.

COP, M. Collocations in the Bilingual Dictionary. In: HAUSMANN, F. J.; REICHMANN, O.; WIEGAND, H. E.; ZGUSTA, L. (Orgs.). *Wörterbücher. Dictionaries. Dictionnaires*. Ein Internationales Handbuch zur Lexikographie. v.3. Berlin: Walther de Gruyter, 1991. p. 2775-2778.

CORPAS PASTOR, G. Tratamiento de las colocaciones del tipo A+S/S+A en diccionarios bilingües y monolingües (español-inglés). In: EURALEX CONGRESO INTERNACIONAL, 4., 1992, *Actas...* Barcelona: Bibliograf, 1992. p. 331-340.

_____. *Manual de Fraseología Española*. Madrid: Gredos, 1996.

_____. Apuntes para el estudio de la colocación. *Lingüística Española Actual*. Madrid, v. XXIII, n. 1, p. 41-56, 2001.

COSERIU, E. Sistema, norma y habla. In: _____. *Teoría del Lenguaje y Lingüística General*. Madrid: Gredos, 1973. p. 11-113.

_____. Las solidariedades léxicas. In: _____. *Principios de semántica estructural*. Madrid: Gredos, 1977. p. 143-161.

_____. *Lições de Lingüística Geral*. Rio de Janeiro: Ao Livro Técnico, 1980.

COWIE, A. P. The Treatment of Collocations and Idioms in Learners' Dictionaries. *Applied Linguistics*. Oxford, v. 2, n. 3, p. 223-235, 1981.

_____. Language as words: Lexicography. In: COLLINGE, N.E. (Ed.). *An Encyclopaedia of Language*. London: Routledge, 1991. p. 671-700.

_____. Phraseology. In: ASHER, R.E. (Ed.). *The Encyclopedia of Language and Linguistics*. v. 6, Oxford: Pergamon, 1994. p. 3168-3171.

CRYSTAL, D. *The Cambridge encyclopedia of language*. Cambridge: Cambridge University, 1997.

DBE. DÍAZ FOUCHES, O. *Diccionario bilingüe escolar portugués-español, espanhol-português*. São Paulo: Special Book Service, 2005.

DBLA. *Dicionário Básico Larousse Ática Espanhol-Português, Português-Espanhol*. São Paulo: Ática, 2001.

DEA. SECO, M. *Diccionario del español actual*. Madrid: Aguilar, 1999. 2v.

DEMONTÉ, V. El adjetivo: clases y usos. La posición del adjetivo en el sintagma nominal. In: BOSQUE, I.; DEMONTÉ, V. (Eds.). *Gramática descriptiva de la lengua española*. Madrid: Espasa, 1999, p. 129-215.

DiBU. MORENO, F.; GONZÁLEZ, N. M. *Diccionario Bilingüe de Uso español-portugués/português-espanhol*. Madrid: Arco Libros, 2003.

DICE. ALONSO RAMOS, M. *Diccionario de colocaciones del español*. A Coruña: [s.n.], 2004. Disponível em: <http://www.dicesp.com/>. Acesso em 18 abr. 2007.

DiCP. *Diccionario Collins Pocket Español Portugués*. Barcelona: Collins, 2003.

DRAEe. REAL ACADEMIA ESPAÑOLA. *Diccionario de la Real Academia Española*. Versión 1.0. Madrid: Espasa Calpe, 2003. 1 CD-ROM.

DUBOIS, J.; DUBOIS, C. *Introduction à la lexicographie: le dictionnaire*. Paris: Larousse, 1971.

DUBOIS, J. et al. *Dictionnaire de linguistique et des sciences du langage*. Paris: Larousse, 1999.

DUBSKY, J. Debilitamiento del valor comunicativo del verbo español. In: WOTJAK, G. (Ed.). *Estudios de fraseología y fraseografía del español actual*. Frankfurt: Vervuert; Madrid: Iberoamericana, 1998. p. 239-244.

DUEe. MOLINER, M. *Diccionario de uso del español*. Versión 2.0. Madrid: Gredos, 2001. 1 CD-ROM.

DUNKER, D. Collecting Collocations. In: EURALEX INTERNATIONAL CONGRESS, 10., 2002, Copenhagen. *Proceedings...* Copenhagen: Center for Sprogteknologi, 2002. p. 521-530.

DURÃO, A. B. *Análisis de errores en la interlengua de brasileños aprendices de español y de españoles aprendices de portugués*. Londrina: UEL, 2004.

ENCYCLOPÉDIE Philosophique Universelle. v. II, tome 1. Paris: Presses Universitaires de France, 1989.

ETTINGER, S. Formación de palabras y fraseología en la lexicografía. In: HAENSCH et al. *La lexicografía. De la lingüística teórica a la lexicografía práctica*. Madrid: Gredos, 1982, p. 233-258.

FARGHAL, M.; OBIEDAT, H. Collocations: A neglected variable in EFL. *IRAL*. Heidelberg, v. XXXIII, p. 315-331, 1995.

FARIAS, V. S. O exemplo como informação discreta e discriminante em dicionários semasiológicos de língua portuguesa. *Alfa*. Araraquara, v. 52, n. 1, 2008. Aceito para publicação.

FERRATER MORA, J. *Diccionario de Filosofía*. Barcelona: Ariel, 1994.

_____. *Dicionário de filosofia*. São Paulo: Martins Fontes, 1996.

FIALHO, V.R. Proximidade entre línguas: algumas considerações sobre a aquisição do espanhol por falantes nativos de português brasileiro. *Espéculo. Revista de estudios literarios*. Madrid, n. 31, 2005. Disponível em: <http://www.ucm.es/info/especulo/numero31/falantes.html>. Acesso em: 25 jan. 2008.

FIRTH, J.R. Modes of meaning. In: _____. *Papers in linguistics 1934-1951*. London: Oxford University, 1957. p. 190-215.

FONTENELLE, T. Collocations et traitement automatique du langage naturel. In: GROSSMAN, F; TUTIN, A. (Eds.). *Les Collocations: analyse et traitement*. Amsterdam: De Werelt, 2003. p. 75-87.

FORNARI, M. K. A problemática dos americanismos: teorias de divisão dialetal, análise de dados e uma proposta para a marcação diatópica do dicionário de falsos amigos espanhol - português. *Revista Letras*. Curitiba, n. 70, p. 233-246, set./dez. 2006a. Disponível em: <http://ojs.c3sl.ufpr.br/ojs2/index.php/letras/article/view/4698/6398> Acesso em: 20 nov. 2007.

_____. Concepção e desenho das Partes Introdutórias do Dicionário de falsos amigos Espanhol-Português. Trabalho apresentado ao XVIII Salão de Iniciação Científica, Porto Alegre, 2006b.

FOULQUIÉ, P. *Diccionario del lenguaje filosófico*. Barcelona: Labor, 1967.

GABRIELATOS, C. *Collocations: Pedagogical implications, and their treatment in pedagogical materials*. University of Cambridge, 1994. Disponível em: <http://www.gabrielatos.com/Collocation.pdf>. Acesso em: 18 fev. 2008.

GALISSON, R.; COSTE, D. *Dictionnaire de didactique des langues*. Paris: Hachette, 1976.

GARCÍA-PAGE, M. Breves apuntes sobre el adverbio en *-mente*. *Verba*. Santiago de Compostela, v. 20, p. 311-340, 1993.

_____. El adverbio colocacional. *Lingüística Española Actual*. Madrid, v. XXIII, n. 1, p. 89-105, 2001.

- GASS, S. M.; SELINKER, L. *Second language acquisition: An introductory course*. Hillsdale: Lawrence Erlbaum, 1994.
- GILI GAYA, S. *Curso superior de sintaxis española*. Barcelona: Bibliograf, 1979.
- GLÜCK, H. *Lexikon Sprache*. Stuttgart: Metzler, 2000.
- HAENSCH, G et al. *La lexicografía*. De la lingüística teórica a la lexicografía práctica. Madrid: Gredos, 1982.
- HAENSCH, G. Aspectos prácticos de la elaboración de diccionarios. In: HAENSCH, G et al. *La lexicografía*. De la lingüística teórica a la lexicografía práctica. Madrid: Gredos, 1982.
- HALLIDAY, M. A. K. Categories of the Theory of Grammar. *Word*. New York, v. 17, n. 3, p. 241-292, 1961.
- _____. Lexis as a linguistic level. In: BAZELL, C. E. et al. (Eds.). *In Memory of John Firth*. London: Longman, 1966. p. 18-162.
- HANAY, M. Types of bilingual dictionaries. In: STERKENBURG, P. (Ed.). *A Practical Guide to Lexicography*. Amsterdam: John Benjamins, 2003, p. 145-153..
- HANKS, P. Primeiro debate. In: COLOQUIO GALEGO DE FRASEOLOXÍA, 1., 1997, Santiago de Compostela. *Actas...* Santiago de Compostela: Centro Ramón Piñeiro para a Investigación en Humanidades, 1998. p. 101-114.
- HARTMANN, R. R. K. *Teaching and researching lexicography*. London: Longman, 2001.
- _____.; JAMES, G. *Dictionary of Lexicography*. London: Routledge, 1998.
- HAUSMANN, F. J. Un dictionnaire des collocations est-il possible? *Travaux de Linguistique et de Litterature*. Strasbourg, v. 17, n. 1, p. 187-195, 1979.
- _____. Kollokationen im deutschen Wörterbuch: Ein Beitrag zur Theorie des lexikographischen Beispiels. In: BERGENHOLTZ, H.; MUGDAN, J. (Eds.). *Lexikographie und Grammatik*. Tübingen: Max Niemeyer, 1985. p. 118-129.
- _____. Le dictionnaire de collocations. In: HAUSMANN, F.J.; REICHMANN, O.; WIEGAND, H.E.; ZGUSTA, L. (Orgs.). *Wörterbücher. Dictionaries. Dictionnaires*. Ein Internationales Handbuch zur Lexikographie. v.1. Berlin: Walther de Gruyter, 1989. p. 1010-1019.
- _____. Von der Unmöglichkeit der kontrastiven Lexikologie. In: KROMANN, H-P.; KJAER, A. L. (Orgs.). *Von der Allgegenwart der Lexikologie*. Kontrastive Lexikologie als Vorstufe zur zweisprachigen Lexikographie. Tübingen: Max Niemeyer, 1995. p. 19-23.
- _____. O dicionario de colocacións. Criterios de organización. In: COLOQUIO GALEGO DE FRASEOLOXÍA, 1., 1997, Santiago de Compostela. *Actas...* Santiago de Compostela: Centro Ramón Piñeiro para a Investigación en Humanidades, 1998a. p. 63-82.

- _____. Primeiro debate. In: COLOQUIO GALEGO DE FRASEOLOXÍA, 1., 1997, Santiago de Compostela. *Actas...* Santiago de Compostela: Centro Ramón Piñeiro para a Investigación en Humanidades, 1998b. p. 101-114.
- _____. La lexicographie bilingue en Europe: peut-on l'améliorer? In: FERRARIO, E.; PULCINI, V. (Eds.). *La lessicografia bilingue tra presente e avvenire*. Vercelli: Mercurio, 2002. p. 11-31.
- _____; WIEGAND, H. E. Theory of Monolingual Lexicography I: Components and Structures of Dictionaries. In: HAUSMANN, F.J.; REICHMANN, O.; WIEGAND, H.E.; ZGUSTA, L. (Orgs.). *Wörterbücher. Dictionaries. Dictionnaires. Ein Internationales Handbuch zur Lexikographie*. v.1. Berlin: Walther de Gruyter, 1989. p. 328-360.
- HEID, U. Deux approche lexicographiques et leur application dans un outil informatisé. *Terminologie et Traduction*. Luxembourg, n. 2, v. 3, p. 523-548, 1992.
- _____. Using Lexical Functions for the Extraction of Collocations from Dictionaries and Corpora. In: WANNER, L. (Ed.). *Lexical functions in lexicography and natural language processing*. Amsterdam: John Benjamins, 1996. p. 115-146.
- _____. Proposte per la costruzione semi-automatica di un dizionario elettronico delle collocazioni. In: DE MAURO, T.; LO CASCIO, V. (Eds.). *Lessico e Grammatica: Teorie Linguistiche e applicazioni lessicografiche*. Roma: Bulzoni, 1997. p. 47-61.
- HEIDERMAN, W. Prefácio. In: HUMBOLDT, W von. *Linguagem, Literatura e Bildung*. Florianópolis: UFSC, 2006.
- HERBST, T.; KLOTZ, M. *Lexikografie*. Paderborn: F. Schöningh, 2003.
- HouE. *Dicionário eletrônico Houaiss da língua portuguesa*. Versão 1.0. [S.l.]: Objetiva, 2001. 1 CD-ROM.
- HUMBOLDT, W von. *Linguagem, Literatura e Bildung*. Florianópolis: UFSC, 2006.
- HUSSEIN, R. F. Collocations: the missing link in vocabulary acquisition amongst EFL learners. *Papers and Studies in Contrastive Linguistics*. Poznan, v. 26, p. 123-136, 1990.
- IRSULA, J. Colocaciones sustantivo-verbo. In: WOTJAK, G. (Ed.). *Estudios de lexicología y metalexicografía del español actual*. Tübingen: Max Niemeyer, 1992. p. 159-167.
- ISENBERG, H. Cuestiones fundamentales de tipología textual. In: BERNÁRDEZ, H. (Org.). *Lingüística del texto*. Madrid: Arco, 1987. p. 95-129.
- JAMES, C. *Errors in language learning and use: exploring error analysis*. Harlow: Longman, 1998.
- JAROŠOVÁ, A. Problems of semantic subdivisions in bilingual dictionary entries. *International Journal of Lexicography*. Oxford, v. 13, n. 1, p. 12-28, 2000.

- JONES, S; SINCLAIR, J. M. English lexical collocations. A study in computational linguistics. *Cahiers de lexicologie*. Paris, v. 24, p. 15-61, 1974.
- JOTA, Z. S. *Dicionário de lingüística*. Rio de Janeiro: Presença, 1976.
- KAUFFMANN, C. *O corpus do jornal: variação lingüística, gêneros e dimensões da imprensa diária escrita*. 2005. Dissertação (Mestrado em Lingüística Aplicada e Estudos da Linguagem)-Pontifícia Universidade Católica de São Paulo, São Paulo, 2005. Disponível em: http://www.pucsp.br/pos/lael/lael-inf/def_teses.html. Acesso em: 15 set. 2006.
- KENNEDY, G. *An introduction to corpus linguistics*. London: Longman, 1999.
- KOIKE, K. *Colocaciones léxicas en el español actual: análisis formal y léxico-semántico*. 2000. Tese (Doutorado em Filologia)-Universidad Nacional de Educación a Distancia, Madrid, 2000.
- KROMANN, H-P.; RIIBER, T.; ROSBACH, P. Principles of Bilingual Lexicography. In: HAUSMANN, F. J.; REICHMANN, O.; WIEGAND, H. E.; ZGUSTA, L. (Orgs.). *Wörterbücher. Dictionaries. Dictionnaires. Ein Internationales Handbuch zur Lexikographie*. v.3. Berlin: Walther de Gruyter, 1991. p. 2711-2728.
- LADO, R. *Introdução à lingüística aplicada*. Petrópolis: Vozes, 1972.
- _____. Procedures in Comparing Two Grammatical Structures. In: ROBINETT, B.W.; SCHACHTER, J. (Eds.). *Second language learning*. Contrastive analysis, error analysis and related aspects. Ann Arbor: University of Michigan, 1983. p. 15-19.
- LANDAU, S. *Dictionaries: the art and the craft of lexicography*. Cambridge: CUP, 2001.
- LANGAGES. Paris: Larousse, v. 30, n. 121, mar. 1996.
- LAUFER, B.; LEVITZKY-AVIAD, T. Examining the Effectiveness of 'Bilingual Dictionary Plus' – A Dictionary for Production in a Foreign Language. *International Journal of Lexicography*. Oxford, v. 19, n. 2, p. 135-155, 2006.
- LIN, T-J. *La adquisición y el uso del artículo por alumnos chinos*. Tese (Doutorado em Lingüística Aplicada)- Universidad de Alcalá, Alcalá de Henares, 2003. Disponível em: <http://www.mec.es/redele/biblioteca2005/lin2.shtml>. Acesso em: 16 jan. 2008
- LIPKA, L. *English lexicology*. Tübingen: Narr, 2002.
- LIPSKI, J. M. *El español de América*. Madrid: Cátedra, 1994.
- LO CASCIO, V. Semantica lessicale e i criteri di collocazione nei dizionari bilingui a stampa ed elettronici. In: DE MAURO, T.; LO CASCIO, V. (Eds.). *Lessico e Grammatica: Teorie Linguistiche e applicazioni lessicografiche*. Roma: Bulzoni, 1997. p. 63-88.
- LYONS, J. *Semantics*. v. 2. Cambridge: Cambridge University, 1977.

MARELLO, C. Les différents types de dictionnaires bilingues. In: BÉJOINT, H.; THOIRON, P. (Org.). *Les dictionnaires bilingues*. Lovain-la-nouve: Duculot, 1996. p. 31-52.

MARTÍNEZ ALBERTOS, J. L. *Curso general de redacción periodística*: periodismo en prensa, radio, televisión y cine, lenguaje, estilos y géneros periodísticos. Barcelona: Mitre, 1983.

MAURER STROH, P. *Towards a Bilingual Adjective-Noun Collocation Dictionary of English and German*. 2004. Tese (Doutorado em Filologia)-Universität Klagenfurt, Klagenfurt, 2004. Disponível em: <http://ubdocs.uni-klu.ac.at/open/hssvoll/AC04445625.pdf>. Acesso em: 14 out. 2007.

MARCO Común Europeo de Referencia para las Lenguas: aprendizaje, enseñanza, evaluación. Madrid: Anaya, 2002. Disponível em: http://cvc.cervantes.es/obref/marco/cvc_mer.pdf. Acesso em: 12 jan. 2008.

MEL'ČUK, I. Théorie de langage, théorie de traduction. *Meta*. Montreal, v. XXIII, n. 4, p. 271-302, 1978.

_____. *Dictionnaire Explicatif et combinatoire du français contemporain*. Researches lexicosémantiques. v. I. Montréal: Université de Montréal, 1984.

_____. *Dictionnaire Explicatif et combinatoire du français contemporain*. Researches lexicosémantiques. v. II. Montréal: Université de Montréal, 1988.

_____. Phrasemes in Language and Phraseology in Linguistics. In: EVERAERT, M. et al. (Eds.). *Idioms: Structural and Psychological Perspectives*. New Jersey: Lawrence Erlbaum, 1995. p. 167-232.

_____. Les collocations: définition, rôle et utilité. In: GROSSMAN, F; TUTIN, A. (Eds.). *Les Collocations: analyse et traitement*. Amsterdam: De Werelt, 2003. p. 23-31.

MELO, J.M. *Gêneros jornalísticos na Folha de São Paulo*. São Paulo: FTD, 1992.

MENUZZI, S. Sobre a sincategorematicidade adjetival. *Letras de Hoje*. Porto Alegre, v. 27, n. 3, p. 85-99, set. 1992.

MiFF. FLAVIAN, E.; FERNÁNDEZ, G. E. *Minidicionário Espanhol-Português/Português-Espanhol*. São Paulo: Ática, 2001.

MiPD. PEREIRA, H. B. C. *Michaelis: pequeno dicionário espanhol-português, português-espanhol*. São Paulo: Melhoramentos, 1996.

MITCHELL, T. F. Linguistic "goings-on": collocations and other lexical matters arising on the syntagmatic record. In: _____. *Principles of Firthian Linguistics*. London: Longman, 1975. p. 99-136.

MORENO DE ALBA, J. G. *El español en América*. 2. ed. México: Fondo de cultura económica, 1988.

- NESSELHAUF, N. *Collocations in a Learner Corpus*. Amsterdam, Philadelphia: John Benjamins, 2005.
- NEUBERT, A. Fact and Fiction of the Bilingual Dictionary. In: EURALEX CONGRESO INTERNACIONAL, 4., 1992, *Actas...* Barcelona: Bibliograf, 1992. p. 29-42.
- NEVES, M. H. M. A delimitação das unidades lexicais: o caso das construções com verbo-suporte. *Palavra*. Rio de Janeiro, v. 1, n. 5, p. 98-114, 1999.
- OTANI, H. Investigating Intercollocations - Towards an archaeology of texts. *International Journal of Lexicography*. Oxford, v. 18, n. 3, p. 1-24, 2005.
- PAZOS BRETaña, J. M. *Detección automatizada de fraseologismos*. 2005. Tese (Doutorado)- Universidad de Granada, Granada, 2005. Disponível em: <http://hera.ugr.es/tesisugr/15476935.pdf>. Acesso em: 27 fev. 2008.
- PENADÉS MARTÍNEZ, I. ¿Colocaciones o locuciones verbales? *Lingüística Española Actual*. Madrid, v. XXIII, n. 1, p. 57-88, 2001.
- PÉREZ HERNÁNDEZ, C. *Explotación de los corpora textuales informatizados para la creación de bases de datos terminológicas basadas em conocimiento*. Madrid: CSIC/EliEs, 2002. Disponível em <http://elies.rediris.es/elies18>. Acesso em: 10 set. 2006.
- PÉREZ LAGOS, M. F. A propósito del *Diccionario Ideológico de la lengua española VOX*. Su didactismo. In: AYALA CASTRO, M. C. (Coord.). *Diccionarios y Enseñanza*. Alcalá: Universidad de Alcalá, 2001. p. 179-192.
- PETERS, H. Lexicografía traduccional en la colocación de los adjetivos. In: WOTJAK, G. (Ed.). *Estudios de lexicología y metalexicografía del español actual*. Tübingen: Max Niemeyer, 1992. p. 90-98.
- PINKER, S. *O instinto da linguagem: como a mente cria a linguagem*. São Paulo: Martins Fontes, 2004.
- PIT CORDER, S. *Error analysis and interlanguage*. Oxford: Oxford University, 1981.
- POLGUÈRE, A. *Typologie des entités lexicales d'une base de données explicative et combinatoire*. 2005. Trabalho apresentado à Journée de l'ATALA: Interface lexicogrammaire et lexiques syntaxiques et sémantiques, Paris, 2005. Disponível em: <http://www.olst.umontreal.ca/pdf/ATALAAPol-2005.pdf>. Acesso em: 12 dez. 2007.
- PÖLL, B. *Portugiesische Kollokationen im Wörterbuch: ein Beitrag zur Lexikographie und Metalexikographie*. Bonn: Romanistischer Verlag, 1996.
- PORZIG, W. *El mundo maravilloso del lenguaje*. Madrid: Gredos, 1970.
- REAL ACADEMIA ESPAÑOLA. *Gramática de la lengua española*. Madrid: Espasa-Calpe, 1931.
- _____. *Esbozo de una nueva gramática de la lengua española*. Madrid: Espasa-Calpe, 1973.

RÔA, M. J. Uma proposta de análise contrastiva. *Letras de hoje*. Porto Alegre, v. 54, n. 4, p. 114-129, dez. 1983.

ROBERTS, R.P. Le traitement des collocations et des expressions idiomatiques dans les dictionnaires bilingues. In: BÉJOINT, H.; THOIRON, P. (Org.). *Les dictionnaires bilingues*. Lovain-la-nouve: Duculot, 1996. p. 181-197.

RUIZ GURILLO, L. Cómo integrar la fraseología en los diccionarios monolingües. In: CORPAS PASTOR, G. (Ed.) *Las lenguas de Europa: estudios de fraseología, fraseografía y traducción*. Granada: Comares, 2000. p. 261-274.

RUNDELL, M. Dictionary use in production. *International Journal of Lexicography*. Oxford, v. 12, n.1, p. 35-53, 1999.

SANTAMARÍA PÉREZ, M.I. Tratamiento de las unidades fraseológicas en la lexicografía bilingüe español-catalán. 2000. Tese (Doutorado em Letras)-Universidad de Alicante, Alicante, 2000. Disponível em: http://descargas.cervantesvirtual.com/servlet/SirveObras/68094941545465830532279/006698_1.pdf. Acesso em: 27 fev. 2008.

SALVADOR, G. Las solidariedades lexemáticas. *Revista de Filología de la Universidad de La Laguna*. Tenerife, n. 8/9, p. 339-365, 1989-90.

SAUSSURE, F. *Curso de Lingüística geral*. São Paulo: Cultrix, 2004.

SECO, M. Problemas formales de la definición lexicográfica. In: _____. *Estudios de lexicografía española*. Madrid: Paraninfo, 1987a. p. 15-34.

_____. El “contorno” en la definición lexicográfica. In: _____. *Estudios de lexicografía española*. Madrid: Paraninfo, 1987b. p. 35-45.

_____. Los pilares de un diccionario moderno. In: _____. *Estudios de lexicografía española*. Madrid: Gredos, 2003. p. 81-90.

SELINKER, L. Interlanguage. In: ROBINETT, B.W.; SCHACHTER, J. (Eds.). *Second language learning*. Contrastive analysis, error analysis and related aspects. Ann Arbor: University of Michigan, 1983, p. 6-14.

SIEPMANN, D. Collocation, Colligation and Encoding Dictionaries. Part I: Lexicological Aspects. *International Journal of Lexicography*. Oxford, v. 18, n. 4, p. 409-443, 2005.

SIMÕES, T.C.S. O uso das preposições locais no processo de aquisição formal da língua alemã como segunda língua. 2006. Dissertação (Mestrado em Letras)-Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2006.

SINCLAIR, J. M. *Corpus, concordance, collocation*. Oxford: Oxford University, 1991.

SINGLETON, D. Lexical partnerships. In: _____. *Language and the lexicon: An introduction*. London: Arnold, 2000. p. 47-62.

- SMADJA, F. Retrieving Collocations from Text: Xtract. *Computational Linguistics*. Arlington, v. 19, n. 1, p. 143-177, 1993.
- SWALES, J.M. *Genre analysis*. English in academic and research settings. Cambridge: Cambridge University, 2000.
- SZENDE, T. Problèmes d'équivalence dans les dictionnaires bilingues. In: BÉJOINT, H.; THOIRON, P. (Org.). *Les dictionnaires bilingues*. Lovain-la-nouve: Duculot, p. 111-126, 1996.
- TAGNIN, S.E.O. *O jeito que a gente diz*. São Paulo: Disal, 2005a.
- _____. Um dicionário de colocações verbais? Para quê? In: BERBER SARDINHA, T. (Org.). *A língua portuguesa no computador*. Campinas, SP: Mercado de Letras; São Paulo: Fapesp, 2005b. p. 197-214.
- TRASK, R.L. *Dicionário de linguagem e lingüística*. São Paulo: Contexto, 2006.
- VAN DIJK, T. A. *La ciencia del texto: un enfoque interdisciplinario*. 4.ed. Barcelona: Paidós, 1996.
- VANDRESEN, P. Lingüística contrastiva e ensino de línguas estrangeiras. In: BOHN, H. I.; VANDRESEN, P. (Orgs.). *Tópicos de lingüística aplicada: o ensino de línguas estrangeiras*. Florianópolis: UFSC, 1988. p. 75-94.
- VARANTOLA, K. Linguistic corpora (database) and the compilation of dictionaries. In: STERKENBURG, P. (Ed.). *A Practical Guide to Lexicography*. Amsterdam: John Benjamins, 2003, p. 228-239.
- VERDÍA, E. Comentarios al Marco común europeo de referencia para las lenguas. *Mosaico*. [S.l.], n. 9, p. 4-7, 2002.
Disponível em: <http://www.mec.es/sgci/be/es/publicaciones/mosaico/mosaico9/mos9a.pdf>.
Acesso em: 05 fev. 2008.
- VERLINDE, S.; SELVA, T.; BINON, J. Les collocations dans les dictionnaires d'apprentissage: repérage, présentation et accès. In: GROSSMAN, F; TUTIN, A. (Eds.). *Les Collocations: analyse et traitement*. Amsterdam: De Werelt, 2003. p. 105-115.
- YAVAS, M. A análise de erros e suas limitações. *Letras de hoje*. Porto Alegre, v. 42, n. 4, p. 112-124, 1980.
- WANNER, L. *Lexical functions in lexicography and natural language processing*. Amsterdam: John Benjamins, 1996.
- WARDHAUGH, R. The Contrastive Analysis Hypothesis. In: ROBINETT, B.W.; SCHACHTER, J. (Eds.). *Second language learning*. Contrastive analysis, error analysis and related aspects. Ann Arbor: University of Michigan, 1983. p. 6-14.
- WEINRICH, H. *Lenguaje en textos*. Madrid: Gredos, 1981.

_____. *Grammaire textuelle du français*. Paris: Didier, 1989.

WELKER, H.A. *Dicionários: Uma pequena introdução à Lexicografia*. Brasília: Thesaurus, 2004.

WHORF, B.L. *Lenguaje, pensamiento y realidad*. Barcelona: Barral, 1971.

WILLIAMS, G. Les collocations et l'école contextualiste britannique. In: GROSSMAN, F; TUTIN, A. (Eds.). *Les Collocations: analyse et traitement*. Amsterdam: De Werelt, 2003. p. 33-44.

WOUDEN, T. van der. Prologomena to a Multilingual Description of Collocations. In: EURALEX INTERNATIONAL CONGRESS, 5., 1992, *Proceedings...* Part II. Tampere: University of Tampere, 1992. p. 449-456.

ZAGREBELSKY, M.T.P. Una guida alle guide all'uso dei dizionari bilingui italiano-inglese/inglese-italiano. In: FERRARIO, E.; PULCINI, V. (Eds.). *La lessicografia bilingue tra presente e avvenire*. Vercelli: Mercurio, 2002. p. 139-147.

ZGUSTA, L. *Manual of Lexicography*. The Hague: Mouton, 1971.

ZIMMER, Rudolf. *Äquivalenzen zwischen Französisch und Deutsch: Theorie - Korpus - Indizes: Ein Kontextwörterbuch*. Tübingen: Max Niemeyer, 1990.

ZULUAGA, A. Primeiro debate. In: COLOQUIO GALEGO DE FRASEOLOXÍA, 1., 1997, Santiago de Compostela. *Actas...* Santiago de Compostela: Centro Ramón Piñeiro para a Investigación en Humanidades, 1998. p. 101-114.

_____. Los "enlaces frecuentes" de María Moliner. Observaciones sobre las llamadas colocaciones. *Philologie im Netz*. [S.l.], v. 22, p. 56-74, 2002. Disponível em: <http://web.fu-berlin.de/phn/phn22/p22t3.htm>. Acesso em: 16 out. 2007.

APÊNDICE A

DOCUMENTAÇÃO DO CORPUS NO MICROSOFT ACCESS

Có	Código	Nome do jornal	País de origem	Data de coleta	Tokens	Types	Nº de textos	Observações	Referências Bibliográficas
1	CLAD1	El Clarín	Argentina	1/10/2006	67468	11732	132	Inclui um caderno sobre viaç	http://www.clarin.com/diario/2006/10/01/index_e
2	DIA01	El Diario	Bolívia	2/10/2006	30617	6503	75	Inclui a Revista "Ciencia & C	http://www.eldiario.net/noticias/m061002/titulare
3	MER01	El Mercurio	Chile	3/10/2006	82796	12781	199	Inclui a Revista "Ya" (para a	http://diario.elmercurio.com/2006/10/03/_portad
4	TIE01	El Tiempo	Colômbia	4/10/2006	25225	5870	65	Inclui uma seção sobre saú	http://www.eltiempo.com/tiempoimpreso/edicion
5	UNIE01	El Universo	Equador	5/10/2006	37299	7958	125		http://www.eluniverso.com/2006/10/05/0001/1/pc
6	MUN01	El Mundo	Espanha	6/10/2006	89602	13266	181	Inclui os suplementos "Su vi	http://www.elmundo.es/papel/2006/10/06/
7	UNIM01	El Universal	México	7/10/2006	67523	11396	145	Inclui um caderno sobre aut	http://www.eluniversal.com.mx/hemeroteca/edic
8	PRLO1	Prensa Libre	Guatemala	8/10/2006	31266	6983	69	Inclui a "Revista D" (especia	http://www.prensallibre.com/pl/2006/octubre/08/i
9	COM01	El Comercio	Peru	9/10/2006	39571	8178	104	Inclui o suplemento "Dia 1"	http://www.elcomercio.com.pe/edicionimpre
10	UNIV01	El Universal	Venezuela	10/10/2006	43558	8995	143		http://www.eluniversal.com/2006/10/10/indice.st
11	CLAD2	El Clarín	Argentina	11/10/2006	45477	9374	113	Inclui um caderno de inform	http://www.clarin.com/diario/2006/10/11/index_e
12	DIA02	El Diario	Bolívia	12/10/2006	41647	8053	100	Inclui a Revista "Femenina"	http://www.eldiario.net/noticias/m061012/titulare
13	MER02	El Mercurio	Chile	13/10/2006	68230	11812	202	Inclui a Revista "Wiken" (ca	http://diario.elmercurio.com/2006/10/13/_portad
14	TIE02	El Tiempo	Colômbia	14/10/2006	42074	8367	83	Inclui uma seção de literat	http://www.eltiempo.com/tiempoimpreso/edicion
15	UNIE02	El Universo	Equador	15/10/2006	61306	11288	147	Dia das eleições no Equado	http://www.eluniverso.com/2006/10/15/0001/1/pc
16	MUN02	El Mundo	Espanha	16/10/2006	85383	13675	162	Inclui o suplemento "Aula" (http://www.elmundo.es/papel/2006/10/16/
17	UNIM02	El Universal	México	17/10/2006	74575	11604	161		http://www.eluniversal.com.mx/hemeroteca/edic
18	PRLO2	Prensa Libre	Guatemala	18/10/2006	18982	4817	66		http://www.prensallibre.com/pl/2006/octubre/18/i
19	COM02	El Comercio	Peru	19/10/2006	40328	8512	110		http://www.elcomercio.com.pe/edicionimpre
20	UNIV02	El Universal	Venezuela	20/10/2006	50278	9770	161		http://www.eluniversal.com/2006/10/20/indice.st
21	CLAD3	El Clarín	Argentina	21/10/2006	69582	12022	156	Inclui a "Revista N" (sobre a	http://www.clarin.com/diario/2006/10/21/index_e
22	DIA03	El Diario	Bolívia	22/10/2006	41745	7908	103	Inclui a Revista "Tú Guía" (s	http://www.eldiario.net/noticias/m061022/titulare
23	MER03	El Mercurio	Chile	23/10/2006	72549	12161	208	Inclui a "Revista del Campo"	http://diario.elmercurio.com/2006/10/23/_portad
24	TIE03	El Tiempo	Colômbia	24/10/2006	24221	5552	66		http://www.eltiempo.com/tiempoimpreso/edicion
25	UNIE03	El Universo	Equador	25/10/2006	32425	7178	140		http://www.eluniverso.com/2006/10/25/0001/1/pc
26	MUN03	El Mundo	Espanha	26/10/2006	78693	12878	164	Inclui o suplemento "Aula" (http://www.elmundo.es/papel/2006/10/26/
27	UNIM03	El Universal	México	27/10/2006	79906	12154	164		http://www.eluniversal.com.mx/hemeroteca/edic
28	PRLO3	Prensa Libre	Guatemala	28/10/2006	22990	5345	78		http://www.prensallibre.com/pl/2006/octubre/28/i
29	COM03	El Comercio	Peru	29/10/2006	44762	9317	92		http://www.elcohttp://www.elcomercio.com
30	UNIV03	El Universal	Venezuela	30/10/2006	50807	9959	168		http://www.eluniversal.com/2006/10/30/indice.st

Registro: 1 de 60

Modo folha de dados

NUM

Microsoft Access - [Organização do corpus : Tabela]

Arquivo Editar Exibir Inserir Formatar Registros Ferramentas Janela Ajuda

7 de 60

Modo folha de dados

Referências Bibliográficas

Obs: Digite uma pergunta

Có	Código	Nome do jornal	País de origem	Data de coleta	Tokens	Types	Nº de textos	Observações	Referências Bibliográficas
31	CLAD4	El Clarín	Argentina	31/10/2006	45022	9040	98	Inclui o caderno "Mujer" (sol	http://www.clarin.com/diario/2006/10/31/index_e
32	DIA04	El Diario	Bolívia	1/11/2006	40492	7920	103	Inclui a Revista "Cine Mundic	http://www.eldiario.net/noticias/m061101/titular_e
33	MER04	El Mercurio	Chile	2/11/2006	56900	10411	174		http://diario.elmercurio.com/2006/11/02/_portad
34	TIE04	El Tiempo	Colômbia	3/11/2006	24771	5524	62		http://www.eltiempo.com/tiempoimpreso/edicion
35	UNIE04	El Universo	Equador	4/11/2006	31209	6849	107	Inclui uma seção sobre agro	http://www.eluniverso.com/2006/11/04/0001/1/pc
36	MUND4	El Mundo	Espanha	5/11/2006	110225	16333	150	Inclui os suplementos: "Maç	http://www.elmundo.es/papel/2006/11/05/
37	UNIM04	El Universal	México	6/11/2006	94439	13752	186		http://www.eluniversal.com.mx/hemeroteca/edic
38	PRLD4	Prensa Libre	Guatemala	7/11/2006	22941	5577	68		http://www.prensalibre.com/pl/2006/noviembre/0
39	COM04	El Comercio	Peru	8/11/2006	32929	7531	106		http://www.elcomercio.pe.com.pe/edicionimpre
40	UNIV04	El Universal	Venezuela	9/11/2006	50827	10009	176		http://www.eluniversal.com/2006/11/09/indice.st
41	CLA05	El Clarín	Argentina	10/11/2006	54395	10390	132	Inclui o caderno "Si" (sobre	http://www.clarin.com/diario/2006/11/10/index_e
42	DIA05	El Diario	Bolívia	11/11/2006	33685	6609	92	Inclui a Revista "Autonitro" (http://www.eldiario.net/noticias/m061111/titular_e
43	MER05	El Mercurio	Chile	12/11/2006	129965	17939	234	Inclui a "Revista de Libros" (http://diario.elmercurio.com/2006/11/12/_portad
44	TIE05	El Tiempo	Colômbia	13/11/2006	34075	6586	97		http://www.eltiempo.com/tiempoimpreso/edicion
45	UNIE05	El Universo	Equador	14/11/2006	34747	7470	125		http://www.eluniverso.com/2006/11/14/0001/1/pc
46	MUND5	El Mundo	Espanha	15/11/2006	77803	12881	173	Inclui o suplemento "Aula" (http://www.elmundo.es/papel/2006/11/15/
47	UNIM05	El Universal	México	16/11/2006	89166	12603	186		http://www.eluniversal.com.mx/hemeroteca/edic
48	PRLD5	Prensa Libre	Guatemala	17/11/2006	23192	5751	73		http://www.prensalibre.com/pl/2006/noviembre/1
49	COM05	El Comercio	Peru	18/11/2006	45808	9082	122		http://www.elcomercio.pe.com.pe/edicionimpre
50	UNIV05	El Universal	Venezuela	19/11/2006	58108	11084	169		http://www.eluniversal.com/2006/11/19/indice.st
51	CLA06	El Clarín	Argentina	20/11/2006	38860	8191	88	Não inclui suplementos	http://www.clarin.com/diario/2006/11/20/index_e
52	DIA06	El Diario	Bolívia	21/11/2006	37767	7459	104	Inclui a Revista "Nuevos Hor	http://www.eldiario.net/noticias/m061121/titular_e
53	MER06	El Mercurio	Chile	22/11/2006	55123	9717	143		http://diario.elmercurio.com/2006/11/22/_portad
54	TIE06	El Tiempo	Colômbia	23/11/2006	29982	7029	73	Inclui uma seção sobre viaje	http://www.eltiempo.com/tiempoimpreso/edicion
55	UNIE06	El Universo	Equador	24/11/2006	37877	7790	133		http://www.eluniverso.com/2006/11/24/0001/1/pc
56	MUND6	El Mundo	Espanha	25/11/2006	79703	13414	160	Inclui o suplemento "Salud"	http://www.elmundo.es/papel/2006/11/25/
57	UNIM06	El Universal	México	26/11/2006	97362	14210	184		http://www.eluniversal.com.mx/hemeroteca/edic
58	PRLD6	Prensa Libre	Guatemala	27/11/2006	24149	5858	63		http://www.prensalibre.com/pl/2006/noviembre/2
59	COM06	El Comercio	Peru	28/11/2006	37939	8063	103		http://www.elcomercio.pe.com.pe/edicionimpre
60	UNIV06	El Universal	Venezuela	29/11/2006	49405	9676	151		http://www.eluniversal.com/2006/11/29/indice.st

Registro: 7 de 60

ANEXO A

LISTA DAS FUNÇÕES LÉXICAS DE MEL'ČUK (1988)

LISTE DES FONCTIONS LEXICALES STANDARD SIMPLES

Les fonctions lexicales (FL) dans la liste ci-dessous sont présentées dans le même ordre d'apparition que dans les articles de dictionnaire. Cette liste reprend celle donnée dans le DEC 1:7-13, avec deux changements: d'une part, nous avons rajouté les FL omises par mégarde dans l'ancienne liste; d'autre part, nous avons éliminé quelques exemples et explications en ne laissant que le minimum nécessaire pour faciliter la compréhension au lecteur.

À la fin de cette liste, le lecteur trouvera un index des FL.

1. **Syn**, **Syn_▷**, **Syn_c**, **Syn_∩**: synonyme exact et quasi-synonymes (▷ signifie «plus large», c signifie «plus étroit», tandis que ∩ dénote une intersection des sens).

Syn(*espoir*) = *espérance*

Syn_▷(*vénération*) = *respect I*

Syn_c(*respect I*) = *vénération*

Syn_∩(*secondar*) = *assister II.1*

N.B.: Les indices **c**, **▷** et **∩** sont utilisés également avec d'autres FL, telles que **Conv**, **Anti**, **S₀**, **S₁**, **S₂**, etc. Leur seul rôle est de signaler au lecteur l'existence d'un décalage sémantique entre la valeur escomptée de la FL et le sens du lexème indiqué qui fait partie de cette valeur. Pour voir le véritable sémantisme d'un synonyme (antonyme, ...) plus large, plus étroit ou à intersection, il faut consulter l'article de dictionnaire de ce dernier. Ainsi, si dans l'entrée de **ASSISTER II.1**, on trouve un synonyme à intersection **SECONDER**, il faut voir l'entrée de ce dernier pour trouver quelle composante est commune aux deux et quelles sont leurs différences. Cette démarche correspond à la stratégie générale du DEC, qui est la suivante:

Si, dans l'entrée de la lexie L₁, on trouve la référence à la lexie L₂ qui, elle, a sa propre entrée dans le DEC, L₂ n'est munie d'aucune information lexicographique, car cette dernière est fournie dans l'entrée de L₂.

2. **Conv_{ijk}**: conversif, c'est-à-dire un lexème qui dénote une relation converse de la relation exprimée par le mot clé de cette FL; les indices montrent l'ordre des actants syntaxiques associés au conversif relativement à l'ordre de base (l'ordre

des actants associés au mot clé du conversif), qui est toujours «123».

Conv₂₁(*être derrière*) = *être devant*

[*Je me trouvais derrière une dame* = *Devant moi se trouvait une dame*]

Conv₂₁(*plus*) = *moins*

Conv_{3214∩}(*vendre I.1*) = *acheter I*

3. **Anti**, **Anti_▷**, **Anti_c**, **Anti_∩**: antonyme exact et quasi-antonymes.

Anti(*respect I*) = *irrespect*

Anti_▷(*désespoir I*) = *espoir*

Anti_c(*mépris I*) = *respect I*

Anti_∩(*aider I.a,b*) = *gêner*

Anti se combine très facilement avec d'autres FL pour former des FL complexes.

4. **Contr**: terme contrastif.

Contr(*cœur I.4a*) = *tête I.4*

Contr(*acier*) = *velours*

5. **Epit**: épithète courante sémantiquement vide.

Epit(*océan*) = *immense*

6. **Gener**: mot générique dont la combinaison avec un dérivé syntaxique (voir le n° 8 plus bas) du mot clé est synonyme du mot clé ou qui peut conclure les séries du type X₁, X₂, X₃, ... et les autres **Gener(X)**.

Gener(*colère I*) = *sentiment [de colère]*

Gener(*république*) = *état [républicain]*

7. **Figur**: métaphore codifiée par la langue dont la combinaison avec le mot clé est un synonyme (plus étroit) du mot clé.

Figur(*fumée*) = *rideau [de fumée]*

Figur(*haine I*) = *feu [de la haine]*

8. **S₀**, **V₀**, **A₀**, **Adv₀**: dérivés syntaxiques — nom, verbe, adjectif, adverbe.

S₀(*acheter I*) = *achat I.a*

V₀(*promesse I*) = *promettre I*

A₀(*école I.1a*) = *scolaire*

Adv₀(*honnête*) = *honnêtement*

9. **S₁**, **S₂**, **S₃**, ...: nom typique pour le premier, deuxième, troisième, ... actant du mot clé.

S₁(*crime*) = *criminel*

S₂(*crime*) = *victime [du crime]*

LISTE DES FONCTIONS LEXICALES

- S₁**(*acheter I*) = *acheteur I*
S_{2c}(*acheter I*) = *marchandise*
S₃(*acheter I*) = *vendeur I*
S₄(*acheter I*) = *prix I*
10. **S_{instr}**, **S_{loc}**, **S_{mod}**, **S_{res}**: nom typique pour le circonstant d'instrument, de lieu, de moyen, de mode et de résultat.
- S_{instr}**(*peindre*) = *pinceau*
S_{loc}(*lutter*) = *arène [de la lutte]*
S_{mod}(*peindre*) = *peinture*
S_{mod}(*écrire*) = *écriture*
S_{res}(*copier*) = *copie*
11. **Sing**: «un quantum régulier <une portion régulière> de ...».
- Sing**(*flotte*) = *bateau*
Sing(*riz*) = *grain [de riz]*
12. **Mult**: «ensemble régulier de ...».
- Mult**(*bateau*) = *flotte*
Mult(*chien*) = *meute*
Mult(*abeille*) = *essaim*
13. **Cap**: «chef de ...».
- Cap**(*avion*) = *capitaine*
Cap(*théâtre*) = *directeur*
14. **Equip**: «équipe de ...».
- Equip**(*avion*) = *équipage*
Equip(*théâtre*) = *troupe*
15. **Germ**: «le germe de ...».
- Germ**(*colère I*) = *ferment, levain [de la colère]*
16. **Centr**: «le centre de ...», «le milieu de ...».
- Loc_{in}Centr**(*hiver, nuit, mêlée*) = *au cœur de [l'hiver, la nuit, la mêlée]*
 [Pour **Loc_{in}** voir n° 29.]
17. **Culm**: «culmination de ...».
- Culm**(*joie I*) = *comble II [de la joie]*
Culm(*colère I*) = *paroxysme [de la colère]*
18. **A₁**, **A₂**, **A₃**, ...: modificateur typique pour le premier, deuxième, troisième, ... actant du mot clé.
- A₁**(*mépris I*) = *plein, rempli¹ II.2a [de mépris]*
A₂(*mépris I*) = *couvert [de mépris]*
A₁(*chercher*) = *en quête de*
A₂(*diriger*) = *sous la direction de*
19. **Able₁**, **Able₂**, ...: «tel qu'il peut ...», «tel qu'on peut le ...», etc.
- Able₁**(*peur*) = *peureux*
Able₂(*peur*) = *effrayant*
20. **Qual_i**: «tel qu'il entraîne **Able_i** avec une haute probabilité».
- Qual₁**(*tromper*) = *malhonnête*
Qual₂(*tromper*) = *naïf*
21. **Magn**: «très», «intense/intensément», «à un degré élevé».
- Magn**(*mémoire I.I*) = *prodigieuse, excellente, étonnante, d'éléphant*
Magn(*bruit*) = *infernal, de tonnerre*
Magn(*remercier*) = *vivement, chaleureusement, de tout cœur*
AntiMagn(*vent I.I*) = *doux, faible, léger*
22. **Plus**, **Minus**: «plus», «moins»; ces FL ne s'emploient que dans une combinaison avec d'autres FL.
- IncepPredMinus**(*joie I*) = *diminuer, faiblir*
 [pour **Incep** et **Pred** voir n°s 36 et 31]
IncepPredMinus(*vent I.I*) = *se calmer, mollir*
23. **Ver**: «tel qu'il doit être», «correct».
- Ver**(*peur*) = *justifiée*
Ver(*appareil*) = *exact, précis*
AntiVer(*compliment I.I*) = *immérité*
AntiBon₁(*promesse I*) + **AntiVer**(*promesse I*) = *à la légère, en l'air*
24. **Bon**: «bon» — expression qu'on emploie comme une louange standard codifiée par la langue.
- Bon**(*conseil*) = *précieux*
Bon(*compliment I.I*) = *recherché, bien tourné*
25. **Pejor**: «pire» [= **MinusBon**].
- CausPredPejor**(*joie I*) = *altérer, gâcher [la joie]*
 N.B.: **IncepPredPejor** = **Degrad**.
26. **Pos₁**, **Pos₂**, ...: «évaluation positive» — expression qu'on emploie comme expression standard de l'évaluation positive du premier, deuxième, ... actant du mot clé.
- Pos₂**(*opinion I*) = *favorable, avantageuse*
Pos₂(*critique*) = *favorable, élogieuse*
27. **Adv₁**, **Adv₂**, ...: adverbe typique pour caractériser le comportement du premier, deuxième, ... actant du mot clé, autrement dit, l'adverbe qui signifie «étant le premier, deuxième, ... actant de».
- Adv₁**(*joie I*) = *avec [joie]*
Adv₂(*joie I*) = *à [la joie]*
28. **Instr**: préposition régissant le mot clé et signifiant «au moyen de».
- Instr**(*bras I.Ia*) = *à [bras]*
Instr(*mains I.a*) = *de, dans, avec [la <les> main<s>]*
29. **Loc_{in}**, **Loc_{ab}**, **Loc_{ad}**: préposition régissant le mot clé et signifiant une localisation ou un déplacement spatial, temporel ou abstrait, autrement dit, la préposition qui signifie «se trouvant dans» (**Loc_{in}**), «se déplaçant de» (**Loc_{ab}**), «se déplaçant à» (**Loc_{ad}**).
- Loc_{in/ad}**(*gare*) = *à [la gare]*
Loc_{in}(*personnel*) = *au sein du [personnel]*
Loc_{ab}(*gare*) = *de [la gare]*
Loc_{ab}(*1970*) = *depuis [1970]*
Loc_{ad}(*front I.a*) = *en plein [front]*

LISTE DES FONCTIONS LEXICALES

30. **Propt**: préposition signifiant «à cause de».
Propt(*peur*) = *de, par* [*peur*]
Propt(*respect* *I*) = *par* [*respect*]
31. **Pred**: «prédicat — verbalisateur des FL adjectivales [= «être»]; n'apparaît qu'en combinaison avec d'autres FL. [Voir les exemples donnés pour la FL **Minus** au n° 22.]
32. **Oper₀**, **Oper₁**, **Oper₂**, ...: verbe sémantiquement vide qui prend C₀ comme son SG et, dans le cas où la situation C₀ a des actants, le nom du premier, deuxième, ... actant de la situation C₀ comme son complément d'objet (CO) principal.
Oper₀(*vent* *I.1*) = *faire* [*du vent*]
Oper₁(*attention*) = *faire*
Oper₂(*attention*) = *attirer*
Oper₁(*conseil*) = *donner*
Oper₂(*conseil*) = *recevoir*
Oper₁(*liste*) = *dresser*
Oper₂(*liste*) = *figurer* [*sur une liste*]
33. **Func₀**, **Func₁**, **Func₂**, ...: verbe sémantiquement vide qui prend C₀ comme son SG et, dans le cas où la situation C₀ a des actants, le nom du premier, deuxième, ... actant de C₀ comme son CO principal.
Func₀(*silence*) = *régner*
Func₁(*aide* *1b,2*) = *venir, provenir* [*de qqn*]
Func₂(*liste*) = *contenir, comprendre* [*qch*]
N.B.: Dans certains cas, **Func₀** peut être réalisé en français par une construction impersonnelle avec un SG «scindé en deux»:
La neige tombe [= **Func₀**] *≈ Il tombe* [= **Func₀**] *de la neige.*
Une telle possibilité est explicitement indiquée dans l'entrée du mot clé (sous **Func₀**).
34. **Labor_{ij}**: verbe sémantiquement vide qui prend le nom de l'actant *i* comme son SG, le nom de l'actant *j* comme son CO principal, et C₀ comme son deuxième CO.
Labor₁₂(*horreur*) = *avoir* [*qch/qqn en horreur*]
Labor₁₂(*soin*) = *entourer* [*qqn de soins*]
35. **Involv**: verbe qui décrit l'action de la situation C₀ sur un objet qui n'est pas un des participants de C₀; il prend C₀ comme son SG et le nom signifiant l'objet qui subit l'action de C₀ comme son CO principal.
Involv(*vent* *I.1*) = *agiter, secouer* [*un arbre*]; *cingler, brûler* [*le visage de Pierre*]; *plier, incliner, courber* [*les roseaux*]
Involv(*odeur*) = *remplir* *II.3* [*la pièce*]
36. **Incep**, **Cont**, **Fin**: «commencer», «continuer», «cesser»; ces FL s'emploient le plus souvent en combinaison avec les FL **Oper_i**, **Func_i** ou **Labor_{ij}**.
IncepOper₁(*forme*) = *prendre*
ContOper₁(*influence*) = *garder*
FinOper₁(*influence*) = *perdre*
- IncepFunc₀**(*vent* *I.1*) = *se lever, s'élever*
IncepFunc₁(*colère* *I*) = *s'emparer de* [*qqn*]
FinFunc₀(*vent* *I.1*) = *se calmer*
37. **Caus**, **Liqu**, **Perm**: «causer» («faire en sorte que quelque chose ait lieu»), «liquider» et «permettre», respectivement; ces FL s'emploient le plus souvent en combinaison avec les FL **Oper_i**, **Func_i** ou **Labor_{ij}**.
CausOper₁(*désespoir* *I*) = *pousser, réduire* [*qqn au désespoir*], *jeter* [*qqn dans le désespoir*], *frapper* [*qqn de désespoir*]
CausFunc₀(*difficulté*) = *créer, poser*
LiquFunc₀(*assemblée*) = *dissoudre*
LiquFunc₂(*attention*) = *détourner* [*l'attention de qqn de qch*]
Perm₁Fact₀(*colère* *1,2*) = *s'abandonner* [*à la colère*]
Caus₂Func₂(*attention*) = *accaparer*
Liqu₁Func₀(*peur*) = *surmonter, vaincre* [*sa (propre) peur*]
Liqu₁Func₀(*colère* *1,2*) = *étouffer* [*sa (propre) colère*]
38. **Real₁**, **Real₂**, ...: verbe ayant le sens «réaliser» qui prend C₀ comme son CO principal et le nom du premier, deuxième, ... actant de C₀ comme son SG.
Real₁(*problème*) = *résoudre*
Real₁(*piège*) = *tendre*
Real₂(*piège*) = *tomber* [*dans un piège*]
Real₂(*ordre*) = *exécuter*
39. **Fact₀**, **Fact₁**, **Fact₂**, ...: verbe ayant le sens «se réaliser» qui prend C₀ comme son SG et le nom du premier, deuxième, ... actant de C₀ (si C₀ a un actant) comme son CO principal.
Fact₀(*souçon*) = *se confirmer*
AntiFact₁(*mémoire* *I.1*) = *trahir* [*qqn*]
Fact₂(*mémoire* *I.1*) = *garder, retenir* [*qch*]
40. **Labreal_{ij}**: verbe ayant le sens «réaliser» qui prend les deux actants de la situation comme son SG et son CO principal, et C₀ comme son deuxième CO.
Labreal₁₂(*scie*) = *couper* [*qch avec une scie*]
Labreal₁₂(*mémoire* *I.1*) = *conserver* [*qch en mémoire*]
Syntactiquement,
Real : **Fact** : **Labreal** = **Oper** : **Func** : **Labor**;
autrement dit, **Real₁** est **Oper₁** + «réaliser», tandis que **Fact₀**, c'est **Func₀** + «se réaliser», etc.
41. **Manif**: «se manifester dans ...»; C₀ est le SG.
Manif(*opinion* *3b*) = *s'exprimer*
MagnManif(*joie* *I*) = *jaillir, éclater*
MagnManif(*désespoir* *I*) = *éclater*
Très souvent, **Manif** apparaît avec **Caus₁**, avec laquelle elle forme une FL complexe:
Caus₁Manif(*excuse*) = *présenter*
Caus₁Manif(*opinion* *I*) = *exprimer, formuler*

LISTE DES FONCTIONS LEXICALES

42. **Prepar**: «préparer (pour l'usage ou le fonctionnement)».

PreparFact₀(fusil) = charger

PreparFact₀(voiture) = mettre au point; faire le plein

43. **Prox**: «être sur le point de ...».

ProxOper₁(désespoir I) = être au bord < à la limite > [du désespoir]

ProxFunc₀(orage I) = (s') approcher

44. **Degrad**: «se dégrader», «devenir pire».

Degrad(cœur I.1a) = faiblir, flancher

Degrad(lait) = tourner

Degrad(vin) = aigrir

45. **Nocer**: «nuire à ...».

Nocer(moustique) = piquer

Nocer(voiture) = faucher, happer

46. **Obstr**: «fonctionner avec difficulté».

Obstr(moteur) = avoir des ratés

47. **Excess**: «fonctionner d'une façon anormalement excessive».

Excess(cœur I.1a) = palpiter, accélérer

Excess(moteur) = s'emballer

48. **Son**: «émettre le son typique».

Son(plancher) = craquer

Son(chien) = aboyer

49. **Imper**: formule exclamative qui exprime l'ordre, la prière, etc. autrement que par une forme impérative du verbe.

Imper(secourir) = Au secours!

Imper(tirer) = Feu!

50. **Perf**: action complétée.

S₁Perf(s'évader) = [un] évadé

S₁Perf(marier) = époux

51. **Result**: ce qui résulte d'un événement.

Result₁Perf(promettre I) = // être obligé [de V_{inf}] ['ce-lui qui a promis se trouve obligé de faire ce qu'il a promis']

52. **Sympt**: «symptôme» – expression verbale signifiant un symptôme physique d'une émotion, d'un état, etc.

Excess^{motor}(dents) – **Sympt₂₁₃**(colère I) = grincer des dents (de colère)

remuer involontairement la tête I.1a... –

Sympt₂₁₃(avoir sommeil) = dodeliner de la tête

Degrad(parole) – **Sympt₂₃**(étonnement) = être muet (d'étonnement)

Excess(cheveux) – **Sympt₁₃**(horreur) = ses cheveux se dressent (d'horreur)

Certaines fonctions lexicales sont munies d'exposants qui en précisent le sémantisme. Ce sont les exposants suivants:

actual : 'actuellement'

usual : 'usuellement'

Ces deux exposants s'emploient surtout avec les FL **S_i**, **Fact_i**, **Excess**, **Degrad**.

quant : 'par rapport à la quantité de ...'

temp : 'par rapport à la durée de ...'

Ces deux exposants s'emploient surtout avec la FL **Magn**.

color : 'par rapport à la couleur'

dim : 'par rapport à la dimension (taille)'

fulg : 'par rapport à l'éclat'

motor : 'par rapport aux mouvements'

stat : 'par rapport à la position verticale'

trem : 'par rapport au tremblement'

t' : 'par rapport à la température'

Ces sept exposants s'emploient surtout avec les FL **Degrad** et **Excess**.

ANEXO B

LISTA DE ETIQUETAS MORFOSSINTÁTICAS DO VISL

WORD CLASS TAGS

N Nouns

PROP Proper nouns (names)

SPEC Specifiers (defined as non-inflecting pronouns, that can't be used as prenominals):

e.g. indefinite pronouns, nominal quantifiers, nominal relatives

DET Determiners (defined as inflecting pronouns, that can be used as prenominals):

e.g. articles, attributive quantifiers

PERS Personal pronouns (defined as person-inflecting pronouns)

ADJ Adjectives (including ordinals, excluding participles which are tagged V PCP)

ADV Adverbs (both 'primary' adverbs and derived adverbs ending in '-mente')

V Verbs (full verbs, auxiliaries)

NUM Numerals (cardinals)

KS Subordinating conjunctions

KC Coordinating conjunctions

IN Interjections

EC Hyphen-separated prefix ("elemento composto", category being phased out)

INFLECTION TAGS

Gender: M (male), F (female), M/F [for: N', PROP', SPEC', DET, PERS, ADJ, V PCP, NUM]

Number: S (singular), P (plural), S/P [for: N, PROP', SPEC', DET, PERS, ADJ, V PCP, V VFIN, INF, NUM]

Case: NOM (nominative), ACC (accusative), DAT (dative), PIV (prepositional), ACC/DAT, NOM/PIV [for: PERS]

Person: 1 (first person), 2 (second person), 3 (third person), 1S, 1P, 2S, 2P, 3S, 3P, 1/3S, 0/1/3S [for: PERS, V VFIN, V INF]

Tense: PR (present tense), IMPF (imperfecto), PS (perfeito simples), MQP (mais-que-perfeito), FUT (futuro), COND (condicional) [for: V VFIN]

Mood: IND (indicative), SUBJ (subjunctive), IMP (imperative) [for: V VFIN]

Finiteness: VFIN (finite verb), INF (infinitive), PCP (participle), GER (gerund) [for: V]

SYNTACTIC TAGS

@SUBJ> @<SUBJ subject

@ACC> @<ACC accusative (direct) object

@DAT> @<DAT dative object (only pronominal)

@PIV> @<PIV prepositional (indirect) object

@ADVS> @<ADVS adverbial object (place, time, duration, quantity)

@ADVO> @<ADVO adverbial object (object-related)

@SC> @<SC subject predicative

@OC> @<OC object predicative

@ADVL> @<ADVL adverbial

@PASS> @<PASS agent of passive

(All above clause arguments [@SUBJ, @ACC, @DAT, @PIV, @ADVS, @ADVO, @SC, @OC, @PASS] and the adverbial complements [@ADVL] attach to the nearest main verb to the left [<] or right [>].)

@ADVL 'free' adverbial phrase (in non-sentence expression)

@NPHR 'free' noun phrase (in non-sentence expression without verbs)

@VOK 'vocative' (e.g. 'free' addressing proper noun in direct speech)

@>N prenominal adjunct (attaches to the nearest NP-head to the right, that is not an adnominal itself)
 @N< postnominal adjunct (attaches to the nearest NP-head to the left, that is not an adnominal itself)
 @N<PRED postnominal (in-group predicative) or predicate in small clause introduced by 'com/sem' (rare, e.g. 'com a mão na bolsa', 'sem o pai ajudando, não conseguiu')
 @APP identifying apposition (always after NP + comma)
 @>A prepositioned adverbial adjunct
 (attaches to the nearest ADJ/PCP/ADV or attributive used N to the right)
 @A< postpositioned adverbial adjunct (rare, e.g. 'caro demais') or dependent/argument of attributive participle (with function tag attached, e.g. @A<ADVL or @A<SC)
 @PRED> 'forward' free predicative (refers to the following @SUBJ, even when this is incorporated in the VP)
 @<PRED 'backward' free predicative (refers to the nearest NP-head to the left, or to the nearest @SUBJ to the left)
 @P< argument of preposition
 @S< sentence anaphor ('não venceu o que muito o contrariou')
 @FAUX finite auxiliary (cp. @#ICL-AUX<)
 @FMV finite main verb
 @IAUX infinite auxiliary (cp. @#ICL-AUX<)
 @IMV infinite main verb
 @PRT-AUX< verb chain particle (preposition or "que" after auxiliary)
 @CO coordinating conjunction
 @SUB subordinating conjunction
 @KOMP< argument of comparative (e.g. "do que" referring to 'melhor')
 @COM direct comparator without preceding comparative
 @PRD role predicator (e.g. "work as", "function as")
 @FOC> @<FOC focus marker ("gosta é de peixe.")
 @TOP topic constituent ("Esse negócio, não gosto dele.")
 @#FS- finite subclause (combines with clausal role and intraclausal word tag, e.g. @#FS-<ACC @SUB for "não acredito que seja verdade")
 @#ICL- infinite subclause (combines with clausal role and intraclausal word tag, e.g. @#ICL-SUBJ> @IMV in "consertar um relógio não é fácil")
 @#ICL-AUX< argument verb in verb chain, refers to preceding auxiliary (the verb chain sequence @FAUX - @#ICL-AUX< is used, where both verbs have the same subject, @FMV - @#ICL-<ACC is used where the subjects are different)
 @#AS- averbal (i.e. verb-less) subclause (combines with clausal role and intraclausal word tag, e.g. @#AS-<ADVL @ADVL> in "ajudou onde possível")
 @AS< argument of complementiser in averbal subclause

SECONDARY TAGS

Subclass tags

<artd> definite article (DET)
 <arti> indefinite article (DET)
 <quant> quantifier pronoun (DET: <quant1>, <quant2>, <quant3>, SPEC: <quant0>) or intensifier adverb
 <dem> demonstrative pronoun (DET: <dem>, SPEC: <dem0>)
 <poss> possessive pronoun (DET)
 <refl> reflexive personal pronoun ("se" PERS ACC, "si" PERS PIV)

<si> reflexive use of 3. person possessive
 <reci> reciprocal use of reflexive pronoun (= "um ao outro")
 <coll> collective reflexive ("reunir-se", "associar-se")
 <diff> differentiator (DET) (e.g. "e outros temas", "a mesma diferença")
 <ident> identator (DET) (e.g. "o próprio usuário", "a si mesmo")
 <rel> relative pronoun (DET, SPEC)
 <interr> interrogative pronoun (DET, SPEC)
 <post-det> typically located as post-determiner (DET @N<)
 <post-attr> typically post-positioned adjective (ADJ @N<)
 <ante-attr> typically pre-positioned adjective (ADJ @>N)
 <adv> can be used adverbially (ADJ @ADVL)
 <ks> relative adverb used like a subordinating conjunction
 <kc> conjunctive adverb (pois, entretanto)
 <det> determiner usage/inflection of adverb ("ela estava toda nua.")
 <foc> focus marker adverb (also forms of "ser")
 <prp> relative adverb used like a preposition
 <KOMP> <igual> "equalling" comparative (ADJ, ADV) (e.g. "tanto", "tão")
 <KOMP> <corr> correlating comparative (ADJ, ADV) (e.g. "mais velho", "melhor")
 <komp> <igual> "equalling" particle referring to comparative (e.g. "como", "quanto")
 <komp> <corr> "correlating" particle referring to comparative (e.g. "do=que")
 <SUP> superlative
 <setop> operational adverb (eg. "não", "nunca", "já", "mais" in "não mais")
 <dei> discourse deictics (e.g. "aqui", "ontem")
 <card > cardinal (NUM)
 <NUM-ord> ordinal (ADJ)
 <NUM-fract> fraction-numeral (N)
 <cif> cipher (<card> NUM, <NUM-ord> ADJ)
 <sam-> first part of morphologically fused word pair ("de" in "dele")
 <-sam> last part of morphologically fused word pair ("ele" in "dele")
 <*> 1. letter capitalized
 <*1> left quote attached
 <*2> right quote attached
 <hyfen> hyphenated word
 <ABBR> abbreviation
 <prop> noun, adjective etc. used as name (upper case initial in mid-sentence)
 <n> adjective or participle used as a "noun", typically as head of a nominal phrase
 <fmc> finite main clause heading verb
 <co-acc>, <co-advl>, <co-app>, <co-dat>, <co-fmc>, <co-ger>, <co-inf>, <co-oc>, <co-
 pcv>, <co-postad>, <co-postnom>, <co-pred>, <co-prenom>, <co-prparg>, <co-sc>, <co-
 subj>, <co-vfin> co-ordinator tags indicating what is co-ordinated: @ACC, @ADVL, @APP,
 @DAT, main clauses, GER, INF, @OC, PCP-@IMV, @A<, @N<, @PRED, @>N, @P<,
 @SC, @SUBJ, VFIN (ordered list matching the <co-...> tags)

Valency tags

<vt> monotransitive verb with accusative object
 <vi> (<ve>) intransitive verb (ergative verb)
 <vtd> ditransitive verb with accusative and dative objects
 <PRP^vp> monotransitive verb with prepositional object (headed by PRP)
 <PRP^vtp> ditransitive verb with accusative and prepositional objects
 <vK> copula verb with subject predicative

<vtK> copula verb with object predicative
 <va> transitive verb with adverbial argument:
 <va+LOC>, <va+DIR>, <vta+LOC>, <vta+DIR>, <vt+QUANT> transitive verb with NP as
 quantitative adverbial object (e.g. "pesar")
 <vt+TID> transitive verb with NP as temporal adverbial object (e.g. "durar")
 <vU> "impersonal" verbs (normally in the 3S-person, e.g. "chove")
 <x> auxiliary verb with infinitive (tagged @(F)AUX - @#ICL-AUX<)
 <x+PCP> auxiliary verb with participle (tagged @(F)AUX - @#ICL-AUX<)
 <x+GER> auxiliary verb with gerund (tagged @(F)AUX - @#ICL-AUX<)
 <PRP^xp> auxiliary verb with (prepositional) auxiliary particle and infinitive
 (tagged as @(F)AUX - @PRT-AUX< - @#ICL-AUX<)
 <xt> "auxiliary" verb with infinitive clause subject in the accusative case, and ACI-
 constructions,
 (both tagged as @(F)MV - @SUBJ> - @#ICL-ACC)
 <PRP^xtp> "auxiliary" verb with accusative object and prepositional object containing an
 infinitive clause with its (unexpressed) subject being identical to the preceding accusative
 object ,
 (tagged as @(F)MV - @<ACC - @<PIV - @#ICL-P<)
 <vr> reflexive verbs (also <vrp>, <vaux-r>, <vaux-rp>)
 <vq> "cognitive" verb governing a 'que'-sentence
 <qv> "impersonal" verb with 'que'-subclause as subject predicative ("parece que", "consta
 que")
 <+interr> "discourse" verb or nominal governing an interrogative subclause
 <+n> noun governing a name (PROP) (e.g. "o senhor X")
 <+num> noun governing a number (e.g. "cap. 7", "no dia 5 de dezembro")
 <num+> "unit" noun (e.g. "20 metros")
 <+INF> governs infinitive (N, ADJ)
 <+PRP> governs prepositional phrase headed by PRP, e.g. <+sobre>
 <PRP+> (typically) argument of preposition PRP
 <+que> <+PRP+que> nominal governing a 'que'-subclause (N, ADJ)

ANEXO C
EXEMPLO DE TEXTO ETIQUETADO PELO VISL

pALABRA=DE=EXPERTO [PALABRA=DE=EXPERTO] <*> <heurtag> <G> PROP
M/F S/P @SUBJ>
gestión [gestión] <*> <2> N F S @N<
responsable [responsable] <2> ADJ M/F S @N<
<María Daverio. Investigadora de la Univ. Nac. de la Patagonia>
 [María Daverio. Investigadora de la Univ. Nac. de la Patagonia
>] <heur> N M S @SUBJ>
la [la] <*> <art> <dem> DET F S @>N
práctica [práctica] N F S @SUBJ>
turística [turístico] ADJ F S @N<
se [se] <refl> PERS M/F 3S/P ACC/DAT @ACC>
ha [haber] <x+PCP> <xque> <de^vp> <vdb> <vr> <vt> V PR 3S IND VFIN
@FAUX
manifestado [manifestar] <vdb> <vr> <vt> <vq> V PCP M S @IMV @&ICL-
AUX<
durante [durante] PRP @<ADVL
las [la] <art> <dem> DET F P @>N
últimas [último] ADJ F P @>N
décadas [década] N F P @P<
como [como] PRP @N<
un [un] <quant2> <arti> DET M S @>N
factor [factor] N M S @P<
de [de] <+INF> PRP @N<
transformación [transformación] N F S @P<
de [de] <+INF> <saml> PRP @N<
el [el] <art> <dem> <sam2> DET M S @>N
espacio [espacio] N M S @P<
. .
el [el] <*> <art> <dem> DET M S @>N
contacto [contacto] <entre> <event> <state> <'contact'> N M S
@SUBJ>
con [con] <+INF> PRP @N<
bienes [bien] N M P @P<
de [de] <+INF> PRP @N<
la [la] <art> <dem> DET F S @>N
naturaleza [naturaleza] N F S @P<
y [y] KC @CO
la [la] <art> <dem> DET F S @>N
cultura [cultura] N F S @SUBJ>
motivan [motivar] <vr> <vt> V PR 3P IND VFIN @FMV
la [la] <art> <dem> DET F S @>N
elección [elección] <act> <occ> <+a> N F S @<ACC
de [de] <+INF> <saml> PRP @N<
el [el] <art> <dem> <sam2> DET M S @>N
destino [destino] N M S @P<
turístico [turístico] ADJ M S @N<
. .
la [la] <*> <art> <dem> DET F S @>N
ampliación [ampliación] N F S @SUBJ>
de [de] <+INF> PRP @N<
la [la] <art> <dem> DET F S @P<
<<
mirada [mirada] <2> N F S @P<
turística [turístico] ADJ F S @N<
>>
demanda [demandar] <vdb> <vr> <vt> V PR 3S IND VFIN @FMV
nuevas [nuevo] ADJ F P @>N
experiencias [experiencia] N F P @<ACC

y [y] KC @CO
 destinos [destino] N M P @<ACC
 singulares [singular] ADJ M/F P @N<
 ,
 especialmente [especialmente] <deadj> <normal> ADV @>A
 atractivos [atractivo] ADJ M P @>N
 y/o [y/o] <heur> N M S @<PRED
 poco [mucho] <aquant> ADV @>A
 conocidos [conocer] <vdb> <vr> <vq> <vt> V PCP M P @N<
 [conocido] ADJ M P @N<
 .
 el [el] <*> <art> <dem> DET M S @>N
 análisis [análisis] N M P "análisis" N M S @NPHR
 de [de] <+INF> PRP @N<
 los [el] <art> <dem> DET M P @>N
 modelos [modelo] <Hprof> N M P @P<
 turísticos [turístico] ADJ M P @N<
 en [en] <+top> <+vej> <+GER> PRP @N<
 el [el] <art> <dem> DET M S @>N
 contexto [contexto] N M S @P<
 de [de] <+INF> PRP @N<
 los [el] <art> <dem> DET M P @>N
 nuevos [nuevo] ADJ M P @>N
 paradigmas [paradigma] N M P @P<
 ,
 en=particular [en=particular] <normal> ADV @ADVL>
 el [el] <art> <dem> DET M S @>N
 caso [caso] N M S @SUBJ>
 de [de] <+INF> PRP @N<
 tierra=del=Fuego [Tierra=del=Fuego] <*> <top> <+> PROP F S @P<
 ,
 antártida=e=Islas=del=Atlántico=Sur
 [Antártida=e=Islas=del=Atlántico=Sur] <*> <heurtag> <G> PROP M/F S/P
 @APP
 ,
 muestra [mostrar] <vdb> <vr> <vq> <vt> V PR 3S IND VFIN @FMV
 la [la] <art> <dem> DET F S @>N
 necesidad [necesidad] N F S @<ACC
 de [de] <+INF> PRP @N<
 una [un] <quant2> <arti> DET F S @>N
 gestión [gestión] <2> N F S @P<
 responsable [responsable] <2> ADJ M/F S @N<
 de [de] <+INF> PRP @N<
 los [el] <art> <dem> DET M P @>N
 flujos [flujo] N M P @P<
 en [en] <+top> <+vej> <+GER> PRP @N<
 estos [este] <dem> DET M P @>N
 espacios [espacio] N M P @P<
 emergentes [emergente] ADJ M/F P @N<
 .
 asimismo [asimismo] <*> <normal> ADV @ADVL>
 ,
 exige [exigir] <vdb> <vr> <vt> <vq> V PR 3S IND VFIN @FMV
 atender [atender] <vr> <vt> V INF @IMV @&ICL-<ACC
 el [el] <art> <dem> DET M S @>N
 impera [imperar] <heurval:> <vt> <vi> V PR 3S IND VFIN @<ACC
 - [hyfen] PU
 tivo [tivo] <DERS -ivo [ATTR Rare]> ADJ M S @<PRED
 de [de] <+INF> PRP @<ADVL
 un [un] <quant2> <arti> DET M S @>N
 ordenamiento [ordenamiento] N M S @P<

territorial [territorial] ADJ M/F S @N<
 que [que] <rel> SPEC M/F S/P @SUBJ> @&FS-N<
 contemple [contemplar] <vr> <vt> V PR 1/3S SUBJ VFIN @FMV
 la [la] <art> <dem> DET F S @>N
 compatibilidad [compatibilidad] N F S @<ACC
 socio [socio] N M S @<ACC
 ecológica [ecológico] ADJ F S @N<
 .
 es [ser] <*> <x+PCP> <vK> <ser> V PR 3S IND VFIN @FMV
 preciso [preciso] ADJ M S @<SC
 distinguir [distinguir] <de^vtp> <vdb> <vr> <pcp-irreg> <vt> V INF
 @IMV @&ICL-<SUBJ
 entre [entre] PRP @<ADVL
 «
 bienes [bien] N M P @P<
 patrimoniales [patrimonial] ADJ M/F P @N<
 »
 y [y] KC @CO
 «
 productos [producto] N M P @<ACC
 turísticos [turístico] ADJ M P @N<
 »
 ;
 si [si] KS @&FS-ADVL> @SUB
 estos [este] <dem> DET M P @SUBJ>
 no [no] ADV @ADVL>
 aseguran [asegurar] <vdb> <vr> <vq> <vt> V PR 3P IND VFIN @FMV
 la [la] <art> <dem> DET F S @>N
 preservación [preservación] N F S @<ACC
 de [de] <+INF> PRP @N<
 aquellos [aquel] <dem> DET M P @P<
 estamos [estar] <va+LOC> <va+QUAL> <x+GER> <vdb> <vr> <vK>
 <vk+ADJ> <vk+N> <vt> V PR/PS 1P IND VFIN @FMV
 frente=a [frente=a] PRP @<ADV
 una [un] <quant2> <arti> DET F S @>N
 inversión [inversión] <2> N F S @P<
 de [de] <+INF> PRP @N<
 la [la] <art> <dem> DET F S @>N
 relación [relación] <+con> <ac> <'relation'> N F S @P<
 entre [entre] PRP @N< @<ADVL
 medios [medio] <num> N M P @P<
 y [y] KC @CO
 fines [fin] N M P @P<