

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

**A AQUISIÇÃO DE MORFEMAS EM INGLÊS COMO L2:
Uma Análise Dos Padrões Evolutivos Através Do BELC
(*Brazilian English Learner Corpus*)**

TESE DE DOUTORADO

ALINE PACHECO

PORTO ALEGRE

2010

**UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE LETRAS – PROGRAMA DE PÓS-GRADUAÇÃO EM LETRAS
ESTUDOS DA LINGUAGEM
ESPECIALIDADE: TEORIA E ANÁLISE LINGUÍSTICA**

**A AQUISIÇÃO DE MORFEMAS EM INGLÊS COMO L2:
Uma Análise Dos Padrões Evolutivos Através Do BELC
(*Brazilian English Learner Corpus*)**

Tese de Doutorado em Teoria e Análise Linguística apresentada como requisito para a obtenção do título de Doutor pelo Programa de Pós-Graduação em Letras pela Universidade Federal do Rio Grande do Sul

ALUNA: ALINE PACHECO

PROFESSOR ORIENTADOR: Dr. SERGIO DE MOURA MENUZZI

PORTO ALEGRE

2010

“If one does not know to which port he is sailing, no wind is favourable to him”.

Seneca

AGRADECIMENTOS

Ao longo destes quase seis anos envolvida diretamente com o doutorado, posso dizer que enfrentei alguns ventos desfavoráveis, mas como sabia o porto onde ancorar, também tive muitos ventos favoráveis a mim.

Quando resolvi recomeçar a minha vida acadêmica e pensar num projeto de doutorado, fui pega de surpresa por uma tempestade horrível com a saída do Sergio da PUC e ventos que pareciam levá-lo para muito longe! Mas sempre depois da tempestade, a calmaria: quando já estava na UFRGS cursando cadeiras como aluna especial, um vento bem bonzinho trouxe o Sergio de volta para cá!

Ao Sergio eu agradeço a pronta disponibilidade em me orientar e em iluminar as minhas idéias com sua indiscutível inteligência e sabedoria.

Os meus mais sinceros agradecimentos a alguns professores da UFRGS: à Prof. Ana Maria Becker Maciel, Prof. Luciene Juliano Simões e Prof. Marília dos Santos Lima que tão bem me receberam como aluna especial, em especial à Prof. Sabrina Pereira de Abreu, que gentilmente me acolheu como sua orientanda por um período curto de tempo.

Um vento favorável soprou algumas idéias da colega Simone Sarmento sobre Linguística de Corpus no meu ouvido quando o meu barco estava no meio da tempestade e aí eu descobri a riqueza da pesquisa em corpus de aprendiz e resolvi conceber o BELC. À minha querida e para sempre colega Simone, eu agradeço as dicas que mostravam a existência de outros rumos de navegação a serem pensados, não necessariamente tomados.

Outro vento bem favorável fez com que colegas e coordenadoras da PUC me ajudassem muito, sempre, desde o momento em que comecei a cursar os créditos à coleta dos textos do BELC. Serei eternamente grata às professoras Cristina Perna e June Campos, assim como aos colegas Anamaria Kurtz de Souza Welp, Adriana Rossa, Carlos Rossa, Clarice Lamb, Silvana Silveira, Ana Eliza Bocorny e à nossa querida Vera Muller!

Agradeço a pronta disponibilidade da Ana Bocorny, pelo carinho na leitura do meu texto.

Agradeço imensamente às duas queridas ex-alunas Ingrid Aline dos Santos Guerra e Natália Ramos Ferreira, sem as quais o BELC não teria ficado pronto para a análise em tempo regulamentar de um curso de doutorado. Muito obrigada, queridas, pelo carinho com o meu trabalho em cada momento da ajuda prestada!

Indiretamente, agradeço também aos 424 sujeitos da pesquisa cujos textos constituem o meu corpus.

Bastando de agradecimentos de caráter acadêmico, dedico agora essa parte aos agradecimentos de ordem mais pessoal, pois não foram só ventos favoráveis que fizeram o meu barquinho navegar até aqui: para isso, ele teve que ter uma estrutura razoavelmente sustentável.

Agradeço aos cuidados e ao carinho da querida Dra. Mari. Agradeço também à Luciana, que não me deixou passar fome e nem viver no meio da sujeira.

Agradeço aos meus sogros, Peter e Ielba, e cunhados Gustavo, Janaína e Jamile, pela constante mentalização positiva e à tia Elisa e ao Bacci pelo constante interesse pelo trabalho.

Aos meus pais, José Irineo e Vera, eu agradeço em especial o fato de sempre terem me incentivado a andar com as próprias pernas e sempre terem acreditado e confiado em mim. Ao meu pai, que muito erroneamente se considera acomodado, eu devo a ambição de querer saber o porquê e o como das coisas. À minha mãe, eu devo a disciplina, sem a qual, sabidamente, não se vai muito longe em nada. Ao meu tão querido irmão Neto eu agradeço a motivação constante e o reconhecimento do meu esforço com conhecimento de quem me acompanhou desde sempre.

Ao meu marido Henrique, eu agradeço o amor, o carinho, a paciência, mas principalmente o respeito por mim como profissional, sem o qual certamente eu não

teria investido tanto na minha carreira. Agradeço o incansável apoio ao me ouvir, ao ler, ao tentar entender sobre o assunto, ao dividir os momentos acadêmicos difíceis comigo, com quem eu divido, nesse momento, a indescritível alegria e satisfação de terminar um trabalho como uma tese de doutorado.

Last but not least, agradeço à minha tão amada filha Helena, que na sabedoria dos seus 7 anos, soube entender, à sua maneira, que a mamãe precisava estudar muito para ser doutora, mas não médica. A Helena, como toda a filha de doutoranda, desde os três aninhos escreve “pojetos” e até agora já escreveu várias teses sobre temas variados ao meu lado no seu *notebook* da Xuxa. Espero que um dia ela sinta algum orgulho de mim por esse trabalho, a fim de amenizar a culpa que sinto pelos vários momentos de ausência e indisponibilidade física e mental.

Enfim, agradeço a todos que contribuíram de alguma maneira para a elaboração e conclusão desta tese.

RESUMO

Este trabalho aborda a aquisição de morfemas em inglês como L2 por aprendizes falantes de português brasileiro como L1 através do BELC – *Brazilian English Learner Corpus* – um corpus de aprendiz elaborado para os fins deste estudo.

Tomamos como referências os estudos de Brown (1973) e de Krashen (1977). Brown observou o seguinte padrão evolutivo com relação à aquisição de morfemas do inglês na condição de L1: 1. *Present Progressive*; 2. *Plural*; 3. *Irregular Past*; 4. *Possessive*; 5. *Uncontracted Copula*; 6. *Article*; 7. *Regular Past*; 8. *Third person singular*; 9. *Uncontracted Auxiliary*. E Krashen (1977) aponta a seguinte ordem para a aquisição de morfemas em L2 (apresentada em estágios): 1. *Present Progressive, Plural, Copula*; 2. *Auxiliary, Article*; 3. *Irregular Past*; 4. *Regular Past, Third person singular, Possessive*.

Há dois estudos que se destacam na literatura ao tentarem propor explicações para a ordem de aquisição de morfemas em L2: Zobl e Liceras (1994) e Goldschneider & DeKeyser (2005). Zobl & Liceras (1994) oferecem uma explicação de natureza sintática para tal ordem, sugerindo que, no processo de aquisição de L1, as categorias lexicais e funcionais emergem de uma maneira específica, isto é, uma por vez. Já em L2, tais categorias emergem de maneira cruzada, pois os aprendizes de L2 dispõem de categorias funcionais desde os estágios de aquisição. Por outro lado, Goldschneider & DeKeyser sugerem uma explicação de natureza multifatorial para uma possível sequência natural na aquisição de morfemas fundamentada, de um modo geral, em aspectos de saliência.

Com o objetivo de verificar o comportamento dos aprendizes brasileiros de inglês como L2, propomos a investigação dos estágios de aquisição de morfemas através do BELC, um corpus elaborado a partir da coleta de textos produzidos por 424 sujeitos, totalizando aproximadamente 103 mil palavras. Os resultados obtidos apontam para a seguinte sequência (também apresentada em estágios): 1. *Copula, Plural*; 2. *Article, Possessive*; 3. *Present Progressive, Auxiliary, Irregular Past*; 4. *Regular Past, Third person singular*.

A sequência apresentada no BELC confirma vários estágios observados numa possível “ordem natural de morfemas” a partir de Krashen (1977). No entanto, duas diferenças são interessantes. O *Present Progressive* surge no terceiro estágio no BELC e no primeiro em Krashen (1977) e em Brown (1973). E o morfema *Possessive* surge no segundo estágio no BELC e no último em Krashen (1977) e, por isso, assemelha-se à posição de Brown (1973). Assim, as explicações oferecidas por Zobl e Liceras (1994) e Goldschneider & DeKeyser (2005) não conseguem explicar totalmente os resultados obtidos na pesquisa. Os resultados mostram que talvez fatores como instrução e a própria natureza da tarefa de coleta (escrita) podem ser apontados como possíveis intervenientes.

ABSTRACT

This dissertation deals with the acquisition of morphemes in L2 English by Brazilian Portuguese speakers through BELC – Brazilian English Learner Corpus. The study by Brown (1973) is used as a reference for the acquisition of L1 English. Brown observed the following developmental pattern in morpheme structures: 1. Present Progressive 2. Plural 3. Irregular Past 4. Possessive 5. Uncontracted copula 6. Article 7. Regular Past 8. Third person singular; 9. Auxiliary uncontracted. Krashen (1977) is used as a reference for the acquisition of grammatical morphemes in L2 English. He observed the following order of morphemes in acquisition (presented in stages): 1. Present Progressive, Plural, Copula, 2. Auxiliary, Article 3. Irregular Past, 4. Regular Past, Third Person Singular, Possessive.

Two studies attempt to offer explanations for a possible order. On the one hand, Zobl and Liceras (1994) offer an explanation that is based upon syntactic assumptions, suggesting that in the process of L1 acquisition, lexical and functional categories emerge in a specific way – one at a time. In L2 acquisition, such categories emerge cross-categorically, for the fact that L2 learners already have functional categories available from the early stages of acquisition. On the other hand, Goldschneider & Dekeyser (2005) suggest a multifactorial explanation for a possible natural sequence based on general aspects of salience.

In order to draw a developmental pattern of Brazilian learners of L2 English, we conduct an investigation about the stages of acquisition of morphemes through BELC, a corpus drawn from the collection of texts produced spontaneously by 424 subjects, which totaled approximately 103.000 words. The results indicated the following sequence (also presented in stages): 1. Copula, Plural, 2. Article, Possessive, 3. Present Progressive, Auxiliary, Irregular Past, 4. Regular Past, Third person singular. By comparing our study and the selected reference studies, we could observe several similarities and some differences in the stages of acquisition. Broadly, we can say that the sequence observed in BELC confirms a possible "natural order of morphemes" from Krashen (1977). However, two differences are interesting: the Present Progressive emerges in the third stage in BELC and in the first in Krashen (1977) and Brown (1973). And Possessive can be observed in the second stage in BELC and in the last stage in Krashen, a very similar position to Brown (1973). Thus, the explanations offered by Zobl and Liceras (1994) and Goldschneider & Dekeyser (2005) could not completely account for the results observed in the study. Factors such as instruction or the nature of the task (which was written) could have interfered in the results.

LISTA DE ABREVIATURAS

- AC:** Análise Contrastiva
- AE:** Análise do erro
- ART:** *Article*, ou Artigo
- ASL:** Aquisição de segunda língua
- AUX:** *Auxiliary*, ou verbo auxiliar BE no presente (*AM, IS, ARE*)
- BELC:** *Brazilian English Learner Corpus* (o corpus desta pesquisa)
- Br-ICLE:** *Brazilian International Corpora of Lerner English*
- BSM:** *Bilingual Syntax Measure*
- CAH:** *Contrastive Analysis Hypothesis*
- CC:** *Contracted Copula* - Cópula contraída, ou seja, verbo de ligação BE (*AM, IS, ARE*) utilizado como 'M, 'S, 'RE, respectivamente.
- CECL:** *Center for English corpus Linguistics*
- CLAWS:** *Constituent Likelihood Automatic Word-Tagging System*
- COP:** *Copula*, ou verbo de ligação (*AM, IS, ARE*)
- CREED:** *Construction-based, Rational, Exemplar-driven, Emergent e Dialectic.*
- DET:** *Determiner*, ou determinante
- EFL:** *English as a Foreign Language*
- ESL:** *English as a Second Language*
- GU:** Gramática Universal
- HAC:** Hipótese da Análise Contrastiva
- ICLE:** *International Corpora of Lerner English*
- IL:** *Interlanguage*, ou interlíngua
- INFL:** *Inflection*, ou constituinte flexional
- IRREG PAST:** *Irregular Past*, ou Passado irregular
- LAD:** *Language Acquisition Device*, ou dispositivo de aquisição de linguagem
- L1:** Língua 1
- L2:** Língua 2
- LE:** Língua estrangeira
- NP:** *Noun Phrase*, ou sintagma nominal
- OULC:** *Oxford University Learning Center*

PDP: Processamento de Distribuição Paralela

POS: *Part-of-speech*

PLU: *Plural*

POSS: *Possessive*, ou Possessivo 's

PREP: *Preposition*, ou Preposição

REG PAST: *Regular past* ou Passado regular

SOC: *Suppliance Obligatory Context*, ou contexto de preenchimento obrigatório

3rd P-S: *Third Person Singular* ou Terceira pessoa do singular

TPR: *Third Person Regular*, morfema

UG: *Universal Grammar*

UNC COP: *Uncontracted Copula*, ou Copula não contraída

V-ING: Verbo mais *ING*

VP: *Verb Phrase*, ou sintagma verbal

LISTA DE TABELAS

Tabela 1:	Ordem de aquisição de morfemas do inglês em Brown (1973).....	42
Tabela 2:	Ordem de aquisição de morfemas em De Villiers & De Villiers (1973).....	44
Tabela 3:	Ordens da aquisição de inglês como L2.....	47
Tabela 4:	Ordem de aquisição de Izumi & Isahara (2004).....	49
Tabela 5:	Comparação de Dulay & Burt (1974) e Wagner (2007).....	53
Tabela 6:	Porcentagens de correção de estudos incluídos na meta-análise.....	61
Tabela 7:	Contagem para cada morfema a partir de cada determinante.....	67
Tabela 8:	Ordem de aquisição de Brown (1973) & De Villiers & De Villiers (1973).....	74
Tabela 9:	Ordens de aquisição de inglês como L2 (a partir de Zobl & Liceras, 1994).....	75
Tabela 10:	Ordem de aquisição dos morfemas gramaticais em Brown (1973) e em Krashen (1977).....	75
Tabela 11:	<i>Design considerations for building learner corpora</i>	87
Tabela 12:	Sujeitos da pesquisa classificados por nível do curso.....	93
Tabela 13:	Classificação de proficiência segundo o OULC.....	95
Tabela 14:	Classificação de proficiência desta pesquisa.....	95
Tabela 15:	Informações sobre os textos.....	96
Tabela 16:	Estrutura geral do BELC.....	99
Tabela 17:	Códigos dos morfemas para a análise.....	109
Tabela 18:	Tabela sobre o morfema TPR – <i>Third Person Regular (+s)</i> ...	110
Tabela 19:	Tabela sobre o morfema TPR – <i>Third Person Regular (+s)</i> com registro.....	114
Tabela 20:	Tabela matriz.....	119
Tabela 21:	Morfema <i>Art (a, the)</i>	119
Tabela 22:	Morfema <i>Past Regular</i>	120

Tabela 23:	Morfema <i>Past Irregular</i>	121
Tabela 24:	Morfema <i>Present Progressive</i>	122
Tabela 25:	Morfema <i>Uncontracted Copula</i>	123
Tabela 26:	Morfema <i>Contracted Copula</i>	124
Tabela 27:	Morfema <i>Plural</i>	125
Tabela 28:	Morfema <i>Third Person Singular Regular</i>	126
Tabela 29:	Morfema <i>Third Person Irregular</i>	127
Tabela 30:	Morfema <i>Uncontracted Auxiliary</i>	128
Tabela 31:	Morfema <i>Contracted Auxiliary</i>	129
Tabela 32:	Morfema <i>Possessive</i>	130
Tabela 33:	Morfema <i>Prepositions in/ on</i>	131
Tabela 34:	Ordem de aquisição de morfemas a partir dos níveis de aprendizagem investigados no BELC.....	133
Tabela 35:	Sequência na aquisição de morfemas em blocos.....	134
Tabela 36:	Resultados de Brown (1973), Krashen (1977) e BELC.....	135
Tabela 37:	Resultados padronizados de Brown (1973), Krashen (1977) e BELC.....	137
Tabela 38:	Classificação dos morfemas analisados por Zobl & Liceras (1994) a partir do entendimento dos autores.....	140

LISTA DE FIGURAS

Figura 1:	Interface proposta na tese.....	20
Figura 2:	Morfema <i>Article</i>	120
Figura 3:	Morfema <i>Past Regular</i>	121
Figura 4:	Morfema <i>Past Irregular</i>	122
Figura 5:	Morfema <i>Present Progressive</i>	123
Figura 6:	Morfema <i>Uncontracted Copula</i>	124
Figura 7:	Morfema <i>Contracted Copula</i>	125
Figura 8:	Morfema <i>Plural</i>	126
Figura 9:	Morfema <i>Third Person Sing Regular</i>	127
Figura 10:	Morfema <i>Third Person Irregular</i>	128
Figura 11:	Morfema <i>Uncontracted Auxiliary</i>	129
Figura 12:	Morfema <i>Contracted Auxiliary</i>	130
Figura 13:	Morfema <i>Possessive</i>	131
Figura 14:	Morfema <i>Prepositions</i>	132

SUMÁRIO

1	INTRODUÇÃO	17
2	AQUISIÇÃO DE MORFOSSINTAXE EM L2	2517
2.1	MORFOSSINTAXE – MORFEMAS GRAMATICAIIS E CATEGORIAS FUNCIONAIS	25
2.1.1	Morfemas	26
2.1.2	Morfemas Gramaticais e Categorias Funcionais	29
2.1.3	Estrutura Sintática e Categorias Funcionais	31
2.1.4	Árvores e Categorias Funcionais	33
2.1.5	Categorias Funcionais e Movimentos	37
2.2	ESTUDOS SOBRE A ORDEM DE AQUISIÇÃO DOS MORFEMAS	41
2.2.1	Estudos em L1	42
2.2.2	Estudos em L2.....	46
2.3	RESUMO	54
3	ANÁLISE DA AQUISIÇÃO DA MORFOSSINTAXE EM L2	56
3.1	CRÍTICAS AOS ESTUDOS TRADICIONAIS DE AQUISIÇÃO DE MORFEMAS EM L2	56
3.2	EXPLICAÇÕES PARA A ORDEM DE AQUISIÇÃO DOS MORFEMAS	58
3.2.1	Goldschneider & DeKeyser (2005)	59
3.2.2	Zobl & Liceras (1994)	70
3.3	QUESTÕES A SEREM INVESTIGADAS	77
4	METODOLOGIA	79
4.1	LINGUÍSTICA DE CORPUS	79
4.1.1	Linguística de Corpus e Corpus de Aprendiz	79
4.1.2	Corpora de Aprendizes	83
4.1.3	Compilação e Análise de Corpora.....	87
4.1.3.1	Considerações sobre etiquetagem de erros.....	88
4.2	O CORPUS BELC – <i>BRAZILIAN ENGLISH LEARNER CORPUS</i>	93
4.2.1	A Coleta dos Dados.....	93
4.2.1.1	Os sujeitos	93
4.2.1.2	Procedimentos	94
4.2.1.2.1	Apresentação da pesquisa e identificação dos informantes	94

4.2.1.2.2	Nivelamento	94
4.2.1.3	Textos	96
4.2.1.4	A coleta dos dados propriamente dita	96
4.2.1.4.1	Planejamento da coleta	96
4.2.1.4.2	Considerações sobre a coleta	98
4.2.2	Estrutura do BELC.....	99
4.2.3	Panorama Geral do Corpus.....	101
4.2.4	Procedimentos de Análise do BELC	108
4.2.4.1	Considerações gerais sobre a etiquetagem	108
4.2.4.1.1	Procedimentos de etiquetagem	109
4.2.4.1.2	Material	109
4.2.4.1.3	Critérios de análise	110
4.2.4.1.4	Procedimentos específicos.....	112
4.2.4.1.5	Análise quantitativa	114
5	RESULTADOS E ANÁLISE	117
5.1	APRESENTAÇÃO DOS RESULTADOS.....	117
5.1.1	Ocorrência dos Morfemas	118
5.2	ANÁLISE	134
5.2.1	Comparação dos Resultados	135
5.2.2	A Perspectiva de Zobl & Liceras (1994).....	139
5.2.2.1	Breve retomada da perspectiva teórica.....	139
5.2.2.2	Aplicação aos resultados do BELC.....	142
5.2.3	A perspectiva de Goldschneider & DeKeyser (2005)	146
5.2.3.1	Breve retomada da perspectiva teórica.....	146
5.2.3.2	Aplicação aos resultados do BELC.....	147
5.3	OUTRAS QUESTÕES A SEREM CONSIDERADAS	151
5.3.1	O Fator Instrução	151
5.3.2	Tarefa escrita	153
	CONCLUSÃO.....	156
	REFERÊNCIAS	164
	ANEXOS	174
	Anexo 1: Teste de proficiência	
	Anexo 2: Answer sheet	
	Anexo 3: Instrução – Professores	
	Anexo 4: Parte I – Apresentação de pesquisa	
	Anexo 5: Parte II – Instruções (aluno)	
	Anexo 6: Parte III – Instruções (aluno)	
	Anexo 7: Tabela modelo de relação dos sujeitos participantes	
	Anexo 8: Relação de turmas	
	Anexo 9: Exemplos do BELC etiquetado	

1 INTRODUÇÃO

Este trabalho trata da aquisição¹ de morfemas em inglês como L2² a partir de um corpus de falantes brasileiros. Propõe-se aqui uma reavaliação dos resultados dos estudos dos morfemas conduzidos de maneira mais intensa nos anos 70 e 80, reavaliação esta baseada no estudo de aprendizes de inglês como L2 falantes de português brasileiro como L1. Mais especificamente, analisamos a ordem de aquisição de alguns morfemas estudados por Brown (1973) e Krashen (1977) a partir de um corpus de aprendizes elaborado primeiramente para servir a esta pesquisa, o BELC – *Brazilian English Learner Corpus*. Com este estudo, procuramos determinar o padrão evolutivo dos nossos aprendizes com relação a uma possível sequência natural na aquisição de morfemas.

A escolha do tema desta tese justifica-se por várias razões. A principal delas é a dificuldade dos alunos brasileiros no processo de aquisição da morfossintaxe do inglês, além da dificuldade dos professores em saber com clareza como se dá este processo. Não sabemos de estudos que contemplem estas questões, no sentido de determinarem estágios evolutivos dos aprendizes brasileiros no processo de aquisição morfossintática.

E acreditamos, como Kwon (2005), que:

knowledge of this learning hierarchy is very important to teachers and educational policy makers. For example, awareness of the order of acquisition that is natural to L2 learners may help curriculum developers to devise teaching sequences that are more adequate. (p. 16)
*Moreover, natural order studies remain relevant because they highlight the necessity for a change in our understanding of the mechanisms that govern language development. Most importantly, natural order studies have important implications for second language teaching. (p. 17)*³

¹ O termo “aquisição” será usado no decorrer deste trabalho indistintamente do termo “aprendizagem”, apesar de estarmos cientes das diferenças entre eles, especialmente aquelas apontadas por Krashen (1981) entre “*learning*” e “*acquisition*”, que serão inclusive mencionadas nesta tese.

² Nesta pesquisa, o inglês será referido como segunda língua (doravante L2), pois será abordado tanto como língua estrangeira como segunda língua, tendo em mente as colocações de Ellis (1994). Este autor propõe que o termo segunda língua seja associado à ideia de língua adicional, sem que se faça diferença entre segunda, terceira, quarta língua ou língua estrangeira. Sabe-se que o Brasil não possui uma segunda língua oficial e, portanto, o que se tem aqui é um ambiente de aquisição de uma língua estrangeira.

³ Manteremos as citações em inglês no corpo do trabalho, sem tradução, pois pressupomos que o leitor esteja familiarizado com o idioma.

Desta maneira, o principal objetivo desta tese é analisar a evolução do processo de aquisição, por aprendizes brasileiros, de um conjunto de morfemas (mais precisamente, de quatorze morfemas)⁴ em inglês como L2, a fim de: (i) estabelecer uma ordem de aquisição; (ii) determinar fatores que possam contribuir para esta ordem; (iii) fazer isso a partir de um corpus cujo banco de dados é constituído de produção linguística espontânea, isto é, de produção não estimulada para fins específicos. Esperamos, com este trabalho, contribuir para a teoria geral de aquisição de L2, além de contribuir para a prática pedagógica do ensino de L2 no Brasil. Nesse sentido, esperamos oferecer suporte teórico para que o professor compreenda melhor os estágios evolutivos experienciados pelo aluno brasileiro, bem como para que esta compreensão resulte na elaboração de materiais didáticos específicos a este público-alvo.

Além dos objetivos acima mencionados, de caráter mais teórico, esta tese também tem como objetivo prático a constituição de um corpus de referência para o estudo da produção linguística de aprendizes brasileiros do inglês como L2. A nosso ver, a construção do corpus BELC é um projeto inovador. Com ele, propomo-nos a organizar e codificar um corpus de textos escritos espontaneamente - isto é, os dados que constituem o corpus não foram coletados de maneira estimulada ou de outro modo dirigida à produção das estruturas sintáticas que são aqui estudadas. Assim, o corpus coleta um conjunto rico de textos, que poderão ser utilizados por outros estudos, com foco em outras estruturas ou em outras habilidades.

Como dissemos, este estudo contempla a análise da língua do aprendiz, tradicionalmente referida como uma *interlíngua* (IL). Este conceito, como sabemos, surgiu no fim dos anos 60 e início dos anos 70, proposto de forma independente por autores como Corder (1967) e Selinker (1972).

Selinker (1992) sustenta que o termo "interlíngua" preencheu uma lacuna semântica: a Análise Contrastiva discutia dois sistemas linguísticos, referidos tradicionalmente pelos termos NL (para *native language*, ou língua materna) e TL (*target language*, para língua-alvo), quando, na verdade, são três sistemas envolvidos num processo de "aprendizagem de uma língua estrangeira". Como

⁴ Veremos, no decorrer do trabalho, que estes morfemas serão, primeiramente, reduzidos a treze a partir da junção das preposições *in* e *on*. Num segundo momento, a análise será reduzida a nove morfemas, em consonância com as pesquisas de referência como Brown (1973) e Krashen (1977).

disse o autor, “*with the third unnamed, it seemed to me to be always implied*” (p. 222).

Gass & Selinker (2008) esclarecem um pouco mais esta noção: “*The basic assumption is that learners create a language system, known as interlanguage (IL). This concept validates learners’ speech, but as a system of its own with its own structure*” (p. 14). Para eles, o sistema é composto de vários elementos que podem não ser oriundos nem da língua materna nem da língua-alvo. Os autores completam: “*What is important is that the learners themselves impose structure on the available linguistic data and formulate an internalized system*” (p. 14).

White (2003) nos diz que a língua do aprendiz de L2 é sistemática e que os erros por ele produzidos não são aleatórios, mas, sim, sugerem um comportamento orientado por regras. E complementa: “*Such observations led to the proposal that L2 learners, like native speakers, represent the language they are acquiring by means of a complex linguistic system*” (p. 1).⁵

Propomos, então, a análise da aquisição dos morfemas em inglês por aprendizes brasileiros a partir de um corpus de produção linguística espontânea. Assim, podemos dizer que o presente trabalho é delimitado pela interface de três grandes domínios de investigação linguística: a **morfossintaxe** – por analisar os morfemas gramaticais e a estrutura sintática que subjaz a eles; a **aquisição de L2** – por analisar o processo de aquisição, por aprendizes brasileiros de inglês como L2, daqueles elementos gramaticais; e a **linguística de corpus** – por utilizar um corpus de produção espontânea, o BELC, criado primordialmente para esta tese a partir de princípios da linguística de corpus. Para visualizar claramente o domínio de estudo constituído por esta interface, é útil representá-la como na figura a seguir:

⁵ Utilizaremos, ao longo do trabalho, o termo *interlíngua* para nos referirmos, de um modo geral, ao sistema linguístico do aprendiz tal como é tradicionalmente entendido, conforme Ellis (1994), Gass & Selinker (2008) e as referências citadas acima.

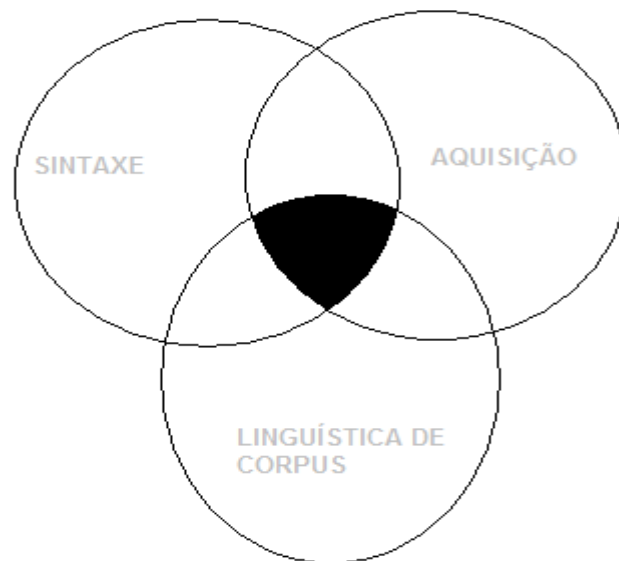


Figura 1: Interface proposta na tese

Sabemos da grande dificuldade em trabalhar com três áreas tão expressivas, cada uma com suas especificidades e prioridades. Mesmo assim, mantivemos a proposta do estudo principalmente porque acreditamos na relevância desta interface no sentido de enriquecer a compreensão do processo de aquisição de L2.

Quanto à estrutura, o presente trabalho se apresenta da seguinte maneira:

O capítulo 2 apresenta alguns fundamentos teóricos desta tese. Num primeiro momento, apresentaremos sucintamente alguns elementos básicos da teoria morfológica contemporânea, especialmente no que diz respeito às noções de "morfema", de "morfema gramatical" e de "categoria funcional", que são elementos teóricos fundamentais para a discussão posterior da tese. Depois disso, apresentaremos algumas pesquisas realizadas sobre a aquisição de morfemas em inglês como L1, como a de Brown (1973). Esta pode ser considerada uma das principais referências nos estudos sobre a aquisição de morfemas, já que foi o trabalho que descobriu a existência de uma ordem característica na aquisição materna do inglês. Especificamente, Brown observou a seguinte sequência de emergência de morfemas: 1. *Present Progressive*; 2. *Plural*; 3. *Irregular Past*; 4. *Possessive*; 5. *Uncontracted Copula*; 6. *Article*; 7. *Regular Past*; 8. *Third person singular*; 9. *Uncontracted Auxiliary*.

Ainda no capítulo 2, veremos também algumas pesquisas sobre a aquisição de morfemas em inglês como L2, especialmente a de Krashen (1977) – que é tida como uma referência fundamental no estudo da aquisição de L2. Entre suas observações, Krashen revela que a aquisição de inglês como L2 também se caracteriza por ordem de emergência de morfemas, muito semelhante à descoberta por Brown: 1. *Present Progressive, Plural, Copula*; 2. *Auxiliary, Article*; 3. *Irregular Past*; 4. *Regular Past, Third person singular, Possessive*⁶. Também veremos os resultados obtidos pelas pesquisas de Wagner (2007), Izumi & Isahara (2004), McEnery, Tono & Xiao (2006), entre outras.

O capítulo 3 fecha o foco em torno do estudo da aquisição de morfemas do inglês como L2, discutindo algumas propostas teóricas recentes sobre este processo. Apresentaremos algumas críticas feitas a estes estudos. Uma das principais é a de as propostas correntes – apesar de oferecerem uma descrição muito completa e detalhada do processo de aquisição de morfemas gramaticais, inclusive estabelecendo estágios e uma sequência razoavelmente constante de aquisição – não oferecem uma explicação satisfatória para o porquê de tais estruturas seguirem a ordem identificada. Neste capítulo, após uma discussão sucinta de vários trabalhos importantes da literatura, dedicamo-nos à discussão mais detalhada de dois estudos que são considerados, nesta mesma literatura, de extrema relevância para a tentativa de explicação do fenômeno: o de Goldschneider & Dekeyser (2005) e o de Zobl e Liceras (1994).

A análise de Goldschneider & DeKeyser (2005) nos diz que os morfemas são adquiridos por diferentes sujeitos numa sequência bastante semelhante devido a diferentes fatores, como a saliência do morfema, sua frequência, complexidade semântica, regularidade morfofonológica e complexidade sintática. Como veremos, os autores extraem dados de uma seleção de pesquisas realizadas sobre o tema a fim de compor uma meta-análise. Nesta meta-análise, procuram determinar a probabilidade que cada morfema tem de ser adquirido numa dada ordem, probabilidade que emerge dos cinco fatores acima mencionados. Os autores concluem que uma “análise multifatorial” poderia explicar de modo natural a sequência de aquisição de morfemas do inglês como L2. Consideram que sua

⁶ Utilizaremos os nomes dos morfemas em inglês ou em português, indistintamente, confiando que o leitor está familiarizado com a terminologia gramatical das duas línguas.

análise é particularmente interessante porque os fatores considerados podem, em última instância, ser manifestações de uma noção geral de "saliência linguística".

Já Zobl & Liceras (1994) propõe uma análise basicamente sintática para a sequência de emergência dos morfemas. Procuram oferecer uma explicação unificada do fenômeno, baseada em princípios linguísticos abstratos, seguindo o modelo da gramática universal assumido na teoria de princípios e parâmetros (CHOMSKY, 1981). Os autores nos dizem que, em L1, os morfemas emergem numa ordem caracterizada pela aquisição das categorias lexicais precedendo a das categorias funcionais.⁷ Em L2, os aprendizes já disporiam destas categorias; logo, o processo de aquisição seria caracterizado pela complexidade das operações sintáticas envolvidas na gramática dos morfemas – especialmente, no que diz respeito a necessidade da operação de movimento; assim, postulam que os morfemas livres precederiam os morfemas presos na aquisição de L2. Dessa forma, para Zobl & Liceras, aspectos sintáticos explicariam a ordem observada no processo de aquisição de morfemas em inglês como L2.

O capítulo 4 descreve a metodologia da pesquisa empírica que desenvolvemos. Fazemos, primeiramente, uma revisão dos estudos com corpora, discutindo questões fundamentais relativas à linguística de corpus e à pesquisa com corpora de aprendizes. Nesta perspectiva, um corpus pode ser definido, em linhas gerais, como um conjunto de dados linguísticos que sejam representativos dos sujeitos alvos da pesquisa e que sejam organizados eletronicamente para descrição e análise (BERBER SARDINHA, 2004); no entanto, a característica mais importante na concepção da linguística de corpus é a de que os dados tenham um caráter espontâneo. Ou seja, os textos não devem ser obtidos através de tarefas estimuladas, que orientem a produção de uma estrutura linguística específica. Veremos alguns exemplos de corpora de aprendizes, dentre os quais destacamos o ICLE, o *International Corpus of Learner English*, organizado pela Prof. Sylviane Granger, da Bélgica, com aproximadamente três milhões de palavras de textos escritos por falantes de quinze nacionalidades diferentes (GRANGER, 2003). Discutimos, ainda, aspectos da compilação e da análise de corpora, bem como da etiquetagem de erros, a fim de esclarecermos melhor os fundamentos que guiam a pesquisa aqui apresentada.

⁷ Esclarecemos que termos mais específicos serão melhor entendidos e ilustrados ao longo da tese, assim que se fizerem efetivamente presentes.

Num segundo momento, apresentamos o BELC, o corpus que elaboramos e que fornece os dados desta tese. Descrevemos a coleta dos dados com relação aos sujeitos, aos procedimentos utilizados (nivelamento, tipo de textos solicitados); descrevemos também a estrutura do BELC, bem como apresentamos um panorama geral do corpus e os procedimentos metodológicos utilizados para extrairmos os dados a partir dele. O BELC é constituído de 902 textos escritos por 424 sujeitos de oito níveis de aprendizagem, todos alunos do curso de inglês da Pontifícia Universidade Católica do Rio Grande do Sul – isto é, alunos que são oriundos dos mais diversos cursos de graduação daquela Universidade e cursam inglês como disciplina complementar ou eletiva. A produção destes aprendizes totalizou pouco mais de cento e três mil palavras (precisamente, 103.810 palavras). O corpus é representativo de acordo com os princípios adotados pela linguística de corpus, bem como por critérios estatísticos básicos.⁸

Os textos produzidos pelos alunos foram digitalizados, sendo que sua forma foi mantida na íntegra, ou seja, não tiveram qualquer tipo de alteração, nem mesmo em sua ortografia. Ainda na apresentação do BELC, discutiremos os procedimentos de etiquetagem adotados, bem como os critérios de análise e coleta dos dados – incluindo o cálculo de frequência de uso correto de morfemas nos textos.⁹

O capítulo 5 apresenta os resultados obtidos e a análise deles. A primeira seção mostra como os dados foram obtidos, especialmente no que diz respeito à porcentagem de correção de uso pelos sujeitos. Em seguida, apresentamos tabelas e figuras que mostram o padrão evolutivo do processo de aquisição de morfemas do inglês pelos aprendizes brasileiros. Mostramos a evolução de cada morfema no processo de aquisição que nos propomos analisar. Estabelecemos em 90% a proporção de uso correto que se pode considerar como indicação definitiva de que o morfema como adquirido¹⁰. A partir disso, apresentamos a tabela com a sequência apresentada pelos nossos aprendizes, que, em resumo, é a seguinte: 1. *Copula, Plural*; 2. *Article, Possessive*; 3. *Present Progressive, Auxiliary, Irregular Past*; 4.

⁸ Segundo os princípios desta metodologia de pesquisa, um corpus é considerado representativo com o número mínimo de cem mil (100.000) palavras.

⁹ Como veremos com mais detalhes no capítulo 4, o trabalho que propomos nesta tese não enfoca os erros dos aprendizes, e sim, os acertos, uma vez que o critério utilizado contempla o preenchimento apropriado dos contextos onde os morfemas são esperados. Para isso, naturalmente, devemos considerar o preenchimento não apropriado dos contextos, e isso poderia ser caracterizado como erro de natureza morfossintática. No entanto, a nossa proposta não pretende analisar esses "erros".

¹⁰ Oportunamente, veremos mais detalhes sobre os critérios de escolha da porcentagem de 90% como ponto de corte para considerarmos o morfema adquirido.

Regular Past, Third Person Singular. Como veremos, os dados confirmam, ainda que parcialmente, as sequências identificadas em outros estudos, em especial aqueles que utilizamos como referência, isto é, Brown (1973) e Krashen (1977).

Na segunda seção, propomos a análise dos dados obtidos. Em outras palavras, discutimos o padrão evolutivo dos aprendizes brasileiros, comparando-o à ordem de aquisição de morfemas encontrada nos estudos de Brown e Krashen. Assim, estabelecemos uma sequência que, acreditamos, fornece uma referência para o tipo de aprendiz aqui estudado – brasileiros que adquirem inglês estudando em condições regulares dos cursos de idiomas brasileiros; e também estamos em condição de apontar semelhanças e diferenças com as principais observações na aquisição materna e de inglês como L2. A seguir, discutimos possíveis explicações para os dados encontrados, procurando verificar, especialmente, até que ponto as propostas teóricas de Goldschneider & DeKeyser (2005) e Zobl & Liceras (1994) podem ser aplicadas a nossos resultados. Em síntese, veremos que nenhum dos dois trabalhos consegue cobrir de maneira satisfatória nossos resultados; hipotetizamos que um dos fatores a ser considerados em estudos futuros é o efeito da "ordem de instrução".

Acreditamos que esta tese possa oferecer uma contribuição significativa para o ensino de inglês como L2, especialmente no Brasil: seus resultados podem dar suporte teórico para os professores e pesquisadores acompanharem os estágios evolutivos dos aprendizes em seu processo de aquisição de estruturas morfossintáticas. Além disso, queremos ainda, com esta tese, mostrar a utilidade do estudo de corpora de aprendizes como fonte de dados espontâneos pertinentes à área de aquisição de L2.

2 AQUISIÇÃO DE MORFOSSINTAXE EM L2

O objetivo deste capítulo é apresentar o *background* teórico para o tema desta tese: a aquisição de morfemas em inglês como L2. Para tanto, o capítulo é dividido em duas seções:

Primeiramente, faremos a apresentação dos elementos fundamentais em teoria morfológica e sintática – especialmente os relativos às noções de “morfema” e de “categorias funcionais” – que utilizaremos nesta tese para nossa análise da aquisição dos morfemas em aquisição de segunda língua (ou ASL). Esta apresentação dará ênfase particular às concepções contemporâneas da relação entre morfemas gramaticais e categorias funcionais. Isso inclui apresentação de elementos de análise sintática, como as representações em árvores e as operações de movimento de morfemas (especialmente, a operação de “descida de afixo”, ou *affix-lowering*, do inglês).

Na segunda seção, apresentamos alguns estudos sobre a aquisição de morfemas em inglês como L1 (BROWN, 1973; DE VILLIERS & DE VILLIERS, 1973) e como L2 (DULAY & BURT, 1974; KRASHEN, 1977; PICA, 1983; ISUMI & ISAHARA, 2004). Veremos que os estudos parecem apontar para uma ordem natural de aquisição de morfemas.

2.1 MORFOSSINTAXE – MORFEMAS GRAMATICAIS E CATEGORIAS FUNCIONAIS

Spencer (1991) nos diz que morfossintaxe pode ser entendida como uma das funções da morfologia. Ou pode, ainda, ser entendida como a interface entre dois componentes gramaticais distintos, a morfologia e a sintaxe, sendo a primeira o componente responsável pela estrutura interna das palavras e a segunda, o responsável pelas estruturas formadas por combinações de palavras – sintagmas e sentenças. Assim, poderíamos definir uma “categoria morfossintática” como aquela que contempla as dimensões de um paradigma morfológico – como a flexão de número, por exemplo – voltado para alterações na estrutura da palavra que

dependem ou têm consequências para a estrutura sintática em que a palavra será usada.

Nosso estudo trata da aquisição da morfossintaxe em L2; mais especificamente, está orientado para a aquisição, por aprendizes brasileiros, dos aspectos da gramática do inglês que se manifestam pelo uso de morfemas gramaticais como as formas *-s* de plural e *-ing* da construção progressiva. Por esta razão, procuraremos delimitar melhor, nas seções a seguir, aquilo que se entende por “morfemas gramaticais”, bem como procuraremos apresentar, ainda que sucintamente, o modo como a teoria gramatical contemporânea – especialmente as abordagens gerativas¹¹ – concebe a organização sintática na qual tais morfemas atuam.

2.1.1 Morfemas

A partir da linguística estruturalista americana, os morfemas têm sido definidos como as “unidades mínimas de significado” (cf. SPENCER, 1991). Ou, ainda, são “os menores elementos com significado numa língua” (BUBENIK, 1999).

Os tipos de morfemas podem ser distinguidos entre si de diversas maneiras. Uma das classificações mais tradicionais é aquela que distingue entre os “morfemas lexicais” e os “morfemas gramaticais”. Chamamos de morfemas lexicais aqueles que atuam na expansão do léxico, como o morfema *ship* em *friendship*. E morfemas gramaticais são os morfemas que têm “papel gramatical” – morfossintático – na língua. Por exemplo, o *-s* da conjugação de terceira pessoa do presente simples em inglês, como em *She works*. Trata-se, portanto, normalmente de morfemas *flexionais* – flexão de número, concordância, tempo, caso, etc. Para Lyons (1979), a flexão é uma mudança na forma da palavra para exprimir a sua relação com outras palavras na frase. A flexão descreve as declinações dos substantivos, dos adjetivos e dos

¹¹ Estas abordagens estão vinculadas à chamada “teoria da gramática universal”, cuja primeira formulação explícita – em particular, de sua relação com os estudos de aquisição da linguagem – apareceu Chomsky (1965). Esta tese não se propõe a discutir teorias linguísticas ou de aquisição de linguagem. Por isso, assume-se que o leitor disponha de conhecimentos básicos da abordagem aqui adotada, que é a versão da “teoria da gramática universal” baseada no modelo “princípios e parâmetros”. Ver, para uma apresentação autorizada desta versão, Chomsky (1981) e (1988), e, quanto ao estudo de aquisição de L2 sob esta perspectiva, ver White (1989), (2003) e (2005).

pronomes, as conjugações dos verbos. Já a derivação se refere a vários processos pelos quais se formam novas palavras ou raízes existentes. Por exemplo: substantivos formados de verbos, como *singer* a partir de *sing*. Nessa perspectiva, entendemos que morfemas lexicais são aqueles geralmente associados a propriedades de derivação, ao passo que morfemas flexionais ou gramaticais, associados a propriedades de flexão.

Ainda, independentemente da definição acima, podemos diferenciar aqueles morfemas que podem ser caracterizados como palavras independentes daqueles que são somente subpartes de palavras. Os primeiros são tradicionalmente conhecidos como morfemas livres, porque podem aparecer sozinhos, e os segundos, morfemas presos, porque nunca podem aparecer sozinhos. Os morfemas livres, normalmente, correspondem a classes abertas de palavras, também conhecidas como “categorias lexicais”. São “abertas” porque não formam um sistema limitado de oposições e podem ser facilmente expandidas, como é o caso das classes dos substantivos e dos verbos. Mas há morfemas livres que pertencem a classes fechadas e que, tal como os “morfemas gramaticais”, pertencem a um sistema limitado de oposições com função importante na estrutura gramatical da língua: estas classes fechadas de morfemas livres são chamadas, especialmente em modelos correntes de teoria gerativa, de “categorias funcionais”. Incluem, por exemplo, os determinantes (artigos definidos, indefinidos, pronomes indefinidos adjetivos, entre outros elementos) e os complementizadores (as conjunções integrantes). As categorias funcionais, por formarem sistemas limitados de oposições gramaticais, raramente admitem novos membros e, quando os admitem, fazem-no após um longo processo de gramaticalização (cf. FROMKIN, RODMAN & HYAMS, 2007).

Tipicamente, uma palavra morfologicamente complexa contém um morfema central, que oferece o significado lexical básico, idiossincrático à palavra, e outros morfemas que servem para modificar este significado básico de várias maneiras. Por exemplo, a palavra *disagreements* possui, como morfema básico e específico a ela, *agree* “concordar”; além deste morfema básico, possui três outros morfemas presos: *dis-*, *-ment*, e *-s*. Destes três outros morfemas, dois criam novas palavras: *dis-*, aplicado à raiz verbal *agree*, cria o verbo *disagree* “discordar”; e *-ment* cria o substantivo *disagreement* “desacordo” a partir de *disagree*. O morfema *-s*, por outro lado, tem uma função diferente: não cria novo vocábulo, mas apenas modifica

semanticamente *disagreement*. Podemos dizer, assim, que *agree* é a raiz e que os outros três morfemas são afixos; quanto à forma, *dis-* é um prefixo porque precede a raiz, e *-ment* e *-s* sufixos porque sucedem a raiz; quanto à função, *dis-* e *-ment* são afixos “derivacionais”, e *-s* é um afixo flexional – ou “puramente gramatical” (SPENCER, 1991).

Portanto, com relação à função, tradicionalmente se distingue entre dois tipos principais de operação morfológica, a flexão e a derivação. A flexão se refere às variantes de uma mesma palavra: mesmo sendo flexionada, uma palavra não altera o conceito básico que denota, o que se revela, por exemplo, pelo fato de não alterar sua categoria sintática. Por exemplo, assim como o *-s* de plural em inglês não cria novas palavras, o acréscimo do *-s* aos verbos também não: aplicado ao verbo *work* serve apenas para permitir a combinação gramatical deste verbo com um sujeito de 3ª. pessoa do singular, como em *He works*. Já a derivação é uma operação diferente, cuja finalidade é a formação de palavras, isto é, a criação de novas palavras a partir de outras já existentes. Isso tipicamente, mas não necessariamente, induz a uma mudança em categorias sintáticas: por exemplo, a adição do sufixo *-ly* ao adjetivo *happy* “feliz” cria o advérbio *happily* “de modo, maneira, etc., feliz”. Por isso, às vezes podemos dizer que os morfemas derivacionais também são “morfemas lexicais” ou “do léxico” – sua função é operar na expansão do léxico ou vocabulário, criando novas unidades conceituais na língua. Esta é a função de morfemas presos como, por exemplo, *-ish*, *-ous*, *-ship*, *-ity*, *-n*, em inglês. Já os morfemas flexionais, como já mencionamos, são também chamados de “morfemas gramaticais” ou ainda “morfossintáticos” porque sua função é adaptar, morficamente, a palavra ao ambiente gramatical em que ocorrerá na frase. São exemplos destes últimos, em inglês, os afixos *-ing* do particípio progressivo dos verbos, *-s* de plural dos substantivos, *-s* de 3ª. pessoa do presente dos verbos, ou, ainda, *-ed* de passado dos verbos.

Além das distinções de função acima mencionadas entre os diferentes tipos de morfemas, é preciso ainda ter em mente, em qualquer estudo que envolva os elementos morfológicos de uma língua, que os morfemas, embora sejam unidades de sentido das línguas, normalmente possuem também alguma forma concreta de exponência - isto é, uma realização fonética ou fonológica. E, quando se considera esta forma concreta de realização, descobre-se que, na verdade, um mesmo morfema – uma mesma unidade de sentido na composição das palavras – pode se

materializar de um, dois ou mais modos distintos. Isso, é claro, é manifestação da independência parcial que há entre significado e forma, razão pela qual, segundo Spencer (1991), a teoria morfológica muito se ocupa em como se dão as diferentes possibilidades de mapeamento entre forma e função.

Esta possibilidade de múltipla realização de forma de um mesmo morfema é conhecida como **alomorfia**. Por exemplo, a terminação de plural regular em inglês, que é normalmente representada genericamente pela forma escrita do afixo, isto é, por “-s” apenas, pode se manifestar em três pronúncias diferentes, /s/, /z/ e /əz/ (em que “/.../” significa que se trata da transcrição fonológica de uma forma). Uma vez que estes três elementos são manifestações concretas de um morfema, eles são chamados de *morfes* (LYONS, 1979). Assim, podemos dizer que os morfes /s/, /z/ e /əz/ são alomorfes de -s; e dizemos, ainda, que este morfema de plural do inglês exhibe alomorfia. Este fenômeno é claro, deve ser levado em conta no estudo da aquisição dos morfemas; afinal, é concebível, por exemplo, que morfemas que possuem inúmeras realizações alomórficas sejam mais difíceis de serem adquiridos que os morfemas sem grande variação de forma.

2.1.2 Morfemas Gramaticais e Categorias Funcionais¹²

Fromkin, Rodman & Hyams (2007) definem morfemas gramaticais como morfemas presos que têm uma função estritamente gramatical. Eles marcam propriedades como tempo, número, gênero, case, etc., e nunca alteram a categoria gramatical da palavra à qual se prendem.

Já para Cook (2001), morfemas gramaticais são morfemas presos como –*ing*, ou palavras gramaticais como o artigo *the* – isto é, palavras que desempenham um papel mais específico na estrutura sintática da frase do que palavras de conteúdo como *horse*.¹³ Portanto, nesta última visão, contrariamente à visão mais tradicional, também “palavras gramaticais” como *the* são incluídas na classe dos

¹² Ou Flexionais. Em inglês, eles são referidos como *grammatical morphemes* ou *inflectional morphemes*. Para os fins desta tese, adotaremos morfemas gramaticais.

¹³ Cook (2001) distingue palavras de conteúdo de palavras de estrutura (do inglês, respectivamente, *content words* e *structure words*), definindo “palavras de conteúdo” como aquelas cujo significado pode ser encontrado num dicionário e que pertencem a classes abertas, podendo ser encontradas em grande número; já “palavras de estrutura” seriam aquelas cujo é normalmente descrito numa gramática, sendo limitadas em número – por exemplo, o artigo *the*, a conjunção *that* ou a preposição de infinitivo *to*.

“morfemas gramaticais”. Isso expressa a intuição de que, em algum grau e com respeito a certos processos, elementos gramaticais como o afixo de plural –s e os artigos, ou o afixo de passado –ed e verbos auxiliares como *will*, são elementos com a “mesma natureza” gramatical. Adotando, ainda, a terminologia da teoria sintática gerativista, são, do ponto de vista sintático, “categorias funcionais” – independentemente de sua classificação puramente mórfica em afixos ou palavras livres. Assim, Fromkin, Rodman & Hyams (2007) nos dizem que há oito morfemas gramaticais em inglês, que são os abaixo listados:

1. *Plural “s”*
2. *Progressive “-ing”*
3. *Third person sing “s”*
4. *Reg Past ED*
5. *Past Participle EN*
6. *Possessive ‘s*
7. *Comparative ER*
8. *Superlative EST¹⁴*

Já na perspectiva da teoria gerativista, seriam ainda incluídas na lista acima “palavras”, isto é, “morfemas livres”: os auxiliares *will*, *be*, etc., junto com o regular past –ed, que pertencem à categoria funcional “INFL”; o artigo *the* e os demonstrativos (*this*, *that*, etc.), mais o afixo possessivo ‘s, que pertencem à categoria funcional “DET”; e assim por diante.

É importante observar que, em estudos de aquisição da linguagem, é esta a concepção – de que, por exemplo, artigos e a flexão de plural –s ou o afixo genitivo ‘s tem similaridades funcionais e gramaticais – que subjaz ao termo “morfemas gramaticais”: afinal, em todos os estudos que mencionamos acima, quando se estuda a ordem de emergência dos “morfemas gramaticais”, incluem-se nas listas não apenas afixos, mas também palavras gramaticais como o auxiliar *be* e os artigos. Não por acaso, as teorias contemporâneas sobre a “aquisição dos morfemas gramaticais” foram buscar subsídios para entender este processo na teoria

¹⁴ Embora a autora não afirme, acreditamos que ela esteja limitando os morfemas gramaticais a morfemas gramaticais presos e, por isso, não considerando outros como o *COP BE*, por exemplo, que, em todos os outros trabalhos revisados nesta tese, é considerado um morfema gramatical.

gerativista – que fornece uma visão integrada destes elementos gramaticais (o que veremos ilustrado em Zobl & Licerias, 1994).

Para a teoria gerativista corrente, como se lê em Haegeman (2007), a propriedade definitória dos núcleos funcionais – ou "morfemas gramaticais" – é que eles não atribuem papéis temáticos, contrariamente a núcleos que pertencem às "categorias lexicais".¹⁵ Por exemplo, no caso das categorias funcionais associadas ao verbo, que são manifestações do núcleo INFL, a função é colocar o evento expresso pelo verbo numa localização temporal. Vejamos a seguir mais alguns dos elementos básicos da teoria sintática que adotaremos nesta tese e que subjazem à concepção das categorias funcionais que assumiremos.

2.1.3 Estrutura Sintática e Categorias Funcionais

Sabemos que as frases têm uma estrutura que pode ser representada por meio de árvores de estrutura frasal cujos nós são rotulados por categorias sintáticas.¹⁶ Uma **categoria sintática** é uma classe paradigmática de expressões, isto é, um conjunto de expressões que podem ocupar a mesma posição sintática e que, por isso, podem substituir umas às outras sem afetar a gramaticalidade da sentença. Por exemplo, “*a police officer*” é uma expressão que pertence a categoria sintática denominada de "sintagma nominal" (do inglês *noun phrase*) e, por isso, compartilha a distribuição sintática, bem como outras propriedades, com outros elementos da categoria – como, por exemplo, os sintagmas nominais *the officer*, *the men who is a police officer*, os nomes próprios como *Bill* e *John*, e os pronomes, como *he* ou *him*.

Categorias sintáticas podem possuir diferentes "níveis de expansão": podem ser formadas por uma palavra (por exemplos, *Bill* e *He* são sintagmas nominais formados por apenas uma palavra) ou mais de uma palavra (como no caso de *a police officer*). Independentemente de serem formados por uma ou mais palavras,

¹⁵ Ou seja, segundo a teoria dos papéis temáticos, os núcleos funcionais não atribuem conteúdo referente a informações sobre as propriedades lexicais destes elementos.

¹⁶ Para uma introdução elementar à estrutura sintagmática da frase, ver Fromkin, Rodman & Hyams (2007), cap. 4. Para introduções à teoria gerativista contemporânea, especialmente a chamada "Teoria de Princípios e Parâmetros", ver Chomsky (1981), Chomsky (1988), Mioto, Silva e Lopes (2005) e Haegeman (2007).

possuem sempre um "núcleo", que é a palavra que determina o tipo sintático da categoria: por exemplo, sintagmas nominais são "nominais" porque seu núcleo – a palavra mais importante, mais necessária sintaticamente – é uma "palavra nominal" (um substantivo, um pronome).

As categorias sintáticas podem ter como núcleo as chamadas "categorias lexicais", que pertencem às classes abertas e podem atribuir papel temático a seus argumentos, e também as "categorias funcionais", que possuem uma função gramatical em geral relacionada às flexões de alguma categoria sintática lexical. As categorias lexicais são N (do inglês *noun*, ou nome/substantivo), V (verbo), P (preposição), Adj (adjetivo) e Adv (advérbio); estas categorias, quando expandidas em categorias sintagmáticas, resultam, respectivamente, nos sintagmas nominais ou NPs (do inglês *noun phrase*), sintagmas verbais ou VPs (*verbal phrase*), sintagmas adjetivais ou AdjPs (*adjective phrase*), sintagmas preposicionais ou PPs (*prepositional phrase*) e sintagmas adverbiais ou AdvPs (*adverbial phrase*).

Já as categorias funcionais reconhecidas em sintaxe gerativa são as seguintes: os determinantes ou Dets, que incluem os artigos, os demonstrativos, os possessivos e o marcador genitivo 's em inglês (ver especialmente ABNEY, 1986); as "flexões verbais" ou INFL, que em inglês incluem os verbos auxiliares, os morfemas *-ed* de passado e *-s* de (3ª. pessoa do singular) de presente e a preposição *to* de infinitivo em inglês; e, finalmente, os complementizadores ou COMP, que incluem as conjunções *that*, *if* e a preposição *for* quando introduz orações infinitivas. Cada uma destas categorias funcionais é núcleo de seu próprio sintagma: respectivamente, o sintagma determinante ou DP (*determiner phrase*), o sintagma flexional ou IP (*inflectional phrase*), e o sintagma complementizador ou CP (*complementizer phrase*).

As categorias funcionais são vistas, especialmente na teoria de princípios e parâmetros, como o *locus* de variação paramétrica, isto é, da variação que dá origem às diferenças sistemáticas entre as gramáticas das línguas. São, por isso, também o *locus* de possíveis diferenças nos processos de aquisição das línguas, ou das diferenças entre a aquisição de L1 e L2 (WHITE, 2005; ver, por exemplo, a discussão do trabalho de Zobl e Liceras no próximo capítulo). A variação entre línguas, no que diz respeito às categorias funcionais, pode ser encontrada: (a) na própria presença ou não de alguma delas, já que nem todas as categorias são realizadas em todas as línguas (ver, por exemplo, FUKUI e SPEAS (1986) para a

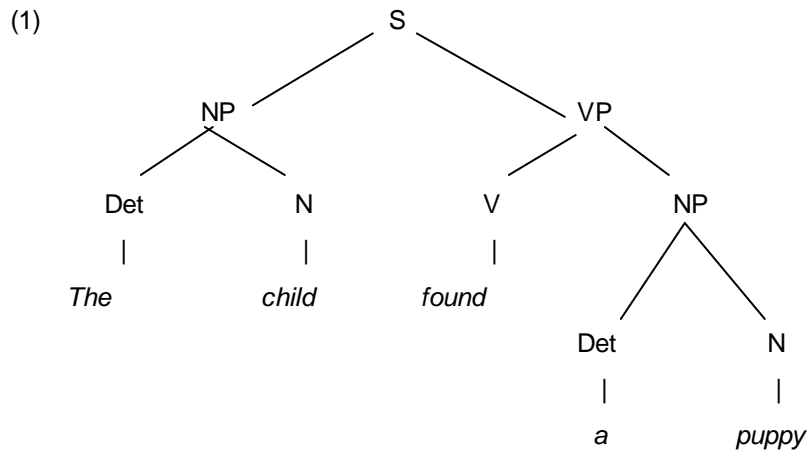
ausência de artigo em japonês); (b) na presença ou ausência de certos traços formais de alguma destas categorias em alguma língua (por exemplo, a ausência de traços de gênero em *Article* em inglês); e, ainda, levando em conta os recentes desenvolvimentos minimalistas, (c) é possível que haja também variação quanto à "força" de um traço, o que tem como consequência que as línguas, em virtude disso, poderão variar quanto às possibilidades de movimento. Por exemplo, inglês apresenta "descida de afixo" ou movimento de -INFL-para-V porque seus traços de concordância são "fracos"; português apresenta "subida do verbo ou movimento de V-para-INFL porque seus traços de concordância são "ricos"; ver Hornstein, Nunes e Grohman (2005).¹⁷

2.1.4 Árvores e Categorias Funcionais

A fim de ilustrar didaticamente o tipo de informação sobre as categorias sintáticas que está por trás da construção de uma frase, considere a seguinte estrutura de constituintes simplificada de uma sentença, que mostra não só a organização linear das palavras que a compõem, mas também o modo como se combinam em constituintes:¹⁸

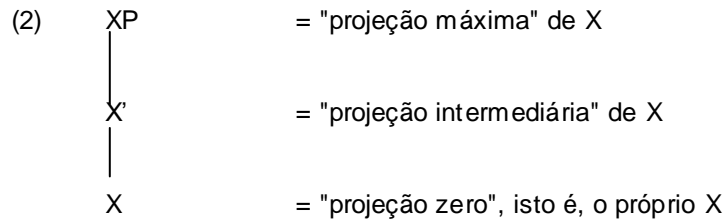
¹⁷ Na perspectiva minimalista, a questão da "força" dos traços diz respeito à relevância que estes podem ter para que uma estrutura gramatical tenha sucesso ao ser processada por um dos dois níveis fundamentais da "interface" gramatical: o nível mais abstrato, denominado forma lógica, doravante FL (do inglês, *logical form*) e o nível mais concreto, a forma fonológica, doravante FF (do inglês, *phonological form*). Na verdade, trata-se de um modo de reconceber a motivação das operações gramaticais de movimento dentro da teoria gerativa: a partir de Chomsky e Lasnik (1995), movimento deixou de ser visto como uma operação acionada "diretamente pela morfologia" das categorias sintáticas – isto é, o verbo deixa de se mover para "prender seus afixos", assim como os DPs deixam de se mover para "receber Caso". Ao invés disso, as categorias se movem para checar traços formais para os quais são especificadas lexicalmente. Estes traços, se não forem checados, resultam em má formação seja de FF ou de FL. Segundo Chomsky (1995), os traços formais de uma categoria podem ser parametrizados com o valor forte ou fraco. Se são fortes, sua checagem deve ocorrer antes de *spell-out*, a operação que envia para a FF as informações necessárias para a representação fonética, superficial, da frase. Assim, quando os traços são fortes, o movimento é visível em FF. Se são fracos, serão checados somente em FL, e o movimento não precisa ser anterior a *spell-out*; portanto, não será visível em FF. Assim, a diferença que vemos nas línguas com relação à ordem de palavras pode ser explicada pela diferença paramétrica quanto à força dos traços de categorias como DET e INFL. Ver Hornstein, Nunes e Grohmann (2005) para mais detalhes, e a seção 2.1.5 abaixo para exemplos.

¹⁸ A estrutura em (1) está sendo utilizada aqui apenas para fins ilustrativos. Trata-se da análise da estrutura da sentença feita ainda no chamado "modelo padrão" da gramática gerativa (CHOMSKY, 1965) e não envolve todos os elementos e relações que as análises contemporâneas assumem para a frase. Logo abaixo, apresentaremos análises mais atualizadas.



Esta representação ilustra os aspectos básicos do conhecimento que o falante tem sobre os elementos que compõem a frase e as estruturas que estes elementos estabelecem: a sentença como um todo é formada, minimamente, pela união de um sintagma nominal (NP em (1)) e um sintagma verbal (VP em (1)). Tais sintagmas são formados pela combinação de constituintes menores, numa certa ordem linear: NP é formado por DET e N. Há uma palavra que determina a categoria do sintagma maior : N determina a categoria do NP, etc. Dentro da linha gerativa de investigação linguística, podemos dizer que as árvores são representações que codificam os fatos fundamentais da sintaxe da frase – o que, tecnicamente, se chama de "estrutura sintagmática".

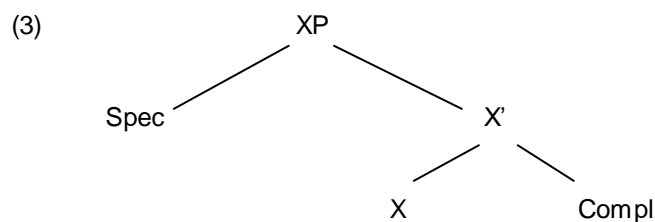
A teoria gerativa contemporânea propõe que a estrutura sintagmática da frase resulta de princípios universais, normalmente expressos por meio de alguma versão da chamada Teoria X-Barra. Os princípios desta teoria, diante de propriedades como as que ilustramos acima e outras descobertas pelos gerativistas (ver referências da nota 29 e autores lá citados), estipulam as propriedades que passamos a mencionar: (a) todo o constituinte deve ser construído a partir de um **núcleo**; (b) considerando a estrutura básica de um constituinte, é preciso reconhecer, além do próprio núcleo, pelo menos dois outros "níveis de projeção" dele. Tecnicamente, utilizando-se X como uma variável para as diferentes categorias lexicais, podemos representar esta "projeção em três níveis" do seguinte modo (ver MIOTO, SILVA E LOPES 2005, p. 47):



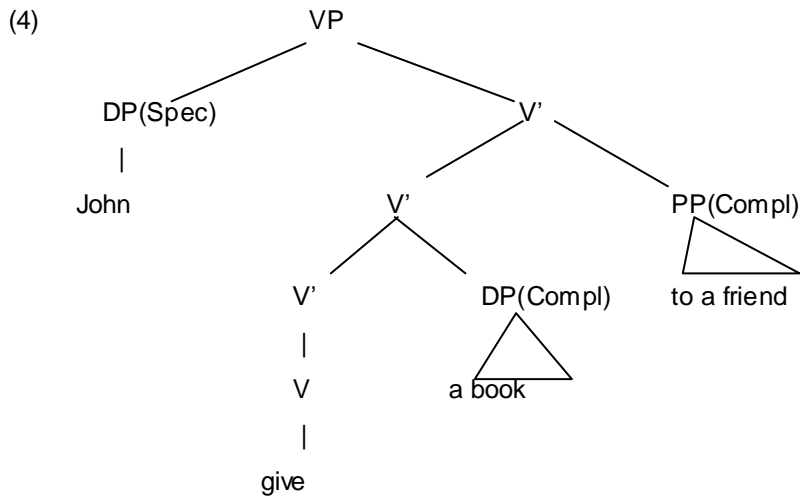
em que X é alguma das categorias sintáticas:

N, V, Adj, Adv, P, mas também DET, INFL, COMP.

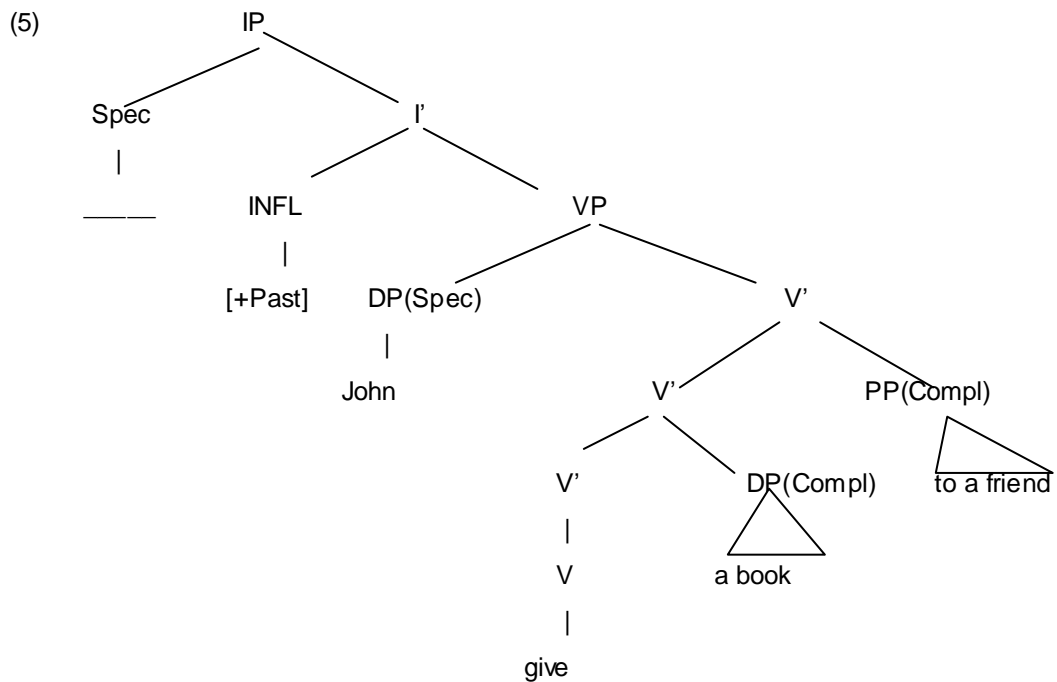
(c) O que determina os três níveis de projeção é a existência de relações gramaticais universais entre um núcleo e outros constituintes; mais especificamente, dependendo das propriedades lexicais do item lexical que ocupa a posição de núcleo, todo sintagma pode conter um (ou mais) complemento(s) (Compl) do núcleo, que universalmente dão origem à(s) projeção(ões) intermediária(s) do núcleo, e um especificador (Spec), que é um elemento que "fecha" a projeção do núcleo, fazendo com que atinja sua "projeção máxima":



Um exemplo de uma estrutura sintagmática resultante destes princípios seria, por exemplo, a expansão de um verbo em um VP – que seria um dos constituintes básicos da sentença, mas não a própria sentença, na teoria de princípios e parâmetros:



A hipótese fundamental da teoria gerativa contemporânea que nos interessará aqui é que a estrutura dos sintagmas não é apenas uma projeção X-barra das categorias lexicais como NPs e VPs. Também as categorias funcionais participam da estrutura dos sintagmas projetando, por sobre as projeções máximas das categorias lexicais, suas próprias projeções X-barra. Assim, na concepção gerativista a partir da teoria de princípios e parâmetros, a estrutura da sentença não é como em (1) anteriormente. É, na verdade, como (5) abaixo, isto é, é a projeção da categoria funcional INFL, que toma como complemento um VP e possui uma posição de especificador que, após movimento, será ocupada pelo DP sujeito:



Assim, na concepção gerativista contemporânea, "morfemas gramaticais" são unidades sintáticas ativas, com caráter de núcleo como outras categorias sintáticas. Por isso, do ponto de vista desta teoria, estas unidades devem ser estudadas também como elementos que pertencem a um conjunto paradigmático definido por sua função sintática, e não apenas por sua "forma". Isto é, além de serem "formas presas" do ponto de vista da morfologia, são também "núcleos funcionais" do ponto de vista sintático. Se, sob o ponto de vista da morfologia, devem ser estudados juntos com outros afixos flexionais, do ponto de vista sintático, devem ser estudados juntos com outros elementos que, unidos a um VP, resultam numa frase completa: por exemplo, em inglês, pertencem a mesma categoria sintática que o auxiliar *will*, ou os auxiliares modais como *must* e *can*.

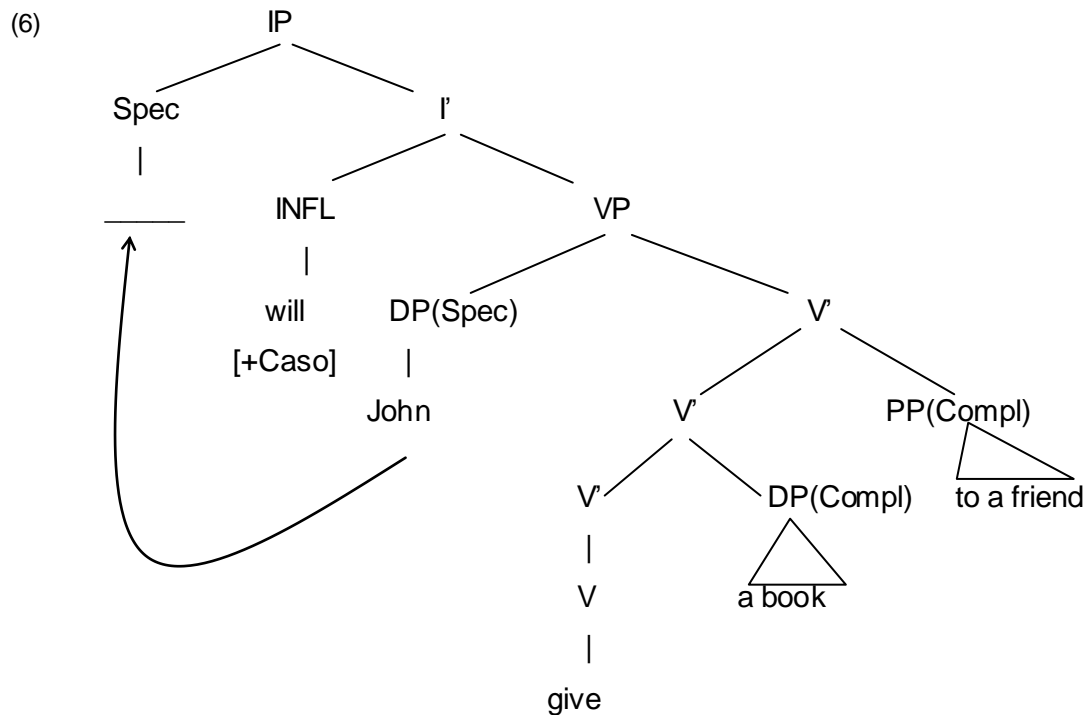
Assim, quando se estuda a aquisição dos "morfemas flexionais", é preciso levar em conta este seu caráter "híbrido", complexo, de "formas dependentes" com caráter de "núcleo sintático funcional". Como veremos na próxima seção, esta complexidade tem consequências – que serão importantes no estudo da aquisição destes elementos em L2.

2.1.5 Categorias Funcionais e Movimentos¹⁹

Como já mencionado anteriormente, para a sintaxe gerativista contemporânea, as categorias funcionais – seja pelo seu caráter morfológico de afixos, conforme as primeiras versões da teoria das categorias funcionais, seja pelo conjunto de "traços formais" que carregam, como no minimalismo – acionam o movimento de diferentes elementos da frase, intervindo diretamente na estrutura superficial das sentenças das línguas. Assim, há um consenso generalizado de que as sentenças do inglês, por exemplo, tem como representação gramatical básica algo como a estrutura em (5) anteriormente; e a estrutura que acaba sendo pronunciada em FF resultada de (5) por meio de uma operação de movimento que leva o DP *John* para a posição de Spec-de-IP, seja para receber Caso nominativo

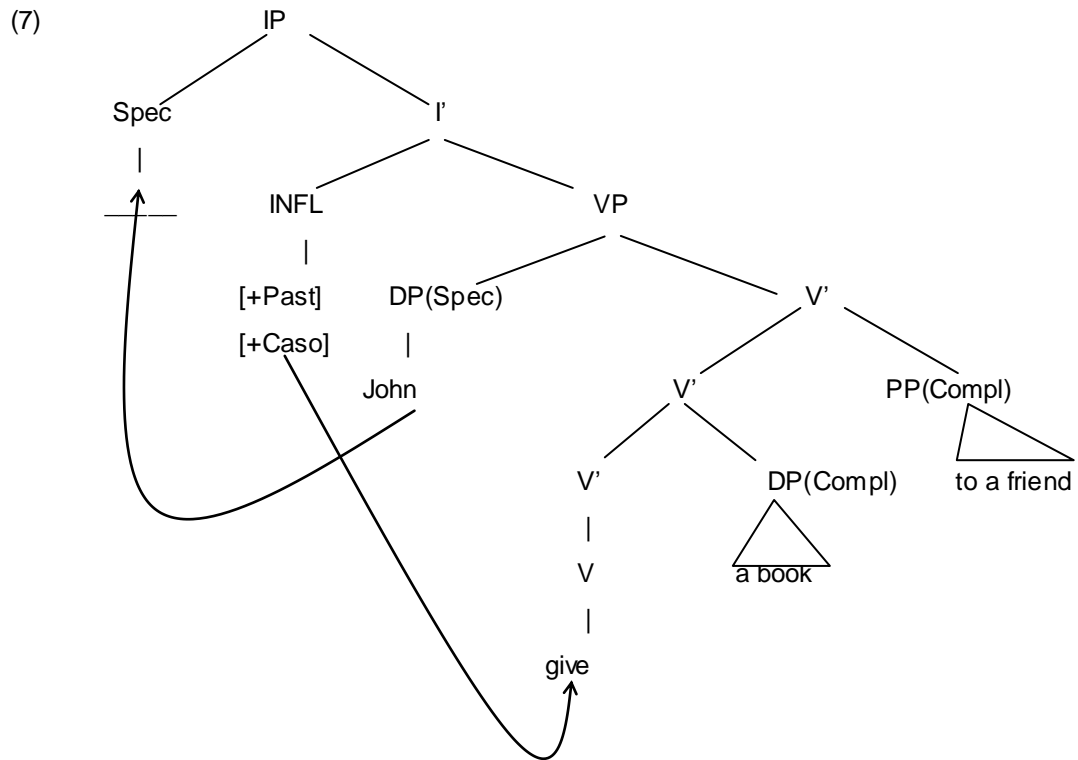
¹⁹ Sabemos que há uma série de operações de movimentos relacionados à flexão verbal. No entanto, como o assunto desta tese é, em linhas gerais, a aquisição de morfemas em inglês como L2, nos limitaremos a analisar o movimento que é mais específico para este estudo: a descida do afixo.

de INFL [+Tense] (como na teoria de princípios e parâmetros), seja para checar este traço formal de INFL (como no minimalismo). O efeito deste movimento pode ser vislumbrado se colocamos no lugar da INFL [+Past] acima um auxiliar, como em (6) abaixo:



Note-se que, em (6), não há operação "reunindo" INFL com *give* porque ambas as categorias são manifestadas por palavras independentes e, por isso, podem permanecer em suas posições próprias na sintaxe de superfície. O mesmo não se pode dizer de (5), em que a informação temporal [+Past] aparece em INFL, sintaticamente autônoma em relação ao verbo. Mas, na frase superficial, isso não pode ser assim, pois a informação se manifesta no verbo, que se realiza como *gave* em (7).

Por isso, a teoria gerativa presume que algum tipo de operação fará com que [+Past], ao fim, seja unido ao verbo *give*. Especificamente, no modelo de princípios e parâmetros prévio ao surgimento do minimalismo, a ideia era a de que, porque a morfologia flexional do inglês é fraca, INFL "descia" até o verbo – operação conhecida como "Descida do Afixo" (do inglês *Affix Hopping*). Assim, na derivação da estrutura superficial de (5) acima, haveria dois movimentos: o do DP sujeito para Spec-de-IP, para receber Caso, e o de INFL [+Past] para V:



Numa análise minimalista, [+Past] seria um traço de INFL que deveria ser checado por V; mas, sendo um "traço fraco", não seria checado antes de *spell-out* e, por isso, não haveria "movimento visível" do verbo para INFL. (Ver nota 17 acima.)

As operações envolvendo a relação entre verbo e INFL adquiriram importância significativa na sintaxe gerativa a partir do trabalho clássico de Pollock (1989), que as utilizou para dar conta de diferenças sistemáticas entre a ordem de palavras do inglês e do francês. Surgiu o chamado "parâmetro de movimento do verbo", pelo qual as propriedades morfológicas ou formais de INFL seriam responsáveis pela variação paramétrica entre línguas – e, por isso, pertinentes à aquisição da linguagem.

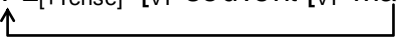
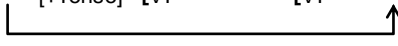
Resumidamente, uma das diferenças básicas entre o francês e o inglês pode residir no fato de que o francês permite que advérbios apareçam entre seu verbo flexionado e seu objeto direto, enquanto que isso não é possível em inglês. Os exemplos clássicos de Pollock são reproduzidos abaixo:

(8) Francês:

- a) **Jean souvent mange des pommes.*
- b) *Jean mange souvent des pommes.*

- (9) Inglês:
 a) *John often eats apples.*
 b) **John eats often apples.*

Na versão de Pollock, as duas línguas compartilhariam a sintaxe básica da frase, e a diferença fundamental entre elas seria, esquematicamente, a seguinte: (a) em francês, porque INFL [+Tense] é um "traço forte", há movimento do V para INFL, como em (10); (b) em inglês, INFL [+Tense] é "fraco" e, por isso, há descida do afixo, como em (11):

- (10) [_{IP} Jean [_{I'} INFL_[+Tense] [_{VP} souvent [_{VP} mange des pommes]]]]

 (11) [_{IP} John [_{I'} INFL_[+Tense] [_{VP} often [_{VP} eats apples]]]]


Na versão minimalista da análise, a diferença entre o francês e o inglês seria o ponto em que checariam o traço [+Tense] de INFL: (a) em francês, porque este traço formal é "forte", tem de ser checado antes de *spell-out* e, por isso, o movimento do verbo é visível; (b) em inglês, porque o traço é "fraco", não precisa ser checado antes de *spell-out*; por isso, só será checado em FL e o movimento (idêntico ao do francês) só se dará após *spell-out* – razão pela qual não é visível.

Neste ponto, é importante ressaltar: "morfemas gramaticais" são elementos que compartilham propriedades com palavras gramaticais – que também são núcleos funcionais – como também participam intensamente da sintaxe da frase. Nesta perspectiva, portanto, espera-se que seu desenvolvimento no processo de aquisição – seja em L1, seja em L2 – interaja de modo complexo tanto com o sistema morfológico quanto com o sistema sintático das línguas.

2.2 ESTUDOS SOBRE A ORDEM DE AQUISIÇÃO DOS MORFEMAS²⁰

Nesta seção, discutiremos alguns dos principais estudos sobre a ordem de aquisição dos morfemas. Primeiramente, apresentaremos algumas pesquisas realizadas sobre o assunto em L1. Como veremos, a pesquisa de Brown (1973) é uma referência importante para o estudo dos morfemas, pois é a partir dela que se postula a possibilidade da existência de uma ordem natural de aquisição destas estruturas. Em função disso, outras investigações foram feitas, tanto em inglês como L1 como em inglês como L2, e várias semelhanças (e também algumas diferenças) foram encontradas entre os dois processos.

Como vimos anteriormente, morfemas são tradicionalmente entendidos como unidades mínimas de significado, que podem ser lexicais ou gramaticais: a forma *played*, por exemplo, contém uma base lexical (*play*) e um morfema gramatical (*ed*). Tais morfemas gramaticais são o principal foco dos estudos dos morfemas.

Os estudos sobre a aquisição de morfemas, iniciados na década de 70 com os estudos de Brown (1973) e de De Villiers and De Villiers (1973) em L1, e com o trabalho de Dulay and Burt (1974) sobre L2, tinham a premissa básica de que deve haver uma ordem consistente na qual os aprendizes de L1/L2 adquirem proficiência no uso de morfemas. Conhecidos também como “*Natural Order Studies*”, tais trabalhos foram extremamente relevantes às pesquisas na área de aquisição de inglês (como L1 e L2), principalmente por terem identificado uma ordem mais ou menos previsível na aquisição de quatorze principais morfemas do inglês.

A principal motivação para o desenvolvimento destes estudos era a de demonstrar semelhanças nos padrões evolutivos na aquisição de morfemas por sujeitos em diferentes ambientes de aquisição/aprendizagem. Com isso, buscava-se fortalecer as evidências a favor de uma teoria inatista (gerativa) e contra os modelos behavioristas dos anos 50 baseados nas ideias de Bloomfield (1933) e Lado (1957).

Desta maneira, os estudos de aquisição de morfemas se tornaram parte da “*Natural Order Hypothesis*”, proposta sugerida principalmente por Stephen Krashen (1977), cuja ideia era de estabelecer uma ordem de surgimento das construções gramaticais que pudesse ser previsível nos processos de aquisição/aprendizagem.

²⁰ A partir do termo convencional em inglês, *morpheme studies*.

Segundo Kwon (2005), que apresenta uma revisão histórica dos estudos sobre a aquisição dos morfemas, as pesquisas conduzidas nos anos de 1970 preocupavam-se mais em estabelecer uma sequência na aquisição de morfemas, ao passo que, nos anos de 1980, a preocupação maior era com a identificação de variáveis que pudessem ser determinantes desta sequência. Já a partir dos anos de 1990 as pesquisas tenderam a revisitar os dados já obtidos a fim de explicá-los a partir do desenvolvimento de teorias linguísticas (GASS & SELINKER, 1994; ZOBL & LICERAS, 1994; GOLDSHNEIDER & DEKEYSER, 2005).

Vejamos a seguir uma discussão mais detalhada de algumas das principais pesquisas sobre aquisição de morfemas em L1 e L2.

2.2.1 Estudos em L1

Conforme Kwon (2005), a história dos estudos sobre a aquisição dos morfemas gramaticais pode ser entendida a partir da pesquisa de Brown (1973), que mostra dados longitudinais da fala espontânea de três crianças aprendendo inglês americano como L1. Os resultados revelam que as crianças adquiriram os morfemas estudados praticamente na mesma ordem, embora não na mesma idade. Vejamos uma síntese dos resultados de Brown:

Tabela 1: Ordem de aquisição de morfemas do inglês em Brown (1973)²¹

ORDEM	MORFEMA
1	<i>Present progressive (ing)</i> ²²
2/3	<i>in, on</i>
4	<i>Plural</i>
5	<i>Past irregular</i>
6	<i>Possessive ('s)</i>
7	<i>Uncontracted Copula (am, is, are)</i>
8	<i>Articles</i>
9	<i>Past regular</i>
10	<i>Third person sing (-s)</i>
11	<i>Third person sing irregular</i>
12	<i>Uncontracted auxiliary (am, is, are)</i>
13	<i>Contracted copula</i>
14	<i>Contracted auxiliary</i>

²¹ Versão de tabela cf. Kwon (2005).

²² Utilizaremos tanto os nomes dos morfemas em inglês quanto a tradução destes nomes.

A pesquisa de Brown é até hoje referência nos estudos sobre a aquisição de morfemas, não só por seus resultados, mas também por ter sido a primeira a enfatizar as questões que se colocam com a descoberta de uma dada ordem. A principal delas, de acordo com Brown, é a seguinte: “*the constancy is there. Some factor or some set of factors caused these grammatical morphemes to evolve in an approximately consistent order in these children*” (1973, p. 272).

Em busca de uma resposta para esta questão, Brown analisa o possível papel de fatores como frequência, complexidade gramatical e complexidade semântica no condicionamento da ordem descoberta.

Com relação ao fator frequência, Brown compara a frequência das estruturas produzidas pelos três sujeitos das pesquisas com a encontrada na produção dos pais²³. O autor conclui que frequência não é uma variável significativa: “*In sum, there is no clear evidence at all that parental frequencies influence the order of development of the form we have studied. I am prepared to conclude that frequency is not a significant variable*” (BROWN, 1973, p. 368).

Já quanto à Complexidade Semântica e à Complexidade Gramatical, Brown analisa seus dados à luz de teorias linguísticas transformacionais da época, concluindo que haveria, sim, evidência de que estes fatores são determinantes na ordem de aquisição. Mas ele também afirma que tal evidência era um tanto prematura e que mais pesquisas sobre o assunto seriam necessárias no futuro.

Especificamente, Brown entende que a questão da complexidade, seja ela semântica ou sintática, está ligada às representações formais de um determinado conhecimento. Assim, o grau de complexidade das representações formais associadas a uma expressão fixará o grau de complexidade da própria expressão. A fim de determinar tais representações, Brown segue Jacobs & Rosebaum (1968), Chomsky (1965) e Katz & Postal (1964) e descreve em detalhe os 14 morfemas selecionados: suas características gramaticais e semânticas em termos de traços como [+/- nominal], [+/- singular], [+/- comum], [+/- concreto], [+/- animado], [+/- progressivo], etc, e também em termos do modo como determinam a aquisição de regras de transformação.

Com base nessa análise, Brown conclui: “*There is an approximately invariant order of acquisition for the 14 morphemes we have studied, and behind this*

²³ Em seu estudo de referência, Brown (1973) também registrou e analisou a produção dos pais na interação com a criança, possibilitando, assim, a comparação de estruturas.

invariance lies not modeling frequency but semantic and grammatical complexity” (BROWN, 1973, p. 379).²⁴

O trabalho de Brown também é importante por estabelecer certos padrões metodológicos. Entre eles, está o conhecido critério “SOC” (*Suppliance in Obligatory Context*), que é inclusive utilizado nesta tese.²⁵ Segundo Brown:

Grammatical morphemes are obligatory in certain contexts, and so one can set an acquisition criterion not simply in terms of output, but in terms of output-where-required. Each obligatory context can be regarded as a kind of test item which the child passes by supplying the required morpheme or fails by supplying none or one that is not correct (1973, p. 255).

Desta maneira, Brown enfatiza que, para a análise de um determinado morfema, faz-se necessário levarmos em conta não só onde ele é obrigatório e corretamente empregado, mas ainda onde ele seria esperado e não é empregado. Isso fornece uma maneira de mensurar/quantificar a performance dos aprendizes que contempla os contextos onde os morfemas seriam esperados em linguagem espontânea.

A principal crítica ao estudo de Brown é que sua pesquisa se limita a estudar dados de apenas três crianças. Então, a fim de responder a este mesmo tipo de questionamento, De Villiers & De Villiers (1973) desenvolveram um estudo transversal com a fala espontânea de 21 crianças no qual investigam os quatorze morfemas de Brown. Os resultados são bastante semelhantes aos de Brown, como podemos ver no quadro de síntese a seguir:

Tabela 2: Ordem de aquisição dos morfemas em De Villiers & De Villiers (1973)

MORFEMA
1. <i>Present progressive (ing)</i>
2. <i>Plural</i>
3. <i>Past irregular</i>
4. <i>Possessive ('s)</i>
5. <i>Contracted Copula (am, is, are)</i>
6. <i>Past regular</i>
7. <i>Third person sing (-s)</i>
8. <i>Uncontracted Cop</i>
9. <i>Contracted auxiliary (am, is, are)</i>
10. <i>Uncontracted auxiliary</i>

²⁴ No capítulo 3, veremos que Goldschneider & DeKeyser (2005) apontam igualmente frequência, complexidade semântica e complexidade sintática, dentre outros, como fatores determinantes para uma suposta ordem de morfemas.

²⁵ No capítulo 4, descreveremos com mais detalhes como aplicamos o critério de análise utilizado por Brown na metodologia de análise do BELC.

As principais semelhanças são as seguintes:

1. O *-ing* é o primeiro morfema a ser adquirido, segundo os dois estudos.
2. O plural vai em segundo lugar em De Villiers & De Villiers, e em terceiro em Brown.
3. O passado irregular vem em terceiro em De Villiers & De Villiers, e em quarto em Brown.
4. O possessivo *'s* vem em quarto em De Villiers & De Villiers, e em quinto em Brown.

Desconsiderando as preposições (aqui apresentadas somente na sequência de Brown), podemos dizer que as quatro primeiras posições são iguais.

Mas há também diferenças nos resultados dos dois estudos:

5. Em Brown, a cópula contraída é o penúltimo sufixo na ordem de aquisição, precedendo apenas o auxiliar contraído, ao passo que, em De Villiers & De Villiers, cópula contraída aparece em quinto na ordem, precedendo vários outros morfemas.
6. Em Brown, cópula não-contraída precede cópula contraída, passado irregular e *3rdP-sing*, e em De Villiers & De Villiers, cópula não-contraída vem depois destes morfemas.

Como veremos adiante, estes resultados, vistos à luz de estudos contemporâneos, apresentam ainda outros fatores importantes. Zobl & Liceras (1994) observam, por exemplo, que há algumas generalizações que sugerem uma explicação de natureza sintática.

1. A clara separação entre o *Aux Be* e o afixo *-ing*, já que a forma contínua se apresenta no topo e *Aux Be* – especialmente sua forma não-contraída, que é exemplo de Infl livre²⁶ – permanece em última posição.
2. *Copula Be* precede o *Aux Be*, embora as duas formas sejam idênticas e a análise corrente assuma que ambas se movam para Infl.
3. Artigo e Possessivo se apresentam juntos, o que sugere que ambos são manifestações de Det, um livre e outro preso.

²⁶ Isto é, que não precisaria se mover para operações de descida de afixo, por exemplo, como vimos em 2.1.5.

4. Passado regular e *3rdP-sing* se apresentam juntos no final da hierarquia, sendo ambos expoentes de Infl e sujeitos à descida de afixo (cf. seção 2.1.5).

Estas observações fazem parte de uma análise que tenta oferecer uma explicação para a semelhança na ordem dos morfemas a partir da teoria das categorias funcionais, que mencionamos na seção anterior. Veremos mais detalhes sobre isso no capítulo 3 e na próxima seção.

2.2.2 Estudos em L2

Segundo Kwon (2005), os primeiros estudos sobre aquisição de morfemas em L2 mostraram que, apesar de algumas pequenas variações, a ordem dos morfemas era muito semelhante entre aprendizes de inglês como L2. Mostraram, ainda, que a ordem em L2 era bastante similar àquela encontrada em L1 (BROWN, 1973 e DE VILLIERS & DE VILLIERS, 1973). Estes são os resultados de estudos pioneiros de Dulay & Burt (1974), que conduziram um estudo com crianças falantes de espanhol e chinês de Bailey et al. (1974), que estudaram adultos aprendizes de inglês como L2 falantes de espanhol e chinês como L1.

Outras pesquisas conduzidas nos anos 70 também chegaram a resultados semelhantes, independentemente de variáveis como L1 ou idade (LARSEN-FREEMAN, 1975), num estudo transversal com adultos, e Hanuka (1976), com falantes de japonês como L1. O ambiente no processo de aquisição/aprendizagem e também a questão da instrução foram investigados. Perkins & Larsen-Freeman (1975) desenvolveram um estudo comparando grupos com instrução (*time 1*) e sem instrução (*time 2*) e os resultados mostraram pouca diferença. Pica (1983) conduziu um estudo bastante interessante com adultos, falantes de espanhol do México, divididos em três grupos: o primeiro grupo era de aprendizes num ambiente formal de sala de aula, o segundo, um grupo que vivia num ambiente onde o inglês era L1 e não tinha instrução, e o terceiro, um grupo que tinha instrução e vivia num ambiente natural de fala do inglês. A conclusão foi que as diferentes condições de aquisição/aprendizagem não afetaram a ordem na aquisição dos morfemas de

maneira significativa. De fato, há muitas semelhanças nos resultados encontrados em diferentes tipos de pesquisa. Considere a tabela abaixo, que resume os resultados de algumas destas pesquisas:

Tabela 3: Ordens da aquisição de inglês como L2²⁷

Krashen (1977)	Andersen (1978)	Bailey et al. (1974)	Larsen-Freman (1975) – Time 1	Larsen-Freman (1975)- Time 2	Dulay & Burt (1973)
<i>V-ING</i>	<i>COP</i>	<i>V-ING/COP</i>	<i>V-ING</i>	<i>COP</i>	<i>PLU</i>
<i>PLU/COP</i>	<i>ART</i>	<i>ART/AUX</i>	<i>COP</i>	<i>V-ING</i>	<i>V-ING</i>
<i>AUX/ART</i>	<i>AUX/V-ING</i>	<i>PLU</i>	<i>ART</i>	<i>AUX</i>	<i>COP</i>
<i>IRREG PAST</i>	<i>PLU</i>	<i>REG PAST</i>	<i>AUX</i>	<i>ART</i>	<i>AUX</i>
<i>REG PAST/3rdP-S</i>	<i>IRREG PAST</i>	<i>3rdP-S</i>	<i>PLU</i>	<i>PLU</i>	<i>ART</i>
<i>POSS</i>	<i>POSS</i>	<i>POSS</i>	<i>REG PAST</i>	<i>POSS</i>	<i>IRREG PAST</i>
	<i>REG PAST</i>		<i>3rdPS</i>	<i>IRREG PAST</i>	<i>3rdPS</i>
	<i>3rdP-S</i>		<i>IRREG PAST</i>	<i>3rdPS</i>	<i>POSS</i>
			<i>POSS</i>	<i>REG PAST</i>	

Como se vê, em todas as pesquisas, o sufixo *-ING* e o *Cop Be* estão entre os primeiros morfemas gramaticais a serem aprendidos por diferentes aprendizes de inglês como L2; e o *Poss-’s*, o *’s* de *3rdP Sing* e o passado regular entre os últimos.

Entre os resultados acima sintatizados, salientamos a ordem dada por Krashen (1977), que tem sido a mais utilizada como referência para pesquisas.

A importância de seus estudos se deve ao fato de que ele organizou várias pesquisas feitas até então (BROWN, 1973; DE VILLIERS & DE VILLIERS, 1973; HANUKA, 1974; ROSANSKY, 1976; DULAY & BURT, 1974; LARSEN-FREEMAN, 1975; ANDERSEN, 1977; entre outras). Por outro lado, é importante observar que tais pesquisas diferem essencialmente em suas propostas metodológicas: algumas são estudos com crianças, outras com adultos; algumas são longitudinais, outras transversais; há pesquisas que envolvem ambientes com instrução, outras estudam ambientes naturais de aquisição. Assim, a síntese de Krashen pode ter a vantagem de ser baseada num amplo conjunto de resultados; ao mesmo tempo, pode ter eliminado pequenas diferenças que poderiam ser significativas. Veremos, neste trabalho, que isso é possível.

²⁷ Tabela elaborada por mim a partir de Zobl e Liceras (1994) e Izumi e Hisahara (2004).

Ao fazer seu estudo, Krashen propunha-se a verificar a possível sequência de morfemas a partir de estudos variados a fim de analisar algumas questões relativas ao seu modelo do “monitor”.²⁸

Kwon (2005) faz a importante observação de que, mesmo que Krashen tenha encontrado as semelhanças em resultados de diferentes tipos de estudo, as questões metodológicas devem ser levadas em consideração. Ou seja, diferenças de método podem levar a resultados diferentes. Por exemplo, Kwon observa que as tarefas orais são as mais frequentemente utilizadas nas pesquisas e que o instrumento mais utilizado foi o BSM (*Bilingual Syntax Measure*), um teste sintático originalmente elaborado para elicitado dados de crianças, através de figuras que apresentam contextos em que a criança seja estimulada a produzir uma determinada estrutura linguística.

Ainda como exemplo, as pesquisas de Dulay & Burt (1973) e Bailey, Madden & Krashen (1974), cujos resultados mencionamos acima, utilizaram este método, considerado muito consistente para calcular ocorrências. Várias críticas foram feitas ao BSM, entre elas a de que o método não foi originalmente elaborado para determinar a ordem de aquisição dos morfemas, além de outras críticas próprias de métodos de elicitado direta de dados. Por outro lado, as pesquisas de Andersen (1977), Larsen-Freeman (1975) e Krashen (1977) não utilizaram o BSM e obtiveram resultados parecidos às que usaram o método. Isso argumentaria a favor da unificação dos resultados feita por Krashen (1977).

Ainda, Krashen *et al.* (1978) utilizaram tarefas escritas a partir de textos livres, com e sem restrições de tempo, e também obtiveram resultados semelhantes às pesquisas que usaram tarefas orais.

Mas nem sempre essa coerência de resultados, mesmo com diferentes métodos, foi confirmada. Por exemplo, Larsen-Freeman (1975) desenvolveu um estudo que utilizava vários tipos de tarefas de elicitado de dados, tarefas escritas e orais: o BSM, um teste de repetição de frases baseados em figuras, um teste de *listening*, um teste de interpretação de leitura através de um *cloze* de múltipla escolha e um teste de escrita com o preenchimento de lacunas. Os resultados obtidos a partir das tarefas orais foram, no geral, parecidos novamente com os de

²⁸ De modo geral, este modelo prevê que os adultos “adquirem” (subconscientemente) alguns aspectos da L2 e aprendem (conscientemente) outros. Assim, a aprendizagem estaria disponível apenas como um “monitor”, ou melhor, um “editor” do *output* do sistema adquirido (KRASHEN, 1977; KRASHEN, 1985)

Dulay & Burt (1974). Mas os resultados das tarefas escritas mostraram diferenças, como a aquisição do morfema *plural -s* e *3rdPS* antes do que outros. Estas diferenças foram atribuídas por Ellis (1994) a questões inerentes às condições de produção da tarefa escrita, que tem influências sócio e psicolinguísticas.

Assim, como previne Kwon (2005), uma questão importante a ser levada em conta na discussão de diferenças com relação às ordens de aquisição encontradas é o método de coleta de dados: métodos diferentes podem levar a resultados diferentes.

Outra consideração importante pode ser ilustrada pelo trabalho de Izumi & Isahara (2004). Numa pesquisa muito semelhante à que é proposta nesta tese, os autores investigaram a ordem de aquisição de morfemas em inglês como L2 por falantes de japonês como L1 utilizando para isso o corpus de aprendiz NICT JLE (cuja descrição mais detalhada veremos no capítulo 4). Neste estudo, a contribuição mais relevante parece ser com relação à importância do papel desempenhado pela L1 no processo de aquisição de inglês como L2. Os autores propõem duas hipóteses para o estudo: a primeira sustenta que os morfemas são aprendidos por todos os aprendizes numa sequência basicamente comum, não interessando L1, idade, ambiente de aquisição, etc; a segunda atribui diferenças nas sequências a diferenças entre a L1 e a L2. Foi utilizado o chamado *Group Score Method* (GSM) como método para calcular os resultados, a partir de Dulay & Burt (1973), estudo que, é importante lembrar, analisa dados de aprendizes espanhóis e chineses. Os resultados são sintetizados abaixo:

Tabela 4: Ordem de aquisição de Izumi & Isahara (2004)

1.	<i>Possessive's</i>
2.	<i>Progressive – ING</i>
3.	<i>Copula BE</i>
4.	<i>Third person present singular tense –s</i>
5.	<i>Plural-s</i>
6.	<i>Auxiliary BE</i>
7.	<i>Irregular past tense</i>
8.	<i>Article</i>

Há diferenças na ordem geral obtida por Izumi & Isahara com relação à ordem obtida por Dulay & Burt (1973) conforme tabela 3 acima. Entre elas:

- a. Os artigos estão em oitavo lugar na ordem de aquisição estabelecida por Izumi & Isahara, ao passo que, em Dulay & Burt (1973), estão em quinto, muito abaixo dos verbos auxiliares no passado.
- b. O possessivo 's está na primeira posição, enquanto que, em Dulay & Burt (1973), está na oitava.
- c. O plural está em quinto; em Dulay & Burt (1973), em primeiro.

Izumi e Isahara (2004) atribuem estas diferenças a dois fatores. O primeiro é a interferência da L1, claramente evidenciada pela ordem de aquisição de dois morfemas: o morfema “artigo”, e isso é explicado pelo fato de que o japonês não possui artigos – e este morfema foi o último na sequência. Além disso, o morfema Poss 's, foi o primeiro na ordem obtida. Segundo os autores, o uso muito frequente da forma “A's B” deve ser oriundo de uma tradução direta da expressão “A no B” – em japonês, “no” é uma partícula possessiva. O segundo fator seria o método, que não considera qualquer tipo de erro (principalmente os erros de supergeneralização), concluindo que a L1 pode afetar a ordem de aquisição dos morfemas.

Podemos citar ainda outro estudo que contempla um corpus de aprendizes de inglês falantes de japonês como L1: o estudo de McEnery, Xiao & Tono (2006). Neste estudo, os autores propõem uma revisão dos estudos dos morfemas a partir da aplicação de métodos da linguística de corpus aos corpora de aprendizes.²⁹ McEnery, Xiao & Tono (2006) acreditam que os estudos dos morfemas são um bom ponto de partida se quisermos saber o quão eficiente os corpora de aprendizes podem ser na tarefa de descrever a língua do aprendiz.

O corpus utilizado na pesquisa é o Longman Learner's Corpus,³⁰ e o subcorpus selecionado contém dados de redações escritas de falantes nativos de japonês (num total de aproximadamente um milhão de palavras). O estudo procura verificar a afirmação de que os aprendizes de L2 adquirem morfemas gramaticais numa ordem universal usando dados de corpus de aprendiz.

Metodologicamente, os autores propõem uma tentativa de reanalisar os resultados obtidos nos estudos dos morfemas através da utilização de ferramentas modernas de organização e análise de dados, como o WordSmith Tools, o CLAWS

²⁹ O estudo é particularmente interessante pela proposta metodológica. A esse respeito, veremos mais detalhes no capítulo 4.

³⁰ Veremos maiores informações sobre esse corpus e sobre pesquisa em corpora de aprendizes em geral no capítulo 4 desta tese.

(*Constituent Likelihood Automatic Word-Tagging System*), etc., muito conhecidas em linguística de corpus (que veremos mais detalhadamente no capítulo 4). O estudo envolve um grande trabalho manual de etiquetagem, pois é preciso encontrar todos os exemplos corretos e incorretos de morfemas relevantes nos dados dos aprendizes. O procedimento de etiquetagem seguiu os seguintes passos:

1. O corpus original foi pré-processado de maneira que as informações possam ser convertidas num formato apropriado para ser utilizado no WordSmith (versão 4).³¹
2. Todo o subcorpus foi etiquetado utilizando um etiquetador POS (*part-of-speech*, parte ou categoria do discurso)
3. Os erros morfológicos foram manualmente etiquetados utilizando a informação POS como um guia.

A porcentagem de correção de uso dos morfemas individuais foi obtida a partir dos números das etiquetas de erros e das etiquetas das POS relevantes. Os resultados são comparados com resultados de pesquisas anteriores. Para poder compará-los, no entanto, os autores etiquetaram os dados da mesma maneira que Dulay & Burt (1973), isto é, utilizando o BSM (*Bilingual Syntax Measure*), que analisa somente o contexto obrigatório que pede o uso do morfema. Finalmente, são considerados adquiridos os morfemas que obtêm 90% de porcentagem de correção.

Já os aprendizes foram divididos em quatro grupos, segundo seu nível de proficiência: *Elementary*, *Intermediate*, *Advanced* e *Proficient*. A ordem de aquisição dos morfemas é determinada a partir desses níveis.

Os resultados obtidos por McEnery, Xiao & Tono (2006) mostram a seguinte ordem: *Copula BE*, *Auxiliary BE*, *Possessive 's*, *Progressive -ING*, *Plural -S*, *3rd person -s*, *Irregular Past*, *Article*. Comparando com os resultados de Dulay & Burt (1973) – apresentados na Tabela 3 acima – observam-se algumas diferenças. A mais notável delas, segundo McEnery, Xiao & Tono, é quanto à posição do artigo, que, assim como as outras diferenças observadas, pode ser atribuídas à

³¹ Para obter informações técnicas mais específicas sobre como esta conversão foi feita, ver McEnery, Xiao & Tono (2006). Embora o texto apresente instruções sobre como proceder neste tipo de proposta de pesquisa, não julgamos este tipo de detalhe relevante para a discussão nesse momento.

interferência do japonês como L1. Assim, os resultados deste estudo parecem confirmar os resultados apresentados por Izumi & Isahara (2004).

As diferenças observadas nos resultados apontam, segundo McEney, Xiao & Tono (2006), para duas possibilidades: (i) que o conceito de uma ordem universal deve ser reconsiderado, necessitando de estudos paralelos que forneçam resultados robustos ou (ii) que os dados do corpus de aprendiz devem ser reexaminados, pois inúmeras variáveis malclassificadas podem interferir nos resultados. Por exemplo, no estudo de McEney, Xiao & Tono (2006), a classificação da proficiência dos aprendizes pode ser questionada, pois já estava dada no corpus utilizado que, por sua vez, não foi elaborado especificamente para os fins da pesquisa que eles propuseram.

Por fim, mencionamos mais um estudo recente que confirma dois aspectos mencionados na discussão dos trabalhos que já apresentamos: a importância dos resultados obtidos na ordem de aquisição de morfemas e a importância da L1. Trata-se do estudo de Wagner (2007), que investiga a aquisição de quatro morfemas em inglês como língua estrangeira,³² a saber: *Plural*, *Present Progressive*, *3rdP-Sing* e *Possessive -'s*.

O estudo se baseia numa pesquisa transversal envolvendo quatorze alunos franceses aprendendo inglês como LE como parte do seu currículo do ensino médio. Estes alunos tinham entre treze e quatorze anos e estavam aprendendo inglês há pelo menos três anos. Foram coletados dados de produção oral e escrita elicitados através de tarefas orientadas. A autora propõe como objetivos investigar: (i) até que ponto a interferência interlinguística ajuda ou atrapalha a aquisição destes morfemas e (ii) as implicações de ensinar morfemas gramaticais, ou gramática em geral, numa aula de língua estrangeira.

A autora utilizou os dados de Dulay & Burt (1974) como referência e obteve os seguintes resultados:

³² Cf. a sigla usada em inglês, EFL, que representa *English as a Foreign Language*, ou inglês como língua estrangeira. Um dos objetivos da autora é saber se os resultados de sua pesquisa serão iguais aos resultados da maioria das pesquisas sobre a aquisição de morfemas em inglês como segunda língua, isto é, ESL (*English as a Second Language*).

Tabela 5: Comparação de Dulay & Burt (1974) e Wagner (2007)

DULAY & BURT (1974)	WAGNER (2007)
<i>Plural</i>	<i>Possessive</i>
<i>Present Progressive</i>	<i>Plural</i>
<i>3rdP- sing</i>	<i>Present Progressive</i>
<i>Possessive</i>	<i>3rdP-sing</i>

Uma das diferenças entre os dois estudos é que, em Wagner (2007), o Possessivo ocupa a primeira posição na ordem de aquisição, ao passo que, em Dulay & Burt (1974), ocupa a última. Wagner atribui isso principalmente ao tipo de tarefa na coleta dos dados, que era o preenchimento de lacunas. Segundo Wagner (2007), isso pode ter facilitado a produção do *Poss-s* por parte dos alunos.

Com relação à interferência linguística, Wagner nos diz:

This negative transfer of grammatical properties between the two languages exhibited by the students in this study indicates that interlanguage interference does occur, though the study cannot accurately reflect the extent to which it occurs. The results of this study cannot be definitely used to determine if interference occurs more or less often than researchers had previously thought (p. 31).

A autora afirma isso a partir da constatação que o *Plural –s* é semelhante à estrutura do francês, e que, por isso, seria uma transferência positiva, assim como o *Present Progressive –ING* que, por não existir exatamente com a mesma forma em francês, seria uma estrutura saliente para o aprendiz. Wagner conclui que, apesar da posição do morfema *Poss-s* se mostrar diferente do que em Dulay & Burt (1974), os outros morfemas evidenciam a mesma sequência e, por isso, podemos dizer que alunos de inglês como LE apresentam a mesma ordem de aquisição que aprendizes de inglês como L2, e que seus dados apontam ainda para interferência interlinguística.

Poderíamos descrever várias outras pesquisas sobre o assunto³³. No entanto, acreditamos que os estudos que mencionamos sejam suficientes para

³³ Citaremos brevemente duas outras pesquisas a fim de registrarmos observações sobre estes estudos, que são bastante mencionados na literatura sobre o assunto. Andersen (1978) conduz um estudo cujo foco é metodológico, propondo uma escala implicacional para analisar a possibilidade de uma ordem de morfemas em inglês como L2 a partir de diferentes L1s. Andersen define a escala implicacional como a possibilidade de que a ocorrência de determinadas estruturas possa implicar outras, e oferece um modelo implicacional que tenta contemplar variação individual e sistematicidade na aquisição e no uso de uma segunda língua, pois uma das críticas aos estudos da época era a de que as pesquisas eram em sua grande maioria transversais e não longitudinais (HANUKA, (1976), ROSANSKY, (1976)). O autor pesquisou o uso de treze (13) morfemas em

termos um panorama das principais investigações sobre a aquisição de morfemas em inglês como L2 que possam ser relevantes para esta tese.

2.3 RESUMO

Este capítulo apresenta uma revisão teórica das principais questões que serão discutidas nesta tese.

Vimos questões teóricas associadas a aspectos gramaticais deste estudo, em particular uma discussão da noção de “morfema gramatical” e sua relação com a teoria gramatical contemporânea.

Na segunda seção, pudemos ver algumas pesquisas relevantes no estudo da aquisição de morfemas em inglês, especialmente porque procuraram determinar a ordem de aquisição destas formas (BROWN, 1973; DE VILLIERS & DE VILLIERS, 1973 – inglês como L1, e DULAY & BURT, 1973; KRASHEN, 1977, ISUMI & ISAHARA, 2004).

A partir disso, podemos estabelecer algumas considerações de caráter mais particular para esta tese. Podemos propor algumas perguntas:

- A ordem na aquisição dos morfemas em inglês como L2 por parte de aprendizes cuja L1 seja o português brasileiro confirmaria, em linhas gerais, a ordem observada em estudos de referência como Krashen (1977)?

oitenta e nove (89) aprendizes falantes de espanhol como L1. Ao revisar este estudo, podemos notar que, apesar de tentativas de explicação já um pouco ultrapassadas em termos de teoria linguística, o autor faz colocações que parecem prever o advento teórico conforme colocam Zobl e Licerias (1994): na análise, Andersen divide os morfemas em (i) V e NP e (ii) livres e presos, afirmando que a ordem parece resultar de operações simultâneas de escalas subjacentes gramaticalmente relevantes (da tradução livre de “*simultaneous operation of simple grammatically-relevant underlying scales*”, em Andersen (1977), p.221). O autor nos diz que talvez a distinção entre V/NP seja uma distinção real sobre como o aprendiz processa a língua: enquanto o aprendiz está construindo a consciência de um aspecto nominal como Plural, ele provavelmente está também armazenando informações do mesmo sintagma nominal sobre o uso de artigos. Ainda, o texto de Rosansky (1976) também tem um enfoque metodológico e uma crítica a estudos cujos resultados não sejam de pesquisas longitudinais, tanto que Andersen salienta “*Since my study is not a longitudinal study, it is not appropriate to speak of AUX-ING and Pal-PaR as being acquired together, but the analogy seems valid, nonetheless*” (ANDERSEN, 1978, p. 248).

- A fim de tentar identificar uma possível sequência na aquisição de morfemas em inglês como L2 a partir destes aprendizes, qual seria o melhor método de coleta: um método de produção estimulada ou um método de produção espontânea?

- Se optarmos por um método de estudo que contemple a produção espontânea dos aprendizes e se encontrarmos diferenças no processo de aquisição que estamos nos propondo a analisar, estas diferenças poderiam ser atribuídas ao método de estudo?

- E se pudermos observar que há, efetivamente, uma ordem natural, quais as possíveis razões para isso? Ou seja, fatores que supostamente explicariam uma dada ordem natural na aquisição de morfemas em inglês como L2 poderiam igualmente explicar a ordem encontrada em nosso estudo?

Sendo assim, passemos aos capítulos seguintes, que analisarão algumas respostas a estas perguntas.

3 ANÁLISE DA AQUISIÇÃO DA MORFOSSINTAXE EM L2

Apresentamos, neste capítulo, uma discussão teórica mais específica, concentrando-nos basicamente em duas abordagens particulares para a análise de aquisição de morfemas gramaticais em L2 em inglês. Nosso intuito aqui é apresentar, de modo detalhado, as duas abordagens que estaremos comparando em nosso estudo empírico.

Na primeira seção, fazemos uma breve revisão das principais críticas tecidas aos estudos tradicionais da aquisição dos morfemas. Em seguida, apresentamos as duas propostas de análise que nortearão esta tese: a primeira, de Goldschneider & DeKeyser (2005), pode ser definida como uma abordagem “multifatorial”, ou seja, uma proposta que considera vários fatores diferentes para explicar a ordem de morfemas no processo de aquisição de L2. Já a segunda proposta de análise é sugerida por Zobl & Licerias (1994) e concentra-se essencialmente na estrutura gramatical associada aos morfemas, em particular na interação entre propriedades lexicais dos morfemas gramaticais (se morfemas livres ou presos, se são de categorias Det, Infl, etc.) e propriedades de arquitetura sintática da frase. Na última seção, levantamos as questões específicas que serão investigadas nesta tese.

3.1 CRÍTICAS AOS ESTUDOS TRADICIONAIS DE AQUISIÇÃO DE MORFEMAS EM L2

Pudemos ver que os resultados das pesquisas com relação à ordem dos morfemas nos processos de aquisição de inglês na condição de L1 e L2 não são idênticos, mas são bastante semelhantes. Naturalmente, várias críticas foram feitas, algumas das quais já mencionamos no capítulo anterior. Kwon (2005), por exemplo, aponta as seguintes:

a. Alguns dos resultados obtidos podem ser artefatos do *Bilingual Syntax Measure* (BSM) ou dos demais instrumentos ou métodos de coleta de dados.

- b. Os morfemas estudados pertencem a uma variedade muito grande de fenômenos gramaticais (ZOBL & LICERAS, 1994).
- c. Os morfemas representam uma parte muito pequena de qualquer língua para servir de base a conclusões gerais (LARSEN-FREEMAN & LONG, 1991).
- d. Os morfemas com significados diferentes não deveriam ser agrupados como se dissessem respeito a uma única estrutura gramatical.
- e. As variações individuais são ocultadas por dados do grupo.
- f. O quadro geral do uso de determinada forma por parte do aprendiz não é levado em consideração.
- g. A ordem de correção não pode ser equiparada com a ordem de aquisição.
- h. O método de cálculo não considera o uso superestendido (*oversuppliance*) de morfemas em contextos não obrigatórios.
- i. É difícil generalizar os resultados porque praticamente só o inglês foi investigado (COOK, 1993).
- j. Os estudos são quase todos transversais.

Tais críticas são feitas não somente à descrição oferecida pelas pesquisas realizadas sobre o assunto, mas também às tentativas de explicação que algumas delas oferecem. E, apesar das críticas, Larsen-Freeman & Long (1991) concluem que:

...the morpheme studies provide strong evidence that IL³⁴ exhibit common accuracy/ acquisition orders. Contrary to what some critics have alleged, there are in our view too many studies conducted with sufficient methodological rigour and showing sufficiently consistent general findings for the commonalities to be ignored. As the hunter puts it, there is something moving in the bushes (LARSEN-FREEMAN & LONG, 1991, p. 92).

Desta maneira, mesmo considerando as diferenças encontradas nas pesquisas, relacionadas à idade, instrução, metodologia, etc., a relativa coerência nos dados levou pesquisadores a admitirem que a ordem de aquisição dos morfemas é um fenômeno que merece atenção, e, assim, é preciso investigar os fatores que possam ser pertinentes.

³⁴ *Interlanguage*, ou “interlíngua”, cf. lista de abreviaturas.

3.2 EXPLICAÇÕES PARA A ORDEM DE AQUISIÇÃO DOS MORFEMAS

São vários os fatores considerados determinantes para a ordem de aquisição dos morfemas. Goldschneider & DeKeyser (2005) afirmam que a transferência da L1 é o mais comumente citado, seguido do fator frequência.

Em sua revisão histórica sobre os estudos dos morfemas, Kwon (2005) analisa alguns fatores determinantes para a explicação das semelhanças encontradas na ordem de aquisição dessas estruturas. Segundo esta autora, tais fatores são conhecidos como “*putative determinants*”, ou determinantes hipotéticos. A autora analisa três de maneira mais detalhada, a saber, complexidade semântica, frequência no *input* e transferência da L1, por considerá-los mais importantes.³⁵

Entre os fatores mais frequentemente mencionados em estudos de aquisição de L1, especialmente na discussão sobre a pesquisa de Brown (1973), estão a saliência (a diferença na percepção do morfema *ing* em *talking*, por exemplo, comparada à percepção do morfema *ed* em *talked*: o primeiro é, sabidamente, mais saliente que o segundo); silabidade (se um morfema é silábico); falta de exceção (o possessivo *'s*, por exemplo, nunca tem exceção, ao passo que os verbos no passado *ed* tem exceções no passado irregular); e outras complexidades de ordem semântica ou sintática.

Em L2, os fatores também são variados e incluem, por exemplo, saliência, regularidade morfofonológica, complexidade sintática, frequência, complexidade semântica, transferência da L1 e variação individual.

Todas as pesquisas sobre a aquisição da ordem de morfemas que mencionamos até aqui postulam algum ou alguns destes fatores como possíveis determinantes dessa ordem. Para os fins desta tese, a discussão se concentrará nas propostas de Goldschneider & DeKeyser (2005) e Zobl & Liceras (1994), já que estes trabalhos adquiriram certa proeminência na discussão recente, conforme Ellis & Barkhuizen (2005): “*Researchers switched attention from description to explanation by focusing on factors that accounted for the accuracy order. Two key studies in this respect were Zobl & Liceras (1994) and Goldschneider & DeKeyser (2001)*” (p. 76).

³⁵ Na verdade, Kwon oferece uma revisão teórica a partir de pesquisas sobre os estudos dos morfemas que consideram esses fatores como determinantes para uma possível sequência no processo de aquisição.

3.2.1 Goldschneider & DeKeyser (2005)

Goldschneider & Dekeyser (2005) fazem uma meta-análise dos dados de outros estudos sobre a aquisição dos morfemas e procuram oferecer uma explicação multifatorial para os resultados. Os autores argumentam que, apesar do vasto número de pesquisas feitas sobre o assunto desde os anos de 1970 e da indiscutível contribuição descritiva delas, nenhuma conseguiu explicar satisfatoriamente a semelhança encontrada na ordem dos morfemas por meio de um único fator. De um modo geral, Goldschneider & Dekeyser (2005) sugerem que ou (i) a ordem de aquisição dos morfemas deve indicar algo inato ou (ii) ela pode ser explicada a partir da interação entre características inerentes dos morfemas. Afirmam ainda que, se tal semelhança puder ser explicada a partir de uma combinação de fatores, isso pode fornecer evidência para uma teoria de aquisição de segunda língua mais abrangente, eliminando a busca por uma explicação única que dê conta deste fenômeno.

Após revisarem várias pesquisas, os autores concluem que as explicações oferecidas para uma possível ordem na aquisição de morfemas podem ser divididas em duas linhas: de um lado, as que consideram que esta ordem se dá a partir de aspectos inatos (por exemplo, DULAY, BURT & KRASHEN (1982), ZOBL (1985)); de outro lado, uma perspectiva de análise que considere mais de um fator determinante.

Para Goldschneider & DeKeyser (2005), a possibilidade de que a ordem de aquisição de morfemas seja determinada, em grande parte, por propriedades dos próprios morfemas serve de base para a meta-análise proposta por eles. O estudo examina a combinação das seguintes propriedades: saliência perceptual, complexidade semântica, regularidade morfofonológica, categoria sintática e frequência. Os autores não excluem a possibilidade de que outros fatores externos aos morfemas, como a interferência da L1, também contribuam para a ordem observada. Para eles, se grande parte das descobertas puder ser explicada com base nestas cinco propriedades, então estaremos mais próximos de uma teoria adequada da ordem de aquisição.

Os dados de Goldschneider & DeKeyser (2005) foram retirados de 12 estudos desenvolvidos entre 1973 e 1996³⁶, envolvendo um total de 900 sujeitos (adultos e crianças) e 29 línguas como L1. Os resultados de todos os estudos foram calculados a partir do critério SOC (cf. seção 2.2.1)³⁷. Goldschneider & DeKeyser (2005) atribuem pontos a cada morfema. Em geral, os pontos mais altos são atribuídos de acordo com a presença de um fator ou subfator que facilite a aquisição de um dado morfema³⁸. Assim, quanto mais alto o valor numérico atribuído a um morfema, mais rapidamente ele deve ser adquirido³⁹.

Apresentamos a seguir uma tabela contendo os resultados de alguns dos estudos selecionados:

³⁶ Foram selecionados aqueles estudos que apresentavam um grupo de morfemas semelhantes, que priorizassem a produção oral, a fim de que resultados relativos a elementos da mesma natureza pudessem ser analisados.

³⁷ Para lembrar, o critério SOC (*Suppliance in Obligatory Context*) leva em conta todos os contextos onde os morfemas seriam esperados e não somente os contextos onde eles são empregados corretamente.

³⁸ Ao que nos parece, Goldschneider & DeKeyser (2005) atribuem pontos a fim de codificar suas expectativas hipotéticas de uma maneira, poderíamos dizer, “arbitrária”, uma vez que os critérios para isso não são esclarecidos de maneira satisfatória no texto original.

³⁹ Esta parece ser a lógica da análise dos autores. No entanto, há exceção para dois fatores: complexidade semântica e regularidade morfofonológica. Nestes fatores, os pontos mais altos são atribuídos àqueles morfemas que seriam hipoteticamente adquiridos mais tarde. Os autores definem isso como correlação negativa, sem maiores detalhes.

Tabela 6: Porcentagens de correção de estudos incluídos na meta-análise⁴⁰

Cód. do estudo	2	7	8	10	12	13	14	18	22	23	24	25
N*	73a	120c	80a	22a	45c	22a	66a	422c	18a	26c	24a	6m
-ING	84.38	67.33	93.75	97.00	78.67	84.00	88.67	97.87	96.33	99.04	85.58	91.00
Plural	78.63	67.67	93.33	75.00	64.33	71.00	92.33	73.06	80.33	54.31	59.90	80.00
Poss	34.63	34.33	63.75	75.00	44.00	-	40.67	81.99	-	23.08	35.71	74.00
Art	78.38	62.67	85.83	82.00	76.67	69.00	81.33	73.55	89.66	96.08	87.66	89.00
Reg Past	-	20.67	71.25	61.00 82.00	50.67	64.00	64.67	53.93	51.00	-	61.71	64.00
3rd Sing	39.88	07.67	55.00	60.00	18.33	36.00	25.67	66.61	36.66	15.39	44.38	45.00

N*= número de sujeitos na pesquisa, sendo *a* para adulto, *c* para criança, *m* para misturado e – para sem informação.

Podemos ver, por exemplo que, no estudo (2), o resultado da análise de regressão múltipla feita por Goldschneider & DeKeyser (2005) é o seguinte: a ordem de aquisição de morfemas é: *-ING, Plural, Article, 3rdP Sing e Possessive*.

Os fatores considerados na meta-análise⁴¹ são os seguintes:

Saliência Perceptual: Refere-se a quão fácil é ouvir ou perceber uma dada estrutura. Goldschneider & Dekeyser nos alertam que “*already vague, the definition*

⁴⁰ A tabela original nos mostra todos os estudos revisados, que somam vinte e cinco (25). Os doze (12) estudos selecionados por Goldschneider & DeKeyser (2005) são: (2) Bailey et al., 1974; (7) Fathman, 1975; (8) Fuller, 1978; (10) Houck et al, 1978; (12) Kjarsgaard, 1979; (13) Krashen et al., 1977; (14) Krashen et al., 1976; (18) Mace-Matluck, 1977; (22) Pica, 1983; (23) Riddle, 1993; (24) Rosado, 1986; (25) Rosansky, 1976.

⁴¹ Salientamos o fato de a pesquisa ser uma meta-análise, isto é, uma pesquisa utilizando dados de resultados de outras pesquisas calculados a partir de uma fórmula estatística. Entendemos que os resultados provenientes deste tipo de trabalho não são tão interessantes se comparados a trabalhos com resultados de pesquisas mais específicas.

of salience can be further confused because salience can refer solely to the characteristics of input itself, or to other factors that cause some parts of the input to become salient" (p. 48). Como veremos a seguir, Goldschneider & DeKeyser (2005) aderiram ao primeiro sentido da noção de "saliência".

Vários pesquisadores sustentam que este conceito desempenha um papel importante na ordem de aquisição dos morfemas. Brown (1973), por exemplo (conforme vimos na seção 2.2.1), afirma que "*some role for salience is guaranteed. The child will not learn what he cannot hear*" (p. 410). Larsen-Freeman (1976) considerou a saliência como um possível fator a condicionar a ordem de aquisição, mas depois o rejeitou a favor da frequência no *input*. Gass & Selinker (1994) também destacam o papel da saliência perceptual. Em todos estes estudos, a suposição é a de que, quanto mais perceptualmente saliente é um morfema, mais cedo ele será adquirido. Goldschneider & DeKeyser (2005) adotam um conceito de "saliência" composto de três subfatores:

- número de fones: refere-se à substância fonética de um morfema; quanto mais fones num morfema, mais perceptível e saliente ele deve ser;
- silabidade: aqueles morfemas que contêm uma vogal devem ser mais salientes do que aqueles que contêm somente consoantes
- sonoridade: morfemas que são mais sonoros devem ser mais salientes do que surdos.

Os autores estabelecem um valor diferente para cada subfator, para poder determinar o quão perceptualmente saliente um morfema pode ser. Assim, para o fator silabidade, o valor zero (0) foi atribuído àqueles morfemas que não tenham vogal e o valor um (1) é atribuído se o morfema contém uma vogal.

Complexidade semântica: É um conceito que se refere à medida de quantos significados são expressos num determinado morfema. Por exemplo, o plural -s expressa apenas número, ao passo que o morfema -s de terceira pessoa do singular expressa pessoa, número e tempo presente. Brown (1973) também considerou este critério e estabeleceu uma hierarquia de complexidade, pela qual formas com mais significados deveriam ser aprendidas mais tarde do que formas mais simples (cf. seção 2.2.1). Outros trabalhos também consideraram este critério:

Larsen-Freeman (1976), Gass & Selinker (1994), Andersen (1978), Krashen, Madden & Bailey (1975).

Goldschneider & DeKeyser (2005) deixam claro que há um consenso com relação à dificuldade de estabelecer teorias sofisticadas sobre “complexidade semântica”. Os próprios autores enfatizam: “*Today, as for Brown (1973), there is no general theory of semantic complexity that makes it possible to assign complexity values*” (1973, p. 50 e 369). Por concordarem com Brown, os autores seguem o mesmo critério dele em sua análise, estabelecendo uma hierarquia de complexidade semântica a partir do número de significados que um morfema pode ter e atribuindo um ponto a cada significado, para fins de cálculo.

Regularidade morfofonológica: Este conceito se refere ao grau com o qual os morfemas são ou não influenciados pelo ambiente fonológico, o que se revela por variação fonológica condicionada (alorfia) e contractibilidade (quando os fatores determinantes também são sintáticos). Segundo os autores, este fator é o menos definido na literatura e só foi incluído no estudo por causa das críticas de Cook (1993), que diz que a maioria dos estudos dos morfemas desconsidera a alorfia.

A premissa é a de que quanto mais morfofonologicamente regular um morfema, mais cedo ele deve ser adquirido. O cálculo envolve outros dois fatores: número de alternâncias fonológicas (morfemas com mais alternâncias devem ser adquiridos mais tarde, como, por exemplo, as diferentes variações fonéticas para o morfema *ed* [t], [d], [ɪd]) e homofonia com outros morfemas (morfemas que são homófonos a outros morfemas devem ser adquiridos mais tarde do que outros que não tem nenhuma correspondência).

O cálculo estabelece o valor um (1) para morfemas em que não há homofonia com outros morfemas e dois (2) para aqueles morfemas que apresentam homofonia com outros, como o *Plural –s*, o *Poss –s*, e *3rd P Sing –s*.⁴²

Categoria sintática: Goldschneider & DeKeyser (2005) limitam o conceito de categoria sintática às características de cada morfema a partir da Teoria de Categorias Funcionais (cf. capítulo 2, seção 2.1.2). Os autores afirmam que essa perspectiva diverge bastante de como a complexidade gramatical era analisada no passado, como em Brown (1973), por exemplo, cuja análise é fundamentada num modelo de representações sintáticas baseado no livro *English Transformational*

⁴² Como mencionamos anteriormente, os autores definem isso como correlação negativa. Isto é, o morfema que tiver pontos mais altos seria esperado mais tarde.

Grammar, de Jacobs & Rosebaum (1968). Goldschneider & DeKeyser afirmam que a teoria sintática evoluiu muito desde então e que não há sustentação para a ideia de que há uma correspondência direta entre o número de transformações sintáticas e o grau de dificuldade/facilidade na aquisição de um dado morfema. No entanto, os autores consideram que Krashen, Madden & Bailey (1975) obtiveram resultados interessantes observando separadamente morfemas relacionados a NPs e morfemas associados a Vs, descobrindo que a maior parte das diferenças entre a ordem de L1 e L2 era devido a diferenças em morfemas relacionados a V, enquanto morfemas associados a NP eram adquiridos em ordens muito semelhantes. Goldschneider & DeKeyser (2005) associam isso à análise feita por Zobl & Liceras (1994), que veremos na seção a seguir, associando a ordem de aquisição de um dado morfema à sua categoria sintática.

Para explicar a ordem de aquisição de morfemas, Zobl e Liceras propõem uma divisão destes em categorias lexicais e funcionais; além disso, propõem que a divisão em morfemas livres e presos também é pertinente. Eles postulam um padrão em que os itens lexicais são adquiridos antes dos funcionais e, dentro destes grupos, morfemas livres adquiridos antes de morfemas presos. A partir destas considerações, Goldschneider & DeKeyser (2005) tentam estabelecer uma maneira de prever quais os morfemas ocorrerão primeiro, através de um cálculo com atribuição de pontos a cada nível da hierarquia sintática proposta por Zobl & Liceras (1994). Os pontos mais altos eram atribuídos às formas a serem adquiridas primeiro e os pontos mais baixos às formas a serem adquiridas mais tarde, como ilustramos abaixo:

CATEGORIA LEXICAL

Morfema livre ----- 4 pontos

Morfema preso ----- 3 pontos

CATEGORIA FUNCIONAL

Morfema livre ----- 2 pontos

Morfema preso ----- 1 ponto

Frequência: O conceito se refere ao número de vezes que uma determinada estrutura está presente no *input* do aprendiz. Ou seja, segundo este fator, quanto

mais frequente um determinado morfema ocorrer na fala endereçada ao aprendiz, mais fácil e rapidamente será adquirido.

Apesar de este fator ser mencionado em quase todas as pesquisas que tratam do assunto, alguns trabalhos mostram que a frequência no *input* não é decisiva para a ordem de aquisição dos morfemas. Brown (1973) não encontrou correlação clara entre a frequência de estruturas na fala dos pais e a ordem de aquisição nas crianças estudadas. Larsen-Freeman (1976) conclui, por outro lado, que a frequência dos morfemas na produção do falante nativo é o principal determinante para a ordem dos morfemas na produção de aprendizes de inglês como L2; mas a própria autora, na mesma obra, coloca que sua conclusão pode ser precipitada, porque seus dados se baseavam na comparação da frequência dos morfemas dos pais participantes da pesquisa de Brown com vinte e quatro (24) alunos adultos de inglês como L2. Dulay & Burt (1978) não concordam que a frequência possa ser o fator que explique a ordem dos morfemas baseados na premissa de que os morfemas gramaticais comuns ocorrem com maior frequência em inglês do que a maioria dos itens lexicais mais comuns e, ainda assim, esses morfemas são adquiridos tardiamente se comparados a palavras de conteúdo.

Goldschneider & DeKeyser (2005) utilizam as frequências da fala dos pais dadas por Brown (1973) como parâmetro para o cálculo que propõem. Os autores ponderam que os dados de Brown são de pais para aprendizes de L1 e, por isso, poderiam não ser pertinentes para aprendizes de inglês como L2 nem para crianças aprendendo L2. Porém, argumentam que os morfemas em questão são bastante elementares e, por isso, não se espera muita variação de uma situação de *input* para outra.

Transferência da L1: Este é o último dos seis fatores que podem determinar a ordem de aquisição dos morfemas, segundo Goldschneider & DeKeyser . No entanto, o real papel da L1 no processo de ASL tem gerado muita discussão e controvérsia. Dulay & Burt (1973) reportam que somente 3% dos erros feitos por crianças na sua pesquisa eram devidos à interferência da L1. Andersen (1978) atribui um papel maior em seu estudo com falantes de espanhol aprendendo artigos e o *Poss-s* em inglês como L2 e conclui que “*L1 transfer clearly is a factor that must be taken into consideration as one of the factors that could interact with morpheme acquisition and accuracy orders*” (p. 267).

Goldschneider & DeKeyser (2005) não consideram este fator nos cálculos da sua pesquisa, argumentando que, em sua proposta, que envolve uma meta-análise, torna-se praticamente impossível controlar as variáveis das vinte e nove L1s envolvidas no corpus selecionado. Os autores colocam, no entanto, que parece ser consensual que as estruturas da L1 desempenham papel importante na aquisição de L2 e concluem que, se entendermos que a saliência num sentido amplo contribui fortemente para uma determinada ordem de aquisição de morfemas, as semelhanças ou diferenças entre a L1 e a L2 poderão marcar tais saliências e, desta maneira, constituir fator determinante na incorporação de morfemas numa dada sequência.

Tendo definido os fatores que julgam relevantes e oriundos de quantificação de seus escores, Goldschneider & DeKeyser (2005) fazem uma análise de regressão múltipla,⁴³ a fim de determinar quanto da variação total da ordem de aquisição de morfemas em aprendizes de inglês como L2 pode ser atribuída à combinação dos cinco (5) fatores selecionados, a saber, saliência perceptual, complexidade semântica, regularidade morfofonológica, categoria sintática e frequência.

Como a pesquisa faz uma meta-análise, os resultados são percentuais de performance média de desempenho correto com relação a morfemas de cada grupo selecionado.

Reproduzimos abaixo a tabela com o escore obtido por Goldschneider & Dekeyser (2005) para cada morfema a partir da pontuação proposta para cada fator:

⁴³ Tradução livre do inglês *multiple regression*.

Tabela 7: Contagem para cada morfema a partir de cada determinante^{44 45}

	Pres Prog	Plural –s	Poss-s	Articles	Reg Past	3rdP – s
1	2	1.33	1.33	1.66	1.33	1.33
2	1.7996	-.5959	-.5959	.5840	-.5959	-.5959
3	1	.33	.33	1	.33	.33
4	1.2910	-.6455	-.6455	1.291	-.6455	-.6455
5	11	5.33	5.33	10.33	3.66	5.33
6	1.3683	-.4922	-.4922	1.1485	-1.0402	-.4922
7	2	1	1	1	2	3
8	2	3	3	4	3	3
9	-1.5811	0	0	-1.5811	0	0
10	1	2	2	1	1	2
11	-.9129	-.9129	-.9129	-.9129	-.9129	.9129
12	3	3	1	2	1	1
13	160	147	71	552	44	89

LEGENDA

Linhas
1. escore bruto ⁴⁶ do número de fones
2. escore –z do número de fones
3. escore bruto de silábico ou não silábico
4. escore –z de silábico ou não silábico
5. escore bruto da sonoridade
6. escore –z da sonoridade
7. escore da complexidade semântica
8. escore bruto do número de alternâncias fonológicas
9. escore –z do número de alternâncias fonológicas
10. escore bruto da homofonia
11. escore –z da homofonia
12. escore da categoria sintática
13. frequência

⁴⁴ Tabela adaptada a partir da que é apresentada por Goldschneider & DeKeyser (2005), p. 74, apêndice D. Tradução livre.

⁴⁵ Enfatizamos que esta tabela apresenta os resultados já calculados a partir dos critérios. Ou seja, não está claro ao longo do artigo a exata atribuição de pontos a cada morfema a partir de cada critério. Por exemplo, não sabemos com clareza porque no critério de complexidade semântica, o morfema *Reg Past* tem dois pontos ao passo que *3rdP sing* tem três, assim como não temos informações precisas sobre a atribuição de pontos a partir da silabidade de cada morfema.

⁴⁶ Sabemos que “escore bruto” e “escore –z” são termos utilizados em linguagem estatística. Para melhor interpretarmos os resultados apresentados pelos autores, vamos entender escore bruto como um valor individual observado numa determinada variável de medição e escore –z como escore padronizado, que se refere à posição de um determinado escore com relação aos demais numa distribuição. Salienciamos que isso não está claro no artigo original.

Desta forma, podemos ver os escores de pontuação obtidos por Goldschneider & Dekeyser (2005). Os escores indicam a probabilidade com que, por hipótese, um morfema será adquirido primeiro que outro a partir de um determinado fator. Assim, temos diferentes índices para cada morfema dentro de cada fator.

Por exemplo: com relação ao número de fones (linha 2), temos um escore bruto que mostra que o morfema do *Present Progressive –ING* apresenta um valor mais alto do que os outros. Ou seja, a substância fonética deste morfema, por hipótese, é significativa, fazendo que ele seja mais saliente e, por isso, ele seria adquirido anteriormente aos outros especificamente a partir deste valor.

Já com relação ao fator categoria sintática (linha 12), o morfema *–ING* recebe 3 pontos, pois é caracterizado como lexical e preso; por isso, deve emergir antes no processo de aquisição que o morfema *Possessive*, que é de categoria funcional e preso (e obtém o valor 1). Os dados de Krashen (1977) confirmam isso, pois *–ING* precede *Possessive* (cf. Tabela 3).

O fator frequência (linha 13) prevê que o morfema *Article* esteja no topo da sequência de aquisição, seguido por *Present Progressive* e *Plural-s*. Em seguida, observaríamos *3rd-P sing*, *Possessive* e *Reg Past*. Os dados de Krashen (1977) mostram que essa ordem é válida, à exceção de *Article*, que não é o primeiro a ser adquirido, mas sim o segundo.

Com relação ao fator complexidade semântica, o morfema *3rd-P Sing* tem três (3) pontos e, a partir desse critério, seria adquirido em último lugar, seguido dos morfemas *Reg Past* e *–ING*, com dois pontos cada.

Em linhas gerais, para Goldschneider & DeKeyser (2005), os resultados desta meta-análise indicam que a combinação dos cinco fatores mencionados acima poderia explicar porque alguns morfemas emergiriam antes do que outros.

Os autores dizem que os cinco fatores investigados têm algo em comum e não são parte de um conjunto heterogêneo: eles constituiriam aspectos da saliência no sentido amplo do termo. Isto é, para os autores, o que facilitaria o processo de indução de estruturas gramaticais de elementos do *input* seria a saliência em vários níveis (fonológico, morfológico, sintático, semântico e numérico).

O primeiro fator – saliência perceptual – é o exemplo mais claro, pois número de fones, silabidade e sonoridade contribuem para uma estrutura ser perceptualmente distinta no seu ambiente. Complexidade semântica dificulta a relação entre forma e significado: o *–s* de terceira pessoa representa pessoa,

número, tempo e aspecto e apenas quando estes quatro componentes de significado coincidem este morfema aparece. Regularidade morfofonológica pode ser entendida como um tipo de saliência da relação forma-significado, uma forma mais abstrata de saliência; a categoria sintática também pode refletir saliência, no sentido de que morfemas livres são mais perceptíveis do que os presos e, naturalmente, morfemas lexicais mais salientes do que funcionais. E a frequência também contribui para a saliência, pois quanto mais frequente for um morfema, mais destacado e, portanto, mais saliente ele será.

Goldschneider & DeKeyser (2005) argumentam que este entendimento da noção de “saliência” pode prever a ordem de aquisição de maneira satisfatória, não sendo necessário postular qualquer tipo de predisposição inata específica ou de modelo sintático abstrato. Os autores nos dizem que, embora alguns estudos prévios tenham tentado mostrar que algum fator único é responsável pela ordem natural dos morfemas (LARSEN-FREEMAN, 1976; ZOBL & LICERAS, 1994; ZOBL, 1995), nenhuma destas tentativas respondeu satisfatoriamente à pergunta básica: Por que aprendizes de inglês como segunda língua de diferentes idades, com tipos diferentes de exposição ao inglês e diferentes L1s, parecem adquirir certos morfemas gramaticais basicamente na mesma ordem? E, segundo os autores, devido ao fato de não terem respondido a esta pergunta, as pesquisas mudaram de direção.⁴⁷

Uma das sugestões dos autores é a de que sua análise seria compatível com uma abordagem de linhas mais conexionistas de investigação, pois, para eles, os vários aspectos dos morfemas podem ser vistos como múltiplas pistas num modelo de competição⁴⁸. Se interpretarmos os resultados nesta perspectiva, ou seja, associando o paradigma conexionista de análise linguística e o modelo da competição, poderíamos pensar que os morfemas possuem propriedades inerentes e competem entre si. E aquelas, que se salientam mais, são adquiridas primeiro.

Passemos à outra proposta de explicação para uma ordem na aquisição de morfemas.

⁴⁷ Não concordamos totalmente com esta posição dos autores, pois, como veremos na seção seguinte, Zobl & Liceras (1994) oferecem uma resposta a essa pergunta, é claro, dentro de sua perspectiva teórica.

⁴⁸ Segundo este modelo, as línguas dispõem de “pistas” diferentes, ou seja, propriedades linguísticas de nível sintático, fonológico, etc., que seriam mais salientes ao aprendiz no momento da aquisição (cf. McWHINNEY, 1997).

3.2.2 Zobl & Liceras (1994)

Zobl & Liceras (1994) fazem uma reanálise dos estudos dos morfemas conduzidos nos anos de 1970, a partir de avanços na teoria linguística. Em particular, tomam como ponto de partida o trabalho das últimas décadas sobre categorias funcionais e sua aquisição em L1 e L2, trabalho que evoluiu de tal forma a possibilitar o retorno a esses estudos e oferecer uma explicação baseada em princípios (*a principled explanation*) para as diferenças na ordem de aquisição de morfemas em inglês como L1 e L2.

Na teoria linguística assumida por Zobl & Liceras (CHOMSKY, 1991; POLLOCK, 1989), a estrutura frasal é projetada a partir daqueles elementos que participam nas relações de marcação de papel-theta (predicados e seus argumentos) e a partir de categorias funcionais, que compreendem morfemas livres como modais, auxiliares, determinantes, complementizadores, e morfemas presos, como os afixos nominais e verbais. Tal teoria inspirou várias hipóteses com relação à aquisição de L1 e L2. Mais especificamente para a aquisição de L1, foi proposto que categorias funcionais emergem de acordo com uma escala maturacional.

Radford (1990) oferece uma contribuição bastante significativa às pesquisas com relação ao desenvolvimento de uma escala maturacional no processo de aquisição de linguagem. O autor conduz um estudo detalhado das gramáticas iniciais das crianças a partir de um corpus de mais de cem mil enunciados de fala espontânea. Radford demonstra que a característica fundamental das primeiras estruturas produzidas pelas crianças é que elas são essencialmente lexicais e temáticas por natureza. Tais enunciados evidenciam a aquisição de categorias lexicais e não funcionais, explicando porque as crianças não dominam determinantes, possessivos, pronomes, expletivos e propriedades associadas ao caso, ligação, concordância, tempo, auxiliares, infinitivos, etc.

No caso do inglês como L2, duas hipóteses se apresentam para Zobl & Liceras (1994): (i) categorias funcionais estariam presentes desde o início da via L1, através da gramática universal; ou (ii) categorias funcionais deveriam ser gradualmente construídas com base no *input*.

Zobl & Liceras procuram oferecer uma explicação teórica unificada para os diferentes resultados encontrados em estudos sobre a aquisição de morfemas nos anos de 1970: para eles, as pesquisas nos anos de 1990 sobre as categorias funcionais e sua aquisição em L1 e L2 evoluíram de tal forma a possibilitar uma melhor compreensão dos resultados obtidos e um retorno ao estudo sobre o assunto. Os resultados que veremos a partir da revisão destes estudos mostram duas ordens para a aquisição de morfemas, uma em L1 e outra em L2. Estas ordens evidenciam, segundo Zobl & Liceras (1994) diferenças associadas às posições mais altas em L2 do verbo *Be* como *Aux* e *Cop*, bem como diferenças associadas à posição mais baixa do *Poss-s*, o que veremos com mais detalhes abaixo.

Zobl & Liceras não consideram frequência como um fator que possa explicar a ordem da aquisição de morfemas, em consonância com Brown (1973). Eles concordam com Krashen, Madden & Bailey (1975), que sustentam que um determinado conhecimento linguístico prévio possa ser uma explicação para tal ordem. Zobl & Liceras colocam, ainda, que este tipo de hipótese (um conhecimento sintático prévio, de natureza inata) antecipa, de alguma maneira, a posição que considera que categorias funcionais estão disponíveis para o aprendiz de L2.

Eles propõem que:

(i) Categorias funcionais, em L1, são adquiridas numa ordem específica (*category specific*), com categorias nominais precedendo categorias verbais; em L2, por outro lado, a ordem é cruzada (*cross-categorical*); assim, pode-se concluir que as categorias funcionais estão disponíveis desde o início de processo de aquisição.

(ii) Em L1, os morfemas gramaticais parecem ser iguais aos morfemas lexicais na implementação de categorias funcionais. Isso seria compatível com a visão de que categorias funcionais amadurecem em L1, o que permite ao aprendiz ser “sensível” às influências lexicais e flexionais da língua. No entanto, em L2, os morfemas gramaticais não precisam desempenhar um papel tão determinante graças à disponibilidade de projeções funcionais. A ordem dos morfemas em L2 sugere que morfemas livres, ao invés de afixos, representam as primeiras manifestações das categorias funcionais em L2. Ainda, o movimento de afixo parece desempenhar um papel-chave no desenvolvimento dos afixos com uma função sintática: segundo Zobl & Liceras (1994), o movimento é a generalização mais significativa que explica a ordem, em L2, dos morfemas presos (nominais e verbais)

Poss -s, Reg Past e 3rd PSing -s. Como vimos na seção 2.3.5, o movimento do afixo representa uma opção na parametrização de mova-alfa, um princípio da GU.

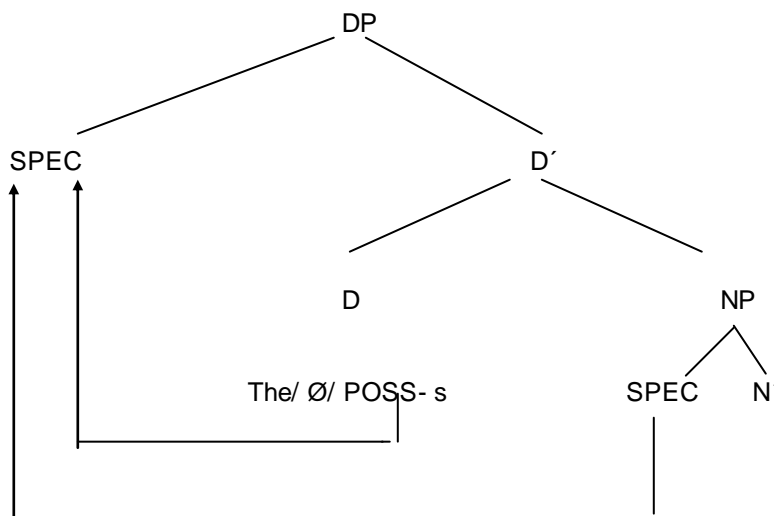
(iii) A ordem dos morfemas não parece estar associada à hipótese de que categorias funcionais são adquiridas gradualmente a partir do *input* (cf. Vainikka & Young-Scholten, 1993); mas sim à hipótese de que os aprendizes de L2 não precisam aprender as categorias como tal e sim os seus expoentes específicos de cada língua.

Os autores então propõem uma reanálise dos resultados dos estudos dos anos de 1970 à luz da teoria de categorias funcionais. Com respeito à ordem de aquisição em L1, foram selecionados os estudos de Brown (1973) e De Villiers & De Villiers (1973), e com relação à ordem em L2, os estudos de Bailey *et al.* (1974), Larsen-Freeman (1975), Krashen (1977) e Andersen (1978).

Zobl & Liceras (1994), baseados em Pollock (1989), assumem que determinantes e afixos de tempo e concordância encabeçam suas próprias projeções, constituindo, respectivamente, DP e IP. Também assumem que AUX BE encabeça sua própria projeção VP e sobe para Infl. Além disso, pressupõem que verbos principais não sobem para Infl, mas que *tense* e *agreement* descem até os verbos principais.⁴⁹ Por último, assumem que os artigos evidenciam a presença de projeções DP.

Assim, os autores adotam as seguintes análises para DPs e IPs em inglês⁵⁰:

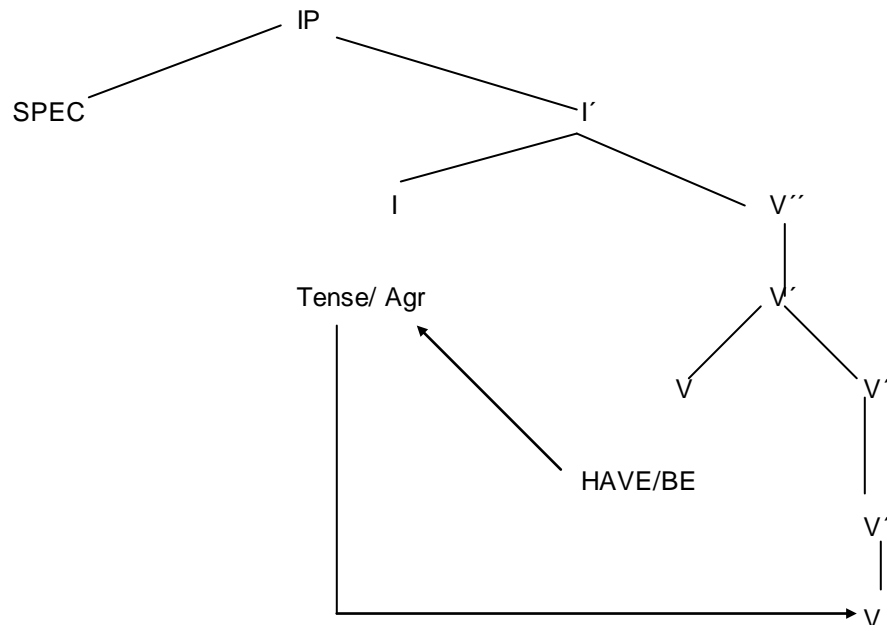
(14) Projeção DP



⁴⁹ Como já foi visto no capítulo 2, seção 2.1.5.

⁵⁰ Projeções apresentadas por Zobl & Liceras (1994)

(15) Projeção IP



A primeira figura representa a estrutura DP. Nela, os determinantes e *Poss-s* são gerados em D. Spec de NP se move para Spec de DP. Na segunda figura, que representa a estrutura IP, tanto o *Aux Be* quanto o *Aux Have* subirão até a posição núcleo de IP, isto é, a posição I, a não ser que um modal esteja presente. Mas quando não há modais ou outro auxiliar, *tense* e *agreement* descem até o verbo principal. Como vimos no capítulo 2, esse movimento de descida do afixo é necessário em inglês porque, ao contrário do francês e do espanhol, os traços de Infl em inglês não são “fortes” para fazer com que o verbo suba além de VP.

A questão da acurácia se impõe naturalmente como um problema metodológico, dada a dificuldade em determinar uma ordem de aquisição a partir da emergência de um morfema X e de seu uso correto. Os autores mencionam não somente a dificuldade em considerarmos um morfema adquirido a partir do momento em que ele começa a ser produzido corretamente, mas também o problema em estabelecermos uma hierarquia a partir disso, uma vez que os morfemas não surgem numa sequência perfeita. Por isso, Zobl & Liceras (1994) justificam que a interpretação dos dados sobre morfemas será feita de forma conservadora, ou seja,

considerando a maneira como os morfemas foram adquiridos em blocos (no topo, no meio e embaixo da hierarquia) e não ortodoxamente como morfema um, dois, etc.

A fim de melhor acompanharmos a explicação dada por Zobl & Liceras (1994), repetiremos abaixo a ordem de aquisição encontrada nas diferentes pesquisas por eles utilizadas quanto à L1:

Tabela 8: Ordem de aquisição de Brown (1973) e De Villiers & De Villiers (1973)⁵¹

Ordem de Brown (1973)	Ordem de De Villiers & De Villiers (1973)
V-ing	V-ing/ PLU (ral) – S
PLU (ral) – S	IRREG. PAST
IRREG. PAST	ART (icle)
POSS(essive) –S	POSS(essive) –S
Uncontr COP (ula) BE	Cont. COP(ula) BE
ART (icle)	REG(ular) PAST/ 3rd P(erson) – S
REG (ular) PAST	Uncontr COP (ula) BE
3rd P(erson) – S	Contr. AUX
Uncont AUX	Uncontr. AUX

Conforme já dissemos, as principais observações são, segundo Zobl & Liceras (1994):

1. A clara separação entre o *Aux Be* e o afixo *-ing*, já que a forma contínua se apresenta no topo e a forma não contraída do *Aux Be*, permanece em última posição.

2. A forma *Cop Be* precede o *Aux Be*, embora as duas formas sejam idênticas e a análise corrente assuma que ambas se movam para Infl.

3. *Art* e *Poss* se apresentam juntos, o que representa, uma evidência de D livre e preso.

4. *Reg. Past* e *3rdP-sing* se apresentam juntos no final da hierarquia. Ambos são expoentes flexionais de Infl e estão sujeitos à descida de afixo (“*affix-lowering*”).

Apresentamos a seguir a tabela que sintetiza a ordem de aquisição de morfemas nos vários estudos sobre L2 compilados por Zobl e Liceras (1994):

⁵¹ Tabela extraída de Zobl & Liceras (1994). Note que quatorze (14) morfemas apresentados na tabela original foram reduzidos a nove (9) e este é o modelo padrão utilizado na maioria dos estudos de aquisição de morfemas.

Tabela 9: Ordens da aquisição de inglês como L2 (a partir de Zobl & Liceras, 1994)

Krashen (1977)	V-ing/ PLU/ COP		AUX/ ART			IRREG. PAST		REG. PAST	POSS
Andersen (1978)	COP		ART	AUX/ V-ing	PLU	IRREG. PAST	POSS	REG. PAST	3rd P-S
Bailey, Madden & Krashen (1974)	V-ing/COP		ART/AUX	PLU		REG. PAST		3rd P-S	POSS
Larsen-Freeman (1975)	Time 1: V-ing Time 2: COP	COP V-ing	ART AUX	AUX ART	PLU PLU	REG. PAST POSS	3rd P-S IRREG. PAST	IRREG. PAST 3rd P-S	POSS REG. PAST

As semelhanças que podem ser observadas, desta vez, são:

1. *Cop Be* precede *Aux Be*, que também se apresenta nas primeiras posições. Lembre que, em L1, tais morfemas aparecem nas últimas posições.

2. *Art* precede *Poss-s* em todos os estudos, e *Poss-s* se apresenta nas últimas posições em todos os estudos. Lembre que, em L1, *Art* e *Poss-s* aparecem juntamente.

3. *Poss-S*, *3rdP-S* e *Reg. Past* estão agrupados nas últimas posições.

4. *Aux Be*, de um lado, e *Reg Past* e *3rdP-S*, do outro, estão em extremos opostos da hierarquia, apesar de todos os três serem expoentes de Infl.

A partir da comparação destas tabelas, Zobl e Liceras (1994) oferecem uma discussão bastante aprofundada sobre o assunto, estabelecendo algumas diferenças entre a aquisição de morfemas como L1 e L2.

Vejamos a seguir uma tabela simplificada com os resultados das duas pesquisas mais utilizadas como referência nos estudos sobre a ordem de aquisição de morfemas, a fim de melhor podermos sintetizar e analisar os dados:

Tabela 10: Ordem de aquisição dos morfemas gramaticais em Brown (1973) e em Krashen (1977)⁵²

BROWN (1973)	KRASHEN (1977)
V- ing	ING
PLU – S	PLURAL
Irreg Past	COP
Poss – S	AUXILIARY
Unc COP BE	ARTICLE
ART	IRREG PAST
REG Past	REG PAST
3 rd P – S	III SING
Uncont AUX	POSSESSIVE

⁵² Tabela apresentada por Zobl & Liceras (1994)

As principais observações dos autores são:

1. Em L2, as projeções de categorias funcionais já estão disponíveis desde os estágios preliminares de aquisição, o que é evidenciado através dos expoentes lexicais de *Det (article)*, *Aux Be* e *Cop*. Em L1, os núcleos flexionais (*Poss – s*, *3rdPS* e *Reg Past*) desempenham um papel semelhante ao marcarem projeções funcionais.

2. Em L2, os núcleos de projeções funcionais se manifestam simultaneamente (*cross-categorically*, isto é, cruzadas categoricamente, com categorias variadas, surgindo ao mesmo tempo), sugerindo que a aquisição de tais projeções não é de categoria específica (*category-specific*). Em L2, *Cop*, *Aux Be* e *Det (Article)* aparecem no topo do ranking de todos os estudos mostrados e os afixos (como *3rd-PS* e *Reg Past*) aparecem no final. Já em L1, *Cop*, *Det (Article)* e *Aux Be* estão distribuídos ao longo da hierarquia, sugerindo que as diferentes projeções funcionais emergem em diferentes momentos. *Pos-s* pode ser visto no meio da hierarquia, assim como *Article*, *3rdPS* e *Reg Past* no fim da sequência antes de *Aux Be*. Portanto, em L1, nota-se uma clara separação das flexões de IP e de DP.

3. A ordem em L2 pode ser explicada através da distinção entre morfemas livres e presos. De um lado, os núcleos lexicais estão no topo da hierarquia, enquanto que os núcleos flexionais expressos por afixos (por exemplo, *3rd-PS*) estão no fim. Este contraste entre os morfemas livres e presos sobrepõe-se ao contraste entre o movimento de núcleos lexicais (*Aux* e *Cop* sobem para Infl) e o movimento de afixo. Desta maneira, o movimento de afixo, mais especificamente, pode representar, para Zobl & Liceras (1994), uma generalização no processo de formação gramatical de aprendizes de L2. A disponibilidade de categorias funcionais desde os estágios mais elementares de aprendizagem, manifesta através de seus expoentes lexicais⁵³, parece adiar a aquisição de afixos⁵⁴.

⁵³ Zobl & Liceras (1994) parecem entender expoentes lexicais como palavras e não como morfemas presos.

⁵⁴ Ao que nos parece, Zobl & Liceras (1994) utilizam o termo “adiar” querendo dizer que, mesmo que categorias funcionais já estejam disponíveis, a aquisição de afixos parece ser deixada para mais tarde, o que pode ser ilustrado através de *3rd PSing* e *Poss’s* nas últimas posições da sequência.

Assim, podemos ver que a teoria das categorias funcionais, para Zobl & Liceras (1994), parece oferecer uma explicação para as diferenças entre as ordens de aquisição de morfemas em L1 e L2, o que torna possível observarmos paralelos sintáticos entre morfemas livres e presos nas categorias. Quanto aos dados referentes à ordem de aquisição de morfemas, podemos observar um agrupamento natural de morfemas (i) em termos de categorias DP e IP, (ii) em termos de movimento de núcleos (lexical ou de afixo) e (iii) em termos da distinção entre morfemas livres e presos. Isso, segundo Zobl & Liceras (1994), nos leva a duas descobertas: a primeira, com relação à aquisição de L1, diz respeito ao desenvolvimento de projeções funcionais através de categorias específicas; a segunda, com relação à ASL, diz respeito (i) ao desenvolvimento de projeções funcionais através de categorias cruzadas (isto é, categorias funcionais e lexicais emergindo simultaneamente), e (ii) ao papel de liderança desempenhado por morfemas livres na implementação de projeções funcionais, bem como (iii) ao papel do movimento na aquisição tardia de afixos sintáticos. Estes últimos não parecem ter a mesma função de “desencadeamento” de estruturas em L2 do que em L1 (já que, por exemplo, *Poss-s*, uma instância de DP preso, aparece no final da hierarquia em L2 e em L1, surge no meio da sequência). Ainda, o movimento dos núcleos lexicais parece acontecer mais cedo e de maneira independente do movimento de afixo.

Os autores concluem que sua análise proposta para a aquisição da ordem dos morfemas em L1 e L2 responde às críticas que sustentam que a sequência de aquisição dos morfemas não apresentaria generalizações, não sendo, desta maneira, reveladora dos processos que intervêm na aquisição de L2.

3.3 QUESTÕES A SEREM INVESTIGADAS

Como vimos ao longo deste capítulo, várias pesquisas sobre aquisição de morfemas em inglês, tanto na aquisição de L1 quanto na de L2, mostram que ambos os processos parecem caracterizados por uma sequência constante. No entanto, como já foi mencionado na introdução desta tese, não sabemos de nenhum estudo que contemple os estágios evolutivos na aquisição de morfemas em inglês como L2 por aprendizes brasileiros.

Por isso, surgem naturalmente algumas questões:

- Como seria o processo evolutivo destes sujeitos com relação à aquisição de morfemas do inglês?
 - Quais os morfemas que emergiriam primeiro neste contexto? E quais apareceriam por último?
 - Seria essa sequência semelhante àquela já postulada em pesquisas de referência, como a ordem de Krashen (1977)?
 - Quais seriam as razões para uma ordem semelhante ou diferente?
 - Uma abordagem multifatorial, que atribua à saliência um papel fundamental no processo de aquisição, como proposto por Goldschneider & DeKeyser (2005), poderia dar conta do processo observado nos aprendizes brasileiros?
 - Ou poderíamos explicar melhor a ordem encontrada em nossos aprendizes como resultado da pronta disponibilidade de categorias funcionais já no início do processo de aquisição de L2?

Deste modo, dispomos não somente de algumas questões norteadoras para esta tese, como também de alguns padrões evolutivos na aquisição de morfemas que devem servir de referência para nossa investigação.

4 METODOLOGIA

Este capítulo está organizado da seguinte forma:

Na primeira seção, veremos algumas questões teóricas sobre linguística de corpus e corpus de aprendiz, questões principalmente associadas a definições e perspectivas de pesquisa a partir de um corpus. Em seguida, veremos alguns exemplos de corpora de aprendizes que têm sido desenvolvidos e que servem de parâmetro para a nossa pesquisa. Ainda, são apresentadas questões sobre compilação e análise de corpora e também sobre etiquetagem de erros, aspectos fundamentais para a elaboração do BELC, o corpus desta pesquisa.

A segunda seção descreve especificamente este corpus. Nela, veremos detalhes sobre o BELC, o corpus de aprendiz que foi construído para servir de base à pesquisa desta tese e a outros projetos futuros. Primeiramente, veremos detalhes sobre a coleta dos dados, no que se refere aos sujeitos da pesquisa e a procedimentos executados para a obtenção dos dados que constituem o corpus. Depois, conheceremos a estrutura do BELC: número de textos, palavras e códigos de identificação. Na subseção seguinte, poderemos ter um panorama geral do BELC a partir de uma primeira leitura realizada por mim, com exemplos de algumas estruturas observadas respectivamente organizadas por níveis. Por fim, na subseção 4.2.4, veremos procedimentos de análise do BELC, isto é, considerações gerais sobre etiquetagem, material utilizado, critérios e procedimentos específicos, e cálculo dos resultados.

4.1 LINGUÍSTICA DE CORPUS

4.1.1 Linguística de Corpus e Corpus de Aprendiz

McEnery & Wilson (1996) caracterizam a LC como uma metodologia de pesquisa que se propõe a estudar a linguagem em exemplos de uso real.

Berber Sardinha (2004) define linguística de corpus (LC) da seguinte maneira:

A Linguística de Corpus ocupa-se da coleta e da exploração de corpora, ou conjunto de dados linguísticos textuais coletados criteriosamente, com o propósito de servir para a pesquisa de uma língua ou variedade linguística. Como tal dedica-se à exploração por meio de evidências empíricas, extraídas por computador (p. 3).

Ainda segundo Berber Sardinha (2004), corpus é:

Um conjunto de dados linguísticos (pertencentes ao uso oral ou escrito da língua, ou a ambos), sistematizados segundo determinados critérios, suficientemente extensos em amplitude e profundidade, de maneira que sejam representativos da totalidade do uso linguístico ou de algum de seus âmbitos, dispostos de tal modo que possam ser processados por computador, com a finalidade de propiciar resultados vários e úteis para a descrição e análise (p. 18).

Podemos observar alguns aspectos que se sobressaem na caracterização de um corpus, como a amostragem, a representatividade, a autenticidade e a forma legível para máquina.

Sarmiento (2008), numa análise sobre verbos modais a partir de um corpus especializado em inglês de aviação, estabelece considerações bastante interessantes sobre pesquisas com uso de corpora. A autora nos apresenta um panorama histórico sobre linguística de corpus enfatizando a importância e o impacto do desenvolvimento de estudos baseados em extratos reais de fala. Sarmiento aponta a organização do corpus Brown, feita por Nelson Francis e Henry Kucera entre 1961 e 1964, como um marco para o desenvolvimento de estudos empíricos em contraposição a estudos formalistas, abordagem dominante na época, dado o advento das teorias gerativas de Noam Chomsky. Sabemos que esta abordagem enfoca a competência linguística do falante (isto é, aquilo que ele sabe) como objeto de estudo e não o seu desempenho linguístico (isto é, aquilo que o falante efetivamente faz ao usar uma dada língua). Já a LC enfoca o desempenho dos falantes de uma língua como objeto de estudo, considerando a ocorrência de itens linguísticos em situação de uso real, e não em exemplos intrateóricos, idealizados ou elicitados para descrever ou explicar alguma estrutura. No entanto, apesar de salientar a dicotomia entre a abordagem de pesquisa baseada em LC e a

abordagem gerativa, Sarmiento reconhece a possibilidade das duas linhas serem complementares uma a outra e contribuirão para a pesquisa linguística.⁵⁵

Assim, a LC torna possível um tipo de descrição linguística que pode ser muito proveitosa para os profissionais do ensino de língua estrangeira, já que esses terão acesso a informações sobre os estágios de aquisição de determinadas estruturas como língua estrangeira de forma sistematizada, ou seja, organizadas de maneira a poderem ser investigadas eletronicamente.

Segundo Granger (2002), *“The area of linguistic enquiry known as learner corpus research, which has only existed since the late 1980’s, has created an important link between the two previously disparate fields of corpus linguistics and foreign/second language research”* (p. 4).

Para tais pesquisas, existem os corpora de aprendizes, ou *learner corpora*, de falantes não-nativos, obviamente, que buscam descrever tendências e estruturas observadas no processo de aquisição de LE.

Segundo Leech (1998, *apud* PRAVEC, 2002) *“A learner corpus is a computerized database of the language produced by foreign language learners”* (p.81). Ainda conforme Leech, um corpus de aprendiz é uma fonte para qualquer um que queira investigar como as línguas são aprendidas.

Granger (2002) adota a definição de *learner corpora* de Sinclair (1996):

Computer learner corpora are electronic collections of authentic FL/SL textual data assembled according to explicit design criteria for a particular SLA/FLT purpose. They are encoded in a standardized and homogeneous way and documented as to their origin and provenance (p. 7).

Para Granger (2002), o principal objetivo na compilação de um corpus é coletar informações objetivas que possam ajudar na descrição da língua do aprendiz, uma vez que, ao examinar um corpus de erros devidamente identificados, por exemplo, um pesquisador não tem acesso somente aos erros do aprendiz, mas também à interlíngua deste.

⁵⁵ A proposta desta tese é um exemplo desta possibilidade. Como vimos no capítulo 2 e 3, alguns estudos mostrados são de orientação gerativa, como Zobl & Liceras (1994), por exemplo. A explicação para a ordem dos morfemas oferecida pelos autores é inclusive considerada como uma das possibilidades de explicação para os nossos resultados, que, por sua vez, serão oriundos de dados de um corpus de fala espontânea.

Cobb (2003) aponta para questões importantes sobre a dificuldade na análise da interlíngua do aprendiz, que é, na verdade, uma fonte de informação para professores e pesquisadores por apresentar sistematicidade na estrutura e não somente erros aleatórios, sustentando a pesquisa baseada em corpora como grande aliada na investigação destes aspectos:

The current trend to electronic submission of classroom writing relieves the data problem, defines steps in the acquisition of literacy and may even contribute to the growth of theory. ... A corpus provides the empirical means to discover what advanced learners know and do not know about their L2 (p. 393).

Cobb (2003) aponta ainda para a falta de dados nas pesquisas sobre interlíngua e ainda para a falta de métodos de pesquisa mais “sensíveis” à interlíngua do aprendiz. Mais especificamente,

Even if commonalities in advanced learning exist, these are likely to be in areas beyond morpho-syntax, such as lexis, discourse and pragmatics, where we have a few characterizations of native speaker performance (NS), let alone of the consequences by which learners arrive there (p. 395). Particularly promising is the comparison of two corpus types, which can reveal not only what is in a learner corpus, but what's not in it (p. 395).

Para Berber Sardinha (2004), o corpus de aprendiz não somente redefine o conceito de corpus (uma vez que a linguagem investigada não pertence à variedade nativa), mas também levanta questões bastante importantes à natureza do tipo de pesquisa a ser conduzida. Questões como, por exemplo, a frequência do uso de determinadas palavras em diferentes níveis do processo de aquisição de LE comparada com a frequência dos falantes nativos.⁵⁶

A LC, enquanto metodologia, muito contribui para tal proposta de pesquisa, uma vez que propõe a investigação de uma determinada estrutura linguística de natureza sintática, semântica, lexical, etc., baseada em produção linguística espontânea, através de um corpus organizado de maneira tal para poder ser estudado via uso de ferramentas computacionais. Por isso, segundo Berber

⁵⁶ Salientamos a posição de Sarmiento (2008) com relação à dicotomia falante nativo (FN) versus falante não nativo (FNN): parece haver um consenso de que as variações linguísticas dos falantes nativos assumem um status superior às variações produzidas por falantes não nativos, quando, por vezes, não há insucesso comunicativo em algumas situações, e sim, diferenças que não interferem na interação. Por isso, as pesquisas com o propósito de comparar corpora de aprendizes e corpora de falantes nativos devem verificar a verdadeira interferência das diferenças (palavra preferida pela autora ao invés de erros ou desvios) nas situações comunicativas.

Sardinha (2004), a LC aplicada ao ensino de línguas deve ser pensada como instrumento de caráter descritivo, e não normativo, pois visa a entender melhor a interlíngua dos aprendizes e, conseqüentemente, melhorar as ações de ensino.⁵⁷

Neste contexto, a questão da análise do erro⁵⁸, na perspectiva da LC, é algo positivo, no sentido de buscar o entendimento do erro, e não classificá-lo negativamente como desvio do padrão da fala nativa. O trabalho apresentado nesta tese está integralmente de acordo com essa proposta. Isto é: o erro aqui não é visto com qualquer tipo de preconceito normativo, e sim, como um uso não apropriado da forma padrão que faz parte do processo evolutivo de ASL. Tanto que o critério utilizado na metodologia, como veremos ainda neste capítulo, contempla os contextos onde os morfemas devem ocorrer, e procura determinar o comportamento dos aprendizes diante destes contextos a fim de determinar o padrão evolutivo dos aprendizes brasileiros de inglês como L2 com relação à aquisição de morfemas em inglês como L2. Por isso, os erros não são prioritários em nossa análise, e, sim, os acertos, tanto que consideramos a porcentagem de correção. E, a partir destes índices, pretendemos propor um padrão evolutivo que possa ser característico do nosso aprendiz.

Vejamos, a seguir, alguns exemplos de corpora de aprendizes.

4.1.2 Corpora de Aprendizes

Tono (2003) afirma que o interesse em pesquisa com corpora de aprendizes é notável e crescente e dá inúmeros exemplos de projetos de pesquisa com corpora

⁵⁷ Em consonância com Berber Sardinha, enfatizamos que este trabalho se propõe a compreender o processo evolutivo dos aprendizes brasileiros no que se refere a morfemas a fim de poder oferecer melhores condições aos instrutores de inglês e pesquisadores sobre o entendimento do processo de ASL. Isso poderia possibilitar uma postura diferenciada do instrutor/pesquisador, pois, a partir do conhecimento desses estágios evolutivos, os professores estariam mais capacitados em sala de aula para conduzirem atividades que orientassem melhor o aluno com relação à produção de formas inapropriadas.

⁵⁸ Ellis (1994) afirma que o estudo do erro é feito através da Análise do Erro (AE). Esta ideia surgiu nos anos de 1970, substituindo a Análise Contrastiva (*Contrastive Analysis*), que tentava prever os erros dos aprendizes através da comparação das línguas envolvidas no processo de aprendizagem (como já vimos brevemente no capítulo 2). Até então, os erros em L2 eram considerados formas indesejadas, enquanto que, num contexto de aquisição de L1, os erros eram vistos como fatos dentro de uma fase transitória, e os erros em L1 na fase adulta eram enganos. Assim, os erros em L2 eram vistos como um aspecto negativo da língua do aprendiz.

de aprendizes que estão sendo desenvolvidos atualmente em vários países desde o início dos anos de 1990, alguns dos quais veremos a seguir.

Segundo Pravec (2002), corpora de aprendizes são predominantemente encontrados na Europa e na Ásia e um dos mais importantes corpora de aprendizes é o ICLE (*International Corpora of Learner English*). São textos produzidos por alunos universitários de inglês de nível avançado com diferentes L1s.

O ICLE⁵⁹ é uma proposta bastante interessante de investigação na área de aquisição de inglês como LE. Este corpus é desenvolvido na Universidade de Louvain-la-Neuve, na Bélgica, no CECL – *Center for English Corpus Linguistics*, e é coordenado pela professora Sylvianne Granger. Caracteriza-se por um projeto que envolve textos de aprendizes universitários de inglês na condição de EAP – *English for Academic Purposes*, de dezesseis diferentes línguas maternas. A última versão, lançada em 2009, é constituída de três milhões e setecentas mil (3.700.000) palavras a partir de seis mil e oitenta e cinco (6.085) ensaios escritos (cf. Granger, em recente comunicação pessoal).

Tal corpus conta com a colaboração de várias instituições universitárias de outros países (como Bulgária, República Tcheca, Holanda, Finlândia, França, Alemanha, Itália, Polônia, Rússia, Espanha, Suécia, e outros). O *Center for English Corpus Linguistics* (CECL) estimula a pesquisa em aplicações potenciais de corpora de aprendiz a materiais pedagógicos e recursos de aprendizagem.

Granger⁶⁰ afirma que este tipo de corpus possibilita a análise, descrição e comparação da língua dos aprendizes e de suas tendências. Ela nos coloca que um corpus de aprendizes pode ser entendido como um “*yardstick to measure the distance between learner performance and target language*”. E atenta para o fato de que a pesquisa com corpora de aprendizes torna possível um panorama da língua destes sujeitos e que as implicações para o ensino permitem delinear “*learner needs, teaching objectives and teachability, and what you are going to select or ignore*”. (Granger, X ELC, 2010).

O *Longman Learners' Corpus* (cf. já mencionado no capítulo 2) foi desenvolvido pela *Longman Corpus Network*, contendo dez (10) milhões de palavras em textos escritos produzidos por aprendizes de inglês de várias nacionalidades, de

⁵⁹ <<http://cecl.fltr.ucl.ac.be/Cecl-Projects/Icle/icle.htm>>

⁶⁰ Alguns dados que citamos são de conferências recentes dadas por Sylvianne Granger no X ELC (Encontro de Linguística de Corpus), 2010.

todos os níveis de aprendizagem, devidamente classificados a partir de sua proficiência. Esse corpus teve sua origem no início dos anos de 1990 (GRANGER, 2003) e, segundo Tono (2003), é disponível comercialmente.

O *Cambridge Learners Corpus* (CLC) também é outro corpus de aprendiz bastante importante, com dez milhões de palavras de textos de alunos de todos os níveis.

No Brasil, as pesquisas com corpora avançaram bastante nos últimos anos. Especificamente com relação a corpora de aprendizes, podemos citar o BrICLE e o CoMAprend. O Br-ICLE, disponível no endereço <<http://www2.lael.pucsp/corpora/bricle/index.htm>>, é um projeto coordenado pelo Prof Dr. Tony Berber Sardinha, (BERBER SARDINHA, 2004), numa parceria entre a Pontifícia Universidade Católica do São Paulo e o CECL, e tem seu acesso parcialmente restrito. O BrICLE é atualmente constituído de aproximadamente cem mil palavras de textos argumentativos produzidos por falantes de PB como L1, aprendizes de inglês como L2 de nível avançado. No entanto, a última versão do ICLE, lançada em 2009, não conta com dados do Br-ICLE (cf. Tagnin & Fromm, 2009).⁶¹

O CoMAprend (Corpus Multilíngue de Aprendiz) é um projeto muito interessante desenvolvido pela prof. Stella Tagnin e por Guilherme Fromm, na Universidade de São Paulo. Este corpus é constituído de redações de aprendizes brasileiros de cinco línguas estrangeiras: alemão, espanhol, francês, inglês e italiano, alunos de cursos de graduação e extracurriculares. Tagnin e Fromm (2009) apontam várias possibilidades a partir desta proposta: (i) pesquisas horizontais, como a comparação entre alunos de um mesmo nível ou turma; (ii) pesquisas verticais, acompanhando um aluno especificamente ou um grupo; (iii) investigações sobre estratégias de redação; (iv) avaliação de métodos ou materiais didáticos e (v) pesquisas translinguais, ou seja, verificar se as dificuldades dos alunos com relação a algum aspecto ou estrutura são específicas a uma ou a várias línguas.

Tagnin e Fromm (2009) descrevem com detalhes o projeto, que contou com financiamento da agência CAPES em dois momentos para a elaboração e atualização de seu banco de dados e do site. O corpus é totalmente digitalizado e é aberto à participação de voluntários. O fato de existir corpora de aprendizes de

⁶¹ Sabemos que, neste momento, o projeto está em desenvolvimento. Os coordenadores (comunicação pessoal) afirmam que estão em busca de sujeitos voluntários para atingirem o número de duzentas mil palavras, índice solicitado por Granger para que um corpus possa ser integrado ao ICLE.

inglês como L2 falantes de PB como L1 é bastante interessante para os fins da pesquisa que esta tese propõe.

Nesta perspectiva, podemos também citar o CABrI (Corpus de Aprendizizes Brasileiros de Inglês), pela Prof. Dr. Deise P. Dutra, da Universidade Federal de Minas Gerais. Este é um corpus ainda em desenvolvimento que deverá ser integrado ao Br-ICLE, e é constituído de 80 textos e 40.808 palavras.⁶² Sabemos que os sujeitos são aprendizes universitários que produziram textos argumentativos, seguindo, portanto, o perfil de sujeitos do Br-ICLE, que são aprendizes de nível avançado de inglês.

Há ainda um corpus de aprendizes de inglês sendo desenvolvido por Cristina Kindermann, da UNIP (Universidade Paulista). Em Kindermann (2009), a autora relata a elaboração de um corpus a partir de sujeitos alunos do curso de letras e justifica a relevância do estudo como uma ferramenta que possa fazer com que os próprios alunos, igualmente professores de inglês, possam ser mais flexíveis a diferentes tipos de análise linguística.

Além disso, Stella Tagnin ainda coordena o projeto COMET (Corpus multilíngue para ensino e tradução), uma parceria da USP com o BrICLE, que tem por objetivo servir de suporte a pesquisas linguísticas, principalmente nas áreas de tradução, terminologia e ensino de línguas. O projeto é amplamente descrito em Tagnin (2003) no endereço <www.fflch.usp.br/dlm/comet>.

No entanto, não encontramos nenhum estudo baseado em corpora de aprendizes de inglês como L2 falantes de PB como L1 que contemple a produção de aprendizes⁶³ desde o nível inicial, fato este que revela uma lacuna nas pesquisas, dada a importância da análise da produção do aprendiz desde o estágio evolutivo mais elementar de seu processo de aprendizagem.

Em seu trabalho, Izumi, Uchimoto & Isahara (2005) apresentam o NICT JLE (*Japanese Learner English Corpus*) e o JEFLL (*Japanese EFL Learner Corpus*), sendo o primeiro, um corpus com dois milhões (2.000.000) de palavras de textos orais e o último com um milhão de palavras a partir de textos escritos e orais produzidos por aprendizes de inglês como LE de todos os níveis. Eles fazem uma

⁶² Informações obtidas a partir de www.corpuslg.org/Elc/Programa/Resumos/Deise.html, 2009.

⁶³ Aprendizizes de inglês “geral”, isto é, sujeitos de diferentes áreas de atuação estudantil e profissional, e não aprendizes de inglês dos cursos de Letras, o que, a princípio, deve fazer alguma diferença.

breve revisão sobre a noção de análise e etiquetagem de erros, noções estas que veremos de maneira genérica na seção seguinte.

4.1.3 Compilação e Análise de Corpora

Há uma série de questões envolvidas na coleta, no armazenamento e no processamento dos dados para a construção de um corpus.

Berber Sardinha (2004) cita alguns critérios para a coleta de material. Podemos citar critérios com relação aos informantes, como caracterização e controle dos grupos etários, contexto educacional, nível de proficiência, etc. Também há critérios com relação aos textos, ou seja, eles podem ser escritos, não-técnicos, argumentativos, ter extensão fixa de 500 palavras, podem ser textos integrais, etc. Há outras variáveis importantes que são imediatamente relevantes a esta pesquisa, como tempo de exposição a uma dada língua, língua materna, local de nascimento, conhecimento de outras línguas. E ainda critérios com relação a tópicos e ambiente de tarefa.⁶⁴

Tono (2003) apresenta uma tabela sobre considerações a serem feitas na elaboração de um corpus de aprendiz, a qual reproduzimos abaixo:

Tabela 11: *Design considerations for building learner corpora*⁶⁵

<i>FEATURE: LANGUAGE-RELATED</i>	<i>FEATURE: TASK-RELATED</i>	<i>FEATURE: LEARNER RELATED</i>
<i>Mode: written/spoken</i>	<i>Data collection: cross-sectional/ Longitudinal</i>	<i>Internal-cognitive: age/ cognitive style</i>
<i>Genre: letter/diary/fiction/essay</i>	<i>Elicitation: spontaneous/ prepared</i>	<i>Internal-affective (motivation/ attitude)</i>
<i>Style: narration/ argumentation</i>	<i>Use of references: dictionary/ source text</i>	<i>L1 background L2 environment (ESL/ EFL/ level of school)</i>
<i>Topic: general/leisure/etc</i>	<i>Time limitation: fixed/free/homework</i>	<i>L2 proficiency (standard test score)</i>

⁶⁴ Os critérios levados em conta especificamente para esta pesquisa serão detalhados na próxima seção.

⁶⁵ Esta tabela contém informações muito interessantes e organizadas sobre como elaborar um corpus de aprendiz e, por esse motivo, serve como base para a organização do BELC, o corpus de aprendiz desta tese. Veremos mais detalhes sobre a utilização destas ideias aplicadas ao BELC na seção 4.2.

Para Granger (2002), a coleta e armazenamento de um corpus devem estar associados aos objetivos de pesquisa:

SLA/ FLT Purpose: a learner corpus is collected for a particular SLA or FLT purpose. Researchers may want to test or improve some aspects of SLA theory, for example, by confirming or disconfirming theories about transfer from L1 or the order of acquisition of morphemes, or they may want to contribute to the production of better FLT tools or methods (p. 10).

A autora aponta duas abordagens metodológicas a partir da exploração de corpora de aprendiz: *Contrastive Interlanguage Analysis* e *Computer-Aided Error Analysis*. A primeira se concentra na comparação de dados entre línguas envolvidas num dado contexto de aquisição e a segunda se caracteriza pela análise de erros envolvendo em alto grau de padronização com o auxílio de recursos computacionais (o que é bastante diferente das pesquisas sobre análise do erro propostas na década de 70).⁶⁶ Para tanto, há ferramentas que facilitam tal análise. Por exemplo, os erros podem ser anotados, isto é, codificados conforme uma determinada informação linguística (McENERY & WILSON (1996), BERBER SARDINHA, (2004)). Dentro deste processo de anotação, há o *Part-of-Speech (POS) tagging* (a etiquetagem segundo a parte do discurso no qual a palavra pertence) e o *error tagging*, ou etiquetagem de erros.

4.1.3.1 Considerações sobre etiquetagem de erros

O centro coordenado por Granger elaborou uma ferramenta de etiquetagem de erros chamada *Error Editor*, que viabiliza a pesquisadores etiquetarem seus erros a partir dos textos de aprendizes a fim de, posteriormente, elaborarem listas de erros típicos de aprendizes. A ideia é de que isto possa, em outro momento, facilitar o desenvolvimento de programas computacionais que possam ajudar no processo de aprendizagem, no que concerne à gramática, ortografia, verificação de estilo e outros.

⁶⁶ Tais pesquisas, conforme apontam Ellis & Barkhuizen (2005), além de atribuírem um peso maior ao erro como um desvio da forma padrão, eram mais específicas a possíveis estruturas contrastivas entre determinadas línguas envolvidas num dado estudo.

O *Error Editor* envolve uma série de etapas, entre elas a correção manual dos erros do aprendiz e classificação desse erro dentro de uma “*error tag*” (ou etiqueta de erro). Tal etiqueta consiste num código maior e uma série de códigos menores. Há sete grandes categorias (*major categories*), que são: formal; gramatical, léxico-gramatical, lexical, registro, ordem de palavra/palavra faltando/palavra redundante e estilo. Tais códigos são, então, seguidos de subcódigos e assim obtêm-se catálogos de erros típicos de aprendiz. Após esse processo, os erros podem ser pesquisados a partir destes códigos (GRANGER, 2002; GRANGER, 2010, X ELC)

Sabemos, ainda, que há algumas outras ferramentas disponíveis.

O WordSmith Tools é um programa de análise de textos criado por Mike Scott e disponibilizado em 1995, cujo objetivo, de um modo geral, é analisar como as palavras se comportam em textos (BERBER SARDINHA, 2004). É amplamente utilizado em pesquisas na área de linguística de corpus, que oferece recursos como listador de palavras, concordanciador, análise de frequências e palavras-chave, entre outros. O programa existe nas versões de 1 a 5 e se apresenta em versões limitadas (demo) gratuitas na rede ou também pode ser adquirido.

O CLAWS (*Constituent Likelihood Automatic Word-tagging System*) é um etiquetador (POS *tagger*) muito utilizado desde os anos de 1980, desenvolvido pela UCREL (*University Centre for Computer Corpus Research on Language*), um centro de pesquisa da Universidade de Lancaster. Disponível na rede de modo limitado (somente 300 palavras), ele pode também ser adquirido.

Pravec (2002) salienta ainda que não há um único padrão para a compilação, etiquetagem e organização de um corpus de aprendiz, ou para as ferramentas usadas para acessar esse corpus, uma vez que cada um foi criado para servir a diferentes propósitos (de pesquisa ou pedagógicos), o que não impede que os corpora possam ser posteriormente usados por outros pesquisadores para fins diferentes.

Dentre os vários corpora disponíveis, o *Cambridge Learner Corpus* (CLC) é o único que possui uma parte codificada para erros de aprendiz, o que viabiliza a pesquisa dos erros mais frequentes em inglês de aprendiz. No entanto, segundo Izumi, Uchimoto & Isahara (2005), o método usado na etiquetagem do corpus não é claramente informado, assim como o próprio acesso a ele é restrito. Os autores dizem que a etiquetagem é feita por somente um anotador desde 1993, quando

começou, a fim de evitar variáveis. Sabe-se que o CLC cobre 80 tipos de erros e que o anotador escolhe uma etiqueta apropriada para cada erro principalmente a partir da identificação de qual parte da fala (POS) ela envolve e o quanto ela difere da forma correta (se é por redundância, omissão ou substituição).

Vários corpora de aprendiz, incluindo o NICT JLE (*Japanese Learner English Corpus*) são anotados com etiquetas de erros predeterminadas através da categorização de erros prováveis a partir de regras gramaticais canônicas já existentes.

Outro exemplo de corpus de aprendiz etiquetado com erros é o JEFLL (*Japanese EFL Learner Corpus*). De acordo com Izumi, Uchimoto & Isahara (2005), tal corpus conta com dados oriundos de produção escrita e oral e concentra-se na análise de erros de tipo específico, tais como artigos, formas plurais de substantivos, terceira pessoa do singular na conjugação de presente simples e outros a nível morfológico, gramatical e lexical, selecionados a partir dos interesses da pesquisa. A etiquetagem desse corpus contém três tipos de informação: POS, regras gramaticais/lexicais/morfológicas e forma correta.

Tono (2003) nos coloca que vários pesquisadores têm optado por fazer a etiquetagem manualmente devido à falta de precisão nos processadores automáticos disponíveis. No entanto, duas questões devem ser observadas: a primeira com relação à validade (a autora afirma que o esquema deve incluir pelo menos dois aspectos: classificação de categoria linguística e taxionomia de modificação da forma-alvo (omissão, adição, etc.); e a segunda com relação à confiabilidade, uma vez que o mesmo tipo de erro pode ser corrigido de várias maneiras. Para isso, Tono (2003) menciona a construção de um manual para a etiquetagem de erros que, na época, vinha sendo construído por Granger, e conclui: “*Sharing such tools will better facilitate the standardization of corpus annotation in the future*” (TONO, 2003, p. 804).

Granger⁶⁷, então, apresenta “*The Louvain Error tagging manual*”, com cinquenta e seis (56) etiquetas padronizadas para erros de diversas naturezas. A etiquetagem de erros, a partir desta orientação, deve seguir seis (6) princípios. Por exemplo, o princípio 3 diz: “*Place the correction immediately after the erroneous Word/phrase. For retrieval purposes, the corrected form is preceded and followed by*

⁶⁷ X ELC, 2010.

a \$ sign. E.g. some (XNUC) advices \$advice\$.” XNUC é um dos 56 códigos de etiquetagem que quer dizer que a natureza do erro é de *Lexico-grammar, Nouns, Uncountable*.^{68 69}

McEnery, Xiao & Tono (2006), na pesquisa que desenvolveram sobre morfemas utilizando dados de corpus de aprendiz (como vimos no capítulo 2), apresentam duas possibilidades de etiquetagem para os propósitos de sua pesquisa. Interessantemente, a seção na qual McEnery, Xiao & Tono dão detalhes sobre os procedimentos de etiquetagem dos erros é intitulada “*Error tagging: a proper way and a dirty way*”. Os autores definem como *proper way*, ou maneira apropriada, o critério de análise que ocupa muito tempo e é extremamente trabalhoso, no qual, de um modo geral, os dados devem ser ajustados manualmente para poderem ser lidos pela máquina. Este método, apesar de demandar muito trabalho, é confiável e os dados etiquetados podem ser reutilizados para pesquisas afins. O outro método, no qual os autores se referem como *dirty way*, é mais rápido no sentido de não demandar tanto trabalho manual, mas os dados não podem ser utilizados em outras pesquisas. Os autores dão detalhes sobre os procedimentos computacionais no texto original, mas, de um modo geral, podemos dizer que este critério de etiquetagem sugere a seleção de informação, já que os dados são muitos. Assim, o pesquisador, após submeter seus dados na máquina através da ferramenta de análise apropriada (no caso da pesquisa dos autores, o *Word Smith Tools* versão 4), selecionaria uma frase em cada dez para a contagem. Por isso, o método é referido como *dirty way*.

Como podemos ver, a questão de etiquetagem de erros é bastante complexa.

Segundo McEnery & Gabrielatos (2006), uma crítica à etiquetagem é que ela impõe categorias preestabelecidas baseadas em um modelo teórico, podendo, assim, haver uma interferência dessas categorias sobre os dados. Nessa mesma

⁶⁸ O trabalho de Granger é notável por sua complexidade e rigor. Note-se que, sabidamente, a autora conta com uma numerosa equipe de colaboradores, dentro e fora do Centro, e grande disponibilidade de recursos, o que favorece o desenvolvimento do projeto.

⁶⁹ Na análise do BELC, não utilizamos qualquer padrão exposto no manual de Granger não somente porque ele só foi disponibilizado em 2009, mas, principalmente, porque, como já colocamos, a metodologia do estudo que propomos tem um objetivo muito pontual, que é o de estabelecer um padrão evolutivo de estruturas morfêmicas a partir de um corpus de aprendizes. Isso faz com que tenhamos que pensar em critérios particulares de análise e etiquetagem, aos quais os códigos do manual em questão não poderiam ser aplicados. Veremos mais adiante, ainda neste capítulo, que tivemos que criar códigos e critérios particulares para a nossa pesquisa.

linha, podemos citar Leech (1997), que determina algumas máximas com relação à prática do uso de corpora, dentre elas a de que o corpus puro (do inglês, *raw corpus* – isto é, o texto sem qualquer tipo de anotação) deve poder ser sempre recuperável e a de que os usuários devem ter acesso aos procedimentos de etiquetagem e a tudo associado a estes processos, como, por exemplo, o porquê da escolha de determinado critério, etc.

De acordo com isso, resolvemos elaborar o BELC de modo a armazenar seus dados num banco de dados original, sem etiquetagem, somente com as informações de identificação acima do texto – um cabeçalho (cf. veremos na seção 4.2.2), que deverão ser assim mantidas independentemente do propósito de utilização. Para os fins da pesquisa que estamos propondo nesta tese, procedemos à etiquetagem com vistas a levantar dados sobre a ocorrência de morfemas previamente selecionados. Consequentemente, temos duas versões do BELC: uma sem qualquer tipo de anotação e outra devidamente etiquetada para a análise desta pesquisa.

Assim, o BELC foi elaborado a partir de leituras de incontáveis fontes, sendo as principais mencionadas neste capítulo. Passamos a seguir à segunda seção, que propõe a descrição do BELC no que se refere, em linhas gerais, a sua estrutura e a procedimentos de análise de dados específicos a ele.

4.2 O CORPUS BELC – *BRAZILIAN ENGLISH LEARNER CORPUS*

4.2.1 A Coleta dos Dados

4.2.1.1 Os sujeitos

Como informantes para a pesquisa, selecionamos alunos do curso de inglês geral oferecido pela Faculdade de Letras da Pontifícia Universidade Católica do Rio Grande do Sul. Os sujeitos eram graduandos ou graduados das mais diversas áreas, que fazem o curso como disciplina eletiva (complementar para o número de horas no currículo de graduação) ou como curso regular de línguas. O curso tem um total de oito níveis e o material didático utilizado era a série *American Inside Out*, do livro *Elementary* até o *Upper Intermediate*.

No total, 424 sujeitos participaram da pesquisa, distribuídos da seguinte maneira:

Tabela 12: Sujeitos da pesquisa classificados por nível do curso

NIVEL	INFORMANTES
1	42
2	61
3	86
4	62
5	40
6	38
7	59
8	36

Tivemos um total de vinte e quatro turmas onde o instrumento foi aplicado.⁷⁰

⁷⁰ A relação das turmas pode ser consultada no anexo 8 deste trabalho.

4.2.1.2 Procedimentos

O instrumento de coleta de dados é constituído de quatro partes, que serão descritas a seguir:

4.2.1.2.1 Apresentação da pesquisa e identificação dos informantes

Os sujeitos: (i) foram brevemente informados sobre o assunto da pesquisa que estão participando, (ii) assinaram o documento de autorização para uso dos dados e (iii) preencheram a ficha de identificação,⁷¹ onde informam questões pessoais que possam ser relevantes à pesquisa, como local de nascimento, data, sexo, tempo de exposição ao inglês, domínio de outro idioma e outros.

4.2.1.2.2 Nivelamento

Dada a grande diferença existente entre os alunos no que se refere à proficiência linguística, decidimos aplicar um teste de nivelamento a fim de atenuar esta diferença como variável na análise da produção escrita destes sujeitos. Para tanto, pesquisamos alguns testes já existentes e selecionamos o *Placement Test* da *Oxford University Learning Center (OULC)*.⁷²

A escolha deste teste se deve a várias razões: além de ser um teste neutro, no sentido de não fazer parte do material didático utilizado pelos sujeitos, ele é um “teste-modelo” com relação ao propósito que vários alunos apresentam (o de estudar inglês no exterior por algum tempo). Entendemos que outros testes de proficiência disponível poderiam ser mais complexos para boa parte dos informantes.

O teste tem 50 questões que devem ser respondidas de 30 a 40 minutos, conforme instruções originais, que foram respeitadas na aplicação.

⁷¹ Uma cópia destes documentos está no anexo 4 desta tese.

⁷² Este teste encontra-se em anexo (anexo 1) no trabalho.

Com relação à classificação, o teste dispõe, originalmente, de três níveis classificatórios diferentes:

Tabela 13: Classificação de proficiência segundo o OULC

SCORE	CLASSIFICAÇÃO DO OULC
0-30	<i>Too low</i>
31- 0	<i>English for Social or Academic Purposes</i>
41-50	<i>Advanced</i>

Para os fins desta pesquisa, consideramos mais prudente mudar esta classificação e estabelecer quatro diferentes níveis de proficiência, ao invés de três, dado o grande número de sujeitos participantes da coleta e a grande diferença em proficiência linguística neles observada, e assim criamos o seguinte nivelamento:

Tabela 14: Classificação de proficiência desta pesquisa

SCORE	CLASSIFICAÇÃO DESTA PESQUISA
0-20	<i>Beginner (B)</i>
21-30	<i>Pre-Intermediate (P)</i>
31- 40	<i>Intermediate (I)</i>
41-50	<i>Advanced (A)</i>

Desta maneira, os sujeitos foram classificados de duas formas: segundo o nível de inglês do curso que estão fazendo (níveis 1 a 8) e, também, segundo o nível de proficiência que ficaram classificados a partir do teste de nivelamento (quatro níveis, *beginner*, *pré-intermediate*, *intermediate* e *advanced*). Para isso, utilizamos a seguinte codificação:⁷³

- letra N (para nível) + X (número do nível do curso, podendo variar de 1 a 8)
- letra P (para *placement*) + B, P, I ou A (para um dos níveis classificatórios desta pesquisa, conforme expusemos acima).

⁷³ A codificação completa com ilustração será vista na seção seguinte.

4.2.1.3 Textos

Conforme já vimos na seção anterior relativamente aos pressupostos teóricos sobre LC, um corpus é essencialmente uma coletânea de material escrito produzido espontaneamente por falantes de uma dada língua. Sendo assim, nos vimos confrontados com o fato de solicitarmos um tipo de produção escrita dos sujeitos que, em sua proposta temática, motivasse a ocorrência de morfemas, mas que fosse, ao mesmo tempo, natural e espontânea.

Por isso, pensamos na redação de três parágrafos, com a distribuição temática e quantitativa da seguinte forma:

Tabela 15: Informações sobre os textos

	TEMA	TAMANHO
Texto 1	Informações pessoais em primeira pessoa	Aproximadamente 100 palavras
Texto 2	Informações pessoais em terceira pessoa	Aproximadamente 100 palavras
Texto 3	Informações sobre uma viagem que o sujeito tenha realizado	Aproximadamente 200 palavras

4.2.1.4 A coleta dos dados propriamente dita

4.2.1.4.1 Planejamento da coleta

A coleta dos dados para esta pesquisa foi feita em dois momentos no ano de 2008. No primeiro semestre (de maio a junho) e no segundo semestre (de setembro a novembro). A necessidade de coletar os dados em duas etapas diferentes deu-se devido ao número insuficiente de informantes numa só aplicação.

A coleta dos dados constituiu-se em três partes:

- Parte 1: apresentação da pesquisa, informações sobre os sujeitos e teste de nivelamento.
- Parte 2: redação de dois parágrafos (sobre os temas 1 e 2, cf. Tabela 15).
- Parte 3: redação de mais um parágrafo (sobre o tema 3).

Os professores de cada turma aplicaram as tarefas e receberam uma folha de instruções, para seguirem determinadas normas padronizadas para a pesquisa,⁷⁴ como a restrição de tempo (30 minutos para cada tarefa) e a não utilização de dicionários e qualquer outro corretor digital.⁷⁵ Ainda, os professores não responderiam perguntas dos alunos com relação à ortografia, vocabulário ou gramática.

Os professores aplicaram os instrumentos, distribuindo as cópias das folhas de instruções sobre cada uma das tarefas, conforme anexo. Os alunos eram solicitados a lerem atentamente as instruções e, em seguida, partirem para a redação dos textos, que deveriam ser enviados por e-mail (no caso da coleta digitalizada) ou entregues ao professor responsável no final do tempo estimado, que era de 30 minutos.

A coleta foi idealizada para ser realizada num período de três semanas, a fim de mantermos um espaço curto entre as aplicações, e isso foi possível em grande parte das turmas, em especial na primeira edição de aplicação, isto é, no primeiro semestre. No entanto, algumas turmas tiveram o intervalo de aplicação alterado em mais de uma semana por problemas de calendário dos professores e também devido ao fato de alguns alunos terem apresentado certa resistência com relação à atividade.

Foi feito um cronograma de aplicação para organizarmos da melhor maneira possível a distribuição das cópias-matrizes⁷⁶ e dos espaços dos laboratórios de pesquisa utilizados na coleta digitalizada dos textos. A coleta foi idealizada para dispor de dados unicamente digitalizados, para facilitar o armazenamento dos mesmos em termos de tempo, organização e outras variáveis. Todavia, como já vimos acima e veremos a seguir, alguns alunos apresentaram resistência ao

⁷⁴ A folha de instruções também está em anexo (anexos 2 e 3).

⁷⁵ Por isso, os alunos eram orientados a escreverem seus textos diretamente no corpo do e-mail e não em arquivo Word separadamente, a fim de preservar os textos de corretores automáticos.

⁷⁶ As cópias-matrizes são as folhas contendo as instruções para a elaboração dos textos, que eram reutilizadas por diferentes grupos.

realizarem a pesquisa, pelo fato de terem que se deslocar das salas, irem até o laboratório e redigir textos. Assim, em algumas turmas, a coleta foi feita manualmente em sala de aula e os dados foram digitalizados por uma monitora particular contratada por mim, autora da pesquisa, monitora esta com conhecimentos de inglês que foi devidamente orientada a digitar os dados sem qualquer tipo de alteração, nem de ordem ortográfica, dada à prioridade científica das informações neles obtidas.

Desta maneira, organizamos toda esta informação num banco de dados constituído de pastas digitalizadas, organizadas por níveis. Podemos ver dados mais detalhados sobre a estrutura do BELC a seguir na seção 4.2.2.

4.2.1.4.2 Considerações sobre a coleta

Em linhas gerais, a coleta transcorreu de maneira positiva, pois, apesar das dificuldades alheias aos fatores da pesquisa, como por exemplo, ausência dos alunos no dia da coleta, cronograma rígido durante o período letivo que facilmente disponibilizasse a aplicação do instrumento durante as aulas, etc, conseguimos coletar um corpus com mais de cem mil palavras, o que é considerado um corpus efetivamente representativo (McENERY & WILSON (1996)).

A coleta do primeiro semestre foi diferente da do segundo, principalmente em função de uma certa resistência por parte dos alunos em participarem da pesquisa. Isto foi uma variável que surpreendeu a todos, inclusive aos professores que auxiliaram na coleta dos textos.⁷⁷ Atribuimos tal resistência a alguns fatos: (i) alunos de língua estrangeira em geral não gostam de escrever (ii) medo de se expor numa pesquisa (iii) descrença na contribuição científica da pesquisa, considerando a atividade uma “perda de tempo”.

Sendo assim, organizamos a coleta do segundo semestre de maneira diferente, a fim de evitarmos mais resistência por parte dos sujeitos.

Como já mencionamos acima, no segundo semestre, por exemplo, não fizemos a aplicação em três semanas seguidas, para que as datas não fossem

⁷⁷ Tagnin & Formm (2009) também observam esta resistência na elaboração do corpus CoMAprend.

previsíveis pelos alunos, embora acreditemos que isso não tenha sido uma variável que possa ter causado interferência ou qualquer alteração diferente nos dados para os propósitos da nossa pesquisa.

Com relação ao teste de nivelamento, alguns informantes o consideraram bastante longo, principalmente os de níveis iniciais. Já esperávamos tal tipo de comentário, mas mesmo assim, mantivemos o tempo sugerido pelo OULC, que era de 30 minutos.

4.2.2 Estrutura do BELC

Apresentamos, a seguir, uma tabela que mostra em detalhes os textos que compõem o BELC, respectivamente com o número de sujeitos participantes e o número de palavras em cada texto e nível.

Tabela 16: Estrutura geral do BELC

	TEXTO	SUJEITOS	PALAVRAS	TOTAL TEXTOS	TOTAL PALAVRAS
NÍVEL 1	Texto 1	50	4306	137	10.907
	Texto 2	47	3084		
	Texto 3	40	3517		
NÍVEL 2	Texto 1	45	4836	140	13.820
	Texto 2	42	3753		
	Texto 3	53	5231		
NÍVEL 3	Texto 1	49	4823	143	14.315
	Texto 2	47	3718		
	Texto 3	47	5774		
NÍVEL 4	Texto 1	53	5700	124	13.371
	Texto 2	42	3588		
	Texto 3	29	4083		
NÍVEL 5	Texto 1	32	4743	86	12.127
	Texto 2	30	3372		
	Texto 3	24	4012		
NÍVEL 6	Texto 1	31	4096	77	10.711
	Texto 2	27	2876		
	Texto 3	19	3739		
NÍVEL 7	Texto 1	52	7364	127	18.458
	Texto 2	45	5274		
	Texto 3	30	5820		
NÍVEL 8	Texto 1	23	2969	68	9.901
	Texto 2	21	2041		
	Texto 3	24	4891		
TOTAL				902	103.610

A fim de melhor analisarmos os dados, organizamos o BELC de duas maneiras: uma com o corpus geral e outra com o corpus devidamente dividido em oito subcorpora correspondentes aos oito níveis de ensino.

Com relação à identificação dos textos, elaboramos um código contendo informações sobre os mesmos, que está posicionado acima de cada um deles ao longo do corpus, conforme o modelo a seguir:

NXGXSXXPXTX

Onde :

N = Nível, podendo ser 1, 2, 3, 4, 5, 6, 7 e 8.

G=Grupo, pois tivemos mais de um grupo de cada nível. As informações particulares de cada grupo (professor, horário, etc.) estão numa tabela em anexo (anexo 8).

S =Sujeito: o numero de identificação do sujeito no grupo, também conforme lista identificadora.⁷⁸

P = Proficiência, podendo ser B (básico), P (pré-intermediário, I(intermediário) ou A (avançado).

T = Texto, podendo ser Texto 1 (sobre informação pessoal), Texto 2 (sobre informação pessoal sobre outra pessoa) ou Texto 3 (sobre alguma viagem já feita).

Abaixo, podemos ver um texto do corpus devidamente identificado como exemplo:

N1G3S09PBT2

My boyfriend is XXXX, he live in Canoas in a house the my mother in law, he work in my house, is web designer, and don't like dancing, is vey beautifull and friend. He love computer and have difficult in sleep in the night. He like a dog, and he don't like a cat. Is married a 3 years, and l'am a two girlfriend in you life! He love childrens!

O código identifica o texto de nível 1, do grupo 3, do sujeito 09, cuja proficiência é Básica e o texto é sobre o tema 2. Como podemos ver, o texto é rico em material linguístico para análise, o que veremos nas seções seguintes.

⁷⁸ Estas listas não se encontram em anexo por razões óbvias de espaço. No entanto, disponibilizamos ao leitor um modelo no anexo 7.

Desta maneira, podemos dizer que o BELC é constituído de 102.708 palavras (103.610 palavras menos 902 códigos identificadores acima de cada texto, que são considerados palavras na contagem do computador).

4.2.3 Panorama Geral do Corpus

Nesta seção, podemos ter uma noção geral sobre o BELC a partir de uma análise bastante preliminar dos textos. Após a leitura de todo o corpus, foi feito um levantamento das principais características observadas nos oito níveis do curso.

Salientamos que, como os alunos não podiam fazer uso de dicionários e nem de corretores digitalizados na redação dos textos, há vários erros ortográficos e palavras em português entre aspas ao longo dos dados, principalmente do corpus do nível 1. Acreditamos que, desta maneira, asseguramos maior autenticidade aos textos.

Sendo assim, faremos, primeiramente, algumas observações gerais organizadas por nível. Ao lermos os corpora, pudemos imediatamente observar:

Nível 1

- a. Muitos erros ortográficos, que não foram e nem serão alterados para garantir a autenticidade do corpus.
- b. Parágrafos curtos, padronizados, básicos, sem muita variação.
- c. Emprego errôneo de verbos, por exemplo, *I have twenty years old*, ou *I'm study Biology*, ou omissão de auxiliares, como *I studying*.
- d. Erros de conjugação incorreta, como *He have*, *She's have*.
- e. Omissão de conjugação no passado.
- f. Omissão do emprego do infinitivo verb+ING ou TO+verb depois de verbos como *like*, *hate*, etc.
- g. Emprego errôneo de preposições em geral, não somente *IN*, *ON* e *AT*, mas também *OF*.
- h. Uso indevido do artigo *THE* em geral.
- i. Confusão nos pronomes sujeito e adjetivos possessivos, como *I have a brother*, *your name is...*; *Her is a student*.

Nível 2

- a. Diminuição considerável de erros ortográficos ou uso de palavras em português.
- b. Parágrafos maiores, textos mais consistentes e mais bem estruturados. Os principais erros observados são de natureza verbal.
- c. Confusão no emprego de sufixos para conjugação verbal, como *He is study, He like, She study psychology and working at...*
- d. Alternância no uso de flexões do presente, como, num mesmo texto, *He study, he loves, we meet*, e mesmo no passado, como *I went to.../ eat, I went to... but I don't like*.
- e. Confusão no emprego de verbos auxiliares em frases negativas, como *doesn't is, she haven't brothers, don't have visit yet, I'm not went*.
- f. Generalização do sufixo *ed* para marcação de conjugação no passado de verbos irregulares, como *speaked, slepted, knowed, eated, falled, swimmmed, drinked, seed* ou conjugação dupla como *founded*.
- g. Uso de estruturas de presente simples para conjugação no passado, como *we goes there, my brothere doesn' go together* (para *didn't go together*), *we takes an airplane, we going of car*.
- h. Uso de auxiliar no presente e verbo conjugado no passado, como *we don't liked*.
- i. Uso de auxiliar *BE* para marcar passado, como *We were fun a lot, I was dance, I were play tennis*.
- j. Confusão de formas verbais, como o uso de *gone* para passado ao invés de *went*.

Nível 3

- a. A proficiência da maioria dos alunos sobe para o nível P (pré-intermediário).
- b. Em geral, observamos uma menor quantidade de erros gramaticais e ortográficos, e falta de vocabulário. No entanto, estes problemas ainda existem e não estão somente associados a sujeitos com proficiência B.
- c. Emprego ainda errado de verbos como *I have 29 years old*.
- d. Ausência da marcação de infinitivo em verbos que seguem “*like, love, hate*”, como *I like play, He likes read*.

- e. Incidência de erros preposicionais, como *in the PUC, I am a sudent Chemistry* (sem o *of*).
- f. Ausência de flexão em casos como *He like, he live, she study*, ou em frases negativas como *He don't like, He don't live*, ou ainda a omissão de COP⁷⁹ como em *She studying*, ainda ocorrem, mas com frequência bem mais baixa.
- g. Erros com relação à ordem de adjetivos, como *She is a person nice, ... who is a person very important to me, some places interesting, the stores have items very cheaps and greats*.
- h. Uso indevido do intensificador *very*, como *I like very her, I very love my sister*.
- i. Ausência de flexão no AUX passado, como *He don't went*.
- j. Alternância de flexão nos verbos no passado num mesmo texto, como "*We went to..., we are only four days there, we does not know the places*".
- k. Flexões anômalas no passado, como *I standed up early, we eatan, we telled, maked*.
- l. Uso de BE no passado para marcar conjugação, como em *I and my friends was go out, we was visit*.

Nível 4

- a. Nível de proficiência: quase todos são pré-intermediário, alguns intermediário, com poucas exceções de básico.
- b. Notável redução de erros gramaticais mais elementares: menos ocorrências de exemplos como *I have 20 years old, I likes, He lives with your girlfriend, I learned with he, I like him family, but I don't see their usually, with him girlfriend*.
- c. Omissão de flexão em terceira pessoa.
- d. Estruturas inapropriadas que são possivelmente derivadas da interferência do PB, como *Like of, think to, I will be of vacation; It's impossible to count the story*.

⁷⁹ COP= copula, verbo de ligação BE.

- e. Erros referentes à falta de padrões verbais: *Love go out, I liked play; Started work, learned fly.*
- f. Alternância de formas no presente e no passado para descrever tempo passado, como “*I went to São Vicente do Sul, we go to the center*” são muito frequentes.
- g. Formas anômalas no passado, como *I knowed; I don't liked; I eaten, I gone, buyed, That we were go to Floripa if.*
- h. *I do my things, like: do the ironong, watch TV, “fly is the best thing”.*

Nível 5

- a. Ainda observamos a ocorrência de erros ortográficos.
- b. Erros gramaticais básicos, ainda que em menor frequência, continuam ocorrendo, como *I am not have, She is don't stop on weekends.*
- c. Erros referentes a padrões verbais, como omissão do TO em *I want get married, I never want come back, I hope come back; I hope some day live in this country.*
- d. Flexão inapropriada no passado para exprimir tempo presente, como em *I went to PUC by train, I sudied everyday; when I arrived, I have dinner with my mother;* e uso de formas anômalas, como em *The man who drived us, I swimed in the sea.*
- e. Começam a aparecer erros mais “sofisticados”, digamos assim, ou seja, erros de inapropriação lexical/sintática/semântica que já não são tão básicos, como *This difficult that I have times to go out,* para “*this makes it difficult for me to have free time*”.
- f. Omissão ou excesso do *it*: *In my opinon is all a question of, because is a calm an interesting place, I go to the disco beacause I like and talk to my friends; Because the city it's beautiful.*
- g. Erros preposicionais mais evidentes, como *In Monday, on PUC, in the weekend, at Saturdays, on September.*
- h. Persistência, ainda, do uso do verbo *have* para expressar erroneamente idade, como *I have 20 years old, she have 19 years old.*
- i. *Always I can, I too work in the afternoon.*

- j. Uso inapropriado na expressão *used to* (no presente): *We use go out, we use to do things on the weekends.*
- k. Confusão com formas pronominais: *She lives with shes parents, have he as a, she name is, I love she.*
- l. Flexão inapropriada em terceira pessoa ainda permanece: *She like, she have a lot of friends, He study, when she aren't.*
- m. Erros possivelmente oriundos de interferência do PB: *I can know many countries; I born.*
- n. *I go to talk about, We pass a lot of time together, to pass the summer vacation.*
- o. Uso inapropriado de *Present Perfect*: *The best trip I have ever gone; I have never seen snow before; It rains two days.*

Nível 6

- a. Erros preposicionais: *At Friday; at next august; in the same street; on UFPR; in the same day; on Paris; on winter; on the morning; in the beach; at 5 years ago; ou ainda I worry so much with; fascinated for.*
- b. Confusões com formas verbais: *Hoping that they were hire me; I borned in Lajeado; Lented; leaved; feeled; we gone; rided; It didn't happened; didn't wanted.*
- c. Uso inapropriado do *have* para expressar idade é bem menos frequente, mas ainda ocorre: *I have 20 years; I have 21 years old; when I had 4 years.*
- d. Confusão com o uso de *other*: *In other state; in other country; another countries.*
- e. Ausência do ING após *like/love* é bem menor, mas ainda ocorre: *I love go out; I love drink; I love play tennis* e uso inapropriado de *Ing* com *like*: *I wasn't liking; I'm liking.*⁸⁰
- f. Confusão com adjetivos no plural: *Differents people; individuals project; differents things* e marcação de plural em substantivos não contáveis: *Listen to musics, other stuffs, a dark hair.*

⁸⁰ Salientamos que formas como *be + liking* são consideradas inapropriadas com relação à forma padrão do inglês, mas ocorrem em inglês americano, como evidenciado no COCA (*Corpus of Contemporary American English*) pelo endereço <<http://www.americancorpus.org>>.

- g. Ausência de flexão em terceira pessoa, como em *My mother always complain; It make me happy; Brazil don't have; he don't study; she have a boyfriend; He have worked; He have ever seen; Everybody say; everybody like.*
- h. Confusão com formas pronominais, como em *He opened your business; something that does we better; spent more money than I.*
- i. Ausência de flexão ING após preposição: *Before go to bed; consist in wake up; very interested in learn; a city for buy.*
- j. Exemplos de erros que evidenciam dificuldades na aplicação correta de *Present Perfect*: *We are together since 2006; When I have been there, because has had a lot of time I didn't saw her.*
- k. Erros variados que parecem ser oriundos de interferência do PB: *We lunch together; He went study; She seems doesn't like to study; I'll do a turn in my life; He actually is living; I'm actually coursing; I pretend to travel in the future.*

Nível 7

- a. Erros preposicionais, como *on week; in january 22; in the same day of my birthday; on puhrs; on the PUCRS; on next semester; on 2004.*
- b. Uso inapropriado do verbo *have* para expressar idade bastante frequente, como *I have 24 years, She has 22 years old, My boyfriend has 24 years old; who is making 21 years; My father have 49 years old.*
- c. Omissão do expletivo *it* como sujeito e objeto: *because is very important; because is very rainy; but some times is pretty hard; I enjoyed very much; I'm enjoying too; but is very boring;*
- d. Padrões verbais inapropriados, como *I like eat; she likes do; He like knew and learn; i'm looking forward to visit them.*
- e. Ausência de flexão em 3ª pessoa ainda observada: *it don't have importance; He don't want; he always watch; He study.*
- f. Confusões com tempos verbais, como em: *in my free time I went to Shopping Mall; my routine is wake up at seven o'clock and goes to my stage, if I had more time, I will tell you; We are together for two years.*

- g. Confusão com formas verbais, como em *I was borned; they still living in camaqua; hadn't think; try to found; She has borned; I knowed Italy, I usually goes; we talks; I don't counted; I was not realized; I don't be sad.*
- h. Problemas no uso da voz passiva, como *I have a cat name Lara; My best Friend Call Felipe; a huge company, what's call UNIVIAS; we have a dog calls Thors.*
- i. Erros variados provavelmente oriundos da interferência do PB, como *friday's afternoon to saturday's morning; I suggest everybody to know this places; I go surfing always I can; I dream in travel; where I feel me better; I spend a lot of money in cds; I'm actually coursing; I like a lot of go out; my mom is biologist.*
- j. Erros variados possivelmente oriundos de supergeneralizações, como *I am older from them; I very agree with many things he said, and it seems the distance not to exist; I wanna that we can stay together.*

Nível 8

- a. Como já é esperado, observamos poucas ocorrências de erros mais elementares de natureza variada, talvez até únicas, como *I love study, She likes read; I selled; I am proud of they; I would like continue; I'm planning going; didn't got selected; he dance flamenco, he never forget; he don't like to vacth much TV; were borned, when is raining; if had money, I went.*
- b. No entanto, observamos nesse corpus uma grande ocorrência de erros (i) com ausência de artigo, como em *I work as public employee, I was portuguese teacher; My mother is housewife; she's great person; He was dentist; who is musician* e (ii) com ausência de ING após a preposição, como em *for study; for show in class, for go to the work, His plans is go to the beach*, erros já não tão esperados para nível 8.
- c. Problemas com falsos cognatos: *after that I pretend to start; some parents came too.*

- d. Pequenas confusões com preposições: *on this last months, On my free times; In the last day, In this weekend.*
- e. Uso indevido de pronomes relativos, como em *I have a boyfriend, which I love very much; I have a girlfriend which I really love!*
- f. Confusão com tempos perfeitos, como a omissão em *We are married since 2003*, a falta da flexão completa: *my relationships had turn better*, mas principalmente a falta de conjugação *Past Perfect* ou uso inapropriado dele, como em *I have never gone to the beach and I wanted to see the sea*, ou *I had visited...* sem nenhuma outra referência no passado.
- g. Ocorrência de erros apropriados ao nível, como *despite of not having enough; but how I was not knowing; but the efforts worth a lot.*

Podemos concluir que os exemplos são bastante interessantes e remetem a várias ideias sobre a evolução dos aprendizes no processo aquisição/aprendizagem de L2. Na seção seguinte, veremos algumas perspectivas de análise de estudo para este corpus.

4.2.4 Procedimentos de Análise do BELC

4.2.4.1 Considerações gerais sobre a etiquetagem

Na seção anterior (4.1), pudemos ver considerações de caráter mais teórico sobre alguns procedimentos referentes à etiquetagem de corpora de aprendizes. A seguir, será feita uma descrição sobre como procederemos com relação à etiquetagem do BELC.

Como já foi dito anteriormente, o BELC, na condição de *learner corpus*, não deve se limitar à pesquisa desta tese. Em outras palavras, a análise que conduziremos nesta pesquisa não deve ser a única na qual ele será utilizado como *learner corpora*, em longo prazo. Portanto, para a proposta da nossa pesquisa, optamos por limitar o procedimento de etiquetagem a partir das categorias que

estamos correntemente analisando, que são os quatorze (14) morfemas selecionados a partir dos estudos sobre a ordem de aquisição dos morfemas.

4.2.4.1.1 Procedimentos de etiquetagem

O BELC foi etiquetado manualmente por mim, a autora da pesquisa, com o auxílio de uma professora devidamente orientada para a tarefa e qualificada para isso, graduada em Letras – Inglês, com ampla experiência em pesquisa.

A análise para a etiquetagem foi realizada a partir da leitura do corpus e o preenchimento de tabelas formuladas especificamente para esta pesquisa.

4.2.4.1.2 Material

Foram elaboradas quinze tabelas: uma genérica, que contém os códigos de cada morfema e outras quatorze, uma para cada morfema que estamos analisando, conforme reprodução abaixo:

Tabela 17: Códigos dos morfemas para a análise

CÓDIGO	MORFEMA
PP	<i>Present Progressive (V-ING)</i>
A	<i>Article (A, AN, THE)</i>
CC	<i>Contracted Copula (AM, IS, ARE)</i>
CA	<i>Contracted Auxiliary (am, is, are)</i>
PI	<i>Past Irregular</i>
P	<i>Possessive ('s)</i>
PR	<i>Past Regular</i>
UC	<i>Uncontracted Copula (am, is, are)</i>
TPR	<i>Third Person Sing Reg(+s)</i>
TPI	<i>Third Person Irregular (has)</i>
PREP	<i>In, on (PREP)</i>
PL	<i>Plural (s)</i>
UA	<i>Uncontracted Aux (am, is, are)</i>

A partir desta tabela, temos limitado o número e o tipo de morfemas que analisaremos nesta pesquisa. Por razões de ordem prática, manteremos estes

morfemas como foco, uma vez que a proposta do estudo desta tese é estabelecer uma comparação entre a ordem natural de aquisição dos morfemas a partir dos estudos sobre eles realizados (cf. capítulos 2 e 3), por mais interessantes que sejam as outras estruturas produzidas pelos aprendizes.

Tabela 18: Tabela sobre o morfema TPR - *Third Person Regular (+s)*

SOC (<i>Supp Oblig Context</i>)				
NÍVEL	<i>Supplied (S)</i>	<i>Absent (A)</i>	<i>Overuse(OU)</i>	Total Ocorrências
1				
2				
3				
4				
5				
6				
7				
8				

Seguindo o mesmo modelo da tabela acima, possuímos outras doze, uma para cada morfema analisado, também conforme códigos da tabela 17.

4.2.4.1.3 Critérios de análise

Como pudemos ver ao longo do capítulo 2, a pesquisa de Brown (1973) é referência nos estudos sobre a ordem de aquisição dos morfemas, sejam eles como L1 ou L2. Da mesma forma, utilizamos várias ideias de Brown (1973) nesta tese, inclusive o critério utilizado para a análise dos dados.

O critério *Suppliance in Obligatory Context*, o SOC, ou contexto de preenchimento obrigatório, foi estabelecido por Brown e pressupõe que, para a análise de como o aprendiz procedeu, o ponto de partida seja o contexto onde o morfema é esperado e, então, se foi preenchido ou não. Ou seja, não são considerados somente os contextos morfossintáticos onde um dado morfema foi corretamente empregado, mas também onde ele não foi e deveria ter sido.

Segundo Ellis & Barkhuizen (2005):

Obligatory Occasion Analysis (BROWN, 1973) constitutes a method for examining how accurately learners use specific (usually grammatical)

features. Like EA [error analysis], it involves a comparison between the forms used by learners and the target language norms... (p. 74).

Desta maneira, este foi o critério escolhido para analisarmos os dados do BELC. Como o propósito é estabelecer um panorama sobre os padrões evolutivos dos aprendizes brasileiros de L2 acerca da ordem de aquisição dos morfemas por eles evidenciado nos textos de produção espontânea que constituem o BELC, optamos por utilizar a mesma metodologia da grande maioria das pesquisas semelhantes (cf. capítulos 2 e 3).

Sendo assim, consideraremos o contexto onde o morfema é obrigatório como ponto de partida para o registro, de modo que marcamos:

* Em S (*supplied*) se o morfema tiver sido empregado corretamente.

* Em A (*absent*) se o morfema não tiver sido empregado num contexto onde deveria ter sido.

* Em OU (*overuse*) se o morfema tiver sido empregado de forma superestendida, como, por exemplo, she *wented*.

Alguns autores como Izumi & Isahara (2004), Goldschneider & DeKeyser (2005) e Kwon (2005) mencionaram que uma das críticas aos resultados de algumas pesquisas sobre os estudos dos morfemas, algumas inclusive conduzidas por eles, é que este critério não leva em conta os erros de *overuse* do morfema, isto é, os erros referentes ao emprego do morfema em contextos supergeneralizados, uma vez que o emprego do morfema mesmo em contextos errados pode ser um indicativo de incorporação e evolução para a aquisição. Nesta pesquisa, consideraremos as três possibilidades de procedimento dos alunos na análise e registro do corpus BELC.⁸¹ Partimos a seguir para uma ilustração mais pontual de como estes registros foram feitos.

⁸¹ Como veremos a seguir na seção 4.2.4.1.3, para os cálculos específicos desta pesquisa, serão considerados somente o preenchimento correto no contexto obrigatório e o total de ocorrências. O não-preenchimento e o preenchimento incorreto serão registrados, mas não serão utilizados na fórmula do cálculo.

4.2.4.1.4 Procedimentos específicos

Com o objetivo de obtermos dados numéricos com relação à ocorrência dos morfemas, analisamos o BELC da seguinte maneira: num primeiro momento, realizamos a leitura de um determinado parágrafo na íntegra, a fim de termos um panorama sobre o texto do aprendiz. Num segundo momento, passamos à etiquetagem das estruturas-alvo, isto é, os morfemas que o texto dispõe e seguimos os seguintes passos:

1. Identificação do contexto de preenchimento obrigatório do morfema.
2. Identificação do procedimento do aprendiz diante deste contexto, dentro das alternativas: preenchimento obrigatório (*supplied*), não preenchimento (*absent* ou ausência) do morfema, preenchimento superestendido (*overuse*).
3. Etiquetagem do item acima no corpus, conforme código do morfema e código do procedimento do aluno. (cf. tabelas acima e ilustração a seguir).
4. Registro do procedimento do aprendiz na tabela apropriada do morfema.

Por exemplo, a partir do texto:

N1G3S09PBT2

My boyfriend is XXXX, he live in Canoas in a house the my mother in law, he work in my house, is web designer, and don't like dancing, is vey beautifull and friend. He love computer and have difficult in sleep in the night. He like a dog, and he don't like a cat. Is married a 3 years, and I am a two girlfriend in you life! He love childrens!

Procedemos da seguinte forma:

1. Leitura do parágrafo e identificação dos morfemas envolvidos.
2. Leitura mais específica e identificação do procedimento do aprendiz.
3. Etiquetagem no corpus, conforme códigos da tabela 18, como ilustramos abaixo:

My boyfriend is (UC-S) XXXX, he live (TPR – A) in (PREP-S) Canoas in (PREP-S) a (ART-A) house the my mother in law, he work (TPR – A) in (PREP-S) my house, is (A-S) web designer, and don't like dancing, is (UC-S) vey beautifull and friend. He love (TPR – A)computer and have (TPI – A) difficult in sleep in the night(PREP-A). He like (TPR-A) a dog, and he don't like a cat. Is (UC-S) married a 3 years, and I am (UC-S) a two girlfriend in you life! He love (TPR – A) childrens (PL-OU)!

Sendo assim, segundo a ilustração de etiquetagem acima, podemos entender que só foram analisados os morfemas que estamos considerando nesta pesquisa, a partir de como o aprendiz procedeu. Como por exemplo:

UC – S: *Uncontracted Copula - Supplied*

TPR – A: *Third Person Regular – Absent*

PREP – S: *Preposition – Supplied*

PL – OU: *Plural- Overuse*

E assim por diante.

4. Registro do procedimento do aprendiz nas tabelas apropriadas.

De maneira mais específica, após a leitura e etiquetagem de cada texto, registramos manualmente⁸² o número de ocorrências e procedimentos do aprendiz de cada morfema na tabela apropriada. Por exemplo, a partir do texto ilustrativo acima, registramos na tabela TPR:

⁸² O trabalho de leitura e etiquetagem de 902 textos foi, sabidamente, árduo. Apesar disso, optamos por fazê-lo manualmente para garantir a veracidade dos dados. Mesmo se houvesse um software disponível perfeitamente programado para etiquetar os morfemas selecionados para esta pesquisa, ainda assim, poderíamos pensar que a análise com relação ao critério SOC poderia apresentar falhas. Isto é, o software poderia etiquetar os morfemas nas suas ocorrências, apropriadas ou não, mas o critério que estamos utilizando no cálculo contempla também as não ocorrências. E seria muito difícil estabelecermos todos esses contextos para treze (13) morfemas. Por isso, preferimos a etiquetagem manual para os fins desta pesquisa.

Tabela 19: Tabela sobre o morfema TPR - *Third Person Regular (+s)* com registro

SOC (<i>Supp Oblig Context</i>)				
NIVEL	<i>Supplied (S)</i>	<i>Absent (A)</i>	<i>Overuse (OU)</i>	Total Ocorrências
1		5		
2				
3				
4				
5				
6				
7				
8				

Assim, ao final da correção de cada nível, temos os números com relação às frequências e aí procedemos aos cálculos.

4.2.4.1.5 Análise quantitativa

Vários trabalhos foram considerados na decisão de como calcular os dados deste corpus (principalmente BROWN (1973), GOLDSCHNEIDER & DEKEYSER (2005), PICA (1983)), mas, a partir das considerações de Ellis & Barkhuizen (2005), a fórmula escolhida⁸³ foi:

$\frac{\text{Número de preenchimento correto}}{\text{Total de contextos obrigatórios}} \times 100 = \text{acurácia (em porcentagem)}$

Ao final dos procedimentos de etiquetagem e registro dos morfemas no corpus BELC, temos um número X de porcentagem do uso correto de morfemas pelos sujeitos dos oito níveis de aprendizagem que investigamos. E assim, partimos para a análise dos dados propriamente dita, quando temos condições de estabelecer um padrão evolutivo sobre a ordem de aquisição dos morfemas dos aprendizes que estamos pesquisando.

⁸³ Tal fórmula foi escolhida por razões de ordem eminentemente práticas, em detrimento a outras que consideravam o uso superestendido do morfema (cf. PICA, 1984, e ELLIS & BARKHUIZEN (2005)).

Para isso, há outra questão relevante: quando considerar um morfema “adquirido/aprendido?” A discussão sobre como determinar a aquisição de um morfema pode ser dividida em duas linhas de pensamento.

Numa primeira perspectiva, Brown (1973) estipula a frequência de 90% de correção como padrão para considerar um morfema adquirido. E estipula ainda que tal morfema só pode ser considerado “adquirido” quando a porcentagem de 90% de correção for mantida três vezes consecutivas (no caso de Brown, em três coletas consecutivas, em nosso caso, em três níveis). Isto porque Brown observou que, por vezes, a porcentagem de correção atingia os 90% e baixava de uma coleta para outra, o que pode ser um indicativo do estágio de aprendizagem da curva em U. Como sabemos, os aprendizes evidenciam a forma correta num primeiro momento como resultado de imitação ou reprodução automática (por exemplo, verbos irregulares no passado, como *went*), num segundo momento, utilizam outras estruturas alternativas (como *goed*) e, num terceiro momento, utilizam a forma correta já incorporada (*went*).

Numa segunda perspectiva, conforme o que foi inicialmente proposto por De Villiers & De Villiers (1973), o ponto de corte é relativo ao conjunto de todos os dados. Isto é, como nos explicam melhor Ellis & Barkhuizen (2005), um morfema é considerado “adquirido” a partir de uma escala implicacional baseada em cálculos relativos sobre a ocorrência dos morfemas nos contextos obrigatórios. A escala implicacional é um procedimento que estabelece até que ponto uma ordem proposta de aquisição é estatisticamente válida e é baseado numa contagem dicotômica de um dado morfema como adquirido ou não adquirido. Dentre os passos básicos para estabelecer esta escala estão a construção de uma matriz onde os morfemas seriam classificados horizontalmente de “difícil” a “fácil” a partir do número de aprendizes que adquiriram cada morfema e verticalmente segundo o número de morfemas que cada aprendiz adquiriu. Consideramos esta escala muito complexa para ser aplicada ao nosso estudo, principalmente porque ela é feita considerando a performance de cada aprendiz para cada morfema e, na pesquisa que propomos, temos quatrocentos e vinte e quatro (424) sujeitos e treze morfemas.

Sendo assim, optamos por utilizar, nesta tese, a perspectiva de Brown (1973), que foi amplamente utilizada nos estudos dos morfemas, principalmente porque ela parece ser mais clara e objetiva. Além disso, dados a serem obtidos de

frequências relativas são mais difíceis de serem calculados e estão mais sujeitos a variáveis que podem influenciar os resultados finais.

No entanto, faremos uma “adaptação” à perspectiva de Brown ao nosso contexto. Utilizaremos a porcentagem de 90% de correção como ponto de corte para considerarmos um morfema adquirido sem que seja necessária a aparição deste número por três vezes consecutivas. Ou seja, a partir do momento que obtivermos 90% de correção no uso de um morfema, ele poderá ser considerado adquirido na nossa análise.⁸⁴ Julgamos este o critério mais adequado para a análise que propomos, uma vez que temos 424 sujeitos aprendizes de inglês como L2 (ao contrário de Brown, que analisou três sujeitos aprendizes de inglês como L1).

Ainda, a decisão por este critério para o ponto de corte é fundamentada em Ellis & Barkhuizen (2005):

Traditionally (see Brown 1973), the level is set at 90 percent. Thus, if a learner achieves an accuracy score of 90% or higher, the morpheme is considered “acquired” but if the score is lower (even 89%) it is deemed “not acquired”. Implicational scaling requires the analyst to score morphemes, on the basis of their accuracy scores as either 1 (acquired) or 0 (not acquired). The choice of 90% as the acquisition level is based on two assumptions; first, it corresponds to the level achieved by native speakers, who typically failed to perform at 100per cent accuracy themselves (i.e. they make “slips”) (p. 85).

Passamos, a seguir, ao capítulo que apresentará os resultados da pesquisa e a análise dos mesmos à luz dos pressupostos teóricos que desenvolvemos nos capítulos 2 e 3. Ou seja, a partir dos resultados obtidos, teremos condições de comparar os padrões evolutivos dos aprendizes brasileiros de inglês no que se refere à ordem natural na aquisição de morfemas em inglês – à sequência comumente entendida como padrão de desenvolvimento dos morfemas, a fim não somente de identificarmos semelhanças e diferenças entre as ordens, mas de tentarmos atribuir a elas alguma explicação, seja esta de ordem sintática (cf. ZOBL & LICERAS, 1994), seja ela de ordem multifatorial (cf. GOLDSCHNEIDER & DEKEYSER (2005).

⁸⁴ Ressaltamos, assim, que “adquirido”, aqui, será entendido como aquele morfema que obteve um alto grau de desempenho igual ou superior a 90%.

5 RESULTADOS E ANÁLISE

A partir da elaboração do corpus BELC, organizamos um banco de dados que pudesse nos oferecer informações específicas sobre os morfemas produzidos pelos sujeitos. Este capítulo apresenta os resultados obtidos e a análise destes dados. Na primeira seção, apresentamos os resultados através de tabelas e figuras, que podem ilustrar o padrão evolutivo com relação à aquisição de morfemas em inglês como L2 por aprendizes brasileiros. Na segunda seção, conduzimos a discussão desses resultados. Num primeiro momento, propomos a comparação da ordem de morfemas que obtivemos a partir do BELC com as sequências de aquisição dos estudos de referência – Brown (1973) em inglês como L1 e Krashen (1977) como L2. Num segundo momento, procuramos discutir possíveis explicações para a sequência que encontramos, baseados no que sugerem Zobl & Liceras (1994) e Goldschneider & DeKeyser (2005). Ao final do capítulo, discutimos brevemente sobre dois fatores: a instrução e a natureza da tarefa (escrita), considerando que tais fatores possam ter tido alguma interferência nos nossos resultados.

5.1 APRESENTAÇÃO DOS RESULTADOS

Como vimos no capítulo 4, mais especificamente na seção 4.2.4.1.4 sobre procedimentos de análise do BELC, os textos foram etiquetados para fornecerem informações sobre a aquisição dos morfemas que selecionamos para este estudo. Basicamente, a análise se deu em quatro etapas.⁸⁵

⁸⁵ Ressaltamos, mais uma vez, que toda a investigação foi feita e revisada manualmente, desde a etiquetagem do corpus até o preenchimento de tabelas e o cálculo, a fim de evitarmos variáveis associadas a erros da máquina.

⇒ Etapa 1: Levantamento de ocorrências

Obtivemos o número de ocorrências dos morfemas a partir do procedimento do aluno diante do contexto obrigatório (isto é, preenchimento do morfema correto, ausência do morfema ou preenchimento errado).

⇒ Etapa 2: Cálculo, feito a partir da fórmula:

$\frac{\text{Número de preenchimento correto}}{\text{Total de contextos obrigatórios}} \times 100 = \text{acurácia (em porcentagem)}$

⇒ Etapa 3: Identificação do estágio de aquisição⁸⁶

Quando tivemos uma porcentagem de correção no uso dos morfemas de 90%, consideramos o morfema adquirido.

Assim, passemos à apresentação dos resultados:

5.1.1 Ocorrência dos Morfemas

As tabelas e figuras que apresentaremos nesta seção mostram os resultados obtidos a partir da análise que conduzimos segundo as etapas que acabamos de descrever. Na tabela abaixo, podemos ver uma matriz de como fizemos isso: registramos a ocorrência dos morfemas no contexto onde eles seriam obrigatórios, classificando a manifestação do morfema neste contexto: *supplied* (preenchido), *absent* (ausente) e *overuse* (uso superestendido). A partir da leitura e da etiquetagem do corpus, obtivemos o número destas ocorrências e então calculamos a porcentagem de uso correto, dividindo o número de contextos preenchidos pelo número total de ocorrências, e multiplicamos este número por 100 (cf. cálculo explicitado acima na seção anterior).

⁸⁶ Como já mencionamos anteriormente no final do capítulo 4, a questão sobre quando considerar um morfema "adquirido" é bastante específica para cada estudo. Entenda-se, a partir disso, que, apesar do uso de termos como "estágios de aquisição" ou "morfema adquirido", estamos referindo aos dados de maneira específica ao nosso estudo, e não querendo estabelecer estágios definitivos no processo de aquisição de L2. Então, "adquirido" deve ser entendido aqui como dispendo de um alto grau de desempenho, o que consideraremos 90%.

Tabela 20: Tabela matriz

<i>SOC (Supp Oblig Context)</i>					
NÍVEL	<i>Supplied (S)</i>	<i>Absent (A)</i>	<i>Overuse(OU)</i>	Total Ocorrências	Porcentagem de correção
1					
2					
3					
4					
5					
6					
7					
8					

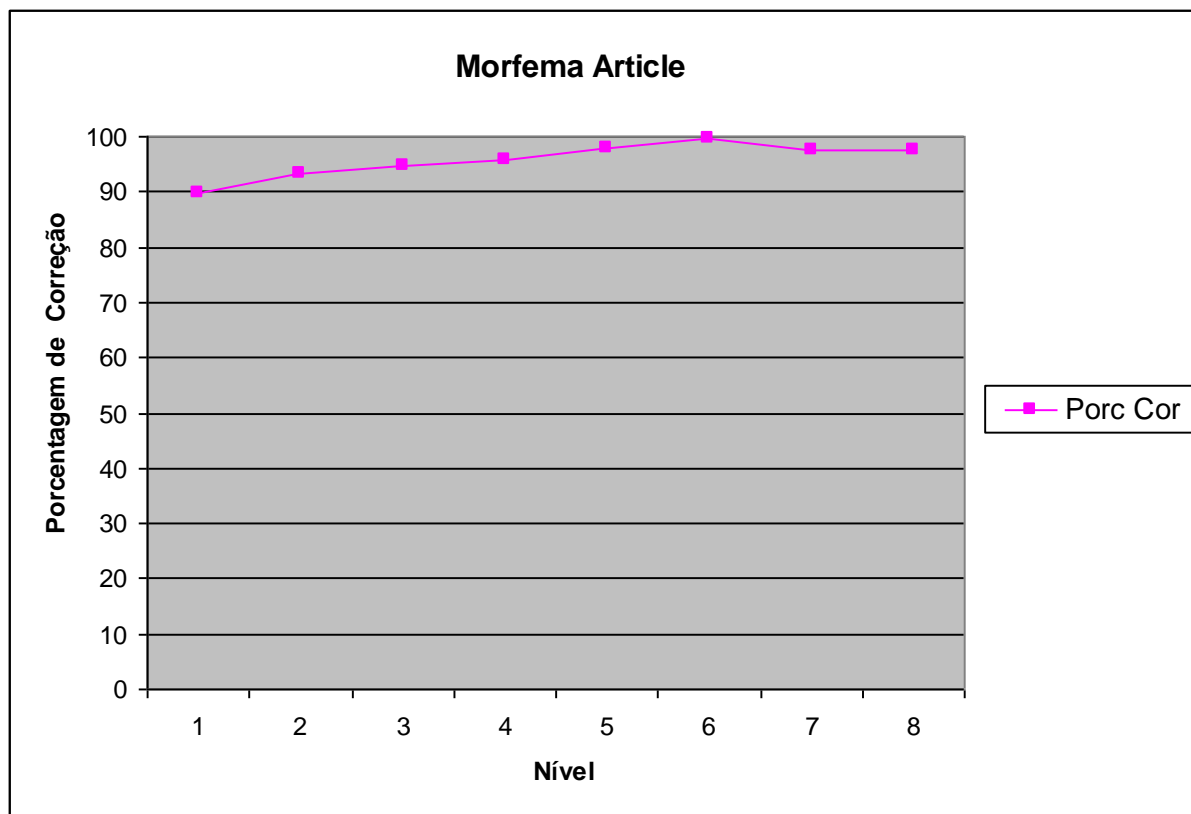
Assim, obtivemos treze⁸⁷ tabelas, as quais apresentamos abaixo, juntamente com as respectivas figuras, que poderão mostrar o padrão evolutivo de cada morfema a partir dos índices de correção em cada um dos níveis de aprendizagem investigados no BELC.⁸⁸

Tabela 21: Morfema Art (*a, the*)

<i>SOC (Supp Oblig Context)</i>					
NÍVEL	<i>Supplied (S)</i>	<i>Absent (A)</i>	<i>Overuse (OU)</i>	Total Ocorrências	Porcentagem de correção
1	278	32	0	310	89.6 %
2	369	26	0	395	93.4 %
3	476	26	0	502	94.8 %
4	460	21	0	481	95.6 %
5	451	10	0	461	97.8 %
6	444	2	0	446	99.5 %
7	923	24	0	947	97.4 %
8	411	10	0	421	97.6 %

⁸⁷ Os quatorze morfemas são aqui reduzidos a treze pois “*in*” e “*on*” são considerados juntamente na categoria preposição.

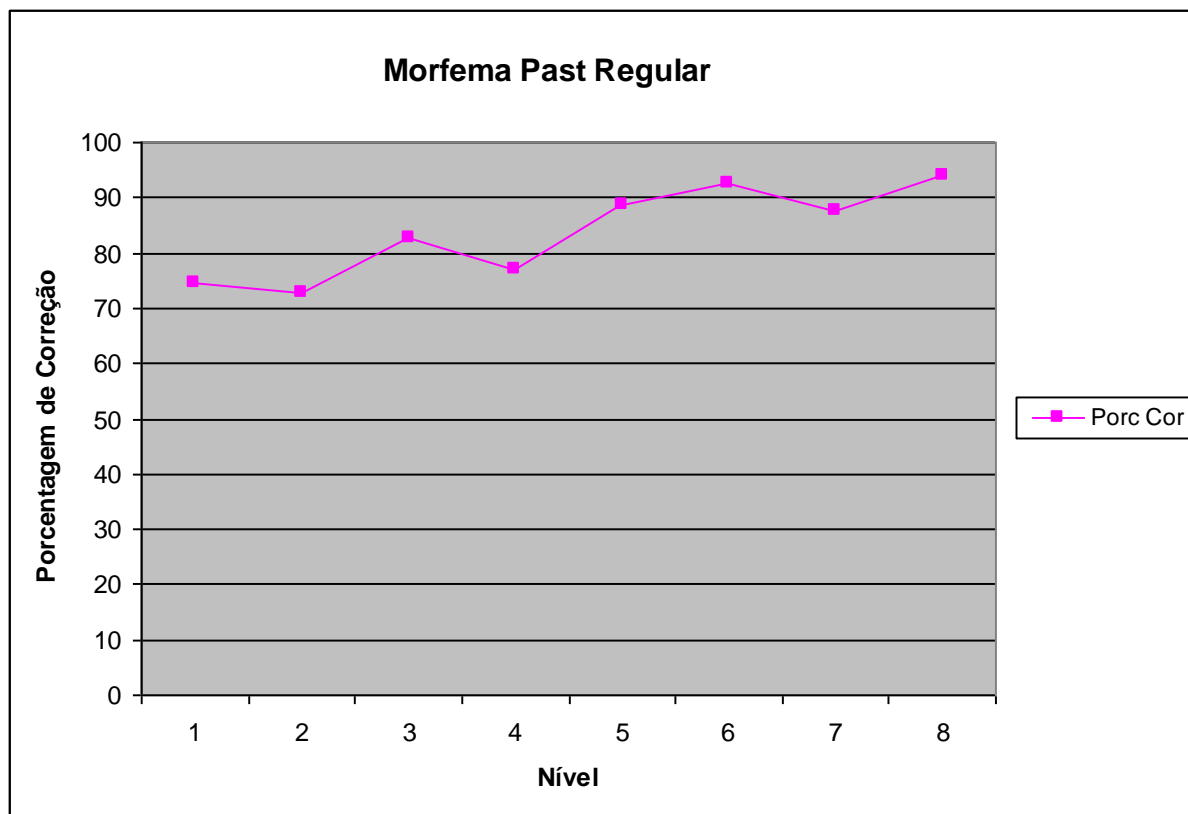
⁸⁸ O anexo 9 da tese disponibiliza amostras do BELC etiquetado, a fim de que o leitor possa entender melhor o que foi considerado em cada categoria para cada morfema.

Figura 2: Morfema *Article*

Como podemos ver, os dados com relação a estes morfemas são bastante claros: a porcentagem de correção já começa alta no primeiro nível: 89.6%. E é de 93,4% (portanto, acima de 90%) no nível 2. Assim, consideramos tal morfema adquirido neste nível. Após isso, os índices de correção se mantêm sempre acima de 90%, com um pico de 99.5% no nível 6, atingindo, no nível 8, o índice de 97.6%

Tabela 22: Morfema *Past Regular*

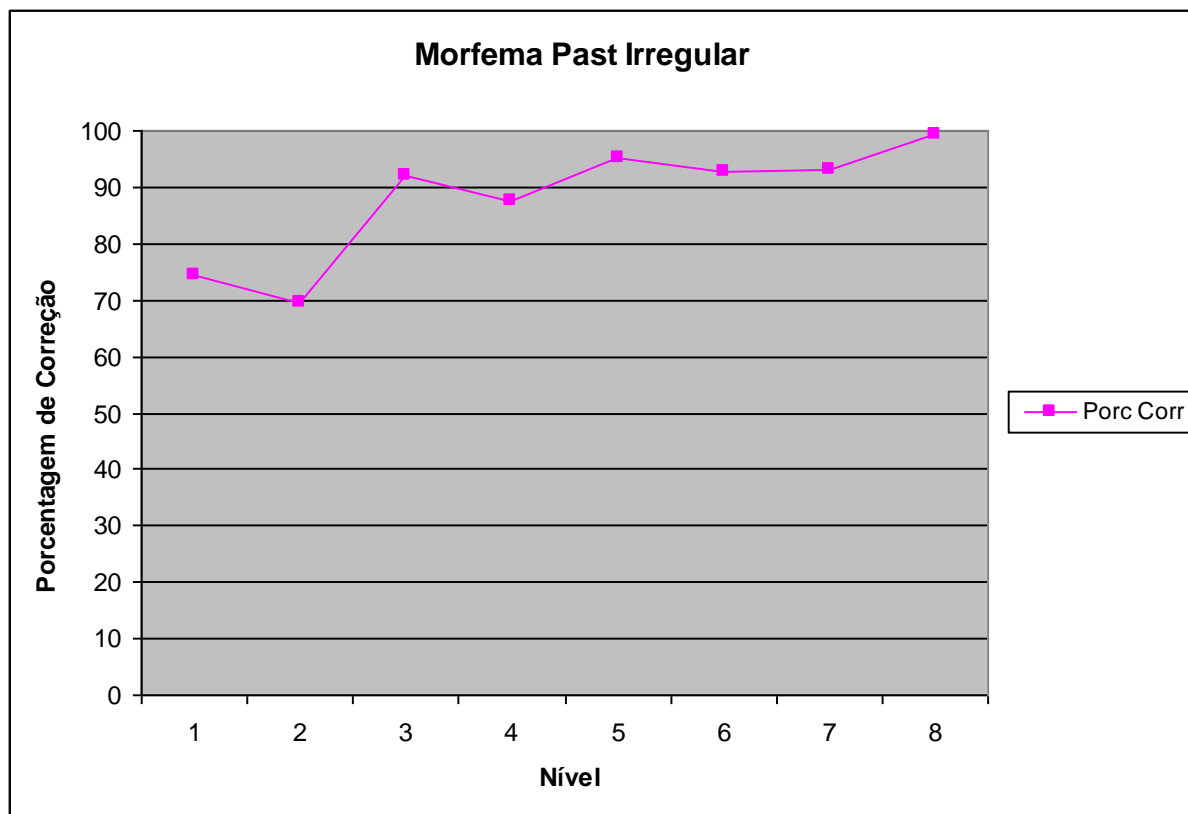
NÍVEL	SOC (Supp Oblig Context)			Total Ocorrências	Porcentagem de correção
	Supplied (S)	Absent (A)	Overuse (OU)		
1	76	25	1	102	74.5 %
2	110	37	4	151	72.8 %
3	139	29	0	168	82.7 %
4	102	29	1	132	77.2 %
5	119	15	0	134	88.8 %
6	103	8	0	111	92.7 %
7	182	24	2	208	87.5 %
8	62	4	0	66	93.9 %

Figura 3: Morfema *Past Regular*

Com relação ao morfema *Past Regular*, os dados são bastante interessantes: os sujeitos evidenciam um índice de 74.5% de acertos no nível 1. No nível 2, o índice cai para 72.8%, mas depois entra em elevação e atinge 92.7% no nível 6, fechando em 93.9% no nível 8.

Tabela 23: Morfema *Past Irregular*

<i>SOC (Supp Oblig Context)</i>					
NÍVEL	<i>Supplied (S)</i>	<i>Absent (A)</i>	<i>Overuse (OU)</i>	Total Ocorrências	Porcentagem de correção
1	193	66	0	259	74.5 %
2	224	98	0	322	69.5 %
3	334	29	0	363	92.0 %
4	232	33	0	265	87.5 %
5	341	17	0	358	95.2 %
6	233	18	0	251	92.8 %
7	389	28	0	417	93.2 %
8	181	1	0	182	99.4 %

Figura 4: Morfema *Past Irregular*

Já o morfema *Past Irregular* apresenta um comportamento um pouco diferente: os sujeitos apresentam uma porcentagem de correção de 74.5% no nível inicial, assim como o morfema *Past Regular*. No entanto, após uma queda de 69.5% no nível 2, o índice de correção atinge 92% no nível 3, se mantendo estável (com exceção do nível 4, onde o índice é de 87.5%) até o nível 8, atingindo lá a marca de 99.4%.

Tabela 24: Morfema *Present Progressive*

<i>SOC (Supp Oblig Context)</i>					
NÍVEL	<i>Supplied (S)</i>	<i>Absent (A)</i>	<i>Overuse (OU)</i>	Total Ocorrências	Porcentagem de correção
1	15	2	0	17	88.2 %
2	33	5	0	38	86.8 %
3	43	2	0	45	95.5 %
4	60	1	0	61	98.3 %
5	57	1	0	58	98.2 %
6	72	0	0	72	100 %
7	123	3	0	126	97.6 %
8	45	0	0	45	100 %

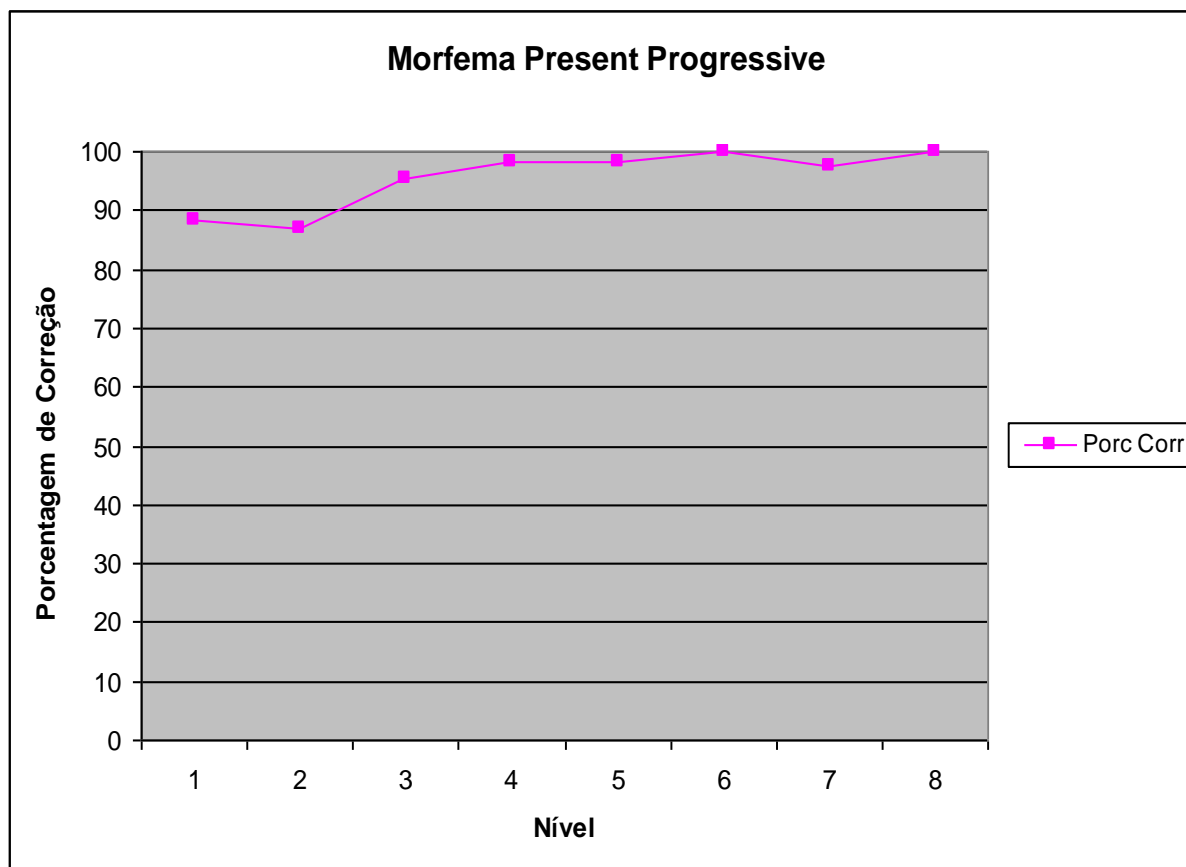


Figura 5: Morfema *Present Progressive*

Os índices de correção do morfema *Present Progressive* (-ING) são de 88.2% no nível 1. No nível 2, há um pequeno declínio para 86.8%, e já no nível 3, observamos o índice de 95.5%. No nível 6, a porcentagem de correção atinge 100%, e, após uma queda para 97.6% no nível 7, se mantém em 100% no nível 8.

Tabela 25: Morfema *Uncontracted Copula*

<i>SOC (Supp Oblig Context)</i>					
NÍVEL	<i>Supplied (S)</i>	<i>Absent (A)</i>	<i>Overuse (OU)</i>	Total Ocorrências	Porcentagem de correção
1	372	38	0	410	90.7 %
2	450	39	0	489	92 %
3	484	21	0	505	95.8 %
4	481	14	0	495	97.1 %
5	351	6	0	357	98.3 %
6	327	7	0	334	97.9 %
7	595	9	0	604	98.5 %
8	223	0	0	223	100 %

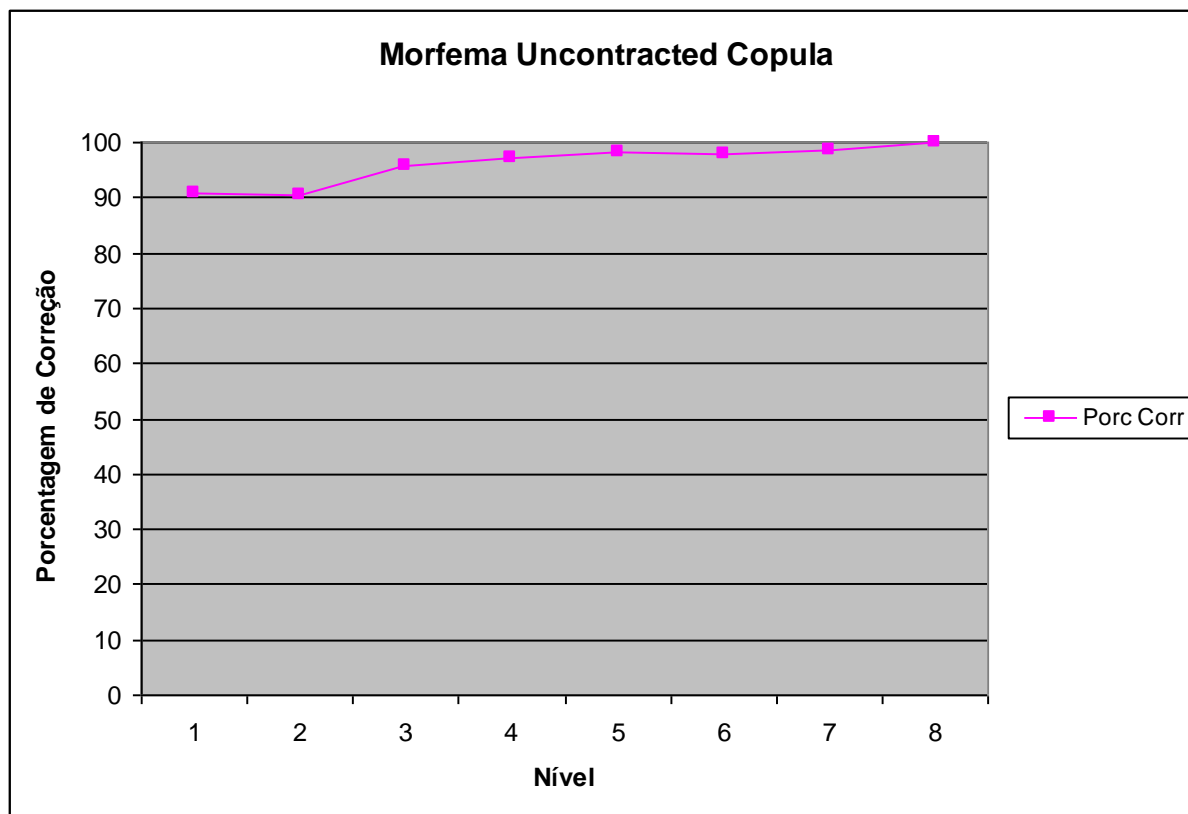


Figura 6: Morfema *Uncontracted Copula*

Os dados do morfema *Uncontracted Copula* nos mostram índices de correção acima de 90% desde o nível 1, se mantendo em ascensão até o nível 8, onde atinge 100%.

Tabela 26: Morfema *Contracted Copula*

<i>SOC (Supp Oblig Context)</i>					
NÍVEL	<i>Supplied (S)</i>	<i>Absent (A)</i>	<i>Overuse (OU)</i>	Total Ocorrências	Porcentagem de correção
1	92	2	0	94	97.8%
2	136	1	0	137	99.2%
3	116	1	0	117	99.1%
4	127	0	0	127	100%
5	94	1	0	95	98.9%
6	62	1	0	63	98.4%
7	164	0	0	164	100%
8	55	0	0	55	100%

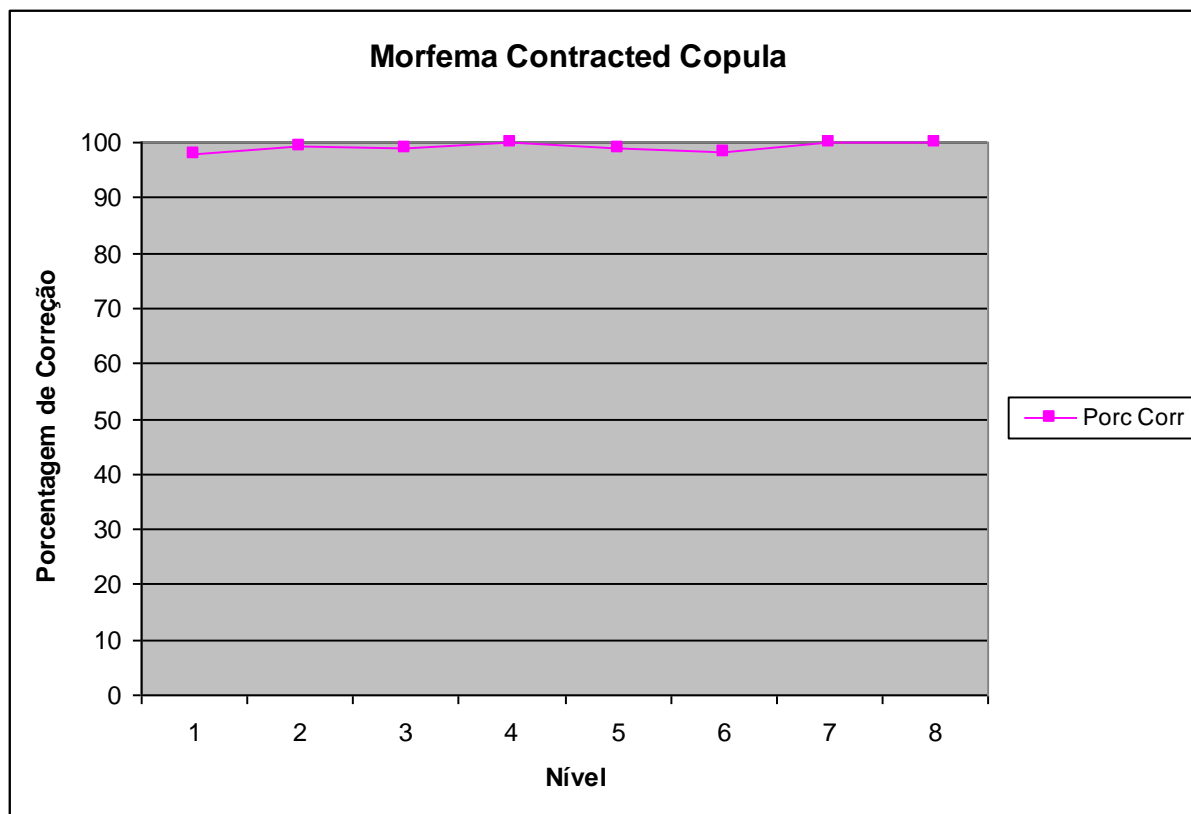
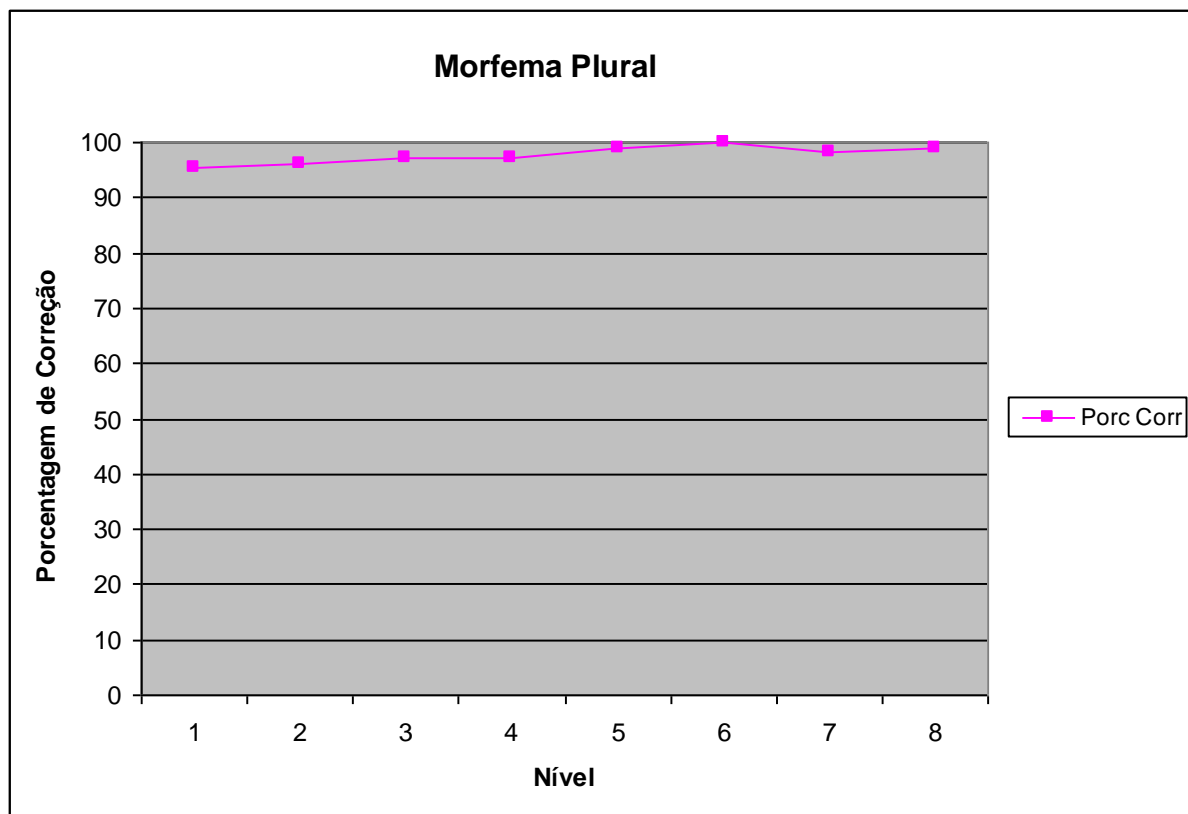


Figura 7: Morfema *Contracted Copula*

Da mesma forma que o morfema anterior, o morfema *Contracted Copula* também tem índices de correção altos desde o nível elementar, iniciando com 97.8% e atingindo a marca dos 100% duas vezes.

Tabela 27: Morfema *Plural*

<i>SOC (Supp Oblig Context)</i>					
NÍVEL	<i>Supplied (S)</i>	<i>Absent (A)</i>	<i>Overuse (OU)</i>	Total Ocorrências	Porcentagem de correção
1	374	11	15	400	93.5 %
2	457	3	16	476	96.0 %
3	430	7	6	443	97.0 %
4	408	6	5	419	97.3 %
5	339	4	0	343	98.8 %
6	320	0	0	320	100 %
7	589	9	0	598	98.4 %
8	236	2	0	238	99.1 %

Figura 8: Morfema *Plural*

O comportamento do morfema *Plural*, segundo os dados do BELC, mostra que os índices de correção estão sempre acima de 90%, atingindo os 100% já no nível 6, e, embora não mantenha esta porcentagem absoluta nos níveis 7 e 8, mantém-se muito próxima dela.

Tabela 28: Morfema *Third Person Singular Regular*

SOC (Supp Oblig Context)					
NÍVEL	<i>Supplied</i> (S)	<i>Absent</i> (A)	<i>Overuse</i> (OU)	Total Ocorrências	Porcentagem de correção
1	60	87	0	147	40.8 %
2	64	75	0	139	46.0 %
3	79	58	0	137	57.6 %
4	84	48	0	132	63.6 %
5	85	24	0	109	77.9 %
6	85	22	0	107	79.4 %
7	127	12	0	139	91.3 %
8	56	3	0	59	94.9 %

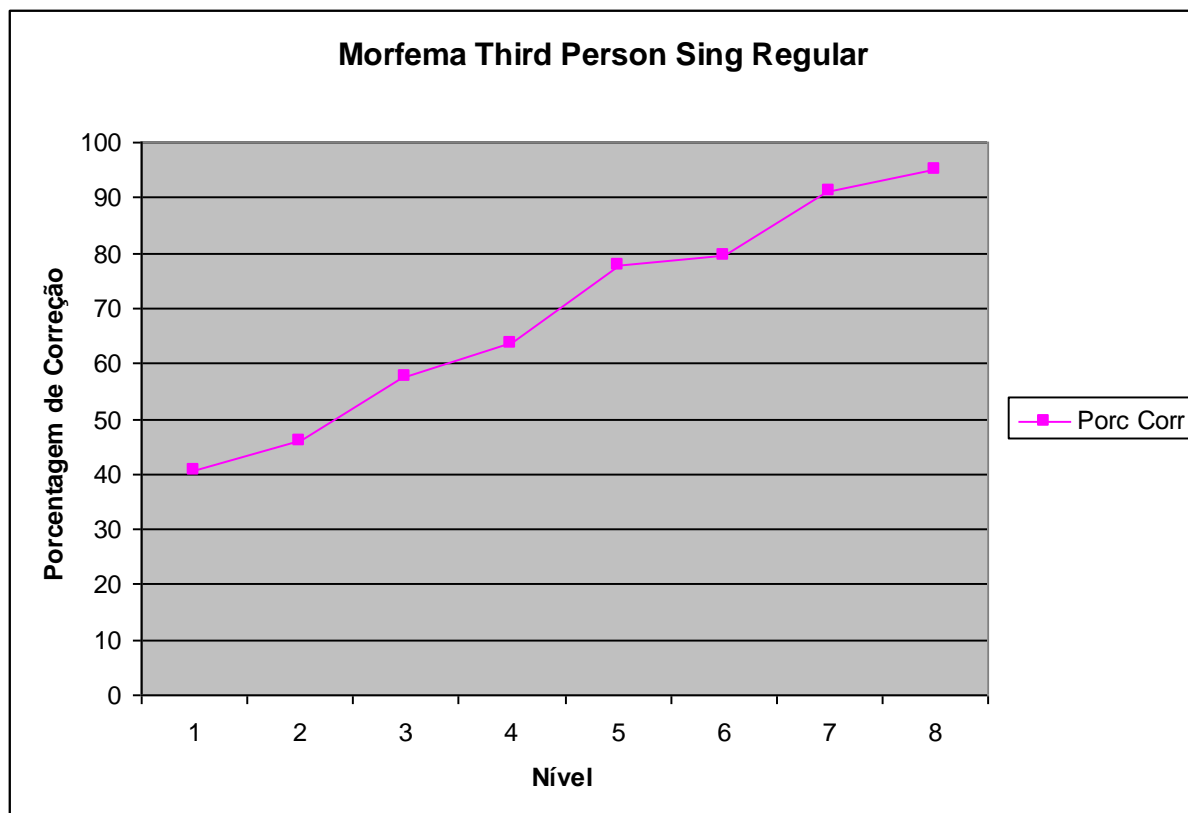


Figura 9: Morfema *Third Person Sing Regular*

Os dados referentes a este morfema são muito interessantes: o índice de correção no nível 1 é de 40.8%, se mantém baixo no nível 2, com 46%, mas está sempre em ascensão, atingindo 91.3% no nível 7. Um dado interessante é que a porcentagem de correção no nível 8 é de 94.9%, onde naturalmente poderíamos esperar 100% de acerto.

Tabela 29: Morfema *Third Person Irregular*

SOC (<i>Supp Oblig Context</i>)					
NÍVEL	<i>Supplied (s)</i>	<i>Absent (a)</i>	<i>Overuse (ou)</i>	Total Ocorrências	Porcentagem de correção
1	6	25	0	31	19.3 %
2	38	22	0	60	63.3 %
3	17	16	0	33	51.5 %
4	17	7	0	24	70.8 %
5	15	8	0	23	65.2 %
6	15	11	0	26	57.6 %
7	30	4	0	34	88.2 %
8	21	2	0	23	91.3 %

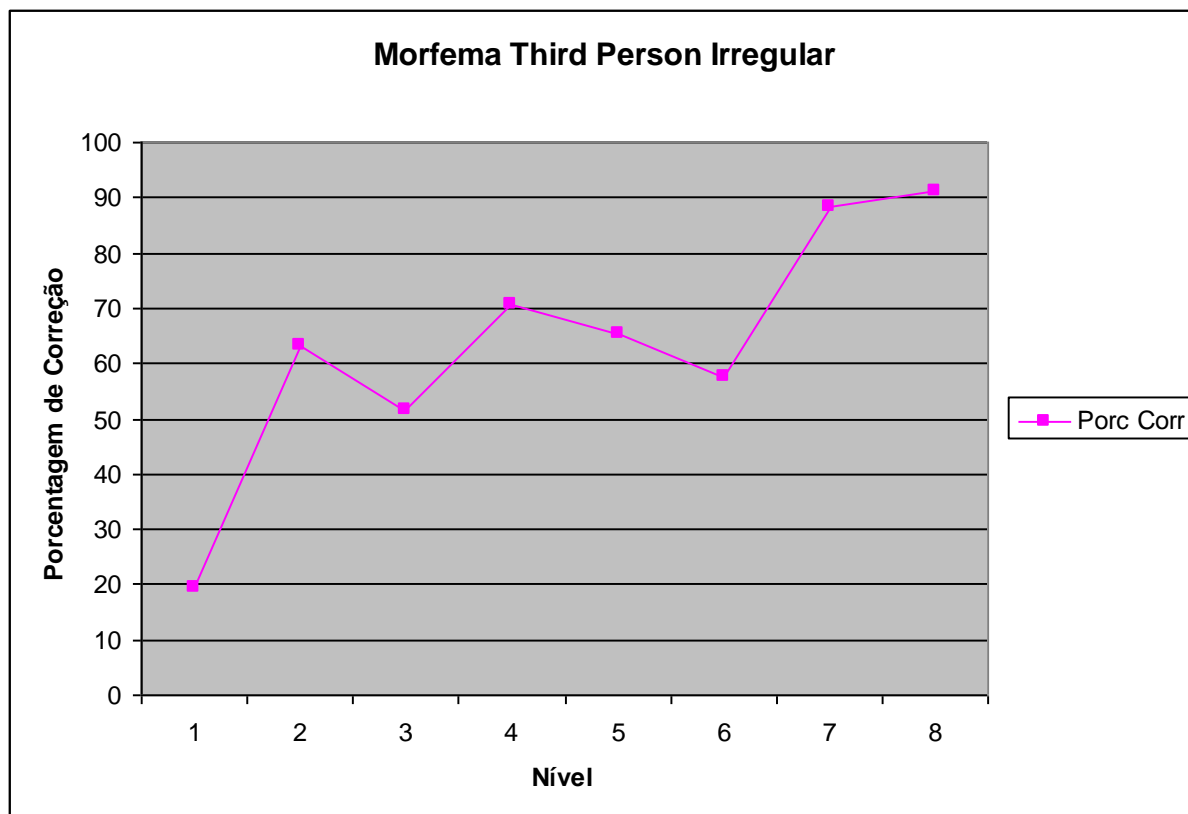


Figura 10: Morfema *Third Person Irregular*

E o morfema *Third Person Irregular* apresenta índices de correção muito baixos no nível 1: 19,3%, que salta para 63,3% já no nível 2. Após oscilações entre os níveis 3 a 7 com média de 60%, ele atinge 88,2% no nível 7 e 91,3% no nível 8.

Tabela 30: Morfema *Uncontracted Auxiliary*

SOC (<i>Supp Oblig Context</i>)					
NÍVEL	<i>Supplied (s)</i>	<i>Absent (a)</i>	<i>Overuse (ou)</i>	Tota Ocorrências	Porcentagem de correção
1	6	2	0	8	75%
2	23	3	0	26	88.4 %
3	19	2	0	21	90.4 %
4	29	2	0	31	93.5 %
5	30	1	0	31	96.7 %
6	35	1	0	36	97.2 %
7	49	2	0	51	96%
8	11	0	0	11	100%

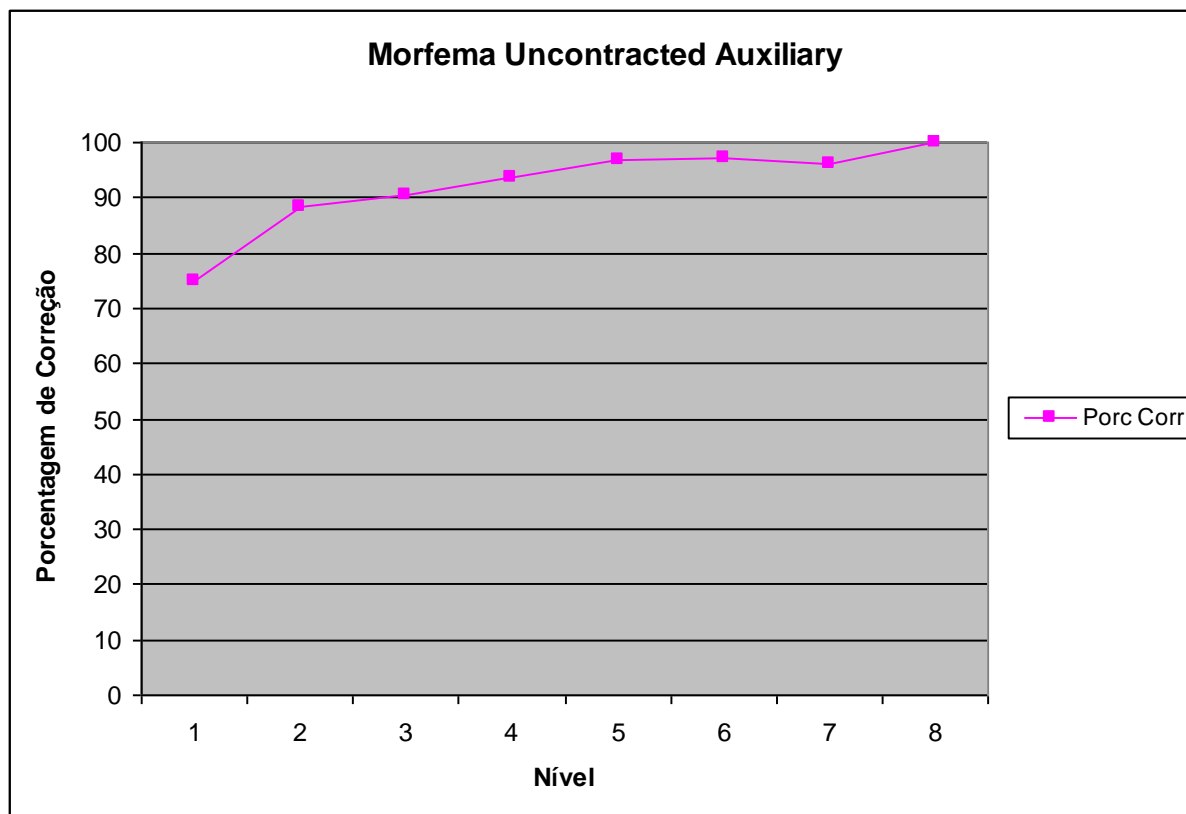


Figura 11: Morfema *Uncontracted Auxiliary*

O morfema *Uncontracted Auxiliary* inicia com um índice inicial razoavelmente alto de correção: 75% e seu padrão evolutivo se mantém ascendente até o nível 8, quando atinge 100% de correção.

Tabela 31: Morfema *Contracted Auxiliary*

SOC (Supp Oblig Context)					
NÍVEL	<i>Supplied (s)</i>	<i>Absent (a)</i>	<i>Overuse (ou)</i>	Total Ocorrências	Porcentagem de correção
1	3	2	0	5	60 %
2	14	2	0	16	87.5 %
3	20	1	0	21	95.2 %
4	23	3	0	26	88.4 %
5	29	1	0	30	96.6 %
6	21	0	0	21	100 %
7	60	1	0	61	98.3 %
8	18	0	0	18	100 %

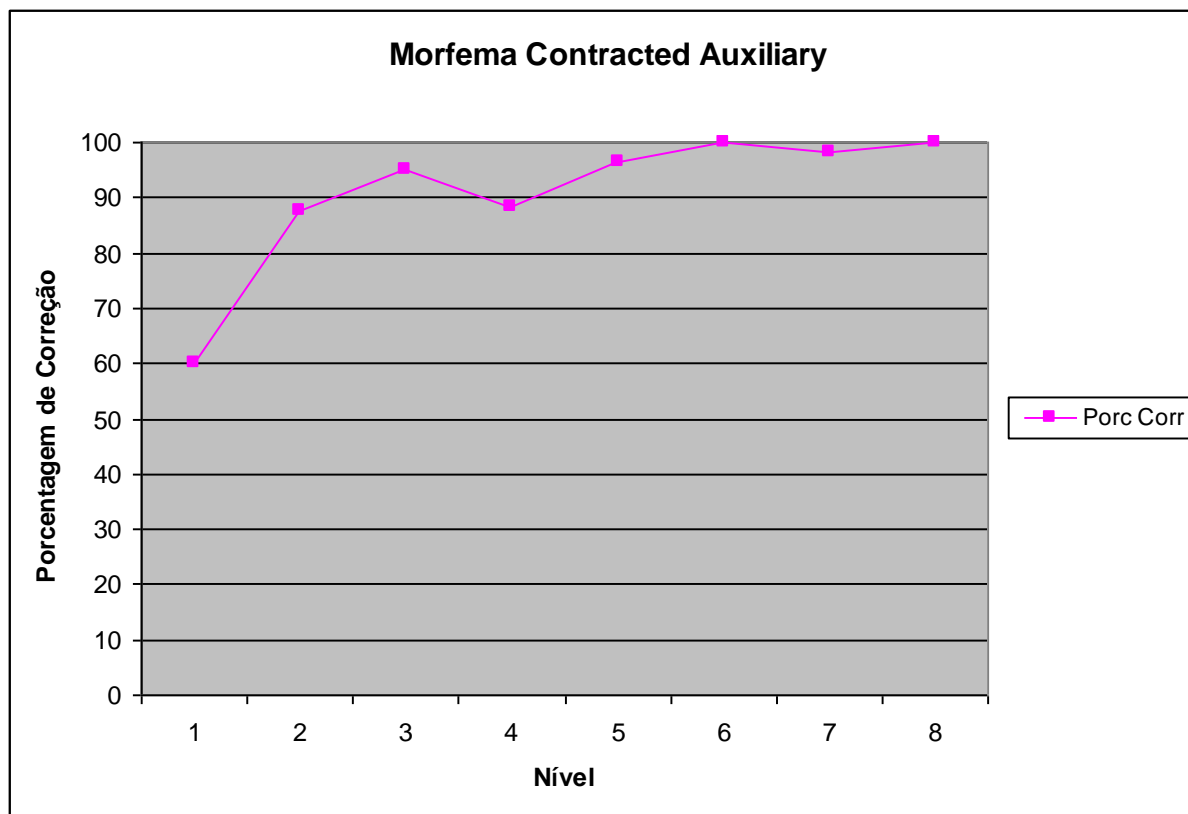
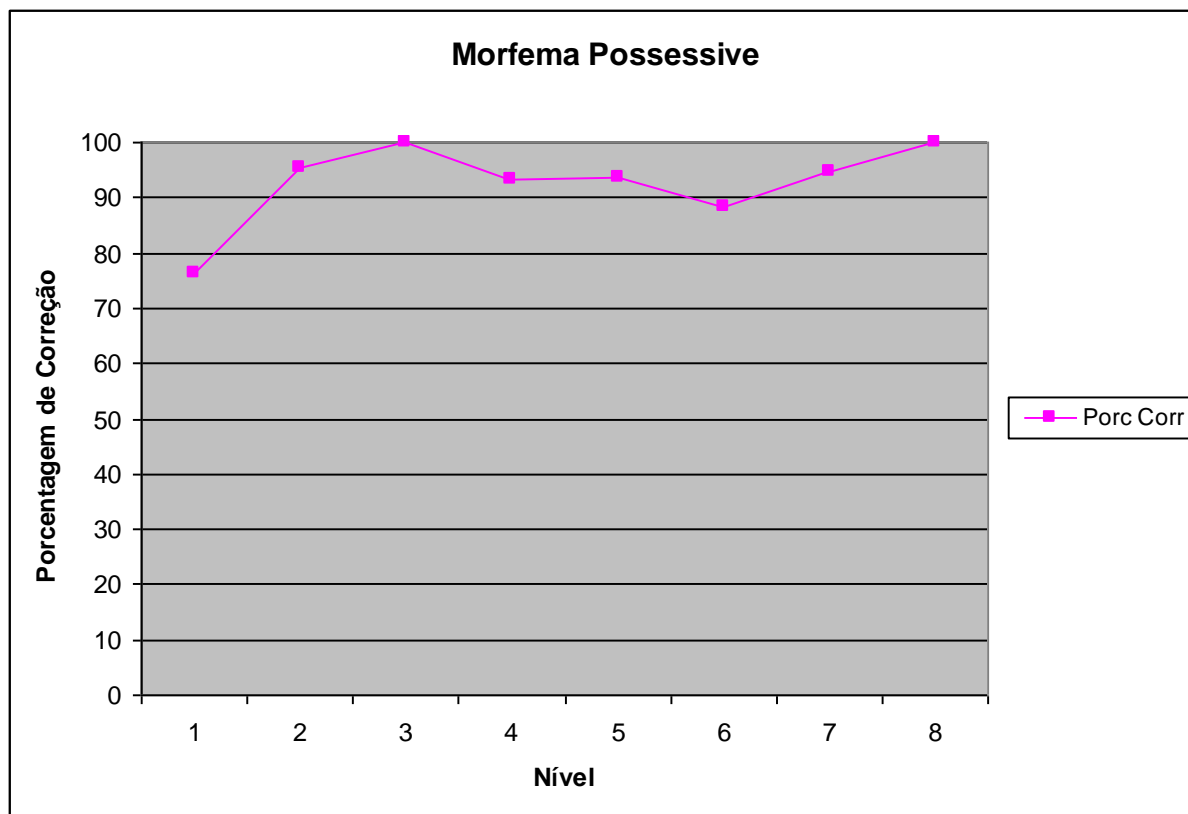


Figura 12: Morfema *Contracted Auxiliary*

Este morfema apresenta 60% de correção no nível 1, índice que sobe para 87.5% já no nível 2 e 95.2% no nível 3. Após oscilações na casa dos 90%, o índice de correção atinge os 100% no nível 8.

Tabela 32: Morfema *Possessive*

SOC (Supp Oblig Context)					
NÍVEL	<i>Supplied (s)</i>	<i>Absent (a)</i>	<i>Overuse (ou)</i>	Total Ocorrências	Porcentagem de correção
1	13	4	0	17	76.4%
2	21	1	0	22	95.4%
3	16	0	0	16	100%
4	23	3	0	26	88.4 %
5	15	1	0	16	93.7%
6	15	2	0	17	88.2%
7	18	1	0	19	94.7%
8	12	0	0	12	100%

Figura 13: Morfema *Possessive*

Os dados com relação a este morfema também são bastante interessantes: no nível 1, o índice de correção é de 76.4%, subindo para 95.4% no nível 2 e 100% no nível 3. A partir do nível 4 até o nível 7 observamos um declínio, onde os índices se mantêm na casa dos 90%. Mas o uso correto é de 100% no nível 8.

Tabela 33: Morfema *Prepositions in/on*

SOC (Supp Oblig Context)					
NÍVEL	<i>Supplied (s)</i>	<i>Absent (a)</i>	<i>Overuse (ou)</i>	Total Ocorrências	Porcentagem de correção
1	220	4	0	224	98.2 %
2	296	5	0	301	98.3 %
3	319	1	0	320	99.6 %
4	239	1	0	240	99.5 %
5	235	1	0	236	99.5 %
6	192	1	0	193	99.4 %
7	374	1	0	375	99.7 %
8	175	2	0	177	98.8 %

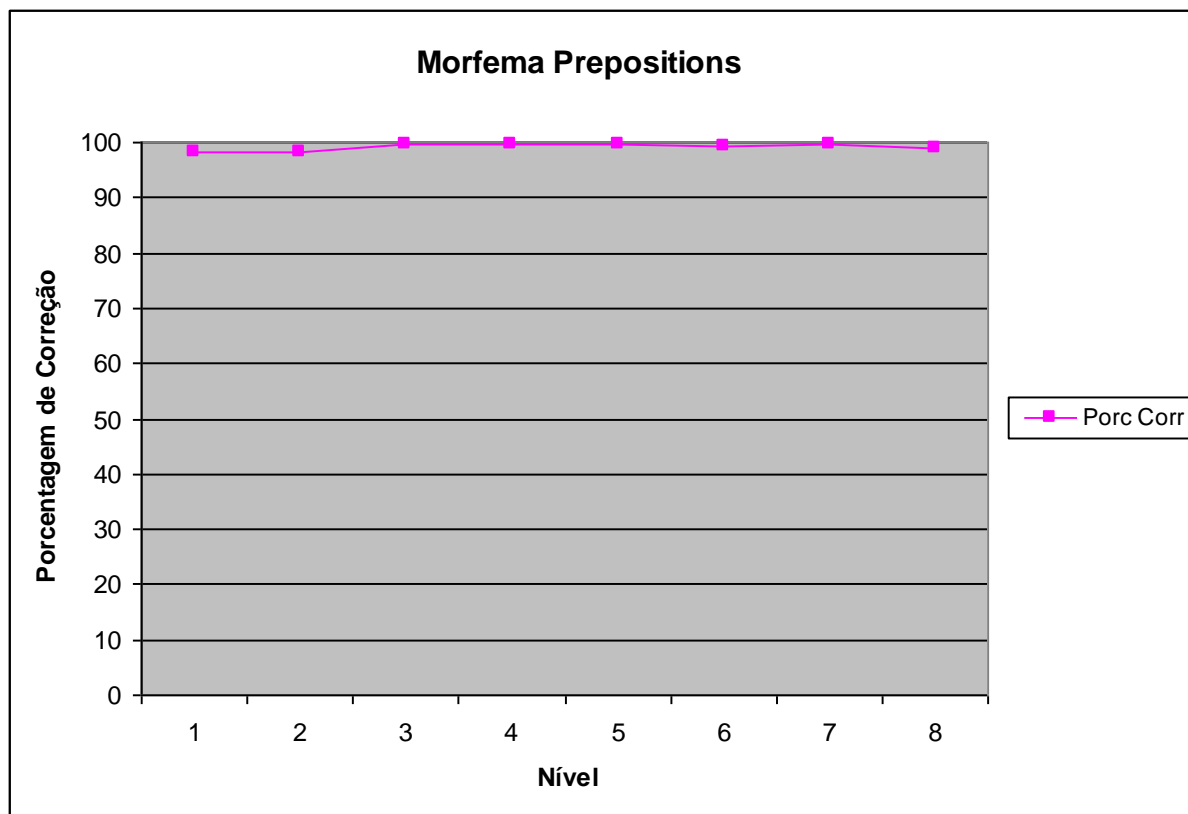


Figura 14: Morfema *Prepositions*

Com relação às preposições, podemos observar dados que apontam para um alto índice de correção por parte dos nossos sujeitos, partindo de 98.2% no nível 1 e seguindo em ascensão até o nível 8, onde atinge 98.8%.

Assim, a partir destas tabelas e figuras, podemos ter um panorama geral sobre o padrão evolutivo dos aprendizes brasileiros de inglês como L2 com relação à aquisição de morfemas. Vejamos a seguir a tabela com a ordem de aquisição evidenciada por estes sujeitos:⁸⁹

⁸⁹ Lembramos que estamos considerando o aparecimento do índice de correção de 90% por uma vez, conforme já mencionamos anteriormente, como ponto de corte para considerarmos o morfema adquirido.

Tabela 34: Ordem de aquisição de morfemas a partir dos níveis de aprendizagem investigados no BELC

NIVEL	MORFEMA
1	<i>Unc Copula</i> <i>Plural</i> <i>Preposition</i> <i>Contract Copula</i>
2	<i>Article</i> <i>Possessive</i>
3	<i>Past Irregular</i> <i>Present Progressive (-ING)</i> <i>Contract Aux</i> <i>Uncontract Aux</i>
4	
5	
6	<i>Past Regular</i>
7	<i>Third Person Regular</i>
8	<i>Third Person Irregular</i>

A ordem evidenciada pelos aprendizes de inglês brasileiros é bastante interessante. Nela, podemos observar:

(i) No nível 1, consideramos adquiridos os morfemas de *Copula* (tanto *contracted* quanto *uncontracted*), quanto o morfema de *Plural* e os morfemas preposicionais *in* e *on*.

(ii) No nível 2, se apresentam os morfemas *Article* e *Possessive*. O aparecimento desse último num estágio elementar é semelhante à posição observada na ordem de Brown (1973), mas é uma diferença fundamental com relação à sequência padrão proposta por Krashen (1977), onde ele aparece em último, conforme veremos em seguida na análise dos dados.

(iii) No nível 3, surgem os morfemas *Past Irregular*, *Present Progressive* e *Auxiliary (Contracted & Uncontracted)*. O morfema *Present Progressive* num terceiro estágio evolutivo é uma diferença com relação à ordem de Krashen (1977) a ser considerada em nossa análise. Note que, nas sequências de Krashen e de Brown (1973), este morfema aparece em primeira posição.

(iv) O morfema *Past Irregular* é observado no terceiro nível, ao passo que o *Past Regular* no sexto. Como veremos na análise dos dados, isso também pode ser observado em outras pesquisas em L2.

(v) Por fim, os morfemas de *Third Person*, tanto regular quanto irregular, apresentam-se nos níveis 7 e 8 respectivamente, isto é, no último estágio evolutivo, como também apontam os estudos revisados sobre o assunto, conforme veremos em nossa análise.

Podemos, ainda, organizar estas informações numa tabela como a sequência proposta por Krashen (1977),⁹⁰ onde a sequência de aquisição é mostrada em blocos de morfemas considerados adquiridos num mesmo estágio.⁹¹

Tabela 35: Sequência na aquisição de morfemas em blocos

<i>COPULA (Contracted & Uncontracted)</i>
<i>PLURAL</i>
<i>PREPOSITION</i>
<i>ARTICLE</i>
<i>POSSESSIVE</i>
<i>PRESENT PROGRESSIVE (-ING)</i>
<i>AUXILIARY (Contracted & Uncontracted)</i>
<i>PAST IRREGULAR</i>
<i>PAST REGULAR</i>
<i>THIRD PERSON REGULAR</i>
<i>THIRD PERSON IRREGULAR</i>

Note que, nesta tabela, condensamos os morfemas de *Contracted Copula* e *Uncontracted Copula* só como *Copula*, pois ambos podem ser considerados adquiridos no nível 1, assim como os morfemas *Contracted Auxiliary* e *Uncontracted Auxiliary* como *Auxiliary*, que se manifestam juntamente no nível 3, como podemos ver na Tabela 34.

Na seção seguinte, propomos uma análise destes resultados, principalmente à luz das teorias propostas para explicar uma possível sequência natural de aquisição de morfemas propostas por Zobl e Liceras (1994) e Goldschneider & DeKeyser (2005), conforme vimos nos capítulos 2 e 3.

5.2 ANÁLISE

A partir da apresentação dos resultados no capítulo anterior, podemos proceder à análise e discussão dos dados. Utilizando os dados das tabelas e figuras apresentadas, confrontaremos, primeiramente, a ordem de morfemas obtida a partir do corpus BELC. Poderemos, então, identificar semelhanças e diferenças ao compararmos com outras ordens de aquisição, em especial com as ordens obtidas

⁹⁰ Cf. Tabela 3, capítulo 2.

⁹¹ Cf. já colocamos a partir de Zobl e Liceras (1994), seção 3.2.2.

por Brown (1973) e Krashen (1977), pesquisas de referência para este estudo. Num segundo momento, tentaremos oferecer explicações possíveis a tais semelhanças/diferenças, principalmente à luz dos trabalhos de Goldschneider & DeKeyser (2005) e Zobl & Licerias (1994). Como já vimos especificamente no capítulo 3, Goldschneider & DeKeyser (2005) consideram múltiplos fatores para explicar o fenômeno, ao passo que Zobl & Licerias (1994) explicam o fenômeno a partir de categorias sintáticas semelhantes.

5.2.1 Comparação dos Resultados

A tabela apresentada a seguir nos mostra os resultados dos estudos de Brown (1973) e Krashen (1977) em paralelo com os resultados obtidos através do BELC.

Tabela 36: Resultados de Brown (1973), Krashen (1977) e BELC

BROWN (1973)	KRASHEN (1977)	BELC
<i>V- ing</i>	<i>ING</i>	<i>COP</i>
<i>PLU – S</i>	<i>PLURAL</i>	<i>PLURAL</i>
<i>Irreg Past</i>	<i>COP</i>	<i>PREP</i>
<i>Poss – S</i>	<i>AUXILIARY</i>	<i>ARTICLE</i>
<i>Unc COP BE</i>	<i>ARTICLE</i>	<i>POSSESSIVE</i>
<i>ART</i>	<i>IRREG PAST</i>	<i>PRES PROG – ING</i> <i>AUX</i> <i>IRREG PAST</i>
<i>REG Past</i>		<i>REG PAST</i>
<i>3^o P – S</i>	<i>REG PAST</i> <i>III SING</i>	<i>THIRD P S REGULAR</i>
<i>Uncont AUX</i>	<i>POSSESSIVE</i>	<i>THIRD P S IRREGULAR</i>
<i>Contracted Copula</i>		

Ao observarmos estes dados de maneira genérica, podemos notar muitas semelhanças e também algumas diferenças. Assim, a fim de podermos melhor estabelecer este paralelo, é necessário, primeiramente, organizarmos os morfemas que estamos comparando.

No estudo de Brown, são considerados 14 morfemas (cf. Tabela 1, capítulo 2) e a partir deles, estabelecemos os morfemas a serem investigados no BELC. No entanto, o estudo de De Villiers & De Villiers (1973) (cf. Tabela 2, capítulo 2) considera somente dez (10) morfemas e a partir disso, o escopo dos morfemas

analisados foi automaticamente reduzido a nove (9). Assim, os diferentes estudos sobre aquisição de morfemas foram variando nas estruturas consideradas, como podemos ver na Tabela 3.

Krashen (1977), por sua vez, considera nove morfemas a partir de Brown (1973), apesar de ser mais genérico com relação a dois deles: *Copula* e *Auxiliary*. Brown considera “*Uncontracted Copula Be*” e “*Uncontracted Auxiliary*”, ao passo que Krashen é mais genérico, utilizando “*Copula*” e “*Auxiliary*”.

Já o BELC etiquetado para esta pesquisa apresenta dados de quatorze (14) morfemas, tendo sido primeiramente reduzido a treze (13) ao considerarmos juntamente as preposições *in* e *on*, depois reduzido a onze (11) ao condensarmos *Contracted* e *Uncontracted Copula* e *Contracted* e *Uncontracted Auxiliary*. Para os fins desta tese, propomos a redução da análise do BELC a nove morfemas, (i) retirando o morfema Preposição⁹² da análise comparativa, uma vez que ele não é considerado nos dados que selecionamos acima de Brown (1973) e Krashen (1977) e (ii) condensando os morfemas *Third Person Singular Regular* e *Third Person Singular Irregular*, igualmente aos estudos de referência.⁹³

Ainda, a fim de podermos estabelecer uma análise de maneira mais harmoniosa, faremos algumas modificações nas referências aos morfemas nas tabelas que seguem: utilizaremos os mesmos termos de Krashen (1977) com relação aos nomes dos morfemas nos outros dois estudos (o de Brown e o do BELC). Assim, apresentamos abaixo uma tabela padronizada para podermos prosseguir a nossa análise:

⁹² No BELC, o morfema é adquirido já no primeiro nível, igualmente aos estudos de Brown, onde ele aparece na segunda posição, o que pode ser considerado um estágio inicial.

⁹³ Ratifico a escolha de ter feito a etiquetagem do BELC a partir dos treze morfemas originais de Brown para não limitar o escopo de pesquisa nesta tese. Ou seja, para o estudo aqui proposto, é recomendável que mantenhamos os mesmos morfemas para podermos estabelecer considerações mais bem fundamentadas. No entanto, os dados utilizados em pesquisas sobre os estudos da ordem natural de morfemas são, muitas vezes, oriundos de análises estatísticas a partir da junção de resultados de outras pesquisas e, nem sempre, há informação sobre todos os morfemas. A fim de evitar esse problema em possíveis análises futuras, preferimos levantar dados sobre os treze morfemas.

Tabela 37: Resultados padronizados de Brown (1973), Krashen (1977) e BELC

BROWN (1973)	KRASHEN (1977)	BELC
<i>ING</i>	<i>ING</i>	<i>COP</i>
<i>PLURAL</i>	<i>PLURAL</i>	<i>PLURAL</i>
<i>IRREG PAST</i>	<i>COP</i>	
<i>POSSESSIVE</i>	<i>AUXILIARY</i>	<i>ARTICLE</i>
<i>COP</i>	<i>ARTICLE</i>	<i>POSSESSIVE</i>
<i>ARTICLE</i>	<i>IRREG PAST</i>	<i>ING</i>
<i>REG PAST</i>	<i>REG PAST</i>	<i>AUXILIARY</i>
<i>III SING</i>	<i>III SING</i>	<i>IRREG PAST</i>
<i>AUXILIARY</i>	<i>POSSESSIVE</i>	<i>REG PAST</i>
		<i>III SING</i>

Isto posto, vejamos, a seguir, uma breve descrição sobre semelhanças e diferenças encontradas nestes resultados a partir de cada morfema, partindo da sequência de Krashen (1977):

1.O morfema –*ING*:

Na pesquisa de Brown, este morfema ocupa o primeiro lugar. A pesquisa de Krashen nos mostra que ele é adquirido no primeiro estágio de aquisição e em nossa pesquisa, os dados do BELC mostram que ele é adquirido no terceiro estágio de aquisição.

2.O morfema *PLURAL*:

Este morfema é adquirido em segundo lugar em Brown e, no primeiro estágio em Krashen e no BELC.

3. O morfema *COPULA*:

Os dados nos mostram que o morfema *Copula* é adquirido no primeiro estágio em Krashen e no BELC e em quinto lugar (na modalidade *Uncontracted*) em Brown.

4.O morfema *AUXILIARY*:

O morfema *Auxiliary* é adquirido no segundo estágio em Krashen, no terceiro estágio no BELC e, em nono lugar, em Brown.

5.O morfema *ARTICLE*:

Em Krashen, este morfema aparece no segundo estágio, bem como no BELC. Em Brown, ele ocupa a sexta posição na ordem de aquisição.

6.O morfema *IRREGULAR PAST*:

Este morfema aparece no terceiro estágio em Krashen e no BELC e na terceira posição do ranking de Brown.

7.O morfema *REGULAR PAST*:

Krashen nos mostra a aquisição deste morfema no quarto estágio evolutivo, igualmente ao BELC, e poderíamos dizer semelhantemente a Brown, que o coloca em sétimo na ordem.

8.O morfema *III SINGULAR*:

O morfema de *Third Person Singular* aparece no quarto estágio em Krashen e no BELC e na oitava posição em Brown.

9.O morfema *POSSESSIVE*:

Este morfema surge no quarto estágio de aquisição em Krashen e no segundo no BELC. Em Brown, ele surge em quarto lugar, numa posição semelhante ao BELC.

Temos, assim, um panorama geral sobre semelhanças e diferenças na aquisição de morfemas a partir dos estudos que estamos investigando.

Em síntese, poderíamos dizer que nossos resultados confirmam, ainda que parcialmente, uma ordem semelhante a dos estudos de referência. As semelhanças ocorrem com os morfemas *Plural* (no primeiro estágio em todos), *Article* (na posição intermediária), *Reg Past* e *III SING* nas posições mais baixas da sequência.

Entretanto, listamos duas diferenças significativas:

(i) O morfema *Possessive*: no BELC ele é observado no terceiro estágio de aquisição, semelhantemente a Brown (1973) – L1 – e não a Krashen (1977) – L2.

(ii) O morfema *-ING* que, em Brown e Krashen, ocupa a primeira posição e, no BELC, a terceira. Este é um problema crucial a ser observado em nossos resultados, pois, por analogia, ele não deveria ser esperado nesta posição.

A seguir, propomos a análise destas semelhanças e diferenças a partir dos dois referenciais teóricos que apontamos no capítulo 3, que são Zobl & Liceras (1994) e Goldschneider & DeKeyser (2005).

5.2.2 A Perspectiva de Zobl & Licerias (1994)

No capítulo 3, apresentamos as principais considerações que estes pesquisadores fazem sobre os estudos dos morfemas, sugerindo uma explicação unificada para a possível sequência observada na aquisição de morfemas nos estudos realizados nos anos de 1970 e de 1980 a partir do advento da teoria linguística, mais especificamente dos estudos sobre categorias funcionais.

5.2.2.1 Breve retomada da perspectiva teórica

Em linhas gerais, os autores tentam provar que uma possível ordem natural não deve ser coincidência e que as diferenças observadas entre os estudos em L1 e L2 podem ser reduzidas a contrastes simples.

Assim sendo, a teoria das categorias funcionais é o principal determinante da hierarquia de morfemas observada. Esta teoria pressupõe alguns agrupamentos naturais de morfemas em termos de:

(i) Conjunto de membros de uma categoria: DP (*Determiners*, como *Article* e *Possessive*) e IP (*Reg Past*, *III Sing*)

(ii) Movimento de núcleo: que pode ser lexical (como o movimento de *Cop Be* e *Aux Be*) ou afixional (como em *III Sing*, *Reg Past*)

(iii) Distinção entre morfemas livres e presos. Morfemas livres: categorias como modais, auxiliares, determinantes. Morfemas presos, que são afixos nominais e verbais, tais como *Possessive* e *III Sing*, respectivamente.

(iv) Categoria lexical: salientamos que os autores não mencionam explicitamente esta categoria. No entanto, só conseguimos compreender bem a explicação oferecida por eles se assumirmos que eles devem pressupor que *-ING* é uma categoria lexical. Assim, com o acréscimo do morfema *-ING*, existe a possibilidade de redundância da informação, pois a forma de expressão do *Present Progressive* é uma forma composta (*Be + verb + ING*), como em *I am leaving*. Ao

contrário, formas como *I will leave* (modal Will) ou *He leaves* (morfema III Sing) não têm a possibilidade de redundância da informação temporal, pois são simples.⁹⁴

Ainda, a fim de acompanharmos melhor a contribuição teórica oferecida por Zobl & Liceras (1994), apresentamos a seguir uma tabela própria que criamos para um melhor entendimento da classificação dos morfemas em termos de definições dentro da teoria das categorias funcionais apresentadas pelos autores tal como eles parecem pressupor esta classificação.

Tabela 38: Classificação dos morfemas analisados por Zobl & Liceras (1994) a partir do entendimento dos autores

MORFEMA	CATEGORIA MÓRFICA			CATEGORIA SINTÁTICA	
	AFIXO (preso)	PALAVRA (livre)	FORMA IRREGULAR	FLEXÃO IP	FLEXÃO DP
-ING	X			X	
PLURAL	X				X
COPULA		X		X	
AUXILIAR		X		X	
ARTICLE		X			X
IRREG PAST			X	X	
REG PAST	X			X	
III SING	X		X ⁹⁵	X	
POSSESSIVE	X				X

Os autores afirmam que, em L1, a ordem depende da categoria e não da diferença entre morfema livre e preso, com os expoentes de DP precedendo expoentes de IP. Isto é, os morfemas em L1 parecem emergir numa ordem de categoria específica, apresentando primeiro as categorias nominais e depois, as categorias verbais. Isso é exemplificado com a manifestação de *Article* (enquanto *Determiner*), juntamente com *Poss's* (enquanto DP) separados de *Auxiliary*, que, por sua vez, se manifesta em último. Da mesma forma, os afixos *III Sing* e *Reg Past* estão agrupados no fim da hierarquia, juntamente a *Auxiliary* (por sua vez, Infl), exibindo o mesmo contraste entre categorias específicas e categorias variadas que os morfemas livres.

Já em L2, segundo Krashen (1977) e os outros estudos apresentados por Zobl & Liceras, a presença de *Copula* (IP), *Auxiliary* (IP) e *Determiner* (*article*, DP)

⁹⁴ Uma das críticas tecidas a esse estudo é feita por Goldschneider & Dekeyser (2005) que dizem, entre outras coisas, não entenderem porque -ING é considerado afixo lexical e oposto a afixo sintático (p.66).

⁹⁵ Se considerarmos também a forma irregular *has*.

nas primeiras posições da sequência de aquisição⁹⁶ evidenciam que a ordem é de categorias variadas, sugerindo que os aprendizes já dispõem de categorias funcionais desde os primeiros estágios de aquisição. Assim, a ordem parece depender da diferença entre morfemas livres e presos, e do contraste entre o expoente afixional preso *versus* livre, ambos se manifestando em extremos da hierarquia, com os expoentes livres no topo e os presos no fim. Este contraste entre morfemas livres e presos justapõe-se com o contraste entre movimento de núcleo lexical (*Auxiliary* e *Copula* sobem para a posição de Infl) e movimento de afixo. O movimento, em particular o movimento de afixo, pode assim representar uma generalização significativa no processo de formação gramatical de aprendizes de L2. A disponibilidade de projeções funcionais desde cedo manifesta-se através de seus expoentes lexicais como *Determiner (article)*, *Auxiliary* e *Copula*, parece adiar a aquisição de afixos.

No entanto, nem todos os afixos são adquiridos tardiamente em L2. Podemos ver *-ING* e *Plural*, afixos lexicais, nas primeiras posições. Para Zobl & Licerias, as diferenças evolutivas entre os afixos lexicais e sintáticos sustentam que a teoria das categorias funcionais, e não propriedades de saliência, são o principal determinante da hierarquia.

Segundo Zobl & Licerias (1994), como *III Sing* e *Reg Past* estão agrupados no fim da hierarquia em L1 e em L2, isso sugere que, em ambos os processos, a descida do afixo pode ser responsável pela aquisição tardia destas formas. Entretanto, os autores notam que *Aux BE* está em posição baixa, igualmente, em L1. Mas *Aux BE* é um elemento que sofre movimento para INFL, e não descida do afixo. Então, para L1, a hipótese de que o fator relevante seja descida do afixo não parece satisfatória; o que parece importar, assim, é a aquisição do sistema de INFL como um todo.

Por outro lado, Zobl & Licerias observam que *Poss-'s* é adquirido simultaneamente ao Determinante (*article*) em L1. Por isso, eles assumem que *Poss 's* envolveria descida de afixo, e isso não aconteceria com *article*, a outra instância de *determiner*. Assim, Zobl & Licerias (1994) concluem que uma explicação baseada

⁹⁶ É importante salientarmos que, por tomarem a ordem dos morfemas em blocos a partir da emergência simultânea de alguns deles no processo de aquisição, podemos notar que os autores consideram, em sua análise, os primeiros dois blocos como estágios iniciais e os dois últimos como estágios mais tardios ou finais.

na descida de afixo não seria adequada para L1. Trata-se, de novo, da aquisição do sistema de uma categoria funcional como um todo – neste caso, do sistema D.

Antes de aplicarmos estas explicações aos resultados evidenciados pelo nosso corpus, salientamos, como uma observação crítica ao estudo de Zobl & Liceras (1994), que suas asserções acerca da descida de afixo não são consensuais na literatura. Por exemplo, na análise difundida por Abney (1986), *Poss 's* é um Det que aparece em sua posição de base, e não sofre qualquer descida de afixo. Em linhas gerais, podemos dizer que o artigo parece carente em argumentação com relação a esse ponto. E essa observação faz bastante diferença na nossa análise, pois estabelece que, ao nosso ver, Zobl & Liceras (1994) não dariam conta de seus próprios resultados com relação ao morfema *Possessive*. Ou seja, para explicar o porquê de este morfema estar nas últimas posições (cf. Krashen, 1977), os autores assumem, na sua análise, que ele sofreria descida de afixo, ao passo que, como vimos, isso não é consensual. Se ele sofre descido de afixo, pode ser entendido e explicado conforme as suposições de Zobl & Liceras (1994). Se não sofre tal movimento, não. Veremos, mais adiante, as consequências disso para os nossos resultados.

5.2.2.2 Aplicação aos resultados do BELC

Como podemos ver, as explicações oferecidas por Zobl & Liceras (1994) para a ordem dos morfemas vista nos estudos selecionados por eles são interessantes no sentido de tentarem explicar, através de um fator de natureza única – sintática – a sequência de morfemas evidenciada por aprendizes de L1 e L2. Porém, elas podem ser aplicadas apenas parcialmente à ordem que podemos observar no BELC.

De um modo geral, podemos ver que estas explicações podem servir para interpretar algumas semelhanças observadas entre as pesquisas de referência e o nosso corpus: o BELC mostra resultados que parecem evidenciar a presença de categorias cruzadas no processo de aquisição de morfemas em inglês como L2, com categorias DP nos estágios iniciais, evidenciadas por Plural, Artigo e Possessivo e também categoria IP no topo, ilustrada através do morfema *Copula*.

Com relação a expoentes afixionais, podemos dizer que algumas categorias que não dependem de afixos se manifestam na posição mais alta da sequência, como *Copula* e *Article*, e as mais dependentes de afixo, nas posições mais baixas da hierarquia apresentada pelo BELC, como *III Sing*, *Reg Past*.⁹⁷ No entanto, os morfemas *-ING* e *Possessive* apresentam posições diferentes do que em Krashen (1977), o que não seria esperado se assumirmos essa perspectiva teórica.

Em nosso corpus, *Copula* e *Plural* aparecem no primeiro bloco da sequência, assim como em Krashen (1977). *Copula* é uma instância de morfema livre e também de IP. *Plural*, nesta concepção, é um exemplo de afixo lexical, que, apesar de ser um morfema preso, então poderia ser esperado no estágio inicial.

No segundo estágio evolutivo, observamos *Article* e *Possessive*, ambos exemplos de DP, morfemas livre e preso, respectivamente. Aqui temos um dado muito interessante evidenciado pelos nossos aprendizes: se observarmos na Tabela 37, veremos que a posição do morfema *Possessive* no BELC é semelhante à da ordem de Brown (1973) e não à de Krashen (1977). Como acabamos de ver, Zobl e Liceras (1994) nos dizem que, em L1, *Possessive* e *Article* estão juntos devido à emergência da categoria DP e em L2, separados porque *Possessive* é uma instância de DP preso, e, a partir da teoria das categorias funcionais, este morfema deveria ser esperado no fim do processo, juntamente com *III Sing* e *Reg Past*. Assim, esta perspectiva teórica tal como é entendida por Zobl & Liceras também não explica o resultado obtido no BELC com relação ao *Possessive*. Todavia, como vimos na seção anterior, se considerarmos que *Possessive* não sofre descida de afixo – de acordo com a análise tradicional (cf. Abney, 1986) e ao contrário de Zobl & Liceras (1994) - poderíamos atribuir a isso a posição deste morfema no BELC. Isto é, *Possessive* surgiria no segundo estágio de aquisição, juntamente com *article*, o que assemelha nossos resultados aos de Brown (1973). No entanto, salientamos que esta não é a análise feita por Zobl & Liceras (1994), logo, tal análise não esclarece a ordem de *Possessive* junto a determinantes no BELC.

Entretanto, algumas pesquisas em L2 mostraram resultados mais semelhantes aos nossos. Wagner (2007, cf. capítulo 2, seção 2.2.2) observou o morfema *Possessive* na primeira posição, mas desconsiderou os resultados devido a possíveis problemas metodológicos. No caso da nossa pesquisa, não poderíamos

⁹⁷ Como já falamos na seção anterior, Zobl & Liceras consideram *-ING* e *Plural* formas lexicais de afixo, contrapondo a afixos sintáticos, como *Reg Past* e *III Sing*.

simplesmente desconsiderar os resultados obtidos por problemas metodológicos. Primeiramente, porque o método escolhido para a coleta é baseado em dados de fala espontânea. Segundo, porque, mesmo se tivesse havido algum problema com relação a algum erro de etiquetagem ou computação da forma de preenchimento no contexto obrigatório, os índices de correção não seriam tão altos e tão estáveis desde o primeiro nível (cf. Tabela 32, Figura 13).

Izumi & Isahara (2004) também nos mostram o morfema *Possessive* em primeiro lugar e atribuem tal posição a aspectos de transferência do japonês, que, segundo os autores, apresenta estruturas semelhantes (cf. capítulo 2, seção 2.2.2). Como sabemos, o português brasileiro não apresenta uma estrutura semelhante ao inglês, pois expressa posse através da preposição “de” ou das contrações “do/da/dos/das”.

Por isso, teríamos que pensar em outros fatores para explicarmos a ocorrência do morfema *Possessive* no terceiro estágio. Um dado interessante que se apresenta é a semelhança na posição deste morfema no BELC e em Brown (1973): nestes dois corpora, *Possessive* se apresenta com *Article*, dois expoentes de DP. Essa informação é bastante relevante se pensarmos numa analogia entre as duas amostras: ambas são correspondentes, pois há um agrupamento uniforme de sujeitos em L2 e L1, o que é drasticamente diferente de Krashen (1977) e outros estudos (ANDERSEN, 1978; GOLDSCHNEIDER & DEKEYSER, 2005), que desenvolveram uma análise a partir de dados de outras pesquisas, aplicando métodos estatísticos e excluindo morfemas ausentes em todos os estudos, o que faz com que os resultados fiquem um pouco enfraquecidos se comparados a dados de pesquisas com coletas orientadas. Portanto, a análise de Zobl & Liceras não esclarece a ordem de *Possessive* junto a determinantes no BELC.

Prosseguindo nossa análise, os resultados do BELC nos mostram os morfemas *ING*, *Auxiliary* e *Irregular Past* no terceiro estágio evolutivo.

Particularmente com relação ao morfema *-ING*, temos um problema com os dados do BELC, devido ao contraste com as pesquisas de referência deste estudo (BROWN (1973) e KRASHEN (1977)). Ambas apontam este morfema no primeiro estágio evolutivo, assim como Wagner (2007) e Izumi e Isahara (2004) o colocam em segunda posição. Como já colocamos, se assumirmos a análise de Zobl & Liceras (1994), esperaríamos este morfema nas primeiras posições, o que não acontece no BELC. Assim, esta concepção teórica também não oferece explicações

adequadas aos nossos resultados com relação ao morfema *-ING*. Por isso, é necessário pensarmos em outros fatores aos quais poderíamos atribuir a posição do morfema *-ING* na sequência de aquisição do nosso corpus.

Semelhantemente aos nossos resultados, apontamos Andersen (1978), que nos mostra *-ING* em terceira posição juntamente com *Auxiliary*, assim como o BELC. O foco do estudo de Andersen é metodológico, pois o autor propõe uma escala implicacional para analisar a possibilidade de uma ordem de morfemas em inglês como L2 a partir de diferentes L1s (cf. capítulo 2, seção 2.2.2). Andersen nos diz que o fato de *-ING* aparecer junto com *Auxiliary* parece lógico, pois ambos são os dois componentes do *Present Progressive* e constituem pares de morfemas logicamente relacionados. O autor reforça a sua idéia a partir de Rosansky (1976), que caracteriza como *functional acquisition* casos como este, onde os morfemas que servem a mesma função são adquiridos juntos ou próximos.⁹⁸

Na análise desta tese, poderíamos igualmente atribuir a manifestação concomitante de *Auxiliary* e *-ING* ao fato de que ambos os morfemas constituem um conjunto para expressar corretamente o *Present Progressive*. Nesta perspectiva, podemos pensar no papel do fator instrução. Embora estejamos utilizando os termos aquisição e aprendizagem indistintamente, em consonância com a literatura revisada nesta tese, sabemos que eles são processos de naturezas distintas. O fato de que os nossos sujeitos são originalmente aprendizes de inglês faz com que pensemos no contexto de sala de aula, onde há uma instrução explícita e insistente para que eles assimilem a estrutura completa de *Present Progressive/Present Continuous*. Este é um fator que deve ser considerado como altamente interferente em nossos dados, o que veremos nas considerações finais deste capítulo.

Igualmente presente no terceiro estágio de aquisição no BELC, o morfema *Irregular Past* se assemelha a *Reg Past* e *III Sing* no sentido de ser uma categoria IP, que, embora irregular, também é um afixo, por isso, emergiria nos últimos estágios de aquisição, à luz da perspectiva teórica de Zobl e Liceras.⁹⁹

⁹⁸ Ainda, Andersen menciona claramente que há vários possíveis determinantes para a ordem de aquisição e que o pesquisador deve lidar com a sistematicidade e a variabilidade destes possíveis determinantes, que podem ser categoria sintática e tipo de morfema (conforme mencionados acima), frequência no *input*, semelhança com a L1, complexidade semântica e sintática e saliência perceptual, fatores todos mencionados nos capítulos 2 e 3, principalmente na análise de Goldschneider e Dekeyser (2005), que será retomada a seguir.

⁹⁹ Como o leitor deve lembrar, pesquisamos a evolução do processo de aquisição de treze morfemas no BELC, mas, para melhor compararmos estes resultados com as pesquisas de referência, desconsideramos alguns deles nesta parte da análise. Um dos morfemas desconsiderados é o

Dessa forma, podemos dizer que as explicações oferecidas por esta perspectiva teórica, que, em linhas gerais, assume fatores exclusivamente sintáticos como determinantes para uma possível ordem no processo de aquisição de morfemas em inglês como L1 e L2, não se aplicam plenamente aos nossos resultados. Os autores não conseguiriam explicar porque, por exemplo, *-ING* emerge no terceiro estágio e *Possessive*, no segundo. Ainda, é necessário um conjunto de pressuposições teóricas que não são postas claramente pelos autores para podermos entender algumas explicações propostas por eles nas hierarquias observadas nos estudos de referência (BROWN, 1973; KRASHEN, 1977).

A seguir, vejamos a interpretação dos nossos resultados segundo as considerações de Goldschneider & DeKeyser (2005).

5.2.3 A perspectiva de Goldschneider & DeKeyser (2005)

No capítulo 3, pudemos ver a abordagem utilizada por estes autores para explicar a ordem observada nos estudos dos morfemas com maiores detalhes. Retomamos, a seguir, as principais considerações deste estudo a fim de, posteriormente, tentarmos aplicá-las à nossa pesquisa.

5.2.3.1 Breve retomada da perspectiva teórica

Na tentativa de oferecer uma explicação abrangente para a sequência observada no processo de aquisição de morfemas em L2,¹⁰⁰ os autores selecionaram doze (12) estudos que apresentavam os mesmos morfemas investigados (a saber, *Present Progressive*, *Plural*, *Possessive*, *Articles*, *Regular Past*, *III Sing*) e atribuíram uma possível ordem à combinação de cinco (5) fatores diferentes, que são: frequência, saliência perceptual, regularidade morfofonológica, complexidade semântica e categoria sintática (cf. capítulo 3, seção 3.2.1). Observe-se, aqui, que se trata de reanalisar dados originados de estudos de natureza

Third Person Irregular, ou seja, a forma *has* a partir da flexão de terceira pessoa do verbo *have*. Interessantemente, este morfema é o último a ser adquirido, conforme podemos ver na tabela 35. Poderíamos igualmente atribuir essa manifestação tardia a propriedades de dependência de afixo, cf. Zobl & Licerias (1994).

¹⁰⁰ Salientamos que o estudo de Goldschneider & DeKeyser não tenta explicar dados sobre a ordem de morfemas em L1 como a de Brown (1973).

diferentes, incluindo estudos experimentais, outras meta-análises, etc. Portanto, são dados “não uniformes” com relação aos informantes e ao método de coleta.

A partir disso, os autores conduziram uma meta-análise de estudos já existentes: uma análise de regressão múltipla, que, a partir da aplicação de testes estatísticos, poderia oferecer resultados que pudessem prever quais os morfemas a serem adquiridos primeiro que outros pelas suas possíveis características inerentes associadas aos fatores acima.

Como vimos no capítulo 3, os resultados mostram que a ordem pode ser atribuída à combinação desses cinco fatores, e, segundo Goldschneider & DeKeyser, a partir da Tabela 7, é bastante difícil determinar qual dos cinco fatores é o mais importante. Eles concluem que frequência, saliência perceptual, regularidade morfofonológica, complexidade semântica e categoria sintática constituem aspectos de saliência no sentido amplo da palavra, e que alguns morfemas seriam adquiridos antes por serem intrinsecamente mais sobressalentes que outros.

Apesar de entendermos essa concepção como muito ampla e não detalhada em termos de explicações precisas para uma dada ordem de morfemas, tentaremos estabelecer algumas associações entre as explicações oferecidas e a ordem encontrada em nosso corpus.

Portanto, vejamos a análise dos resultados do BELC a partir destas ideias:

5.2.3.2 Aplicação aos resultados do BELC

Como já vimos na tabela 6, o estudo de Goldschneider & DeKeyser (2005) contempla apenas seis morfemas. Por isso, não podemos estabelecer previsões a partir deste trabalho sobre os morfemas *Copula*, *Auxiliar* e *Irregular Past*.

Além disso, o texto não fornece uma hierarquia clara de morfemas que possa ser comparada a outros resultados. Sabemos que o objetivo dos autores é levantar fatores que expliquem uma dada ordem. Porém, entendemos que seria muito interessante se tivéssemos uma sequência de morfemas claramente apresentada e também uma tabela com a atribuição de valores especificamente utilizada para a meta-análise.

Se observarmos a tabela 7, que propõe uma análise dos possíveis fatores determinantes de uma possível ordem de morfemas precisamente direcionada a prever quais os morfemas que deveriam ser esperados primeiro a partir de cada fator, podemos concluir que:

(i) O morfema *Plural*, presente no primeiro estágio em Krashen (1977) e no BELC, se destaca em relação aos outros a partir dos fatores frequência (está em terceira posição em relação aos outros morfemas) e categoria sintática (tem valor 3, pois é tido como uma categoria lexical presa, cf. capítulo 3, seção 3.2.1)¹⁰¹. Assim, existe a possibilidade de este morfema ser adquirido primeiro devido a esses fatores.¹⁰²

(ii) Com relação ao morfema artigo, os dados da tabela 7 nos dizem que, conforme o fator saliência perceptual, este morfema se salienta em relação aos demais nos índices referentes ao escore bruto sobre o número de fones, escore bruto de silábico ou não silábico (como não-silábico), sonoridade (estando somente abaixo do *-ING*) e frequência, sendo o mais frequente de todos com índice de 552. Outro dado interessante é que este morfema apresenta índice 4 na linha 8, que se refere ao escore bruto de alternâncias fonológicas, o mais alto comparado aos outros.¹⁰³ Entendemos que isso contribui para que este morfema seja adquirido primeiro que outros, por exemplo, principalmente com relação ao fator frequência.

(iii) O morfema *Possessive* se assemelha a *Reg Past* e a *III Sing* nos índices apresentados, especialmente a *III Sing*. Se comparados, podemos notar diferença mais significativa somente no fator complexidade semântica, onde o primeiro obteve valor 1 e o segundo valor 3 a partir da análise dos autores e, a partir da descrição no capítulo 3, seção 3.2.1, os morfemas com índice mais alto nesse critério devem ser adquiridos mais tarde.

¹⁰¹ Este morfema também se destaca em relação ao fator saliência perceptual, especificamente com relação à homofonia, pois se assemelha a *Possessive* e *III Sing*. No entanto, com relação a esse fator, há o que os autores chamam de correlação negativa: os valores mais altos indicam que o morfema deveria ser adquirido mais tarde.

¹⁰² Salientamos, mais uma vez, que os autores não oferecem qualquer tipo de explicação específica como a que acabamos de oferecer. Por isso, é difícil entender e descrever com precisão as pressuposições originais dos autores.

¹⁰³ Infelizmente, não sabemos detalhes desta atribuição, pois a consulta a um estatístico apenas explica o índice em termos de como foi calculado e não como foi obtido, e isso não está claro no artigo.

Esta observação é bastante relevante aos nossos resultados, pois um dos dados do BELC mostra *Possessive* antes de *III Sing*. Segundo as observações de Goldschneider & DeKeyser (2005), *Possessive* deve ser adquirido anteriormente ao *III Sing* (pelo menos a partir da previsão deste fator) porque é menos complexo semanticamente. Como sabemos, esse resultado é diferente da maioria das pesquisas sobre aquisição de morfemas em L2, em especial a de Krashen (1977), que é tida como pesquisa de referência.

Assim, podemos concluir que as previsões com relação ao morfema *Possessive*, mais especificamente com relação ao fator complexidade semântica, poderiam explicar os nossos resultados, mas não os de Krashen.¹⁰⁴

(iv) Já o morfema *-ING* apresenta índices altos de previsibilidade de aquisição: os escores referentes ao número de fones são claramente mais altos do que com relação a outros morfemas, assim como o critério sonoridade, todos integrantes do fator saliência perceptual. Este morfema também se destaca em relação aos outros quanto aos fatores categoria sintática e frequência. Assim, segundo a previsão de Goldschneider & DeKeyser (2005), este morfema se salienta perante os demais e, por isso, deve ser esperado no início do processo de aquisição. Isso explica os dados de Krashen (1977), mas não os do BELC, pois, como já vimos, o nosso corpus apresenta *-ING* no terceiro estágio de aquisição.

A análise de Goldschneider & Dekeyser não contempla o fator transferência da L1, embora os autores mencionem (cf. capítulo 3, seção 3.2.1) que este é um fator muito relevante. Mesmo se considerarmos este fator, não poderíamos atribuir a diferença na posição do morfema *-ING* entre a hierarquia apresentada no BELC e por Krashen (1977), pois a estrutura em inglês é semelhante à do português brasileiro.

No entanto, julgamos que a contribuição de Goldschneider & Dekeyser é válida no sentido de pensar em fatores múltiplos aos quais podemos atribuir uma sequência natural na aquisição de morfemas em inglês como L2.

Em síntese, os fatores determinantes para a ordem de aquisição de morfemas postulados por Goldschneider & Dekeyser (2005), embora vagos, são interessantes não somente no sentido de tentarem explicar o porquê de uma dada hierarquia no processo de aquisição de morfemas em inglês como L2, mas,

¹⁰⁴ Novamente, salientamos que isso não é mencionado no artigo.

principalmente, no sentido de postularem explicações de natureza multifatorial para isso. Ou seja, ao passo que Zobl & Liceras tentam explicar uma possível sequência de morfemas a partir de fatores estritamente sintáticos, Goldschneider & Dekeyser propõem uma abordagem mais ampla, considerando fatores de naturezas diferentes que possam dar conta de uma ordem de estruturas morfêmicas.

Entretanto, como vimos, nenhuma das propostas conseguiu explicar exatamente a ordem de aquisição de morfemas obtida através do BELC. Sabemos que o objetivo principal do trabalho era o de obter uma sequência no processo de aquisição de morfemas em inglês como L2 a fim de podermos estabelecer um padrão evolutivo deste processo a partir dos nossos aprendizes, para, num segundo momento, podermos comparar este padrão a outros estudos já realizados e, então, pensar em possíveis explicações para uma ordem obtida. Isso foi feito.

Conduzimos uma extensa revisão bibliográfica nas áreas de aquisição de L2, morfossintaxe e linguística de corpus, e procuramos delimitar nosso trabalho contemplando as interfaces desses grandes domínios. Após muita preparação, procedemos à coleta de dados, e, num segundo momento, à elaboração de um banco de dados cuidadosamente organizado que constitui o BELC. Em seguida, etiquetamos o corpus para os fins desta pesquisa e extraímos os dados a fim de obtermos uma hierarquia de morfemas que serviria de base à nossa análise. Então, pudemos comparar nossos resultados com dados das pesquisas de referência e observamos algumas diferenças, especialmente no que se refere aos morfemas *Possessive* e *-ING*. Como vimos, os dois estudos que selecionamos para tentarmos explicar uma dada ordem de aquisição de morfemas não nos parecem suficientemente consistentes ao oferecerem explicações claras e precisas acerca do assunto, e, por isso, conseqüentemente, não conseguem explicar totalmente os resultados do BELC.

Portanto, podemos pensar em outros fatores aos quais poderíamos atribuir a hierarquia obtida em nossos resultados. Após analisarmos os fatores já especulados nas pesquisas utilizadas nesta tese, pensamos em considerar duas questões: o fator instrução e a natureza da tarefa de coleta dos dados (tarefa escrita). Desta forma, propomos, a seguir, uma breve análise sobre os possíveis efeitos destas questões em nossos resultados.

5.3 OUTRAS QUESTÕES A SEREM CONSIDERADAS

Nesta seção, apresentamos brevemente uma discussão sobre fatores que possam ser relevantes aos resultados que obtivemos após a análise do BELC. Dado o enfoque bem definido do estudo desta tese, é impossível, inapropriado e até mesmo desnecessário que nos estendamos em detalhes sobre questões de natureza teórica mais específicas que fundamentam conceitos como “instrução” e “diferenças entre tarefas”.

Assim, tentaremos, a seguir, associar noções básicas sobre esses tópicos aos nossos resultados.

5.3.1 O Fator Instrução

Ellis (1994) nos diz que o termo “instrução formal” foi, por muito tempo, utilizado para referir ao ensino de gramática, e que é necessário que esse escopo seja aumentado, considerando tipos de instrução formal aquelas com objetivos cognitivos, como por exemplo: (i) instrução centrada na língua (aspectos lexicais, gramaticais, discursivos e de pronúncia e (ii) instrução centrada no aluno (uma adequação da instrução ao aluno) e aquelas com objetivos metacognitivos (treinamento de estratégias).

De um modo geral, isso nos leva a refletir sobre até que ponto funcionaria a instrução direcionada em ensinar regras linguísticas específicas.

Ellis (1994) nos diz ainda que há um vasto número de pesquisas que investigam se a instrução formal resulta em melhor aprendizagem de L2 e que estas pesquisas podem ser de quatro tipos: (i) as que investigam se aprendizes que receberam instrução formal alcançam níveis mais altos de proficiência que aqueles que não foram orientados; (ii) as pesquisas que consideram os efeitos da instrução formal no grau de correção no qual os aprendizes usam elementos linguísticos específicos; (iii) aquelas que estudam se a instrução formal afeta a ordem ou sequência de aquisição e (iv) as que investigam se os efeitos da instrução são duradouros.

Naturalmente, delimitamos a nossa breve análise às pesquisas do grupo (iii).

A questão que se impõe naturalmente é: os aprendizes de sala de aula, que presumidamente receberam algum tipo de instrução gramatical, manifestam uma ordem de aquisição diferente daqueles aprendizes que estão num ambiente natural de aquisição?

Dois estudos que já mencionamos nos capítulo 2 (seção 2.2.2) obtiveram resultados que mostram que a instrução tem efeito mínimo ou insignificante numa possível sequência de aquisição.

Perkins & Larsen-Freeman (1975) conduziram um estudo com doze sujeitos venezuelanos através de duas tarefas: a primeira de tradução e a segunda baseada num filme sem diálogos. Na primeira tarefa, as ordens de morfemas foram diferentes antes e depois da instrução, mas na segunda tarefa, a sequência observada não foi diferente. A conclusão é que a instrução formal parece apenas ter um efeito mínimo na ordem dos morfemas evidenciada em uso de linguagem espontânea.

Pica (1983) desenvolveu uma pesquisa que investiga a produção de morfemas gramaticais em inglês como L2 por dezoito (18) falantes nativos de espanhol em três diferentes condições de exposição: (i) só com instrução; (ii) ambiente natural e (iii) misto (instrução e ambiente natural). Os sujeitos foram expostos a conversas individuais com uma pesquisadora – logo, o método utiliza dados orais. Os resultados sugerem que as diferentes condições de exposição ao inglês como L2 não alteram significativamente a ordem de correção na qual os morfemas gramaticais são produzidos.

Ao examinarmos mais cautelosamente os dados de Pica com relação ao morfema *-ING*, podemos ver que esta estrutura está no topo da hierarquia dos três grupos investigados, diferindo dos nossos resultados.

Nesta perspectiva, temos ainda Trosborg (1994):

Summing up, formal instruction seems to have only a negligible effect on the morpheme order manifest in spontaneous language use. The morpheme order is either the same or it differs only in short term, or only in one or two features which may have been overlearned. In addition, the order may differ when the learner's attention is focused on form, rather than meaning (p. 72).

No entanto, Ellis (1994) nos diz que os resultados obtidos por estudos como os de Pica (1983) e Perkins & Larsen-Freeman (1975) são evidências fracas de que a instrução formal não tem nenhum efeito numa possível sequência de aquisição

dos morfemas. Há questões metodológicas sobre as pesquisas: como já vimos, estudos longitudinais são considerados ideais para este tipo de investigação, cf. Rosansky (1976), porque mediriam a aquisição em si e não a correção na performance dos aprendizes.

Segundo Ellis (1994) há, ainda, uma questão teórica sobre aquisição de L2 que subjaz estas considerações: os aprendizes não adquirem a gramática da L2 como um conjunto de “entidades acumuladas”, e sim trabalham com vários traços simultaneamente, escolhendo gradualmente as relações de forma e significado que estas estruturas codificam. Assim, espera-se que o aprendiz seja guiado por uma seleção natural de itens que estabeleçam as relações linguísticas mais relevantes no contexto comunicativo com a ajuda do contexto linguístico ao qual está sendo exposto.

Analisando o material didático utilizado pelos aprendizes, o livro *American Inside Out*, vemos que a estrutura de *Present Progressive* é somente apresentada aos alunos no nível 2 e revisada nos níveis 3 e 4, estágios onde identificamos o índice de 90% de correção na produção dos nossos aprendizes. Ainda, a estrutura *Possessive* é apresentada já no nível 1, na segunda unidade. Por isso, acreditamos que, apesar de pesquisas não considerarem a instrução como um fator tão significativo, os resultados do BELC evidenciam fortes indícios de que os nossos aprendizes possam ter sido influenciados pela exposição formal da estrutura.¹⁰⁵

5.3.2 Tarefa escrita

Outro fator muito interessante a ser considerado em nossos resultados é a natureza do método de coleta de dados. Vimos no capítulo 2, mais especificamente na seção 2.2 sobre estudos dos morfemas em L1 (2.2.1) e em L2 (2.2.2), que o método é uma questão que pode ser considerada como interferente em estudos sobre uma possível sequência na aquisição de morfemas em inglês. Como vimos, foram observadas algumas diferenças nas várias pesquisas feitas, mas nada que

¹⁰⁵ Como alternativa para futuras pesquisas a partir desta tese, poderíamos propor um estudo comparativo entre a sequência que obtivemos e a ordem no desenvolvimento dos conteúdos a partir do material didático utilizado pelos sujeitos. Assim, poderíamos estabelecer semelhanças e diferenças teoricamente mais consistentes com relação ao efeito da instrução formal, uma vez que isso sairia do escopo deste trabalho.

causasse grande impacto na elaboração de uma ordem natural, tanto que o estudo de Krashen (1977) apresenta esta sequência utilizada como referência e, como sabemos, tal sequência é oriunda de um estudo proveniente de várias outras pesquisas de orientações metodológicas diferentes. Sobre o assunto, Ellis (1994) nos diz: *One possibility, then, is that different orders exist for the oral and written learner language. This finding is not problematic where production is concerned, as speaking and writing are influenced by different sociolinguistic and psycholinguistic conditions.* (p. 91)

Os resultados do BELC foram obtidos a partir de textos escritos. Logo, são de natureza de produção escrita.

Entendemos que, em tarefas escritas, mesmo que tenham um tempo controlado para desenvolverem a tarefa (como foi o caso do BELC, onde os alunos tinham 30min para a produção da tarefa, cf. capítulo 4, seção 4.2.1.4), os sujeitos dispõem de condições mais favoráveis para elaborar o texto. Há mais tempo para estruturarem melhor o conteúdo a ser comunicado: os sujeitos podem reler o texto e fazer quaisquer alterações que julguem necessárias. Por isso, há uma tendência natural que textos escritos apresentem menos erros do que textos orais.

Analisando especificamente a estrutura do *Present Progressive*, entendemos que o fato de os alunos serem expostos formalmente a este conteúdo no nível 2 somado à característica da tarefa – coleta de dados escritos – possa ter influenciado os nossos resultados: os alunos dominariam a estrutura depois do período de instrução formal, numa tarefa escrita, onde necessitariam da forma completa *Be + verbo + ING*. Os nossos dados confirmam ainda que o morfema *-ING* emerge no terceiro estágio juntamente com a forma Auxiliar, diferentemente de outras pesquisas, como a de Krashen (1977), onde *-ING* aparece no primeiro estágio e *Auxiliary* no segundo.

Como professores de inglês, sabemos que há uma forte tendência dos alunos a produzirem, em níveis iniciais, orações como “*She working*”, com a estrutura incompleta do *Present Progressive*, mas com a forma mais significativa em termos de conteúdo, o *ING*. Isso confirmaria os dados vistos em Krashen (1977) e em várias outras pesquisas, que colocam o morfema *-ING* no topo da hierarquia. No entanto, a grande maioria destas pesquisas, conforme vimos nos capítulos 2 e 3, utiliza dados de produção oral. Como os nossos dados são de produção escrita, isso pode ser diferente.

Ainda, o alto desempenho dos sujeitos com relação ao morfema *Possessive* também pode ter sido influenciado pela natureza da tarefa, pois, ao redigirem a oração, os sujeitos têm melhores condições até mesmo de visualizarem a forma e acrescentarem o 's ao nome que querem indicar posse.

Em resumo, neste capítulo, apresentamos os resultados em nossa pesquisa e propusemos a discussão e análise dos dados obtidos. Como vimos, a hierarquia observada na aquisição de morfemas pelos sujeitos desta pesquisa é bastante semelhante aos dois estudos de referência utilizados: Brown (1973) e Krashen (1977). Nossos resultados mostram: (i) *Copula, Plural* (ii) *Article e Possessive* (iii) – *ING, Auxiliary e Irregular Past* e (iv) *Regular Past, Ill Sing.* Esta sequência confirma parcialmente os resultados obtidos por Brown (1973) e Krashen (1977), pois, como vimos, há diferenças com relação aos morfemas –*ING* e *Possessive*. Vimos que a análise de Zobl & Liceras (1994) e Goldschneider & DeKeyser (2005) podem explicar parcialmente os nossos resultados. Ainda, levantamos alguns outros fatores – instrução e natureza da tarefa (escrita) – com possíveis intervenientes em nossa sequência. Dessa forma, passaremos, a seguir, às considerações finais do trabalho.

CONCLUSÃO

Ao longo desta tese, propusemo-nos a discutir a aquisição de morfemas em inglês, na condição de L2, por aprendizes brasileiros. A fim de obtermos um quadro mais claro sobre esse processo, propomos, de um lado, a apresentação da literatura referente ao tema e, de outro, a caracterização do padrão evolutivo específico de sujeitos brasileiros.

Para isso, o trabalho seguiu o seguinte roteiro:

No capítulo 2, tecemos algumas observações de caráter teórico com relação às estruturas que são abordadas na tese – os morfemas gramaticais do inglês e as categorias funcionais correspondentes. Ainda, apresentamos algumas das principais pesquisas que enfocam a aquisição de morfemas em inglês como L1 e como L2.

Como vimos, a pesquisa de Brown (1973) é utilizada por grande parte da literatura da área como referência no estabelecimento da ordem de aquisição de morfemas em inglês como L1. Por isso, adotamos, igualmente, a sequência de morfemas gramaticais apresentada por Brown para o nosso estudo. Lembremos que a ordem de aquisição descoberta por Brown é a seguinte: 1. *Present Progressive*; 2. *Plural*; 3. *Irregular Past*; 4. *Possessive*; 5. *Uncontracted Copula*; 6. *Article*; 7. *Regular Past*; 8. *Third Person Singular*; 9. *Uncontracted Auxiliary*. Do mesmo modo, os resultados de Krashen (1977) também assumem o *status* de referência nos estudos sobre a ordem de aquisição de morfemas do inglês como L2, razão pela qual também utilizamos este trabalho como uma das bases de nossa discussão. Segundo este autor, a ordem básica de aquisição de morfemas do inglês como L2 é a seguinte: 1. *Present Progressive, Plural, Copula*; 2. *Auxiliary, Article*; 3. *Irregular Past*; 4. *Regular Past, Third Person Singular, Possessive*. Além disso, vimos que outros estudos apresentam resultados que, de um modo geral, confirmam a existência de uma ordem natural na aquisição de morfemas.

Além dos resultados fundamentais de Brown e Krashen, utilizamos, nesta tese, como análises de referência para discutirmos nossos resultados, dois trabalhos que concentraram muito da atenção recente na literatura sobre a aquisição de morfemas em inglês como L2, a saber, Zobl & Liceras (1994) e Goldschneider & DeKeyser (2005). Como vimos, o artigo de Zobl & Liceras toma como *framework* de referência a chamada "teoria gerativa"; já o estudo de Goldschneider & DeKeyser

propõe uma análise próxima das abordagens conexionistas, enfatizando propriedades do *input* a que o aprendiz tem acesso. No capítulo 3, discutimos de modo detalhado estes dois estudos, procurando comparar a maneira como eles interpretam os resultados da literatura com relação à ordem natural de aquisição dos morfemas do inglês como L2, e como tentam explicar esta ordem.

Vimos que Goldschneider & DeKeyser (2005) propõem que a ação de vários fatores é o que daria conta da ordem que investigaram; por isso, sua análise pode ser caracterizada como multifatorial. Os fatores propostos seriam: saliência perceptual do morfema; complexidade semântica; regularidade morfofonológica; categoria sintática e frequência. Mencionamos, também, que estes autores entendem que todos estes fatores são, na verdade, diferentes manifestações de uma "força básica", que identificam por meio de uma noção geral de "saliência". Como dissemos, a proposta nos parece interessante porque oferece flexibilidade de análise, já que admite que mais de um fator pode ter um papel importante na ordem de aquisição dos morfemas observada nos estudos de referência. Por outro lado, também nos parece que a proposta perde um pouco de sua objetividade conceitual quando os autores consideram que os cinco fatores mencionados são "manifestações" de um conceito amplo de "saliência".

Já com relação à proposta de Zobl & Liceras (1994), vimos que eles sugerem que a ordem de aquisição de morfemas em inglês pode ser explicada, basicamente, por fatores sintáticos. De acordo com os autores, o processo de aquisição em L1 pode ser caracterizado pela hipótese de que a aquisição das categorias lexicais (N, V, Adj, etc.) precede a das categorias funcionais (Det, INFL, COMP). Já em L2, os autores apontam para o fato de que o aprendiz já dispõe – porque as adquiriu no processo de aquisição nativa – das estruturas funcionais; assim, em L2, o processo de aquisição de morfemas é, antes, orientado pela distinção entre morfemas livres e presos, sendo os primeiros, para os autores, aqueles observados no topo da sequência.

Como procuramos apontar no capítulo 5 – ao qual voltaremos logo abaixo – ambas as abordagens podem dar conta de certos aspectos dos dados que analisamos nesta tese; mas nenhuma delas cobre todo o conjunto de fatos revelados, de modo que somos obrigados a recorrer a hipóteses adicionais.

Quanto aos dados desta tese, já enfatizamos antes que são provenientes de um amplo corpus de produções de aprendizes, o BELC – *Brazilian English Learner*

Corpus. Este corpus, bem como a metodologia utilizada para a sua constituição, foram apresentados no capítulo 4. As principais características do corpus é que se trata de um conjunto de textos produzidos por mais de 400 alunos brasileiros, todos universitários, distribuídos em 8 diferentes níveis de aprendizagem. O corpus oferece, a nosso ver, um quadro bastante representativo das possibilidades evolutivas do conhecimento gramatical do inglês neste tipo de aprendiz.

No capítulo 5, apresentamos, então, os dados obtidos com o BELC e propusemos a discussão e a análise destes resultados. Como vimos, a hierarquia observada na aquisição de morfemas pelos sujeitos do BELC é, de certa forma, semelhante às hierarquias obtidas nos dois principais estudos de referência utilizados, Brown (1973) e Krashen (1977). Nossos resultados mostram a seguinte ordem de aquisição dos morfemas do inglês por aprendizes brasileiros: (i) *Copula e Plural*; (ii) *Article e Possessive* (iii) *-ING, Auxiliary e Irregular Past*, e (iv) *Regular Past e III Sing*. Como vimos, esta sequência confirma, ainda que parcialmente, uma ordem natural do padrão evolutivo na aquisição de morfemas em inglês como L2. A principal diferença com relação aos dados de Krashen e Brown é a ausência do morfema *-ING* no primeiro estágio de aquisição: nos dados obtidos com o BELC, este morfema emerge apenas no terceiro estágio. Há, além disso, uma outra diferença interessante relativa à ordem de aquisição do morfema *Possessive*: em Krashen (1977), pode ser visto no final da hierarquia; no BELC, ele se apresenta no segundo estágio de aquisição, posição semelhante à observada nos resultados de Brown (1973) – que dizem respeito ao inglês como L1, e não à aquisição do inglês como L2. Vale a pena lembrar, neste ponto, que vimos também que outros estudos mostram o morfema *Possessive* sendo utilizado em níveis altos em estágios iniciais de aquisição de L2. Observamos, por outro lado, que tais estudos descartaram a importância deste resultado: Izumi & Isahara (2004) atribuem-no à interferência da L1 de seus sujeitos (japoneses), e Wagner (2007) desconsidera o resultado por razões metodológicas (seus dados foram elicitados experimentalmente, isto é, não refletem produções espontâneas dos aprendizes). Nenhuma destas justificativas poderia explicar os nossos resultados; ao contrário, parece que nossos resultados colocam uma nova perspectiva para os estudos acima mencionados, que deveriam ter seus dados reavaliados.

Ainda no capítulo 5, buscamos verificar se as principais análises que discutimos ofereceriam uma explicação para a ordem que encontramos – e, especialmente, para os fatos relativos aos morfemas *-ING* e *Possessive*.

Tentamos, primeiramente, explorar as contribuições e propostas de Zobl & Liceras (1994). Como vimos acima, para estes autores, os aprendizes de uma L2 já disporiam de estruturas correspondentes às categorias funcionais e os estágios de aquisição em L2 seriam reflexos, inicialmente, de princípios de análise de morfemas livres e presos. Especificamente: os morfemas livres seriam adquiridos antes e os presos, depois; ainda, os "afixos lexicais" também viriam antes dos "afixos sintáticos". Os autores consideram *-ING* e *plural* como "afixos lexicais", e *III Sing* como "flexão sintática". Como vimos na seção 5.2.2, esta abordagem não parece capaz de explicar claramente, nem na própria análise de Zobl & Liceras, por que o morfema *Possessive* aparece por último em Krashen (1977), uma vez que é DP, um afixo nominal e não poderia ser comparado a *III Sing* e *Reg Past*, pois são afixos sintáticos. No BELC, *Possessive* aparece junto a *Article*, no segundo estágio evolutivo. Sabemos que ambos os morfemas são instâncias da categoria Det. Logo, no que diz respeito a tal categoria, nossos resultados parecem se assemelhar aos dados de Brown (1973). Além disso, as explicações oferecidas por Zobl & Liceras a este estudo sobre a aquisição de morfemas em L1 possibilitam ser aplicadas ao nosso estudo em L2 com relação a *Article* e *Possessive*.

Consequentemente, as considerações feitas pelos autores explicariam parcialmente os nossos resultados. Tal como eles entendem o papel das categorias funcionais em processos de aquisição – especialmente, a hipótese de que seriam pertinentes apenas em processos de aquisição de L1 –, não poderíamos explicar a ocorrência do morfema *Possessive* no segundo estágio juntamente com *Article*, bem como a do morfema *-ING* no terceiro estágio, junto a *Auxiliary*. Tais fatos poderiam indicar que há alguma relação entre estágios e categorias funcionais na aquisição de morfemas do inglês por aprendizes brasileiros.

Na seção 5.3.3, procuramos explorar nossos resultados a partir das propostas de Goldschneider & DeKeyser (2005). Como dissemos antes, estes autores se propõem a explicar a ordem de aquisição dos morfemas a partir de fatores como saliência perceptual, complexidade semântica, regularidade morfofonológica, complexidade sintática e frequência. Concretamente, para fins de análise, os autores atribuem "pontos" correspondentes a estes fatores, de modo que

a classificação de cada morfema dentro destas categorias resulta num certo "valor"; o qual, por sua vez, indicaria a posição do morfema na "ordem de prioridade" de aquisição. Especificamente aplicado aos nossos resultados, esse método explicaria porque o morfema *Possessive* surge num estágio anterior ao encontrado em outras pesquisas: devido a suas propriedades de saliência, *Possessive* seria mais facilmente observado do que outros. Ou melhor, a análise de Goldschneider & DeKeyser (2005) aponta uma diferença do morfema *Possessive* comparado a *III Sing* e *Regular Past* com relação ao fator complexidade semântica: ao primeiro, Goldschneider & DeKeyser (2005) atribuem 1 ponto, e aos outros, respectivamente, 2 e 3, e, segundo os autores, os morfemas mais complexos – aqueles que têm o escore mais alto – devem ser adquiridos mais tarde. Por isso, a partir desta perspectiva, *Possessive* deveria ser esperado antes dos outros.

Tendo concluído que nenhuma das análises consideradas oferece uma abordagem completamente satisfatória para nossos resultados, procuramos pensar em outros fatores que pudessem oferecer uma explicação alternativa. Sugerimos, ao fim do capítulo 5, que o fator *instrução* deve ser considerado, por exemplo, com relação à ocorrência do morfema *-ING* no terceiro estágio no BELC. Outro fator que apontamos é a natureza da tarefa utilizada para a coleta dos dados do BELC: tratava-se de uma tarefa de produção de um texto escrito, com duração de cerca de 30 minutos, o que poderia oferecer aos alunos condições de produção nas quais as habilidades metalinguísticas teriam chance de atuar; assim, os alunos estariam em melhores condições de produzir formas mais corretas. Além disso, os alunos já foram apresentados à estrutura completa do *Present Progressive*, que seria *Auxiliary* e *ING* no nível 2 do curso. Assim, os nossos dados parecem apontar para o fato de que os sujeitos tendem a utilizar a estrutura completa por já terem sido expostos a ela e por terem que desempenhar uma tarefa escrita, que naturalmente estimularia o emprego de construções linguísticas mais ricas.

Em resumo, em relação aos nossos resultados, concluímos que as explicações dos referenciais teóricos selecionados não são totalmente satisfatórias porque parecem fragmentadas. Isto é, não podemos atribuir nossos resultados a uma única perspectiva de análise. Como vimos, os estudos de referência que utilizamos para isso (ZOBL & LICERAS, 1994 e GOLDSCHNEIDER & DEKEYSER, 2005), juntamente com os fatores interferentes que apontamos ao final do trabalho

(instrução e tarefa escrita),¹⁰⁶ só poderiam explicar os nossos dados separadamente, tomando a ordem de cada morfema isoladamente. Do ponto de vista teórico, a explicação seria satisfatória se pudesse dar conta dos resultados a partir de uma única perspectiva de análise. No entanto, acreditamos que a contribuição de diferentes pontos de vista tenha enriquecido muito nossa discussão.

Poderíamos, futuramente, conduzir um estudo que proponha a comparação dos nossos dados à sequência de conteúdos do material didático utilizado pelos sujeitos, a fim de podermos traçar algumas semelhanças e/ou diferenças e possíveis conclusões a partir disso. Ainda, outra possibilidade de investigação que se apresenta é a análise do nível de proficiência dos sujeitos com relação a cada nível e a cada morfema. Poderíamos propor a análise do fator proficiência como interferente ou não nos resultados obtidos no BELC.

Ao concluir este trabalho, é importante, ainda, mencionarmos suas implicações pedagógicas.

A primeira delas diz respeito ao fato de termos caracterizado um padrão evolutivo que é representativo do processo de aquisição de morfemas gramaticais do inglês por aprendizes brasileiros, e representativo da situação de aprendizagem característica de cursos regulares de idioma no Brasil. A simples existência de uma ordem característica de aquisição de morfemas, específica de aprendizes brasileiros e do ensino de inglês no Brasil, pode oferecer ao professor de inglês atuando no Brasil um melhor entendimento dos estágios evolutivos de aquisição gramatical de seus alunos. Isso pode auxiliá-lo a ordenar melhor os conteúdos e as atividades, conduzindo os aprendizes de uma maneira mais consciente nas tarefas propostas e elaborando exercícios que possam estimulá-los de forma mais orientada, mais apropriada. É claro que este tipo de informação forneceria subsídios importantes para análise de materiais didáticos já existentes e para a elaboração de novos materiais didáticos.

Além dos resultados específicos da investigação empírica que aqui fizemos, não podemos deixar de mencionar o que nos parece ser a indiscutível contribuição que significa a elaboração do BELC para novas pesquisas sobre a aprendizagem de inglês como L2 no Brasil. Por ser um corpus de aprendiz organizado com base em princípios da linguística de corpus, abarcando um conjunto amplo de dados – mais

¹⁰⁶ Embora não tenhamos proposto uma análise mais profunda destes fatores.

de 400 sujeitos, mais de 100 mil palavras, mais de 900 textos de aprendizes - contemplando 8 níveis de aprendizagem de inglês como L2, o BELC sem dúvida é um recurso importante para quem se interessa pelo estudo das habilidades linguísticas de aprendizes brasileiros do inglês. As informações nele contidas podem ser exploradas de diversas formas e, considerando diferentes perspectivas; pode-se, por exemplo, estudar aspectos semânticos, lexicais, discursivos, etc., presentes nos textos, além de, é claro, estruturas sintáticas que são apresentadas a partir de uma fonte genuína de língua de aprendiz.

A análise que propusemos enfoca a aquisição de morfemas, que são, em sua essência, estruturas morfossintáticas relevantes do inglês. Sabemos da condição da língua inglesa como língua franca e do extenso número de dialetos e variação linguística que este idioma naturalmente assume em virtude desta condição. Sabemos, além disso, do advento dos métodos de ensino que priorizam efetivamente a comunicação desde meados dos anos de 1980. No entanto, não podemos ignorar o fato de que, para mantermos um grau de desempenho comunicativo satisfatório, é preciso observar alguns padrões linguísticos formais. Por exemplo, para relatarmos uma experiência sobre o que fizemos no último fim de semana, a utilização de advérbios não nos parece suficiente para expressarmos o tempo da ação – e, muito menos, suficiente para fazer isso de maneira "fluente". É desejável que utilizemos os verbos marcados no passado para expressarmos satisfatoriamente determinados aspectos, e, para isso, devemos fazer uso de morfemas. Situações comunicativas com algum grau de sucesso podem até ser estabelecidas sem o uso correto de construções gramaticais da língua; mas é claro que este tipo de objetivo não deve se constituir, consensual e regularmente, no objetivo de médio e longo prazo do ensino de língua inglesa. Entendemos que os profissionais do ensino de inglês não trabalham para que os alunos se expressem apenas de forma a comunicarem suas idéias mais básicas, de modo rudimentar, utilizando elementos linguísticos mínimos que assegurem apenas algum grau de transparência da informação. Como bons profissionais do ensino de inglês, devemos almejar que nossos alunos se expressem da melhor maneira possível, com a liberdade de ação que deve caracterizar qualquer agente linguístico pleno, e isso requer a observação de construções gramaticais e padrões morfossintáticos corretos. Dessa crença deriva não somente a motivação deste trabalho, mas a certeza de sua contribuição para aqueles profissionais que acreditam que, quanto

mais entendermos o caminho trilhado pelo nosso aluno, melhor poderemos ajudá-lo nesse processo.

REFERÊNCIAS

AARTS, B. & McMAHON, A. *The handbook of english linguistics*. Oxford: Blackwell, 2006.

ABNEY, S. *The english noun phrase in its sentencial aspect*. Doctoral dissertation, MIT, Cambridge, MA, 1986.

ANDERSEN, R. The impoverished state of cross-sectional morphology acquisition/accuracy methodology. *Working Papers on Bilingualism*, 14, 49–82, 1977

ANDERSEN, R. W. An implicational model for second language research. *Language Learning*, n. 28, p. 221-82, 1978.

ARCHIBALD, J. *Second language acquisition and linguistic theory*. Oxford: Blackwell, 2000.

BAILEY, N., MADDEN, C. & KRASHEN, S.D. Is there a “natural sequence” in adult second language learning? *Language Learning*, n. 24, p. 235-43, 1974.

BERBER SARDINHA, T. *Linguística de corpus*. São Paulo: Manole, 2004.

_____. *O corpus de aprendiz Br-ICLE*, 2001, disponível online em <<http://lael.pucsp.br/Tony/2001/bricle-interc.pdf>>.

BLOOMFIELD, L. *Language*. New York: Holt, 1933.

BROWN, R. *A first language*. Cambridge, MA: Harvard University Press, 1973.

BUBENIK, V. *An introduction to morphology*. Lincom Europa, Meuenchen, 1999.

CÂMARA JÚNIOR, J. Mattoso. *Princípios de lingüística descritiva*. Rio de Janeiro: Livraria Acadêmica, 1969.

CHOMSKY, N. *aspects of the theory of syntax*. Cambridge: MIT Press, 1965.

_____. *Lectures on Government and Binding*. Dordrecht: Foris, 1981.

_____. *Language and problems of knowledge*. The Managua Lectures. Cambridge, Massachusetts: MIT Press, 1988.

_____. *Principles and parameters theory*. Berlin: W. de Gruyter, 1991.

_____. *The minimalist program*. Cambridge: MIT Press, 1995.

CHOMSKY, N. & LASNIK, H. The theory of principles and parameters. In: CHOMSKY, N. (eds.). *The minimalist program*. Cambridge: MIT Press, p.13-127, 1995.

COBB, T. Analysing late interlanguage with learner corpora: québec replications of three european studies. *The Canadian Modern Language Review*, n. 59, v. 3, p. 393-423, 2003, disponível online em <<http://www.er.uqam.ca/nobel/r21270/cv/LC3.htm>>.

COOK, V. *Chomsky's universal grammar*. Oxford: Basil Blackwell, 1988.

COOK, V. *Linguistics and second language acquisition*. New York: St Martin's Press, 1993.

_____. *Second language learning and language teaching*. London: Arnold Publishers, 2001.

CORDER, P. The significance of learners' errors. *International Review of Applied Linguistics*, n. 5, p. 161-70, 1967.

_____. *Error analysis and Interlanguage*. Oxford: Oxford University Press, 1981.

De VILLIERS, J. & De VILLIERS, P. A cross-sectional study of the acquisition of grammatical morphemes in child speech. *Journal of Psycholinguistic Research*, n. 2, p. 267-78, 1973.

DOUGHTY, C & LONG, M. *The handbook of second language acquisition*. United Kingdom: Blackwell Publishing, 2005.

DULAY, H. C. & BURT, M. K. Should we teach children syntax? *Language learning*, n. 23, p. 245-58, 1973.

_____; _____. Natural sequences in child second language acquisition. *Language Learning*, n. 24, p. 37-53, 1974.

_____; _____. Some remarks on creativity in second language acquisition. In: *Second language acquisition research: issues and implications*. New York: Academic Press, p. 65-89, 1978.

_____; _____ & KRASHEN, S. *Language two*. New York: Oxford University Press, 1982.

ELLIS, N. The associative-cognitive CREED. In: VAN PATTEN, B. & WILLIAMS, J. *Theories in second language acquisition – an introduction*. New York: Routledge, 2007.

ELLIS, R. *The study of second language acquisition*. Oxford: Oxford University Press, 1994.

_____. *Instructed second language acquisition*. Oxford: Basil Blackwell, 1990.

_____; BARKHUIZEN, G. *Analysing learner language*. Oxford: Oxford University Press, 2005.

_____ & LAPORTE, N. Contexts of Acquisition: Effects of formal instruction and naturalistic exposure on second language acquisition. In: GROOT, A. M. B. de & KROLL, J. F. (eds). *Tutorials in bilingualism: psycholinguistic perspectives*. Mahwah/NJ: Erlbaum, p. 53-83, 1997, disponível em: <http://books.google.com.br/books?hl=pt-BR&lr=&id=IEEDufXc-nsC&oi=fnd&pg=PA53&dq=Ellis+%26+Laporte,+1997&ots=nmCiQ2tay8&sig=ljqNCfT0zWfIyez-OPQXWewWAv4#v=onepage&q=&f=false>

EUBANK, L. On the transfer of parametric values in L2 development. *Language acquisition*, v. 3, p. 183- 208, 1993.

FATHMAN, A. Language background, age and the order of acquisition of English structures. In: BURT, M.K. & DULAY, H. C. (eds.). *On TESOL '75: new directions in second language learning, teaching and bilingual education*. Washington, DC: TESOL, p. 33-43, 1975.

FLYNN, S. *Contrast and construction in a parameter-setting model of L2 acquisition*. *Language Learning* (37), p. 19-62, 1987.

FROMKIN, V.; RODMAN, R & HYAMS, N. *An introduction to language*. Boston, MA: Thomson Wadsworth, 2007.

FUKUI, N. & SPEAS, M. *Specifiers and projection*. Manuscript, MIT, 1986.

FULLER, J. An investigation of natural and unmonitored morpheme difficulty order by non-native adult students of English. Unpublished Doctoral Dissertation, Florida State University, Tallahassee, FL, 1978. In: GOLDSCHNEIDER, J. & DEKEYSER, R. 2005. Explaining the “natural order of L2 morpheme acquisition” in English: a meta-analysis of multiple determinants. *Language Learning*, n. 55, Supplement 1, p. 27-75, 2005.

GASS, S & SELINKER, L. *Second language acquisition: an introductory course*. Hillsdale, Nj: Erlbaum, 1994.

_____; _____. *Second language acquisition: an introductory course*. 3.ed. Routledge: New York, 2008.

GASSER, M. Connectionism and universals of second language acquisition. *Studies in second language acquisition*, (12), p. 179- 99, 1990.

GOLDSCHNEIDER, J. & DEKEYSER, R. 2005. Explaining the “natural order of L2 morpheme acquisition” in English: a meta-analysis of multiple determinants. *Language Learning*, n. 55, Supplement 1, p. 27-75, 2005.

GRANGER, S. A. Bird’s eye view of learner corpus research. In: GRANGER, S., HUNG, J., and TYSIN, P.S. (eds). *Computer learner corpora, second language acquisition and foreign language teaching*. Amsterdam: John Benjamins Publishing Company, 2002.

_____. *Computer-aided error analysis*. System, n. 26, p. 163-74, 1998.

_____. The international corpus of learner english: a new resource for foreign language learning and teaching and second language acquisition research. In: www.eric.ed.gov/ERICWebPortal/recordDetail?accno=EJ677136, 2003.

HADLEY, A. O. *Teaching language in context*. Boston: Heinle & Heinle, 1996.

HAEGEMAN, L. *Thinking syntactically: a guide to argumentation and analysis*, Malden, Blackwell, 2007.

HANUKA, K. A case study of a japanese child learning english as a second language. *Language Learning*, n. 26, p. 321-51, 1976.

HAZNEDAR, B. Missing surface inflection in adult and child L2 acquisition. *Language Acquisition*, v. 3, p. 257-325, 2003.

HORNSTEIN, N.; NUNES, J.; GROHMANN, K. *Understanding minimalism*. Cambridge: Cambridge University Press, 2005.

HOUCK, N.; ROBERTSON, J. & KRASHEN, S. On the domain of the conscious grammar: morpheme orders for corrected and uncorrected ESL students transcriptions. *TESOL Quarterly*, n. 12, p.335-9, 1978.

HUSTON, S. *Corpora in applied linguistics*. United Kingdom: Cambridge University Press, 2002.

IZUMI, E.; UCHIMOTO, K & ISAHARA, H. *Error annotation for corpus of japanese learner english*. Kyoto, Japan, 2005.

_____; ISAHARA, H. Investigation into language learner's acquisition order based on an error analysis of a learner corpus. *IWLeL: An Interactive Workshop on Language Learning*, 2004.

JACOBS, R. & ROSEMBAUM, P. *English transformational grammar*. Massachussets: Blaisdell, 1968.

KATZ, J. & POSTAL, P. *An integrated theory of linguistic descriptions*. Cambridge.: MIT Press, 1964.

KINDERMANN, C. Linguística de corpus e as novas tecnologias de informação e comunicação (ITCs): uma interface necessária na formação do professor de língua inglesa sob uma perspectiva crítica, social e educacional. *Revista Domínios da Linguagem*, 2009. Disponível em: <www.dominiosdelinguagem.org.br/pdf/09-07-09/Texto%207.pdf>.

KRASHEN, S. Some issues relating to the monitor model. In: BROWN, H.; YORIO, C. & CRYMES, R. (eds): *On TESOL '77*. Washington DC: TESOL, 1977.

_____. *Second language acquisition and second language learning*. Pergamon Press: New York, 1981.

_____. *The input hypothesis: issues and implications*. New York: Longman, 1985.

_____; MADDEN, C. & BAILEY, N. Theoretical aspects of grammar sequencing. In: BURT, M. K. & DULAY, H. C. (eds). *New directions in second language learning, teaching and bilingual education*. Washington DC: TESOL, 1975, p. 44-54.

_____; SFERLAZZA, V.; FELDMAN, L. & FATHMAN, A. Adult performance on the SLOPE test: more evidence for a natural sequence in adult second language acquisition. *Language Learning*, n. 26, p. 145-51, 1976.

_____; & SCARCELLA, R. *Research in second language acquisition. Selected papers of the Los Angeles second language acquisition forum*. Rowley, Massachussets: Newbury House Publishers, 1978.

_____; SCARCELLA, R. & LONG, M. *Child adult differences in second language acquisition*. Rowley, Massachusstes: Newbury House Publishers, 1982.

KWON, E. The "natural order" of morpheme acquisition: a historical survey and discussion of three putative determinants. Teachers College, Columbia University working Papers. *TESOL & Applied Linguistics*, V. 5, n. 1, 2005.

LARDIERE, D. Mapping features to forms in second language acquisition. In: ARCHIBALD, John. *Second language acquisition and linguistic theory*. Oxford: Blackwell, 2000.

LARSEN-FREEMAN, D. E. The aquisition of grammatical morphemes by adult ESL students. *TESOL Quarterly*, n. 9, p. 409-19, 1975.

_____. An explanation for the morpheme acquisition order of second language learners. *Language Learning*, n. 26, p. 125-34, 1976.

_____. Making sense of frequency. *Studies in second language acquisition*, n. 24, p. 275-85, 2002.

_____.& LONG, M. *An introduction to second language acquisition research*. London: Longman, 1991.

LEECH, G. Introducing corpus annotation. In: GARSIDE, R.; LEECH, G & McENERY, A (eds). *Corpus annotation. Linguistic information from computer text corpora*. London: Longman, p. 1-18, 1997.

LONG, M. The least a second language acquisition theory needs to explain. *TESOL Quarterly*, n. 24, p. 649- 66, 1990.

LYONS, J. *Introdução à linguística teórica*. São Paulo: Nacional, 1979.

MACE-MATLUCK, B. The order of acquisition of certain oral english structures by native speakers of spanish, cantonese, tagalog and ilokano learning english as a second language between the ages of five and ten. Unpublished doctoral dissertation, University of Texas at Austin, 1977. In: GOLDSCHNEIDER, J. & DEKEYSER, R. Explaining the “natural order of L2 morpheme acquisition” in English: a meta-analysis of multiple determinants. *Language Learning*, n. 55, Supplement 1, p. 27-75, 2005.

McWHINNEY, B. Second language acquisition and the competition model. In: GROOT, A. & KROLL, J. (eds) *Tutorials in bilingualism: psycholinguistic perspectives*. Mahwah, NJ: Erlbaum, p.113-42, 1997.

McCLELLAND, J. Parallel distributed processing: implications for cognition and development. Cap. 2, p. 8-45. In: MORRIS, R. (ed). *Parallel Distributed processing: implications for psychology and neurobiology*. Oxford: Clarendon Press, 1989.

McENERY, T. & GABRIELATOS, C. English corpus linguistics. In: AARTS, B. & McMAHON, A. *The handbook of english linguistics*. Oxford: Blackwell, 2006.

_____; & WILSON, A. *Corpus linguistics*. Edingurgh: Edinburgh University Press. 1996.

McENERY, T; XIAO, R & TONO, Y *Corpus-based language studies : An advanced resource book* .London : Routledge. .2006

MEISEL, J. The acquisition of syntax negation in French and German: contrasting first and second language development. *Second Language Research*, n. 13, p. 227-63, 1997.

- MIOTO, C.; SILVA, M. C.; LOPES, R. *Novo manual de sintaxe*. Florianópolis: Insular, 2005.
- ODLIN, T. Cross-Linguistic Influence. In: DOUGHY, C & Long, M. *The Handbook of Second Language Acquisition*. Oxford: Blackwell Publishing, 2003.
- PACHECO, A. *A aquisição dos pronomes him, her e it em inglês como L2: evolução do processo e aspectos de transferência do português brasileiro como L1*. Dissertação de Mestrado. Pontifícia Universidade Católica do Rio Grande do Sul. 2000.
- PERKINS, K. & LARSEN-FREEMAN, D. The effect of formal language instruction on the order of morpheme acquisition. *Language Learning*, n. 25, p. 237-43, 1975.
- PICA, T. Adult acquisition of English as a second language under different conditions of exposure. *Language Learning*, n. 33, p. 465-97, 1983.
- _____. Methods of morpheme quantification: their effect in the interpretation of second language data. *Studies in Second Language Acquisition* 6, p. 69-78, 1984.
- PINKER, S. *Language learnability and language development*. Cambridge: Massachussets: Harvard University Press, 1996.
- _____. & PRINCE, A. "Rules and Connections in Human Language". Cap. 9 (p. 182-99). In: MORRIS, R. (ed). *Parallel distributed processing: Implications for psychology and neurobiology*. Oxford: Clarendon Press, 1989.
- POLLOCK, J. Verb movement, UG, and the structure of IP. *Linguistic Inquiry*, v. 20, p. 365-424, 1989.
- PRAVEC, N. *Survey of learner corpora*. IAME, International Computer Archive of Modern and Medieval English, n. 26, abr. 2002.
- PREVOST, P. & WHITE, L. Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research*, v. 16, n. 2, p. 103- 33.
- RIDDLE, P. The relationship of the bilingual syntax measure to English morpheme development measure. Unpublished doctoral disseration, Temple University, Philadelphia, 1993. In: GOLDSCHNEIDER, J. & DEKEYSER, R.

Explaining the “natural order of L2 morpheme acquisition” in English: a meta-analysis of multiple determinants. *Language Learning*, n. 55, Supplement 1, p. 27-75, 2005.

ROSADO, L. The role of the monitor model in the acquisition sequence of twelve English morphemes by adult English as a second language learners from four different language groups. Unpublished doctoral dissertation, Texas A & I University, Kingsville, TX, 1986. In: GOLDSCHNEIDER, J. & DEKEYSER, R. Explaining the “natural order of L2 morpheme acquisition” in English: a meta-analysis of multiple determinants. *Language Learning*, n. 55, Supplement 1, p. 27-75, 2005.

ROSANSKY, E. Methods and morphemes in second language acquisition research. *Language Learning*, n. 26, p. 409-25, 1976.

RUMELHARDT, D. & McCLELLAND, J. “On learning the past tense of English verbs”. Cap. 18 (p. 216-71). In: RUMELHARDT, D & McClelland, J. (eds). *Parallel distributed processing: explorations in the microstructure of cognition*. V. 2: Psychological and Biological Models. Cambridge, Mass: MIT Press, 1986.

SARMENTO, S. *O uso dos verbos modais em manuais de aviação em inglês: um estudo baseado em corpus*. Tese de doutorado. Porto Alegre: Universidade Federal do Rio Grande do Sul, 2008.

SCHLATTER, M.; GARCEZ, P.; SCARAMUCCI, M. O papel da interação na pesquisa sobre aquisição e língua estrangeira: implicações para o ensino e avaliação. *Letras de Hoje*, n. 137, . p. 345-78. Porto Alegre: EDIPUCRS, 2004.

SELINKER, L. Interlanguage. *International Review of Applied Linguistics*, n. 10, p. 209- 31, 1972.

_____. *Rediscovering Interlanguage*. London: Longman, 1992.

SINCLAIR, J. EAGLES. *Preliminary recommendations on Corpus Typology*. <<http://www.ilc.pi.it/EAGLES96/corpusTyp/corpusTyp.html>>, 1996.

SPENCER, A. *Morphological theory*. Oxford: Blackwell Publishers, 1991.

TAGNIN, S. & FROMM, G. CoMAprend – a experiência da construção de um corpus de aprendiz para estudos. *Domínios de Linguagem*, p. 01- 17, 2009.

TONO, Y. Learner corpora: design, development and applications. In: *Proceedings of the 2003 corpus linguistics*, p. 800-809, 2003.

TOWELL, R & HAWKINS, R. *Approaches to Second Language Acquisition*. Great Britain: Longdunn Press, Bristol, 1994

TROSBORG, A. *Interlanguage Pragmatics – requests, complaints and apologies*. Berlin: Walter de Gruyter & Co., 1994.

VAINIKKA, A & YOUNG SCHOLTEN, M. Gradual development of L2 phrase structure. *Second Language Research*, v. 12, p. 140-76, 1993.

VAN PATTEN, B & WILLIAMS, J. *Theories in second language acquisition – an introduction*. New York: Routledge, 2007.

WAGNER, J.. *Grammatical morpheme acquisition in english as a foreign language*. Master's thesis. Oakland University, 2007. Disponível em: <www.ielanguages.com/documents/papers/MastersSeminar.pdf>.

WHITE, L. Second language acquisition and universal grammar. *Studies in second language acquisition*. N. 12, p. 121-33, 1990.

_____. *Second language acquisition and universal grammar*. Cambridge: Cambridge University Press, 2003.

_____. On the nature of interlanguage representation: universal grammar in the second language. In: DOUGHTY, C & LONG, M. *The handbook of second language acquisition*. Oxford: United Kingdom: Blackwell Publishing, 2005.

_____. Linguistic theory, universal grammar and second language acquisition. In: VAN PATTEN, B. & WILLIAMS, J. *Theories in second language acquisition – an introduction*. New York: Routledge, 2007.

ZOBL, H. Converging evidence for the “acquisition-learning” distinction. *Applied Linguistics*, 16, p. 35-56, 1995.

_____; LICERAS, J. Functional categories and acquisition orders. *Language learning*, v. 44, n. 1, 1994.

ANEXOS

English Placement Test – Oxford English Learning Center

NÃO ESCREVA NESTA FOLHA. Use a folha de respostas do final.
***Responda as perguntas abaixo selecionando a melhor alternativa.**

1. **How many peoplein your family?**
I. ARE THEY II. IS IT III. ARE THERE IV. IS
2. **What time is it?**
I. Ten and a quarter II. Ten minus the quarter III. A quarter past two IV. Fifteen and a quarter
3. **I get up at 8 o'clock_____ morning.**
I. in the II. in III. the IV. at the
4. **How much_____ where you live?**
I. do houses cost II. does houses cost III. does cost houses IV. do cost houses
5. **Where are you going_____ Friday?**
I. at II. in III. on IV. the
6. _____ **come to my party next Saturday?**
I. do you can II. can you to III. can you IV. do you
7. **What _____ in London last weekend?**
I. you were doing II. did you do III. you did IV. did you
8. **Is your English improving?**
I. I hope it II. hoping III. I hope so IV. I hope
9. **I'm going to Sainsbury's_____ some food.**
I. buy II. for buy III. to buy IV. for to buy
10. **Oxford is the most attractive city_____.**
I. I've ever seen II. that I see III. I've never seen IV. that I ever saw
11. **Oxford isn't _____ Bath.**
I. as beautiful as II. so beautiful than III. so beautiful that that IV. as beautiful than
12. **He was mowing the lawn when I _____ him yesterday.**
I. saw II. had seen III. was seeing IV. have seen
13. **Last Tuesday I _____ to the Passaport Office.**
I. must gone II. must go III. had to go IV. had go
14. **What were you doing at 7:30 on Wednesday evening? I _____ TV.**
I. was watching II. watched III. was watched IV. watching
15. **What time _____ to bed during the week?**
I. do you go II. are you go III. do you going IV. you are going
16. **Do you like Oxford? Yes,_____.**
I. I like II. So do I III. I does IV. I do
17. **I'm afraid I haven't got _____.**
I. any scissors II. scissor III. some scissors IV. a scissor
18. **This book is mine and that book is _____.**
I. yours II. your III. your's IV. you're
19. **Would you mind_____ me that pencil?**
I. to pass to II. pass III. passing IV. that you should pass
20. **I live in Oxford now. I _____ to France for a long time.**
I. don't been II. didn't come III. haven't been IV. don't come
21. **I don't understand. What language _____?**
I. speak you II. you speak III. you are speaking IV. are you speaking
22. **She came to Britain _____.**
I. four days ago II. at four days III. before four days IV. since four days
23. **My mother never _____ out in the evenings.**
I. go II. Goes III. is going IV. going
24. _____ **Oxford?**
I. Since when you live in II. How much time you are living in III. How long have you been living in
IV. How long time are you living in
25. _____ **car is the red Ford?**
I. Whose II. To whom III. Who's IV. Of who

26. I'm sorry. I haven't done my report _____.
I. up to the now II already III. until the present IV. yet
27. My friend doesn't speak Chinese. I don't _____.
I. also II. Neither III either IV. too
28. That's the house _____.
I. in which Mr Brown lives II. in which Mr Brown lives in that III. Mr Brown lives in IV. Mr Brown lives in that
29. If _____.
I. you come to my office, I'd pay you II. you shall come to my office, I'll pay you III. you came to my office, I would to pay you IV. you come to my office, I'll pay you
30. She asked me how big _____.
I. is your house II. my house was III. was my house IV. is my house
31. My friend let _____ his bike yesterday.
I. to borrow II. me borrowing III. me to borrow IV. me borrow
32. _____, what would you spend it on?
I. when you had a lot of money II. If you had a lot of money III. If you would have a lot of money IV. If you shall have a lot of money
33. I _____ smoking last year, but I didn't.
I. ought to give up II. ought to have given up III. ought given up IV. oughted to give up
34. I'm _____ the film on Wednesday.
I. look forward to see II. looking forward to seeing III. look forward seeing IV. looking forward seeing
35. I'm not _____ grammar.
I. interested to learn II. interested in learning III. interested to learning IV. Interest in learning
36. The film was very good. It's _____.
I. worth seeing II. worth to see III. worthwhile to see IV. worthwhile see
37. I have difficulty _____ English.
I. to write II. writing III. about writing IV. to writing
38. When I lived in France, I _____ a lot of wine.
I. was use to drinking II. was used to drink III. used to drink IV. used to drinking
39. I wish _____ Russian.
I. I could speak II. I would speak III. I can speak IV. I'll be able to speak
40. What will you do when _____ studying?
I. you're finishing II. You'll have finished III. you've finished IV. you're going to finish
41. The Chancellor _____ the new wing yesterday, but it still isn't finished.
I. had to open II. has to have opened III. was to have opened IV. had to have opened
42. I'd rather _____ English than Swedish.
I. you should learn II. you learnt III. that you might learn IV. you learn
43. No sooner _____ in through the door than the phone rang.
I. I had walked II. was I walking III. had I walked IV. I was walking
44. We're having the party at _____.
I. the house of Deborah II. the Deborah's house III. Deborah's IV. house of Deborah
45. If he hadn't known the boss, he _____ the job.
I. wouldn't get II. hadn't got III. wouldn't have got IV. wouldn't had got
46. I'd sooner _____ a car than a motorbike.
I. him to buy II. that he buy III. he bought IV. he should buy
47. I need to go to _____ toilet.
I. the II. a III. ---- IV. some
48. It's time _____ some work.
I. for to do II. she would do III. she did IV. she were to do
49. It's now 9 o'clock and the train _____ arrive at 8:15.
I. had to II. must III. was due to IV. is going to
50. We regret _____ that the course has been cancelled.
I. to tell II. telling III. to have said IV. to say

ANEXO 2: Answer Sheet

ANSWER SHEET

Data: _____ Nível: _____ Código: _____

Nome: _____

-
1. I. () II. () III. () IV. ()
 2. I. () II. () III. () IV. ()
 3. I. () II. () III. () IV. ()
 4. I. () II. () III. () IV. ()
 5. I. () II. () III. () IV. ()
 6. I. () II. () III. () IV. ()
 7. I. () II. () III. () IV. ()
 8. I. () II. () III. () IV. ()
 9. I. () II. () III. () IV. ()
 10. I. () II. () III. () IV. ()
 11. I. () II. () III. () IV. ()
 12. I. () II. () III. () IV. ()
 13. I. () II. () III. () IV. ()
 14. I. () II. () III. () IV. ()
 15. I. () II. () III. () IV. ()
 16. I. () II. () III. () IV. ()
 17. I. () II. () III. () IV. ()
 18. I. () II. () III. () IV. ()
 19. I. () II. () III. () IV. ()
 20. I. () II. () III. () IV. ()
 21. I. () II. () III. () IV. ()
 22. I. () II. () III. () IV. ()
 23. I. () II. () III. () IV. ()
 24. I. () II. () III. () IV. ()
 25. I. () II. () III. () IV. ()
 26. I. () II. () III. () IV. ()
 27. I. () II. () III. () IV. ()
 28. I. () II. () III. () IV. ()
 29. I. () II. () III. () IV. ()
 30. I. () II. () III. () IV. ()
 31. I. () II. () III. () IV. ()
 32. I. () II. () III. () IV. ()
 33. I. () II. () III. () IV. ()
 34. I. () II. () III. () IV. ()
 35. I. () II. () III. () IV. ()
 36. I. () II. () III. () IV. ()
 37. I. () II. () III. () IV. ()
 38. I. () II. () III. () IV. ()
 39. I. () II. () III. () IV. ()
 40. I. () II. () III. () IV. ()
 41. I. () II. () III. () IV. ()
 42. I. () II. () III. () IV. ()
 43. I. () II. () III. () IV. ()
 44. I. () II. () III. () IV. ()
 45. I. () II. () III. () IV. ()
 46. I. () II. () III. () IV. ()
 47. I. () II. () III. () IV. ()
 48. I. () II. () III. () IV. ()
 49. I. () II. () III. () IV. ()
 50. I. () II. () III. () IV. ()

ANEXO 3: Instrução – Professores

Professor: _____ Turma: _____ Alunos: _____

Caro colega!

Muito obrigada pela colaboração!

O título da pesquisa é “Erros morfossintáticos em inglês como L2: uma análise baseada em *Learnner Corpus*” e o objetivo é construir um corpus de aprendiz a partir da nossa atuação profissional.

A coleta será feita em três partes, que, por razões metodológicas, devem ser aplicadas na sequência de três semanas.

A primeira parte, na sua turma agendada para _____, é formada por (i) uma apresentação da pesquisa, (ii) por um breve questionário sobre informações gerais, (iii) por um termo de autorização formal dos alunos para o uso das informações e (iv) por um *Placement Test* com 50 questões. Esta parte poderá ser feita em sala de aula, e pediria que, se possível, fosse aplicada na metade da aula (nem no início, nem no final) para evitar variáveis que possam interferir no desempenho dos alunos.

O material a ser utilizado nesta parte consta de:

1. Uma folha com os itens (i), (ii) e (iii);
2. Uma cópia do *Placement Test*, de 4 folhas, que NÃO DEVERÁ SER PREENCHIDA pelos alunos, pois, por razões de economia em geral, SERÁ UTILIZADA NOVAMENTE em outros horários.
3. Uma *Answer Sheet*, onde os alunos deverão marcar as respostas.

As cópias-matrizes para uso dos alunos serão entregues a você antes do início da aula, já que estão sendo reutilizadas.

Peço que o material seja entregue imediatamente após a aplicação (ou a mim, ou depositado em um dos meus escaninhos no 4° ou no 5° andar).

A segunda parte, agendada para _____, deverá ser aplicada no Laboratório de _____, que já deverá estar previamente reservado por mim para a atividade. Esta parte é a de coleta de dados propriamente dita. A tarefa é a redação de dois parágrafos, de aproximadamente 150 palavras cada, sobre informações pessoais em geral, em primeira e terceira pessoa, respectivamente. Solicito que as instruções para os alunos sejam lidas juntamente com os mesmos a fim de evitarmos alguns problemas de interpretação, principalmente o de que os textos deverão ser redigidos NO ESPAÇO DO E-MAIL E NÃO EM WORD.

O material a ser utilizado nesta parte consta somente de uma folha de instruções para os alunos, já que os textos deverão ser digitalizados. Esta folha deverá se recolhida pelo professor, pois também será utilizada novamente em outras aplicações.

A terceira e última parte, agendada para _____, é semelhante a anterior. Os alunos serão solicitados a redigirem um parágrafo de aproximadamente 200 palavras contando sobre uma viagem. Da mesma forma como na segunda parte da coleta, o material a ser utilizado consta somente de uma folha de instruções para os alunos. Esta folha também deverá se recolhida pelo professor, pois será utilizada novamente em outras aplicações.

Alguns problemas previstos são:

*Falha técnica no computador/ servidor/ provedor: pavor pedir para os alunos imprimirem seus textos;

*falha na impressora: pedir para os alunos escreverem manualmente seus textos.

Peço que quaisquer problemas ou imprevistos sejam registrados, a fim de que possam ser formalmente considerados em outro momento da pesquisa.

As ausências alternadas serão posteriormente analisadas. Ou seja, todos os alunos em aula deverão fazer a tarefa, mesmo que não tenham participado de alguma das partes.

Mais uma vez, muito obrigada pela disponibilidade.

Aline

ANEXO 4: Parte I: Apresentação da Pesquisa

Apresentação:

O seguinte trabalho tem como objetivo coletar dados para a pesquisa “Erros morfossintáticos em inglês como L2: uma análise baseada em *learner corpus*”, desenvolvida pela prof. Aline Pacheco, aluna do curso de doutorado em Estudos da Linguagem da Universidade Federal do Rio Grande do Sul, sob a orientação do Prof. Sergio de Moura Menuzzi.

Todas as informações prestadas serão usadas para fins de pesquisa. Para tanto, asseguramos o caráter sigiloso e ético das mesmas.

O trabalho é constituído de três etapas e cada etapa terá duração de 30 min. Desde já, agradecemos a colaboração.

Informações pessoais:

Nome: _____

Data de Nascimento: ____/____/____ Local: _____

Sexo: _____

Ocupação: _____

Ano de início de exposição à Língua Inglesa: _____

Onde: (___) escola regular (___) curso de línguas (___) atual local de ensino

Fala alguma outra língua estrangeira? _____ Qual? _____

Há quanto tempo? _____

Costuma estudar inglês em tempo extraclasse? _____

TERMO DE CONSENTIMENTO

Autorizo o uso das informações contidas neste instrumento de coleta de dados para os fins de pesquisa em lingüística aplicada.

ANEXO 5: Parte II: Instruções Aluno

PARTE II

Instruções:

1. Você será solicitado a escrever dois parágrafos curtos, que deverão ser enviados para os seguintes endereços eletrônicos:

aline.pacheco@puhrs.br

aline.pacheco@terra.com.br

2. Solicitamos que, acima dos textos, você escreva os seguintes dados de identificação:

Nome: _____ Nível: _____

3. Os textos NÃO deverão ser escritos em Word, mas no próprio espaço para textos do e-mail.

4. Para os fins da pesquisa, estes textos não serão alterados. No entanto, para seu interesse, eles serão corrigidos e reenviados a você. Se não tiver interesse em recebê-los de volta, manifeste-se junto à identificação.

5. O tempo para a tarefa será de 30 minutos.

Tarefa:

1. Escreva um parágrafo sobre você: informações pessoais, atividades diárias, gostos, descrições em geral, família/amigos, lazer, experiências interessantes, etc.

(100-150 palavras)

2. Escreva um parágrafo sobre alguém que você conhece bastante: informações pessoais, atividades diárias, gostos, descrições em geral, família/amigos, lazer, experiências interessantes, etc.

(100-150 palavras)

Parte III

Instruções:

6. Os textos que você será solicitado a escrever deverão ser enviados para os seguintes endereços eletrônicos:

aine.pacheco@pucrs.br

aline.pacheco@terra.com.br

7. Solicitamos que, acima dos textos, você escreva os seguintes dados de identificação:

Nome: _____

Nível: _____

8. Os textos NÃO deverão ser escritos em Word, mas sim, no próprio espaço para textos do e-mail.

9. Pedimos a gentileza de não fazer qualquer tipo de consulta: ao material, a qualquer dicionário ou ao professor.

10. O tempo para a tarefa será de 20 a 30 minutos.

Tarefa:

1. Escreva um parágrafo sobre uma viagem que você fez. Fale sobre o lugar que você foi, com quem, quando, como você foi, coisas que você fez e que não fez, coisas que gostou e que não gostou, algo ou algum episódio interessante sobre a viagem ou sobre o lugar (mínimo de 200 palavras).

ANEXO 7: Tabela modelo de relação dos sujeitos participantes

NIVEL:		PROFESSOR:		GRUPO:		COL:2008/0X	
SUJEITO	NOME XXXXXXXXXXXXX	PLACEM	T1	T2	T3		
01							
02							
03							
04							
05							
06							
07							
08							
09							
10							
11							
12							
13							
14							
15							
16							
17							
18							
19							
20							
21							
22							
23							
24							
25							
26							

ANEXO 8: Relação de turmas

RELAÇÃO DE TURMAS

NIVEL	PROFESSOR XXXX	HORARIO	GRUPO	PERIODO X
1	Ana Welp	4ABCD	1	2008/01
1	Aline	2ABCD	2	2008/02
1	Simone	2JK	3	2008/02
1 (?)	Aline	3JK	4	2008/01
2	Aline	4ABCD	1	2008/01
2	Ana Boc	SAT	2	2008/01
2	Aline	3FG	3	2008/02
3	Simone	3FG	1	2008/01
3	Vera	4ABCD	2	2008/01
3	Silvana	3EE	3	2008/01
3	Simone	3FG	4	2008/02
4	Simone	3JK	1	2008/01
4	Aline	SAT	2	2008/01
4	Aline	FRI	3	2008/01
4	Aline	3EE	4	2008/02
5	Vera	FRI	1	2008/01
5	Silvana	3EE	2	2008/02
5	Aline	3JK	3	2008/02
6	Ana Welp	SEX	1	2008/01
6	Vera	3EE	2	2008/01
6	Carlos	2JK	3	2008/02
7	Simone	FRI	1	2008/01
7	Clarice	3EE3JK	2	2008/01
7	Vera	3EEFRI	3	2008/02
8	Clarice	2JK	1	2008/01
8	Adriana	2JK	2	2008/02

ANEXO 9: Exemplos do BELC etiquetado

N1G1S04PBT2

He name is (UC-S) Gutter. He (UC-A) 23 years (PL-S). He works (TPR-S) in the (PREP-A) Ceee. He likes (TPR-S) partys (PL-S), music, computers (PL-S) and travel. He's family is (UC-S) big. He is (UC-S) single and (A-A) student in Puc. He live (TPR-A) in (PREP-S) Porto Alegre.

N1G1S02PBT2

Her name's (CC-S) Nara. She works (TPR-S) in (PREP-S) a (A-A) office. She's (CC-S) 45 years (PL-S) old. She have (TPI-A) one son, I. She have (TPI-A) much friends (PL-S). She's (CC-S) my mather. She's (CC-S) my friend and my mather.

N1G1S08PBT2

Her First Name is (UC-S) João and her last name is (UC-S) Jordani, h e is (UC-S) my father. He lives (TPR-S) with me in (PREP-S) one apartment in (PREP-S) Porto Alegre. He wakes (TPR-S) early today and makes (TPR-S) my breakfast. He starts (TPR-S) to works at 8h A.M. and come (TPR-A) back to home in lunch hour, to eats with my and my brothers (PL-S). After it, he backs to your job, he is (UC-S) a (A-S) manager and works (TPR-S) all the (A-S) afternoon, goes out of job at 6h P.M. When, he buys (TPR-S) the (A-S) broad and go (TPR- A)home. He watchs (TPR-S) TV and (TPI-A) dinner with our family, and at 11h P.M. he goes (TPR- S) to bed. In (PREP- A) Friday Nights (PL-S) he plays (TPR-S) soccer with you friend and her hobby is (UC-S) kitchens (PL-S). Your favourite food is (UC-S) barbecue and your favourite drink is (UC-S) Beer.

N1G1S12PBT2

I have a (A-S) girlfriend, she's (CC-S) (A-A) odontologist. She (UC-A) has 24. She likes (TPR-S) to dancing, eat and travel. Her name is (UC-S) Cristina. She has (TPI-S) one brother and her parents (PL-S) don't live together. She is (UC-S) honest and ambitious. She works (TPR-S) in (PREP-S) (A-A) odontologist clinic in (PREP-S) the São Leopoldo and the someclinic in (PREP-S) the Porto Alegre. She's (CC-S) great. Heheeh. she likes (TPR-S) to give roses (PL-S) and see tv.

N1G1S11PBT2

My Father is (UC-S) João Miranda, he is (UC-S) 44. He teatches (TPR-S) in PUC at ningth. He is (UC-S) married with my mother Susana sience 1984. He haves (TPI-A) tree doughthers, Mariana, Marina e Leticia. On (PREP-S) teh weekens he likes (TPR-S) watches soccer and does borbucue.

N1G1S05PPT2

Jenyffer is (UC-S) my my sister. She's (CC-S) 26 years (PL-S) old. She lives (TPR-S) in (PREP-S) Santa Catarina. She is (UC-S) a (A-S) student. She lives (TPR-S) with your husband Pedro, and they have two sons (PL-S), Mauricio and bruno. She work's with my cousin José in (PREP-S) the hospital. She like's drink soda and eat cake. She's (CC-S) very fat. She doesnt like walk in (PREP-S) the (A-S) parks (PL-S) with me. She loves (TPR-S) watch TV. Her favorite film is (UC-S) Lagoa Azul. In the (A-S) holidays (PL-S) she go to the (A-S) beach with your friends (PL-S). And in the (A-S) weekends (PL-S) eat (TPR-A) barbecue with your family. She call (TPR-A) me every nights (PL-OU). Now, she's (CC-S) pregnancy and very happy.

N3G4S13PPT3

I went (PI-S) to europe in (PREP-S) july, 2006. I stayed (PR-S) in (PREP-S) europe for 6 months (PL-S) and I lived (PR-S) in (PREP-S) London and Florence. I studied english and italian for 4 months (PL-S) and 2 months (PL-S) I did a (A-S) tour in (PREP-S) europe. I didn't work, only studied and did a (A-S) tour. When I stayed (PR-S) in (PREP-S) london I lived (PR-S) with Porto Alegre's friends (PL-S) and in (PREP-S) florence I lived (PR-S) only. In (PREP-S) florence a did more friends (PL-S) I think because a stayed (PR-S) only. I liked (PR-S) this experience, because a lived (PR-S) only and I needed (PR-S) do clean of my house and buy my food for exemple. When I stayed (PR-S) in (PREP-S) florence I did a (A-S) course of cooking and I liked (PR-S), I prepared (PR-S) very italian specialits.

N3G4S06PBT3

In the last summer, I went (PI-S) to pinhal beach with my parents (PL-S), grandparents (PL-S), sister, brother-in-law and my boyfriend. The weather was (UC-S) (PI-S) bad, because are raining (PP-S). In the (A-S) good days (PL-S), we went (PI-S) to the (A-S) beach, and played (PR-S) voleyboll. In the (A-

S) bad days (PL-S) I stayed (PR-S) in home, we played (PR-S) games (PL-S). We go out all the (A-S) nights (PL-S), we went (PI-S) for the centre city. In (PREP-S) _the centre we ate (PI-S) ice-cream. We went (PI-S) in another citys (PL-S), Tramandaí and Imbé. In (PREP-S) Imbé, we went (PI-S) visited my aunt and bought (PI-S) things (PL-S). We went (PI-S) visited relatives (PL-S) in (PREP-S) cidreira, my aunt and my cousin. I didn't like of weather, but I love (PR-A) stay with my family and because of this, the vacation was (PI-S) very nice.

N3G4S04PPT3

Two years (PL-S) ago, me and my classmates (PL-S) of school went (PI-S) to Ilha do Mel-PR in (PREP-S) a (A-S) end of year travel. There, we went (PI-S) to the beach and some partys. The beach is (UC-S) so wonderful: blue water, white sand, green vegetation and some rocks (PL-S) make the (A-S) beach the (A-S) best i ever seen. The (A-S) partys (PL-S) is (UC-A) so animated, and you don't pay nothing to enter in the (A-S) party, you only pay the (A-S) drinks (PL-S) that you buy. The (A-S) only way to go to the (A-S) island is (UC-S) by boat: you pay the (A-S) pass and you cross the (A-S) river, and in (PREP-S) 30 minutes (PL-S) you're (CC-S) in (PREP-S) the (A-S) island. We stay (PR-A) there for a (A-S) week, and every days (PL-OU) we went (PI-S) to the (A-S) beach and every night we went (PI-S) to the partys. Sometimes, we went (PI-S) to dinner out, and we eat (PI-A) very good food, like camarao with potatos (PL-S) and rice, it is (UC-S) very good!! We come (PI-A) to Paraná by bus, and cross (PR-A) the river by boat. The travel by bus take (PI-A) us 10 hours (PL-S), and we sleep (PI-A) all the (A-S) time!!

N4G1S12PBT3

In summer past, I to travel (PR-A) to Tramandaí. I was (UC-S) (PI-S) com my boyfrendy. I sllep, eat and dance muito. I stay (PR-a) in (PREP-S) (A-A) hotel 5 stars (PL-S). I suwin in (PREP-S) Tramandaí.

I to travel (PR-A) to Tramandaí by car.

N4G1S05PPT3

One year ago, in the (A-S) same time, I, my girlfriend, her mother and her two little brothers (PL-S) went (PI-S) to their farm by bus. The (A-S) trip lenghted 6 hours (PL-S). Her father's gone one day before. The (A-S) city knowns by São Vicente do Sul. The (A-S) farm is (UC-S) very big and beautiful. There are (UC-S) kilometers of rice cultivation and many wild animals (PL-S). We did (PI-S) a lot of things (PL-S), play (PR-A) soccer, barbecue, swim (PI-A) and fish (PI-A) in the (A-S) lake, have (PI-A) fun with dogs (PL-S)... In spite of that we did (PI-S) many things (PL-S), we relax (PR-A) and our stress gone (PI-A) away because the (A-S) life there, are (UC-A) very calm. We stayed (PR-S) there for two weeks (PL-S) and after we went (PI-S) to Rivera (Uruguay) for buying a (A-S) copple of things (PL-S) because are (UC-S) cheap.

We returned (PR-S) to Porto Alegre one day after by car with her father. The (A-S) journey finished (PR-S) when I was (PI-S) at home. I hope I will return there very soon.

N4G1S09PPT3

My last travel was (PI-S) very cool! I went (PI-S) to São Luis in (PREP-S) Maranhão and visited (PR-S) my relatives (PL-S) and my sweet grandmother. I went (PI-S) to the (A-S) beach and the (A-S) sand there is (UC-S) more clear and thin that here. I went (PI-S) alone to there by air in (PREP-S) seven hours (PL-S) to travel in the year 2006. There I went (PI-S) out to my old friends (PL-S) and cousins (PL-S) every weekends (PL-OU) and some days (PL-S), I traveled (PR-S) with my ancles for all state and Pará state too. I don't liked poverty in (PREP-S) the (A-S) interior of state, peoples (PL-OU) in (PREP-S) precarius houses (PL-S) and much childrens (PL-OU) working (PP-S) for a (A-S) few money. In (PREP-S) Pará I insided the florest and knowed (PR-OU) big trees (PL-S) and knowed Tocantins, Tapajós, Xingu and Amazônia rivers (PL-S), and I saw (PI-S) very big fishes in (PREP-S) the (A-S) boats of fishing. I liked (PR-S) very much this vacation!

N4G1S06PBT3

The last travel (trip?) that I made (PI-S) it was (UC-S) (PI-S) in (PREP-S) the (A-S) last month, I gone (PI-A) for my hometown Cachoeira do Sul. I go (PI-A) for Cachoeira, especially to visit my mother, because she lives (TPR-S) there. Generally, I travel alone, but sometimes I travel with my brhother Ricardo, that he lives (TPR-S) withme, here in (PREP-S) Poa. I like of Cachoeira, because there I have my friends (PL-S) of when I was (UC-S) (PI-S) (A-A) child, however I have, there too, my family grandmother, uncles (PL-S), cousins (PL-S), etc. Cachoeira do Sul is (UC-S) a (A-S) small city, with a (A-S) population of (more or less) 90.000 habitants (PL-S). Who live (TPR-A) there, they (?) can to

have a (A-S) good life, because it is (UC-S) very calm city. My next trip for Cachoeira it will be in (PREP-S) july, because I already will be of vacation of the classroom at Puc, so I will have more time for to stay in (PREP-S) the (A-S) mother's (P-S) house, and this is (UC-S) very good! (I think...).

N5G1S08PPT1

Hi my name is (UC-S) XXXXX and i study here in (PREP- S)puc ciencias aeronauticas I am (UC-S) 20 years (PL-S) I leave in (PREP-S) Porto Alegre, but I born in (PREP-S) Arroio dos Ratos. I love flying , play soccer I have a (A-S) motocicle I like to get out at night with my friends (PL-S) I live with my parents (PL-S) I have one sister.And my gretest dream one day is (UC-S) to provide a home for my parents (PL-S) and be an (A-S) iternational master.

N5G1S02PIT1

My name's (CC-S) XXXXX, but people call me Pati since I was born. I'm (CC-S) twenty ears old, and I live in (PREP-S) Porto Alegre with my parents (PL-S) and my older sister. I have two others sisters (PL-S), but now they're (CA-S) living (PP-S) in (PREP-S) Sao Paulo, together.I also live with a (A-S) little cat, wich I adopted (PR-S) from the (A-S) streets (PL-S) last year. I just love animals (PL-S) and, if I could, I'd rescue them all from the (A-S) streets (PL-S) right to my house. I study Pharmacy at PUCRS since 2005, and next year I'm (CA-S) gonna graduate. I really enjoy what I'm (CA-S) studyng, but I'm (CC-S) still not sure what I am (UA-S) going to do in (PREP-S) the (A-S) future. I have a (A-S) boyfriend called Lucas, we met (PI-S) about five years (PL-S) ago, and our relationship is doing (PP-S) well since then. We have already travelled (PR-S) to the USA together, and it was (UC-S) (PI-S) one of the (A-S) most exciting trip I've ever done...

N5G1S06PIT1

Hi! My name is (UC-S) XXXXX, i'm (CC-S) 23 years (PL-S) old. I live in (PREP-S) São Leopoldo with my mom and my brother. I'm (CC-S) in (PREP-S) the (A-S) second year of medicine school at PUCRS, and i like very much of my college. I went (PI-S)to Puc every day of the week by the tren and then a bus. Sometimes is (UC-S) very boring, because i got (PI-S) very tired and so didn't study very well as i wanted (PR-S) . I didn't come live in (PREP-S) Porto Alegre yet because it's (CC-S) very expensive. Although the fact of the distance of my house and my college, i like to live in (PREP-S) São Leopoldo, because it's (CC-S) there that lives (TPR-S) my friends (PL-S) and my family, except my dad, who lives (TPR-S) in (PREP-S) Santa Cruz. I have a (A-S) sister too, from my father's (P-S) part.

N5G1S03PPT1

My name is (UC-S) XXXXX, I'm (CC-S) 23. I live in (PREP-S) Porto Alegre, study at PUC and work as a (A-S) trainee in two places (PL-S), a (A-S) clinic and a (A-S) school. And I work publisher courses of Cognitive-behavior therapy. I am not have very much friends (PL-S), but I have a especial friends (PL-S). I think that I work to much, and this difcult that I have frre times to go out, or dance, or walk on a park, or drink somethink with other girls (PL-S). Well, I am (UC-S) in (PREP-S) a (A-S) moment so happy of my life. In te middle of this year, I am (UA-S) going to gratuated on Psychologic, and last week, I was (PI-S) invited to enter on a Clinic Psychologic that I always want. I'm (CC-S) so glad, and nervous! And, the other moment: I will get married on (PREP-A) december! I think that is (UC-S) the (A-S) most important choice of my life, because I want get married to stay forever on side this person. My family is (UC-S) so glad to, they are (UC-S) really envolved with this, help with the preparatives of the (A-S) party and my boyfriend bought (PI-S) our appartment, and he want (TPR-A)surprise-me on(PREP-S)the engaged day.

N5G2S16PBT1

Well, let me think... My name is (UC-S) XXXXX and I'm (CC-S) nineteen years (PL-S) old. I live with my parents (PL-S) in (PREP-S) a (A- A) apartament next shopping Iguatemi. I like to live here because is (UC-S) a (A-S) calm and beautiful place with many trees (PL-S) around there. My favorite kind of music is (UC-S) reggae and my favorite singer is (UC-S) Bob Marley of course! But I'm (CC-S) very interesting in (PREP-S) all kinds (PL-S) of music 'cause I love these so much! In (PREP- S) my spare time I like to stay in (PREP- OU)home on computer talking (PP-S) with my friends (PL-S) about many things (PL-S). I like to dance to! I'm (CC-S) in (PREP-S) (A-A) second year of university and odontology is (UC-S) my dream. I hope turn be a (A-S) good dentist to improving a quality oral health to people. I was born in (PREP-S) Minas Gerais and I live here since I was (UC-S) (PI-S) six. I have no idea what more write about me.

N6G2S08PPT2

So, about some friend I'll write about my ex-girlfriend. Waleska is (UC-S) 19 years (PL-S) old and have (TPI-A) a (A-S) sister called Ohana. She's (CC-S) a (A-S) nice person who likes (TPR-S) party so much. She hadn't already finished (PR-S) school because she seems (TPR-S) doesn't like to study very much, but she is (UA-S) trying (PP-S) to finish it because now she understood (PI-S) how it is (UC-S) important to her life. Actually we're (CC-S) a (A-S) little far of each other because she always goes out with her friends (PL-S) and I'm (CC-S) always busy with my work, the university or my band, but on (PREP-S) weekends (PL-S) I use to drop by her home to say hello and take a (A-S) funny conversation. Just it!

N6G1S01PIT2

My girlfriend is (UC-S) a (A-S) hard-working girl. Her name is (UC-S) Alessandra, and she has (UC-A) 24 years (PL-S) old. She used (PR-S) to wake up early, and is (UC-S) very responsible with her jobs (PL-S) and things (PL-S) to do. She is (UA-S) doing (PP-S) Electronic technician course in (PREP-S) the (A-S) mornings (PL-S), and after her classes (PL-S), she works (TPR-S) on a mall's shop. She is (UC-S) friendly, has (TPI-S) a (A-S) short brown hair and pretty green eyes (PL-S). She likes (TPR-S) to travel and watch movies (PL-S). But she doesn't like when I arrive later in our dates (PL-S).

N6G1S01PIT2

I'll talk about my father. He's (CC-S) a (A-S) short man, but have (TPI-A) a (A-S) huge heart. He's (CC-S) the (A-S) owner of his own business and go out to his job at 8:30 a.m pretty much every day. At work he's stress his self a lot, once he's (CC-S) a (A-S) stressed person. He have (TPI-A) many friends (PL-S) because his business deal (TPR-A) with college students (PL-S) and the (A-S) fact that he's (CC-S) a (A-S) very gentle man helps (TPR-S) on it. On (PREP-S) the (A-S) weekend he likes (TPR-S) to stay at home resting (PP-S) and listening (PP-S) to his blues and rock CDs and DVDs. Once he traveled (PR-S) to Argentina with my mother and went (PI-S) to a (A-S) Tango concert, which he described as one of the biggest concert he have (TPI-A) ever seen.

N6G1S05PIT2

My sister that lives (TPR-S) in (PREP-S) Porto Alegre with me is (UC-S) a (A-S) dentist. She is (UC-S) 25 years (PL-S) old and she is (UC-S) a (A-S) dentist. She came (PI-S) to Porto Alegre before, in (PREP-S) 2001, actually. She used (PR-S) to live in (PREP-S) a (A-S) small apartment and alone. After I came (PI-S), she moved (PR-S) to another apartment to live with me. She works (TPR-S) 8 hours (PL-S) for a (A-S) day and we see each other only at night. She likes (TPR-S) her job and to relax she goes (TPR-S) to the (A-S) movies and goes (TPR-S) travel. She loves (TPR-S) it. Next year she will get married and I will live alone.

N6G1S06PIT2

My sister is (UC-S) thirty years (PL-S) old, and she doesn't live with me and our parents (PL-S) anymore. My sister just gave (PI-S) birth to a (A-S) child, and she, her child and her husband visit my parents (PL-S), my brother and I very often, usually on (PREP-S) Sunday, and her family spend (TPR-A) the (A-S) whole day in (PREP-S) my family's (P-S) house, we lunch together, and in (PREP-S) the (A-S) afternoon we take a (A-S) walk or we just stay at home and chat. She's (CA-S) working (PP-S) at a (A-S) hospital for five years (PL-S) now, and she love (TPR-A) her job. She has (TPI-S) many friends (PL-S), that she used (PR-S) to go out a lot with, in (PREP-S) the (A-S) past, but now that she is (UC-S) a (A-S) mother, she spends (TPR-S) most of her time at home taking (PP-S) care of her baby. She is (UC-S) very happy with her life right now, you can tell just looking (PP-S) at her eyes (PL-S).

N6G1S03PPT2

Well, I will talk about my father. He is (UC-S) 50 years (PL-S) old. He works (TPR-S) in (PREP-S) a (A-S) factory. He is (UC-S) married, and he has (TPI-S) 2 daughters (PL-S), I and my sister Rafaela. He is (UC-S) a (A-S) lovely person, he doesn't like read books (PL-S), just see the journal's TV. He loves (TPR-S) eat cheese and drink wine. Normally he call (TPR-A) me during the (A-S) week, and he always ask (TPR-A) me when I will come back to Lajeado. In (PREP-S) our house we have a (A-S) funny dog, my father always say (TPR-A) that he doesn't like dogs (PL-S), but he always buy (PI-A) some kids for Fredy, our dog. Well, I think my father passed by different things (PL-S) in (PREP-S) your life, because he borned in (PREP-S) Marcorama, and after, when he was (PI-S) 7 years (PL-S) old he went (PI-S) study in (PREP-S) Lajeado, where he know (PI-A) my mom. He started (PR-S) work very early, he has (TPI-S) 6 brothers (PL-S) and 3 sisters (PL-S) that lives in (PREP-S) Garibaldi.

So I think he is (UC-S) a (A-S) great people that has (TPI-S) many qualities (PL-S) and good thinks (PL-S) too show me, I think I still have many thinks (PL-S) to learn with he.

N7G2S05PPT3

I have been to Gramado 3 times (PL-S), but the (A-S) last was (PI-S) the (A-S) best. It was (PI-S) 2006, December. My grandfather died (PR-S) that month and all my family was (PI-S) very sad, and we were (UC-S) (PI-S) in Christmas's Time. So, my father decided (PR-S) that if we stayed (PR-S) at home it will (PI-S) be worse for all, then he talked (PR-S) to me, my mom and my sister that was (UC-S) (PI-S) time to stop cry and try have a (A-S) good Christmas. We went in December, 22 I think. There was (UC-S) (PI-S) so beautiful, the weather was (UC-S) (PI-S) great, not so hot like in (PREP-S) Porto Alegre. The (A-S) people there are (UC-S) more polite, they talk with you like they known you. We went to a (A-S) house, the (A-S) same that we stayed (PR-S) the (A-S) 2 times (PL-S) before. The city is (UC-S) clear. You don't see a (A-S) paper in the (A-S) ground, or the (A-S) people put the (A-S) things (PL-S) out the (A-S) garbage. We walked (PR-S) on (PREP-S) the (A-S) street at 11 o'clock at night, if you do this here, probably something will happen to you. On (PREP-S) the Christmas's night we went (PI-S) to Cafe Colonial. That was (UC-S) (PI-S) really amazing. When I saw some many kinds (PL-S) of food I didn't know what to do. I wanted (PR-S) eat everything, like children when saw candies (PL-S). Very funny. After this, we went to the (A-S) church. I don't use to go anymore, but that time was (PI-S) good. I think something really good, something the give peace. The (A-S) day after was (PI-S) great too, we went to the (A-S) park, walked (PR-S) on the (A-S) city, buy (PI-A) some chocolates, eat (PI-A) some of then. And we came (PI-S) back home, to the (A-S) normal things (PL-S). I think that year we can have a (A-S) Christmas.

N7G2S01PBT3

The paragraph I will write, will be about my trip from Portugal to Spain. It was (PI-S) a 2 years (PL-S) ago, and by that time I was (PI-S) working to my company cliente in (PREP-S) Europa. The (A-S) trip started (PR-S) with a (A-S) map, a (A-S) rented car, five work collegues, some of them portuguese guys (PL-S), and some ideas (PL-S) about the (A-S) path. The map was (PI-S) useless, 'coz it was (PI-S) the (A-S) Italy map. The (A-S) guy who brought, one of those portuguese guys (PL-S), doesn't look inside, only readed (PR-OU) the (A-S) word 'MAP something...' in the (A-S) cover and grab (PR-A) it. We started (PR-S) the (A-S) jorney in (PREP-S) Portaugal and the (A-S) inicial path, we all knowed well and the (A-S) map was (PI-S) not consulted, but as we arrived (PR-S) the (A-S) first crossroad, the (A-S) map have said to get the (A-S) "valpolicella" road. The (A-S) strange thing in (PREP-S) that situation was (PI-S) the (A-S) words the (A-S) guy spoke (PI-S) to defend him self: "Unless, it was (PI-S) usefull to show us that we don't have any idea about how to go Spain". So, finally, on (PREP-S) my trip to Spain, I knowed (PI-A) Italy.

N7G2S03PPT3

In (PREP-S) December I went (PI-S) to pass the (A-S) reveillon in Torres with a (A-S) couple of friends (PL-S), my boyfriend and my best friend. We were (UC-S) (PI-S) right there 2am and eat (PI-A) a (A-S) hot dog. After download our bags (PL-S) in (PREP-S) the (A-S) apartment of my grandmother and we were (UC-S) (PI-S) given a (A-S) walk on (PREP-S) the (A-S) beach. We all know the (A-S) holiday on the (A-S) beach and fun. We went (PI-S) to several beautiful hills (PL-S), and we know the (A-S) lighthouse where people go to fly and parachute. Our holiday was (UC-S) (PI-S) very great advantage because the beach, the (A-S) days (PL-S) were (UC-S) (PI-S) beautiful and we had (PI-S) many options (PL-S) to exit. We take (PI-A) pictures (PL-S) of beautiful beaches and places (PL-S) we visit (PR-A). We have done and learned (PR-S) to cook many things (PL-S). We walk (PR-A) too, know the (A-S) pond in the Guitar that put the sun is (UC-S) even more beautiful to see. At the (A-S) turn of the (A-S) year witnessed (PR-S) the (A-S) show of Dudu Nobre the side of the (A-S) beach, take (PI-A) bath of (PREP-A) (A-A) sea and drink (PI-A) many champagnes. The (A-S) other day buy (PI-A) many gifts (PL-S) tidy the (A-S) house and returned (PR-S) to Porto Alegre, because on (PREP-S) day 2 we come back to work. It took us almost 4 hours (PL-S) on the trip back because the (A-S) road was (PI-S) too full. The trip was (PI-S) great and we had a lot of fun.