

Fernando Gonçalves Ferreira-Junior

**Em Construção!**  
**Uma investigação acerca da natureza dos processos cognitivos  
envolvidos na aprendizagem de uma segunda língua**

Tese apresentada ao Programa de Pós-Graduação  
em Estudos Lingüísticos da Universidade Federal  
de Minas Gerais, como requisito parcial para a  
obtenção do título de Doutor em Lingüística.

Área de concentração: Lingüística  
Linha de pesquisa: Estudo da Inter-relação entre  
Linguagem, Cognição e Cultura  
Orientador: Prof. Dr. Fábio Alves

**Belo Horizonte**  
**Faculdade de Letras da UFMG**  
**2008**

### Ficha Catalográfica


XXX Ferreira-Junior, Fernando Gonçalves

Em Construção! Uma investigação acerca da natureza dos processos cognitivos envolvidos na aprendizagem de uma segunda língua / Fernando Gonçalves Ferreira-Junior. - Belo Horizonte: Universidade Federal de Minas Gerais, 2008.

250p.

1. Aprendizagem de Línguas. 2. Chunking. 3. Conexionismo.  
4. Construções. 5. Emergentismo. 6. Neurobiologia da Aprendizagem.  
I. Título.

Tese defendida por FERNANDO GONÇALVES FERREIRA JÚNIOR em 09/07/2008 e aprovada pela Banca Examinadora constituída pelos Profs. Drs. relacionados a seguir:



---

**Fábio Alves da Silva Júnior - UFMG**  
**Orientador**



---

**Márcia Cristina Zimmer - UCPel**



---

**Paula Lenz Lima - UECE**



---

**Heliana Ribeiro de Mello - UFMG**



---

**Vera Lúcia Menezes de Oliveira e Paiva - UFMG**



*Para Lucas, meu filho.*

(que tem a mesma idade desta tese e cuja língua encontra-se em plena construção!)



## Agradecimentos

Em primeiro lugar, gostaria de agradecer ao Prof. Fábio Alves pela condução competente da orientação, pelo apoio incondicional em momentos cruciais ao longo do processo de doutoramento e permanente incentivo desde minha defesa de mestrado na Unicamp, em 2000. Mais que orientador oficial de uma tese, será sempre um grande amigo.

Ao Prof. Nick Ellis agradeço pela supervisão da pesquisa conjunta desenvolvida na Universidade de Michigan, Ann Arbor. A convivência intensa com o Nick durante um ano (seu ritmo acelerado, suas reflexões profundas e generosidade inconmensurável) já deixaram marcas no que sou hoje. Conseguimos manter uma sintonia bem fina! Mais que supervisor de pesquisa e co-autores, nos tornamos amigos.

Agradecimentos ainda aos Professores Heliana Mello, Ricardo Souza e Thais Cristófaros-Silva pela participação, leitura crítica e sugestões valiosas durante o Exame de Qualificação de Tese. Agradeço ainda aos membros da Banca de Defesa de Tese pelas sugestões dadas, algumas delas já incorporadas ao texto presente. Desnecessário dizer que possíveis *erros* ou *imprecisões* que permanecem ao longo destas páginas são de minha inteira responsabilidade.

À CAPES, sou grato pela concessão de um ano de bolsa sanduíche (processo BEX # 0043/06-0), e ao CEFET-Ouro Preto agradeço o afastamento integral de minhas atividades de docência por um período de dois anos. Sou grato ainda às colegas Consuelo Perdigão e Silvana Avelar pelo apoio no início de meu curso de doutorado que se traduziu, principalmente, pela diminuição de minha carga horária de aulas.

Um muito obrigado aos colegas Daisy Rodrigues do Vale, Sérgio Elias da Silva e José Luiz Vila Real Gonçalves, pelo importante estímulo dado a mim quando esta tese ainda era um pré-projeto de pesquisa com cerca de dez páginas.

Agradecimentos também às funcionárias do PosLin, sempre prontas em sanar minhas dúvidas.

Do período em que passei em Ann Arbor só trago boas memórias. Impossível citar todas as pessoas que tive a felicidade de conviver, mas torno pública aqui minha mais profunda gratidão à Rosa Peters. Desde minha chegada ao aeroporto de Detroit naquele final de primavera até os últimos momentos em Ann Arbor ela esteve ao meu lado, juntamente com o Brian e a Kelly. Deram-me abrigo e carinho sempre! E você, Rosa, me apresentou Waninne, que, por sua vez, me apresentou a Silvia. Wannine Lima e Silvia Miranda, agradeço pelo prazer de cada encontro que tivemos (quase sempre proporcionados pela *bon-vivant* Wannine!). Agradeço a agradável companhia de Birgit Rieck no inesquecível show do Bob Dylan, naquela noite de muita neve em Detroit (*danke schön!*). Participar do *Dialectic Dining Club* foi também uma experiência agradabilíssima, uma criação do hilário, brilhante e lingüista poliglota Gregory Garretson (*many thanks, Gregory!*).

Do *English Language Institute* (ELI), casa de Charles C. Fries, agradeço e destaco a companhia agradavelmente sueca de Annelie Adel, a delicadeza de Jinyun Ke, a presteza e competência de Lisa Russ, e a simpatia e espíritosidade do ilustre John Swales (que, por já ter ido a minha bela Maceió, às vezes me chamava pelo nome da cidade!). Finalmente, sou grato por ter convivido com a brilhante Diane Larsen-Freeman. Aos demais colegas do ELI, meus mais sinceros agradecimentos!

Aos queridos Pedro Meira e Andréa Melloni, amigos e incentivadores de longa data, agradeço pela acolhida calorosa e pelos bons momentos vividos na charmosa Princeton.

Ao Antonio Carlos Oliveira e Vera Pacheco, muito obrigado por me darem carinho e abrigo quando eu mais precisei durante meu conturbado retorno ao Brasil. Jamais esquecerei! Uma amizade que se mantém sólida desde Campinas.

Meus agradecimentos aos Kehdy (Sérgio, Regina, Fernanda e Rafael), ao Leonardo e ao Chico pelo apoio e momentos agradáveis vividos em Nova Lima e Maceió.

Agradeço a minha querida mãe, Mirian, e aos meus irmãos, Adriana e André, pelo carinho e por sempre torcerem pela minha felicidade, esteja eu onde estiver. Um muito obrigado deste ente familiar exilado!

*Last but not least*, meus sinceros agradecimentos a Àdna! Hoje, dia nove de julho, é um dia realmente especial para nós dois!



## Resumo

Esta tese investiga a natureza dos processos cognitivos básicos envolvidos na aquisição de uma segunda língua (L2) a partir de mecanismos de construção que utilizam fórmulas lingüísticas fixas ou entrincheiradas (i.e., sociocognitivamente determinadas). Processos de *chunking*, que operaram desde o nível neurobiológico e são mediados por processos probabilísticos e de atenção, parecem guiar o desenvolvimento de nosso conhecimento lingüístico. O caráter dinâmico e emergente de tal concepção de aprendizagem requer a adoção do paradigma conexionista, cujo dinamismo é altamente compatível com a noção de aprendizagem, representação e processamento de línguas naturais defendida nesta tese (i.e., a estrutura lingüística vista como um sistema complexo adaptativo). Além disso, uma abordagem de cunho construcionista e a concepção de um *continuum* léxico-gramatical na aprendizagem de línguas orientam as análises empíricas conduzidas nesta tese. O estudo consistiu numa análise da aquisição espontânea de três construções em língua inglesa (VL = verbo + locativo, VOL = verbo + objeto + locativo, e VOO = verbo + objeto + objeto ou bitransitivo) em um subcorpus longitudinal de aprendizes de inglês como L2 no Projeto ESF (PERDUE, 1993). Os resultados revelaram que tais construções são totalmente dominadas pelos verbos altamente freqüentes *go*, *put* e *give*, respectivamente, revelando um padrão de desenvolvimento semelhante para todos os sete aprendizes investigados, independentemente de sua língua nativa. Os dados analisados mostraram, ainda, que a distribuição probabilística observada na estrutura de línguas naturais, por meio da qual algumas palavras revelam-se mais freqüentes, mais genéricas e mais prototípicas, tende, posteriormente, a facilitar a aprendizagem de construções mais abstratas da estrutura argumental de uma língua. Finalmente, uma análise mais qualitativa dos dados encontrados em nosso estudo revelou que 1) protótipos semânticos possibilitam a aprendizagem de estruturas esquemáticas mais complexas de uma L2, resultantes, acima de tudo, de processos cognitivos contínuos de categorização, generalização e indução de regularidades presentes no insumo lingüístico de uma dada língua; 2) que tal processo é dependente da freqüência de *types* e *tokens* com os quais as construções ocorrem no insumo lingüístico; 3) que o desenvolvimento de interlínguas é pautado por um processo lento e gradual, inicialmente calcado em itens bastante concretos até o desenvolvimento de estruturas lingüísticas mais abstratas; e 4) que a aprendizagem de fórmulas ou *chunks* depende da freqüência de *tokens*, ao passo que os processos de esquematização e abstração dependem da freqüência de *types*.

Palavras-chave: 1. Aprendizagem de Línguas. 2. Chunking. 3. Conexionismo. 4. Construções. 5. Emergentismo. 6. Neurobiologia da Aprendizagem.



## *Abstract*

This thesis investigates the nature of some basic cognitive processes involved in the constructing of a second language (L2) and which underlie the acquisition and use of linguistic formulae and entrenched chunks of language (i.e., sociocognitively conventionalized ones). Chunking processes, which operate right from a neurobiological level and are mediated by probabilistic and attentional processes, seem to guide the development of our knowledge of language. The emergent and dynamic nature of such a view is utterly compatible with connectionist models of learning, representation and processing, i.e., the underlying conception of language as a complex adaptive system expressed throughout this thesis. Additionally, a constructionist approach and a conception of a lexis-grammar *continuum* in language learning guide the empirical analyses carried out here. The study comprised an analysis of the spontaneous acquisition of three different types of English constructions (VL = verb + locative, VOL = verb + object + locative, and VOO = verb + object + object or ditransitive) on a longitudinal corpus of naturalistic English learners in the ESF Project (PERDUE, 1993). Results show that these constructions are seeded by the highest-frequent verb exemplars *go*, *put* and *give*, respectively, thus revealing a quite similar developmental pattern for all learners investigated, irrespective of their native languages. Results also demonstrate that such probabilistic distribution observed in the structure of natural languages later facilitates the learning of more abstract verb argument constructions by the learner, due to their first acquiring the most frequent, prototypical, and generic exemplar. Finally, a more qualitative analysis of our data indicates that 1) semantic prototypes optimize the learning of more complex abstract constructions in an L2 and are the result, above all, of continuous cognitive processes of categorization, generalization, and induction of the regularities present in the linguistic input; 2) this process is crucially dependent on type and token frequencies with which particular constructions appear in the input; 3) L2 development is slow and gradual, moving from an initial heavy reliance on concrete items to the development of more abstract linguistic schemas; and, last but not least, 4) the learning of formulae and chunks depends crucially on token frequency, whilst schematization is heavily dependent on type frequency.

Keywords: 1. Language Learning. 2. Chunking. 3. Connectionism. 4. Constructions. 5. Emergentism. 6. Neurobiology of Learning.



## Sumário

|   |     |
|---|-----|
| Introdução  | 18  |
| 1. A neurobiologia da aprendizagem  | 33  |
| 1.1 A motivação   | 33  |
| 1.1.2 Princípios hebbianos  | 34  |
| 1.2 A atenção   | 37  |
| 1.3 A memória   | 42  |
| 1.3.1 Memória de trabalho e memória de longo prazo                          | 42  |
| 1.3.2 Memória procedimental: processos automáticos                          | 44  |
| 2. Conexionismo, probabilidades & chunking no processo de aprendizagem      | 47  |
| 2.1 Desenvolvimento e processamento do conhecimento conexionistas           | 51  |
| 2.2 Modelo de ativação interativa para a leitura                            | 54  |
| 2.3 Modelo de ativação interativa para a percepção da fala                  | 58  |
| 2.4 Probabilidades e o constante cálculo de incertezas                      | 58  |
| 2.5 As ‘irracionalidades’ do processo de aquisição de uma L2                | 62  |
| 2.6 Revisitando o conceito selinkiano de <i>interlíngua</i>                 | 67  |
| 2.7 A interlíngua vista como uma rede conexionista em estruturação          | 70  |
| 2.8 A linguagem como um sistema probabilístico                              | 72  |
| 2.9 Processos de <i>chunking</i>  | 74  |
| 3. Categorias e conhecimento semântico                                      | 79  |
| 3.1 Definindo termos  | 80  |
| 3.2 O impacto das idéias de Eleanor Rosch                                   | 81  |
| 3.3 Níveis de categorização e frequência de ocorrência de protótipos        | 83  |
| 3.4 Protótipos semânticos e aquisição de línguas naturais                   | 86  |
| 4. Construindo uma segunda língua   | 88  |
| 4.1 O léxico: de um dicionário a estados mentais                            | 89  |
| 4.2 O uso de <i>corpora</i> na investigação sobre a aprendizagem de línguas | 94  |
| 4.3 Construções   | 96  |
| 4.4 Construções em L1   | 99  |
| 4.5 Construções em L2   | 102 |
| 4.6 Como construímos o conhecimento gramatical                              | 105 |
| 4.7 Verbos abre-alas ( <i>path-breaking</i> )                               | 106 |
| 4.8 Lei de Zipf e frequência de <i>types</i> e <i>tokens</i> verbais        | 107 |
| 4.9 Em busca da síntese: neurobiologia, conexionismo e construções          | 110 |

|   |     |
|---|-----|
| 5. Análise longitudinal de um corpus de aprendizes de segunda língua  | 112 |
| 5.1 O Projeto ESF   | 114 |
| 5.2 Sujeitos investigados   | 115 |
| 5.3 Justificativa e característica do subcorpus utilizado             | 116 |
| 5.4 Procedimento de análise   | 118 |
| 5.5 Resultados  | 123 |
| 5.6 Curvas individuais de aquisição cumulativa                        | 130 |
| 5.7 Análise longitudinal dos entrevistadores no subcorpus ESF         | 131 |
| 5.8 Discussão: comparando os resultados de frequência verbal          | 136 |
| 5.8.1 Muito <i>type</i> para pouco <i>token</i>                       | 138 |
| 5.8.2 Para além da frequência verbal                                  | 140 |
| 6. Estudo sobre prototipicidade verbal na língua inglesa              | 142 |
| 6.1 Materiais   | 142 |
| 6.2 Sujeitos  | 143 |
| 6.3 Coleta de dados   | 143 |
| 6.4 Resultados  | 144 |
| 6.5 Correlacionando frequência e prototipicidade verbais              | 145 |
| 6.6. Discussão: do genérico para o específico                         | 149 |
| 7. Reflexões para além de frequência e prototipicidade verbais        | 153 |
| 7.1 De volta às redes   | 154 |
| 7.2 Análise qualitativa do uso do verbo <i>go</i> em construções VL   | 159 |
| 7.2.1 Andrea  | 159 |
| 7.2.2 Lavinia   | 169 |
| 7.2.3 Jarnail   | 178 |
| 7.2.4 Ravinder  | 184 |
| 7.2.5 Discussão   | 192 |
| 7.3 Análise qualitativa do uso do verbo <i>put</i> em construções VOL | 195 |
| 7.3.1 Andrea  | 195 |
| 7.3.2 Lavinia   | 201 |
| 7.3.3 Jarnail   | 205 |
| 7.3.4 Ravinder  | 209 |
| 7.3.5 Discussão   | 214 |
| 7.4 Análise do uso das construções VOO                                | 218 |
| 7.4.1 Discussão   | 220 |
| 7.5 De ilhas verbais a um arquipélago construcional                   | 224 |
| Considerações finais  | 227 |

|  |     |
|--|-----|
| Referências bibliográficas   | 233 |
| Anexo A - Material utilizado nos experimentos sobre prototipicidade  | 243 |
| Tabela 1 - Verbos mais frequentes em cada construção   | 242 |
| Tabela 2 - Escala de cinco pontos para a classificação dos verbos  | 242 |
| Tabela 3 - Instruções dadas aos participantes para cada tarefa   | 242 |
| Tabela 4 - Modelo de formulário usado no registro das informações pessoais dos participantes e resultados das classificações | 244 |
| Tabela 5 – Frases usadas no estudo de Gries & Wulff  | 245 |
| Anexo B – Informações sociobiográficas dos sujeitos  | 246 |
| 1. Informações sociobiográficas de Andrea  | 246 |
| 2. Informações sociobiográficas de Lavinia   | 247 |
| 3. Informações sociobiográficas de Santo   | 247 |
| 4. Informações sociobiográficas de Vito  | 248 |
| 5. Informações sociobiográficas de Madan   | 248 |
| 6. Informações sociobiográficas de Ravinder  | 249 |
| 7. Informações sociobiográficas de Jarnail   | 250 |
| Anexo C - CD-ROM com subcorpus ESF e análises de frequência conduzidas   |     |

### Índice de Tabelas

|   |     |
|---|-----|
| Tabela 1 - Exemplos de construções e suas respectivas variações em tamanho e complexidade         | 97  |
| Tabela 2 - Exemplos de construções VL, VOL e VOO analisadas                                       | 100 |
| Tabela 3 - Resumos dos resultados encontrados nos dados de crianças e mães                        | 101 |
| Tabela 4 – Lista de atividades realizadas e seus respectivos códigos                              | 117 |
| Tabela 5 – L1 e L2 dos sujeitos, número total de sessões e datas de início e conclusão no projeto | 118 |
| Tabela 6 - Grau de prototipicidade dos verbos de acordo com as construções nas quais ocorrem      | 145 |

### Índice de Gráficos

|  |     |
|--|-----|
| Gráfico 1 - Total de construções dos aprendizes                  | 124 |
| Gráfico 2 - Construções VL dos aprendizes                        | 125 |
| Gráfico 3 - Curva de aquisição cumulativa VL para os aprendizes  | 126 |
| Gráfico 4 - Construções VOL dos aprendizes                       | 127 |
| Gráfico 5 - Curva de aquisição cumulativa VOL para os aprendizes | 128 |
| Gráfico 6 - Construções VOO dos aprendizes                       | 129 |
| Gráfico 7 - Curva de aquisição cumulativa VOO para os aprendizes | 130 |
| Gráfico 8 - Total de construções dos entrevistadores             | 132 |
| Gráfico 9 - Construções VL dos entrevistadores                   | 133 |
| Gráfico 10 - Construções VOL dos entrevistadores                 | 134 |

|   |     |
|---|-----|
| Gráfico 11 - Construções VOO dos entrevistadores  | 135 |
| Gráfico 12 - Comparação da frequência verbal nas construções em L1 e L2   | 137 |
| Gráfico 13 - Correlação entre prototipicidade (y) e frequência (x) verbais para a construção VL                                 | 147 |
| Gráfico 14 - Correlação entre prototipicidade (y) e frequência (x) verbais para a construção VOL                                | 148 |
| Gráfico 15 - Correlação entre prototipicidade (y) e frequência (x) verbais para a construção VOO                                | 149 |
| Gráfico 16 - Rede de relações das palavras presentes nas 436 construções VL dos aprendizes                                      | 155 |
| Gráfico 17 - Rede de relações das palavras presentes nas 52 construções VL envolvendo o type verbal <i>come</i>                 | 156 |
| Gráfico 18 - Rede de colocações de pronomes e preposições nas 52 construções VL dos aprendizes com o verbo <i>come</i>          | 157 |
| Gráfico 19 - Curva cumulativa de aquisição VL de Andrea   | 161 |
| Gráfico 20 - Curva cumulativa de aquisição VL de Lavinia  | 170 |
| Gráfico 21 - Curva cumulativa de aquisição VL de Jarnail  | 179 |
| Gráfico 22 - Curva cumulativa de aquisição VL de Ravinder   | 186 |
| Gráfico 23 - Rede de colocações mais frequentes em torno do verbo <i>go</i> na construção VL para o grupo de aprendizes         | 193 |
| Gráfico 24 - Comparação das redes de colocações em torno do verbo <i>go</i> em construções VL de aprendizes e entrevistadores   | 194 |
| Gráfico 25 - Curva cumulativa de aquisição VOL de Andrea  | 195 |
| Gráfico 26 - Curva cumulativa de aquisição VOL de Lavinia   | 202 |
| Gráfico 27 - Curva cumulativa de aquisição VOL de Jarnail   | 206 |
| Gráfico 28 - Curva cumulativa de aquisição VOL de Ravinder  | 210 |
| Gráfico 29 - Rede de relações entre palavras nas 152 construções VOL com o verbo type <i>put</i> produzidas pelos aprendizes    | 214 |
| Gráfico 30 - Rede de colocações mais frequentes na construção VOL em torno do verbo <i>put</i> produzidas pelos aprendizes      | 215 |
| Gráfico 31 - Comparação em rede das colocações mais frequentes na construção VOL nos grupos de aprendizes e entrevistadores     | 217 |
| Gráfico 32 - Redes de relações de palavras nas construções VOO (rede superior) e colocações em tais construções (rede inferior) | 221 |
| Gráfico 33- Comparação das redes de colocações mais frequentes na construção VOO de aprendizes e entrevistadores                | 222 |

### Índice de Figuras

|   |    |
|---|----|
| Figura 1 - Classificação hierárquica da memória   | 42 |
| Figura 2 - Rede recorrente simples (SRN) com unidades de contexto   | 53 |
| Figura 3 - Identificação da letra <i>E</i> com base na percepção de um dado conjunto de traços                            | 56 |
| Figura 4 - Identificação de palavras com base na ativação ou inibição de letras e percepção de um dado conjunto de traços | 57 |
| Figura 5 - Curva de erro numa rede treinada em tarefa de previsão de fonema   | 76 |



|  |     |
|--|-----|
| Figura 6 - Visualização esquemática, em 3D, do estado espacial multidimensional gerado pela camada oculta de uma rede conexionista SRN | 93  |
| Figura 7 - Diagrama horizontal com os resultados do experimento de classificação de frases   | 104 |
| Figura 8 - Configuração das línguas fonte e alvo no Projeto ESF  | 114 |
| Figura 9 - Configuração das línguas fonte e alvo no subcorpus analisado  | 115 |
| Figura 10 - Ilustração da seqüência de aquisição da construção VOL   | 225 |

## Índice de Quadros

|  |     |
|--|-----|
| Quadro 1 - Estágios na aquisição do tempo passado do verbo <i>eat</i>  | 65  |
| Quadro 2 - Visão geral dos dados transcritos   | 116 |
| Quadro 3 - Sumário dos verbos mais freqüentes em cada construção usados pelos aprendizes   | 131 |
| Quadro 4 - Sumário dos verbos mais freqüentes usados em cada construção pelos entrevistadores  | 136 |
| Quadro 5 - Comparação do número de verbos usados por aprendizes e entrevistadores  | 138 |
| Quadro 6 - Tarefas nas quais os verbos prototípicos <i>go</i> , <i>put</i> e <i>give</i> ocorrem em suas respectivas construções VL, VOL & VOO | 159 |
| Quadro 7 - Sessões de Andrea nas quais ocorre a construção VL  | 160 |
| Quadro 8 - Sessões de Lavinia nas quais ocorre a construção VL   | 170 |
| Quadro 9 - Sessões de Jarnail nas quais ocorre a construção VL   | 179 |
| Quadro 10 - Sessões de Ravinder nas quais ocorre a construção VL   | 185 |
| Quadro 11 - Sessões de Andrea nas quais ocorre a construção VOL  | 196 |
| Quadro 12 - Sessões de Lavinia nas quais ocorre a construção VOL   | 201 |
| Quadro 13 - Sessões de Jarnail nas quais ocorre a construção VOL   | 205 |
| Quadro 14 - Sessões de Ravinder nas quais ocorre a construção VOL  | 209 |
| Quadro 15 - Todas as construções VOO produzidas pelos aprendizes   | 219 |
| Quadro 16 - Todas as 36 construções VOO dos aprendizes   | 221 |
| Quadro 17 - Ilustração do processo de abstração da construção VOO  | 223 |

## Introdução

*o vento aeiouava*

Guimarães Rosa

Esta tese tratará de uma das habilidades mais complexas, e por isso mesmo difícil, que um adulto pode desenvolver: aprender a se comunicar numa segunda língua<sup>1</sup>. Apesar de complexa, muitas pessoas aprendem, em graus variados de proficiência, uma segunda ou várias outras línguas<sup>2</sup>. Uma língua é algo sempre dinâmico, ativo. Também, ao longo de seu desenvolvimento, uma língua exibe períodos de maior ou menor dinamicidade<sup>3</sup>. Tais estágios observados na aprendizagem de um sistema lingüístico específico serão sempre diferentes, variando de acordo com a experiência (aqui entendida como todo e qualquer *background*) de uma determinada pessoa. Estas diferenças dar-se-ão nos múltiplos níveis de organização sistêmica (i.e., tradicionalmente conhecidos como níveis fonológico, lexical, morfossintático, discursivo, pragmático, etc.), provocando, assim, variações no uso de uma língua por parte das pessoas<sup>4</sup>.

---

<sup>1</sup> Embora a reflexão aqui feita aplique-se tanto à língua materna (L1) como a uma segunda língua (L2), minha preocupação maior é com área de aquisição de segunda língua (ASL), nesta tese também entendida como sinônima de aquisição de língua estrangeira (ALE).

<sup>2</sup> Nos extremos de um suposto contínuo de aprendizagem poderíamos colocar tanto aqueles aprendizes com pouca ou nenhuma intimidade no uso de uma L2, passando por alguns dedicados guias de turismo (que aprendem uma, duas ou até três línguas estrangeiras) até, para nossa admiração, pessoas como o célebre viajante e explorador britânico Sir Richard Francis Burton (poliglota, falante de cerca de 25 idiomas).

<sup>3</sup> Nesta tese, ao me referir a 'interlínguas', evito ao máximo utilizar o termo *fossilização*. Tal termo, em minha opinião, contradiz, de uma certa forma, a idéia aqui defendida de que uma língua é um sistema probabilístico altamente dinâmico. Acima de tudo, na discussão acerca dos processos cognitivos envolvidos na aquisição de línguas, é importante substituir conceitos categóricos por outros que expressem gradientes.

<sup>4</sup> Neste manuscrito, cuja investigação dos processos de aprendizagem se dará desde as suas supostas bases neurobiológicas, veremos que tais diferenças são, em última instância, produtos de cérebros que são sempre também, ontogeneticamente, diferentes uns dos outros nos níveis microestruturais (como defende Schumann, 2004:2). Tais diferenças são o que chamamos, no nível comportamental, de 'aptidões'. Esta discussão será retomada no item 1.1.

A bem da verdade, tais variações estão presentes no uso de toda e qualquer língua, não apenas na aquisição de uma L2. É consenso que há variações intra-ideoletais condicionadas pelo meio, determinando, desta forma, mudanças/alterações sutis nos níveis fonológico, lexical, morfossintático, discursivo e pragmático na produção de um falante nativo dependendo do contexto. Este manuscrito, portanto, concebe a linguagem como um sistema altamente dinâmico estruturado em níveis; o grau de dinamismo do sistema está diretamente relacionado aos padrões probabilísticos de conectividade/interação (i.e., regularidades estruturais) observados nos vários níveis de organização sistêmica.

Apesar de ainda ser considerado tabu em muitos departamentos de Lingüística mundo afora, defendo aqui que a ‘construção’ de uma língua é, em primeiro lugar, um problema psicológico. A aprendizagem de uma língua (i.e., sua gradual e constante estruturação) é resultante, acima de tudo, de processos cognitivos de categorização, generalização e indução de padrões. Tal “estrutura”, importante desde já destacar, é emergente e maleável, e por isto mesmo altamente condicionada pelo contexto<sup>5</sup> (i.e., contexto-dependente).

A concepção de linguagem como um sistema emergente e intrinsecamente ligado aos nossos processos cognitivos gerais, defendida nas páginas que se seguem, resulta de

---

<sup>5</sup> **Contexto**, ao longo desta tese, é entendido como algo social, interativo e altamente dinâmico, equivalente à noção expressa em Goodwin & Duranti (1992). Os autores chamam o fenômeno a ser contextualizado de *evento focal*, em torno do qual se localiza uma *estrutura* (um *frame*, que convencionalmente chamamos de *contexto*) que “fornece os recursos para a sua interpretação adequada” (p.3). Tal concepção do que seja contexto encontra-se na etimologia da própria palavra, no vocábulo latino *contextus* (uma junção ou ajuntamento). Para os autores, a relação entre *evento focal* e *contexto* é similar àquela entre *organismo* e *meio ambiente*. Tal definição permite delimitar os limites entre o contexto e o evento focal. Para tal, devemos levar em conta 1) a perspectiva do *agente do contexto*; 2) o conhecimento sócio-histórico (regras de viver) possuído por tal agente sobre o mundo no qual vive; e 3) a natureza dinâmica, emergente e altamente fugaz de contextos instanciados em eventos específicos. A diferença principal entre evento focal e contexto é a *saliência perceptual*. O evento focal é sempre o enfoque oficial, ou primeiro (veja as análises lingüísticas bloomfieldianas e chomskianas, que atribuem total autonomia à estrutura e à sentença, respectivamente, sendo o contexto completamente ignorado ou relegado como algo não tão importante para análise). Em virtude disto, o evento focal torna-se claramente delimitado, ao passo que o contexto torna-se cada vez mais nebuloso e intangível (uma relação *figure-ground*, termo usado pelos autores). A lingüística, segundo Goodwin & Duranti, tem se beneficiado deste estado de coisas para poder, assim, legitimar seus modelos. Tal concepção de contexto como algo social e interativo subjaz à filosofia emergentista, conexionista e de sistemas complexos, que orienta teoricamente esta tese. Finalmente, para Goodwin & Duranti (1992), tal noção do que vem a ser o *contexto* é fundamental para compreendermos, dentre outras coisas, a “relação entre língua, cultura e organização social”, bem como para entendermos “como a língua é estruturada da maneira como ela é” (p.32).

teorizações recentes a partir da lingüística cognitiva (Langacker, 2008; 1999) e modelos de aquisição de línguas baseados no uso, tanto de cunho conexionista e emergentista, como sugerido por Bybee & McClelland (2005), Elman et.al. (1998), McClelland et. al. (1986), Ellis (1998, 2003), como reflexões inspiradas em teorias do caos e sistemas complexos adaptativos, como se observa nas discussões em Ellis & Larsen-Freeman (2006), Ellis (2006a) Larsen-Freeman (1997), e, finalmente, teorias contrucionistas de aprendizagem e uso da linguagem propostas principalmente por Goldberg (2006, 2003), Goldberg & Casenhiser (2008), Goldberg, Casenhiser & Sethuraman (2004) e Tomasello (2003).

Como sugerido por Seidenberg (1997), o mesmo mecanismo cognitivo responsável pelo uso da língua também parece operar durante a sua aprendizagem. Assim, complementando a noção de linguagem como um sistema dinâmico emergente, é aqui também entendido que todo o sistema se estrutura a partir de ‘construções’, desde as mais elementares como bigramas e trigramas (i.e., o que entendemos como conhecimentos fonotático e ortográfico), passando pelas raízes de palavras, sufixos e prefixos (a morfologia) até construções mais complexas como sintagmas verbais/nominais (i.e., a sintaxe), revelando, desta forma, a noção de um sistema totalmente uniforme, como destaca Ellis (2003) ao comentar concepções contrucionistas recentes de língua, como aquelas difundidas pela lingüística cognitiva:

Tais visões destacam as construções como sendo padrões recorrentes de elementos lingüísticos com uma função lingüística bem definida. Elas podem ocorrer no nível da sentença (como o imperativo, os transitivos direto e indireto, as perguntas afirmativo-negativas) ou outros níveis mais básicos (o sintagma nominal, o sintagma preposicional, etc.). Enquanto a Teoria dos Princípios e Parâmetros rejeita as construções, considerando-as epifenômenos resultantes da interação entre princípios e parâmetros, por um lado, e o léxico, de outro, a lingüística cognitiva – em particular a gramática das construções (Croft, 2001; Goldberg, 1995) – trouxe a questão das construções à baila, ao cogitar, em vez disso, que são as próprias sistematicidades (regras) que emergem a partir da interação entre construções de natureza as mais variadas. (p.64)<sup>6</sup>

---

<sup>6</sup> Minha tradução de: *These focus on constructions as recurrent patterns of linguistic elements that serve some well-defined linguistic function. These may be at sentence level (such as the imperative, the ditransitive, the yes-no question) or below (the noun phrase, the prepositional phrase, etc.). Whereas Government and Binding Theory denied constructions, viewing them as epiphenomena resulting from the interaction of higher-levels principles-and-parameters and lower level lexicon, cognitive linguistics – construction grammar in particular (Croft, 2001; Goldberg, 1995) – has brought them back to the fore, suspecting instead that it is the higher-level systematicities that emerge from the interactions of constructions large and small.*

Portanto, partindo do pressuposto que a linguagem é toda ela construída, buscarei investigar, especificamente, o papel desempenhado pelas fórmulas lingüísticas<sup>7</sup> nos processos de aquisição e uso de uma língua. Como destacam Tomasello (2000, 2003), Skehan (1998) e Nick Ellis (2001, 2002a, 2002b, 2003 e 2004), dentre outros, o uso de fórmulas lingüísticas parece revelar um estágio de desenvolvimento sistêmico (seja na L1 ou L2) no qual construções mais criativas emergem a partir da recombinação de partes de fórmulas fixas (i.e., um “caleidoscópio lingüístico”)<sup>8</sup>.

Tomasello (2000), cuja teorização enfoca a aprendizagem de L1, sugere que, inicialmente, a criança faz uso da ‘imitação’ no processo de aquisição da língua. Esta é uma aprendizagem cultural, acima de tudo; um mecanismo de socialização (e, como veremos mais adiante, no item 1, este é um processo diretamente relacionado com o mecanismo biológico de sociostase). Em seguida, a criança faz uso de uma linguagem calcada em fórmulas fixas e no acúmulo gradual de combinações em torno destas fórmulas (para Tomasello, sempre a partir de um verbo). É com base no aumento gradual no uso destas fórmulas fixas (por exemplo, *I think he’s gone*, *I think it’s here*, *I think my daddy took it*, *I think I saw one*, *It’s a crazy bone*, *I think*), sua co-ocorrência e constantes e crescentes re-combinações (ou variações sobre uma fórmula padrão), que a criança conseguirá induzir aspectos estruturais da língua, ou seja, detectar padrões estruturais similares a partir de esquemas gerais sobre a estrutura de relações, independente das palavras específicas que compõem tais construções. Só então é que níveis abstratos de uso, baseados em categorias, são aprendidos. Por exemplo, a experiência com construções do tipo “*I draw tree*”, “*she kissed me*”, “*I hit Jeffrey*”, “*You hug Mommy*”, “*Jammie kicking ball*”, etc., levam a criança a construir, inicialmente, o esquema “X faz algo com Y”, que, por sua vez, será a base para o desenvolvimento futuro das categorias abstratas sujeito, verbo e objeto (p.161-162).

---

<sup>7</sup> Fórmulas lingüísticas são expressões lexicais fixas e convencionadas. Tais fórmulas são aqui entendidas como pedaços (*chunks*) da língua armazenados na memória, seja como unidades inteiras (ex., ‘eu não tenho a menor idéia’) ou como unidades parcialmente analisadas (‘eu gostaria de falar com \_\_\_\_’). As fórmulas são, assim, elementos lexicais por natureza, como destaca Rod Ellis (1997).

<sup>8</sup> Uma metáfora bastante ilustrativa de um mecanismo desta natureza foi mencionada certa vez em um grupo de discussão da UCSD: a metáfora do caleidoscópio. Ou seja, temos peças relativamente estáveis (semelhantes a *chunks*) que se agrupam em formas totalmente diferentes a cada momento. Tais configurações “emergentes” (altamente sensíveis ao momento ou contexto) seriam a estrutura ou gramática de uma língua. Acredito que tal metáfora resume, de maneira precisa, a discussão desenvolvida ao longo deste trabalho.

Mais adiante veremos que estudos psicolinguísticos recentes com bebês e adultos, envolvendo gramáticas artificiais, corroboram a teorização de Tomasello (2000, 2003) sobre a aprendizagem de construções linguísticas abstratas.

Da mesma forma Skehan (1998) e Ellis (2001) sugerem que o uso da linguagem é pautado basicamente na utilização da memória. Segundo os autores, o uso da língua caracteriza-se essencialmente pela utilização de elementos (pedaços) lexicais dos mais variados tamanhos, combinados de forma a produzir um texto contínuo e coerente. Por exemplo, ambos os autores citam o clássico e influente artigo de Pawley e Syder (1983). Segundo estes últimos, a concepção da linguagem a partir de regras linguísticas seria mais uma imposição do lingüista do que uma realidade incontestável, e que a realidade por nós verificada no uso da linguagem aponta para um uso repetitivo e nada criativo, o oposto da tão propagada capacidade de nosso sistema linguístico de produção contínua de novas (de fato, infinitas) frases no nosso *linguajar* diário, como defende Pinker, (1994:22). Na realidade, como destacam Langacker (2008), Skehan (1998) e Ellis (2001), o que realmente parece ocorrer é que no uso de uma língua, seja L1 ou L2, temos um número bastante reduzido de escolhas<sup>9</sup>.

A observação acima não implica na negação do caráter criativo da linguagem, mas sim reflete a constatação de que muito da nossa produção linguística seja, possivelmente, a repetição (co-ocorrência) de discursos anteriores. O uso natural da língua é calcado em assuntos triviais e previsíveis, portanto altamente repetitivos e nada originais. A originalidade é algo bastante incomum no uso corriqueiro de uma língua. Por exemplo, Langacker (2008) destaca que construímos gradualmente um inventário de expressões (ou fórmulas) lexicalizadas fixas e convencionalizadas, sendo o bom emprego de tais expressões (i.e., a idiomaticidade) o que caracterizaria o que chamamos de falante fluente, ou seja, aquele que expressa certas noções de uma maneira convencional ou usual numa dada língua:

---

<sup>9</sup> Obviamente, discordaríamos de maneira radical da observação acima após ler/ouvir a poesia de um Bob Dylan ou nos depararmos com um neologismo brilhante de um Guimarães Rosa, como vemos na epígrafe que abre esta tese, não acha?

Sem um conjunto substancial de expressões pré-fabricadas, o falar fluente em tempo real dificilmente seria possível. Teóricos têm de forma grosseira exagerado o aspecto original das “frases novas“ (p.19)<sup>10</sup>

Na mesma linha de argumentação, Skehan (1998) destaca

Por motivos vários não poderia ser diferente, já que poderíamos ter muita dificuldade em lidar com uma ‘dieta lingüística’ rica em linguagem criativa (e uma constante profusão de novos significados); é provável que isto se tornasse até uma ameaça. Talvez uma das funções da literatura seja justamente institucionalizar as áreas nas quais o uso da linguagem possa ser exercido e exercitado de maneira criativa e original em seus mais variados graus (p.32).<sup>11</sup>

No entanto, para Skehan, há um sistema duplo de processamento da linguagem regido tanto por regras (i.e., *rule-based*) como por itens lexicais convencionalizados e estocasticamente combinados (i.e., *exemplar-based*) (p.53). Veremos ao longo desta tese que tal distinção, de uma perspectiva conexionista e construcionista, não faz sentido. Ou seja, as “regras” tão propaladas por alguns teóricos em seus modelos de aquisição da linguagem parecem ser justamente um subproduto emergente de uma atividade sistêmica contínua, em processos probabilísticos de busca de padrões e categorização por meio, principalmente, de mecanismos de *chunking*.

Acredito que haja, portanto, primazia e prevalência de um processamento lexical nos processos cognitivos de aquisição e uso de uma língua. Em última instância, é aqui defendido que o processamento/uso de uma língua se dá, principalmente, a partir do emprego dessas estruturas lexicais convencionalizadas, por meio das quais as proposições são comumente expressas – i.e., a aquisição e uso da língua caracterizam-se basicamente pela constante rotinização (ou convencionalização) de elementos lexicais dos mais variados tamanhos. Esta teorização vem sendo recentemente defendida na área de ASL principalmente, e de maneira bastante incisiva, por Nick C. Ellis (1998, 2001, 2002a, 2002b, 2003, 2004, 2006b, 2006c, 2007 e 2008). Tal concepção recebe forte

---

<sup>10</sup> Minha tradução de: *Without a substantial inventory of prefabricated expressions, fluent real-time speech would hardly be possible. Theorists have grossly exaggerated the novelty of “novel sentences”*

<sup>11</sup> Minha tradução de: *In many ways it could not be otherwise since one might well find a diet of creative language, encoding new meanings, rather difficult to handle, and even, possibly, threatening. One of the functions of literature, perhaps, is to institutionalize the areas where creative and original language use can be found and responded to with the intensity it requires.*

influência da teorização de Langacker (2008; 1999) no âmbito da lingüística cognitiva, como comentado no início deste manuscrito. O aspecto formal da linguagem (ou a gramática cognitiva da língua) é criado a partir das relações internas destes elementos semânticos, refletindo, assim, uma perspectiva bastante funcional - ou mesmo “microfuncional”, como coloca Clark (1989), se pensarmos em um processamento de base conexionista - para a aquisição, desenvolvimento e uso da linguagem<sup>12</sup>.

Os modelos e teorias de aquisição, acima discutidos, que enfatizam um papel mais central do léxico, em especial fórmulas lexicais (ou seja, modelos de base semântica), encontram a resistência daqueles que defendem modelos sintáticos de processamento da linguagem, i.e., baseados em regras, e geralmente inspirados na teoria da gramática gerativa defendida pela escola chomskiana (CHOSMKY, 1980,1998). Como sabemos, teorias inatistas nos estudos lingüísticos sustentam a existência de uma Gramática Universal (UG) - i.e., representações inatas - subjacente aos processos de aquisição da linguagem a partir de um mecanismo especial de aquisição (*language acquisition device* ou LAD) (cf. WHITE, 2003; PINKER, 1994). A defesa de um mecanismo deste tipo decorre da suposta “pobreza do insumo lingüístico” e da ocorrência de inúmeros fenômenos estruturais específicos em determinadas línguas, como reafirma White (2003:24)

Entretanto, há uma outra possibilidade de compreensão destas mesmas características estruturais presentes na aquisição da linguagem a partir de uma perspectiva não inatista, como veremos ao longo deste manuscrito.

O’Grady (2003), por exemplo, sugere que a representação nos moldes gerativistas (i.e., representações simbólicas estruturadas hierarquicamente na forma de processos lineares que ocorrem da esquerda para a direita) reflete processos cognitivos globais de composição, e não um mecanismo sintático específico (tipo UG), e destaca:

---

<sup>12</sup> Embora cite ao longo desta tese a visão langackeriana acerca dos processos envolvidos na construção de línguas naturais, não tratarei especificamente da teoria da Gramática Cognitiva (GC) desenvolvida por Langacker. Na verdade, inerentes à teorização da GC de Langacker estão pressupostos básicos não apenas da Lingüística Cognitiva, mas das teorizações emergentistas, conexionistas, de sistemas complexos, etc.



Pelo contrário, tal modelo reflete duas características computacionais mais globais – uma tendência a operar a partir de pares de elementos (também uma característica da faculdade aritmética) e uma propensão a combinar functores com seus argumentos na primeira oportunidade (uma estratégia de redução de armazenamento que eu chamo simplesmente de “eficiência”) (p.47)

O’Grady (2003:49) também sugere que a dificuldade encontrada por aprendizes quanto ao processamento de sentenças relativas (*subject relative vs. object direct relative*, como no estudo citado envolvendo aprendizes de coreano) é determinada principalmente pelo fator “distância estrutural” entre um *gap* e o seu complemento (*filler*) (i.e., *subject relatives* são mais fáceis). Assim, a dificuldade e complexificação observadas na estruturação do sistema lingüístico refletem os vários níveis de dependência estrutural (i.e., dependências de várias ordens).

Os argumentos de O’Grady são compatíveis com a teorização construcionista do sistema lingüístico a partir de processos de *chunking*, como dito acima uma visão principalmente defendida na área de ASL por Nick Ellis (2001, 2002a, 2002b, 2003 e 2004). Acredito que o que subjaz aquilo que O’Grady chama acima de “estratégia de redução de armazenamento” para um desempenho lingüístico eficiente seja, na verdade, um mecanismo de *chunking* (ou mecanismo de aprendizagem estatística, como preferem outros). Mais adiante revisaremos uma literatura recente que aponta nesta direção: MacWhinney (2005, 2001), Elman, Hare & McRae (2005), Bod et al. (2003) Seidenberg & MacDonald (2001), Seidenberg (1997). Toda esta discussão será retomada mais adiante, quando trataremos, respectivamente, da memória, da natureza probabilística de nossos processos cognitivos e dos processos de *chunking* na aquisição de línguas.

É importante fazer duas observações nesta introdução. A primeira diz respeito às diferenças existentes entre a aquisição de uma L1 por parte de crianças e de uma L2 por adultos. Os aspectos mais relevantes que merecem ser apontados parecem ser: 1) o grau de desenvolvimento conceitual e a possibilidade ou não de uso de estratégias e recursos metacognitivos e metalingüísticos por parte do indivíduo; 2) a natureza do insumo e o ambiente de aprendizagem; 3) a possibilidade de transferência lingüística, ou seja, o aprendiz de L2 já possui uma língua e utiliza tal estrutura ao longo de sua

aprendizagem. Entretanto, como destacam MacWhinney (2005) e Ellis (2001, 2002a, 2002b, 2003 e 2004), a natureza do mecanismo de aquisição utilizado (baseado na utilização de fórmulas e construções de base probabilística) é essencialmente a mesma, determinando os mesmos padrões sequenciais de desenvolvimento em ambos os casos<sup>13</sup>.

A segunda observação importante diz respeito à distinção entre *aprendizagem* e *aquisição* nos processos de ASL proposta por Stephen Krashen no início da década de 1980. Segundo Krashen (1982:10) haveria uma diferença fundamental entre os processos de ‘aquisição’ e os de ‘aprendizagem’ de uma L2. A aquisição seria resultante de um processo inconsciente (implícito) por meio do qual a língua é aprendida em decorrência da necessidade real de comunicação, semelhante à experiência vivenciada naturalmente pelas crianças na aprendizagem da L1, ao passo que a aprendizagem seria um processo consciente (explícito) que resulta em saber sobre a língua (i.e., um conhecimento metalingüístico). Para Krashen, o conhecimento oriundo de processos de aprendizagem (um conhecimento explícito e controlado) não levaria o aprendiz a desenvolver uma competência<sup>14</sup> lingüística chomskiana (i.e., um conhecimento implícito, automático e ‘correto’ do ponto de vista gramatical).

As duas bases de conhecimento propostas por Krashen, implícita (adquirida) e explícita (aprendida), seriam totalmente independentes (algo como módulos fodorianos). Ou seja, falar em aquisição e aprendizagem nos termos krashenianos implica não apenas reconhecer a existência de aprendizagem implícita (e seus processos automáticos) e explícita (e seus processos controlados), mas também admitir a impossibilidade de inter-relação/interface entre tais bases de conhecimento e tais tipos de processamento.

---

<sup>13</sup> Larsen-Freeman (1997:145-146), em uma excelente reflexão sobre a linguagem a partir da teoria do caos e sistemas complexos, argumenta que a linguagem é um ‘sistema fractal’ (i.e., um sistema cujo processo de organização estrutural é o mesmo ao longo de seus vários níveis constituintes). Ou seja, concepção totalmente compatível com as noções de aprendizagem/uso de uma língua (e seus respectivos processos de *chunking*) defendidas ao longo destas páginas.

<sup>14</sup> Nesta tese, sempre que o termo competência for mencionado sem referência explícita à competência chomskiana, ficará subentendido que este consiste na própria dinâmica sociocognitiva do sistema interlingual, resultante da constante busca por regularidade entre as partes (categorias) constituintes que emergem continuamente ao longo do desenvolvimento lingüístico; uma competência lingüística dinâmica e dialógica, permanentemente condicionada e moldada pelo insumo e pelo meio. Ou seja, uma noção de competência equivalente àquela proposta originalmente por Hymes (1972) e objeto de reflexão em Ferreira-Junior (2004).

Tal dicotomia é categoricamente rejeitada ao longo deste texto. Os termos ‘aprendizagem’ e ‘aquisição’ são empregados ao longo deste manuscrito como sendo sinônimos. Veremos que modelos conexionistas, teorizações e pesquisas recentes no âmbito da neurobiologia (assim como na lingüística cognitiva) aqui discutidas tornam a conceituação krasheniana totalmente inválida. Apesar de sabermos que há, de fato, duas bases ou sistemas de memória cujas naturezas diferem, teorias e pesquisas recentes têm enfatizado o caráter interativo e altamente dinâmico do processamento da linguagem nos níveis cognitivo e neurobiológico, no qual informações de diferentes natureza (não apenas lingüística) são ativadas a todo instante.

Por exemplo, para O’Reilly & Norman (2002) tais sistemas são complementares e atuam decisivamente na aprendizagem e consolidação da memória. Da mesma forma, Lee (2004:67) e Crowell (2004:99) destacam, respectivamente, que os subsistemas presentes no neocórtex (i.e., gânglios da base, responsáveis pelo processamento automático e memória procedimental/implícita) e os mecanismos presentes no hipocampo (processos controlados e memória declarativa/explicita) operam de forma paralela e interconectados:

Em outras palavras, a atividade paralela e a interconectividade dos sistemas dos gânglios da base com os sistemas do hipocampo deixam claro que ambos os tipos de aprendizagem não-declarativa e declarativa naturalmente vão sempre ocorrer, independente do contexto de aquisição. Entretanto, é provável que a aprendizagem por imersão seja, por natureza, ligeiramente mais procedimental, ao passo que a aprendizagem na sala de aula seja um pouco mais declarativa (CROWELL, 2004, p.99-100)<sup>15</sup>

Portanto, podemos entender as diferenças qualitativas observadas na aquisição de L1 e L2 em termos mais globais, influenciadas principalmente por fatores ontogenéticos e contextos específicos para cada situação de aprendizagem (como destacado anteriormente). É importante ainda ressaltar que a teorização de Krashen, embora

---

<sup>15</sup> Minha tradução de: *In other words, the parallel activity and interconnectivity of the basal ganglia system with the hippocampal system make it clear that both nondeclarative and declarative learning will naturally take place in any language learning scenario. However, it may be the case that immersion learning is slightly more procedural in nature, whereas classroom is slightly more declarative.*

importante por ter re-colocado o insumo como ponto central na pauta das discussões em ASL, foi totalmente influenciada pelas idéias de Noam Chomsky brevemente comentadas anteriormente (i.e., a conceituação de um LAD, de uma gramática universal e as noções de competência e desempenho lingüísticos).

Long & Doughty (2003:866) destacam que a aprendizagem de uma língua é, em última instância, uma atividade cognitiva que envolve processos psicolingüísticos e mudanças nos nossos estados mentais internos; ou seja, uma concepção de que a questão central na investigação em ASL são os processos cognitivos que envolvem representações e computações sobre tais representações. Diante disto, Long & Doughty são enfáticos em afirmar que

Há uma enorme diferença entre tal concepção e a visão prevalente entre alguns lingüistas aplicados que relacionam a área de “ASL” com praticamente qualquer pesquisa sobre o uso de uma língua por parte de falantes não-nativos. Boa parte da pesquisa que se inclui nesta definição mais abrangente possui valor e é séria, mas pouco tem a ver com a maneira pela qual se dá a **aprendizagem** de uma segunda língua – ou, pelo menos, uma conexão precisa ainda ser estabelecida (p.869).<sup>16</sup>

Este trabalho é uma humilde tentativa de estabelecer tal conexão. Busca-se, aqui, contribuir para a teorização recente de que a construção de uma língua baseia-se no uso; sua aprendizagem é uma conseqüência direta dos propósitos e usos que dela fazemos (veja uma excelente síntese em ELLIS & LARSEN-FREEMAN, 2006). Nas páginas a seguir buscarei, inicialmente, enfocar os processos cognitivos envolvidos na aquisição de línguas a partir de seus mecanismos neurobiológicos, para, posteriormente, tratar de modelos cognitivos baseados em fenômenos (comportamentos) lingüísticos observáveis (utilizando, como dados empíricos, uma análise longitudinal de um corpus de aprendizes de língua inglesa como segunda língua).

---

<sup>16</sup> Minha tradução de: *There is a big difference between that conception and the view prevalent among some applied linguists that would equate “SLA” with almost any research having to do with non-natives when they use a second language. Much of the work that would be included under the broader definition is rigorous and valuable, but little of it has anything to do with how people learn a second language – or, at least, a connection has yet to be demonstrated.*

Destaco, ainda, que ao concebermos a aquisição e o uso da linguagem nos termos propostos resumidamente acima, temos, obrigatoriamente, que repensar o modelo clássico vigente de representação do conhecimento, pois a natureza de tal representação determinará o grau de funcionalidade e dinamismo *vis-à-vis* aos usos (ou ativação) dos conhecimentos lingüísticos que possuímos. A nossa concepção acerca do funcionamento da memória torna-se, então, crucial por determinar a plausibilidade psicológica do modelo adotado para a organização, representação e utilização de tais conhecimentos. A concepção clássica de representação (serial, indexada e regida por processos modularizados), defendida principalmente por Fodor (1983) não parece dar conta da natureza dinâmica implícita na concepção da linguagem a partir de "pedaços" lexicais constantemente instanciados em concatenações convencionalmente determinadas. Na verdade, a concepção tradicional de representação do conhecimento é compatível apenas com a não menos ortodoxa concepção da linguagem como sendo um sistema modular e regido por regras, como comentado anteriormente, uma tradição de 2.500 anos no mundo ocidental, como sustentam Oaksford & Chater (2007) (discutirei o enfoque probabilístico para a cognição humana nos capítulos 2 & 3).

Assim, a defesa de um mecanismo de aquisição e uso da linguagem nos moldes defendidos nesta tese torna-se mais consistente quando apoiada em modelos conexionistas para o processamento de sistemas estruturados, do tipo PDP<sup>17</sup> (i.e., processamento distribuído em paralelo), como sugerido por McClelland & Rogers (2003), Rogers & McClelland (2004), Elman et al (1998), McClelland, Rumelhart & Hinton (1986), Clark (1993) e Broeder & Plunkett (1994), em oposição aos modelos simbólicos clássicos de processamento da linguagem, i.e., sintático-analíticos, baseados em princípios da lógica formal e/ou em concepções inatistas como a existência de uma gramática universal da linguagem ou módulos especializados em processar informações em domínios específicos, como sustentado por Pinker (1994) e Fodor (1983).

Como apontam vários teóricos que buscam explicações mais convincentes para a natureza do fenômeno da ASL, a crença em um LAD e em uma Gramática Universal

---

<sup>17</sup> O termo sinônimo "redes neurais" implica uma possível plausibilidade biológica devido à metáfora do cérebro em funcionamento, por meio de neurônios e sinapses, e também é usado nas teorizações de cunho conexionista sobre a estrutura e o funcionamento do sistema cognitivo humano.

diminui a importância da teorização em ASL sobre o processamento cognitivo da linguagem (bem como de todos os demais processos envolvidos em sua aprendizagem e uso).

Busca-se aqui, portanto, investigar os processos cognitivos envolvidos na aquisição de línguas da perspectiva de modelos conexionistas, da psicologia cognitiva e da neurobiologia da aprendizagem, avenidas ainda pouco exploradas na área de aquisição de segunda língua, como destacam Larsen-Freeman (1997), Skehan (1998), Broeder & Plunkett (1994), Schumann (2004) e Ellis (2002b), numa tentativa de ampliar a discussão teórica em ASL no contexto brasileiro.<sup>18</sup>

Finalmente, esta tese é a materialização da própria teorização aqui discutida. É o mais recente fruto de uma (re)construção contínua de estudos anteriores por mim realizados. Principalmente, ela consiste num claro aprofundamento, extensão e reelaboração de idéias (algumas ainda embrionárias) originalmente presentes na minha dissertação de mestrado (FERREIRA-JUNIOR, 2000) e também reelabora alguns argumentos e idéias expressos em outros artigos e escritos publicados desde então (FERREIRA-JUNIOR, 2004, 2005, 2007). Resulta, acima de tudo, do desejo de ampliar meus conhecimentos sobre os processos cognitivos envolvidos na aquisição e uso de línguas naturais. A análise de um corpus longitudinal de aprendizagem espontânea de uma L2 e o estudo experimental relatados nos capítulos 5 & 6 foram realizados durante o período sanduíche de um ano no *English Language Institute* (ELI) da Universidade de Michigan, em Ann Arbor, EUA, sob a supervisão de Nick C. Ellis.

A discussão aqui feita está estruturada da seguinte forma: inicialmente, no capítulo 1, é feita uma breve reflexão sobre a importância da motivação (1.1) e da atenção (1.2) nos processos de aprendizagem, levando em conta suas bases neurobiológicas. Em seguida, no item 1.3, analisaremos mecanismos de memória, também levando em conta suas

---

<sup>18</sup> Destaco aqui o trabalho sobre os fenômenos observados em ASL conduzido por Vera Menezes Paiva e seu grupo na Universidade Federal de Minas Gerais (UFMG) à luz da teorização de sistemas complexos adaptativos e teoria do caos. Igualmente, destaco o volume temático da Revista Brasileira de Linguística Aplicada (RBLA, 2005/2), organizado por Heliana Mello e dedicado ao tema cognição e ASL, em especial o capítulo de Mota & Zimmer (2005). Veja ainda uma excelente reflexão teórica em Zimmer (2008) sobre cognição, conexionismo e ASL. Embora recentes, até onde chega meu conhecimento, tais trabalhos são pioneiros no âmbito da Linguística Aplicada e estudos em ASL aqui no Brasil.

bases neurobiológicas e a importância da motivação e da atenção para a consolidação dos processos de memória.

No capítulo 2 trataremos de modelos conexionistas (e outros modelos estatístico/probabilísticos) de aprendizagem, paradigmas teóricos que vêm sendo recentemente re-elaborados na investigação do papel das probabilidades nos processos cognitivos do raciocínio humano e, em especial, na aquisição e uso de línguas, e sua relação com os mecanismos neurobiológicos de memória implícita discutidos, mais especificamente, no item 1.3.2 do capítulo 1. No item 2.9 discutiremos os processos de *chunking*, cuja gênese podemos observar na discussão feita no capítulo 1 e em modelos conexionistas, e o papel destes processos na estruturação da linguagem. Esta discussão é muito importante, pois o argumento central desenvolvido nesta tese é o de que a estruturação de uma língua (representada pelo seu nível mais “robusto” que é o domínio da sintaxe) é produto da interação e recorrência de fórmulas lexicais e construções ao longo dos seus diferentes e variados graus de abstração e complexidade (i.e., consiste essencialmente na criação de *chunks*), concepção esta encerrada nas epígrafes ao longo dos capítulos desta tese.

No capítulo 3 será feita uma breve discussão sobre conceitos, categorias e protótipos, assim como sobre a importância de entendermos como se dá a estruturação de nosso conhecimento semântico. O foco principal do capítulo é a teorização de Eleanor Rosch sobre o conceito de protótipos e o impacto atual destas idéias na Linguística Cognitiva, em modelos conexionistas e na aprendizagem de línguas. No Capítulo 4 discutiremos a teorização sobre *construções* e modelos construcionistas de aprendizagem de línguas naturais. A aprendizagem de uma língua consiste, em última instância, no desenvolvimento de uma rede de construções guiada por processos de categorização e generalização, nos moldes propostos originalmente por Rosch e reelaborados em modelos conexionistas e emergentistas.

Nos capítulos 5 & 6 são apresentados dados empíricos resultantes de dois estudos desenvolvidos e que, acredito, dão embasamento aos argumentos teóricos aqui defendidos: uma análise longitudinal de um corpus de aprendizes de segunda língua (capítulo 5) e um estudo sobre prototipicidade verbal (capítulo 6). Nestes capítulos são

apresentados os resultados quantitativos dos respectivos estudos, seguidos por suas respectivas análises e conclusões.

No capítulo 7 é feita uma análise qualitativa da produção dos sujeitos do estudo longitudinal descrito no capítulo 5, seguida de uma discussão.

Este trabalho, apesar de revelar uma forte tendência à discussão/reflexão de base teórica, busca também contribuir para embasar ou iluminar as práticas pedagógicas no ensino de línguas estrangeiras. As possíveis implicações pedagógicas desta pesquisa serão discutidas mais especificamente nas Considerações Finais desta tese. Por fim, nas últimas páginas deste manuscrito encontram-se as referências bibliográficas e os anexos.



## 1. A neurobiologia da aprendizagem

*But will SLA adopt a biological perspective? We would hope so. SLA is about learning; learning is mediated by the brain. Therefore, neurobiology is as central as linguistics to our enterprise.*

John Schumann

### 1.1 A motivação

Para o filósofo Bertrand Russell, o desejo seria o desencadeador de toda atividade humana. Em um célebre discurso, quando do recebimento do prêmio Nobel em 1950, Russell destacou que se quisermos saber qual será o comportamento de alguém basta conhecer seus desejos e as forças que regem tal sistema<sup>19</sup>.

Da mesma forma, nas teorias sociointeracionistas acerca dos processos de aprendizagem, influenciadas pelo pensamento de Vygotsky, a motivação assume um papel central. Lantolf & Pavlenko (2001), por exemplo, enfatizam a necessidade da relevância de algo para que ocorra a aprendizagem da L2. Os autores adotam a “teoria da atividade” (um rebento da filosofia vigotskiana), que sustenta que a nossa cognição é condicionada pelo contexto e por nossas práticas sociais. O fazer e o pensar são indissociáveis, e são sempre guiados pelo “como” agimos, “onde” agimos e o “porquê” agimos (p.145). A relevância gera a sensação de sentido da aprendizagem que, por sua vez, gera a ação. O significado subjacente a uma determinada atividade (cognitivamente

---

<sup>19</sup> Cito a seguir o trecho do discurso de Russell: “*All human activity is prompted by desire. There's a wholly fallacious theory, advanced by some earnest moralists to the effect that it is possible to resist desire in the interest of duty and moral principle. I say this is fallacious, not because no man ever acts from a sense of duty, but because duty has no hold on him unless he desires to be dutiful. If you wish to know what men will do, you must know not only, or principally, their material circumstances, but rather the whole system of their desires with their relative strengths*”.

criado) determina a “orientação do indivíduo para aprender ou não” (p.148). Desta relação é criada a motivação para aprender. Veremos, a seguir, que estudos básicos recentes no campo da neurobiologia da aprendizagem de uma L2 corroboram tanto a elucubração filosófica de Russell como as teorizações sociointeracionistas e socioculturais sobre o papel da motivação nos processos de aprendizagem de línguas.

Schumann & Wood (2004) destacam que a motivação, parte integrante da cognição, é indispensável para a aprendizagem. Para que algo seja aprendido é preciso haver inicialmente um desejo (uma motivação). Motivação, portanto, é também algo fundamental na discussão sobre os processos neurobiológicos em ASL, sendo uma das primeiras evidências comportamentais da gênese de processos sólidos de aprendizagem de uma língua. Os autores ressaltam que a aprendizagem profunda e sustentável (APS) difere tanto da aprendizagem canalizada como da aprendizagem observada em estudos psicolinguísticos, por exemplo. A APS é lenta, gradual e responsável por desempenhos considerados proficientes ou especializados, um processo que requer vários anos. A aprendizagem canalizada, por sua vez, é biologicamente determinada, apenas um desabrochar genético, como a visão, a audição, a locomoção, etc. Ou seja, a APS é entendida como processos automatizados que levam o cérebro a tornar-se especialista numa área na qual ele não possui, inicialmente, uma especialização, como jogar xadrez, tocar violino, traduzir, aprender uma L2, etc. (p.23).

### **1.1.2 Princípios Hebbianos**

No tocante ao substrato neural da motivação, Schumann & Wood (2004) destacam que estudos recentes revelam que a motivação é um comportamento cognitivo de alto nível (portanto observável), cujo mecanismo é bastante simples: tudo se resume a constantes avaliações, julgamentos, do estímulo/insumo (*stimulus appraisal*). O princípio neurobiológico subjacente a este mecanismo é o mesmo proposto por Donald Hebb, em 1949 (cf. CLARK, 1989). Ou seja, tal princípio de aprendizagem (conhecido como ‘hebbiano’) postula que se dois neurônios forem ativados simultaneamente a conexão entre eles (sinapse) será reforçada (o que, conseqüentemente, inibirá ativações entre outros neurônios), de forma que, em ocorrências subseqüentes do mesmo estímulo, tais conexões terão um nível de ativação maior dentro do sistema neural, pois se constituem

em configurações neuronais específicas já estabelecidas (ou *cell assemblies*), como destacam Jones (2004:26) e O'Reilly & Norman (2002:506-507).

O mecanismo neurobiológico da motivação é composto de três componentes: a *amígdala (cerebelar)*, o *córtex orbitofrontal* e o nosso *corpo* (i.e., nosso sistema nervoso periférico). Tais avaliações de estímulo baseiam-se essencialmente na noção biológica de *valor*. O valor é uma “propensão do organismo que o leva a certas preferências, capacitando-o a fazer escolhas diante de certas alternativas”, como apontam Schumann & Wood (2004:24).

Há três tipos de valores: homeostático, sociostático e somático. O valor homeostático mantém o organismo equilibrado no tocante aos seus aspectos fisiológicos (temperatura do corpo, respiração, batimento cardíaco, etc.). O valor sociostático promove a interação de um indivíduo com seus semelhantes, é um mecanismo que promove inter-relações no âmbito social. O valor somático é responsável pelo desenvolvimento de preferências e aversões por parte do indivíduo ao longo de sua ontogenia e depende dos dois mecanismos de atribuição de valores acima descritos. Desta forma o organismo, em seu viver, começa a desenvolver uma *memória de valores*, resumida por Schumann & Wood da seguinte maneira:

À medida que um organismo vivencia o mundo, ele vai encontrando as mais variadas situações de estímulo. Ele sente o impacto emocional de tais estímulos no que concerne a sua relevância em relação aos sistemas de valores homeostático, sociostático e somático. O organismo incorpora à sua memória (memória esquemática emocional [Leventhal, 1984] ou memória da categoria de valores [Eldeman, 1992]) as características das situações de estímulo encontradas e sua relevância em relação aos seus objetivos, sua habilidade em adaptar-se, seu significado hedônico e, em organismos mais complexos, seu senso de personalidade. Esta memória então torna-se parte do sistema de valores utilizados em estímulos subsequentes. (p.25)<sup>20</sup>

---

<sup>20</sup> Minha tradução de: *As an organism moves through the world, it encounters various stimulus situations. It experiences the emotional impact of these stimuli in terms of their relevance to homeostatic, sociostatic, and somatic value systems. It incorporates into its memory (schematic emotional memory [Leventhal, 1984] or value category memory [Eldeman, 1992]) the characteristics of the stimulus situation and their relevance to its goals, its ability to adapt, its hedonic sense, and in higher organisms, its sense of self. This memory then becomes part of the value system used in evaluation of subsequent stimuli.*

Da perspectiva neurobiológica, portanto, a motivação para a aprendizagem é gerada nos circuitos límbicos e paralímbicos que envolvem o córtex *orbitofrontal* e a *amígdala* e baseia-se na avaliação (positiva ou negativa) da habilidade/conhecimento a ser aprendido (p.29). Ambas as áreas acima citadas se conectam com uma área motora no cérebro chamada de *gânglios da base*, provocando a liberação de dopamina em uma subárea chamada de *tegmental ventral* (p. 30)<sup>21</sup>. A partir deste ponto, um mecanismo de auto-regulação é estabelecido no tocante à avaliação positiva ou negativa do estímulo/insumo, sempre com base na memória emocional e de valores atribuída a certos objetivos e tipos de experiência. A dopamina liberada serve para “comunicar” ao organismo se o estímulo encontrado é bom ou ruim (i.e., o grau de recompensa baseado na quantidade de dopamina liberada).

Para Schumann & Wood (2004:39-40), este mesmo mecanismo de recompensa subjacente à motivação (que, como vimos acima, é a gênese do estabelecimento de memórias e consolidação da aprendizagem do tipo APS) também opera na dependência de substâncias como álcool, cocaína, anfetaminas, heroína e nicotina. Ou seja, a motivação, como os vícios da vida, resume-se a certos padrões de avaliação de estímulo. Tais substâncias “seqüestram” ou “cooptam” nosso mecanismo natural de avaliação de estímulos (um mecanismo prazeroso e movido à dopamina), usando-o como fonte de prazer não-funcional.

Estes circuitos que produzem a motivação/incentivo possibilitam a geração de atividade motora e cerebral ao ativarem outros circuitos envolvendo o córtex pré-frontal, o núcleo *accubens*, o *pallidum ventral dorsolateral*, o núcleo *pedunculopontine*, a *região locomotora medular medioventral* e o *núcleo reticular gigantocelular*. Como destacam Schumann e Wood:

É por meio deste sistema que o aprendiz se lança à atividade motora e mental. Ele(a) age no seu meio, em termos motores, com o objetivo principal de aprender a língua. O aprendiz pode então ir à aula, estudar vocabulário, fazer exercícios de gramática, ouvir gravações, conversar com falantes nativos, memorizar diálogos ou qualquer outra coisa que

---

<sup>21</sup> Veja mais sobre a função dos neurônios dopaminérgicos em Gregory, 1998, p.199-200

ele(a) entenda como facilitadora da aprendizagem. Essencialmente, é este o sistema subjacente aos comportamentos que são descritos na literatura que trata das estratégias de aprendizagem (p.31).<sup>22</sup>

Finalmente, o constante reforço (avaliação positiva do estímulo) provoca a liberação de um neurotransmissor chamado acetilcolina no córtex cerebral, cuja função é a de facilitar formações sinápticas e seu fortalecimento nos mecanismos de memória (i.e., hipocampo e gânglios da base). Tal discussão sobre os mecanismos de consolidação da memória será feita, inicialmente, no item a seguir, ao tratarmos dos mecanismos cognitivos e neurobiológicos envolvidos nos processos de atenção, estendendo-se até o item 1.3.2, dedicado especificamente aos mecanismos da memória procedimental. Tais processos possibilitam a transformação do insumo em *intake*, promovendo o estabelecimento de memórias e a consolidação da aprendizagem automatizada da língua (ou aprendizagem do tipo APS, acima discutida).

## 1.2 A atenção

Tradicionalmente, durante a aprendizagem de uma L2, acredita-se que provavelmente o insumo lingüístico passe por três etapas até que seja finalmente incorporado à interlíngua: 1) um estágio de percepção primeira do insumo/estruturas (onde fatores como frequência, saliência, conhecimento prévio, etc, são decisivos); 2) em seguida, este insumo promove comparações intra-sistêmicas (onde “lacunas” são identificadas na interlíngua); e 3) culminando, finalmente, na integração destas novas estruturas à interlíngua do aprendiz, algo provavelmente análogo à criação e reestruturação de esquemas, como proposto inicialmente por Bartlett (1932). Esta visão, prevalente na literatura que trata dos supostos processos envolvidos na aquisição de L2, como observamos em Gass (1997), Rod Ellis, (1997) e McLaughlin (1987), parece ser neurobiologicamente bastante plausível, como comentaremos brevemente neste item.

---

<sup>22</sup> Minha tradução de: *It is through this system the learner engages in motor and mental foraging. He or she motorically moves through the environment to do things to acquire the language. The learner may go to class, study vocabulary, do grammar exercises, listen to tapes, talk to native speakers, memorize dialogues or any number of things that she believes will facilitate learning the language. Essentially this system subserves the behaviors that have been described in the learning strategies literature.*

Na área de ASL, Richard Schmidt é um dos teóricos que mais investigam o conceito psicológico da atenção e seu papel em processos de aprendizagem de uma língua. Na definição de Schmidt (2001), o construto ‘atenção’ refere-se a uma variedade de mecanismos cognitivos (*alertness*, orientação, registro pré-consciente ou detecção inconsciente, facilitação e inibição). No construto ‘atenção’ proposto por Schmidt os conceitos *awareness* e *noticing* são importantes. O primeiro refere-se ao processo de tomar consciência de algo e resulta da alocação de atenção a certas formas da L2. No entanto, é possível haver atenção sem que haja *awareness*, o que Schmidt chama de *noticing*. Esta última forma de atenção, portanto, pode ocorrer de forma inconsciente, em processos implícitos. Schmidt destaca que processos de alocação de atenção parecem ser essenciais para a aprendizagem de uma língua e, em tais processos, variáveis importantes como saliência perceptual, frequência e co-ocorrência do insumo, etc. teriam um papel determinante. Ainda, seria também devido a processos que envolvem atenção ao insumo que “lacunas” poderiam ser identificadas na interlíngua do aprendiz.

Como destacam Schmidt (2001) e Ellis (2001), é provável que os processos de atenção subjacentes aos vários estágios de processamento da linguagem (fonológico, morfológico, sintático, pragmático) ocorram em níveis nos quais torna-se impossível haver consciência (i.e., em níveis cognitivos muito baixos, cuja natureza parece ser essencialmente probabilística). Assim, apenas quando o processamento atinge níveis mais robustos em cada nível (ou seja, na forma de palavras, sentenças ou partes de sentenças) é que a atenção (*awareness*) seria possível. Portanto, é importante distinguir atenção da perspectiva do sistema cognitivo (interlíngua) da atenção por parte do aprendiz às formas externas de sua interlíngua. Há níveis de atenção e consciência.

A atenção também é um conceito central na discussão de processos cognitivos controlados e automáticos (explícitos e implícitos). Os processos de atenção parecem “determinar o curso de desenvolvimento, a fluência e a variação em L2” (SCHMIDT, 2001:11), na medida em que influem e determinam os limites de capacidade de processamento da memória humana na busca de controle e seleção de estruturas específicas no uso de uma língua. Para Ellis (2005) é a consciência (ativada por

processos de atenção) que otimiza a aprendizagem. Ou seja, os processos de aprendizagem são iniciados por processos de atenção, que, por sua vez, conduz à consciência e que, inevitavelmente, promoverá todos os tipos de associação possíveis entre nosso consciente e o nosso inconsciente. A Metáfora do Teatro descrita por Ellis (2005:312-313) é bastante útil na compreensão de tais processos: todo o palco do teatro equivaleria à nossa memória de trabalho, sendo a consciência focal (ou o foco momentâneo de atenção) a luz do holofote no palco. No entorno do foco de luz, fora dos limites da atenção ou consciência momentânea, localiza-se uma área de penumbra, onde fazemos associações vagamente conscientes com outros eventos de nossa memória (um estado de percepção mais fraco). O foco de luz no palco irradia (distribui) suas informações para toda a audiência no teatro (i.e., áreas inconscientes da mente/cérebro). Ellis (2005) resume tais processos implícitos e explícitos de atenção ilustrados na metáfora teatral da seguinte forma

Um teatro combina eventos bastante limitados que acontecem no palco com uma vasta audiência, da mesma forma que a consciência envolve informações limitadas que permitem o acesso a um vasto número de fontes de conhecimento inconscientes. A consciência é o órgão de publicidade do cérebro. (p.312)<sup>23</sup>

A tomada de consciência é, portanto, muito importante na consolidação da aprendizagem. Como destaca Baars (1997:304, citado em ELLIS, 2005), ela seria o *solvente universal da mente*.

Tais questões também já fazem parte de teorizações sobre a neurobiologia da aprendizagem, pois os mecanismos de atenção parecem ser fundamentais para o estabelecimento, consolidação e atualização de nossos mecanismos de memória. Os processos de atenção filtram e classificam o insumo recebido de acordo com nossos esquemas na memória de longo prazo (ELLIS, 2002b). Portanto, atenção, percepção, cognição e ação parecem ser processos contínuos e inseparáveis.

---

<sup>23</sup> Minha tradução de: *A theater combines very limited events taking place on stage with a vast audience, just as consciousness involves limited information that creates access to a vast number of uncounscious sources of knowledge. Consciousness is the publicity organ of the brain.*

De acordo com Schuchert (2004), a atenção é um processo que envolve várias partes do cérebro e funções; também, de uma perspectiva neurobiológica, fica mais clara a noção de gradientes de atenção e a idéia de uma competição neural envolvida em processos de atenção. Os mecanismos neurobiológicos da atenção descritos por Schuchert assemelham-se aos processos de competição que vemos no modelo de competição proposto por MacWhinney (2001, 2005). Ou seja, a atenção, de uma perspectiva neurobiológica, resulta de processos cíclicos presentes no constante intercâmbio de informações entre o hipocampo e o neocórtex. A formação de memórias (explícita e implícita) inicia-se com a informação que chega ao hipocampo oriunda de áreas do córtex cerebral responsáveis pela percepção. De acordo com Jones (2004) e O'Reilly & Norman (2002), estas informações são transmitidas e retransmitidas para as várias camadas celulares presentes no neocórtex. Neste processo há competição, pois as arquiteturas neurais presentes no hipocampo e no neocórtex possuem naturezas distintas: a primeira processa itens arbitrariamente relacionados, ao passo que a segunda processa itens associados por meio de algum padrão de similaridade. Ou seja, os mecanismos presentes no hipocampo e no neocórtex, embora sejam complementares e altamente interativos, geram representações distintas, como apontam O'Reilly & Norman (2002:505-506).

Tais processos são influenciados por fatores neurobiológicos e contextuais (relativos ao meio). Alguns fatores neurobiológicos são ontogeneticamente pré-determinados, tais como o tamanho e a qualidade dos campos receptivos neurais (ou seja, características micro estruturais estabelecidas ao longo do desenvolvimento molecular do indivíduo). No entanto, destaca Schuchert (2004:173), os processos neurobiológicos de alocação de atenção são em grande parte determinados pelo meio e pela nossa história pessoal

Portanto, como vimos acima, tanto o meio como nossas experiências com uma determinada língua são fundamentais no processo de aprendizagem. A relação entre o insumo recebido e sua incorporação ao sistema lingüístico do aprendiz (i.e., a transformação deste insumo em *intake*) não é direta, linear. Sabemos que o simples uso da L2 (seja ele na produção ou na compreensão) não implica, necessariamente, em alterações no sistema lingüístico dos aprendizes, como evidenciado no clássico estudo sobre o inglês usado pelo japonês Wes, largamente citado na literatura de ASL (cf. ROD



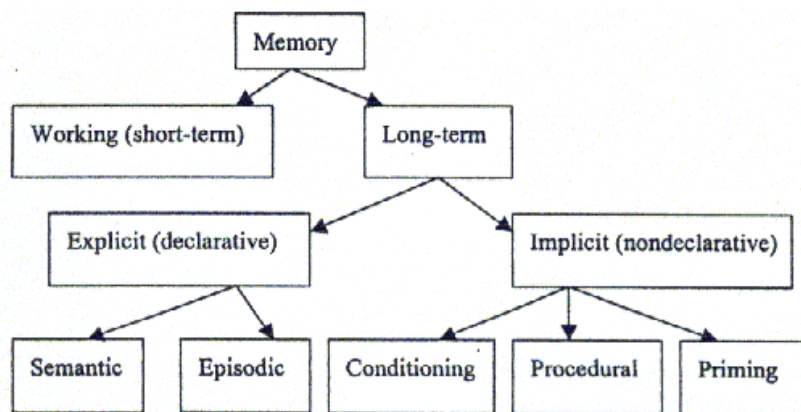
ELLIS, 1997; GASS, 1997). É aqui sustentado, entretanto, que a atenção à forma lingüística é o que provoca, por tornar o insumo saliente, a aprendizagem (i.e., a percepção provoca processos metalingüísticos que desencadeiam, eventualmente, a reestruturação lingüística). Tal percepção da forma lingüística é, por sua vez, consequência direta da repetição/frequência de rotinas lingüísticas ou estruturas convencionadas e entrincheiradas (principalmente no caso de aquisição ‘natural’ de uma L2, como no estudo longitudinal relatado no capítulo 5 mais adiante).

A repetição de rotinas é, assim, fator decisivo para tornar o insumo lingüístico saliente, sendo tal saliência responsável pela percepção da forma lingüística por parte do aprendiz. Ou seja, um processo de automatização que se inicia no uso explícito de fórmulas lingüísticas (nos mecanismos do hipocampo) e, gradativamente, a partir de processos de categorização, vai sendo incorporado ao mecanismo implícito de memória (sistemas do neocórtex).

O desenvolvimento de categorias (na memória implícita) propiciará, por sua vez, processos cada vez mais automatizados. Para Jones (2004:116) e Ellis (2002b:303) a atenção é importante para o registro inicial de traços de memória. Dela depende o estabelecimento de memórias explícitas, como as fórmulas ou *chunks*. Posteriormente, como destacado acima, esta base de conhecimento explícita (calcada em *chunks* ou fórmulas) promoverá a criação de uma base de conhecimento implícita (calcada em padrões e categorias) nos circuitos do neocórtex (i.e., gânglios da base). Este parece ser o caminho neurobiológico dos mecanismos de memória subjacentes às teorizações de Tomasello (2000, 2003) e Ellis (2001, 2002a, 2002b, 2003 e 2004) discutidas ao longo desta tese.

### 1.3 A memória

A Figura 1, reproduzida de Schumann (2004), exibe uma taxonomia padrão dos sistemas de memória<sup>24</sup>. A taxonomia abaixo permite uma descrição da memória tanto em termos temporais como funcionais. Na dimensão temporal podemos subdividir a memória em ‘memória de trabalho/curto prazo’ (MT) e ‘memória de longo prazo’ (MLP). A seguir comento sobre cada um destes dois tipos de memória.



**Figura 1**  
 Classificação hierárquica da memória  
 (adaptação baseada em Fabbro, 1999, In: John Schumann, 2004, p.4)<sup>25</sup>.

#### 1.3.1 – Memória de trabalho e memória de longo prazo.

Para Schumann (2004:4) a memória de trabalho (MT) pode ser definida como uma memória que fica ativa por um período de tempo muito curto (menos de 20 segundos)

---

<sup>24</sup> Acredito não caber, nesta tese, uma discussão detalhada sobre aspectos neuroanatômicos e neurofisiológicos envolvidos nos mecanismos de memória procedimental e declarativa. A terminologia empregada é altamente especializada e requer treinamento e conhecimento profundos em neurologia e neurociências. Entretanto, fica a cargo do leitor, caso tenha interesse numa descrição técnica e mais detalhada de tais processos, consultar a literatura aqui citada de forma resumida e simplificada.

<sup>25</sup> Embora contrariando as normas da ABNT, algumas figuras ao longo desta tese foram mantidas no seu formato original em inglês devido às dificuldades encontradas pelo autor no manejo gráfico destas.

durante a realização de uma tarefa. Ellis (2001:35), por sua vez, define a MT como um sistema especializado em perceber e representar, tanto temporariamente como a longo prazo, informações visuais e auditivas em processos de atenção limitados. Portanto, na definição de Ellis, a memória de trabalho também parece ser, grosso modo, um mecanismo de atenção por excelência (como vimos anteriormente).

A memória de longo prazo (MLP), por sua vez, é bastante duradoura, podendo variar de dias, semanas, meses ou anos. De acordo com a taxonomia da figura 1 acima, a MLP também pode ser descrita em termos funcionais: declarativa (explícita) e não-declarativa (implícita). A memória declarativa registra fatos e eventos, sendo subdividida em memória semântica ou enciclopédica e memória episódica (registros espaciais e temporais específicos). A memória declarativa é processada basicamente pelos mecanismos presentes no hipocampo, como destacado por Crowell (2004).

A memória implícita diz respeito a nossos hábitos, habilidades motoras e perceptivas, bem como à aprendizagem emotiva (SCHUMANN, 2004:5). Este tipo de memória pode ser subdividido em condicionamento, memória procedimental e *priming*. A memória implícita é processada pelos mecanismos do neocórtex. No entanto, a memória procedimental, que, como veremos no próximo item, é fundamental para a discussão sobre mecanismos de *chunking*, é processada, de acordo com Lee (2004), principalmente nos gânglios da base, um subsistema do neocórtex.

Finalmente, como vimos anteriormente, os mecanismos de memória explícita e implícita possuem uma natureza distinta. Como destaca Schumann (2004:5), a memória implícita, que precede a memória explícita tanto em termos filogenéticos como ontogenéticos, é relativamente inflexível e apenas disponível em contextos idênticos ou semelhantes a situação original de aprendizagem. Isto, por sua vez, a torna uma memória mais ‘robusta’ (e, certamente, é por isto que a memória implícita tende a ser preservada na velhice, diferentemente da memória explícita, que se deteriora de maneira acentuada com o passar dos nossos anos).

### 1.3.2 - Memória procedimental: processos automáticos

Automatismo é o nome popular para o que entendemos por memória procedimental. Este é um tipo de memória adquirida pela repetição de uma tarefa (por ex., jogar capoeira, tocar violão ou piano, falar uma língua, etc.), um conhecimento implícito, automático e inconsciente. É nos *gânglios da base* que concentram-se as nossas memórias implícitas; lá elas são apreendidas, reforçadas e instanciadas. Tal processo de automatização pode ser observado no nível neurobiológico em processos de convergência, divergência e reconvergência celular (neuronal). De acordo com Lee (2004), estudos mostram que o insumo (informação) proveniente de múltiplas áreas do neocórtex converge para os neurônios dos gânglios da base (mais especificamente para uma área conhecida por *striatum*) de forma seletiva e precisa. Por exemplo, segundo Lee (2004:49) é comum que “cerca de 10.000 neurônios do neocórtex disparem para um único neurônio do *striatum*”. Ou seja, dez mil pedaços de informações variadas são fundidos e sintetizados em um único neurônio. Este processo de *chunking* celular observado no mecanismo de convergência, divergência e reconvergência produz, gradativamente, “códigos para seqüências de movimento com ordens temporal e espacial específicas” (p.49). Ou, como destaca Lee,

Esta harmonização gradual de populações especializadas de neurônios no *striatum*, a fusão espaciotemporal dos neurônios do *striatum* por meio de TANs (neurônios tonicamente ativados) e a reconvergência da informação em alvos na saída do *striatum* são os mecanismos que produzem os processos de *chunking* (Graybiel, 1998). Por meio de processos de *chunking*, seqüências de ações motora e cognitiva são estabelecidas como rotinas que podem ser, subseqüentemente, executadas como unidades de desempenho. Tal processo de *chunking* é um mecanismo para a aprendizagem e produção de repertórios de ações motora e cognitiva (p.49).<sup>26</sup>

---

26 Minha tradução de: *This gradual tuning of modular populations of neurons in the striatum, the spatio temporal binding of the striatal neurons by TANs, and the reconvergence of the information onto striatal output targets are the mechanisms that bring about chunking (Graybiel, 1998) Through chunking, motor and cognitive action sequences are formed as routines that can be subsequently executed as performance units. This chunking process is a mechanism for the learning and expression of motor and cognitive action repertoires.*

Como vimos na discussão sobre motivação, todo o processo de *chunking* acima descrito é modulado pelo nosso sistema dopaminérgico. Também, tais representações mais complexas (i.e., representações seqüenciais resultantes de processos de *chunking*) serão posteriormente selecionadas e produzidas pelos circuitos de saída dos gânglios da base. Assim, podemos entender o comportamento automatizado (seja ele lingüístico, artístico, etc.) como uma seleção de representações seqüenciais a serem fundidas em *chunks* maiores; visto desta forma, o comportamento é em si um grande *chunk* petrificado e difícil de ser quebrado e/ou modificado. Uma engenharia reversa ou a decomposição do comportamento em seus *chunks* constitutivos tornam-se algo difícil.

Este, portanto, parece ser o mecanismo neurobiológico subjacente à aprendizagem do tipo APS, como destacado anteriormente. Aqui vemos claramente a relação entre motivação e aprendizagem: em termos neurobiológicos, os processos que ocorrem nos gânglios da base acima descritos repercutem na amígdala, e as interconexões presentes nestas duas áreas passam a ser reguladas pelo mecanismo de recompensa mencionado anteriormente (i.e., ativam ou inibem as representações formadas nos gânglios da base).

Segundo Lee (2004:60-63), o papel dos gânglios da base na memória procedimental é corroborado por estudos sobre automatismo, assim como estudos sobre patologias como coréia de Hungtinton, doença de Parkinson, transtorno obsessivo compulsivo (TOC) e afasias, como veremos a seguir.

Evidências fortes do papel dos gânglios da base na aprendizagem de habilidades são encontradas nos casos de afasia, principalmente a afasia paradoxal. Recentemente tivemos um caso conhecido deste tipo de afasia que foi bastante veiculado na imprensa brasileira. Após um acidente, o cantor Herbert Viana, da banda Paralamas do Sucesso, sofreu lesões na região dos gânglios da base. No hospital, ao retomar a consciência, o cantor só comunicava-se em língua inglesa. Como o inglês é sua segunda língua, provavelmente partes de seu processamento ainda não haviam sido tão fortemente automatizadas, sendo utilizados circuitos envolvidos nos mecanismos de memória explícita (como os circuitos envolvendo o hipocampo). Por outro lado, sua língua materna, o português, provavelmente encontrava-se toda ela automatizada nos gânglios

de base. Nestes casos, devido a lesões naquela região, apenas a memória parcialmente automatizada poderia ser acessada, daí o fenômeno da segunda língua vir à tona.

O capítulo seguinte discute modelos conexionistas (e probabilísticos) que são inspirados na neurofisiologia discutida acima. Devido a tal plausibilidade neurobiológica, acredito que sejam bastante adequados para o desenvolvimento da nossa discussão. Iremos agora para o nível conhecido como simbólico-representacional do processo de construção deste domínio cognitivo específico que é a linguagem.

## 2. Conexionismo, probabilidades & *chunking* no processo de aprendizagem de línguas

*Language learning is the associative learning of representations that reflect the probabilities of occurrence of form-function mappings. Frequency is thus a key determinant of acquisition because 'rules' of language, at all levels of analysis from phonology, through syntax, to discourse, are structural regularities that emerge from learners' lifetime analysis of the distributional characteristics of the language input. Learners have to FIGURE language out.*

Nick C Ellis

Se optássemos por um arcabouço teórico baseado em modelos simbólicos clássicos de processamento da linguagem certamente nossa concepção de representação, ou memória, seria estática e localizada (no sentido de possuir um "domicílio" específico e localizável). Ou seja, uma palavra, um conceito, ou um dado conhecimento estariam estocados, armazenados, em um ponto específico de nosso cérebro (talvez numa área ou aparato destinados especificamente ao processamento da linguagem), sendo tais informações acessíveis a partir de um endereço também específico. Como sugerem McClelland, Rumelhart & Hinton (1986), o uso da memória nestes termos consistiria simplesmente em transferir, quando fosse necessário, tal representação para a memória de trabalho<sup>27</sup>.

---

<sup>27</sup> É importante aqui uma observação. Diante do que foi discutido no capítulo anterior sobre o desenvolvimento/criação gradual de uma base de conhecimento implícita e automatizada nos sistemas do neocórtex (mais especificamente nos gânglios da base), o leitor poderá achar totalmente contraditória a presente discussão conexionista de processamento em paralelo e representação distribuída. De fato, o que vimos anteriormente pode ser entendido como o desenvolvimento de um *módulo especializado*. Entretanto, é fundamental frisar, ser modular não implica, necessariamente, ser inato. A questão principal é o tipo de especialização ou modularidade que temos em mente, i.e., se é algo *pré-existente* ou *emergente*. O primeiro tipo reflete a conceituação original proposta por Fodor (e Chomsky), neste projeto criticada. Uma certa especialização/modularização gradativa (emergente) de domínios cognitivos, como observamos tanto na discussão sobre a neurobiologia da aprendizagem como em modelos conexionistas para o processamento de sistemas estruturados, feita em Elman et.al. (1998), por outro lado, é totalmente plausível. Ou seja, há uma diferença importante entre *tornar-se* modular e *ser* modular.

Por outro lado, entretanto, podemos especular que palavras e conhecimentos em geral não possuem endereços específicos, ou seja, não estejam estocados em um local específico (com endereço próprio) na memória, mas encontram-se sim espalhados, distribuídos, ao longo de todo nosso sistema cognitivo. Estas informações ou conhecimentos seriam, como conseqüência de sua natureza fragmentada, constantemente instanciados e reformulados de acordo com os usos e os contextos específicos que encontramos em nossa utilização da linguagem (i.e., instanciados na forma de palavras, expressões lexicais, conceitos, etc.). O ambiente (ou o insumo lingüístico), ou ainda a experiência com o meio, tornar-se-ão, portanto, fundamentais e determinantes para a estruturação e reestruturação do sistema lingüístico, uma vez que a criação e o funcionamento deste sistema dependerão de um processo indutivo de base probabilístico-associativa.

Este último ponto, a aprendizagem ou o estabelecimento de padrões de (co)ocorrência de elementos a partir de um mecanismo de indução, no sentido de aprendizagem implícita, nos termos colocados por Winter & Reber (1994), é bastante controverso na área de aquisição da linguagem (na verdade, não é apenas controverso do ponto de vista psicolingüístico, mais sim epistemológico, ou seja, o velho nó filosófico que opõe o pensamento empirista ao racionalista), pois foi a partir da negação de um sistema de indução de regras que o argumento da “pobreza de estímulo” foi inicialmente proposto na teoria inatista da linguagem de Noam Chomsky, conforme discutido em Chomsky (1998, 1980), Pinker (1994), Harley (1998) e Winter & Reber (1994).

É importante ressaltar, no entanto, que a perspectiva conexionista é, em sua essência, também inatista. No entanto, há duas possibilidades de concebermos tal qualidade ‘inata’ da linguagem: como um produto ou como um processo. Se entendemos inato como produto, aceitamos a existência, desde a nossa concepção, de uma estrutura lingüística pré-determinada, uma “gramática” da linguagem pronta. Seria o equivalente mental de uma gramática normativa, pronta, por assim dizer, para ser lançada no mercado. Ou seja, um produto pronto para ser usado ou ainda um software pronto para ser rodado. Por outro lado, se entendemos que o que é inato na linguagem refere-se ao processo, a tendência é levarmos em conta os estados mentais, buscando uma



explicação dos processos cognitivos e neurobiológicos envolvidos em seu desenvolvimento.

O que parece haver de inato é uma capacidade estrondosa de aprendizagem baseada em processos holísticos de associação e estabelecimento de padrões por parte de nosso sistema cognitivo (como vimos no capítulo sobre a neurobiologia da aprendizagem). Ou seja, numa visão conexionista, apenas os mecanismos responsáveis pelo funcionamento do sistema cognitivo são inatos, i.e., os estados neurobiológicos e neurofisiológicos, estando o sistema cognitivo, por sua vez, subjacente a qualquer forma de aprendizado. Neste sentido há uma grande diferença entre a concepção conexionista de desenvolvimento cognitivo da linguagem e a clássica idéia de desenvolvimento a partir de uma *tabula rasa*.

Nas duas últimas décadas, o número de pesquisas que evidenciam a possibilidade da aprendizagem a partir de modelos de indução de padrões ou regras (como o modelo PDP de aprendizagem do passado dos verbos na língua inglesa, largamente citado na literatura) tem possibilitado o alargamento das teorizações sobre os processos cognitivos envolvidos na aquisição da linguagem, como vemos nas discussões em Elman, Hare & McRae (2005), Elman (2004), Elman et. al, (1998), Harley (1998), McClelland, Rumelhart & Hinton (1986), Winter & Reber (1994), Broeder & Plunkett (1994) e Gomez & Gerken (1999, 2000).

Abrem-se, portanto, novas possibilidades de reflexão sobre a aprendizagem de línguas (L1 e L2), baseadas não em regras lingüísticas inatas e preestabelecidas, mas sim resultantes de processos cognitivos indutivos e gerais (holísticos), passíveis de verificações empíricas, em modelos inspirados na neurofisiologia e biologicamente bastante plausíveis. Ou seja, a maneira pela qual concebemos a natureza de nosso sistema lingüístico poderá determinar decisivamente o nosso entendimento sobre como se dá o aprendizado de uma L2.

Como comentado anteriormente, os modelos PDP de aprendizagem e processamento da linguagem visam à criação de redes estruturadas a partir de mecanismos de indução de padrões - baseados, essencialmente, em regras probabilísticas (estatísticas) e/ou leis da

dinâmica, como aponta Jordan (1999). Em linhas gerais, tais modelos têm uma arquitetura semelhante àquela verificada na neurofisiologia. Ou seja, há unidades e conexões que são a base (estrutura) de processamento destas redes. As redes são, portanto, configurações específicas resultantes das interconexões entre as suas unidades.

O sistema lingüístico (ou a gramática cognitiva da linguagem) é criado a partir deste intrincado sistema de conexões de uma rede conexionista (e, importante destacar, não apenas a partir de conexões entre unidades que contenham informações puramente lingüísticas, mas com base na utilização de informações do sistema cognitivo como um todo). Estas redes neurais se encarregariam de estabelecer as relações de informação através das conexões entre as unidades. As unidades, em suas conexões, podem ser ativadas ou inibidas. A ativação ou inibição de uma unidade dependerá das associações percebidas no insumo (lingüístico, visual, etc.). Bialystok (1994) descreve da seguinte forma o princípio básico de modelos PDP:

Tal princípio baseia-se em associações detectadas no insumo lingüístico... As redes subjacentes são criadas a partir da aprendizagem das frequências distribucionais de formas específicas na língua. Tais redes são a gramática. A aquisição da linguagem, nesta perspectiva, é uma 'análise distribucional guiada por pistas'. O fator principal na aquisição da língua materna pelas crianças, na aquisição de segunda língua por parte de adultos, ou ainda na aquisição da linguagem pelos computadores, diz respeito ao fato de que o insumo fornece informação que é transformada em sistemas associativos (p.124-125).<sup>28</sup>

A seguir, é feita uma síntese sobre o funcionamento de redes conexionistas para o processamento e representação do conhecimento. A discussão baseia-se, principalmente, em Elman et al (1998), Elman (2004) e Ferreira Junior (2005).

---

<sup>28</sup> Minha tradução de: *It is based on associations perceived in linguistic input...The underlying networks are built up through the perception of the distributional frequencies of particular forms in the language. Those networks are the grammar. Language acquisition, on this view, is a 'cue-driven distributional analysis'. The main factor in first language acquisition by children, second language acquisition by adults, or language acquisition by computers, is that the input delivers information that is translated into associative systems.*

## 2.1. Desenvolvimento e processamento do conhecimento da perspectiva conexionista.

Uma rede conexionista é formada, basicamente, por unidades (numa analogia com nossos neurônios) interligadas umas as outras através de conexões (analogia com sinapses), um enfoque, portanto, fortemente apoiado na neurofisiologia humana. O processamento desta rede é totalmente distribuído e em paralelo (daí o termo sinônimo de rede PDP, como comentado na introdução deste manuscrito). Modelos conexionistas, com uma arquitetura interna como a descrita acima, são construídos na forma de programas de computador. Uma vez em funcionamento, o sistema estabelece processos de regulação entre as unidades que compõem a rede (na forma de ajuste das forças/pesos de conexão) provocando, assim, a estruturação do sistema. O 'conhecimento' da rede está justamente nas forças/pesos definidos localmente entre as unidades (i.e, através da constante ativação ou inibição destas). Assim, surge um desempenho baseado em padrões de ativação entre unidades da rede, padrões estes que estabelecem configurações específicas a cada ativação.

Basicamente, há três tipos de unidades em uma rede: as unidades de insumo, as unidades ocultas e as unidades de saída. É na camada oculta que as representações são criadas e o conhecimento é gerado. A rede estabelece associações através da regulação entre o seu peso/forma atual e o peso/forma ideal, este último informado, inicialmente, pela experiência com o meio. Os pesos entre as unidades (e a conseqüente estabilização da rede, ou a aprendizagem propriamente dita) em um modelo conexionista são autoreguláveis. Assim a rede vai, aos poucos, autoregulando-se e refinando positivamente o limiar de ativação (em outras palavras, *refinar* significa corrigir a diferença (o erro) entre produção atual e produção ideal); visto desta maneira, o erro (e a sua constante minimização) torna-se *fundamental* para a rede atingir a produção ideal. A rede atinge a generalização de padrões a partir da exposição gradual ao padrão ideal desejado. A cada produção da rede é feita uma associação com o padrão ideal. O cálculo da diferença entre os pesos refaz o caminho de forma inversa. Esta *retropropagação* faz com que o *cálculo* retorne para as unidades ocultas, de forma que haja uma regulação e um reajuste visando às futuras ativações (este é um processo de constante minimização do erro por parte da rede). Desta forma, a rede aprende a aprender, através da

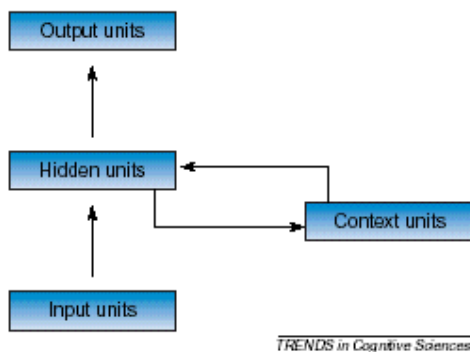
diminuição da sensibilidade ao erro (i.e., torna-se um sistema mais tolerante e menos maleável).

Há uma questão importante nesta discussão. Como destacam Elman et al (1998), as associações de padrões idênticos nos primeiros modelos conexionistas (tipo o Perceptron) só podiam ocorrer em um campo espacial linear, ou bidimensional. Isto fazia com que a rede ignorasse o fato de que, às vezes, formas diferentes podem revelar padrões de comportamento idênticos (e vice-versa). Modelos como o Perceptron, como explicitado a seguir, tinham apenas duas camadas (níveis) de unidades, ou seja, unidades de entrada e unidades de saída, e isto tornava o comportamento da rede algo parecido com o estímulo/resposta do behaviorismo. Em um modelo Perceptron as associações são *puramente mecânicas*. É necessário, portanto, um nível de processamento que gere *representações abstratas*.

Para tentar resolver tal problema (i.e., não cair em um mero 'associacionismo computacional'), Elman et al (1998) e Elman (2004) propõem uma rede do tipo SRN (*simple recurrent network*) que possui, além de uma camada de unidades intermediária (oculta) responsável pelo aspecto "abstração" ou "representações abstratas" numa rede, uma camada ou unidades de *contexto*, adicionando, assim, o aspecto *tempo* ou uma *memória* à rede (devido à recursividade).

Assim, nas redes recorrentes as unidades são informadas não apenas pelo insumo externo a elas, mas também por ativações prévias da própria unidade que ficam "registradas" nas unidades de contexto, aumentando o poder de autoregulação. E isto é uma importante contribuição conexionista para a discussão sobre *autonomia*. A rede produz, assim, insumo que, por sua vez, regulará a própria rede. Também, inerente a uma rede recorrente, é a sua capacidade de 'inferir'.

Este aspecto, consoante com a aprendizagem em vários domínios, reflete nossa capacidade de fazer previsões (certamente consequência direta do estabelecimento de protótipos ou configurações padrão). A Figura 2, reproduzida de Elman (2004:302), exibe uma arquitetura conexionista do tipo SRN:



**Figura 2** – Rede recorrente simples (SRN) com unidades de contexto. As camadas são compostas de uma ou mais unidades. A informação entra pelas camadas de insumo, passa pelas camadas ocultas e chega até as camadas de saída. Além disso, a cada passagem de tempo  $t$ , a camada de unidades ocultas recebe insumo da camada de contexto, que armazena as ativações da camada oculta referentes ao momento  $t-1$ . (Reproduzida de Jeffrey Elman, 2004, An alternative view of the mental lexicon. *Trends in Cognitive Science*, v.8, n.7, p.302).

Processos de associação de padrões numa rede conexionista do tipo SRN tornam-se mais dinâmicos, pois o que é "similar" não mais será definido em um ambiente bidimensional (baseado puramente na *forma*), mas multidimensional (sendo consideradas também relações *espaciais*). Desta maneira, aumenta-se a capacidade de representação e, conseqüentemente, de associação. Podemos dizer que tais redes também exibem um comportamento *inteligente*, pois possuem níveis diferentes de abstração/representação do conhecimento gerado. A rede funciona sempre levando em conta restrições das mais variadas ordens, sejam restrições presentes tanto no insumo ao qual ela está constantemente exposta como nas próprias restrições sistêmicas por ela criadas ao longo de seu desenvolvimento (ou seja, a sua língua interna ou 'interlíngua'). É a partir destas restrições que a rede emerge ou, em outros termos, o sistema é construído. É a partir de processos de restrições de restrições que se dá a regulação do sistema, levando em última instância a uma autoregulação (autonomia) e uma certa consolidação da aprendizagem.

Para concluir, é importante destacar que em uma rede conexionista o importante são as representações locais, i.e., as conexões entre as unidades. É a partir da força inicialmente estabelecida localmente que se dá um processo de equilíbrio da rede. E é a partir do equilíbrio da rede que surgem as representações distribuídas. Ou seja,

representações locais geram representações emergentes (ou ativas) e distribuídas, determinadas sempre, por sua vez, por aquelas mesmas configurações localmente e constantemente (re)definidas através de pesos (forças) entre as conexões.

Tais redes, através de mecanismos de indução e generalização de padrões, conforme a explicação dada acima, passam a comportar-se como se estivessem seguindo regras. Na verdade, o que acontece é que estas redes passam a funcionar com base em configurações estabelecidas nos pesos (níveis de ativação ou inibição) entre as suas conexões. Daí resulta a capacidade de aprendizagem. A aprendizagem é, assim, vista como uma (re)configuração de padrões previamente estabelecidos. Da verificação e confirmação de padrões idênticos de configuração (através da constante ativação e inibição das unidades e suas respectivas conexões), resulta a consolidação da aprendizagem.

A meu ver, o *insight* conexionista encontra ressonância em outras teorizações acerca da aprendizagem, como o modelo proposto por McLaughlin (1987) que sugere que a aprendizagem resulta de processos “controlados” que se tornam “automatizados” (estes últimos resultantes de um mapeamento constante de um determinado insumo a um mesmo padrão de ativação ao longo do tempo), bem como na noção piagetiana de processos cognitivos de “assimilação”, “generalização” e “equilíbrio” em sistemas que tornam-se, gradualmente, autoreguláveis. Esta última noção de autoregulação, através de processos de regulação de regulações, é o que provoca a aprendizagem, i.e., o que leva à construção ou desenvolvimento de novas estruturas cognitivas, como sugeriu Piaget (1980) (embora seja importante lembrar que Piaget pensou estes processos como estanques e lineares, e não algo dinâmico e em paralelo).

## **2.2 Modelos de ativação interativa para a leitura**

Um exemplo bem simples de processamento conexionista, inspirado em McClelland, Rumelhart & Hinton (1986), é o seguinte. Se uma suposta unidade que corresponda à letra “C” na posição inicial de uma palavra for ativado, isso aumentará conseqüentemente o nível de ativação das unidades correspondentes as palavras “COMER”, “CASA”, “CORRER”, etc. (ou seja, todas as palavras que tenham o “C”

como letra inicial), bem como diminuirá o nível de ativação (i.e., inibirá a ativação) de palavras como “MORAR”, “FALAR”, “SALA”, etc., (novamente, todas as demais que não tenham o “C” como letra inicial).

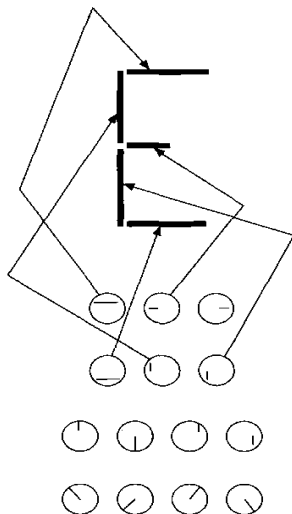
Isto ocorre devido ao fato de as unidades estarem conectadas umas as outras em um mesmo nível através de conexões inibidoras, ou seja, uma vez que uma unidade é ativada ela começa a inibir todas as outras unidades pertencentes àquele nível. Continuando a explicação, se o sistema depara-se com outros tipos de restrições (por exemplo, devido a sua função e contexto específicos, alguma unidade que contenha a informação de que a palavra pertence à categoria ‘verbo’ for ativada) a palavra CASA, anteriormente ativada, sofrerá uma redução em seu nível de ativação devido a esta nova restrição, ao passo que as palavras COMER e CORRER sofrerão um aumento nos seus níveis de conectividade sistêmica<sup>29</sup>.

Na tentativa de ilustrar um pouco mais a discussão e exemplos acima, inspirado num modelo conexionista de leitura chamado de “*Interactive activation model*” para a percepção de traços, letras e palavras (McClelland & Rumelhart, 1981), veremos abaixo duas ilustrações gráficas deste modelo de percepção da linguagem (no nível morfológico) sensível ao (e dependente do) contexto. O modelo é interativo no sentido de utilizar informações ascendentes (*bottom-up*) e descendentes (*top-down*) durante o processo de interpretação, resultante da constante competição de inúmeras **pistas** presentes no insumo (no caso específico deste modelo de leitura, pistas visuais). Como discutido anteriormente, neste modelo a percepção resulta das interações entre ativação e inibição dos detectores de traços visuais, letras, e palavras.

A Figura 3 ilustra a instanciação da letra “E” a partir da identificação de um conjunto de traços (ou seja, a letra *E* é uma configuração emergente baseada num contexto específico)

---

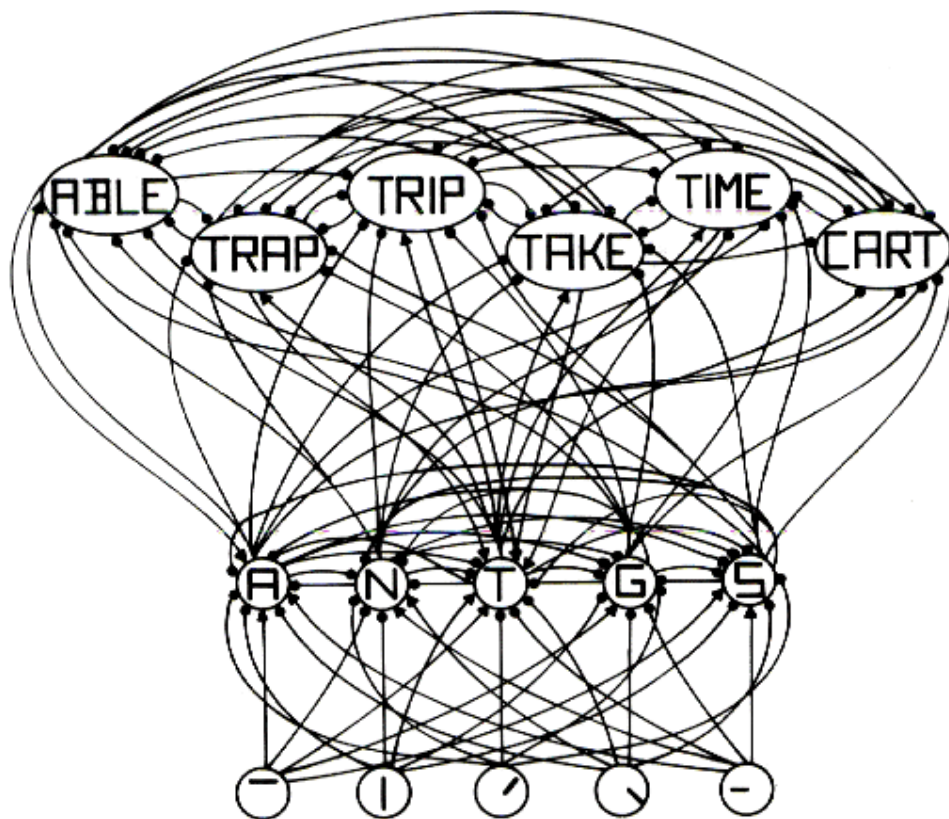
<sup>29</sup> Mais adiante, no capítulo 4, veremos uma proposta recente (Elman, 2004) que redefine de maneira bastante precisa o conceito de palavra de uma perspectiva emergentista.



**Figura 3** - Identificação da letra *E* com base na percepção de um dado conjunto de traços.

A Figura 4 a seguir, que poderíamos entender como um ‘sistema’ mais complexo, em um nível de representação superior, ilustra como um insumo ativa, inicialmente, os detectores de traços visuais. Estes últimos, por sua vez, ativam letras que sejam consistentes com os traços ativados e inibem outras tantas inconsistentes com tal configuração. Num nível de representação mais acima, os detectores de letras ativam o detectores para palavras que sejam consistentes com a configuração instanciada no nível da letra, inibindo também outras tantas palavras inconsistentes com tal instanciação específica no nível da letra. Finalmente, detectores ativados no nível da letra inibem-se mutuamente, retroalimentando também as ativações no nível da letra e dos traços, e assim recalibrando (ou seja, reforçando ou inibindo) as ativações da rede em seus vários níveis de organização. Importante destacar que processos implícitos e explícitos subjazem toda esta discussão. A consciência é uma orquestração de tais processos (veja ELLIS, 2005).





**Figura 4** - Identificação de palavras com base na ativação ou inibição de letras e percepção de um dado conjunto de traços. In: McClelland, Rumelhart & Hinton, 1986, p.22.

Apesar da ilustração e exemplos serem adequados para a nossa discussão no momento, o modelo de rede conexionista dado acima é limitado pelo fato de suas representações serem **localistas**. Suas unidades são dedicadas a um tipo específico de representação (o mesmo problema da teorização simbólica tradicional já criticado aqui). Também, como já discutido anteriormente, em modelos realmente conexionistas as representações são **distribuídas**, ou seja, residem nas forças das próprias conexões da rede. O conhecimento emerge a partir das **pistas** presentes no insumo, que é sempre contextualizado no uso. Esta filosofia (*rationale*) está presente na teorização emergentista do Competition Model de MacWhinney (2001) e em modelos probabilísticos de processamento da cognição e linguagem, como propostos recentemente por Oaksford & Chater (2007), Chater and Manning (2006), Bod, Hay & Jannedy (2003), Christiansen & Chater (2001), e tem como base a aprendizagem

hebbiana discutida nesta tese, mais especificamente no capítulo primeiro dedicado à neurobiologia da aprendizagem.

### **2.3 Modelos de ativação interativa para a percepção da fala**

Um modelo conexionista semelhante, dedicado à percepção da fala, é o modelo TRACE de McClelland & Elman (1985). O modelo TRACE segue os mesmos princípios do modelo de identificação de traços visuais, letras e palavras em processos de leitura, acima discutido. A diferença óbvia é que, neste caso, as camadas constituem-se em traços fonéticos, fonemas e palavras. Outra importante diferença é a presença de uma dimensão temporal para dar conta do aspecto temporal da fala (i.e., o fluxo contínuo da fala, que torna difícil a sua segmentação), problema este que não existe na leitura pelo fato da palavra em um texto ser estática (i.e., apresentar-se por inteiro), como apontado em Christiansen & Chater (2001).

Para inúmeros teóricos discutidos ao longo deste capítulo, os modelos conexionistas e probabilísticos fornecem a estrutura computacional adequada para a testagem (por meio de simulações) das *conjecturas*, para os mais céticos, a) de toda a teorização sobre aquisição da linguagem baseada no uso; b) para a investigação dos processos de instanciação de sistemas emergentes; e c) para lançar luzes sobre como o conhecimento (e toda nossa rede conceitual) pode ser entendido como padrões de ativação convencionalizados e entrincheirados (como defendido ao longo desta tese).

### **2.4 Probabilidades e o constante cálculo de incertezas.**

Os modelos sintático-analíticos de processamento da linguagem não dão conta do uso (*performance*) da língua em situações reais. Obviamente, como explicação de um sistema abstrato e idealizado (i.e., puramente lingüístico), baseado em regras da lógica, especializado e independente, tais modelos cumprem a sua tarefa. No entanto, estas teorizações não parecem levar em conta a realidade dos processos cognitivos globais inerentes à aquisição e o uso de línguas naturais. Tais modelos não se adequam à realidade psicológica (e neurobiológica) evidenciada no uso real da linguagem humana, especialmente no que concerne à capacidade de nossos sistemas de processamento e

representação. Nos sistemas formais, baseados em regras, o processamento vai ficando cada vez mais lento e distante da realidade à medida que o processamento lingüístico fica mais complexo (ou seja, à medida que mais restrições são impostas ao sistema).

Tal constatação decorre do provável fato de que o uso natural, ou automatizado, da língua não seja analítico, mas esteja calcado na memória e em processos probabilísticos de inferência. Utilizamos pedaços (*chunks*) de linguagem na produção lingüística, semelhante ao que ocorre na montagem de um quebra-cabeças. Tais *chunks* são resultado da organização cognitiva de rotinas altamente freqüentes. A idiomaticidade da língua decorre, justamente, da boa utilização destes pedaços lingüísticos convencionalizados. Seu uso implica, acima de tudo, na diminuição da carga de processamento cognitivo. Ou seja, quanto mais o aparato mental se detém em análises lingüísticas, mais lento é o processamento cognitivo da linguagem, ao passo que quanto mais uso se faz de fórmulas lexicais, mais fácil e rápido, portanto automático, torna-se o processamento da língua (esta discussão será retomada mais adiante).

Talvez um dos motivos da discrepância entre os modelos logicistas para o comportamento humano idealizado (historicamente conhecidos na lingüística como *langue* ou competência) e o comportamento mundano humano (*parole* ou desempenho) seja o fato de que nosso sistema cognitivo não seja regido por um sistema lógico. Como destacam Oaksford & Chater (2007), há 2 mil anos e meio o mundo ocidental acredita que a mente (ou nossa cognição) é regida pelos mesmos princípios da lógica. Talvez não sejamos assim tão racionais! Entre o sim e o não existe um vão! Ou talvez seja preciso entender que a razão precisa levar em conta (e seja permanentemente guiada por) um cálculo constante de incertezas, o que não é compatível com os sistemas de lógica tradicionais, baseados no silogismo aristotélico. Nos estudos da linguagem, tal concepção logicista da mente se refletiu mais claramente na sintaxe gerativista e na semântica formal (este último com os conhecidos cursos envolvendo Cálculo de Predicados).

As ‘regras’ gramaticais podem ser apenas *probabilidades lingüísticas de uso*, como colocam com bastante propriedade Chater & Manning (2006). Tais regras envolveriam o que é linguisticamente provável e não apenas o que seria linguisticamente possível.

Por exemplo, a questão da resolução de ambigüidades no *parsing* sintático sempre foi visto na psicolingüística tradicional como sendo algo inerente à *estrutura* lingüística. Veja os exemplos abaixo extraídos de Chater & Manning (2006:336), com suas respectivas traduções em português:

- a) *the girl saw the boy with a telescope* (a garota viu o garoto com um telescópio)
- b) *the girl saw the boy with a book* (a garota viu o garoto com um livro)
- c) *the girl hit the boy with a book* (a garota acertou o garoto com um livro)

O exemplo fornecido em a) contém uma ambigüidade clara que poderia ser facilmente resolvida levando-se em conta outras informações, de natureza não-lingüística (ou seja, aumentando o contexto ao observarmos o acontecimento e as pessoas envolvidas). Mesmo assim, nos limitando somente a esta pobre informação estrutural, em termos probabilísticos pode-se resolver esta ambigüidade levando-se em conta apenas o fator *complexidade estrutural*. A probabilidade de interpretação é maior para o sentido que envolve o número menor de “galhos” na árvore gerativista ou que possua uma “regra sintática” a menos. Assim, para as duas possibilidades de interpretação desta frase descontextualizada, 1) se o telescópio liga-se ao sintagma verbal (ou seja, o ato de *ver* utiliza um telescópio) ou 2) ao sintagma nominal objeto (se o *garoto* carrega um telescópio), a primeira seria probabilisticamente a mais provável por ser menor a distância entre o funtor e seu respectivo argumento (portanto uma relação menos complexa). O’Grady (2003) chama isto de ‘eficiência’, ou seja, um processo cognitivo de redução de armazenamento para lidar com dependências estruturais de várias ordens, como destacado anteriormente nesta tese. A incerteza é algo inerente aos sistemas complexos, e não seria diferente na linguagem. O cálculo de incertezas, portanto, precisa ser probabilístico e calcado tanto em frequência de ocorrência como em eficiência (a famosa ‘lei do menor esforço’). Esta estratégia probabilística de eficiência nada mais é do que um processo de *chunking*.

No exemplo em b) não há, a princípio, ambigüidade, pois livros, diferentemente de telescópios, não são instrumentos usados para avistar algo, embora possamos, metaforicamente, dizer que “podemos ver o mundo com um livro”. A situação torna a se reverter quando mudamos o verbo, no exemplo da frase c), ou seja, um livro pode ser

um instrumento usado para bater em alguém. *Gramáticas lexicalizadas*, que levam em conta a frequência de co-ocorrência de palavras (por exemplo, se na maioria das vezes em que ocorre a ação de *ver* algo o fazemos “com-um-telescópio” ou “com-um-livro”) são uma alternativa bastante promissora na resolução de problemas de *parsing* desta natureza, como sugerem Chater & Manning (2006:337-338). Tal cálculo de contingência, ou probabilidades tradicionais, como vemos no clássico estudo conexionista de Elman sobre identificação de palavras (*Finding structure in time*, 1990), é uma fonte de informação importante nos processos de aprendizagem implícita de uma língua (como veremos mais adiante). Este mesmo raciocínio está presente nas teorizações construcionistas ou modelos de linguagem baseados no uso. Ou seja, uma simples interpretação do sentido mais plausível expresso nas frases em a), b) e c) envolve nosso conhecimento de mundo, a compreensão do meio e contextos sociais, questões pragmáticas, etc. E empregamos, implícita e probabilisticamente, todos estes tipos de conhecimentos na aprendizagem/uso de uma língua.

O cálculo estatístico de contingência de palavras (ou, “diga-me com quem tu andas e direi quem tu és”) é uma das principais características da cognição humana. Como dito anteriormente, a linguagem é um sistema complexo adaptativo (altamente dinâmico). A competição de informações oriundas dos vários níveis da organização do sistema lingüístico (fonotático, lexical, morfossintático, discursivo e pragmático) é constante. Por exemplo, segundo Ellis (2002a), nosso sistema cognitivo é perito em calcular qual alternativa se mostra mais adequada num determinado contexto de uso ao constantemente realizar identificação de padrões sequenciais no insumo e suas probabilidades (como nos exemplos de *parsing* mostrados acima).

Inspirado em Ellis (2003), adaptei para o português o exemplo a seguir: qual seria a próxima letra numa frase em português que comece com “T....” ? Nativos e não nativos proficientes sabem que uma vogal é bem mais provável do que uma consoante, e, dentre as consoantes, *r* é infinitamente mais provável do que as demais (sendo *x*, *z*, *p*, e outras, impossíveis). Qual a primeira palavra provável na frase acima? Possivelmente um verbo, talvez *ter*, já que normalmente um artigo definido ou indefinido precede um substantivo. Tais ‘pistas’ probabilísticas nos ajudam, implicitamente, no uso da língua em nosso dia-a-dia. Pois bem, se “*Ter...*” começa a frase, como ela continua? Com um

artigo ou substantivo, muito provavelmente. Talvez uma preposição. Dificilmente um outro verbo ou um adjetivo. Falantes nativos e proficientes conhecem uma quantidade tremenda de seqüências lingüísticas assim, em todos os níveis. Nós aprendemos inconscientemente estas regularidades presentes no insumo, aprendemos as probabilidades seqüenciais da língua no nível do fonema/letra, palavra e sentença, como destaca Ellis (2003:75). Como comentado na introdução deste manuscrito, estes efeitos de freqüência e recência são importantes diante do caráter gradual e cumulativo da aquisição de uma língua, desde bigramas e trigramas (i.e., o que entendemos como conhecimentos fonotático e ortográfico), passando pelas raízes de palavras, sufixos e prefixos (a morfologia), construções mais complexas como sintagmas verbais/nominais (i.e., a sintaxe) e o discurso; ou seja, dos *tokens* aos *types*<sup>30</sup> ao sistema complexo que é a língua (ELLIS, no prelo; 2002a, 2002b ).

## 2.5 As ‘irracionalidades’ do processo de aquisição de uma L2

A teorização conexionista e probabilística acima discutidas iluminam também a noção tradicional de ‘erro’ lingüístico, sua possível natureza e sua importância para o próprio processo de aprendizagem de línguas (em especial uma L2). No que se segue, discuto, sob a ótica conexionista/probabilística, o conhecido conceito de interlíngua, desenvolvido por Larry Selinker em 1972, conceito este que busca explicar o processo de ASL como sendo algo sistemático e gradual, revelado pela natureza também sistemática dos erros lingüísticos em L2 (refletido por sua instância maior que é a conhecida *fossilização* lingüística). Possíveis semelhanças e diferenças conceituais entre a noção de interlíngua e a filosofia conexionista para a aquisição e processamento da linguagem são apontadas, numa tentativa de compatibilizar o construto interlíngua (i.e., este sistema específico) com redes conexionistas. A presente discussão reelabora os argumentos recentes expostos em Ferreira Junior (2007).

---

<sup>30</sup> Ao longo desta tese usarei na maioria das vezes os termos *token* e *type* em inglês e, em menor proporção, os equivalentes *ocorrência* e *tipo* em português. Além do fato de que, mesmo em língua inglesa, os termos *token* e *type* parecem estar mais entrancheirados entre nós lingüistas, a opção pela terminologia inglesa também evita, a meu ver, quaisquer confusões com seus usos não técnico ao longo de um texto tão extenso (i.e., busca-se uma maior clareza na discussão feita nesta tese).

O Dicionário Houaiss da Língua Portuguesa (2001:1190) fornece as seguintes definições para o verbete *erro*: 1) ato ou efeito de errar 2) juízo ou julgamento em desacordo com a realidade observada; engano 3) qualidade daquilo que é inexato, incorreto 4) desvio do caminho considerado correto, bom, apropriado; desregramento 5) *física* pequena discrepância no valor medido de um observável físico devido a imperfeições dos instrumentos de medida e/ou incorreções do observador 6) *matemática* diferença entre o valor aproximado de uma função ou grandeza e o seu valor real; erro real, erro absoluto.

Todas as definições dadas acima interessam à discussão ora feita, pois, obviamente, definem a palavra 'erro'. No entanto, as definições fornecidas em 5 e 6 tornam-se mais relevantes diante da compreensão de 'erro' da perspectiva conexionista discutida neste trabalho. As definições expressam, de certa forma, a noção logicista clássica de erro como um comportamento *irracional*, um desvio. A concepção logicista ignora que nossos processos cognitivos são dominados por um cálculo constante de 'incertezas', ignorando assim que quem aparentemente 'erra' (uma rede conexionista ou um aprendiz) o faz numa tentativa extremamente racional, portanto acertada, de satisfazer restrições múltiplas com base no estabelecimento de padrões probabilístico-estatísticos (estocásticos) de ocorrência/frequência de uma dada experiência. Esta discussão será retomada mais adiante.

Como destaca Rod Ellis (1997), o 'erro' é um aspecto flagrante e bastante comum em ASL. Vários estudos mostram, de maneira descritiva, os erros presentes na produção oral de falantes não-nativos. Além disso, todos nós cometemos erros na produção lingüística, e tais erros (na maior parte das vezes) tendem a variar (em intensidade e tipo) à medida que vamos ficando mais "proficientes" tanto na língua materna (L1) como na língua estrangeira (L2).

De fato, por um longo período a 'análise de erros' e a 'análise contrastiva' influenciaram teorizações sobre o fenômeno dos erros em ASL. A possibilidade de prever, no processo de aquisição, os casos em que ocorreriam fenômenos como transferência negativa ou 'interferência' de estruturas da L1, provenientes do contraste entre a L1 e a L2 estudada, era algo que tornava a hipótese contrastiva atraente para a aprendizagem (e ensino)

bem-sucedidos de uma L2, como já destacavam Weinreich (1953) e Pit Corder (1969; 1973). No entanto, apontam Widdowson (1996) e Rod Ellis (1997), a explicação do fenômeno do erro como sendo uma simples interferência de estruturas da L1, ou um mau hábito que deveria ser evitado, não era satisfatória do ponto de vista teórico. Além disso, observava-se que os erros em ASL eram sistemáticos e recorrentes, e muitos deles comuns a todos os aprendizes da L2, independentemente de sua L1.

Rod Ellis (1997:17), retomando a teorização de Corder acima mencionada, aponta ainda que é importante diferenciar erro e lapso lingüísticos (*error* e *mistake*, na terminologia inglesa). Os primeiros refletem a natureza (ou estado/estágio) atual de nossa estrutura psicológica da língua (i.e., nosso conhecimento lingüístico) ou ainda lacunas ou inadequações (sistemáticas) em relação a L2 estudada, ao passo que os lapsos são falhas ocasionais na produção, provocadas por fatores afetivos e circunstanciais presentes no uso da linguagem.

Portanto, o erro no processo de ASL é algo bastante sistemático e, por isto mesmo, de certa forma previsível; e, por ser algo consistente e sistemático, os erros indicam que o aprendiz, ou seu sistema lingüístico mental, está seguindo "regras" ou "normas" implícitas, ainda que estas sejam diferentes daquelas da L2 estudada. Mas quais seriam os processos cognitivos subjacentes aos erros? O que, de fato, tais erros representariam? O que tais erros parecem refletir pode ser observado em um importante estudo, agora elevado à condição de clássico por ser largamente citado na literatura sobre aquisição de línguas, conduzido por Rumelhart & McClelland (1986) sobre um fenômeno observado na aquisição do passado de verbos da língua inglesa por crianças.

Rumelhart & McClelland desenvolveram um modelo conexionista para o processamento de verbos da língua inglesa que reproduz o fenômeno da curva em U observado na aprendizagem do passado dos verbos em inglês. O processo de aprendizagem revela, basicamente, três estágios: 1) inicialmente o uso, pela criança, de um número limitado de verbos de alta frequência, geralmente verbos irregulares, e o uso correto do passado de cada verbo (i.e., as crianças parecem usar os verbos separadamente, tratando-os item por item; 2) em um segundo estágio, à medida que a criança entra em contato com mais e mais insumo, ela começa a generalizar, tratando



todos os verbos como regulares. Neste estágio há um processo de supergeneralização de regras/padrões abstratos subjacentes às formas verbais, ou seja, busca de aprendizagem sistemática. Nesta fase ocorrem muitos erros, resultantes de uma busca constante por sistematização através da generalização de padrões diferentes. 3) o último estágio revela a coexistência de formas verbais regulares e irregulares, consequência de um processo de categorização a partir da sistematização de padrões diferentes. O Quadro 1, adaptado de Ellis (1997:23), ilustra bem o fenômeno relatado no estudo aqui descrito:

| <i>Estágio</i> | <i>Descrição</i>   | <i>Exemplo</i> |
|----------------|--|----------------|
| 1              | Os aprendizes inicialmente não usam o verbo no tempo passado.                        | 'eat'          |
|                | Os aprendizes começam a produzir formas do passado irregular dos verbos              | 'ate'          |
| 2              | Os aprendizes supergeneralizam a forma do passado regular dos verbos.                | 'eated'        |
|                | Às vezes os aprendizes produzem formas híbridas.                                     | 'ated'         |
| 3              | Os aprendizes finalmente produzem as formas corretas do passado irregular dos verbos | 'ate'          |

### Quadro 1

Estágios na aquisição do tempo passado do verbo *eat*

O estudo de Rumelhart & McClelland revela duas coisas muito importantes. Em primeiro lugar, parece haver, por parte do sistema lingüístico do aprendiz, um processo contínuo de re(construção) ou (re)estruturação do conhecimento através da busca constante de adequação deste sistema às informações presentes no insumo lingüístico. Nesta busca, o sistema parece seguir regras (exibido pelo seu comportamento). No entanto, tais regras não existem. Ou melhor, o comportamento revela um processo de regularização de insumo e estabelecimento de padrões de base estatístico-probabilística.

O estudo acima relatado também revela que, apesar do fenômeno ser observado em estágios aparentemente estanques, o pleno desenvolvimento ou a plena aquisição das estruturas da linguagem (revelado pela completude da curva em U) é altamente irregular e não-linear. Esta inevitável irregularidade e não-linearidade é justamente o que produz comportamentos erroneamente taxados de erros – quando analisados de uma perspectiva puramente comportamental ou de ‘desempenho’ lingüístico, como podemos observar nos exemplos do quadro 1 acima. Da perspectiva do sistema lingüístico do

aprendiz de uma língua os erros tornam-se acertos, pois, como destacado no início deste trabalho, revelam uma busca lógica por sistematicidade ou regularização.

Finalmente, é importante observar a diferença crucial entre a aprendizagem de itens em separado e a aprendizagem sistemática, ou a aprendizagem das regras/padrões abstratos subjacentes ao uso dos itens lingüísticos (sejam eles sufixos de passado, palavras, estruturas, etc.). E isto é evidenciado por estudos conduzidos com bebês no qual são usadas gramáticas artificiais, como o trabalho de Gómez & Gerken (1999; 2000)<sup>31</sup>.

No estudo de Gómez & Gerken acima mencionado, os bebês, de 1 ano de idade, diferenciaram seqüências gramaticais das agramaticais apesar das mudanças na ordem das palavras. Eles também conseguiram diferenciar as estruturas da gramática na qual haviam sido treinados das outras estruturas de uma outra gramática. Finalmente, os bebês conseguiram também distinguir entre novas seqüências da gramática na qual foram treinados e seqüências produzidas por uma outra gramática, apesar das gramáticas usarem vocabulários distintos. Ou seja, os bebês abstraíram algum aspecto da estrutura gramatical além dos pares de elementos específicos.

Há importantes implicações, para a aquisição da linguagem, decorrentes de estudos como o de Gómez & Gerken aqui relatado. Uma diz respeito ao tipo de mecanismo subjacente ao processo de aquisição. Ou seja, o mecanismo parece não ser específico para a aquisição da linguagem, evidenciado pelo emprego de probabilidades transicionais na identificação das seqüências gramaticais. Outra implicação decorrente do estudo diz respeito aos possíveis níveis de abstração. Para as autoras, é preciso diferenciar entre abstrações baseadas em associação de padrões (*pattern-based abstraction*) e abstrações baseadas em associação de categorias (*category-based abstraction*). Apenas o segundo tipo de abstração possibilitaria a aquisição da sintaxe. O primeiro tipo de abstração baseia-se em relações de identidade perceptualmente limitadas (puramente físicas, como, por exemplo, perceber um padrão de ordenamento do tipo ABA), ao passo que a segunda baseia-se em relações de identidade em

---

<sup>31</sup> Não serão aqui detalhados os procedimentos metodológicos de estudos psicolingüísticos experimentais envolvendo bebês, como os adotados no experimento ora relatado.

categorias abstratas (abstração a partir de elementos já abstratos, ou categorias), como substantivo-verbo-substantivo.

Observe que há uma semelhança *superficial* entre o padrão *ABA* com a ordem de categorias *substantivo-verbo-substantivo*. Ou seja, as relações envolvendo categorias ocorrem em um nível superior de abstração (as quais não podem basear-se simplesmente na percepção (física) de identidade)<sup>32</sup>. Estes estudos são consoantes com a teorização construcionista, como veremos mais adiante no capítulo 4 desta tese.

A habilidade humana de abstrair com base em categorias (abstração de abstrações), como visto acima, é fundamental para a produção lingüística. Veja o exemplo fornecido em Gomez & Gerken (2000:183), sobre um conhecido teste feito com crianças que, quando deparadas com um determinado objeto referido como "isto é um *wug*" pelo pesquisador e, logo em seguida, por ele inquiridas "Agora existem dois deles, então existem dois \_\_\_\_\_?" elas tendem a responder "*wugs*". Ou seja, dominam por abstração a categoria *plural*.

Parece haver, portanto, processos de abstração em dois níveis: 1) o nível da identificação da relação de padrões perceptualmente limitados (físicos); e 2) o nível (superior) da relação de membros de categorias similares identificados, inicialmente, nas relações de padrões 1, ou seja, tornados salientes e, desta forma, identificáveis através de relações estocásticas de estabelecimento de pistas sistematicamente relacionadas.

## **2.6 Revisitando o conceito selinkiano de interlândia**

Um conceito que, de certa maneira, representa um grande avanço nos estudos em ASL no tocante à busca pela compreensão da "natureza do erro" ou do processo de

---

<sup>32</sup> Por exemplo, as frases *dogs eat pizza* e *john loves books* partilham a mesma estrutura de categorias, embora haja diferenças (físicas) entre membros de categorias idênticas como *John* e *books* (Gomez & Gerken, 2000:183). A mesma discussão, como veremos no item 4 adiante, é feita em Elman et al (1998) sobre modelos conexionistas que fazem apenas associações bidimensionais (relações baseadas puramente em *formas*) e modelos que realizam associações tridimensionais (relações baseadas em *forma* e *espaço*) das representações geradas.

fossilização é o conceito de interlíngua (IL), proposto por Larry Selinker em 1972. Em um artigo seminal publicado naquele ano, Selinker aponta (como revelado pelos estudos de análise contrastiva entre L1 e L2) que frases produzidas pela maioria dos aprendizes de uma L2 (em termos de estruturação) diferem daquelas mesmas frases produzidas por falantes nativos da L2 estudada, deduzindo assim que deveria existir um sistema lingüístico em separado para esta L2. Selinker, com isto, apontava para o processo de ASL como sendo algo altamente sistemático e gradual. O autor sugere, refinando *insights* anteriores de Weinreich e Corder, a existência de uma "interlíngua" responsável pelo processo de ASL. Os aprendizes desenvolveriam tal sistema lingüístico único, uma estrutura lingüística mental altamente dinâmica, diferente tanto da língua materna (L1) como da L2 estudada. A 'competência' lingüística do aprendiz, em qualquer estágio do processo de ASL (i.e, sua interlíngua), seria influenciada por uma *estrutura psicológica latente* (*latent psychological structure*) que possuímos, e que permite, na constante busca de sentido na aprendizagem da L2 por parte do aprendiz, que "identificações entre-línguas" sejam feitas (*interlingual identifications*, ou seja, relações entre a sua L1 e a L2 estudada).

Selinker destaca, assim, que tanto o sucesso como o fracasso na aprendizagem de uma L2 dependerão sempre da ativação inicial de nossa *estrutura psicológica latente*, cujas restrições (impostas inicialmente tanto pela sua L1 como pela L2 estudada) condicionam a busca constante de produção de significados/sentidos, gerando, desta forma, um terceiro sistema lingüístico (i.e., a interlíngua). Em última instância, a aquisição de uma língua se revelará como uma constante e permanente "*reorganização do material lingüístico de uma IL*" de forma a ajustar-se às convenções (ou 'regras', como preferem outros) da L2 estudada, como originalmente observado por Selinker (1972:224).

É indispensável frisar que, para Selinker, tal *estrutura psicológica latente* é inata. Mas este inatismo é em relação aos processos cognitivos (a configuração biológica) e não em termos proposicionais (lingüísticos):

É importante destacar que a estrutura latente descrita neste artigo não possui nenhum programa genético, se comparada com a proposta de Lenneberg; não há um equivalente direto com

nenhum conceito do tipo ‘gramática universal’, não há nenhuma garantia de que esta estrutura latente será ativada; não há nenhuma garantia de que tal estrutura venha a se tornar uma estrutura de fato de qualquer língua natural (i.e., não há nenhuma garantia que a aprendizagem será bem sucedida), e é bem possível que exista uma superposição entre tal estrutura latente para a aquisição da linguagem com outras estruturas do nosso intelecto.<sup>33</sup> (1972:212).

No entanto, para explicar os cerca de 5% dos aprendizes que desenvolvem uma competência lingüística e atingem um desempenho idênticos ao de um falante nativo da L2 estudada, Selinker lança mão de uma suposta ativação da Gramática Universal (UG) chomskiana (que conteria todo o conteúdo proposicional) através de um mecanismo de aquisição de línguas, um LAD, presente no conceito análogo de *estrutura lingüística latente* proposto por Lenneberg e adotado por Selinker em seu artigo (1972:211-12).<sup>34</sup>

Colocada desta maneira, a teorização de Selinker, do ponto de vista conceitual, contém uma certa contradição; evidentemente não precisamos recorrer a conteúdos proposicionais inatos para explicarmos a exceção à regra que constituem os 5% dos aprendizes "bem sucedidos" de uma L2. É também importante observar que o artigo de Selinker foi escrito, e o construto interlíngua desenvolvido, sob forte influência do paradigma chomskiano que se estabelecia na época. De fato, em reformulações recentes do conceito de interlíngua, Selinker (1996) reconhece a necessidade de entendermos a competência “interlingual” (*interlanguage competence*) em termos dos mecanismos subjacentes ao seu processo de estruturação, sugerindo também a possibilidade de que modelos conexionistas possam explicar os mecanismos processuais e a natureza altamente variável dos fenômenos lingüísticos que o conceito de interlíngua encerra (cf. SELINKER, 1996: 97).

---

<sup>33</sup> Minha tradução de “It is important to state that with the latent structure described in this paper as compared to Lenneberg’s, there is no genetic timetable; there is no direct counterpart to any grammatical concept such as ‘universal grammar’; there is no guarantee that this latent structure will be activated at all; there is no guarantee that the latent structure will be ‘realized’ into the actual structure of any natural language (i.e., there is no guarantee that attempted learning will prove successful), and there is every possibility that an overlapping exists between this latent language acquisition structure and other intellectual structures”.

A idéia de uma *interlíngua* (principalmente como sugerida nas reformulações recentes de Selinker) é altamente compatível com modelos conexionistas para o processamento de sistemas estruturados. A filosofia norteadora dos modelos conexionistas parece dar conta, ou pelo menos vislumbra a possibilidade de fornecer boas explicações, do processo de aquisição de línguas naturais. O conceito de interlíngua, acima de tudo, redireciona a teorização em ASL para aspectos importantes do processamento cognitivo dos aprendizes.

### ***2.7 A interlíngua vista como uma rede conexionista em estruturação***

Como vimos acima, o mérito principal de Selinker foi o de apontar para o fato de que há, durante o processo de ASL, a criação e o desenvolvimento de um sistema lingüístico autônomo (no tocante às regras de funcionamento a ele subjacentes). Contudo, o conceito 'interlíngua' carece de explicações mais detalhadas sobre a natureza dos processos psicológicos básicos. Selinker (1972:224) fala de "unidades psicológicas" responsáveis pelas "relações interlinguais" (i.e., relações entre a L1, a IL e a L2) que ocorreriam na *estrutura psicológica latente* do aprendiz. Estas unidades não teriam caráter apenas lingüístico, mas um caráter psicológico geral, ou seja, lá ocorreriam relações de toda a natureza. A interlíngua seria, numa possível interpretação conexionista, caracterizada como unidades psicológicas emergentes na constante busca, por parte do aprendiz, de sentido/significado no uso da L2.

É possível interpretar o texto de Selinker, por sinal bastante hermético (como bem destacado por ROD ELLIS, 1997)<sup>35</sup>, à luz da filosofia conexionista. É altamente válido comparar o construto *interlíngua*, e o que o autor chama de *estrutura psicológica latente*, com a arquitetura e o funcionamento de modelos conexionistas. Acredito, portanto, que modelos conexionistas explicam, de uma maneira bastante plausível por partirem de uma base neurofisiológica, o processo de ASL como algo parecido com o

---

<sup>34</sup> Veja White (2003) para uma discussão acerca do conceito interlíngua de uma perspectiva inatista (i.e., chomskiana) no âmbito dos estudos em ASL.

<sup>35</sup> Rod Ellis (1997:127) destaca, na bibliografia comentada ao final de seu livro, que o artigo de Selinker "is not easy to read, but it gave SLA the term 'interlanguage' and it contains a rich seam of theoretical ideas that is still being mined today".

construto selinkiano de uma “interlíngua”, caracterizado pelo que chamei acima de 'unidades psicológicas emergentes'. Como sugerido posteriormente pelo próprio Selinker (1996), também acredito haver fortes semelhanças entre ambas as teorizações.

Quais comparações podemos, então, estabelecer entre a proposta de Selinker e os modelos conexionistas? Em ambos os casos há, inicialmente, um "alvo" externo, que é o insumo em L2 (ou ainda a experiência como um todo)<sup>36</sup>. Com o passar do tempo, e devido a constante complexificação da rede, a informação da L2 passa a determinar ajustes locais cada vez mais finos entre unidades específicas. Este processo de "calibração", que busca minimizar o 'erro', é feito com base em processos probabilístico-estatísticos. A externalização deste processo de calibração se dá na forma de 'erros' no desempenho do sistema (ou na produção lingüística do aprendiz), como vimos na tabela 1 acima.

Os 5% (cálculo altamente questionável) dos aprendizes que atingem o sucesso total em ASL, mencionados por Selinker, podem muito bem ser colocados em um dos extremos de um *continuum* de aprendizagem/desenvolvimento, sem que precisemos nos apoiar em argumentos inatistas. As redes são mais sensíveis a erros no início da aprendizagem. À medida que o treinamento avança, as ativações das unidades são alargadas (em termos de variação e tolerância). Ou seja, vão do centro para os extremos (em escala). Isto diminui a velocidade da aprendizagem e, em última instância, gera o que chamamos em ASL de fossilização (ou seja, exemplos positivos não mais afetam a interlíngua ou o sistema).

Portanto, nenhuma teorização inatista sobre um determinado período crítico de aprendizagem parece ser necessária. Tais “períodos” são uma consequência do próprio desenvolvimento ou estruturação da rede; ou seja, uma vez que certa configuração, ou protótipos (fonéticos, sintáticos, etc..), se estabelece, o sistema praticamente se fecha (como destacado no parágrafo anterior). Em termos neurofisiológicos tal fenômeno

---

<sup>36</sup> Embora passe aqui a idéia de língua como um ‘alvo’, destaco, em tempo, ser este não um alvo estático, mas um em constante movimento e altamente dinâmico!

equivaleria à diminuição observada na neuroplasticidade humana (cf.: Elman et. al., 1998:389).<sup>37</sup>

### ***2.8 A linguagem como um sistema probabilístico***

Como anteriormente discutido, as redes conexionistas estruturam-se com base em um cálculo probabilístico/estatístico da frequência de regularidades detectadas em sua experiência com o insumo (i.e., a frequência de (co)ocorrência de certos padrões no insumo). A aprendizagem da língua, podemos pensar, é, portanto resultante, em última instância, de um processo de regularização de insumo e estabelecimento de padrões de base estatístico-probabilística, como apregoam os pesquisadores em um ramo da lingüística em ascensão neste momento chamado de Lingüística Probabilística.

Bod, Hay & Jannedy (2003:08) destacam que a linguagem, como outros domínios cognitivos, parece ser regida por um processamento probabilístico. Conseqüentemente, destacam os autores, devido ao fato de o processo de aquisição da linguagem ser aparentemente um processo probabilístico, isto faz com que as "generalizações baseadas em inferência estatística tornem-se cada vez mais robustas à medida que aumenta o tamanho da amostra" (p.08). E talvez o *insight* mais importante desta nova maneira de encarar a linguagem, como observou Seidenberg (1997:1601), seja o fato de que este mesmo mecanismo probabilístico que extrai regularidades estruturais no insumo durante o processo de aquisição também esteja presente no uso da linguagem.

Como apontam inúmeros estudos que utilizam modelos conexionistas, a linguagem parece mesmo possuir propriedades probabilísticas. Os modelos conexionistas exibem um desempenho baseado em gradação (contínuo) no qual o limiar de força/peso de ativação estabelece, gradualmente, categorizações; ou seja, os modelos "são capazes de

---

<sup>37</sup> Entretanto, no tocante a aprendizagem de línguas, uma "desfossilização" é algo possível por meio de instrução com foco na forma e um "estímulo exagerado". Como observa Ellis (2004:63), um modelo conexionista desenvolvido por McClelland et al (2002, *apud* Ellis) mostrou-se eficaz no clássico fenômeno da discriminação dos fonemas /l/ e /r/ na aprendizagem do inglês por parte de aprendizes japoneses.



fazer distinções sutis e mesmo categorizar ao longo de dimensões que parecem ser, por natureza, contínuas” (ELMAN et. al.,1998:53).

Alguns estudos importantes que podemos destacar, e que apontam para processos de base estatístico-probabilística, são o estudo de categorização de Elman (1990) que revela 1) a presença de regularidades estatísticas na segmentação fonotática de frases e 2) aponta gradientes na categorização de verbos e substantivos; o estudo de Gómez & Gerken (1999;2000) que aponta o papel das probabilidades transicionais na identificação de padrões em seqüências gramaticais, revelando, assim, que é possível "aprender" padrões sintáticos abstratos; bem como o já citado estudo sobre os verbos irregulares, desenvolvido por Rumelhart & McClelland (1986), que mostra a existência de regularidades e subregularidades dentro da categoria tradicionalmente chamada de "verbos irregulares".

Tais estudos apontam, ao que parece, para o processo de aquisição da linguagem, ou o estabelecimento do sistema lingüístico, como uma consolidação gradativa de processos gradativos de estabelecimento de padrões também gradativos, sempre guiados por um mecanismo estatístico/probabilístico, numa busca constante de aproximação entre a interlíngua do aprendiz e a L2 por ele vivenciada.

No item a seguir farei uma rápida discussão sobre os mecanismos de *chunking* na aprendizagem de uma língua. Com isto, retomo nossa discussão inicial sobre o papel central das fórmulas lingüísticas na aprendizagem de um língua (tema principal deste manuscrito).

## 2.9 Processos de *chunking*<sup>38</sup>

Ellis (2003), citando originalmente William James (1890), destaca o que chama de Lei da Contigüidade, certamente o princípio mais básico de processos de associação:

Objetos uma vez vivenciados conjuntamente tendem a tornar-se associados na imaginação, de forma que quando um deles é lembrado os outros provavelmente também o serão, na mesma ordem de seqüência ou coexistência anteriormente vivenciada. (James, 1890, *apud* Ellis, p.74-75).<sup>39</sup>

Uma compreensão mais clara destes processos básicos de associação é a noção de *chunking* expressa por Newell (1990, *apud* Ellis 2003):

Um *chunk* é uma unidade de organização da memória, construído a partir da junção de um conjunto de elementos previamente já formados na memória (os quais, eles próprios, podem ser *chunks*) que, por sua vez, são fundidos em uma unidade ainda maior. O processo de *chunking* pressupõe a habilidade de construir tais tipos de estruturas de maneira recursiva, provocando assim uma organização hierárquica da memória. O processo de *chunking* parece ser uma característica comum à memória humana (p.76).<sup>40</sup>

Tal princípio, segundo Ellis (2001, 2002a, 2002b, 2003 & 2004), está presente em muitos processos de percepção, aprendizagem, solução de problemas e cognição em geral. Processos de *chunking* permeiam todos os aspectos da nossa representação mental da linguagem. Usuários fluentes de uma língua conhecem (tacitamente) uma quantidade incalculável de seqüências lingüísticas.

---

<sup>38</sup> Ao longo desta tese venho utilizando a expressão ‘processos de *chunking*’ para descrever o fenômeno denominado simplesmente de *chunking* na literatura em língua inglesa. Não é de meu conhecimento, até este momento, uma versão deste termo na língua portuguesa. Um dos objetivos desta tese seria também propor um termo equivalente em nossa língua. Os únicos candidatos foram ‘empedacamento lingüístico’ ou ‘processos de empedacamento lingüístico’, criados a partir do substantivo ‘pedaço’, mas que, para mim, ainda soam muito estranhos. Diante disto, optei por manter a terminologia inglesa.

<sup>39</sup> Minha tradução de: *Objects once experienced together tend to become associated in the imagination, so that when any one of them is thought of, the others are likely to be thought of also, in the same order of sequence or coexistence as before.*

<sup>40</sup> Minha tradução de: *A chunk is a unit of memory organization, formed by bringing together a set of already formed elements (which, themselves, may be chunks) in memory and welding them together into a larger unit. Chunking implies the ability to build up such structures recursively, thus leading to a hierarchical organization of memory. Chunking appears to be a ubiquitous feature of human memory*

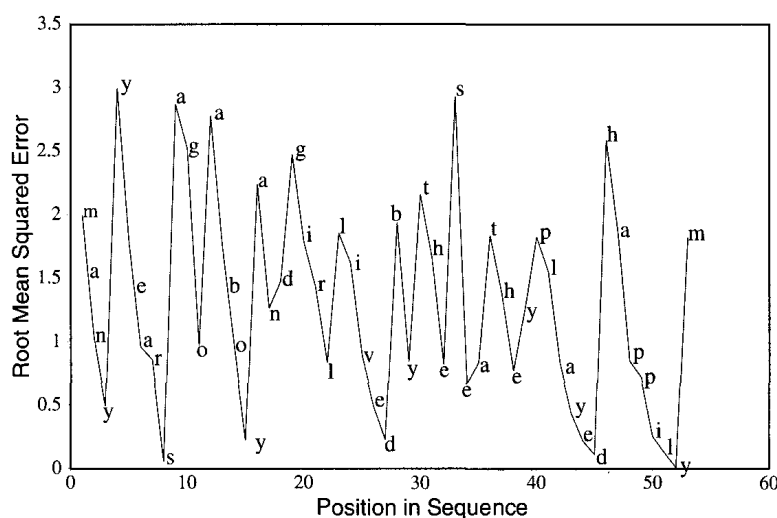
Veremos a seguir um exemplo trivial de processos básicos de *chunking* observados na linguagem, inspirado em um clássico estudo de Elman (1990), intitulado *Finding Structure in Time*, descrito também em Elman et al. (1998). Abaixo transcreverei a citação original de William James dada acima em um fluxo contínuo, sem espaços e sem pontuações:

*objectsonceexperiencedtogethertendto becomeassociatedintheimaginationsothatwhenan  
yoneofthemisthoughtoftheothersarelikelytobethoughtofalsointhesameorderofsequenceor  
coexistenceasbefore*

Vamos supor que estivéssemos querendo aprender a língua inglesa e ouvíssemos a seqüência acima. Se for a primeira exposição à língua, tudo não passará de barulho e seqüência de sons sem nenhuma segmentação (o que acontece quando ouvimos qualquer língua que não conhecemos). O insumo a nossa volta é cheio de regularidades e nosso sistema cognitivo, ao processar o insumo, é perito em detectar tais regularidades e padrões na língua através de um mecanismo indutivo, associativo e probabilístico. À medida que vamos tendo contato com a língua nos mais variados contextos, começamos a segmentá-la (desde níveis bem baixos, por exemplo, no nível do fonema > palavra > frase, etc.) com base na freqüência de ocorrência de certos padrões detectados.

O sistema, então, começa a aprender padrões seqüenciais recorrentes (a partir de cálculos probabilísticos) e começa a inferir seqüências prováveis na língua a qual estamos expostos. Assim, em um nível sublexical, conseguirá prever, ao ouvir no trecho acima transcrito, a seqüência inicial OB, se o que virá será EDIENT, ELISK, ESE, EY, LIGATION, LIGE, LIQUE, LIVION, etc.; ou ainda, se ao ouvir a seqüência maior OBJE, o que virá depois será CT, CTION, CTIONABLE, CTIVE, etc. É assim que nós desenvolvemos nosso conhecimento fonotático. Ou seja, falantes do inglês sabem que em uma palavra que comece com o fonema /t/ a probabilidade que o fonema seguinte seja /h/ ou uma vogal é infinitamente maior do que um /z/ ou outras consoantes, por exemplo. E assim vamos construindo *chunks* com base em restrições variadas. Somos sensíveis a estas combinações probabilísticas de certas seqüências em uma língua.

Continuando o exemplo para além de formas sublexicais, nosso sistema cognitivo também, aos poucos, vai um nível mais acima e começa a diferenciar palavras, pois nas fronteiras das palavras as probabilidades transicionais são menores, da mesma forma que começa a fazer previsões de co-ocorrência de palavras e co-ocorrência de seqüências inteiras de grupos de palavras (processos maiores de *chunking*). Assim as formas vão "emergindo" nos vários níveis (fonológico, lexical, morfossintático, etc.). Tudo isto, desde o processo mais baixo, é feito a partir de um cálculo estatístico/probabilístico contínuo. Veja ilustração (Figura 5) reproduzida de Elman et al (1998) para a frase: *manyyearsagoaboyandagirlivedbytheseatheplayedhappily*



**Figura 5** – Curva de erro numa rede treinada em tarefa de previsão de fonema. O erro é maior no início de uma palavra e diminui à medida que a palavra é processada. In: Elman et al (1998), p.121.

Quanto mais "cresce" a seqüência, quanto mais robusta ela é, maiores serão as restrições impostas ao sistema e, portanto, mais exatas e mais corretas (e mais rápidas) serão as previsões resultantes do processamento estatístico/probabilístico, pois o "contexto", por alargar-se cada vez mais, fornece mais restrições ao sistema em termos de "pistas" durante o processamento, o que torna tal processamento algo mais fácil, rápido ou automático. Visto desta forma, o processamento rápido da linguagem é essencialmente calcado na memória e, portanto, consiste basicamente em um processo formuláico. No

experimento de Elman (1990) citado, por exemplo, processos de *chunking* seriam responsáveis pelo estabelecimento de dependências estruturais cada vez maiores ao longo da aprendizagem da linguagem (resultando, desta forma, em uma complexificação cada vez maior do sistema lingüístico ao longo de vários níveis ou hierarquias, ou seja, fonológico, morfológico, morfossintático, discursivo etc.).

Processos/mecanismos de *chunking* parecem permear o desenvolvimento da competência em muitas habilidades cognitivas desde o nível celular, como discutimos anteriormente na discussão sobre memória procedimental no item 1.3.1. Da mesma forma vimos que nossa memória de trabalho (MT) é limitada. No entanto, a capacidade de informação de nossa MT pode ser aumentada por processos de *chunking*, por meio do qual informações/estruturas que co-ocorrem podem ser “fundidas” para que se tornem, em termos representacionais, uma única unidade ou entidade. Este processo representacional é extremamente útil e indispensável em desempenhos considerados especializados (cf. ELLIS, 2003:76), como jogar xadrez, traduzir/interpretar, falar uma língua, etc. Portanto, tais processos de *chunking* são uma realidade neurobiológica.

Para concluir nossa discussão, é importante destacar que a defesa da repetição de estruturas lexicais convencionalizadas (ou rotinas) nas teorizações em ASL tornou-se algo considerado tabu após o *debacle* do behaviorismo. É amplamente conhecido o impacto, em grande parte negativo, provocado pela corrente behaviorista nas teorizações sobre a aquisição de L2. Entretanto, um *insight* valioso da corrente behaviorista, a saber, a importância crucial atribuída ao insumo lingüístico, foi, segundo Gardner et. al, (1998), descartado pela corrente cognitivista que se estabeleceu a partir das idéias revolucionárias de Chomsky. Todavia, como vimos nas páginas acima, tanto as teorizações sobre a neurobiologia da aprendizagem como os modelos conexionistas para o processamento da informação (e, conseqüentemente, modelos de aquisição e uso da linguagem) têm enfatizado o importante papel desempenhado pelo insumo, e seu constante reforço, na permanente (re)configuração e consolidação das redes de conexões que constituem a nossa memória.

Recapitulando, este trabalho sustenta que o uso e a aquisição da linguagem baseiam-se, principalmente, na utilização da memória, ou seja, um processo pautado basicamente

pela utilização de fórmulas lingüísticas convencionalizadas (i.e., estruturas lexicais dos mais variados tamanhos combinadas de forma a produzir um texto contínuo e coerente). Para entendermos o funcionamento de um sistema lingüístico desta natureza temos que optar por um modelo de representação holístico e dinâmico do tipo conexionista. Nestes modelos, podemos entender a estruturação do nosso sistema lingüístico como resultante da constante repetição e co-ocorrência de estruturas lexicais rotinizadas, provocando a automatização da língua (i.e., a conversão do insumo em *intake*) a partir de processos cognitivos associativos de estabelecimento de padrões estatísticos (de base estocástica) de ocorrência/frequência de tais estruturas lexicais. Como vimos, os modelos conexionistas e processos de *chunking* possuem plausibilidade neurobiológica. E isto reforça o fato de que muito da nossa produção lingüística é calcada, aparentemente, na repetição (ou co-ocorrência) de discursos anteriores, i.e., baseia-se em assuntos triviais e previsíveis, portanto altamente repetitivos e nada originais.

### 3. Categorias & Conhecimento Semântico

*The very structure of language itself is a matter of categorization. Rosch's discoveries regarding the internal structure of categories are no less relevant to the category of language than they are to the categories symbolized by language.*

John R. Taylor

Há mais de três anos tive uma conversa hilariante com um conhecido. Ele, um vendedor ambulante de livros e deficiente físico, estava indignado com o que acontecera numa reunião da qual participara no dia anterior com o propósito de se criar uma associação de deficientes físicos na cidade onde residíamos. Sua indignação estava no fato de terem elegido um sujeito “caolho” para presidir a tal associação, o que o fez protestar veementemente. Em sua opinião um “caolho” não é deficiente e, portanto, não deveria de forma alguma ser o representante maior de associação tão importante. Afinal, segundo ele, lá havia gente “sem braço”, “sem perna”, paraplégicos, dentre outros. Ou seja, “verdadeiros” deficientes. Lembro que conversamos durante um tempo, ali em pé, na calçada defronte ao cinema. Ele, por discordar e expor claramente seu ponto de vista, havia sido expulso da tal reunião na Câmara dos Vereadores ou na Prefeitura, o local não me recordo exatamente. Concluí a conversa em tom de ironia dizendo a ele o que todos nós infelizmente já sabemos: que aqui no Brasil são justamente as pessoas menos capazes que, na maioria das vezes, ocupam os cargos mais importantes!

Pois bem, a tal conversa me veio à mente para iniciar esta breve reflexão sobre uma questão lingüística, e epistemológica acima de tudo, extremamente importante na nossa discussão: os processos de categorização. Como classificamos as coisas<sup>41</sup>? O que nos

---

<sup>41</sup> Nesta tese, devido ao seu enfoque lingüístico, temos em mente, em especial, a questão de como se dá o desenvolvimento de um léxico mental, sendo tais “coisas” as palavras (substantivos e verbos, principalmente) de uma determinada língua. Esta discussão será feita mais adiante ao abordarmos a proposta de léxico mental de Elman (2004).

leva a agrupar certas coisas em classes (ou categorias) idênticas ou distintas? Quais critérios usamos neste processo constante de classificação ao qual estamos expostos? Em suma, como se organiza e qual a natureza da estruturação de nosso conhecimento semântico? Acredito que respostas a algumas destas perguntas já foram fornecidas, ainda que de forma implícita, na discussão feita no capítulo anterior ao discutirmos, principalmente no item 2.1, questões relativas ao desenvolvimento, representação e processamento do conhecimento de uma perspectiva conexionista e emergentista.

Filósofos desde Platão e Aristóteles se ocuparam deste problema. Também psicólogos se preocupam com o problema da categorização de nosso conhecimento semântico desde quando a Psicologia se estabeleceu como ciência, no final do século XIX, como destacam Rogers & McClelland (2004) e Schultz (1975). Mas abordaremos o assunto a partir da psicologia cognitiva, da lingüística cognitiva e da filosofia conexionista. Em primeiro lugar, é importante compreender a noção de conceito em termos gradativos (ou escalar). Tal noção difere radicalmente da noção clássica de conceito e categoria ainda expressa tanto em uma parte da literatura especializada como na percepção do leigo em seu dia-a-dia, como vimos no relato que abre este capítulo. Assim, abordaremos a noção de categorias e conhecimento semântico de uma perspectiva gradativa e probabilística.

### 3.1 Definindo termos

Como aponta Clark (1993:87), a função mais evidente dos conceitos é a de categorizar (i.e., promover processos de classificação). Um **conceito**, de certa forma, é um mecanismo de categorização por excelência, pois permite, por exemplo, que as pessoas atribuam exemplos (*tipos de cachorros*) à uma dada categoria (*cachorro*). Por outro lado, Murphy (2002:5) define **conceito** como sendo uma representação mental de algo, ao passo que uma **categoria** é definida como sendo simplesmente classes de coisas. Finalmente, McClelland & Rogers (2003:311) define o **conhecimento semântico** como sendo o conhecimento sobre objetos e suas propriedades, e das relações entre eles, incluindo o conhecimento sobre os significados de palavras. O conhecimento enciclopédico geral é às vezes também incluído neste conceito.



A noção clássica prevalente e altamente disseminada de conceitos e categorias como sendo o alicerce de nosso conhecimento está entre nós há pelo menos 2.500 anos, como destacam Oaksford & Chater (2007) e Murphy (2002). Tal noção clássica determina que a entidade a ser categorizada precisa, dentre outras coisas, se encaixar dentro de um conjunto apriorístico de atributos ou propriedades consideradas necessárias e suficientes, o que faz com que todos os membros de uma dada categoria possuam o mesmo estatus. Assim, se algo contém todas aquelas definições ou propriedades fornecidas *a priori*, então pertencerá automaticamente a uma dada categoria. Na filosofia, a invenção do silogismo aristotélico pode ser visto como uma das mais antigas teorias de organização do conhecimento baseado em categorias, como destacado por Oaksford & Chater (2007), Rogers & McClelland (2004) e McClelland & Rogers (2003).

Alguns exemplos clássicos são normalmente fornecidos por autores que tratam do tema: o conceito de um *solteirão* e o conceito de *jogos*, exemplificado originalmente por Wittgenstein (cf. ROGERS & McCLELLAND, 2004; MURPHY, 2002; CLARK, 1993). Na concepção clássica, um solteirão deve ser um homem que não seja casado e seja adulto. Para nos certificarmos da validade do conceito, basta checar se o exemplo (uma dada pessoa) encaixa-se nas três definições que encerram o conceito de solteirão. Entretanto, logo fica claro que nem todos os homens não casados e adultos são solteirões. O mesmo ocorre com o conceito de jogos: não é possível haver uma lista de atributos que determine, com total segurança, que uma dada atividade seja um jogo. Há certamente algo que faz com que todos os jogos sejam percebidos como tal, mas não há algo comum a todos eles.

### **3.2 O impacto das idéias de Eleanor Rosch**

O trabalho pioneiro de Eleanor Rosch (1978, e em outros artigos publicados ao longo da década de setenta passada) a fez rejeitar a idéia clássica de que objetos são categorizados com base em atributos fixos ou propriedades apriorísticas consideradas necessárias e suficientes. Rosch estabeleceu 1) a noção de categoria (em termos gradativos) como um número de objetos considerados equivalentes entre si; 2) a idéia de um nível básico de categorização, cuja frequência de ocorrência é fator determinante e,

em consequência disto, 3) a idéia de prototipicidade ou protótipos semânticos<sup>42</sup>. Com isto, de acordo com Murphy (2002:16), Rosch colocou por terra a concepção clássica de categoria. Para Rosch, seguindo Wittgenstein, uma explicação do conceito de jogos precisa levar em conta que há características (ou atributos) que se sobrepõem a todos os tipos de jogos, como os traços físicos ou fisionômicos observados nos membros de uma dada família ou ainda como os ‘fios que se entrelaçam numa corda’ (TAYLOR, 2008:44). Ou seja, assim como não podemos observar em todos os membros de uma família a presença dos mesmíssimos traços em todos os seus entes (por exemplo, um filho tem o olho da mãe, o outro o mesmo tipo de cabelo, a outra o queixo da mãe, e assim por diante) também não podemos verificar propriedades ou atributos que estejam presentes em todos os jogos (por exemplo, uns jogos são jogados com bola e com os pés, outros com as mãos, outros usam raquetes, outros fazem uso de tacos, uns são jogados em dupla, outros em grupos, equipes, etc.). E há exemplos que são difíceis de categorizar, como a nossa capoeira: seria jogo, dança ou uma arte marcial?<sup>43</sup>

Portanto, após a teorização de Rosch, passamos a entender as categorias como sendo gradientes, estando seus membros organizados em agrupamentos nada rígidos e percebidos em diferentes graus de tipicidade; que as categorias (e, portanto, seus membros) ocupam espaços mentais multidimensionais que se sobrepõem umas as outras; e que alguns membros de uma dada categoria ocuparão áreas espaciais mais distantes dos exemplos considerados mais prototípicos. A capoeira, por exemplo, parece ocupar regiões de fronteira das categorias *dança*, *jogo* e *arte marcial*, ao passo que o futebol seria provavelmente, para a média dos brasileiros, o protótipo da categoria *jogo* (i.e., um **protótipo** é o melhor exemplo ou representante de uma dada categoria). Conseqüentemente, um outro aspecto importante da teorização de Rosch é o de que as categorias e seus atributos possuem estreita relação e são mediados pela cultura de um dado grupo lingüístico (ROSCH, 1978).

De acordo com Taylor (2008), a noção de categorização baseada em protótipos proposta por Rosch contribuiu enormemente para o nascimento da própria Lingüística Cognitiva.

---

<sup>42</sup> É importante destacar aqui que a própria Rosch descartou qualquer interesse por modelos de processamento baseados na idéia de protótipos semânticos (Rosch, 1978; Rogers & McClelland, 2004). Isto não invalida a noção proposta, apenas expressa os interesses de pesquisa da autora.

<sup>43</sup> Da mesma forma, a conceito de ‘caolho’ como deficiente físico precisa ser entendido nestes termos!

Em especial, a teoria de protótipos de Rosch estabeleceu de maneira precisa que as categorias possuem uma estrutura interna cuja organização revela a existência de membros mais representativos ou melhores exemplos (i.e., protótipos) dentro de uma dada categoria. Um protótipo semântico é um parâmetro cognitivo a partir do qual classificamos ou categorizamos nossas experiências (seja ela lingüística ou não). Tal classificação se dá com base em semelhança ou função semântica com o protótipo. Quanto maior for a frequência de ocorrência de um membro ou representante de uma dada categoria, mais importante será o seu papel na formação da identidade desta categoria, sendo também maior a probabilidade de ele ser considerado o protótipo (i.e, o melhor exemplo).

### 3.3 Níveis de categorização e frequência de ocorrência de protótipos

Um resquício deixado pela concepção tradicional de categorização, baseada em atributos semânticos dispostos de maneira hierárquica, são os três níveis de categorização chamados de nível subordinado, nível básico e nível superordenado. De acordo com McClelland & Rogers (2003), estes níveis ainda expressam a idéia quilliana bastante difundida de que nosso conhecimento semântico estaria organizado hierarquicamente e que o desenvolvimento conceitual partiria de categorias específicas (nível subordinado) para categorias gerais (nível superordenado).

É sabido, no entanto, que o desenvolvimento conceitual inicial de crianças não se dá nem do específico para o genérico nem do genérico para o específico, mas se inicia a partir de palavras presentes em categorias do nível básico. Este seria possivelmente o nível de entrada para o desenvolvimento conceitual, possibilitando que, posteriormente, o desenvolvimento se dê do genérico para o específico. Parece haver um processo de diferenciação conceitual que segue este padrão de desenvolvimento taxonômico. De acordo com Taylor (2008) e McClelland & Rogers (2003), estudos em L1 indicam que são os exemplares localizados no nível básico de uma dada categoria que são os mais frequentemente usados por crianças e adultos. Um exemplo fornecido em McClelland & Rogers (2003:317) refere-se ao fato de que, normalmente, para moradores de zonas urbanas, as palavras **árvore**, **passarinho** e **cachorro** são as primeiras a serem aprendidas, antes dos termos superordenados **planta** ou **animal**, ou dos termos

subordinados **pinheiro**, **canário** ou **poodle**. E isto parece ser uma verdadeira contradição, pois, como vimos acima, o desenvolvimento conceptual se dá do genérico para o específico. Outro fenómeno importante é a generalização a partir das palavras do nível básico: crianças tendem a chamar muitos animais (especialmente os de quatro patas) de **au-au** (ou simplesmente **cachorro**).

A explicação para isto, segundo Rosch (1978), McClelland & Rogers (2003), Rogers & McClelland (2004) vem do fato de que nosso sistema cognitivo explora representações que correspondem a blocos (*bundles*) de atributos que co-ocorrem no meio que nos cerca. Tais categorias básicas possuem propriedades estatísticas especiais pelo fato de que seus membros partilham muitas propriedades entre si e poucas com exemplares de categorias distintas, por isto são úteis. Murphy (2002) também destaca que estas categorias básicas otimizam os aspectos *informativo* e *distintivo* que expressam, ou seja, não são tão genéricas a ponto de não serem informativas nem são tão específicas a ponto de não serem eficazes em relação ao aspecto distintivo, pois possuem poucas características distintivas. Rogers & McClelland (2004:16-17) dão o exemplo de poltronas: elas possuem muitas propriedades em comum umas com as outras, mas a maioria destas características é partilhada também com outros tipos de cadeira. Por isto, objetos numa mesma categoria básica tendem a partilhar muitas propriedades uns com os outros, e poucas com exemplares de outras categorias, fazendo com que tais categorias básicas sejam consideradas mais úteis em termos comunicativos.

Para Taylor (2008:49), os conceitos de protótipo e de nível básico são interdependentes, ou seja, o primeiro relaciona-se com a organização horizontal das categorias, ao passo que o último remete à organização vertical destas em taxonomias.

McClelland & Rogers (2003) e Rogers & McClelland (2004) destacam que uma característica importante dos exemplares considerados prototípicos em uma dada categoria é o seu alto grau de frequência de ocorrência na categoria. É a partir deste protótipo que outros membros da mesma categoria vão sendo incorporados. Os protótipos abrem o caminho para que ocorra, gradualmente, uma co-variação coerente entre os outros membros que farão parte ou não de uma dada categoria. A co-variação coerente é explicada da seguinte forma em McClelland & Rogers (2003)

Uma co-ocorrência consistente de um conjunto de propriedades entre diferentes objetos. O conceito é diferente de uma correlação simples na medida em que ele geralmente se refere à co-ocorrência de mais de duas propriedades. Por exemplo, ter asas, ter penas, ter ossos pneumáticos e ser capaz de voar co-ocorrem todos consistentemente em pássaros (p.318)<sup>44</sup>

McClelland & Rogers (2003) e Rogers & McClelland (2004) propõem ainda uma explicação conexionista, mecânica, para os processos de categorização e diferenciação que subjazem nosso desenvolvimento conceptual (ou seja, do geral para o específico, e inicialmente centrado em protótipos de alta frequência). A proposta incorpora as contribuições de Rosch e busca eliminar resquícios ainda presentes da concepção clássica de categorização (i.e., hierarquizada e ainda rígida). Não há espaço aqui para discutir a proposta dos autores, mas a argumentação baseia-se nos pressupostos conexionistas básicos discutidos no capítulo anterior, ou seja, representações emergentes e distribuídas, que se constituem, em última instância, em configurações ou padrões específicos de ativação em redes reguladas por princípios hebbianos (i.e., redes neurais, cujo processamento simples entre suas unidades vai, de maneira gradual, constante e probabilística, ajustando o peso entre suas conexões em consequência de sua experiência com o meio). Segundo Clark (1993), os princípios de categorização propostos por Rosch são, em essência, um mecanismo de detecção de padrões, característica principal dos próprios modelos conexionistas discutidos no capítulo anterior. Também, a concepção de léxico mental de Elman (2004) discutida no capítulo seguinte é um excelente exemplo da proposta conexionista sobre a natureza dos processos de categorização e desenvolvimento conceptual aqui discutidos.

Finalmente, uma discussão mais aprofundada sobre conceitos e teorias de categorização foge ao escopo desta tese. Sobre o assunto, o leitor encontrará em Murphy (2002) e Rogers & McClelland (2004) dois livros dedicados exclusivamente ao tema<sup>45</sup>.

---

<sup>44</sup> Minha tradução de: *Consistent co-occurrence of a set of properties across different objects. The concept is distinct from simple correlation in that it generally refers to the co-occurrence of more than two properties. For example, having wings, having feathers, having hollow bones and being able to fly all consistently co-occur in birds.*

<sup>45</sup> A propósito, o próprio título do livro de Murphy (*The Big Book of Concepts*) já é bastante sugestivo.

O que nos interessa a partir deste ponto é estabelecer uma relação direta entre a teorização de Rosch e a discussão feita nos capítulos vindouros.

### **3.4 Protótipos semânticos e aquisição de línguas naturais**

Para Taylor (2008:58-59), uma importante tentativa de se estabelecer a natureza da relação entre palavras (conhecimento semântico) e categorias sintáticas (construções sintáticas) é a pesquisa sobre aquisição de construções conduzida por Goldberg e colegas. Como destaca Taylor, a tarefa de Goldeberg e colegas é essencialmente onomasiológica, ou seja, parte das funções da língua para as suas formas.

Por exemplo, podemos pensar em verbos prototípicos que se tornam associados a uma função comunicativa específica. Estes verbos são altamente freqüentes e, por este motivo, terminam por otimizar a aprendizagem de categorias abstratas (construções sintáticas) por meio de processos de co-variação coerente com outros tipos verbais que surgem numa dada construção. O desenvolvimento construcional, por sua vez, segue os mesmos princípios observados em processos de desenvolvimento conceptual, ou seja, do genérico ao específico, cuja entrada se dá a partir de um protótipo altamente freqüente. Como sugerido por McClelland & Rogers (2003) e Rogers & McClelland (2004), o desenvolvimento conceptual acontece por meio de ciclos de diferenciação, nos quais se observa um nível básico como ponto de partida na estruturação do sistema. Finalmente, as categorias se desenvolvem por meio de processos de chunking, nos quais são criados *clusters* de atributos correlacionados (co-ocorrência de coisas ou co-variação coerente de atributos entre os membros de uma dada categoria).

Pesquisas recentes sugerem que a aprendizagem de categorias é otimizada por um insumo inicial de baixa variação centrado em exemplares prototípicos de alta freqüência. Isto permite que o aprendiz ‘perceba’ as principais características presentes em outros membros de uma dada categoria, promovendo generalizações mais rapidamente, como sustentam Goldberg & Casenhiser (2008) e Goldberg (2006).

Como veremos nos capítulos seguintes, há fortes evidências empíricas de que protótipos semânticos estejam na gênese da aprendizagem de construções sintáticas abstratas de uma língua (seja ela uma L1 ou uma L2). Quando pensamos em categorias lingüísticas, sabemos que, dentre as classes de palavras, o verbo é tido geralmente como a palavra mais importante para a compreensão do sentido global de uma frase (i.e., quem faz o que com quem), como lembra Goldberg (2006, 2003). E isto faz com que logo cedo a criança aprenda um padrão de estrutura argumental baseada no verbo. Veremos que o surgimento de esquemas (sintáticos) construcionais conhecidos por *ilhas verbais* parece ser resultante de processos de categorização nos moldes discutidos neste capítulo.

## 4. Construindo uma segunda língua<sup>46</sup>

*Importantly, when we focus in this way on language use and usage-based operations, we must perforce invoke in our acquisition theory a variety of cognitive and social-cognitive processes that originate from outside the domain of language per se (perception, memory, joint attention, intention-reading, categorization, analogy, and so on)*

Michael Tomasello

Como vimos na discussão do capítulo anterior, parece haver um contínuo entre os processos de representação lexical e o desenvolvimento sintático, ou seja, nossas categorias semânticas possibilitam o desenvolvimento de categorias sintáticas. Do concreto ao abstrato. A estrutura gramatical parece ser um subproduto emergente de configurações lexicais específicas (sociocognitivamente determinadas), como discutimos anteriormente. Tais processos possuem a mesma natureza e são regidos pelos mesmos mecanismos cognitivos de categorização e generalização.

De fato, as pesquisas recentes sobre aquisição de línguas (L1, L2 e línguas adicionais) estão convergindo para um consenso sobre certos aspectos cognitivos centrais ou universais envolvidos na aquisição de uma língua. Outrora separadas por idealizações acerca de uma Gramática Universal (UG) e períodos críticos na aprendizagem, as pesquisas em L1 e L2 estão convergindo e derrubando as divisões há tempos criadas por teorizações maniqueístas sobre a dissociabilidade entre uso e o conhecimento abstrato da estrutura lingüística (i.e., entre ‘*langue*’ e ‘*parole*’ saussuriana, entre ‘competência’ e ‘desempenho’ chomskianos, entre ‘aprender’ e ‘adquirir’ krashenianos).

---

<sup>46</sup> O crédito do feliz título deste capítulo é todo de Nick C. Ellis, que cunhou a expressão (inspirado em Tomasello, 2003) e organizou simpósio homônimo no 15º Congresso da AILA, em 2008.



Da mesma forma, a distinção tradicional entre léxico e gramática não é mais tida como um dogma, com resultados empíricos recentes de pesquisa sobre a aquisição e uso da linguagem, especialmente oriundos de abordagens de cunho construcionista e modelos de aquisição de línguas baseados no uso (*usage-based*) utilizando-se *corpora*, mostrando que a estrutura abstrata das línguas naturais é construída desde os menores pedaços (bastante concretos!) de uma língua. Finalmente, é aqui defendido, como destacado alhures, que são as estruturas lexicais fixas (conhecidas como fórmulas, expressões multi-palavras ou simplesmente *chunks*) que se constituem a gênese deste processo dinâmico de construção. Como vimos anteriormente, esta teorização em ASL tem sido proposta, principalmente, por Nick C. Ellis (2008, 2006, 2004, 2003, 2002a, 2002b, 2001).

#### 4.1 O léxico: de um dicionário a estados mentais<sup>47</sup>

O nosso léxico mental é um bom exemplo de “rede” estruturada nos moldes conexionistas já discutidos. Embora o próprio conceito de palavra seja algo altamente controverso, conforme aponta Singleton (1999:10-14), é certo que as palavras desempenham um papel muito mais importante no desenvolvimento e estruturação de uma língua do que muitos teóricos parecem admitir<sup>48</sup>. Há um verdadeiro abismo entre teorizações ortodoxas que vêem o léxico como um simples dicionário, do qual as palavras são escolhidas para o preenchimento de estruturas sintáticas no uso da linguagem (cf. PINKER, 1994), e reflexões conexionistas recentes que preconizam que palavras são “pistas” ou “estímulos” que provocam mudanças em estados/configurações mentais (cf. ELMAN, 2004). Discutirei mais detalhadamente sobre estas últimas teorizações, por natureza mais abstratas, mais adiante.

Como uma transição suave para a recente alternativa conexionista de Elman, começaremos com uma conceituação de léxico mental que pode ser caracterizada como

---

<sup>47</sup> A discussão feita neste item tem uma relação direta com os processos de categorização discutidos no capítulo 3 e, principalmente, com a proposta conexionista de McClelland & Rogers (2003) e Rogers & McClelland (2004). A concepção de léxico mental de Elman (2004) exibida mais adiante é um epítome da teorização conexionista sobre os processos de desenvolvimento conceptual.

<sup>48</sup> Para uma compreensão maior do papel das construções lexicais na estruturação de uma língua, destaco duas recentes reflexões elegantemente produzidas por Nick C. Ellis (2003, 2001).

um enorme avanço em relação às teorizações ortodoxas que vêem o léxico como um simples dicionário ou depósito de palavras, usadas, principalmente, a serviço de estruturas sintáticas.

Para Meara (1996), um léxico mental pode ser entendido a partir de duas dimensões básicas: tamanho e grau de conectividade. Ao adquirirmos novas palavras na L2 (pela exposição ao insumo) também adquirimos um conhecimento mais amplo sobre as palavras que já conhecemos. Mas o tamanho do vocabulário, por si só, não é um bom indicativo da competência em L2. À medida que o léxico cresce, o *tamanho* torna-se algo de menor importância. Meara enfatiza a importância do grau de organização deste vocabulário. É necessário entender o léxico de uma perspectiva sistêmica, levando em conta não apenas o *tamanho do vocabulário*, mas o *nível de profundidade* deste conhecimento. Isto requer que olhemos para a dimensão da *organização* (ou seja, algo que diferencie alguém que conhece apenas uma longa lista de palavras daquele outro aprendiz cujo vocabulário é estruturado em termos de vastas redes de associação entre as palavras). Tais associações podem ser paradigmáticas, sintagmáticas, situacionais, emocionais, mas todas contribuem, de maneiras diferentes, para o significado de uma palavra. Para Meara, tais tipos de *conexões* são o que precisamente distingue um verdadeiro *léxico mental* de uma simples lista de palavras (p.48).

*Estruturação*, portanto, é uma palavra chave na discussão sobre léxico em L2, sugerindo assim que o grau de estruturação lexical relaciona-se diretamente com o nível de proficiência do aprendiz. Não é que todas as palavras tenham conexões umas com as outras, mas simplesmente que o padrão de conexão precisa ser rico o suficiente para permitir o estabelecimento, de forma fácil e eficaz, de *conexões* entre partes diferentes do léxico. Desta forma, Meara mostra que é possível descrever um léxico em duas dimensões distintas, que nos possibilita distinguir, por exemplo, entre um vocabulário grande e desestruturado e um outro pequeno, mas altamente estruturado (veja que é o nível de estruturação que determina o desempenho do aprendiz na L2, como falado anteriormente). A noção sistêmica de léxico mental proposta por Meara difere radicalmente das teorizações tradicionais. Um léxico mental não pode ser entendido como coleções de itens ou palavras específicos, mas como um sistema (rede) de conexões altamente estruturado.

De uma perspectiva conexionista, ou emergentista, Elman (2004) propõe uma visão ainda mais dinâmica e funcional acerca da natureza da representação e funcionamento de um léxico mental. Elman sugere que palavras não são coisas que contém significados fixos armazenados em nossa memória de longo prazo, como nos leva a crer o conceito tradicional de léxico mental. Para Elman, uma palavra não carrega em si um significado, mas sim “pistas” para significados que, no fundo, se resumem a configurações ou estados mentais específicos. Neste sentido, palavras equivaleriam a estímulos sensoriais que atuam em nossos estados mentais, determinando, assim, o tipo de significado instanciado. Em um sistema desta natureza o tempo é um fator crucial. O desenvolvimento do sistema (no caso, nosso léxico mental) depende da *recorrência* de certos padrões e seus respectivos estímulos. Desta relação, categorias são criadas e generalizações e conexões, de base probabilística, são estabelecidas.

A teorização sistêmica de Meara acima discutida aproxima-se, mesmo que timidamente, dos *insights* conexionistas de Elman. Tamanho e grau de conectividade são dimensões cruciais para que categorias sejam estabelecidas e significados sejam continuamente (re)criados e refinados, possibilitando, por sua vez, que interpretações novas e mais complexas sejam geradas.

Para Elman (2004), as palavras (e seus respectivos significados) são regiões de ativação específicas ao longo do “espaço mental” de uma rede. Variações espaciais implicam em mudanças/distinções de significado, sendo tais espaços preenchidos hierarquicamente de forma esquemática/prototípica. Estas variações espaciais, é importante destacar, equivalem a pesos ou restrições de base probabilística em uma rede conexionista. Dependendo do estímulo (contexto), categorias, subcategorias e níveis extremamente sutis de significado são ativados. Assim, Elman (2004:302) sugere que as “propriedades de uma palavra são reveladas pelos efeitos que estas causam em estados mentais”. Tais efeitos, por sua vez, são regulados pelo contexto, como destacado acima.

A concepção de representação proposta por Elman redefine de maneira precisa a distinção tradicional entre *type* e *token* e os variados níveis de abstração lexemáticos, como discutido em Singleton (1999:10). Como vimos acima, cada sentido de uma

palavra resulta de uma configuração mental específica. Tomemos como exemplo a palavra *carro* na frase “o carro passou e quebrou o carro da criança”. Estas duas ocorrências de *carro*, ou seja, estes dois *tokens*, não serão, em termos mentais, os mesmos *types*. Os estados mentais serão diferentes.

Como destaca Elman (2004), cada configuração mental gerada pela palavra *carro* ocupa um determinado espaço mental (*a bounded region*) reservado aos membros deste lexema. Pequenas variações dentro desta área lexemática implicam em variações sutis de significado. Cada ocorrência da palavra *carro* produzirá, portanto, uma determinada configuração nesta micro-região a ele dedicada. Tal variação, por sua vez, reflete alterações e “instabilidades ao longo de diferentes eixos espaciais” (p.304), correspondentes a informações sobre número, papel temático (sujeito, objeto, etc.), flexão, etc. Um único lexema instanciado em contextos diferentes produzirá significados também diferentes.

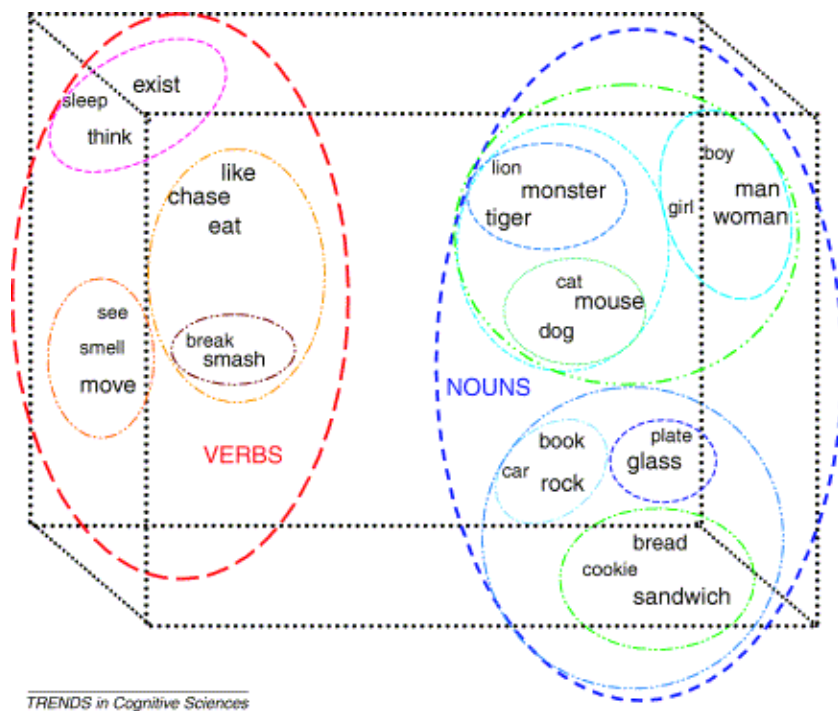
O mesmo tipo de processo acima descrito aplica-se aos mais variados tipos de categorias criadas, como verbos e substantivos por exemplo. Como as representações são esquemáticas, subcategorias são criadas.

Vejamos abaixo um exemplo corriqueiro para a categoria *substantivo* numa representação bidimensional padrão de duas subcategorias de elementos constitutivos (i.e., animados e inanimados), na qual a variação no tamanho das fontes indica relações hierárquicas:

[substantivos [animados [animais [rato, gato, cachorro] [monstro, leão, tigre] [humanos [mulher, garota,], [homem, garoto] ] ] ] [inanimados [comida [sanduíche, biscoito, pão] ] [quebráveis [prato, copo] ] ]].

A representação de uma dada categoria, como visto acima, também pode ser ilustrada em um espaço multidimensional. Neste último, os significados emergem em consequência de sua localização específica em uma dada região espacial, na qual

“diferentes dimensões podem ser alocadas enquanto outras são ignoradas” (p.304) dependendo do contexto. Veja a Figura 6:



**Figura 6** – Visualização esquemática, em 3D, do estado espacial multidimensional gerado pela camada oculta de uma rede conexionista do tipo SRN. O estado espacial é dividido em diferentes regiões que correspondem a categorias gramaticais e semânticas. Relações de aninhamento espacial (por ex., a categoria homem está contida na categoria animado que, por sua vez, está contida na categoria substantivo) refletem relações hierárquicas entre as categorias. In: Jeffrey L. Elman (2004), An alternative view of the mental lexicon. *Trends in Cognitive Science*. Vol 8, N.7, p.304.

Observe que esta última representação permite relações mais flexíveis, de natureza **heterárquicas**<sup>49</sup>, não apenas hierárquicas, ilustrando de forma mais precisa a noção de léxico como uma rede semântica altamente dinâmica.

<sup>49</sup> Segundo Von Goldammer (2003) o termo “heterarquia” é ainda pouco conhecido. Cunhado pelo neurofisiologista e ciberneticista Warren McCulloch, o termo reflete a noção de relações sistêmicas (ou processos) que ocorrem em espaços multidimensionalmente estruturados, ou seja, um complemento à organização hierárquica tradicional totalmente compatível com fenômenos observados no desenvolvimento de redes neurais.

Resumindo, vimos, anteriormente, aspectos básicos de funcionamento de modelos conexionistas e as novas maneiras de compreendermos o desenvolvimento e a estruturação de sistemas complexos possibilitadas por estes modelos. Também vimos uma nova forma de entender como se dá a representação do conhecimento, partindo do pressuposto de que todo conhecimento é, por natureza, distribuído e emergente. Em seguida, numa rápida discussão, abordamos diferentes concepções sobre léxico mental, desde teorizações ortodoxas que concebem o léxico como um simples depósito de palavras na memória de longo prazo, passando por teorizações mais sistêmicas e dinâmicas (porém limitadas em termos explicativos) e culminando na concepção de léxico de Elman acima discutida, para o qual as palavras, em si, não possuem significados, mas servem de pistas, estímulos, para determinados padrões de ativação em espaços mentais altamente dinâmicos.

Ferreira Junior (2005) faz uma analogia entre a estruturação de um léxico mental nos moldes acima descritos e o desenvolvimento de um grupo de discussão na Internet. A discussão é uma tentativa de ilustrar como se dá a emergência de certos fenômenos ao longo do desenvolvimento de sistemas auto-estruturados.

Então, o que esta nova visão de lexico mental tem a ver com o desenvolvimento de categorias abstratas (i.e., aprendizagem da ‘gramática’) durante a aquisição de uma língua? Bom, este é o ‘nível construcional’ no qual e a partir do qual muitos acreditam que algo interessante começa a acontecer (cf. MARCHMAN & THAL, 2005; e a maioria dos outros autores citados ao longo desta tese). É neste nível que efeitos de frequência de palavras (tanto *types* como *tokens*) exercem um papel decisivo na aquisição de categorias semânticas que possibilitam, pois, o desenvolvimento de categorias sintáticas. Esta relação entre léxico e gramática é o foco das investigações empíricas relatadas e discutidas nos capítulos 5, 6 e 7 desta tese.

#### **4.2 O uso de corpora na investigação sobre a aprendizagem de línguas**

O uso de *corpora* na investigação sobre aprendizagem/ensino de L2 não é coisa nova. Desde o final do século XIX já havia preocupações pedagógicas neste sentido. Segundo

McEnery & Wilson (1997), alguns lingüistas e psicólogos já elaboravam, naquela época, listas de vocabulário para o ensino de L2 com base em corpora. Edward Thorndike, psicólogo americano que investigou de forma pioneira certos processos de aprendizagem, deixou importantes contribuições nesta área, conforme destaca Gregory (1998: 775-776).

No entanto, de acordo com McEnery & Wilson (1997), a partir do final da década de 50 (mais especificamente no ano de 1957, ano de lançamento de *Estruturas Sintáticas*) as teorizações de Chomsky fizeram com que o uso de corpora e dados “reais” e “autênticos” de uso de uma língua fossem rejeitados (diante da famosa dicotomia competência *versus* desempenho lingüísticos). Tudo relacionado ao desempenho (*performance*) foi tratado como inútil para os estudos sobre a aquisição linguagem (i.e., o estudo da competência lingüística). Assim, os corpora perderam o valor, pois nada mais são do que exemplos autênticos de uso de uma língua (ou seja, estabeleceu-se um paradigma racionalista nos estudos lingüísticos, em oposição à tradição empirista anterior).

Recentemente, com o desenvolvimento tecnológico e o uso de computadores com uma capacidade de memória cada vez maior, renovou-se o interesse pelos corpora em lingüística, lingüística aplicada, e outras disciplinas. Estabeleceu-se, assim, a Lingüística de Corpus, que, de acordo com Gries (2008), pode ser definida como um conjunto de procedimentos (i.e., método) que envolve a coleta e análise de elementos lingüísticos dos corpora. Por meio do uso de arquivos digitais contendo textos escritos e/ou transcrições minuciosas de linguagem oral, de crianças e adultos (i.e., *corpora*), auxiliados por computadores com uma capacidade cada vez maior de processamento e armazenamento destas informações, está se tornando cada vez mais possível investigar as regularidades estatísticas da estrutura lingüística, principalmente com o desenvolvimento de ferramentas (programas) de concordância, ou concordanceadores, que possibilitam explorar cada vez mais detalhadamente estas grandes coleções eletrônicas de exemplos de língua em uso (tanto escrita como oral). Com isto, estamos observando o (re)estabelecimento gradual do paradigma empirista nos estudos lingüísticos.

De acordo com Gries (2008), a própria natureza da Lingüística de Corpus (que se preocupa essencialmente com os aspectos de frequência e co-ocorrência de palavras e estruturas) já produz, naturalmente, uma grande afinidade com as investigações conduzidas no âmbito da lingüística cognitiva (cujo principal foco de investigação é a língua em uso), fornecendo, assim, os dados empíricos e a matéria prima que compõem o cerne das investigações conduzidas pela lingüística cognitiva.

No tocante à pesquisa no âmbito dos estudos em ASL, os corpora podem ser muito úteis. Gries (2008:413) lista uma série de perguntas que a investigação por meio do uso de corpora pode vir a responder: como se estrutura o insumo lingüístico? Como se estrutura a L1 do aprendiz? Como se estrutura a L2? Qual a diferença entre a estruturação da L1 e da L2? Como se estrutura a interlíngua? Quais tipos de erros os aprendizes cometem? Como sugere Sardinha (2000), o uso de corpora na investigação dos fenômenos em ASL tem propiciado uma desmistificação de muitos pontos acerca da sua natureza (i.e., sua estruturação e sua aprendizagem).

Como vemos na discussão teórica que permeia esta tese, alguns pontos obscuros e cruciais acerca da natureza dos processos de aquisição de línguas já começaram a ser, pelo menos parcialmente, clareados. A investigação sobre o processo de aquisição de algumas construções por meio da análise de corpora em ASL conduzida no âmbito desta tese é mais uma tentativa de lançar luzes sobre algumas questões importantes dos processos envolvidos na aprendizagem de uma L2 e, acredito, toca em alguns pontos relevantes de pesquisa levantados por Gries acima.

### **4.3 Construções**

As *Construções*, mapeamentos recorrentes e entrincheirados entre forma e função, constituem-se unidades básicas em todos os níveis de representação lingüística. Como destacado por Goldberg (2006; 2003), construções são aprendidas (ou induzidas) devido a nossa exposição ao insumo lingüístico e condicionadas por restrições de processamento cognitivo globais e por aspectos pragmáticos. Alguns tipos de construções da língua inglesa, com base em Goldberg (2003:220), são fornecidos na Tabela 1 a título de ilustração para a nossa discussão:



| <b>Construção</b>                           | <b>Forma/Exemplo</b>  | <b>Função</b>   |
|---|---|---|
| Raízes de palavras                          | e.g., <i>book, dog, or</i>  |   |
| Expressões semi-idiomáticas.                | e.g., <i>jog &lt;someone's&gt; memory</i>   |   |
| Construções co-variacionais ou condicionais | Forma: The Xer the Yer (e.g., <i>the more you think about it, the less you understand</i> ) | Significado: ligar variáveis independentes e dependentes. |
| Construção bitransitiva                     | Forma: Subj [V Obj1 Obj2] (e.g., <i>He baked her a carrot cake.</i> )                       | Significado: transferência (pretendida ou realizada)      |

**Tabela 1**

Exemplos de construções e suas respectivas variações em tamanho e complexidade

As construções, desde seus ‘pedaços’ menores como as palavras, são simbólicas. As coisas no mundo se repetem. Como costumava dizer uma eloquente e conhecida ex-senadora conterrânea minha, citando o ministro da propaganda da Alemanha nazista Joseph Goebbels, “uma mentira repetida mil vezes torna-se verdade!”. Pois, como uma mentira repetida *ad nauseam*, as formas da língua se consolidam e tornam-se entrincheiradas com os seus usos. Assim é a natureza das construções. Ou, como bem destaca Tomasello (2003:99):

Quando as pessoas usam repetidamente os mesmíssimos símbolos linguísticos para produzir enunciados entre si em situações “semelhantes”, o que possivelmente surge, com o passar do tempo, é um padrão de uso da língua esquematizado na mente dos falantes como um certo tipo de categoria ou construção (grifo no original).<sup>50</sup>

<sup>50</sup> Minha tradução de: *When people repeatedly use the same particular and concrete linguistic symbols to make utterances to one another in “similar” situations, what may emerge over time is a pattern of language use, schematized in the mind of users as one or another kind of linguistic category or construction.*

A construção bitransitiva X VERBou Y o Z indica transferência (veja exemplo a seguir). Cada ‘pedaço’ (do lexical ao sintático) desta construção possui um significado, significa algo. O aspecto criativo e o caráter infinito da linguagem resultam das possibilidades ilimitadas de (re)combinações de pedaços (chunks) de linguagem em nosso ‘estoque mental’ de construções (i.e., nosso *constructicon*). As construções fundem-se livremente com outras construções para que possamos realizar nossos propósitos comunicativos de expressão de significados diferentes e novos. Como exemplificado em Goldberg (2003:221), a construção bitransitiva abaixo envolve a combinação (ou chunking) de outros seis diferentes tipos de construções. Observe:

- a) [*What did Liza buy the child?*]
- b)
  1. Construções → *Liza, buy, the, child, what, did* (i.e., palavras)
  2. Construção bitransitiva → (what buy the child)
  3. Construção interrogativa → (what did Liza buy the child?)
  4. Construção de inversão do sujeito-auxiliar → (did Liza)
  5. Construção de sintagma verbal → (buy the child)
  6. Construção de sintagma nominal → (what, Liza, the child)

Desta perspectiva, o que seria a nossa tão almejada ‘competência gramatical’? Podemos dizer que tal ‘competência’ pode ser compreendida como um conjunto de construções - um *constructicon* - que construímos gradualmente ao longo de nosso viver lingüístico. Podemos conceber nosso *constructicon*, também, como um gradiente que vai **do concreto ao abstrato** (i.e., que varia em especificidade e complexidade), desde morfemas, passando por palavras, palavras compostas, expressões idiomáticas fixas e semi-fixas, e, finalmente, padrões sintáticos totalmente abstratos. Esta teorização, cada vez mais presente em ambas as áreas de pesquisa citadas, convergem e embasam empiricamente a hipótese de que a linguagem seja inteiramente ‘construída’. Os princípios psicológicos que subjazem tais mecanismos de construção são os mesmos que guiam a aprendizagem de categorias, como discutido no capítulo 3, e que, por sua vez, estão presentes nos processos cognitivos da aprendizagem humana em todos os domínios. Esta é a filosofia subjacente às recentes investigações teóricas e empíricas

conduzidas, principalmente, por Goldberg & Casenhiser (2008), Goldberg (2006, 2003), Goldberg, Casenhiser & Sethuraman (2004), Bybee (2008), Ellis (2007, 2006), Lieven & Tomasello (2008), Gries & Wulff (2005) e Tomasello (2003).

Desta forma, como discutido anteriormente, começamos a construir uma língua com palavras: desde as holófrases (ou seja, enunciados de uma única palavra como *mamãe*, *vê*, *pai*, etc.) passando, por volta dos 18 meses, a enunciados multi-palavras (por exemplo, frases a partir de duas palavras como *bola mesa*, etc). Holófrases e frases multi-palavras são compostas apenas de **pedaços concretos de uma língua** (não há categorias). Então, por volta dos 18 meses, os enunciados multi-palavras das crianças passam a exibir um **padrão mais sistemático**. Tomasello (2003:114-117) chama tais construções de **esquemas pivô** (há uma palavra ou frase que parece estruturar o enunciado), por exemplo: mais leite (*more milk*), mais uvas (*more grapes*), mais suco (*more juice*); pai foi (*Dad gone*), carro foi (*car gone*), neném foi (*baby gone*), etc. Finalmente, enunciados calcados em itens (*item-based*) vão além dos esquemas-pivô ao possuírem marcadores sintáticos como uma parte integrante da construção (ilhas verbais ou *verb islands*). Neste processo, a frequência de ocorrência exerce um papel importante (veja mais adiante sobre os tipos de frequência *token* e *type*).

Em última instância, os aprendizes parecem construir associações entre significado/função e forma (i.e., construções) em dois níveis básicos: 1) categorias centradas em verbos (ou ilha verbais, i.e., <agente> coloca <coisa> <local>) e 2) generalizações sobre verbos específicos, desenvolvendo, assim, um conhecimento acerca dos padrões da estrutura argumental numa dada língua (como vimos no exemplo de Goldberg ilustrado anteriormente sobre as construções usadas em uma construção bitransitiva na língua inglesa).

#### 4.4 Construções na aprendizagem de L1

Quase toda pesquisa empírica guiada pelo referencial teórico da Gramática das Construções enfoca a aquisição e o uso de L1 por crianças e adultos (i.e., investiga o desenvolvimento e o funcionamento do sistema lingüístico de falantes nativos). O grupo

liderado por Adele Goldberg, do departamento de psicologia da Universidade de Princeton, é um dos pioneiros neste tipo de investigação empírica. A exemplo dos próprios verbos construcionais prototípicos *go*, *put* e *give*, que veremos mais adiante, a literatura sobre L1 ‘abocanha’ a maior parte do que sabemos sobre o papel das construções no desenvolvimento e uso de línguas naturais.

Goldberg, Casenhiser & Sethuraman (2004) conduziram um estudo com resultados interessantes sobre a aprendizagem de L1 de uma perspectiva construcionista. O estudo mostra correlações importantes entre padrões lingüísticos formais e o significado destes (veja Tabela 1). Goldberg e colegas analisaram o corpus BATES, que integra a base de dados do Projeto CHILDES, com transcrições de 27 crianças (13 garotos e 14 garotas) entre 20-28 meses de idade. As transcrições possuem duração de quinze minutos e incluem o uso da L1 em três tipos de atividades: brincadeiras (*free play*), contar histórias (*story-telling*) e hora do lanche (*snack time*).

Os pesquisadores também analisaram as conversas transcritas de um grupo de 15 mães com o propósito de identificar os verbos mais frequentes nos três tipos de construção analisados (veja Tabela 2 com as construções investigadas). Como no corpus BATES só havia seis ocorrências da construção VOO, Goldberg e colegas analisaram dois outros *corpora* com diálogos de cinco outras crianças com o propósito de identificar mais construções VOO<sup>51</sup>.

| Forma/Exemplo   | Função/Significado            | Construção (Rótulo)         |
|---|-------------------------------|-----------------------------|
| 1. Subj V Obl path/loc<br><i>e.g. She get in her bed</i>        | X moves Y path/loc            | Movimento intransitivo (VL) |
| 2. Subj V Obj Obl path/loc<br><i>e.g., Pour some milk in it</i> | X causes Y to move Z path/loc | Movimento causado (VOL)     |
| 3. Subj V Obj Obj2<br><i>e.g. She faxed him a letter</i>        | X causes Y to receive Z       | Bitransitivo (VOO)          |

**Tabela 2**  
Exemplos de construções VL, VOL e VOO analisadas

<sup>51</sup> Os outros dois corpora longitudinais analisados em busca de construções VOO foram o corpus Brown (1973), com as crianças Adam, Eve e Sarah, e o corpus McWhinney (1995), com dados das crianças Ross e Mark. Goldberg e colegas apontam que a escassez de construções VOO deve-se ao fato de que tais construções estão começando a ser aprendidas aos 28 meses.

Os resultados obtidos na análise dos dados das crianças fornecem evidências de como as correspondências entre forma e significado nas construções envolvendo deslocamento intransitivo (VL), deslocamento causado (VOL) e bitransitivo (VOO) possuem sua gênese em categorias centradas em um verbo ou o que Tomasello (2000; 2003) chama de hipótese de ‘ilhas verbais’ (*verb island hypothesis*). Goldberg e colegas identificaram o mesmo padrão de desenvolvimento nestas associações entre forma e significado; basicamente, elas possuem as mesmas características apontadas por Tomasello (2000; 2003), já discutidas neste manuscrito. Ou seja, inicialmente as crianças produzem, de maneira bastante conservadora, padrões sintáticos em forma de ‘ilhas verbais’ (um único verbo de alta frequência é quase sempre usado com um padrão construcional específico). Este verbo prototípico ‘abocanha’ a maior parte ou está presente na maioria das vezes que uma dada construção é usada. Os resultados da análise de Goldberg e colegas mostram que as construções VL, VOL e VOO são totalmente dominadas pelos verbos altamente frequentes *go*, *put* e *give*, respectivamente.

Os resultados da análise dos dados das mães também mostram a preponderância de um único verbo sobre os demais usados numa construção específica. Veja Tabela 3 com a síntese dos resultados:

| <b>Construção</b>                | <b>Crianças<br/>(percentagem de tokens)</b> | <b>Mães<br/>(percentagem de tokens)</b> |
|----------------------------------|---|---|
| 1. Subj V Obl path/loc (VL)      | 54% go (121/224)                            | 39% go (136/353)                        |
| 2. Subj V Obj Obl path/loc (VOL) | 31% put (16/51)                             | 38% put (99/259)                        |
| 3. Subj V Obj Obj2 (VOO)         | Adam - 53% give (59/112)                    | 20% give (11/54)                        |
|                                  | Eve - 36% give (4/11)                       |   |
|                                  | Sarah - 43% give (29/67)                    |   |
|                                  | Ross - 43% give (69/160)                    |   |
|                                  | Mark - 32% tell (11/34)                     |   |
|                                  | 29% give (10/34)                            |   |

**Tabela 3**

Resumos dos resultados encontrados nos dados das crianças e das mães

Finalmente, com o tempo, outros verbos vão sendo gradualmente incorporados no grande processo de generalização envolvido na aprendizagem de uma língua, possibilitando a formação, assim, de um conhecimento acerca dos padrões da estrutura argumental da linguagem, conforme sustentam Goldberg (2006, 2003), Goldberg, Casenhiser & Sethuraman (2004) e Tomasello (2003, 2000).

#### 4.5 Construções na aprendizagem de L2

Como comentado no item anterior, quase toda pesquisa empírica sobre construções tem sido conduzida no âmbito tanto da aprendizagem como do uso de uma L1, seja com crianças ou adultos. O trabalho de Gries & Wulff (2005) é praticamente um dos primeiros a fornecer dados empíricos sobre a validade das construções em processos de ASL. Influenciada pelos achados em L1, a pesquisa investigou a validade das construções no processo de ASL por meio de dois experimentos psicolinguísticos com o propósito de investigar: a) *priming* sintático (i.e., a tendência que temos de repetir estruturas sintáticas ouvidas anteriormente), por meio de um estudo de formação de frases; b) certas preferências construcionais utilizando um dado verbo (i.e., o uso de protótipos verbais, como visto no estudo de GOLDBERG, CASENHISER & SETHURAMAN, 2004) com as construções VL, VOL & VOO); e c) tendências observadas em atividades de classificação de frases com base em semelhança semântica (i.e., a categorização de frases com base em padrões construcionais ao invés de uma classificação meramente baseada em verbos).

O Experimento 1 consistiu num estudo transversal (*cross-sectional*) envolvendo 64 alemães, estudantes avançados de inglês como língua estrangeira (LE) (a média de estudo da língua inglesa era de 11.1 anos). Os sujeitos desconheciam os propósitos do experimento do qual estavam participando. Os resultados mostram efeitos de *priming* semelhantes àqueles verificados em falantes nativos da língua inglesa. A correlação destes resultados de efeitos de *priming* com um estudo prévio relatado em Gries & Stefanowitsch (2004), que revelou a tendência de um verbo ‘abre-alas’ ou ‘*pathbreaking*’ ser usado na maior parte das construções analisadas (construções bitransitivas e dativas preposicionadas), aponta para uma forte correlação entre a

presença de protótipos verbais e *priming* sintático, i.e., “verbos que estão associados a uma construção específica no corpus são mais prováveis de serem usados junto com a construção”, como concluem Gries & Wulff (2005:188). Isto parece demonstrar de maneira contundente que no desenvolvimento de construções (pelo menos no caso dos aprendizes avançados) o acesso a padrões construcionais é verbo-dependente. Estes resultados são plenamente consoantes com os resultados empíricos verificados na aprendizagem e uso de L1, como vimos anteriormente.

O experimento 2 consistiu numa atividade de classificação de dezesseis frases com base em semelhança semântica<sup>52</sup>. Este estudo transversal envolveu 22 falantes nativos de alemão, todos eles falantes de língua inglesa em nível avançado (a média de estudo da língua inglesa era de 11.1 anos). Como no experimento anterior, os sujeitos também desconheciam os propósitos do estudo do qual estavam participando. As construções bitransitivas e as dativas preposicionadas usadas no primeiro experimento foram testadas. Os resultados mostram que os agrupamentos válidos de vinte dos 22 sujeitos pesquisados são feitos com base na semelhança (implícita) entre as frases, ao invés de uma simples classificação baseada nos verbos destas frases. Em outras palavras, este resultado indica que esquemas construcionais abstratos (i.e., construções resultantes do pareamento entre significado e estruturas sintáticas) são determinantes e mais confiáveis na percepção de semelhança entre as frases do que uma classificação baseada somente em verbos (ou seja, pareamentos identificados somente pelos verbos)<sup>53</sup>.

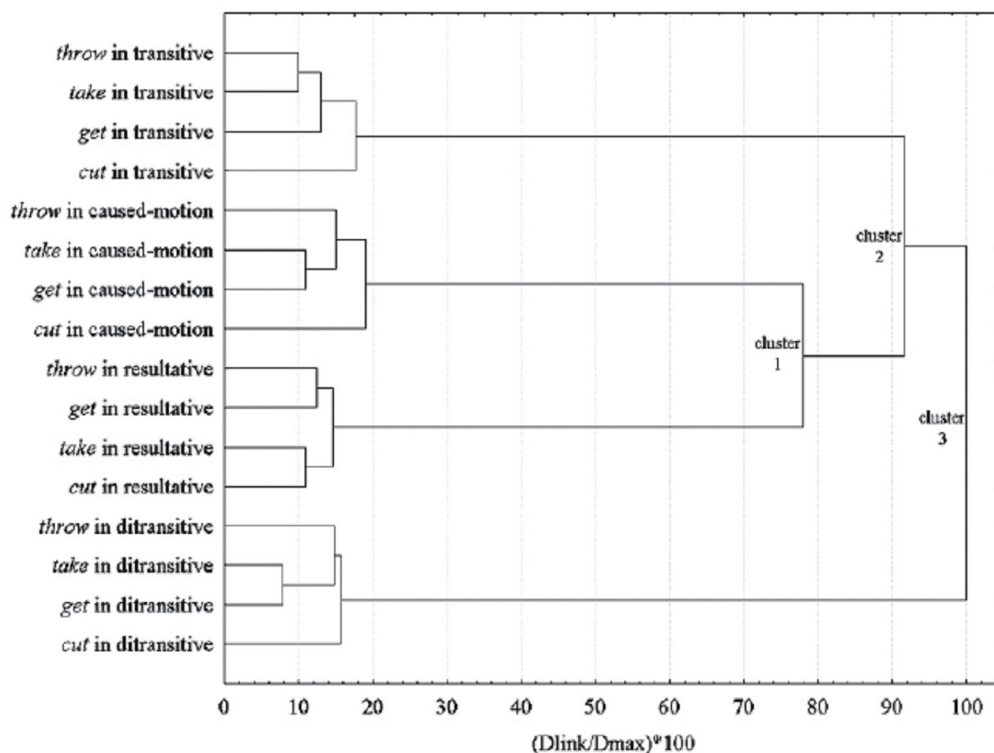
Como comentado anteriormente, estes resultados são consoantes com os dados empíricos verificados na pesquisa em L1, e parecem demonstrar, portanto, que as construções passam a ser pistas mais confiáveis do que os verbos no tocante ao sentido global da frase. Ainda, o resultado da classificação feita pelos sujeitos revela um aninhamento hierárquico (*hierarchical clustering*) das frases, revelando, com isto, as ondas de diferenciação pelas quais as construções passam desde os padrões iniciais

---

<sup>52</sup> Veja a lista de frases usadas no experimento na tabela 5, no Anexo A desta tese.

<sup>53</sup> Os resultados acima encontram ressonância não apenas em outros estudos guiados pela teorização construcionista, mas também em pesquisas envolvendo a aprendizagem de gramáticas artificiais por parte de bebês, conduzidas por Gomez & Gerken (2000; 1999) e outros. Estes estudos apontam, entre outras coisas, para o papel da atenção nos processos de aprendizagem. Os processos de atenção são o que levam à abstração do insumo em dois diferentes níveis indispensáveis para que haja a aprendizagem da sintaxe: abstração baseada em padrões e abstração baseada em categorias. Veja páginas 66 e 67 desta tese.

baseados em verbos, como se observa na produção de crianças e aprendizes adultos pouco proficientes numa L2 (veja a Figura 7).



**Figura 7**

Diagrama horizontal com os resultados do experimento de classificação de frases. In Gries, S., & Wulff, S. (2005). Do foreign language learners also have constructions? Evidence from priming, sorting, and corpora. *Annual Review of Cognitive Linguistics*, 3, p. 193.

É importante destacar a semelhança entre as teorizações sobre o processo gradual de diferenciação exibido no desenvolvimento de categorias semânticas, como sugerido por McClelland & Rogers (2003) e Rogers & McClelland (2004) e discutido no capítulo 3 desta tese, com o processo acima discutido de desenvolvimento de construções abstratas (ou esquemas construcionais). Na verdade, trata-se de um mesmo processo: categorização. Como sabemos, para que haja o desenvolvimento de categorias é necessária a existência de um protótipo (i.e., um *token* de alta frequência) que ocorre em uma dada categoria mais frequentemente. Novos itens (verbos ou simplesmente coisas) são gradualmente incorporados ao uso desta categoria com o passar do tempo (um



processo resultante de uma covariação coerente de atributos destes itens), o que provoca um novo ciclo de desenvolvimento e propicia uma “desconstrução” e uma re-análise dos itens constituintes de uma dada categoria, o que, por sua vez, permite o surgimento de novas categorias.

Por exemplo, como vimos na Figura 7, não são mais os exemplares verbais que promovem o aninhamento hierárquico das frases acima, mas a semelhança entre o tipo de construção (pareamento entre forma e significado/função) destes. No início, quando o sistema não tinha a complexidade construcional exibida, certamente o aninhamento (i.e., a categorização) ocorria com base na semelhança semântica dos verbos. Mas, como vimos anteriormente, os verbos passam a ser não muito confiáveis quando os usos que fazemos com a linguagem tornam-se mais complexos. Voltaremos a este ponto mais adiante.

#### 4.6 Como adquirimos o conhecimento gramatical?

Bybee (2008:220) sugere que nosso conhecimento gramatical resume-se, em última instância, a um conhecimento procedimental:

O conhecimento procedimental está imbricado com eventos neuromotores e consiste num conhecimento de como fazer algo, semelhante a dirigir um carro ou abrir massa de panqueca. Embora estas ações sejam aprendidas por meio do domínio de uma seqüência de ações, com a prática tais ações são embrulhadas em uma única ação. Uma vez automatizadas, torna-se difícil desembrulha-las novamente<sup>54</sup>

Esta organização cognitiva da nossa experiência lingüística torna-se possível por meio de processos de **chunking**. Processos de *chunking* possibilitam a formação de novas estruturas cognitivas, mais compactas, complexas e funcionalmente mais eficazes, como sugerido por Bybee na citação acima. Estes mecanismos de memória operam lentamente e guiam processos de categorização, generalização e diferenciação do conhecimento,

---

<sup>54</sup> Minha tradução de: *Procedural knowledge is bound up with neuromotor events and is knowledge of how to do something, such as drive a car or stir pancake batter. While these kinds of actions are learned through the mastery of a sequence of actions, with practice they are repackaged into a single action. Once automatized, it is difficult to unpack them again (p.220).*

como discutimos anteriormente. Tais seqüências de ações, que são neuromotoras por natureza, possibilitam a criação de construções e levam à gramatização (*grammaticization*), expressão cunhada por Bybee. Desta forma, a gramatização é um fenômeno da memória procedimental.

Nesta mesma linha de argumentação, Ellis (2001, 2003) sugere que tanto os nativos como os não-nativos proficientes de uma língua conhecem uma enorme quantidade de seqüências lingüísticas em diferentes níveis (fonológico, morfológico, sintático); ou seja, é por meio da percepção inconsciente das regularidades presentes no insumo lingüístico que vamos aprendendo, da mesma forma, as probabilidades seqüenciais de uma dada língua nos níveis da letra, palavra e frase. Ellis (2006), em sua elegante explicação dos processos de aquisição de L2 recentemente desenvolvida na forma de um Associative-Cognitive CREED, sugere que os processos de chunking são processos básicos na aprendizagem de línguas.

É por isto que, de uma perspectiva construcionista, a competência gramatical pode ser entendida também como sendo seqüências de padrões cristalizados, entrincheirados. E estes padrões são as próprias construções, i.e., pareamentos entrincheirados de forma e significados/funções de vários graus de especificidade e complexidade, como observado anteriormente.

#### **4.7 Verbos abre-alas (*pathbreaking*)**

Como comentado anteriormente, a análise de *corpora* tem permitido a investigação da natureza e desenvolvimento das regularidades estatísticas presentes na estrutura lingüística. Estudos recentes demonstram que alguns verbos (altamente freqüentes) parecem promover o desenvolvimento de construções da estrutura argumental da linguagem (construções abstratas). Eles são chamados de *verbos prototípicos* ou verbos ‘abre-alas’ em Goldberg (2006) e Goldberg, Casenhiser & Sethuraman (2004), pois se constituem na melhor dica para o significado geral de um enunciado, i.e., quem faz o que para quem.

Por exemplo, pesquisas lideradas por Goldberg e colegas (2006, 2004) sobre a aquisição de L1 fornecem importantes *insights* de como as correspondências entre significado e forma envolvendo as construções *intransitive motion (VL)*, *caused-motion (VOL)* e *bitransitivas (VOO)* têm sua gênese em categorias centradas em um verbo. Este é o mesmo fenômeno que Tomasello (2003) chama de ‘ilhas verbais’ (*verb island*). O padrão de desenvolvimento encontrado nestas associações segue basicamente dois estágios: primeiro as crianças produzem, de maneira bastante conservadora, padrões sintáticos em forma de ‘ilhas verbais’ (um único verbo de alta frequência é quase sempre usado com um padrão construcional específico i.e., *go* para construções VL, *put* para as VOL e *give* para as construções VOO). Com o tempo, outros verbos (semanticamente similares) são incorporados à categoria por meio de processos de generalização, categorizados sempre a partir daqueles verbos prototípicos iniciais, ampliando e enriquecendo, assim, os padrões sintáticos exibidos na estrutura argumental de uma língua.

#### **4.8 Lei de Zipf e o papel das frequências *type* e *token***

Como destacado por Goldberg (2006:76) e Crystal (2003:87), um dos primeiros pesquisadores a investigar a natureza estatística (probabilística) da linguagem foi o filólogo americano George Zipf, em 1935. Zipf sugeriu que palavras altamente frequentes seriam responsáveis pela maior parte dos *tokens* em um dado texto. E como vimos nos estudos relatados acima, os resultados de frequência verbal nas construções VL, VOL e VOO parecem confirmar o que passou a ser conhecido como Lei de Zipf (i.e., há uma preponderância de verbos altamente frequentes em cada uma das construções investigadas). Esta característica do insumo lingüístico ajuda muito o aprendiz durante a comunicação, pois, por serem mais frequentes, estes verbos se tornam uma importante dica na busca e identificação de sentido no uso de uma língua, sentido (significado) este que é influenciado a todo instante por contextos altamente dinâmicos.

De acordo com Lieven & Tomasello (2008), alguns estudos em L1 mostram que a frequência desempenha um papel crucial na aprendizagem de línguas. Por exemplo,

estudos feitos com ordem truncada de palavras (*weird-word order experiments*), nos quais ordenamentos não canônicos são apresentados, revelam que as crianças tendem a corrigir tais ordenamentos, colocando as palavras numa ordem canônica mais facilmente quando há verbos conhecidos do que com verbos novos (desconhecidos). Mas não é assim tão fácil quanto parece. Existem dois tipos diferentes de frequência e cada uma produz efeitos diferentes em nosso sistema lingüístico. Um tipo, a frequência de *tokens*, provoca o ‘entrincheiramento’ da compreensão e uso de pedaços concretos de língua – itens, fórmulas e expressões multi-palavras – como um *chunk* não analisável.

A frequência de *types*, por outro lado, promove generalizações, i.e., tal frequência possibilita a percepção, compreensão e aprendizagem da estrutura interna de um enunciado, ou seja, possibilita que o aprendiz perceba que itens diferentes podem exercer a mesma função no contexto de uma mesma construção (veja os exemplos dados anteriormente). A frequência de *types* é condição necessária para que estruturas abstratas da língua (gramática) se desenvolvam, ou, como concluem Lieven & Tomasello (2008:174)

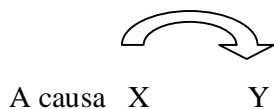
Portanto, a diferença entre a frequência de *tokens* e a frequência de *types* é entre entrincheirar palavras ou sentenças específicas e criar *slots* nos quais possam ocorrer uma gama maior de palavras e sentenças<sup>55</sup>

Finalmente, esta característica probabilística e os efeitos de frequência observados na distribuição do insumo lingüístico facilita a aprendizagem das estruturas abstratas da linguagem (i.e., como dito anteriormente, este verbo/*token* altamente freqüente passa a ser um protótipo semântico por meio do qual outros verbos/*types* serão associados). O contato cada vez maior com um insumo lingüístico altamente variado (e a sua gradual percepção) faz com que o aprendiz perceba a existência de certos padrões na língua (esquemas semânticos) invariáveis, que auxiliam na realização de funções específicas, apesar dos itens que compõem estes esquemas semânticos serem altamente variáveis e cambiáveis.

---

<sup>55</sup> Minha tradução de: *So the difference between token and type frequency is between entrenching specific words or phrases and creating slots in which a range of words or phrases can occur.*

Por exemplo, todas as frases abaixo expressam a idéia de deslocamento ou movimento causado (*caused motion*) e são exemplos de construção VOL, ou seja, descrevem como alguém ou algo (A) provoca o movimento ou deslocamento de alguém ou algo (X) para um novo lugar ou em uma nova direção (Y), veja primeiro a ilustração abaixo:



Os exemplos em *a-f* abaixo expressam o mesmo esquema sintático ilustrado acima (i.e., o mesmo sentido de movimento causado), por mais concretos ou abstratos que pareçam. Veja:

- a) O menino colocou o dedo na boca
- b) O rapaz atirou o gato para cima do telhado
- c) O professor arrotou teoria na cabeça dos alunos
- d) Esta tese joga meus neurônios pra lá e pra cá
- e) O bibliotecário entulhou outra tese dentro de uma estante empoeirada
- f) Felizmente, o PosLin publicou a tese na Internet

O desenvolvimento de esquemas sintáticos ocorre em decorrência do fato de que alguns verbos possuem baixa validade de previsão do significado global de um enunciado. (Os diferentes contextos de uso tornam tais verbos, muitas vezes, bastante ambíguos). Portanto, os aprendizes desenvolverão *construções* mais confiáveis (embora mais complexas e abstratas) que os auxiliem na interpretação do significado linguístico. Por exemplo, nas frases 1 e 2 abaixo fornecidas por Goldberg (2006:106), quando o verbo *get* é usado na construção VOL ele expressa o sentido de movimento causado, ao passo que quando é usado na construção VOO o sentido passa a ser o de transferência:

- 1) *Pat got the ball over the fence* (VOL → movimento causado)
- 2) *Pat got Bob a cake* (VOO → transferência)

Como vimos, as construções surgem diante da necessidade de resolvermos, ou pelo menos evitarmos, problemas durante a comunicação. Ellis (no prelo) destaca que “a aprendizagem de uma língua é uma estimativa baseada numa dada amostra” (*language learning is estimation from sample*). Ou seja, à medida que tal sistema cresce (dos *tokens* aos *types* aos padrões sintáticos abstratos) e vai se tornando, assim, mais e mais complexo, o aprendiz percebe (inconscientemente) a necessidade de evitar problemas de comunicação; há a necessidade de se perceber e entender o mundo que o cerca, bem como de se fazer entender de uma forma a mais eficiente, clara e rápida possível. Como destacado acima, os processos de *chunking* se constituem numa forma bastante eficaz de estruturação de sistemas complexos, sintetizando informações por meio da criação de estruturas recursivas que produzem, por sua vez, processos de subordinação (*embedding processes*), possibilitando, com isto, que recursos de memória e atenção sejam otimizados, liberando-os para que possam ‘captar’ outros aspectos e/ou coisas no/do mundo (ELLIS, no prelo).

#### **4.9 Em busca da síntese: neurobiologia, connexionismo e construções**

Argumentamos nas páginas anteriores que, se a aprendizagem de uma língua é, em última instância, a aprendizagem de construções (das mais concretas até as mais abstratas), e se os processos subjacentes a tal aprendizado são regidos por processos cognitivos gerais (i.e., possuem a mesma natureza dos princípios psicológicos envolvidos no desenvolvimento das categorias do nosso conhecimento semântico, como propostos originalmente por Eleanor Rosch), então torna-se válida a noção de protótipos verbais (ou lingüísticos) na aquisição dos aspectos abstratos de línguas naturais. Os processos de *chunking* possuem plausibilidade neurobiológica (como discutido no capítulo 1 desta tese) e modelos connexionistas e emergentistas (de base probabilística) parecem fornecer uma boa explicação de como se dá o processamento e a representação de nosso conhecimento em um nível simbólico-representacional (como discutimos nos capítulos 2 & 3).

Como vimos neste capítulo, exemplares altamente freqüentes são fundamentais no processo de aprendizagem das categorias abstratas (construções sintáticas), pois auxiliam e possibilitam que outros membros (desde os mais periféricos) sejam

gradualmente associados a uma determinada categoria ou construção, de forma que a própria categoria ou construção passe a existir como entidade abstrata, independente dos itens individuais que possam dela fazer parte.

Finalmente, como é possível depreender de nossa discussão até este ponto, os processos envolvidos em ASL tornam ainda mais complexo todo o processamento cognitivo acima discutido. *Reconstruir* uma língua é ainda mais complicado, pois, ao longo do desenvolvimento da L2, há não apenas um processo básico de indução, mas principalmente de competição com aquelas categorias já construídas e consolidadas na L1 (implicando, assim, em formas alternativas de se construir a mesma realidade). O conceito selinkiano de *interlíngua* encerra todas estas questões.

Estes são os pontos teóricos centrais que embasam esta tese. A seguir, no capítulo 5, apresento os resultados da análise de um corpus de aprendizes de inglês como segunda língua (ou aprendizagem espontânea de uma L2).

## 5. Análise de um corpus longitudinal (Aprendizes de ESL no Projeto ESF)

*Why is longitudinal research essential to the advancement in the field of SLA? The simple but uncontestable answer is that many questions concerning second language learning are fundamentally questions of time and timing.*

Lourdes Ortega & Gina Ibarra-Shea

Este capítulo apresenta os resultados da análise de um corpus longitudinal de aprendizes de inglês como segunda língua<sup>56</sup>. O estudo que gerou o corpus foi concebido e planejado sob total influência da teorização de Larry Selinker (discutida anteriormente) a respeito do desenvolvimento e estruturação de *interlínguas* no processo de aquisição de uma L2, ou seja, a idéia de uma língua do aprendiz ou *learner variety*, como define Perdue (1993), cujos princípios são: 1) a organização cognitiva da L2 de um dado aprendiz é sempre sistemática; e 2) a transição de uma variedade de interlíngua (ou sua configuração específica) para outra ao longo do tempo é também altamente sistemática. Esta é a concepção de processos contínuos e graduais de construção, estruturação e, principalmente, reconstrução de uma língua que permeia a noção de competência lingüística dinâmica e dialógica, altamente condicionada pelo insumo e pelo contexto, como vimos na discussão teórica apresentada neste manuscrito. Portanto, o estudo já nasceu moderno, pois este é um dos princípios mais caros do paradigma lingüístico que permeia esta tese.

---

<sup>56</sup> Tanto a análise de frequência como o estudo sobre prototipicidade verbal relatados nas páginas seguintes foram realizados durante estágio sanduíche de um ano no *English Language Institute* (ELI) da Universidade de Michigan, em Ann Arbor, EUA, sob a supervisão de Nick C. Ellis. O estágio só foi possível graças à bolsa sanduíche da CAPES (BEX # 0043060) e ao afastamento de minhas funções de docência no CEFET Ouro Preto. O leitor encontrará em Ellis & Ferreira-Junior (no prelo, a & b) duas excelentes sínteses (e outros desdobramentos) desta pesquisa conduzida no subcorpus do projeto ESF e relatada nas páginas seguintes.



Pelo fato de o Projeto ESF (acrônimo para *European Science Foundation* ou Fundação Europeia da Ciência)<sup>57</sup> basear-se em dados empíricos longitudinais sobre a aprendizagem espontânea de uma língua estrangeira, será bastante interessante analisar tais dados à luz da teorização de Goldberg sobre o conceito de *construções*. Talvez seja possível identificar, num estudo longitudinal deste quilate, a gênese de certas construções sintáticas. Talvez seja possível verificar se na gênese de tais construções encontramos palavras específicas, fórmulas e expressões fixas, ou rotinas. É isto que buscaremos verificar neste e nos capítulos seguintes.

Uma observação das limitações do subcorpus do projeto ESF aqui analisado torna-se importante neste momento. Embora seja um grande projeto, envolvendo 6 línguas fontes e 5 línguas alvo, apenas uma sub-parte será analisada nas seções a seguir, ou seja, 7 aprendizes de inglês como L2 (destes, de apenas 4 temos os dados completos). O enfoque será igualmente restrito às construções, uso de fórmulas e colocações. Também, mesmo sendo nossa investigação de certa forma pioneira em ASL, tanto por adotar uma perspectiva construcionista como em utilizar uma análise longitudinal deste porte, é necessário mais estudos em ASL que complementem e estendam para outros idiomas os estudos de Liang (2002, apud Goldberg) e Gries & Wulff (2005), ambos enfocando a aprendizagem de construções da estrutura argumental em língua inglesa como L2.

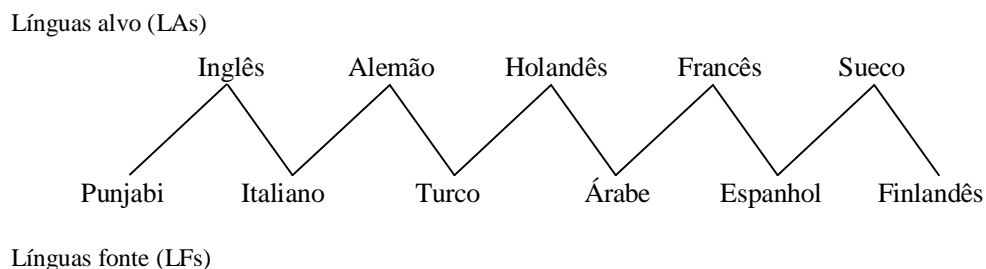
Os outros subcorpora do projeto ESF (com dados preciosos sobre a aprendizagem do alemão, do holandês, do francês e do sueco) se constituem, a meu ver, numa fonte de investigação extremamente rica para pesquisadores em ASL.

---

<sup>57</sup> Diante das três versões encontradas em língua portuguesa para *European Science Foundation*, quais sejam, Fundação Europeia da Ciência, Fundação Europeia de Ciência e Fundação Europeia para a Ciência, decidi optar pela primeira, que, além de ser mais freqüente, conforme resultados de buscas das três expressões realizadas pelo *Google* com filtro em páginas de língua portuguesa, é também a versão oficial que consta na página do Ministério dos Negócios Estrangeiros de Portugal (<http://www.reper-portugal.be/>). Também optei pela ortografia do português brasileiro ao acentuar a palavra Europeia nesta expressão, nos poucos casos em que o acrônimo ESF não é usado ao longo desta tese.

### 5.1 O Projeto ESF

O Projeto ESF, realizado no início dos anos 80, foi um grande projeto de pesquisa desenvolvido durante cinco anos e meio, com o objetivo de estudar a estrutura e a natureza do processo de aquisição de L2 por parte de imigrantes adultos em cinco países europeus: França, Alemanha, Reino Unido, Holanda e Suécia. Foi, portanto, um projeto transnacional, envolvendo seis equipes de pesquisadores nos cinco países mencionados acima (sendo duas equipes na França) e financiado pela Fundação Europeia da Ciência. No período de cinco anos e meio os pesquisadores coletaram dados espontâneos de quarenta imigrantes de sete diferentes nacionalidades, à época recém residentes nos cinco países europeus mencionados. Havia, ao todo, cinco línguas-alvo (LAs) e seis línguas-fonte (LFs), o que possibilitou comparações interessantes sobre o processo de ASL, tais como, *inter alia*, a aquisição de uma LA por falantes de LF de diferentes estruturas (por ex, SOV, SVO, etc.), bem como a aquisição de duas LAs por falantes de uma mesma LF. Veja, na Figura 8, o desenho original que orientou o projeto ESF:



**Figura 8**

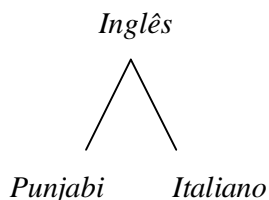
Configuração das línguas fonte e alvo no Projeto ESF (Perdue,1993: xi).

Todo o projeto foi idealizado como um estudo longitudinal envolvendo a gravação da produção dos sujeitos (áudio ou vídeo) ao longo de um período de 30 meses de entrevistas (dois anos e meio). As entrevistas ocorriam num intervalo entre 4 a 6 semanas para os sujeitos do grupo longitudinal, ao passo que para os sujeitos no grupo de controle foram realizadas gravações no início, meio e fim deste período de 30 meses.

Uma boa parte das transcrições do Projeto ESF está agora etiquetada (*tagged*) no formato CHAT (*.cha*) do Projeto CHILDES<sup>58</sup>, podendo o leitor acessar gratuitamente os arquivos disponíveis na *webpage* do *Max Planck Institute for Psycholinguistics* (doravante, MPI), em Nijmegen, Holanda<sup>59</sup>.

## 5.2 Sujeitos investigados

A análise aqui relatada é baseada nos dados de sete aprendizes de inglês no grupo longitudinal, à época do estudo morando na Inglaterra, falantes nativos de italiano (Andrea, Lavinia, Santo e Vito) e punjabi (Madan, Jarnail e Ravinder). Os dados destes sujeitos serão referidos, doravante, como o *subcorpus* do projeto ESF investigado. Veja, também, o Anexo B, que contém informações sociobiográficas importantes de cada sujeito. Como mencionado, apenas parte dos dados transcritos está disponibilizada no *site* do Max Planck, portanto o subcorpus utilizado consiste de quatro conjuntos de dados completos e três incompletos (levando em conta o planejamento original do estudo, que envolvia uma coleta de dados ao longo de 30 meses). A Figura 9 ilustra o subcorpus investigado nesta tese e o Quadro 2 fornece uma visão geral dos dados transcritos:



**Figura 9**  
Configuração das línguas fonte e alvo no subcorpus analisado

---

<sup>58</sup> Veja MacWhinney (2000) para informações detalhadas sobre o projeto *Child Language Data Exchange System* (CHILDES) e as ferramentas (programas) de análise CLAN (KWAL, FREQ, MOR, COMBO, etc.), que foram usadas na nossa análise e são mencionadas no item 5.4 a seguir.

<sup>59</sup> Visite a página do Instituto em <http://www.mpi.nl/> ou siga a rota a seguir para acessar diretamente os dados: [http://corpus1.mpi.nl/ds/imdi\\_browser/](http://corpus1.mpi.nl/ds/imdi_browser/)

| Línguas fonte | Língua alvo | Grupo longitudinal<br>Completos | Grupo longitudinal<br>Incompletos |
|---------------|-------------|---------------------------------|-----------------------------------|
| Italiano      | Inglês      | 2                               | 2                                 |
| Punjabi       | Inglês      | 2                               | 1                                 |

### Quadro 2

Visão geral dos dados transcritos

### 5.3 Justificativa e característica do subcorpus utilizado

A escolha do subcorpus analisado deveu-se, principalmente, 1) a familiaridade do autor com a língua inglesa, o que permitiu 2) a replicação do estudo conduzido por Goldberg, Casenhiser & Sethuraman (2004) com aprendizes de L1, possibilitando, desta forma, uma verificação da validade da teorização construcionista no âmbito da aprendizagem de uma L2. Dada tal configuração (ou seja, a comparação entre grupos de aprendizes & mães no estudo em L1 por um lado, e aprendizes de L2 & entrevistadores no projeto ESF por outro) seria possível identificar tanto a existência de verbos prototípicos em L2 como comparar a frequência de ocorrência destes nas três construções investigadas em ambos os grupos e estudos. Finalmente, partindo do pressuposto de que as motivações funcionais e comunicativas em L1 e L2 são as mesmas, a expectativa é que tal padrão de desenvolvimento construcional seja confirmado nos dados dos aprendizes de L2 do projeto ESF. Este foi o objetivo primeiro da análise conduzida em nosso subcorpus ESF (cf Ellis & Ferreira-Junior, no prelo a).

Como comentado no início deste capítulo, os outros subcorpora do projeto ESF (com dados sobre a aprendizagem do alemão, do holandês, do francês e do sueco), se analisados com as mesmas lentes utilizadas em nosso estudo, poderiam, potencialmente, confirmar os dados aqui verificados sobre a validade das construções em ASL naquelas línguas. Até esta data, não é de meu conhecimento nenhum estudo publicado com análises semelhantes nestes outros subcorpora.

A tabela 4 exibe os dados transcritos que foram analisados e que consistem em sessões envolvendo as seguintes atividades:

| <b>Código</b> | <b>Atividade</b>   |
|---------------|--|
| a             | Conversa livre ( <i>Free conversation</i> )  |
| b             | Produção de vocabulário ( <i>Vocabulary elicitation / Lexical task</i> )                         |
| c             | Encenação ( <i>Role play</i> )   |
| d             | Descrição de figuras ( <i>Picture description</i> )  |
| e             | Dar instruções ( <i>Stage directions</i> )   |
| f             | Assistir/comentar o filme <i>Harold Lloyd</i> ( <i>Film watching / retelling: Harold Lloyd</i> ) |
| g             | Assistir/comentar filmes ( <i>Cultural film watching / commenting</i> )                          |
| h             | Autoconfrontação ( <i>Self-confrontation</i> )   |
| i             | Assistir/contar o filme <i>Tempos Modernos</i> ( <i>watching / Retelling: Modern Times</i> )     |
| j             | Encenação com interlocutor externo ( <i>Role play with external interlocutor</i> )               |
| k             | Passeio acompanhado ( <i>Accompanied outing</i> )  |
| l             | Descrever rotas ou trajetos ( <i>Route description</i> )   |
| m             | Tarefas com relações espaciais: objeto escondido ( <i>Spatial task: hidden object II</i> )       |
| n             | Relato de registros ( <i>Logs retelling</i> )  |
| o             | Comentar foto ( <i>picture commenting</i> )  |
| p             | Relato induzido: comentar foto ( <i>Prompted recall: picture commenting</i> )                    |
| q             | Descrição de apartamento ( <i>Apartment description</i> )  |
| r             | Autoconfrontação com encenação ( <i>Self-confrontation with role play</i> )                      |

**Tabela 4**  
Lista de tarefas realizadas e seus respectivos códigos

Ao todo, 234 sessões transcritas com as atividades dos sete sujeitos foram analisadas. Como observado acima, existem dois conjuntos completos de dados para os falantes de punjabi, Ravinder e Jarnail, e dois para os falantes de italiano, Andrea e Lavinia. Os conjuntos incompletos são de outros dois falantes de italiano, Vito e Santo, e de um falante de punjabi, Madan. A Tabela 5 exibe informações dos sujeitos, seus respectivos números de sessões e datas de início e fim no projeto:

| L1       | L2     | Dados completos | Dados incompletos | Sessões | Datas               |
|----------|--------|-----------------|-------------------|---------|---------------------|
| Italiano | inglês | Andrea          |                   | 28      | 10-01-83 a 15-01-85 |
| Italiano | inglês | Lavinia         |                   | 41      | 23-08-83 a 28-02-86 |
| Italiano | inglês |                 | Santo             | 11      | 28-07-83 a 29-02-84 |
| Italiano | inglês |                 | Vito              | 32      | 14-07-82 a 06-02-85 |
| Punjabi  | inglês | Jarnail         |                   | 49      | 24-11-82 a 24-04-85 |
| Punjabi  | inglês | Ravinder        |                   | 51      | 19-11-82 a 24-04-85 |
| Punjabi  | inglês |                 | Madan             | 22      | 10-11-83 a 20-02-85 |

**Tabela 5**

L1 e L2 dos sujeitos, número total de sessões e datas de início e conclusão no projeto.

#### 5.4 Procedimento de análise

Inicialmente, todos os arquivos transcritos no formato CHAT (.cha) foram descarregados do website do MPI na internet por meio do navegador IMDI BCBrowser 3.0, disponível também gratuitamente no website do MPI (veja na subpasta *Step 1* do CD-ROM em anexo todos os 234 arquivos das sessões). O programa CLAN, desenvolvido no Projeto CHILDES, foi utilizado na análise dos 234 arquivos de sessões transcritas. Veja trecho abaixo de uma transcrição original no formato CHAT:

```
*SAN: i want one # er in@s pianto@s [:=T system].
%snd: "lian11a.wav" 9383 13613
@Turn
*INM: dont know.
%snd: "lian11a.wav" 13613 14203
@Turn
*SAN: er system # for er tape.
%snd: "lian11a.wav" 14193 17397
@Turn
*INM: mhm <> [% laughs].
%snd: "lian11a.wav" 17570 18226
@Turn
*SAN: is very good.
%snd: "lian11a.wav" 18216 19813
@Turn
*INM: yes?
*INM: <> [% laughter].
```

Transcrições no formato CHAT possuem uma sintaxe própria. Por exemplo, os enunciados produzidos por um dado sujeito ou seus entrevistadores iniciam com um asterisco seguido por três letras de identificação (no trecho acima, o código \*SAN identifica uma frase do sujeito Andréa, ao passo que \*INM indica que a frase foi produzida pela entrevistadora (pesquisadora) Margaret Simonot). Assim, o próximo passo foi extrair as linhas (*tier lines*) apenas dos sujeitos em cada uma das sessões por ele realizadas. Para tal, foi usada a ferramenta de análise KWAL, que permite extrair apenas as linhas dos sujeitos. Um único arquivo com todas as *tier lines* de cada sujeito em todas as sessões foi criado. O próximo passo foi ‘limpar’ manualmente este novo arquivo contendo apenas as linhas dos sujeitos de forma a eliminar todo e qualquer comentário ou tradução feitos pelos entrevistadores durante a gravação das sessões, assim como comentários feitos na transcrição dos dados. Veja como ficou o trecho da sessão com o sujeito Andrea exibido acima após tal ‘limpeza’

\*SAN: i want one # er in@s pianto@s.

\*SAN: er system # for er tape.

\*SAN: is very good .

Em seguida, foi realizada uma análise morfossintática nos arquivos ‘limpos’ na etapa descrita acima, utilizando, desta vez, a ferramenta MOR para identificar todos os possíveis verbos usados pelos sujeitos. Todas as palavras identificadas na análise MOR como sendo palavras-tipo-verbo (*verb-like words*) foram, então, checadas e listadas manualmente. Tal procedimento foi adotado devido ao fato de que a ferramenta de análise faz apenas a identificação das várias classes gramaticais as quais uma palavra pode pertencer, ignorando o contexto daquelas ocorrências específicas. Por exemplo, ambas as palavras *house* e *need* podem ser substantivo ou verbo, ao passo que a palavra *open* pode ser um adjetivo, um verbo ou ainda um substantivo (como na frase *life in the open*). Após a checagem manual, *house*, que registrava 337 tokens, não possuía nenhuma ocorrência como verbo, ao passo que *open* registrava 3 tokens como adjetivos e 131 como verbos. Todas as 24 ocorrências de *need*, por outro lado, eram como verbo. Por isto identificamos 405 palavras-tipo-verbo.

Foram então criadas listas no formato *.cut* com todas as palavras-tipo-verbo identificadas no corpus, 405 tipos ao todo, que incluíam também suas possíveis formas no passado (regular e irregular) e formas progressivas (ing). As listas foram usadas em análises de frequência de todos os sete arquivos *.cha* limpos usando a ferramenta *FREQ* do *CLAN*. Foram extraídos de todos os sete arquivos individuais todas as linhas (enunciados) contendo aqueles verbos (e suas formas no passado e progressivo) listados no arquivo *.cut*. Todos os arquivos individuais foram fundidos em um único arquivo *CHAT*. Este grande arquivo *CHAT* listava todas as linhas de todos os sujeitos em todas as sessões nas quais havia ocorrência de cada uma das palavras-tipo-verbo identificadas.

Assim, o arquivo de 29.706 linhas com 14.639 ocorrências (*tokens*) das 405 palavras-tipo-verbo listava: a) a quantidade de vezes em que um dado verbo ocorria no corpus, b) a identificação do sujeito que o produziu, c) as frases nas quais tais verbos apareciam, e d) a linha da qual cada frase tinha sido extraída nos arquivos individuais dos sujeitos. Veja exemplo abaixo:

```

1 buying
  File "C:\Documents and Settings\Fernando\Desktop\ESF Phase 8\Clean Files\SLA.cha": line 5055.
  *SLA:      yes hello good afternoon i am interesting in buying a house .
132 coming
  File "C:\Documents and Settings\Fernando\Desktop\ESF Phase 8\Clean Files\SJA.cha": line 269.
  *SJA:      coming yeah .
  File "C:\Documents and Settings\Fernando\Desktop\ESF Phase 8\Clean Files\SLA.cha": line 276.
  *SLA:      it is coming home .
  File "C:\Documents and Settings\Fernando\Desktop\ESF Phase 8\Clean Files\SMA.cha": line 152.
  *SMA:      i am going [/?] coming in the down <seven star> .
2 dream
  File " C:\Documents and Settings\Fernando\Desktop\ESF Phase 8\Clean Files \SLA.cha": line 4812.
  *SLA:      <no> [?] they dream about .
  File " C:\Documents and Settings\Fernando\Desktop\ESF Phase 8\Clean Files \SLA.cha": line 4820.
  *SLA:      they dream about .

```

Finalmente, o próximo passo consistiu na identificação (manual) de todas as possíveis construções VL, VOL e VOO e seus respectivos verbos. Foram cuidadosamente analisadas, uma a uma, todas as 29.706 linhas (frases) do corpus. A etiquetagem das linhas foi feita manualmente. Os verbos e suas respectivas construções VL, VOL e VOO foram colocados entre colchetes. Primeiro o verbo (type) seguido pelo tipo de construção analisada. Veja exemplo:



- a) SLA: you come out of my house. [come] [VL]  
 b) SMA: charlie say # shopkeeper give me one cigar ## he give it ## he er # he smoking  
 # after two # baby coming . [give] [VOO]  
 c) SRA: come on you ## thats it ### open your bag # put on table as well . [put] [VOL]

Após esta fase de etiquetagem, foram finalmente identificadas 10.448 ocorrências de 234 tipos verbais (*tokens* e *types*). Ao todo, 778 construções foram inicialmente identificadas. Todas as construções foram checadas por um pesquisador independente (i.e., que não fazia parte da pesquisa) buscando-se, com isto, acuidade dos dados<sup>60</sup>. Ao todo, o pesquisador fez observações sobre 158 construções (20,4% do total), classificando-as da seguinte forma: 48 não foram reconhecidas como sendo construções válidas e 16 foram re-classificadas. O pesquisador expressou dúvidas e não soube classificar outras 94 construções. Destas últimas, a maioria dos casos (64 construções) necessitava de uma nova checagem do contexto original para que os rótulos VL, VOL e VOO fossem confirmados (por exemplo, quando envolviam oração sem sujeito, objeto oculto, contextos fragmentados, etc.).

Por outro lado, havia um número de possíveis construções VOL expressas num ordenamento incomum de palavras (diferentes ordenamentos sintáticos). Veja alguns exemplos encontrados para os dois tipos possíveis de construção VOL:

a) VOL → VLO (14 ocorrências)

- 1) \*SJA: *yes # now er # in the put in a /box blue box n s money*
- 2) \*SMA: *# put in there money <##> [>1] close in the box*
- 3) \*SMA: *yeah # he put in the till <the money>*
- 4) \*SMA: *# opposite ## show case # fitting the you know ## glasses <big glasses> you know # i put er inside ## nice thingy you know.*
- 5) \*SMA: *# left ha [/?] right hand side in the corner ## i put in er telly [/?] television.*
- 6) \*SRA: *<yeah> [<1] # first # put in # blue table # < > the chair bring ###*
- 7) \*SMA: *putting in there money ##*
- 8) \*SLA: *oh you can take home them*
- 9) \*SJA: *yes # now er # in the put in a /box blue box n s money*

---

<sup>60</sup> Sou grato ao Carson Maynard, do *English Language Institute* da Universidade de Michigan, por este trabalho de checagem dos dados aqui relatados.

## b) VOL → OVL (16 ocorrências)

- 10) \*SMA: *yeah ## keep it that er cigarette # and # money ## and matches put in the pocket #*
- 11) \*SAN: *and er the [/?] your bag # no [!] the plastic bag # put er on the floor*
- 12) \*SJA: *dust bin n s bag put in*
- 13) \*SMA: *yeah # stand up the chair ### and er one leg put on the table ## yeah # put this money thingy ern this file # first one ##*
- 14) \*SMA: *## telephone put in the table*
- 15) \*SMA: *this bag <he put him> [/?] put in the st [/?] er floor #*
- 16) \*SMA: *yeah ## keep it that er cigarette # and # money ## and matches put in the pocket #*
- 17) \*SMA: *sweets put in the counter left hand <side [/?] no er> [>1] sorry right hand side*
- 18) \*SMA: *you know # big er you know # bag <my marriage you know put in the stuff> [\$ dowry] clothes everything put in there.*
- 19) \*SRA: *a horse # put in there.*
- 20) \*SRA: *bag put in table.*
- 21) \*SRA: *its one chair an put in dining room.*

Após as observações feitas pelo pesquisador independente, foram descartadas 48 construções, outras 16 construções foram reclassificadas e 64 foram revistas juntamente com o orientador da pesquisa, observando cuidadosamente o contexto original (i.e., os arquivos originais). Após discussão, ficou decidido que os casos explicitados acima eram bons exemplos de construção VOL. Outras 13 construções que haviam sido rotuladas pelo pesquisador independente como VL ou VOL por conter objetos ocultos eram, na verdade, construções OVL ou VLO que apenas poderiam ser detectadas quando colocadas em seus contextos originais<sup>61</sup>.

Diferentemente do estudo de Goldberg, Casenhiser & Sethuraman (2004), que analisou as mesmas construções (VL, VOL e VOO) em um corpus de crianças aprendendo L1, o subcorpus do Projeto ESF aqui analisado diz respeito à aquisição de uma L2 por adultos. Os exemplos exibidos acima refletem o que Perdue (1993) chama de ‘variedade básica’ (*basic variety*), após uma análise minuciosa da produção dos aprendizes dos

---

<sup>61</sup> Todo procedimento descrito concisamente acima pode ser checado por meio de consulta aos arquivos relativos a cada fase de análise. Todos os arquivos gerados durante a análise dos dados, com todos os passos do processo desde os 234 arquivos CHAT com as transcrições originais, encontram-se em CD-ROM anexo a esta tese. O leitor também encontrará no CD o programa CLAN, que deverá ser instalado para a leitura dos arquivos.

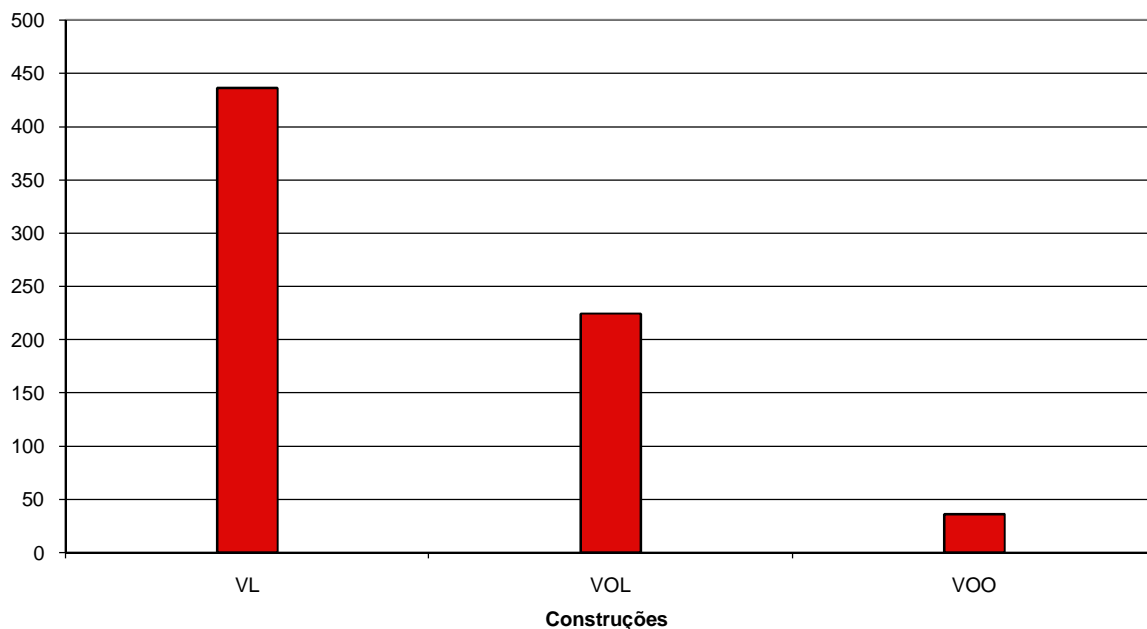
cinco idiomas estudados no Projeto ESF. Conforme Mitchel and Myles (2004:150), a variedade básica é caracterizada por uma estrutura tópico-comentário, conforme a distinção de Givón entre os usos pragmáticos e sintáticos que fazemos numa língua.

Para Perdue (1993), há aspectos semelhantes verificados na interlíngua de todos os sujeitos que participaram do estudo ESF, independentemente de suas línguas alvo e fonte (veja Figura 8, página 114). Neste padrão de básico de desenvolvimento, parece haver uma certa padronização na organização cognitiva da interlíngua dos aprendizes, evidenciado por estágios que vão desde a utilização de recursos pragmáticos, passando pelo emprego de expressões lexicalizadas para, finalmente, atingir o nível complexo da construções sintaticamente estruturadas (mais abstratas).

Embora a própria idéia de que uma variedade básica, pragmaticamente estruturada, observada no desenvolvimento da interlíngua implique um padrão de desenvolvimento universal – ou seja, independentemente da L1 do aprendiz, parece razoável imaginar que os processos iniciais em ASL parasitem nas estruturas ‘sintáticas’ da L1 do aprendiz (seja, por assim dizer, uma *tabula repleta*), como sugere Ellis (no prelo, 2006b, 2006c), e que, muito possivelmente, fenômenos de interferência/transferência lingüística sejam observados na interlíngua dos aprendizes. Como vimos acima, quase todos os casos de construção invertida ocorrem na produção dos falantes de punjabi. De fato, o punjabi é classificado como sendo uma língua SOV e, como tal, parece estar interferindo diretamente na interlíngua dos aprendizes durante a comunicação na língua inglesa (que, como o italiano, é predominantemente uma língua SVO).

## 5.5 Resultados

Após as fases de etiquetagem, checagem e re-checagem mencionadas acima, restaram 696 construções: 436 VL, 224 VOL e 36 VOO. Ao todo, havia 39 diferentes tipos verbais nas construções VL, 24 nas VOL e 9 nas construções VOO. A seguir é apresentado o Gráfico 1 com os resultados das análises acima descritas.

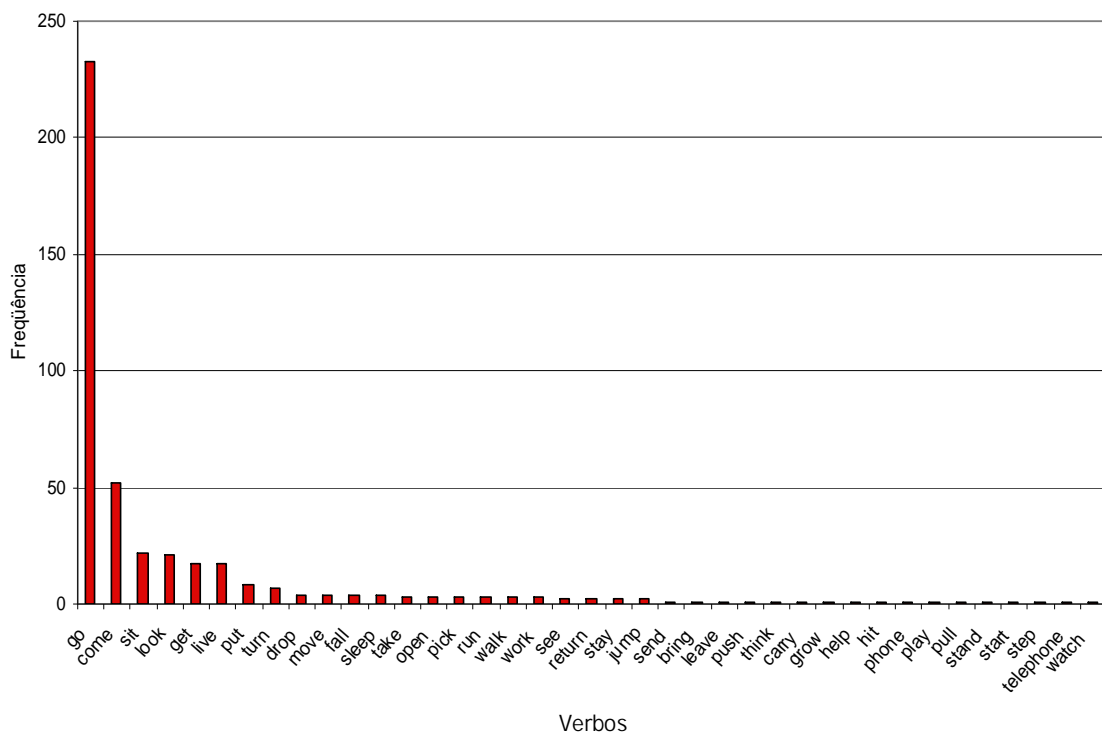


**Gráfico 1**  
Total de construções dos aprendizes

Os gráficos que serão mostrados a seguir exibem detalhadamente cada uma das três construções VL, VOL e VOO com seus respectivos verbos, conforme verificado nos dados dos sete aprendizes. Cada gráfico exibe a distribuição de frequência e ocorrência de *types* numa dada construção.

Iniciaremos com o Gráfico 2, que exibe os achados relativos à frequência de *types* verbais nas construções VL. Ao todo, 38 verbos foram usados nas 436 construções VL identificadas. Observe que há uma clara tendência de um verbo dominar a maior parte das ocorrências da construção VL (nesta construção, o verbo *go*). Observe também que a distribuição das construções ao longo do gráfico é gradativa.

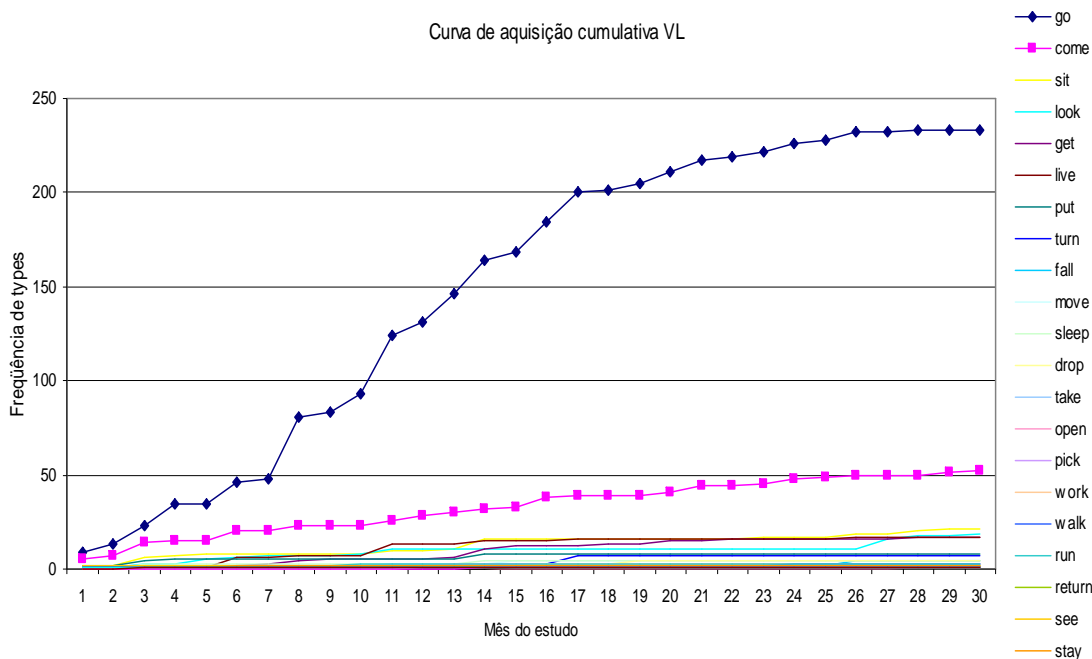
Construções VL dos aprendizes



**Gráfico 2**  
Construções VL dos aprendizes

No Gráfico 3 é apresentada uma curva de aquisição cumulativa dos verbos na construção VL em discussão. Observe que há um crescimento gradual e constante na frequência de uso do verbo prototípico *go*, seguido pelos verbos *come*, *sit*, *look* e *get*.<sup>62</sup>

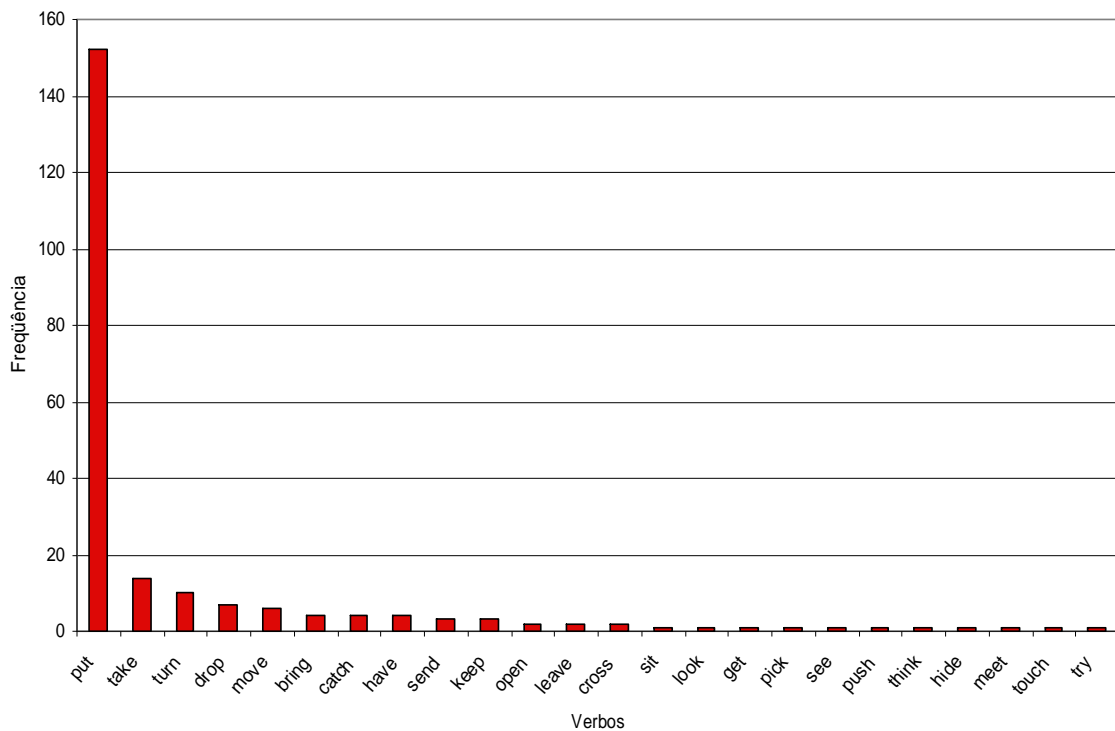
<sup>62</sup> Para facilitar a legibilidade dos gráficos de aquisição cumulativa, apenas os dois verbos mais frequentes em cada construção serão destacados com padrões gráficos diferenciadores. Este será o padrão adotado para este tipo de gráfico ao longo desta tese.



**Gráfico 3**  
Curva de aquisição cumulativa VL para os aprendizes

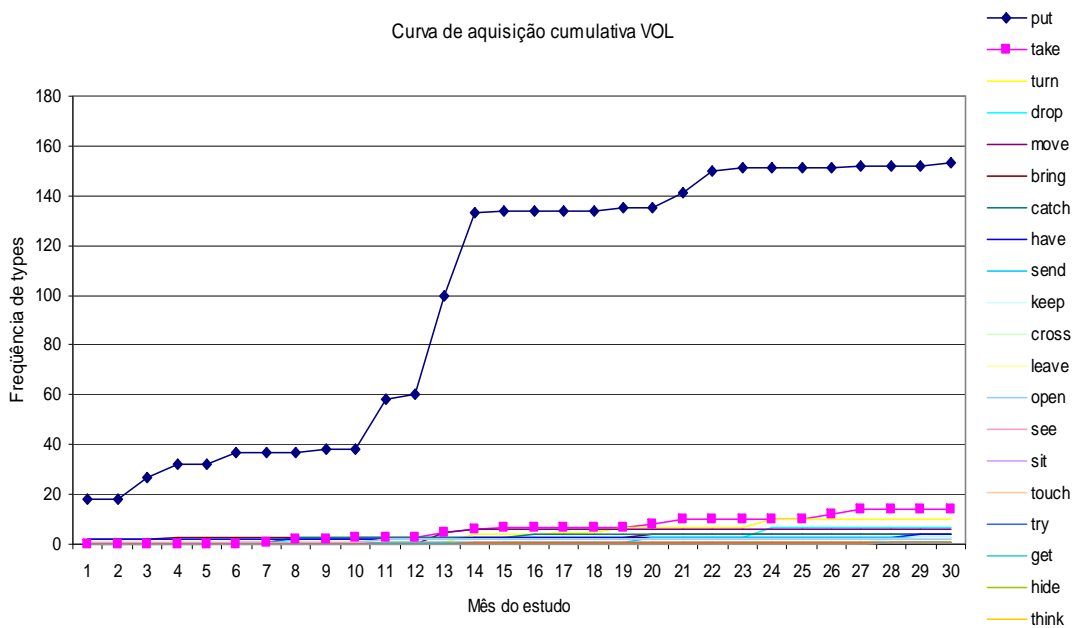
O Gráfico 4 exibe os resultados relativos à construção VOL dos aprendizes no subcorpus analisado e frequência de verbos que ocorrem nesta construção. Ao todo, foi verificada a ocorrência de 24 verbos nas 224 construções VOL identificadas. Novamente observamos a tendência de um verbo dominar a maior parte das ocorrências da construção VL (nesta construção, o verbo *put*). Como no resultado da construção VL exibido anteriormente, observe também que a distribuição das construções ao longo do gráfico é gradativa.

Construções VOL dos aprendizes



**Gráfico 4**  
Construções VOL dos aprendizes

Como exibido anteriormente no Gráfico 3 para a construção VL, veremos a seguir no Gráfico 5 uma curva de aquisição cumulativa dos verbos ao longo do desenvolvimento da construção VOL. Há um crescimento acentuado no uso do verbo prototípico *put*, seguido pelos verbos *take*, *turn*, *drop* e *move*.

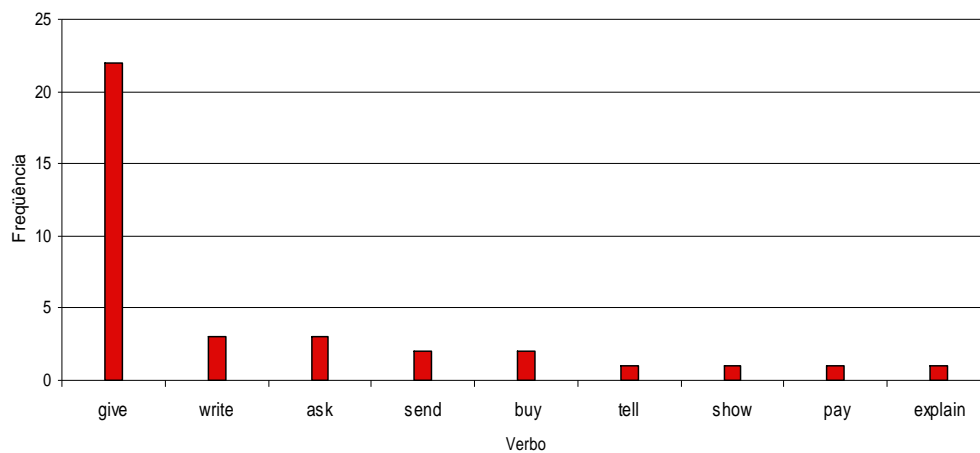


**Gráfico 5**  
Curva de aquisição cumulativa VOL para os aprendizes

O Gráfico 6 exibe os achados relativos ao uso e frequência dos 9 tipos verbais empregados nas 36 construções VOO identificadas no subcorpus ESF. O padrão verificado nas outras duas construções se repete aqui: embora haja um baixo número de construções VOO, há uma clara tendência de um verbo dominar a maior parte das ocorrências da construção (nesta construção, o verbo *give*). Observe também que a distribuição das construções ao longo do gráfico é gradativa.

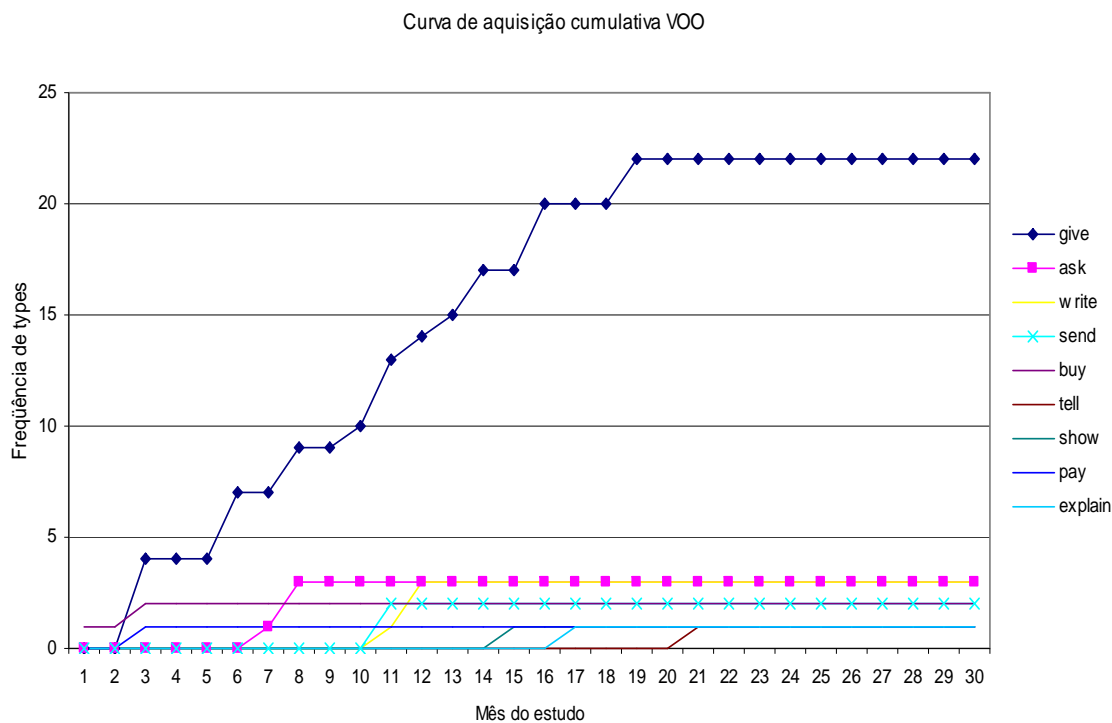


Construções VOO dos aprendizes



**Gráfico 6**  
Construções VOO dos aprendizes

Finalmente, o Gráfico 7 exibe a curva de aquisição cumulativa dos verbos para a construção VOO. Como observado nas construções VL e VOL acima, verifica-se aqui o mesmo padrão de aumento gradual e constante no uso do verbo prototípico *give*, seguido pelos verbos *write*, *ask*, *send* e *buy*.



**Gráfico 7**  
Curva de aquisição cumulativa VOO para os aprendizes

### 5.6 Curvas individuais de aquisição cumulativa

Embora importantes, os gráficos exibidos no item anterior com as curvas de aquisição cumulativa permitem apenas uma visão global do processo de aquisição de protótipos construcionais. Diante disto, analisaremos de maneira mais fina tal processo por meio dos padrões individuais de desenvolvimento dos sujeitos a partir de curvas cumulativas individualizadas, observando o comportamento dos verbos nas três construções investigadas. Além de um retrato mais fiel, as curvas de aquisição de protótipos verbais individualizadas podem nos fornecer informações importantes sobre quais tipos de tarefas promoveram mais facilmente o uso e o desenvolvimento de determinados tipos de verbos/construções. Esta análise, baseada no desenvolvimento dos quatro aprendizes cujos dados estão completos (ou seja, Andrea, Lavinia, Jarnail e Ravinder), será feita no

capítulo 7 desta tese. Finalmente, o Quadro 3 exibe um sumário com os verbos mais freqüentes em cada construção, para cada sujeito, com suas respectivas porcentagens.

| <b>Aprendizes</b> | <b>VL</b>  | <b>VOL</b>                                     | <b>VOO</b>                              |
|-------------------|--|--|---|
| Andrea            | go 52/76 (68%)<br>come 8/76 (10%)                                | put 18/33 (54%)<br>take 7/33 (21%)             | tell 1/2 (50%)<br>explain 1/2 (50%)     |
| Ravinder          | go 32/51 (62%)<br>come 5/51 (9%)                                 | put 25/36 (69%)<br>turn 6/36 (16%)             | give 4/4 (100%)                         |
| Lavinia           | go 70/123 (56%)<br>come 11/123 (8%)                              | put 17/28 (60%)<br>take 4/28 (14%)             | give 8/17 (47%)<br>ask/write 3/17 (17%) |
| Madan             | go 57/123 (46%)<br>come 20/123 (16%)                             | put 87/109 (79%)<br>move 6/109 (5%)            | give 8/11 (72%)<br>buy 2/11 (18%)       |
| Vito              | look 13/28 (46%)<br>come/get/live/sit 3/28 (10%)<br>go 2/28 (7%) | put/keep/take 2/8 (25%)<br>look/push 1/8 (12%) | give 1/1 (100%)                         |
| Santo             | go 13/19 (68%)<br>come 3/19 (15%)                                |  | give 1/1 (100%)                         |
| Jarnail           | go 7/15 (46%)<br>come/sit 2/15 (13%)                             | put 3/9 (33%)<br>cross 2/9 (22%)               |   |

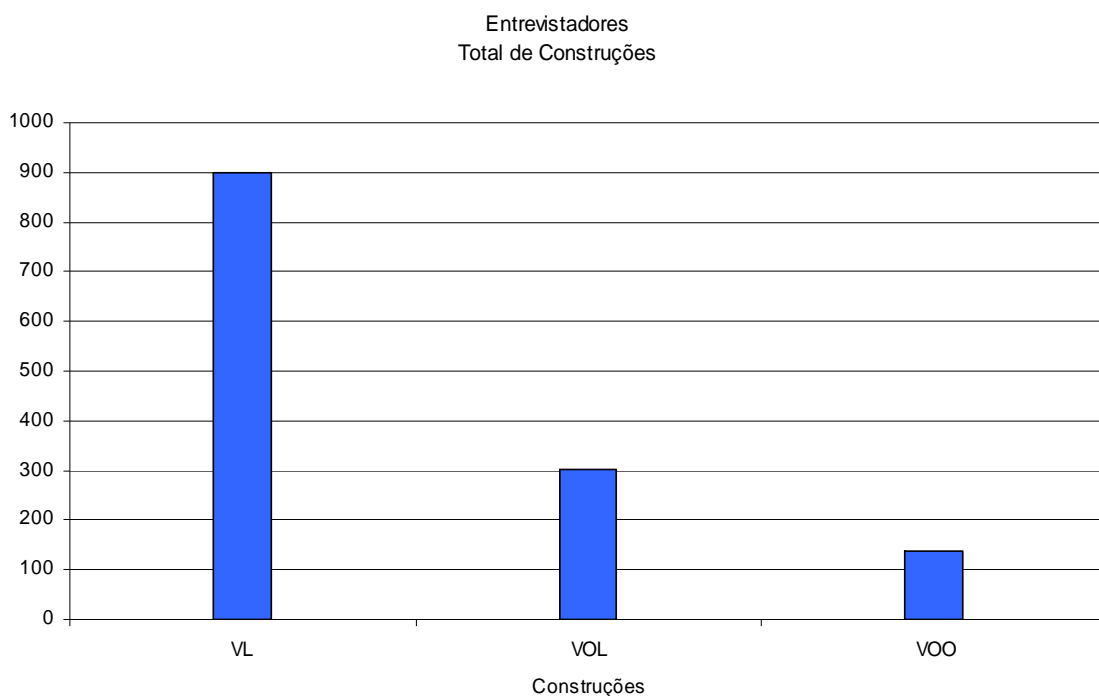
### **Quadro 3**

Sumário dos verbos mais freqüentes em cada construção usados pelos aprendizes

#### **5.7 Análise longitudinal dos entrevistadores no subcorpus ESF analisado**

Seguindo os mesmos procedimentos adotados na análise dos dados dos aprendizes, foi decidido conduzir o mesmo tipo de análise para os 19 entrevistadores que participaram das 234 sessões do nosso subcorpus no projeto ESF. O objetivo desta análise foi obter dados representativos de insumo em L1 de forma a compará-lo, posteriormente, com os resultados dos aprendizes. O procedimento adotado foi o mesmo relatado no item 5.4

acima. Foram identificadas 14.574 ocorrências de 232 tipos verbais (*tokens* e *types*). Ao final da fase de etiquetagem, as construções identificadas foram checadas pelo pesquisador independente. Algumas construções foram descartadas, outras receberam nova classificação, ao passo que algumas construções necessitaram de uma nova observação do contexto original para fins classificação. Ao todo, 1.341 construções foram identificadas: 900 VL, 302 VOL e 139 VOO. Havia 33 diferentes tipos verbais nas construções VL, 33 nas VOL e 12 nas construções VOO. Veja o Gráfico 8 com o total de construções identificadas:

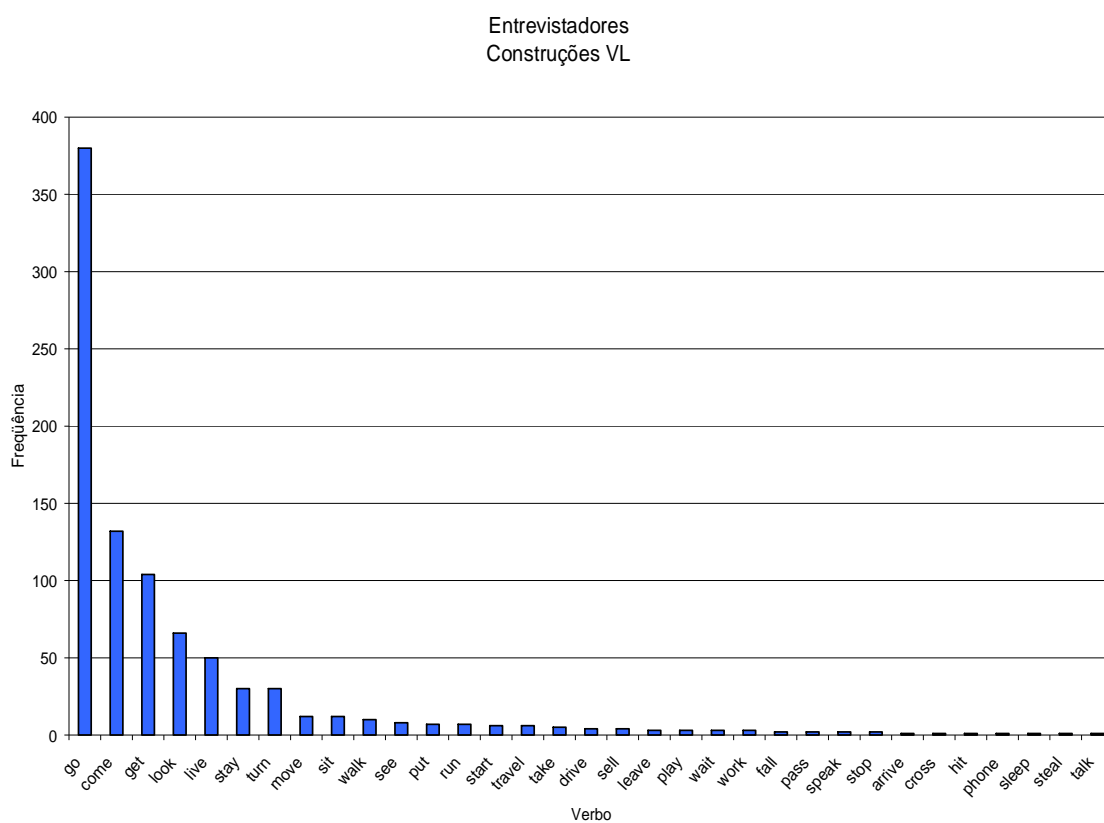


**Gráfico 8**  
Total de construções dos entrevistadores

Ao todo, foram identificados 33 diferentes *types* verbais nas construções VL, 32 nas construções VOL e 12 nas construções VOO. Veja a seguir os gráficos com os resultados de frequência dos verbos encontrados nas construções VL, VOL e VOO.

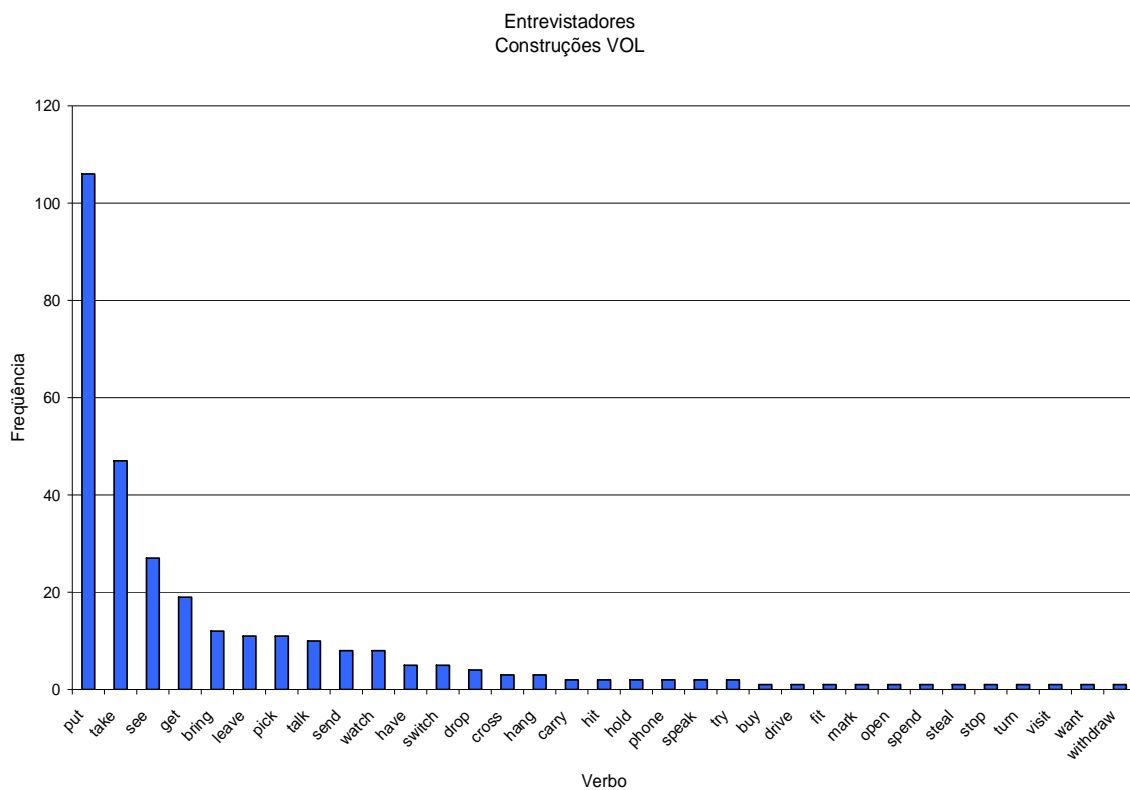
Mais adiante será feita uma comparação destes resultados com aqueles obtidos pelos aprendizes.

O Gráfico 9 exhibe os achados relativos à frequência de *types* verbais nas construções VL por parte dos entrevistadores. Ao todo, 33 verbos foram verificados no uso das 900 construções identificadas. Observe que há uma clara tendência de um verbo dominar a maior parte das ocorrências da construção VL (nesta construção, o verbo *go*). Observe a distribuição gradativa das construções ao longo do gráfico.



**Gráfico 9**  
Construções VL dos entrevistadores

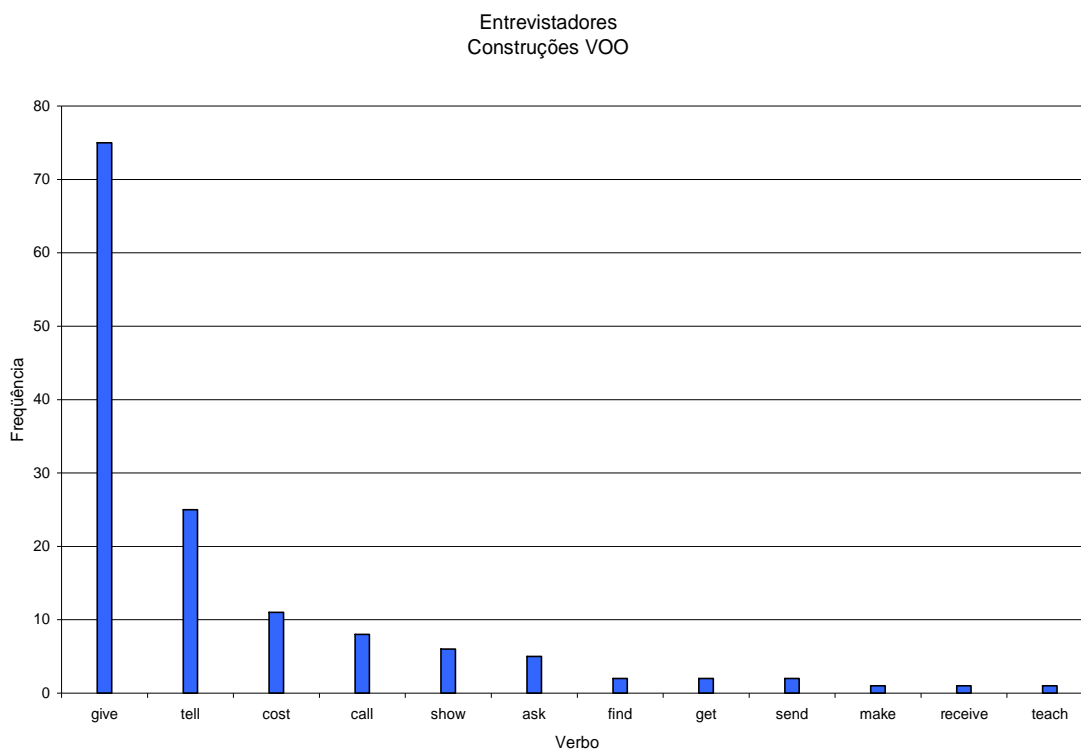
O Gráfico 10 exibe os resultados de frequência de *types* verbais na construção VOL por parte dos entrevistadores. Ao todo, verificamos a ocorrência de 33 diferentes verbos nas 303 construções VOL identificadas. Aqui, novamente, observamos a preponderância de um verbo dominar a maior parte das ocorrências da construção VL (nesta construção, o verbo *put*). Observe também que a distribuição das construções ao longo do gráfico é gradativa, a exemplo da construção VL exibida anteriormente.



**Gráfico 10**  
Construções VOL dos entrevistadores

Finalmente, vemos no Gráfico 11 os resultados relativos à construção VOO dos entrevistadores no subcorpus analisado e a frequência de verbos que ocorrem nesta

construção. Ao todo, foi verificada a ocorrência de 12 diferentes tipos de verbos nas 139 construções VOO identificadas. Apesar da escassez de construções, também é possível observar que há uma clara tendência de um verbo dominar a maior parte das ocorrências da construção VOO (nesta construção, o verbo *give*). Observe também que a distribuição ao longo do gráfico é gradativa, padrão verificado também nos gráficos para as construções VL e VOL exibidos anteriormente.



**Gráfico 11**  
Construções VOO dos entrevistadores

O Quadro 4 exibe um sumário com os verbos mais freqüentes identificados nos dados dos 19 entrevistadores para cada construção, com suas respectivas porcentagens.

| Entrevistadores    | VL                   | VOL                    | VOO               |
|--------------------|----------------------|------------------------|-------------------|
| 19 entrevistadores | go 380/900 (42%)     | put 106/302 (35%)      | give 75/139 (53%) |
|                    | come 132/900 (14%)   | take 49/302 (21%)      | tell 25/139 (17%) |
|                    | get 104/900 (11%)    | see 27/302 (16%)       | cost 11/139 (7%)  |
|                    | look 66/900 (7%)     | get 19/302 (6%)        | call 8/139 (5%)   |
|                    | live 50/900 (5%)     | bring 12/302 (3%)      | show 6/139 (4%)   |
|                    | stay/tum 30/900 (3%) | pick/leave 11/302 (3%) | ask 5/139 (3%)    |

#### Quadro 4

Sumário dos verbos mais freqüentes usados em cada construção pelos entrevistadores

### 5.8 Discussão: comparando os resultados de freqüência verbal

O Gráfico 12 exibe um conjunto de pequenos gráficos com os resultados da análise do subcorpus ESF no grupo de entrevistadores (gráficos na fileira superior) e no grupo de aprendizes investigados (gráficos na fileira inferior) para as construções VL, VOL e VOO, respectivamente. Na comparação, percebemos a estreita semelhança entre freqüência verbal no insumo dos entrevistadores e freqüência verbal na produção dos aprendizes, com os verbos mais prototípicos à frente de cada construção. Ou seja, o mesmo padrão de distribuição estatística dos verbos no insumo em L1 é verificado nos resultados dos aprendizes em L2. Como discutido anteriormente (Capítulo 4) estes resultados são consoantes com os achados de Goldberg e seu grupo e parecem refletir padrões de distribuição estatística do insumo verificados em línguas naturais (i.e., a conhecida Lei de Zipf)



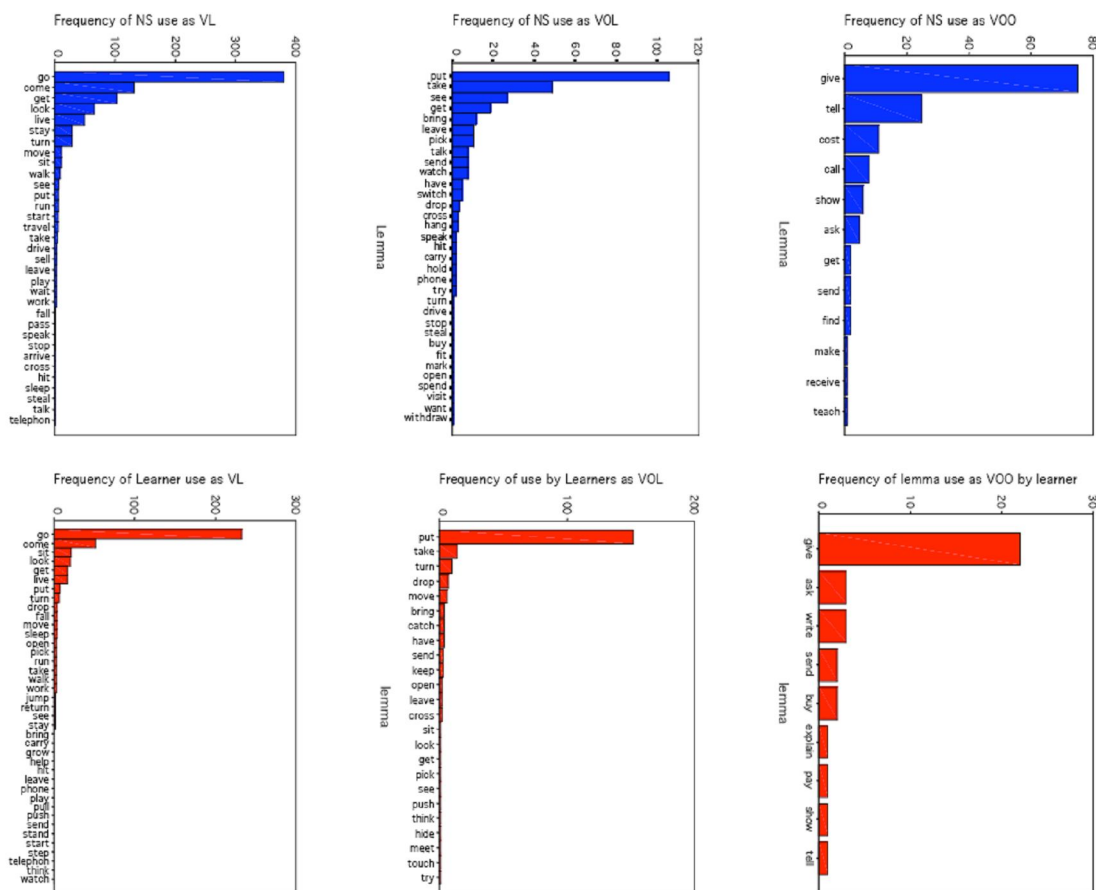


Gráfico 12

Comparação da frequência verbal nas construções dos entrevistadores e nas dos aprendizes. In: Ellis, Nick. & Ferreira-Junior, Fernando. *Construction learning as a function of Frequency, Frequency Distribution, Form & Function*, (Trabalho apresentado na AAAL Conference, Washington, DC, 2008).

Como já discutido, os resultados de frequência verificados em nosso subcorpus encontram respaldo em pesquisas recentes que sugerem que a aprendizagem de categorias é otimizada por um insumo inicial de baixa variação, centrado em exemplares prototípicos (os verbos abre-alas). Isto faz com que o aprendiz ‘perceba’ as principais características presentes em outros membros de uma dada categoria, promovendo generalizações mais rapidamente (Goldberg & Casenhiser, 2008).

### 5.8.1 Muito *type* para pouco *token*

Uma outra observação importante a ser feita sobre os resultados de frequência acima exibidos está relacionada com a quantidade de *types* verbais identificados nos dados dos aprendizes e dos entrevistadores. Observe no Quadro 5 que a quantidade de *types* para os dois grupos não é muito discrepante.

| <i>Types</i> verbais | VL | VOL | VOO |
|----------------------|----|-----|-----|
| Aprendizes           | 39 | 24  | 9   |
| Entrevistadores      | 33 | 33  | 12  |

#### Quadro 5

Comparação do número de verbos usados por aprendizes e entrevistadores

A pequena diferença verificada no emprego dos tipos verbais torna-se, no entanto, extremamente significativa diante da enorme diferença observada na distribuição destes verbos ao longo dos gráficos e o pedaço do *bolo construcional* reservado a cada um deles. A distribuição das construções VL, VOL e VOO entre os verbos em cada construção se mostra mais democrática nos gráficos dos entrevistadores do que nos gráficos dos aprendizes. Se olharmos cuidadosamente o Gráfico 12, tal distribuição é bem menos *graciosa* para os aprendizes.

Por exemplo, considerando a relação verbo/total de construções VL, verificamos que no grupo de entrevistadores o verbo *go* abocanha 42%, *come* 14%, *get* 11%, *look* 7%, *live* 5%, *stay/turn* 3% das construções. No grupo de aprendizes, a porcentagem é a seguinte: *go* abocanha 54%, *come* 11%, *sit* 5%, *look* 4,81%, *get/live* 3,89%. Os demais verbos *put*, *turn*, *drop*, *move*, *fall*, *sleep*, etc, possuem uma porcentagem em torno de 1% e o restante nem a 1% chegam.

Considerando a mesma relação para as construções VOL dos entrevistadores, verificamos que o verbo *put* abocanha 35% das construções, seguido por *take* 21%, *see* 16%, *get* 6%, *bring/pick/leave* 3%, etc. Já para os aprendizes, a mesma construção é abocanhada 68% das vezes pelo verbo *put*, seguido por *take* 6%, *turn* 4%, *drop* 3%, *move* 2%. Os verbos *bring*, *catch*, *have*, *send*, *keep* e *open* possuem uma porcentagem em torno de 1% e o restante nem a 1% chegam.

Finalmente, para a construção VOO, entre os entrevistadores o verbo *give* abocanha 53%, *tell* 17%, *cost* 7%, *call* 5%, *show* 4%, *ask* 3%, etc. Para os aprendizes, a mordida de *give* corresponde a 61%, *ask/write* 8%, *buy/send* 5% das construções. Novamente, uma olhada no Gráfico 12 permite uma compreensão mais rápida do contraste aqui exibido em porcentagens (forma de ilustração que admito ser bastante enfadonha!)

Uma explicação possível para este fenômeno está tanto na idéia de protótipos e gradação observada nas categorias naturais como na própria distribuição zipfiana da linguagem. Os aprendizes nativos possuem um conhecimento mais profundo ou rico das categorias em questão (i.e., construções VL, VOL e VOO) bem com dos elementos que as compõem (i.e., os diferentes graus de tipicidade verbal em cada categoria). Apesar de continuar existindo um protótipo para cada uma destas construções, os falantes nativos e proficientes de uma dada língua não são tão dependentes de tais protótipos e são capazes de fazer uso de outros verbos que co-existem numa mesma categoria (ou seja, que contêm atributos semânticos que possibilitam a expressão de nuances e significados bem específicos, para além do sentido genérico dos protótipos).

Por outro lado, os aprendizes nos estágios iniciais de aquisição de uma língua tendem a se apoiar fortemente no protótipo para cada uma das construções analisadas (VL, VOL e VOO). Tais protótipos de fato abocanham uma parte muito grande das construções analisadas, restando uma parcela muito pequena para os outros verbos, apesar de o número de tipos verbais ser bem semelhante ao dos entrevistadores. E isto é bastante revelador da importância dos protótipos verbais nos estágios iniciais de aprendizagem, como revelam as pesquisas discutidas nos capítulos anteriores desta tese. Os dados encontrados em nosso subcorpus de aprendizes de L2 no projeto ESF confirmam os achados de Goldberg e colegas em relação à aprendizagem da L1.

Finalmente, as diferenças encontradas no emprego dos *types verbais* por parte dos aprendizes e dos entrevistadores nas três construções analisadas parecem também ter uma relação direta com o que Granger (1998) chamou de ‘noção de saliência’ (*sense of salience*). Ou seja, a divisão menos democrática do bolo construcional entre os verbos em cada construção pode ser um bom indicativo de que a noção de saliência dos aprendizes ainda é bastante fraca, e às vezes até ‘enganosa’, quando pensamos no aspecto *idiomaticidade*. No entanto, seria preciso uma análise mais detalhada deste aspecto ao longo do desenvolvimento interlingual dos sujeitos em nosso subgrupo, o que, por ora, extrapola o escopo do presente trabalho.

### 5.8.2 Para além de frequência verbal

Os dados até agora exibidos referem-se, basicamente, a uma análise longitudinal de frequência verbal nas construções VL, VOL e VOO identificadas no subcorpus investigado. Com o intuito de complementar a análise apresentada neste capítulo, que, como vimos, revelou que o desenvolvimento das três construções na língua inglesa é calcado por *tokens* de alta frequência, ou *protótipos verbais*, foi elaborado um pequeno experimento visando avaliar a percepção de prototipicidade verbal por parte de falantes proficientes da língua inglesa (nativos e não nativos). Ou seja, o estudo relatado no capítulo 6 a seguir busca determinar até que ponto os dados encontrados sobre frequência verbal encontram ressonância nas percepções de prototipicidade por parte de falantes proficientes da língua inglesa.

Como destacado por Gries, Hampe & Schönefeld (2005:02), uma análise construcional baseada apenas em frequência verbal (ou percentagem de ocorrência de certos verbos em certas construções), como mostram os dados exibidos até agora, quase sempre pode ser “inútil ou até problemática”. É importante um cálculo de contingência, ou seja, é fundamental conhecer o grau probabilístico de associação entre uma pista e um dado resultado, ou, na discussão em questão, entre uma forma e uma função linguística. Por exemplo, olhos e asas co-ocorrem em pássaros. Sempre que avistamos um pássaro, veremos também seus olhos e suas asas. A frequência de ambas as pistas é, portanto, idêntica. No entanto, as asas são mais confiáveis do que os olhos como pista na

identificação de pássaros quando observamos o cálculo de contingência, o que McWhinney (2001, 2005) chama de grau de 'confiabilidade da pista' ou *cue reliability*. Isto deve-se ao fato de que as asas de um pássaro são uma característica mais distintiva da categoria do que os olhos (ou seja, outros animais em outras categorias também possuem olhos). Assim, os olhos são menos distintivos e portanto menos confiáveis. A melhor pista será, portanto, aquela que sempre que presente num dado evento seja observado também um dado resultado. Por outro lado, sempre que tal pista estiver ausente também devemos observar a ausência de um determinado resultado. É uma questão de grau de probabilidade de pareamento entre uma pista e um evento (por exemplo, vemos olhos ou asas e temos uma experiência com pássaros ou encontramos um verbo que seja relacionado a uma dada função ou significado)<sup>63</sup>.

Embora não sejam feitos aqui cálculos de contingência utilizando métodos estatísticos nos moldes discutidos no parágrafo anterior, os capítulos 6 & 7 seguintes investigam a existência de evidências em nossos dados que sejam convergentes. Ao adotarmos uma perspectiva metodológica adicional por meio de um estudo experimental sobre prototipicidade verbal (capítulo 6) e uma análise qualitativa detalhada acerca do desenvolvimento das construções ao longo do projeto ESF (capítulo 7), buscamos, com isto, reforçar nossos argumentos relativos ao caráter ontológico das construções no âmbito dos estudos em ASL.

---

<sup>63</sup> Numa análise de contingência são geralmente empregados métodos estatísticos para o cálculo de probabilidades, como o  $\Delta P$  (Ellis, 2006a, 2006b), colexemas (Gries, Hampe and Schönefeld, 2005), dentre outros, com o intuito de medir o grau de associação entre um par de eventos. Veja Ellis & Ferreira-Junior (no prelo b) para os resultados de tais análises de contingência baseadas na pesquisa aqui descrita.

## 6. Estudo sobre prototipicidade verbal na língua inglesa

*The pervasiveness of prototypes in real-world categories and of prototypicality as a variable indicates that prototypes must have some place in psychological theories of representation, processing, and learning.*

Eleanor Rosch

Neste capítulo é descrito um estudo realizado com o propósito de aferir o grau de prototipicidade dos verbos mais freqüentes identificados na análise longitudinal discutida no capítulo anterior. O estudo buscou verificar se os resultados de freqüência verbal encontram ressonância na percepção de prototipicidade por parte de falantes proficientes da língua inglesa. O universo pesquisado, como veremos, é bastante limitado e bastante heterogêneo. No entanto, o estudo constitui-se num complemento importante acerca da validade das construções em ASL no âmbito desta pesquisa.

### 6.1 Materiais

Inicialmente, foram feitos três conjuntos de cartões com os verbos mais freqüentes em cada uma das três construções VL, VOL e VOO, identificados na análise longitudinal descrita no capítulo anterior. Desta forma, cada conjunto de cartas VL e VOL continha vinte verbos, ao passo que o terceiro conjunto de cartas continha apenas os nove verbos identificados nas construções VOO (veja Tabela 1, no Anexo A). Também foi confeccionada em cartolina uma escala de cinco pontos, ao longo da qual todos os verbos deveriam ser classificados de acordo com a sua prototipicidade (veja exemplo na Tabela 2, no Anexo A). As instruções para as tarefas envolvendo as três construções pesquisadas eram dadas no início de cada tarefa. Informações sobre os participantes (nome, sexo, idade e tempo de residência) e os respectivos resultados das tarefas de

categorização por ele realizadas eram anotadas pelo pesquisador em um formulário específico (veja Tabela 4, no Anexo A).

## 6.2 Sujeitos<sup>64</sup>

Ao todo, 15 sujeitos considerados proficientes em língua inglesa participaram do estudo. Todos residiam no estado de Michigan, Estados Unidos. Destes, oito eram do sexo feminino e sete do sexo masculino, com idades que variavam de 19 a 48 anos. Onze participantes eram norte-americanos (seis mulheres e cinco homens, com idades entre 19 e 48 anos) e quatro eram estrangeiros residentes (duas mulheres e dois homens, com idades entre 26 e 48 anos). Dentre estes últimos, três eram falantes nativos de português brasileiro e um falante nativo de espanhol da Espanha (veja Tabela 4, no Anexo A).

## 6.3 Coleta de dados

Após algumas explicações iniciais básicas fornecidas pelo pesquisador, eram apresentadas aos participantes, por escrito, as instruções para cada uma das três tarefas envolvendo as construções (veja Tabela 3, no Anexo A). A ordem das três tarefas era aleatória para todos os participantes. As cartas eram embaralhadas e, em seguida, os verbos eram apresentados um a um aos participantes. Este procedimento era repetido para cada uma das três tarefas envolvendo as construções VL, VOL e VOO. Do começo

---

<sup>64</sup> Durante exame de qualificação foi feita uma observação sobre as diferenças entre os dialetos britânico (no subcorpus ESF) e americano (no estudo em foco), o que, teoricamente, teria um impacto conceptual nos processos de categorização lingüística. Acredito, no entanto, que tais diferenças dialetais não influenciem os processos de categorização verbal no nível construcional investigado no experimento aqui relatado. Os verbos mais prototípicos para cada construção encerram e refletem os significados de nossas relações/experiências com o mundo que nos cerca e, por isto, penso que existem protótipos que são compartilhados e subjazem tais construções da língua inglesa, apesar das diferenças dialetais existentes (i.e., nos níveis fonológico, lexical, morfosintático, pragmalinguístico, etc.), algo que os próprios resultados aqui exibidos parecem revelar. Antes de serem britânicos (ou gauleses, ingleses, escoceses, irlandeses) ou norte-americanos (com seus inúmeros dialetos da costa leste a oeste, e desde o norte até o sul estadunidense), a maioria dos sujeitos do estudo pertence ao mundo anglo-saxão e partilham, de certa forma, o mesmo sistema conceptual (i.e., que inclui esquemas ou protótipos fonológicos, lexicais, morfosintáticos, pragmalinguísticos, etc.), sistema este construído a partir, principalmente, da exposição a padrões estatísticos presentes no insumo e redes de associações cognitivas bastante semelhantes, como sugere a própria teorização de Rosch (1978). Além disto, o estudo de Gries & Wulff (2005), envolvendo estudantes alemães proficientes em língua inglesa, reforça ainda mais o caráter ontológico de tais construções da língua inglesa. A meu ver, os resultados aqui apresentados reiteram as idéias whorfianas sobre a interrelação dinâmica entre linguagem, pensamento e cultura!

ao fim de cada tarefa as instruções escritas (veja Tabela 3, no Anexo A) permaneciam à vista dos participantes para consulta. Após feita a classificação de todos os verbos, o pesquisador pedia ao participante que ficasse à vontade para fazer quaisquer alterações antes de iniciar o registro da classificação dada (1,2,3,4 ou 5) a cada verbo apresentado.

#### 6.4 Resultados

Neste estudo, o grau de prototipicidade de um determinado verbo foi medido pela sua nota ponderada, i.e., a soma de todas as notas a ele atribuídas (ao longo da escala de cinco pontos) por cada um dos participantes do estudo (sendo 15 o número de participantes). Desta forma sabemos, por exemplo, que a prototipicidade do verbo *drop* na construção VL é de 2.13, ao passo que a prototipicidade do verbo *come* na mesma construção é de 1.4. Quanto menor o valor atribuído ao longo da escala de 1-5 pontos, maior a prototipicidade do verbo num dado tipo de construção.

A Tabela 6 exhibe o grau de prototipicidade (utilizando uma média simples) para cada verbo nas três construções investigadas. Em seguida, correlacionaremos os resultados de frequência verbal obtidos na análise do corpus longitudinal do Projeto ESF com os resultados de prototipicidade encontrados em nosso experimento. Como podemos observar na tabela, nem sempre o verbo considerado mais prototípico numa dada construção foi aquele que se mostrou mais freqüente (como revelado na análise de frequência do subcorpus ESF). Este fenômeno será discutido detalhadamente mais adiante neste capítulo.

Finalmente, acredito que a correlação aqui discutida possa iluminar ainda mais o papel que verbos altamente freqüentes desempenham nos estágios iniciais de aquisição de uma língua (i.e., ao se constituírem em *protótipos semânticos* que facilitam a aprendizagem de construções como categorias esquemáticas).



| VL     |      | VOL   |      | VOO     |      |
|--------|------|-------|------|---------|------|
| go     | 1    | move  | 1    | give    | 1    |
| run    | 1    | send  | 1.13 | send    | 1.06 |
| walk   | 1    | push  | 1.13 | pay     | 1.66 |
| move   | 1.13 | drop  | 1.53 | tell    | 2    |
| come   | 1.4  | put   | 1.6  | buy     | 2.26 |
| fall   | 1.86 | bring | 1.73 | explain | 2.26 |
| return | 2.06 | take  | 1.93 | show    | 2.53 |
| drop   | 2.13 | turn  | 2.4  | write   | 2.66 |
| turn   | 2.2  | leave | 3.13 | ask     | 3.53 |
| put    | 2.53 | pick  | 3.26 |         |      |
| take   | 2.66 | get   | 3.4  |         |      |
| get    | 3.53 | catch | 3.66 |         |      |
| open   | 3.66 | think | 3.8  |         |      |
| sit    | 3.86 | sit   | 3.93 |         |      |
| live   | 4.2  | open  | 4    |         |      |
| pick   | 4.2  | cross | 4    |         |      |
| look   | 4.4  | keep  | 4.26 |         |      |
| work   | 4.4  | see   | 4.4  |         |      |
| see    | 4.46 | look  | 4.53 |         |      |
| sleep  | 4.6  | have  | 4.66 |         |      |

**Tabela 6**  
Grau de prototipicidade dos verbos  
de acordo com as construções nas quais ocorrem

### 6.5 Correlacionando frequência e prototipicidade verbais no uso de construções

Os gráficos apresentados a seguir correlacionam prototipicidade e frequência no uso das construções VL, VOL e VOO em língua inglesa. Os melhores exemplares de uma dada construção seriam aqueles que exibem uma alta frequência de ocorrência, evidenciada na análise do corpus ESF, juntamente com um alto grau de prototipicidade, como evidenciado em nosso estudo experimental.

Como já destacado em nossa discussão teórica, acredito que haja um contínuo entre os processos de representação lexical e o desenvolvimento sintático, ou seja, nossas categorias semânticas possibilitam o desenvolvimento de categorias sintáticas. Do

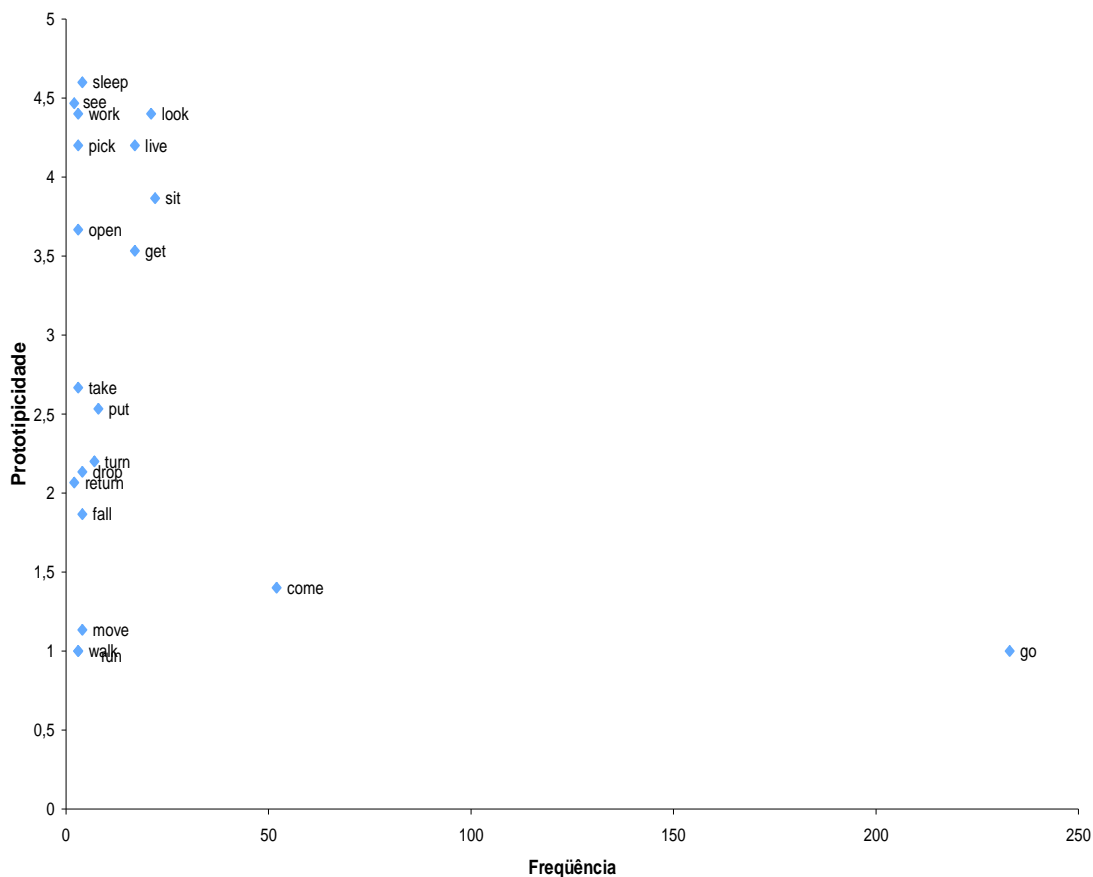
concreto ao abstrato. Tais processos possuem a mesma natureza e são regidos pelos mesmos mecanismos cognitivos de categorização e generalização.

O Gráfico 13 exibe a correlação entre prototipicidade e frequência para os verbos na construção VL. Como vimos na Tabela 6, os cinco verbos mais prototípicos na tarefa experimental foram, respectivamente, *go*, *run*, *walk*, *move* e *come*. Por outro lado, os cinco verbos mais frequentes na análise do corpus ESF foram, respectivamente, *go*, *come*, *sit*, *look* e *get*. Portanto, os melhores exemplares (ou protótipos semânticos) na construção VL são encontrados no canto inferior do lado direito do gráfico de correlação exibido, o que indica que o verbo *go* é considerado um protótipo semântico desta construção, seguido de perto pelos verbos *come*, *move*, *walk* e *run*<sup>65</sup>.

Estes resultados são semelhantes aos resultados obtidos por Goldberg, Casenhiser & Sethuraman (2004) com aprendizes de L1, como discutido no capítulo 4.

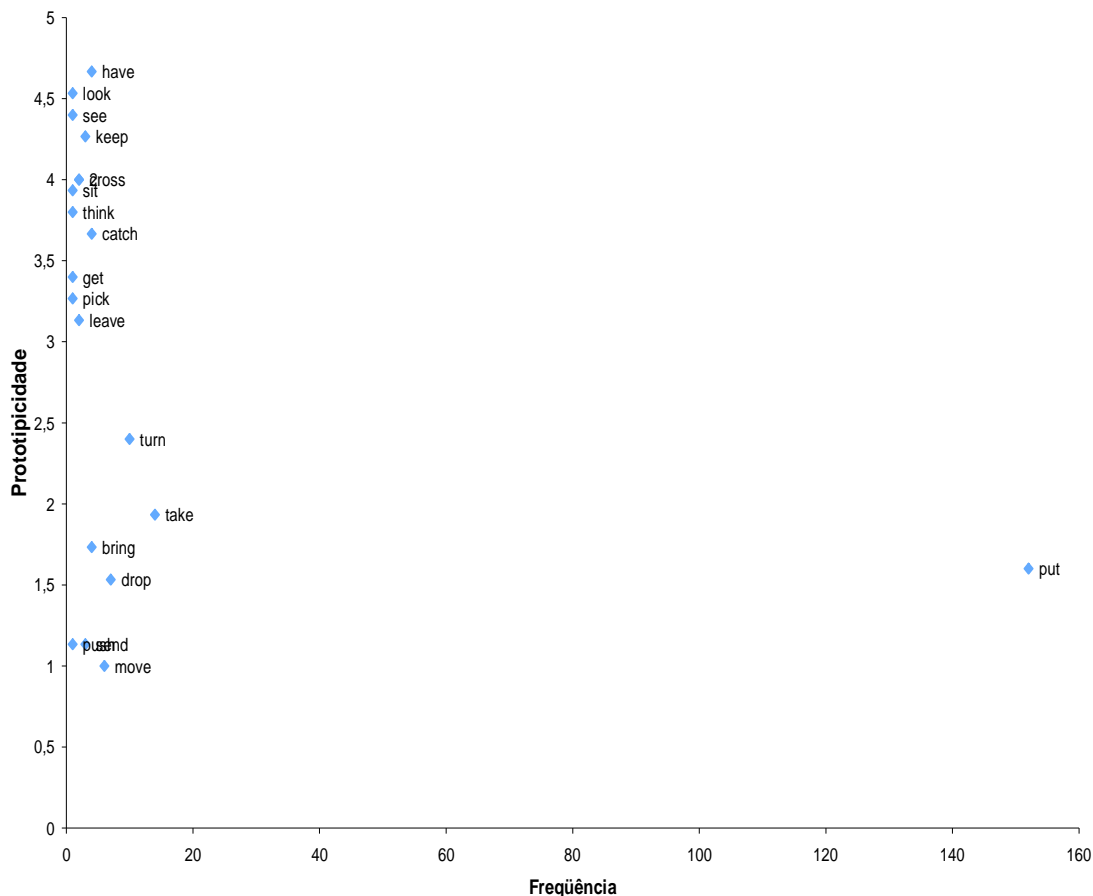
---

<sup>65</sup> Pode ter havido um efeito de *priming* influenciando o resultado do experimento para as construções VL e VOL, devido à presença da palavra **movement** na instrução dada, o que, em tese, teria influenciado (*biased*) a percepção de prototipicidade dos outros verbos na tarefa realizada, sendo privilegiado, assim, o verbo *move*. Veja instruções na Tabela 3 do Anexo A.

**Gráfico 13**

Correlação entre prototipicidade (y) e frequência (x) verbais na construção VL

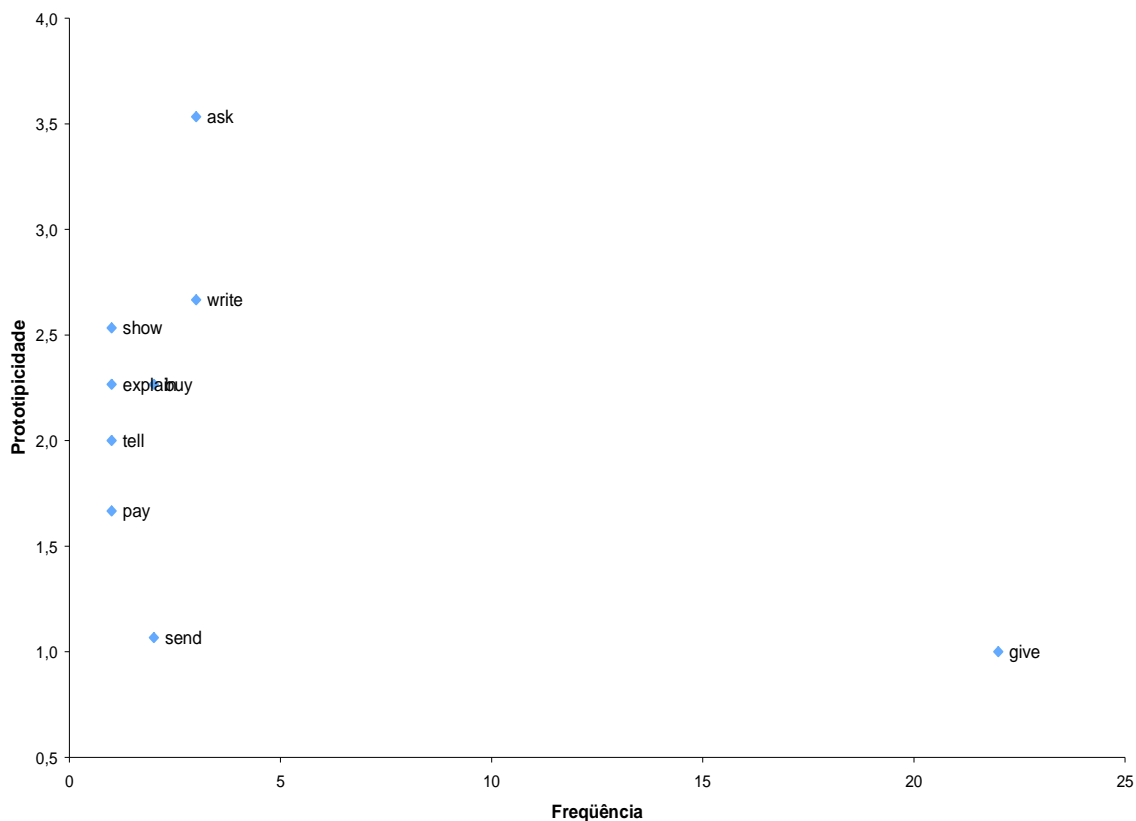
O Gráfico 14 exibe a correlação entre prototipicidade e frequência para os verbos na construção VOL. Como vimos na Tabela 6, os cinco verbos mais prototípicos na tarefa experimental foram, respectivamente, *move*, *send*, *push*, *drop* e *put*. Por outro lado, os cinco verbos mais frequentes na análise do corpus ESF foram, respectivamente, *put*, *take*, *turn*, *drop* e *move*. Desta forma, podemos observar no Gráfico 14 que os melhores exemplares (ou protótipos semânticos) na construção VL são encontrados no canto inferior do lado direito do gráfico de correlação exibido, o que indica que o verbo *put* é considerado um protótipo semântico desta construção, seguidos pelos verbos *move*, *drop* e *push* e *send*. Estes resultados também são semelhantes aos resultados obtidos por Goldberg, Casenhiser & Sethuraman (2004) com aprendizes de L1.



**Gráfico 14**

Correlação entre prototipicidade (y) e frequência (x) verbais na construção VOL

Finalmente, O Gráfico 15 exibe a correlação entre prototipicidade e frequência para os verbos na construção VOO. Como vimos na Tabela 6 acima, os cinco verbos mais prototípicos na tarefa experimental foram, respectivamente, *give*, *send*, *pay*, *sell* and *buy*. Por outro lado, os cinco verbos mais frequentes na análise do corpus ESF foram, respectivamente, *give*, *write*, *ask*, *send* e *buy*. Como vimos nas discussões anteriores, os melhores exemplares (ou protótipos semânticos) na construção VOO são encontrados no canto inferior do lado direito do gráfico de correlação exibido, o que indica que o verbo *give* é considerado o protótipo semântico desta construção, seguido pelos verbos *send*, *pay*, *tell* e *buy*. Estes resultados também corroboram os achados de Goldberg, Casenhiser & Sethuraman (2004) com aprendizes de L1.

**Gráfico 15**

Correlação entre prototipicidade (y) e frequência (x) verbais na construção VOO

### 6.6. Discussão: do genérico para o específico

Como vimos na tarefa de elicitación sobre prototipicidade relatada no início deste capítulo, cujos resultados encontram-se na Tabela 6, nem sempre o verbo considerado mais prototípico numa dada construção é necessariamente o mais freqüente (como revelado na análise de freqüência do subcorpus ESF) quando levamos em conta a classificação de tipicidade feita por falantes proficientes da língua inglesa.

Analisemos primeiro os resultados para a construção VL. Os resultados mostram que, na média, os sujeitos do estudo sobre prototipicidade deram igualmente a nota 1 para os verbos *go*, *run* e *walk*, e as notas 1.13 e 1.4 para os verbos *move* e *come*, respectivamente. Todos estes cinco primeiros verbos, assim como os demais posicionados gradativamente ao longo da tabela, descrevem relações (ou encerram o sentido) de movimento. Entretanto, alguns verbos possuem sentidos mais específicos, menos genéricos, do que outros. Por exemplo, os verbos *run* e *walk*, apesar de expressarem o sentido de movimento, carregam em si também a *maneira* pela qual tal movimento se dá. São exemplos bem mais específicos. Por outro lado, o verbo *go* possui um significado genérico e pode ser aplicado a qualquer tipo de movimento ou deslocamento para um lugar ou outra direção. Veja alguns exemplos em bom português para os verbos *ir*, *correr* e *caminhar*

|                             |                                 |                                    |
|-----------------------------|---------------------------------|------------------------------------|
| 1) <i>ir</i>                | 2) <i>correr</i>                | 3) <i>caminhar</i>                 |
| a) eu fui ao/até o banheiro | a) eu corri para/até o banheiro | a) eu caminhei para/até o banheiro |
| b) eu fui à/até a praia     | b) eu corri para/até a praia    | b) eu caminhei para/até a praia    |
| c) eu fui à missa           | c) eu corri para a missa        | c) eu caminhei para a missa        |
| d) eu fui à aula            | d) eu corri para a aula         | d) eu caminhei para a aula         |
| e) eu fui para/até a casa   | e) eu corri para/até a casa     | e) eu caminhei para/até a casa     |

Como podemos observar, algumas frases em 2 e 3 são bem específicas e refletem situações incomuns ou até estranhas (em 2c e 3c, por exemplo) quando usadas com os respectivos verbos *correr* e *caminhar*. O verbo *ir*, por outro lado, serve para praticamente todos os propósitos quando a intenção é simplesmente expressar mudanças de direção ou local, daí a sua alta frequência.

O mesmo padrão é verificado nos resultados para a construção VOL. Os sujeitos classificaram como exemplos mais prototípicos os verbos *move* (nota 1), *send* e *push* (nota 1.3) e *drop* e *put* (notas 1.53 e 1.6, respectivamente) para a situação na qual alguém é o causador do movimento de algo para um outro local ou em uma outra direção. O mesmo argumento dado acima para o verbo *go* na construção VL em relação ao caráter genérico de *put* pode ser invocado. Ou seja, alguns verbos como *push* ou *drop* são bem específicos na descrição da forma ou maneira de transportar ou deslocar algo. O verbo *put*, por sua vez, expressa de forma direta o sentido comunicativo pretendido

numa construção VOL. Desta forma, devido à sua alta possibilidade de uso em uma gama muito grande de contextos ou situações, ele torna-se um verbo muito eficaz e altamente freqüente.

O verbo *give* na construção VOO foi considerado o exemplo mais prototípico para a função comunicativa de transferir algo a alguém. Entretanto, os verbos *send*, *pay* e *tell* também contêm significados parecidos (e são igualmente percebidos e classificados como tal pelos falantes proficientes) quando a função comunicativa é, geralmente, a transferência de algo de uma pessoa para outra (i.e., construções VOO). No entanto, a preferência em relação ao verbo *give* e a sua alta freqüência de ocorrência devem-se ao fato de *give* ser, como no argumento dado em relação aos verbos *go* e *put*, um verbo altamente genérico em relação ao seus usos e, assim, possuir uma alta possibilidade de uso em situações variadas.

As análises feitas aqui são complementares àquelas conduzidas no final do capítulo 5. O argumento sobre prototipicidade e desenvolvimento de categorias semânticas também se aplica aqui, obviamente. Por exemplo, os falantes nativos (e não nativos proficientes) possuem uma grande quantidade de verbos em seu léxico mental, desde os mais genéricos aos mais específicos (resultado de processos contínuos de categorização e diferenciação, como discutidos anteriormente). No entanto, segundo Goldberg, Casenhiser & Sethuraman (2004), nos estágios iniciais da aprendizagem da L1, as crianças fazem bastante uso dos chamados “verbos leves” (*light verbs*), cujas características principais são sua generalidade semântica e alta freqüência de ocorrência. Estes verbos leves são usados como verbos abre-alas (são geralmente prototípicos). Na língua inglesa, os chamados verbos leves são *bring*, *want*, *make*, *put*, *do*, *give*, *take*, *get*, *go*, dentre outros<sup>66</sup>. Os gráficos exibidos no capítulo 5 sobre freqüência verbal nas construções VL, VOL e VOO dos aprendizes no subcorpus ESF indicam uma forte presença destes verbos, evidenciando que também nos estágios iniciais de aquisição de uma L2 tais verbos otimizam a aprendizagem de esquemas, padrões sintáticos e

---

<sup>66</sup> Observe aqui a semelhança entre a idéia de protótipos semânticos que se localizam em um nível básico de categorização e otimizam os aspectos *informativo* e *distintivo* que expressam, ou seja, não são tão genéricos a ponto de não serem informativos nem são tão específicos a ponto de não serem eficazes em relação ao aspecto distintivo, como visto na nossa discussão sobre a teoria de protótipos feita no capítulo terceiro.

categorias abstratas de uma dada língua (cf. ELLIS & FERREIRA-JUNIOR, no prelo a). Como discutido anteriormente, é importante destacar que tais “verbos leves” continuam a fazer parte do repertório de falantes proficientes, mas com bem menos intensidade, como pudemos demonstrar na discussão do capítulo 5.

No capítulo 7 a seguir será feita uma análise qualitativa dos enunciados dos sujeitos em suas tarefas durante as sessões do projeto ESF. A análise enfocará aspectos da interlíngua dos sujeitos para além do comportamento dos verbos prototípicos nas três construções por eles produzidas ao longo do estudo, enfocando também as **colocações** (co-ocorrências freqüentes) dos outros itens (i.e., pronomes e preposições) que compõem as três construções investigadas nesta tese. Ou seja, o termo **colocações** usado em nossa análise se refere à co-ocorrências freqüentes de elementos lexicais com elementos gramaticais.

As colocações, também entendidas como *colostruções*<sup>67</sup>, *coligações* ou *fraseologismos* (cf. GRIES, 2008; ELLIS, 2008), constituem-se uma ponte entre as dimensões tradicionalmente rotuladas de léxico e sintaxe e, acredito, são resultantes de processos de *chunking* que ocorrem numa dada língua desde as suas partes mais concretas até àquelas mais abstratas. O que se observa, de fato, é um *continuum* construcional que reflete uma granularidade ao longo dos vários níveis construcionais observados em uma dada língua.

---

<sup>67</sup> Embora mais adequado, o termo *colostruções* (*collostructions*) usado em Gries (2008) e Gries, Hampe & Schönefeld (2005) não é adotado na análise feita nesta tese. É certo que, diante dos argumentos e estudos aqui apresentados, o termo funde de maneira bastante criativa as noções de colocação e construção, encerrando, assim, a própria filosofia construcionista de desenvolvimento e estruturação de uma língua. No entanto, este termo é novo e ainda encontra-se em fase de consolidação. Por tal motivo, optei por usar o já bastante consolidado *colocações*.



## 7. Reflexões para além de frequência e prototipicidade verbais

*You shall know a word by the company it keeps*

John R. Firth

Neste capítulo será feita uma análise qualitativa da produção dos sujeitos nas transcrições do projeto ESF, levando-se em conta não apenas os resultados de frequência e prototipicidade verbais exibidos anteriormente, mas também as colocações mais frequentes observadas nas três construções investigadas em nosso subcorpus. A análise busca observar aspectos da interlíngua dos aprendizes relativos à frequência de colocações nas três construções por eles produzidas ao longo do estudo. Os sujeitos analisados são aqueles cujos conjuntos de dados estão completos, de acordo com as informações disponíveis na base de dados do MPI (ou seja, Andrea, Lavinia, Jarnail e Ravinder).

Finalmente, nossa análise buscará verificar se nos dados longitudinais dos quatro sujeitos acima mencionados há indícios de um desenvolvimento da L2 que se inicia em fórmulas fixas e esquemas pivô até o desenvolvimento das construções VL, VOL e VOO identificadas no estudo relatado no Capítulo 5, e que se constituem no foco principal da nossa discussão teórica. O objetivo aqui é identificar as colocações mais frequentes encontradas no entorno dos verbos prototípicos, i.e., envolvendo pronomes e preposições, colocações estas que poderíamos também entender como prototípicas pela sua alta frequência de ocorrência em tais construções. No item 7.6 retomarei esta idéia de desenvolvimento construcional que vejo metaforicamente como um “arquipélago” de colocações emergentes.

## 7.1 De volta às redes<sup>68</sup>

Ke (comunicação, 2007) demonstrou a utilidade da representação em redes na compreensão de vários fenômenos complexos, como o desenvolvimento e estruturação de uma língua. Ke conduziu uma breve análise dos nossos dados de frequência verbal e construções acima relatados em forma de redes de palavras (tais redes, chamadas de *egonets*, foram originalmente desenvolvidas e usadas para a interpretação de relações humanas e sociais complexas). Este tipo de representação é especialmente adequado à discussão teórica que embasa esta tese, baseada em redes conexionistas e no desenvolvimento de construções emergentes. Tal representação também é ideal na discussão sobre a estruturação emergente de nosso léxico mental (como vimos na proposta emergentista de Elman, 2004, discutida no capítulo 4 desta tese).

Em sua análise, Ke construiu redes (VL e VOL) utilizando todas as palavras para cada uma das construções identificadas em nosso subcorpus do Projeto ESF. O objetivo foi uma verificação inicial de colocações adjacentes, fórmulas e esquemas pivô (que seriam, como discutido anteriormente, a gênese de construções abstratas ou estrutura argumental numa dada língua).

Num exemplo de rede deste tipo exibido a seguir, as palavras são as unidades e as colocações adjacentes são os *links*. A espessura dos links entre as palavras reflete a frequência de ocorrência de uma dada colocação. Desta maneira, tem-se uma análise e representação em forma de rede que revela a frequência de co-ocorrência das palavras (e pesos entre tais conexões) nas construções VL, VOL e VOO identificadas em nossos

---

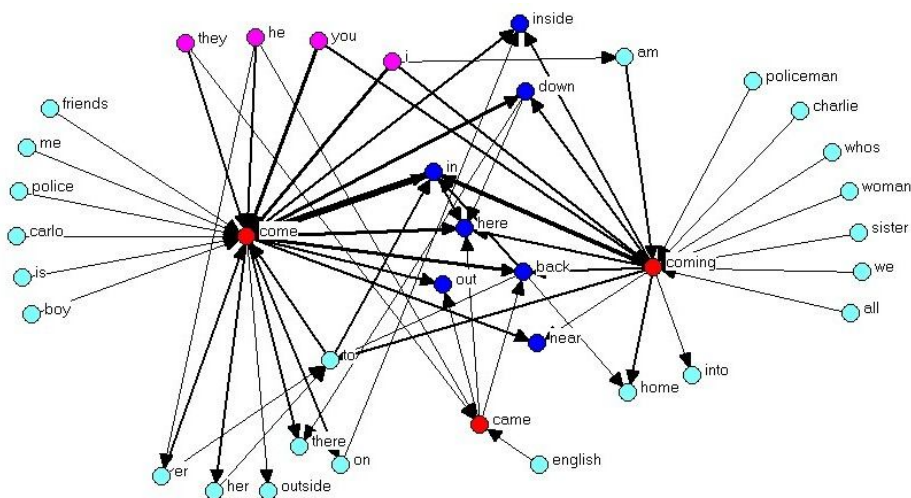
<sup>68</sup> Um lembrete importante: os gráficos em rede exibidos a partir deste ponto não são exemplos de redes conexionistas, como vimos no Capítulo 2. Entretanto, acredito que tal representação gráfica em forma de redes de conexões entre palavras (ou colocações frequentes) torna mais palpável, a meu ver, a apreensão do princípio hebbiano básico de aprendizagem (princípio fundamental da plasticidade sináptica), discutido no capítulo 1 desta tese, que se resume na idéia de que se dois neurônios forem ativados simultaneamente a conexão entre eles (sinapse) será reforçada (o que, conseqüentemente, inibirá ativações entre outros neurônios), de forma que, em ocorrências subseqüentes do mesmo estímulo, tais conexões terão um nível de ativação maior dentro do sistema neural, pois se constituirão em configurações neuronais específicas já estabelecidas ou entrincheiradas. Finalmente, reforço aqui que tais representações não são atômicas, mas **emergentes**, como vimos em nossa discussão teórica.





palavras). Portanto, utilizaremos este tipo de representação em rede ao final da análise qualitativa para cada construção feita ao longo deste capítulo com o intuito de ilustrar o papel das colocações mais freqüentes no entorno dos verbos prototípicos verificados. Com isto, buscaremos ampliar ainda mais o foco nas colocações mais freqüentes em torno do verbo prototípico usado nas construções, utilizando informações resultantes da frequência de co-ocorrência de palavras ao adicionarmos algumas restrições à rede de forma a excluir totalmente os links estatisticamente não significantes (ou seja, abaixo do valor  $\geq 5$  de ocorrência, tal qual exibido no Gráfico 18). Veja que o resultado obtido é um gráfico em rede que nos permite visualizar claramente certos padrões observados nas construções (i.e., construções ancoradas numa colocação verbal de alta freqüência ou ‘ilha verbal’ tomaselliana).

Este refinamento maior exibido no Gráfico 18 envolvendo o verbo *come* na construção VL nos permite, em última instância, visualizar mais claramente um esquema **agente-ação-locativo** (refletido nas categorias sujeito-verbo-preposição), no qual se observa a freqüência de co-ocorrência de certos pronomes no *slot/categoria agente* e certas preposições no *slot/categoria locativo*.



**Gráfico 18**

Rede de colocações de pronomes e preposições nas 52 construções VL com o verbo *come* produzidas pelos aprendizes no subcorpus ESF

Como vimos na discussão de Tomasello (2003, 2000) e outros autores, estes esquemas pivô e ilhas verbais seriam a gênese de construções envolvendo categorias abstratas, como observado na estrutura argumental de uma língua. As demais redes geradas para as construções VOL e VOO com base nos nossos dados do projeto ESF seguem o mesmo padrão verificado para a construção VL exemplificada acima, com seus respectivos verbos prototípicos e colocações específicas. Esta ilustração em redes será adotada ao longo das análises das sessões dos sujeitos no projeto ESF para as construções VL, VOL e VOO.

Nas páginas a seguir será feita uma análise qualitativa da produção dos sujeitos buscando iluminar um possível desenvolvimento da L2 que se inicia em fórmulas fixas e esquemas pivô baseados em verbos prototípicos (como vimos acima na ilustração em rede) até as construções VL, VOL e VOO identificadas. Analisaremos, principalmente, as tarefas nas quais as construções VL, VOL e VOO ocorrem mais freqüentemente com seus respectivos verbos prototípicos, ou seja, *go* na tarefa A (77/233); *put* na tarefa E (79/152); e finalmente *give* na tarefa A (8/22).

O Quadro 6 ilustra a distribuição das construções de acordo com as atividades específicas realizadas ao longo das 234 sessões do grupo

| Tarefa | VL (go) | VOL (put) | VOO (give) |
|--------|---------|-----------|------------|
| a      | 77      | 5         | 8          |
| b      | ---     | 2         | ---        |
| c      | 1       | ---       | ---        |
| d      | 9       | 8         | 2          |
| e      | 8       | 79        | ---        |
| f      | ---     | ---       | ---        |
| g      | 18      | 6         | 2          |
| h      | 9       | 16        | 2          |
| i      | 52      | 9         | 6          |
| j      | 2       | 1         | 2          |
| k      | 4       | ---       | ---        |

|              |            |            |           |
|--------------|------------|------------|-----------|
| l            | 34         | ---        | ---       |
| m            | 2          | 14         | ---       |
| n            | ---        | ---        | ---       |
| o            | 2          | ---        | ---       |
| p            | 9          | 1          | ---       |
| q            | 5          | 11         | ---       |
| r            | 1          | ---        | ---       |
| <b>Total</b> | <b>233</b> | <b>152</b> | <b>22</b> |

### Quadro 6

Tarefas nas quais os verbos prototípicos *go*, *put* & *give* ocorrem em suas respectivas construções VL, VOL & VOO.

## 7.2 Análise qualitativa do uso do verbo *go* em construções VL

Neste ítem será feita uma análise qualitativa enfocando principalmente as sessões nas quais o verbo *go* na construção VL foi mais freqüentemente encontrado, sessões estas que envolviam as atividades do tipo A (i.e., conversa livre). Foi neste tipo de atividade que a maior parte das construções VL ocorreu (o equivalente a 33,1% de todas as ocorrências nos 15 tipos de tarefas nas quais tais construções aparecem no projeto ESF, conforme ilustra o Quadro 6). No entanto, outros tipos de sessões também serão levadas em consideração ao longo da análise dos quatro sujeitos<sup>70</sup>

### 7.2.1 Andrea

Das vinte sessões de Andrea nas quais há ocorrência da construção VL, enfocaremos principalmente as sessões envolvendo a tarefa A. As outras tarefas de Andrea em que ocorrem o uso de tal construção com outros verbos (incluindo *go*) são as tarefas *e*, *k*, *l*, *g*, *i*, *h*, *q*, *j*, *d* & *q*. Veja Quadro 7 com o código do arquivo transcrito, a data de realização da sessão, mês do sujeito no estudo quando da sessão e verbos usados na construção com sua freqüência (em *tokens*) entre parênteses

<sup>70</sup> Os exemplos aqui fornecidos são trechos inalterados das sessões transcritas nos arquivos em formato .cha. Como destacado anteriormente, o leitor encontrará TODAS as sessões em CD-ROM anexo a esta tese. O leitor também encontrará no CD o programa CLAN, que deverá ser instalado para a correta leitura dos arquivos.

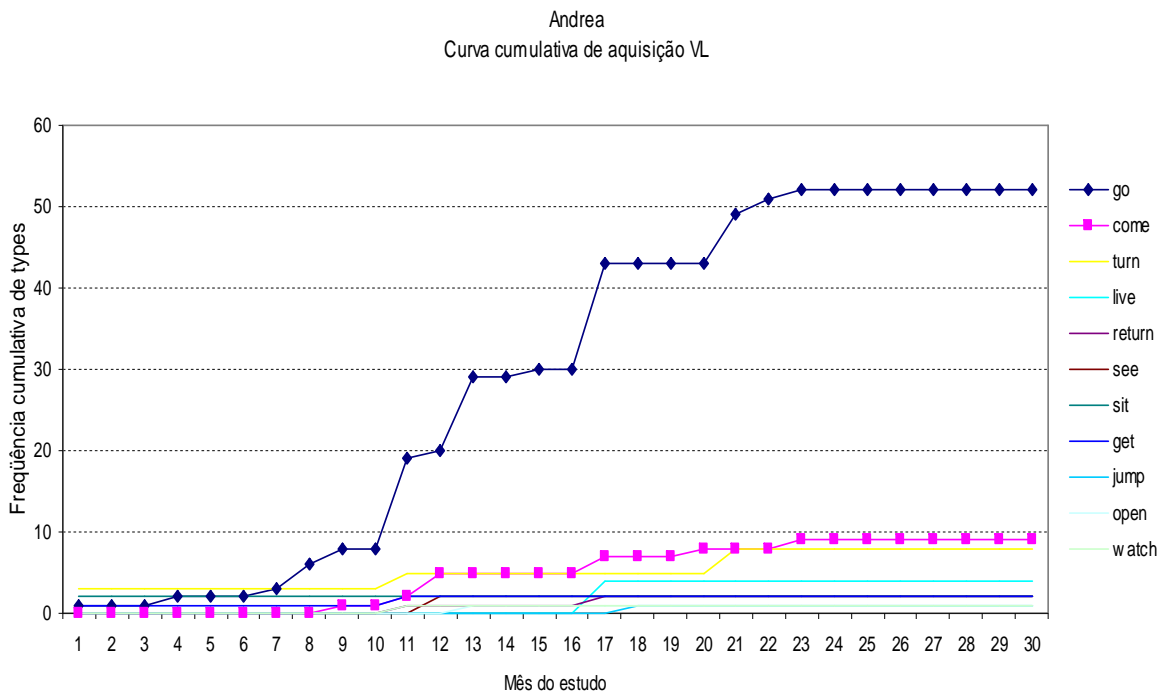
| Arquivo        | Data da sessão   | Mês no estudo | Verbos usados (frequência <i>tokens</i> ) |
|----------------|------------------|---------------|---|
| liean12e.1.cha | 10 janeiro 1983  | 1             | come (3), see (2), go (1), sit (1)        |
| liean14a.1.cha | 10 abril 1983    | 4             | go (1)                                    |
| liean16k.1.cha | 8 julho 1983     | 7             | go (1)                                    |
| liean17i.1.cha | 23 agosto 1983   | 8             | go (3)                                    |
| liean18a.1.cha | 28 setembro 1983 | 9             | go (2)                                    |
| liean22a.1.cha | 16 novembro 1983 | 11            | go (4)                                    |
| liean22g.1.cha | 16 novembro 1983 | 11            | go (6) come (2), open (1), live (1)       |
| liean22e.1.cha | 16 novembro 1983 | 11            | go (1), get (1), sit (1)                  |
| liean11a.1.cha | 13 dezembro 1983 | 12            | return (2), go (1)                        |
| liean24i.1.cha | 20 janeiro 1984  | 13            | go (7), jump (1)                          |
| liean24a.1.cha | 20 janeiro 1984  | 13            | go (2)                                    |
| liean25a.1.cha | 13 março 1984    | 15            | go (1)                                    |
| liean27i.1.cha | 10 maio 1984     | 17            | go (13), live (1), turn (4)               |
| liean31d.1.cha | 24 outubro 1984  | 22            | watch (1)                                 |
| liean32e.1.cha | 21 novembro 1984 | 23            | come (3), go (2)                          |
| liean32a.1.cha | 21 novembro 1984 | 23            | go (1)                                    |
| liean32g.1.cha | 21 novembro 1984 | 23            | go (1)                                    |
| liean32h.1.cha | 21 novembro 1984 | 23            | go (2)                                    |
| liean34q.1.cha | 21 dezembro 1984 | 24            | go (2)                                    |
| liean35j.1.cha | 15 janeiro 1985  | 25            | go (1)                                    |

### Quadro 7

Sessões de Andrea nas quais ocorre a construção VL

O Gráfico 19 ilustra o desenvolvimento de Andrea ao longo das sessões exibidas no Quadro 7. O gráfico revela, por meio de uma curva cumulativa de aquisição, o uso gradual dos verbos nas construções VL por parte de Andrea durante todo o período do projeto, construções estas principalmente ancoradas no verbo prototípico *go*. No que se segue, analisaremos detalhadamente o desenvolvimento de Andrea ao longo de algumas sessões.





**Gráfico 19**  
Curva cumulativa de aquisição VL de Andrea

*Sessão realizada em 10 de abril de 1983*

Andrea encontra-se no quarto mês do projeto. Em sua primeira sessão no estudo ESF (lean12e.cha) realizada em 10 de janeiro de 1983, Andrea parecia desconhecer verbos (especialmente verbos de movimento). Seus enunciados eram praticamente construídos com substantivos (i.e., são, essencialmente, o que poderíamos chamar de enunciados multi-palavras) e havia uma total ausência de *estrutura*, como podemos verificar na análise daquela sessão.

Nesta sessão, que consiste numa conversa livre, de caráter geral, Andrea continua monossilábico na maior parte do tempo. ‘Sim’ e ‘não’ são as palavras mais usadas no diálogo. Andrea utiliza algumas fórmulas fixas. Veja alguns exemplos:

- 1) *is very good.*
- 2) *very expensive.*
- 3) *no not really*
- 4) *i dont know*

As poucas frases produzidas são telegráficas

- 5) *# er i work only # italian restaurant.*
- 6) *one friends english.*
- 7) *yes one friend okay but sometime er er another friends.*
- 8) *is possible -\_ ?*

Também observa-se uma constante interferência da L1 na L2, como no exemplo 10 a seguir no qual Andrea quer dizer que tentou trabalhar como técnico em eletrônica por duas vezes mas não conseguiu. Ao tentar se expressar, Andrea é influenciado pelo verbo italiano *provare* na intenção de usar o verbo ‘tentar’ (*try*, em inglês)

- 9) *# i <prove # prove> [% a uses the english word prove to mean try he is being influenced in his choice by the italian \*provare\* to try] two time.*

### ***Sessão realizada em 8 de julho de 1983***

Andrea encontra-se no sétimo mês do estudo e continua fazendo uso, na maior parte de sua produção, de fórmulas e expressões fixas e semi-fixas. Nesta sessão ele vai um pouco além das frases telegráficas, veja

- 10) *i dont think very good.*
- 11) *i like to know the price of coach for er birmingham.*
- 12) *i like to see the last train for cambridge.*

Andrea quer perguntar se eles irão a Cambridge pela British Rail, e produz algumas frases que mesclam formas que revelam que seu conhecimento lingüístico neste estágio

não possui ainda uma estrutura consolidada, mas caracteriza-se, na verdade, por uma coleção de fórmulas fixas e semi-fixas, como nestes dois exemplos no qual ele expressa a noção de futuro

13) *are we british rail?*

14) *we will take british rail?*

Neste exemplo, vemos a construção VL pela primeira vez

15) *ah yes you dont go to cambridge*

### ***Sessão realizada em 28 de setembro de 1983***

Nesta nona sessão Andrea está mais comunicativo, embora utilize fórmulas fixas e semi-fixas constantemente. Nesta sessão os participantes conversam sobre o trabalho numa padaria e sobre os problemas que geralmente ocorrem com as máquinas utilizadas. Andrea usa construções VL

15) *when er come # inside some people english er # i [/?] # i tell e # of go in the office*  
< > [% laughs].

Boa parte de sua produção nesta sessão continua ainda bastante telegráfica. Também começam a aparecer outros verbos em suas frases e há uma presença constante de advérbios para expressar temporalidade, veja

16) *is big bakery?*

17) *## no eh where i work is # one bakery # small.*

18) *i remember tomorrow.*

19) *because # er ### the first week i work er # there -\_ er there is*

Como vimos no exemplo 18 fornecido, a frase não contém o auxiliar do futuro *will*, que Andrea havia usado há dois meses. Sua interlíngua continua sem muita estruturação, e

advérbios de tempo e expressões adverbiais, como *tomorrow* e *the first week*, são os únicos marcadores de temporalidade (veja Klein, 1993). Andrea também usa, pela segunda vez, uma construção VOL com o verbo *put* para dizer que na Itália eles jogam o pão velho no lixo

18) *er ah we # put in the [/?] the ru [/?] rubbish.*

### Sessões realizadas em 16 de novembro de 1983

Três sessões foram realizadas nesta data. Andrea encontra-se no décimo primeiro mês do estudo. Sua interlíngua, apesar de basicamente monossilábica e calcada em fórmulas, começa a apresentar alguma estruturação. Se olharmos os gráficos de aquisição cumulativa das construções VL e VOL de Andrea, respectivamente os Gráficos 19 e 25, constataremos que é por volta deste mês que há um aumento no uso de construções ancoradas pelos seus verbos prototípicos (*go* e *put*, respectivamente). Nos trechos exibidos a seguir, Andrea fala de uma protesto que houve em Londres. Observe as construções VL

19) *no i [/?] i [/?] i go to the westminster bridge yeah [>1] but there isn't ## people or manifestation@s [:=T demonstration--word influenced by the italin] or.*

20) *i see when i go out of my # work.*

21) *and er ## we need to ## we [/?] we cant to go in [/?] into [>1] yeah into the street where the [/?] there is the manifestation.*

22) *but er we go around.*

Nesta outro trecho da conversa, Andrea dá instruções a sua interlocutora

23) *and er # go [/?] # do you go in er that chairs # in the corner*

Já neste trecho de outra sessão, ele conta uma estória. Muitas fórmulas fixas ainda são usadas, mas Andrea está bem mais comunicativo, e sua interlíngua começa a ficar mais elaborada, como vemos nos exemplos 24 e 25 a seguir.

24) *no i dont know if er the ## blonde girl [/?] blonde woman i dont [>1] know er ring or no.*

25) *she is a friend but id [/?] she dont live er in that flat.*

Várias construções VL são produzidas nesta sessão (repare que na frase 27 ele diz *cooking*, querendo se referir a uma cozinha (*kitchen*). Mais adiante no diálogo ele próprio se corrige e produz a palavra *kitchen*.

26) *the pack is come there # in.*

27) *and er # other woman er go out the [/?] the ^cooking # for er # for see that pack.*

28) *i dont see i dont [/?] ## and er the [% prolonged] [/?] the women goupstairs.*

29) *for # for ## for drink one # coffee -\_ or [>1] tea and in this moment the black boy come # there for the window.*

#### Sessões realizadas em 20 de janeiro de 1984

Andrea já está participando do projeto há um ano e um mês. Mais construções VL são produzidas. Nesta conversa livre ele fala de sua esposa, vida pessoal e da Itália.

30) *uh yeah # she go [>1] to the dentist # for one ## previsit.*

31) *uh but i [/?] # i like to go # in italy for [/?] ### for to see my mother my sister.*

Neste trecho da conversa de uma outra sessão, a tarefa de Andrea consiste em contar o filme de Charlie Chaplin (*Tempos Modernos*) que acabara de assistir

32) *yeah # but during the [% prolonged] # er ### [% very very long pause] the charlie chaplin and the girl # er go ## to the police station.*

Mais construções VL são produzidas nestas sessões

33) *and the boat go into the # sea i dont know.*

- 34) <because er> [% prolonged] he [% prolonged] going to the restaurant
- 35) all right [>1] # and together # they go #### [% very long pause] around the city the town # again
- 36) all right and he go around [/?] #### around to the town?
- 37) and # the [/?] the movie is finish ## with er they #### go ## in one way # on one way
- 38) he working only for # i dont know # for ten minutes < > [% m laughs] and er # and he go #### go outside er this er ## # er

### Sessão realizada em 10 de maio de 1984

Como vimos acima, Andrea está tendo um bom progresso no desenvolvimento de sua interlíngua. Nesta sessão ele já está no estudo há dezessete meses. Treze construções VL com o verbo go foram identificadas. Veja algumas delas

- 39) # er you have to go out of this building
- 40) ## when you are at the other side # you go right -\_
- 41) because er i - make other way for ## go - there
- 42) but -\_ # mm #### you go left and ## at the ri [/?] er right side of this road
- 43) there -\_ < > [% laughs] er straight on of the [/?] the ladder [% # #] when you < > go down
- 44) well er i #### i go out -\_ at london bridge
- 45) when i go out of my house i turn right
- 46) when you go out # erm #### if -\_ you go out from there -\_ [/?]

### Sessões realizadas em 21 de novembro de 1984

Nestas quatro sessões Andrea encontra-se no vigésimo terceiro mês do estudo (ou seja, há quase 2 anos). Nesta primeira conversa Andrea fala de ouvir rádio, ver tv, etc. Veja duas construções VL em, seguida de outras frases

47) *er on the radio ## on [/?] on the car when i drive to the [/?] when i go to the work [/?] to [>1] the job [>2] when i come back.*

48) *no i watch only movie films [>1] ya [% laughs] [<2].*

49) *there are some week # er that the television er # show some italian [% laughs] [/?] some italian movie.*

50) *but i [% hesitant] [/?] ilike very much that movie # is a new # new new new [/?] # new [/?] new - \_ movie.*

51) *american is more difficult i think.*

52) *the [/?] there is a different accent ## in [% prolonged] a English movie.*

53) *some - \_ people er is [/?] are easy for me to [>1] understand and er ### some is more difficult.*

Como podemos observar acima, a interlíngua de Andrea está ficando bem mais estruturada. Estas sessões acima foram umas das últimas sessões de Andrea no projeto ESF. Como vimos, Andrea estava no seu vigésimo segundo mês no estudo. Sua interlíngua exibe mais complexidade (se comparada às sessões primeiras), e ele consegue ser bem mais articulado durante os diálogos. Veja trechos de sua descrição de um filme

54) *well there is this house with er # two ## two people er i think married # one man and one woman < > [% laughs] of course and [% prolonged] there is the man with er ## with one pain # in <two teeth>.*

55) *## erm ### [% long pause] and she tell him # the same # thing ## erm ### that is too long time # that sh [/?] he have to wait # he has to wait ## about three weeks er ## but he can [/?] he can do nothing ## she tell him that is possible to [% prolonged] [/?] # to - \_ go to the dentist after the time ### some dentist working until eight o clock during the evening.*

Veja mais alguns exemplos de construções VL produzidas por Andrea neste estágio

56) *come in margaret ### come - in this way between the [% prolonged] table # the big table and the [% prolonged] little table*

57) *okay ## come - in front of the table*

58) *take the little bag and the other one and er ## go outside of the +/*

59) *## and going # outside of the door*

60) *er he went to the station*

### Últimas sessões, realizadas em 21 de dezembro de 1984 e 15 de janeiro de 1985

Nestas duas últimas sessões de Andrea, realizadas quando ele estava no décimo quarto e décimo quinto meses do projeto, identificamos apenas três construções VL, mas numa análise mais ampla observamos o avanço de sua interlíngua em relação às primeiras sessões. Veja as construções abaixo

61) *## one stip and # at er that stip er # we can go at the first floor # and er here we have er er the toilets er ## man # the mechanic office*

62) *for er to go inside of the bakery*

63) *er # for fifteen days but er < > i cant go there +/*

Observe o nível de estruturação da interlíngua de Andrea nestas duas descrições feitas

64) *thats it ## well ### the bakery # where we make the bread # is er very little for er the kinds of equipment that we have there and we have [/?] we havent got enough space for working er mm # for working well # mm ### [% very long pause].*

65) *### in front of the second chair there is er # a little table ## no er square table square but er no square # square is er # er square a long # no oval # square ## rectangular hm and er on the table there are two [/?] two # newspapers # one near the other # oh under [!] the table under the legs of the table there is a a circu <lar> [?] carpet # mm ## and er # mm ### [% very long pause] at the right [!] hand of the table there is a television this television is on the floor with four little legs ## and er the the screen # the screen of the television is er # is in front of the lamp.*

Vemos, pela análise feita sobre o desenvolvimento da interlíngua de Andrea ao longo das sessões, que há um crescimento gradual de verbos usados na construção VL, que



inicialmente é dominado pelo verbo prototípico *go*. É possível observar claramente um crescimento gradual, mas constante, no uso de palavras soltas, fórmulas e expressões semi-prontas na direção de esquemas mais abstratos e com uma estrutura própria. Analisaremos, a seguir, o desenvolvimento de Lavinia no uso das construções VL ao longo do estudo.

### 7.2.2 Lavinia

Das vinte e seis sessões de Lavinia nas quais há ocorrência da construção VL, apenas sessões envolvendo a tarefa A serão analisadas mais detalhadamente. Veja Quadro 8 com o código do arquivo transcrito, a data de realização da sessão, o mês do sujeito no estudo quando da sessão e verbos usados na construção com sua frequência (em *tokens*) entre parênteses

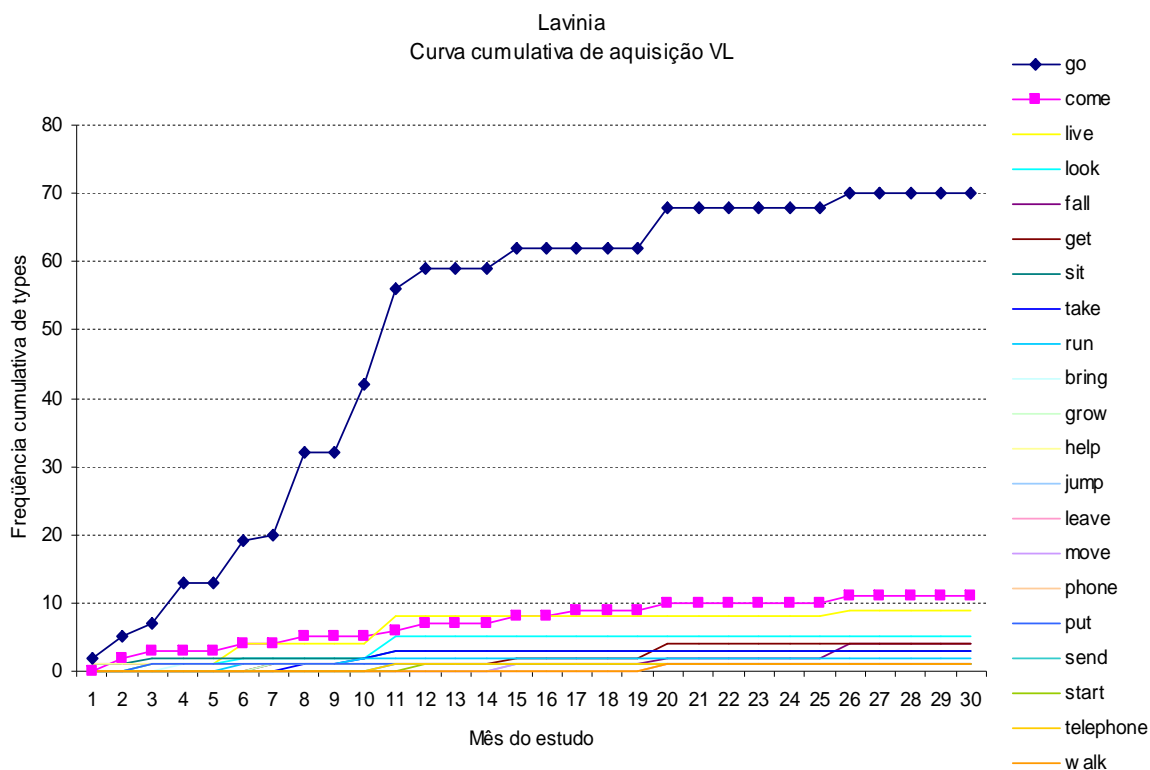
| Arquivo        | Data da sessão    | Mês no estudo | Verbos usados (frequência <i>tokens</i> )        |
|----------------|-------------------|---------------|--|
| liela11a.1.cha | 23 agosto 1983    | 1             | go (1 token), help (1 token)                     |
| liela11d.1.cha | 23 agosto 1983    | 1             | sit (1 token) go (1 token), fall (1 token)       |
| liela12g.1.cha | 22 setembro 1983  | 2             | come (2 tokens)                                  |
| liela12a.1.cha | 22 setembro 1983  | 2             | go (3 tokens)                                    |
| liela13e.1.cha | 18 outubro 1983   | 3             | come (1), look (1) put (1), sit (1 token)        |
| liela13c.1.cha | 18 outubro 1983   | 3             | go (1 token)                                     |
| liela13a.1.cha | 18 outubro 1983   | 3             | go (1 token), live (1 token)                     |
| liela14i.1.cha | 16 novembro 1983  | 4             | go (5 tokens), bring (1 token)                   |
| liela14m.1.cha | 16 novembro 1983  | 4             | go (1 token)                                     |
| liela16a.1.cha | 20 janeiro 1984   | 6             | go (6), come (1), live (3), look (1), run (1)    |
| liela17a.1.cha | 29 fevereiro 1984 | 7             | go (1 token), get (1 token), grow (1 token)      |
| liela18l.1.cha | 27 março 1984     | 8             | go (4 tokens), come (1), take (1 token)          |
| liela18h.1.cha | 27 março 1984     | 8             | go (5 tokens)                                    |
| liela18a.1.cha | 27 março 1984     | 8             | go (3 tokens)                                    |
| liela21p.1.cha | 10 maio 1984      | 10            | go (7 tokens), run (1 token)                     |
| liela21a.1.cha | 10 maio 1984      | 10            | go (1 token), take (1 token)                     |
| liela22g.1.cha | 11 junho 1984     | 11            | go (3 tokens), come (1 token), live (3)          |
| liela22a.1.cha | 11 junho 1984     | 11            | go (8 tokens), live (1), send (1), telephone (1) |
| liela22e.1.cha | 11 junho 1984     | 11            | go (3 tokens), look (3), sit (1), take (1)       |

|                |                  |    |   |
|----------------|------------------|----|---|
| liela23a.1.cha | 13 julho 1984    | 12 | go (3 tokens), come (1 token)                   |
| liela23b.1.cha | 13 julho 1984    | 12 | start (1 token)                                 |
| liela24a.1.cha | 24 outubro 1984  | 15 | go (3), come (1), get (1), jump (1), move (1)   |
| liela26a.1.cha | 6 dezembro 1984  | 17 | come (1)  |
| liela31g.1.cha | 11 maio 1985     | 22 | go (1 token)                                    |
| liela31i.1.cha | 11 maio 1985     | 22 | go (3), come (1), fall (1)                      |
| liela31a.1.cha | 11 maio 1985     | 22 | go (2), get (2), leave (1), phone (1), walk (1) |
| liela25i.1.cha | 13 novembro 1985 | 28 | go (2), fall (2), come (1), live (1)            |

**Quadro 8**

Sessões de Lavinia nas quais ocorre a construção VL

Observe também o Gráfico 20 que ilustra o desenvolvimento de Lavinia ao longo dos meses do estudo em relação ao uso dos verbos detectados nas construções VL, estas ancoradas principalmente no verbo prototípico *go*, por meio de uma curva cumulativa de aquisição



**Gráfico 20**

Curva cumulativa de aquisição VL de Lavinia

### Primeiras sessões, realizadas em 23 de agosto de 1983

Lavinia inicia o estudo com uma capacidade de comunicação em língua inglesa bem maior do que a dos outros participantes. Ela já estava morando em Londres há sete meses quando do início das sessões e tinha facilidade em interagir com falantes nativos em seu dia a dia (confira informações sociobiográficas no Anexo B). Veja a seguir dois exemplos de construções VL produzidas por Lavinia

- 1) *and i go to the nursery and speak with other people*
- 2) *he is go in*

### Sessões realizadas em 22 de setembro de 1983

Já nas sessões deste segundo mês, há várias evidências de que a interlíngua de Lavinia encontra-se num estágio mais sofisticado de desenvolvimento. Com relação à construção VL, desde a primeira sessão o verbo *go* vem sendo usado nesta construção. Veja mais exemplos

- 3) *and [/?] and after er perhaps i go in another school*
- 4) *yeah er yeah he # went to the school*
- 5) *yeah now i go to [/?] to this school*
- 6) *and is not possible for her to come inside # and erm she telephone she phone to [/?] # i dont know the fire guard i dont know the [/?] the er*

### Sessões realizadas em 18 de outubro de 1983

Terceiro mês de estudo e vemos mais algumas construções VL. No entanto, apesar de possuir uma grande capacidade de comunicação neste estágio inicial de aprendizagem da L2, a estruturação dos enunciados produzidos por Lavinia possui também um caráter telegráfico. Além disto, ela também faz bastante uso de fórmulas. Veja exemplos

- 7) <so [/?] dyou want to go out xxx [>1] > [% to e] ?
- 8) i mm i like it to m to study m to study it to m go to the xxx courses
- 9) but havent time because in the morning doing this course # i go home i clean m  
just a little bit er i it er xxx i take my son # i havent had a time.
- 10) i dont know ## im not sure [>1].
- 11) full time?
- 12) i dont know [<1].

### Sessões realizadas em 16 de novembro de 1983

Quarto mês no estudo e Lavínia continua a fazer uso constante de construções VL com o verbo *go*, principalmente. Veja exemplos

- 13) and [<1] the boat go out the port go in +/
- 14) and after and after er # he goes [/?] he going out because er he ### uh its  
difficult < >
- 15) and er # he go [/?] go out quickly because the [>1]
- 16) but the man and the woman er # er i dont know the name er they dont go in the  
prison they < > [% laughs]
- 17) he going # on the ladder
- 18) yeah some prisoners er they want to go out to the prison

### Sessão realizada em 20 de janeiro de 1984

Lavínia encontra-se no sexto mês do estudo. Nesta sessão a conversa gira em torno da comparação entre a vida na Inglaterra e na Itália. Observe algumas construções

- 19) in summer is very good because you can go to the sea everyday

20) *they say is better there underground [/?] the french underground is better is more fast [/?] faster its faster [>1] i dont know # i went to paris but i dont remember*

21) *yeah [<1] maybe i go in italy*

22) *yeah but er i went to madame tussaud*

23) *yeah but they know underground because before they went [/?] this summer to paris*

24) *yeah go out [<1] maybe some girls find apartment for home [/?]*

#### Sessão realizada em 29 de fevereiro de 1984

Neste sétimo mês, Lavinia permanece utilizando muitas fórmulas durante suas conversas (trechos 26 e 27). Uma única construção VL com o verbo *go* é identificada

25) *i went to ^wood green shopping centre the evans*

26) *would you like [/?] ^five?*

27) *how much do you pay a lot?*

#### Sessões realizadas em 27 de março de 1984

Nas construções VL com o verbo *go* mostradas abaixo, podemos observar que Lavinia emprega o passado irregular (e também o passado regular *phoned*) dos verbos de maneira bastante adequada. Também, a partir deste estágio de desenvolvimento, há um aumento e predominância de construções VL ancoradas no verbo *go*

28) *< > after that ## i went to a travel agency*

29) *<yes i went to the bank>*

30) *about erm # about er to find a plane for or i asked are about er # because i had to go in the town <i dont remember the name>*

31) *after that i went to a bakery*

- 32) *and he said why you dont go # to italy*
- 33) *but i phoned to my mother yesterday about this # and er I felt very sad because i know my mother # doesnt want i go in italy very much*
- 34) *i went to the bank and i asked to the clerk about the change xxx*
- 35) *i would like to go next month because my husband # go for [/?] go out of london for ten day*
- 36) *on the stairs you # you go out*
- 37) *you go [/?] after the playground you go [/?] there is # a door # you go through the door*
- 38) *you go straightaway ?*
- 39) *you go straightaway on the left*

#### Sessões realizadas em 10 de maio de 1984

Lavinia encontra-se agora no décimo mês do estudo. A conversa é sobre a ida do filho de Lavinia ao hospital e alguns planos profissionais. O diálogo ocorre durante o almoço. Nos trechos abaixo, envolvendo construções VL com o verbo *go*, há exemplos de uma construção com o verbo *run* e uma outra construção VOL com o verbo *take*. Também a fórmula *I have to* é bastante usada nestas sessões. Observe os trechos

- 40) *after breakfast we run to school because usually its er its nine oclock # or five past nine # i have to go to mm # to xxx # to ## i have to go to school m with my son*
- 41) *after i take him to [/?] into school i # i go to school < > [% interruption]*
- 42) *after the school i i shop [/?] go for shopping [/?] i go shopping or er if i have to go ## in [/?] in another place i go for a walk er something like this i go home mm*
- 43) *i go to school*
- 44) *if there is a nice day we go to the park or we go # for a walk together*
- 45) *monday he went to the dentists for the last time for some filling and now hes stopped until september for a check up*

- 46) *on if mm for example # a monday next monday [>1] # n i go to school*  
 47) *or if i got an appointment go to the doctor or #*  
 48) *outside the police station [>1] one was cleaning the window and the other one  
 [<1] was just er mm # going into the police station*  
 49) *xxx yeah i was erm [/?] i was just erm going out to the park when i see this  
 happen*

### Sessões realizadas em 11 de junho de 1984

Décimo primeiro mês de Lavinia no estudo. Sua interlíngua está visivelmente mais complexa e seu ritmo de desenvolvimento é bem superior ao dos demais sujeitos. Lavinia já utiliza construções variadas envolvendo o passado simples (irregular e regular), o futuro com *will*, o *present perfect*, o *past continuous*, etc. Há uma profusão de usos (às vezes gramaticalmente ‘corretos’ e às vezes com algum ‘erro’ estrutural). Em relação à análise específica das construções VL, há uma preponderância evidente do uso destas com o verbo *go*. Veja exemplos

- 50) *# i go to the electricity board*  
 51) *## my husband went in # some places*  
 52) *and the friend suggest to her # maybe it was better if the [% prolonged] [/?]  
they go downstairs*  
 53) *and when she went out ## the door closed*  
 54) *can you go in the south of london ?*  
 55) *in the meantime she went out of the kitchen*  
 56) *on wednesday we are going to kew garden*  
 57) *she went -\_ in the room er elisa went in the room*  
 58) *she went in the room*  
 59) *two or three times a week # i say i want to go # there -\_ # he say # no -\_ you  
 have to go # in the south of london*  
 60) *ya they [/?] <ma@s [:=T but] > i went # there ma@s [:=T but] they say we  
 have to pay if you dont pay we cut er out[/?]*

### Sessão realizada em 13 de julho de 1984

Lavinia está no décimo segundo mês. Faz uso de muitas fórmulas, como *I think so* e *you know*. Nos dois diálogos analisados aqui, não há muita variação em relação ao padrão verificado nas sessões do mês anterior. Veja os exemplos de construções VL com o verbo *go*

61) *and she told me shes going [/?] hes going to ## to go to the first class*

62) *hes going to school yes*

63) *to go back again*

### Sessão realizada em 24 de outubro de 1984

Décimo quinto mês. A interlíngua de Lavínia continua progredindo, embora de maneira irregular e não linear. Como mencionado anteriormente, ela às vezes faz uso ‘correto’ do passado dos verbos (irregular e regular) e da forma progressiva (ex., *went, was, saw, happened, asked, is working, is showing, are doing*) para, num momento posterior, utilizar formas ‘incorretas’. O uso de construções VL continua ancorado no verbo *go*, como vemos nas ocorrências abaixo.

64) *^oh mh # i dont explain what is it like because i went only to the seaside you know*

65) *and then you see and then you stop again er [>1] and not ## i cant cope a lot you know some people got three children they go off to work er out you know is not # i cant [>2] you know*

66) *i used to go # mm i used to go to my friend house for instance*



### Sessões realizadas em 11 de maio de 1985

Estas sessões foram realizadas sete meses depois, quando Lavinia encontrava-se no vigésimo segundo mês do estudo. O progresso de Lavinia em relação às primeiras sessões é evidente. Veja estes enunciados nos quais a construção VL é usada com o verbo *go*

67) *and not very clear # and then when he went home and the wife ask him # if he phone to the dentist he said yes mm [% hesitation] but he told her he had appointment in three weeks time so the wife said you know # you mm # you can change your appointment because it [/?] its two oclock in the afternoon so <it must take it must leave the [/?] the [/?] the work you know without xxx [/?] is not going to get paid and erm the man you know say no mm # not going to change and things like that*

68) *because maybe the prison was a nice safe place you know he can have ssome food you know things like that so he did that you know he went in the rrestaurant and a big lunch everything and call a policeman and the policeman call the erm ## policeman van i dont know to collect all the the call it er the the*

69) *because the chair the legs of the chair they they they # they go inside you know then they they make some holes*

70) *but er at first i wasnt so upset you know [>1] because # im not interest in this kind of [/?] but when i got there in the office [% #DfI#] they said to me they havent got any course er at the moment they got just just with shorthand in it and er # so i was so disappointed and [>2] i go again to the job centre*

71) *i went there about two weeks ago*

72) *very horrible and they went there # started living together # erm xxx you know*

### Sessão realizada em 13 de novembro de 1985

Esta é a penúltima sessão de Lavinia no estudo e a última na qual há ocorrência da construção VL com o verbo *go*. A interlíngua de Lavinia continua a progredir, embora exiba um padrão de desenvolvimento irregular e não linear, com a recorrência de uso de formas atípicas (i.e., a variedade básica). No entanto, observa-

se que a intensidade deste *backsliding* é gradativamente reduzida ao longo do estudo (pelo menos no que podemos observar nas sessões analisadas). A interlíngua de Lavinia é cada vez mais idiomática e exibe um padrão bem próximo ao de falantes nativos. Veja abaixo dois exemplos de construção VL com o verbo *go*

73) *thats right yes [<1] and he was very hungry and he went in a rrestaurant and [>2] he had a lot of things and when was the time to pay the bill he saw a [/?] a policeman [>3] in the road and he call him*

74) *what happened oh ## mm when he [/?] went er [/?] out of the prison he couldnt find a job*

Nas análises conduzidas nas sessões de Lavínia, verificamos um desenvolvimento mais rápido das estruturas de sua interlíngua. É visível que sua proficiência no uso da L2 é superior a de Andrea, mas algo comum aos dois pode ser claramente notado: o desenvolvimento parte de fórmulas fixas, expressões semi-fixas e evolui gradualmente para esquemas e estruturas sintáticas, cujas palavras em determinados *slots* tornam-se cambiáveis. Outro aspecto importante deste desenvolvimento é a sua irregularidade, como podemos perceber claramente ao longo das sessões. O uso do verbo prototípico *go* nas construções VL é alto e predomina desde o início, mas tende a ser gradualmente acompanhado por outros verbos.

### 7.2.3 Jarnail

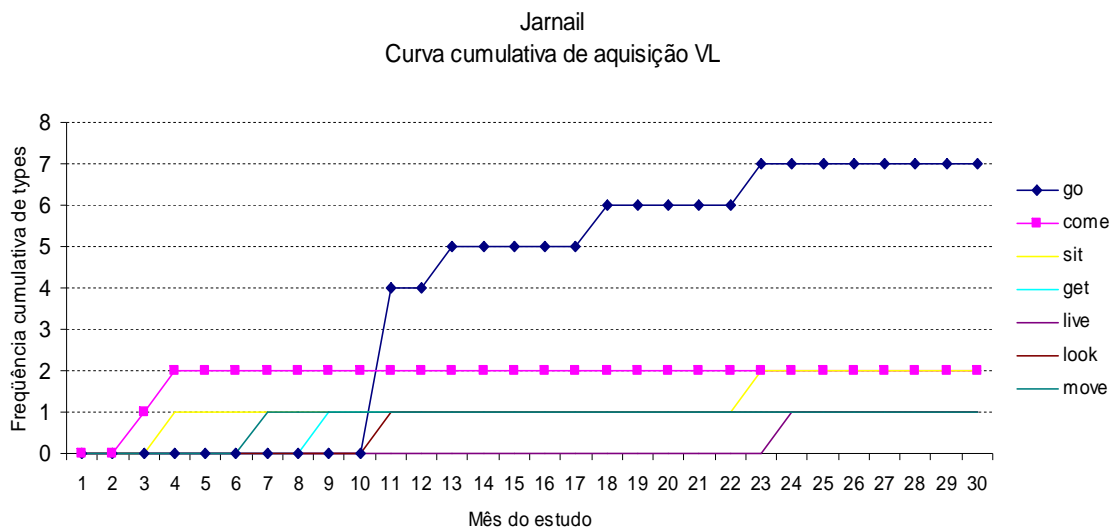
Foram verificadas construções VL em oito sessões de Jarnail, sendo a ocorrência maior com o verbo *go* nas tarefas A. Outras construções VL produzidas por Jarnail envolvem os verbos *come*, *move*, *get*, *sit* e *look*. O Quadro 9 exibe o código do arquivo transcrito, a data da realização da sessão, mês do sujeito no estudo quando da sessão e verbos usados na construção com sua frequência (em *tokens*) entre parênteses

| Arquivo        | Data da sessão    | Mês no estudo | Verbos usados (frequência <i>tokens</i> ) |
|----------------|-------------------|---------------|---|
| lpeja12g.1.cha | 19 janeiro 1983   | 3             | come (1 token)                            |
| lpeja13e.1.cha | 12 fevereiro 1983 | 4             | come (1 token), sit (1 token)             |
| lpeja15a.1.cha | 4 maio 1983       | 7             | look (1 token)                            |
| lpeja16i.1.cha | 13 julho 1983     | 9             | get (1 token)                             |
| lpeja17a.1.cha | 6 setembro 1983   | 11            | go (4 tokens), move (1) live (1)          |
| lpeja26h.1.cha | 11 abril 1984     | 18            | go (1)                                    |
| lpeja33a.1.cha | 30 novembro 1984  | 25            | go (1)                                    |

**Quadro 9**

Sessões de Jarnail nas quais ocorre a construção VL

Veja também o Gráfico 21 que ilustra o desenvolvimento de Jarnail ao longo dos meses do estudo em relação ao uso dos verbos em construções VL, ancoradas principalmente no verbo prototípico *go*, por meio de uma curva cumulativa de aquisição



**Gráfico 21**

Curva cumulativa de aquisição VL de Jarnail

### Sessão realizada em 19 de janeiro de 1983

Jarnail está no terceiro mês do estudo quando da realização desta sessão. A atividade consistia em assistir a um filme e, em seguida, comentá-lo. A produção de Jarnail é telegráfica, totalmente calcada em substantivos. Jarnail praticamente não utiliza verbos. Seus enunciados são praticamente construídos com substantivos (i.e., são, essencialmente, o que poderíamos chamar de enunciados multi-palavras), e há uma total ausência de *estrutura*. Veja seus comentários quando perguntado sobre o filme

- 1) *yes and door open no and stuck door door stuck # erm # telephone and*
- 2) *<hm> [?] ## and police er telephone # and door stuck ###*
- 3) *## no erm the fire telephone number and some other telephone number.*
- 4) *two lady house.*

A única construção VL nesta sessão envolve o verbo *come* (um dos raros verbos usados)

- 5) *police.s ## erm police come here*

### Sessão realizada em 12 de fevereiro de 1983

Jarnail está em seu quarto mês no estudo. Jarnail faz uso principalmente de gestos e algumas expressões de comando para conseguir comunicar suas intenções durante esta sessão. Nesta tarefa (i.e., dar instruções) sua produção é bastante telegráfica. Jarnail utiliza basicamente recursos paralingüísticos juntamente com um substantivo e/ou locativo. Duas construções VL foram identificadas nesta sessão

- 6) *alright you come in*
- 7) *and sit down*

**Sessão realizada em 4 de maio de 1983**

Jarnail está no sétimo mês do estudo. Ainda bastante telegráfica, calcada em substantivos e com total ausência de estrutura, sua interlíngua parece não ter progredido muito desde a última sessão. Veja alguns exemplos de enunciados produzidos neste trecho da sessão no qual ele quer dizer que sua esposa, grávida, parou de trabalhar e fica sozinha em casa (e ele também precisa fazer companhia para ela)

8) *er er six and week and stop.*

9) *school # in my wife in.*

10) *home # me in school n wife in sick*

11) *me and left out.*

Neste outro trecho, ele fala que seu carro às vezes tem problemas para dar partida pela manhã

12) *car alright.*

13) *in morning in trouble.*

14) *first in second in not xxx.*

15) *no in sometime tick tick tick tick.*

Finalmente, identificamos uma construção VL apenas.

16) *<look out in my house>*

**Sessão realizada em 13 de julho de 1983**

Nono mês no estudo. Nesta sessão Jarnail comenta o filme *Tempos Modernos*, de Charlie Chaplin. Embora comece a utilizar alguns verbos, o desenvolvimento de sua interlíngua continua lento e apresenta as mesmas características discutidas nas sessões anteriores

- 17) *er ## two in skinhead # erm policeman er n # wrong an lock.*
- 18) *er n mister charlie on policeman and help.*
- 19) *yes erm mister charlie and policeman and help.*
- 20) *er # policeman er mister charlie and job working.*
- 21) *# and # mister charlie and # working start.*
- 22) *good job ## sorry no understand # mister charlie and working /wrong.*
- 23) *mister charlie and no working.*

Observe uma construção VL invertida (Jarnail é falante de Punjabi)

- 24) *daughter and mister charlie police station in go [/?] go.*

Em seguida, a entrevistadora pergunta a Jarnail se Charlie Chaplin e a garota saem da delegacia (*they go out?*) onde estavam detidos. Jarnail responde com uma construção VL nesta sessão.

- 25) *get out*

### **Sessão realizada em 6 de setembro de 1983**

Este é o décimo primeiro mês de Jarnail no estudo. Vários verbos começam a ser usados, em especial o verbo *go* na construção VL, apesar de a interlíngua de Jarnail permanecer bastante telegráfica no início da sessão. Por exemplo, observe a ausência de estrutura neste trecho no qual ele diz que vai a Índia para o casamento das irmãs e que por isto venderá o carro

- 26) *no erm s january february in this india.*
- 27) *and er my two sister.*
- 28) *s eighteen nineteen year in married.*
- 29) *er n car er sale.*

Mais adiante na sessão, a conversa passa a ser sobre o trabalho de Jarnail e a escola de sua filha. Veja as construções VL (e os primeiros usos do verbo *go* nesta construção). Jarnail produz alguns enunciados bastante estruturados

30) *and from there to the office where i sign they said that I they said that i should go back from there i was directed to the job centre*

31) *at first they sent me to the social security office from there i went to the labour exchange*

32) *october in go india*

33) *then she will go to another school a bigger school*

34) *they lived in caldmore they are appointed by the government*

35) *it all happened because of a change of address i had moved from the lxxx street*

#### **Sessão realizada em 11 de abril de 1984**

Décimo oitavo mês no estudo. Apesar de passados sete meses desde a última sessão, a interlíngua de Jarnail parece não progredir muito. Veja abaixo uma construção VL com o verbo *go*

36) *hes very nice er n s next go er n # railway station*

#### **Sessão realizada em 30 de novembro de 1984**

Esta é a última sessão na qual identificamos uma construção VL. Jarnail estava em seu vigésimo quinto mês no estudo e exibia pouco progresso em sua interlíngua (o desenvolvimento era muito lento, como observado em outras sessões aqui analisada), embora estivesse gradualmente incorporando verbos, fórmulas e alguma estrutura à sua interlíngua. Veja dois exemplos

37) *and the finish in this may.*

38) *its no trouble and its easy.*

A única construção VL identificada é esta

39) *yeah erm you [/?] you go to the big post office for that*

Na análise do progresso de Jarnail ao longo dos meses, podemos perceber um ritmo totalmente diferente daquele observado para Andrea e Lavinia. É certo que houve algum progresso, mas o tempo de estruturação da interlíngua de Jarnail é bem diferente. No entanto, algumas características do desenvolvimento de Jarnail são observadas também em Andrea e Lavinia: a linguagem telegráfica, a total ausência de estrutura, o uso de fórmulas e algumas expressões semi-fixas e, finalmente, indícios de estruturação de construções VL principalmente a partir do décimo mês, com o uso frequente do verbo *go*.

#### 7.2.4 Ravinder

Vinte e duas sessões de Ravinder continham construções VL. O verbo prototípico para esta construção é o *go*, mas as tarefas nas quais ele ocorre mais comumente variam bastante, não havendo uma preponderância de um tipo de tarefa específico (como se verifica nos demais sujeitos). Faremos uma análise por data da sessão, de forma a capturar o desenvolvimento da construção VL com o verbo *go*. Veja o Quadro 10 com o código do arquivo transcrito, a data de realização da sessão, mês do sujeito no estudo quando da sessão e verbos usados na construção com sua frequência (em *tokens*) entre parênteses

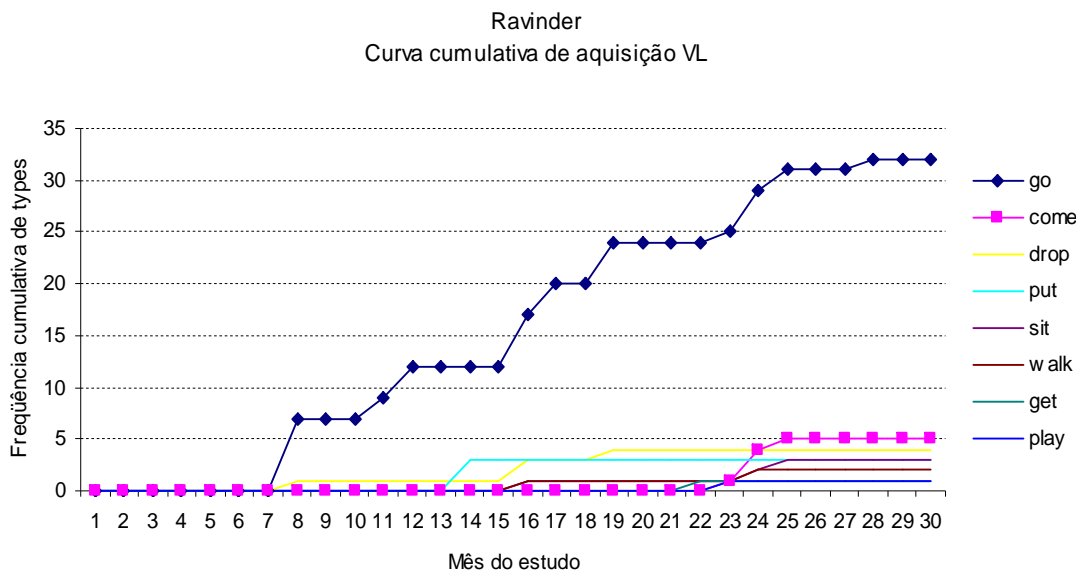


| Arquivo        | Data da sessão    | Mês no estudo | Verbos usados (frequência <i>tokens</i> ) |
|----------------|-------------------|---------------|---|
| lpera12d.1.cha | 20 janeiro 1983   | 3             | play (1 token)                            |
| lpera16i.1.cha | 8 junho 1983      | 8             | go (7 tokens), drop (1 token)             |
| lpera17a.1.cha | 7 setembro 1983   | 11            | go (2 tokens)                             |
| lpera21o.1.cha | 19 outubro 1983   | 12            | go (2 tokens)                             |
| lpera21k.1.cha | 19 outubro 1983   | 12            | go (1 token)                              |
| lpera23m.1.cha | 8 dezembro 1983   | 14            | put (3 tokens)                            |
| lpera24i.1.cha | 14 fevereiro 1984 | 16            | go (5), drop (2), sit (1), walk (1)       |
| lpera25d.1.cha | 29 março 1984     | 17            | go (1 token)                              |
| lpera25a.1.cha | 29 março 1984     | 17            | go (1 token)                              |
| lpera25j.1.cha | 29 março 1984     | 17            | go (1 token)                              |
| lpera27h.1.cha | 10 julho 1984     | 21            | go (1 token)                              |
| lpera27l.1.cha | 10 julho 1984     | 21            | go (1 token)                              |
| lpera31d.1.cha | 24 julho 1984     | 21            | drop (1 token)                            |
| lpera31p.1.cha | 24 julho 1984     | 21            | go (2 tokens)                             |
| lpera32g.1.cha | 17 outubro 1984   | 24            | get (1 token)                             |
| lpera33a.1.cha | 5 novembro 1984   | 25            | come (1 token)                            |
| lpera33q.1.cha | 5 novembro 1984   | 25            | go (1 token)                              |
| lpera34a.1.cha | 19 dezembro 1984  | 26            | go (1 token), come (1 token)              |
| lpera34l.1.cha | 19 dezembro 1984  | 26            | come (2 tokens)                           |
| lpera34i.1.cha | 19 dezembro 1984  | 26            | go (3 tokens), sit (1), walk (1)          |
| lpera35l.1.cha | 22 janeiro 1985   | 27            | go (2) come (1), sit (1)                  |
| lpera36r.1.cha | 24 abril 1985     | 30            | go (1 token)                              |

### Quadro 10

Sessões de Ravinder nas quais ocorre a construção VL

O Gráfico 22 ilustra o desenvolvimento de Ravinder no uso de construções VL, ancoradas principalmente no verbo *go*, por meio de uma curva cumulativa de aquisição



**Gráfico 22**  
Curva cumulativa de aquisição VL de Ravinder

### Sessão realizada em 20 de janeiro de 1983

Ravinder encontra-se no terceiro mês do estudo. Em sua primeira sessão, no dia 19 de novembro de 1982, Ravinder demonstrava uma baixa compreensão do seu entrevistador e limitava-se, quando compreendia algo, a dizer sim ou não. Poucas palavras foram produzidas. Nesta sessão, que consistia em descrever uma foto que mostrava lojas e prédios no entorno de uma praça, os enunciados de Ravinder são semelhantes aos dos outros sujeitos (exceto Lavinia) neste estágio inicial do estudo: uso de fórmulas fixas, enunciados bastante telegráficos, calcados em substantivos (multi-palavras) e com total ausência de estrutura. Veja exemplos

- 1) *i dont know.*
- 2) *fruit shop.*
- 3) *ap [/?] furniture shop er carpet anything.*
- 4) *that one co^ffee - \_ ?*
- 5) *play anything*

Quando perguntado o que a criança na foto estava fazendo ele usa o verbo *play* (um dos poucos verbos usados nesta sessão) numa construção VL

6) *play on the ^dog*

### Sessão realizada em 8 de junho de 1983

Ravinder encontra-se no oitavo mês do estudo. A atividade desta sessão consistia em assistir a um filme (*Modern Times*, de Chaplin) e comentá-lo. Verbos começam a aparecer nos enunciados de Ravinder, especialmente o verbo *go* na construção VL, permitindo uma certa estruturação de sua interlíngua (i.e., para além dos enunciados multi-palavras). Veja alguns exemplos

7) *hes #### drop on the floor # and charlie said girl you go out ## xxx #/ <x>*

8) *# er police #### charlie go to jail .*

9) *## charlie said you go in*

10) *and charlie # go to work # in ship*

11) *hes # go to police station*

12) *hes #### drop on the floor # and charlie said girl you go out ## xxx #/*

13) *hes you ## go to # work and letter xxx*

14) *m er i think <> [\$ laughs] go to back in jail*

### Sessões realizadas em 7 de setembro e 19 de outubro de 1983

Nestas sessões de Ravinder, aos onze e doze meses do estudo, a interlíngua de Ravinder parece estar progredindo lentamente. Ravinder usa mais verbos, utiliza fórmulas fixas (i.e., *its alright, i dont mind*) e uma estruturação canônica das palavras começa a aparecer (ou seja, sintaticamente adequada como o enunciado '*sometime have cup of tea*'). O uso do verbo *go* na construção VL predomina, veja abaixo

15) *tell me how go to walsall town centre # er best way go to ## town centre you tell me.*

16) *# i dunno < > [\$ laughter] hes go to school and*

17) *# xxx she say you go to /school or something # like class*

18) *go to the # police station*

19) *you tell me miss # how you go to birmingham by train*

### Sessões realizadas em 14 de fevereiro e 29 de março de 1984

Estas sessões, entre o décimo sexto e décimo sétimo meses do estudo, envolvem as atividades de relatar/comentar um filme assistido (*Modern Times*, de Chaplin), conversa livre, descrição de fotos/lugares. Ravinder começa a usar bastante o passado de alguns verbos (irregulares, inicialmente, como *said*, *broke*). Veja algumas construções VL abaixo (com preponderância do verbo *go*)

20) *and ## girl is #### and #### after < > its like er xxx ## its everybody push an girl is hot like its everybody push and policeman # in stand an the back # its push # pushin in # charlie and girl and policeman drop on the road < < > [\$ coughs] > [>1]*

21) *drop on the floor*

22) *# and # after plan # go to jail*

23) *and # coming in van # and go to jail*

24) *and < > its charlie said to girl is you # go away alright*

25) *and its [/?] plan is go to back to jail*

26) *its er < > its no plan its go to restaurant*

27) *and < > sit on the # one side ##*

28) *and go # wife # walk on the road*

29) *## go inside <i think> [>1] <ya> [<2]*

30) *and # # second in # # # we go to one # hospital*

31) *i go to # manor hospital*

### Sessões realizadas nos dias 10 e 24 de julho de 1984

Vigésimo primeiro mês. As mudanças na interlíngua de Ravinder perceptíveis. Ele começa a ampliar o leque de preposições, pronomes e fórmulas usados, exibindo claramente uma interlíngua mais bem estruturada em termos sintáticos. Veja alguns exemplos abaixo

32) *yeah <phone call thats why> [>1] he cant go outside.*

33) *i think his friend you know.*

34) *he said you wait here you tell me er which room is er xxx*

35) *<well i dont know> [\$ laughingly] mh what you call you know number.*

36) *i can understand a lot you know because erm # m eh ## i can understand lot what i speak [/?] i cant speak a lot you know m # i cant speak a lot like you know.*

37) *## mh i am not sure but i thinks # er garlic or.*

38) *i want red apple.*

A seguir, algumas construções VL verificadas nestas sessões

39) *er sorry hes gone to school you know # hes go to school # <erm> [>1]*

40) *its m er with somebody er only memory and i go that way*

41) *its go to town centre or something # mh*

42) *thats mohammed go to # m church masjid@s*

43) *hes crying # well drop in the floor*

### Sessões realizadas em 19 de dezembro de 1984

Vigésimo sexto mês. Pouco progresso pode ser observado nesta sessão em relação às sessões anteriores, realizadas em Julho. Cinco meses se passaram e Ravinder mantém o mesmo padrão de estruturação de sua interlíngua (uma progresso altamente irregular e

não-linear). Veja um exemplo de enunciado produzido por Ravinder ao contar o filme *Modern Times*, de Chaplin

44) *hen its [/?] now two children coming its give this chocolate # xxx then its policeman finish he said you pay for it he said no money # and say pay for it and policeman xxx < > [\$ r clears his throat] then its van coming over # and its go in van # and one # its lot people here in van.*

45) *[\$ r clears his throat] its # charlie give to seat to him and stand up he said you know me.*

46) *er he said its im bread man you know # and < > [\$ r clears his throat] # then its ### girl is push the policeman # and drop it in floor.*

Abaixo, algumas construções VL

47) *well < > [\$ ] my sister coming to er in this country from india*

48) *yeah its er [/?] if you [/?] i am coming into mangat house*

49) *yeah its er if you i am coming into mangat house*

50) *### ill go somewhere holiday # so ## if y i go to india*

51) *< > when in [/?] charlie said you go away # its n going m girl said you coming as well*

52) *and # after go inside ###*

53) *its er # two girl stay there and big girl # go away*

54) *mh and after like # sit down the xxx table stand*

55) *and its go as well and then ## its walk down the n road # and ## sitting [/?] sit down on one [/?] # and by house ## its talking and looking # then # its er think wife and husband # its husband go to # that xxx work somewhere*

### Sessão realizada em 22 de janeiro de 1985

Vigésimo sétimo mês. Nesta sessão a tarefa consistia na reprodução de uma foto/imagem e, à medida que Ravinder ia reproduzindo cada item da foto, os entrevistadores pediam para que ele nomeasse os itens ou objetos desenhados. Portanto,

é uma tarefa que envolve compreensão, elicitación de vocabulário e referência espacial (ao relatar a localização e forma dos itens). As construções VL abaixo foram identificadas

56) ### right < > come in downstairs

57) ## < <thats> i think going to town

58) three step m er when you # go in house

59) <girl> [<1] one girl # sit on the floor

### Sessão realizada em 24 de abril de 1985

Esta é a última sessão de Ravinder, realizada em seu trigésimo mês no estudo. A tarefa consistia em relatos de Ravinder sobre suas impressões acerca de seu próprio processo de aquisição da língua inglesa ao longo do projeto ESF. Pela natureza da atividade, há poucos enunciados nesta sessão. Abaixo, a única construção VL identificada

60) *can you come any [/?] go to school in this country*

A análise do desenvolvimento da interlíngua de Ravinder confirma o padrão verificado em Andrea, Lavinia e Jarnail. O desenvolvimento de sua interlíngua é mais rápido do que o de Jarnail, embora contenha características encontradas nos demais: inicialmente mostra-se monossilábico, se apóia também no uso de fórmulas e faz uso de linguagem telegráfica. A utilização de gestos também é comum, e, posteriormente, começa a empregar fórmulas semi-fixas e expressões sintaticamente estruturadas. No que concerne a construção VL, tal construção passa a ser bastante usada a partir do décimo mês com a incorporação constante do verbo *go*. Outros verbos também são utilizados nesta construção, mas com menos intensidade.

### 7.2.5 Discussão sobre o desenvolvimento da construção VL

Como vimos acima, ao fazermos uma análise utilizando gráficos em rede das sessões dos quatro sujeitos investigados para a construção VL com o verbo *go*, é possível perceber que o padrão de desenvolvimento construcional constitui-se num gradual acúmulo de algumas **colocações** altamente freqüentes. Este **alicerce construcional emergente** parece ser guiado por alguns princípios de estruturação da interlíngua baseados, inicialmente, em recursos pragmáticos, seguidos pelo emprego de fórmulas, expressões lexicalizadas e expansão gradual de verbos, atingindo, finalmente, o nível mais complexo de enunciados estruturados em torno de verbos prototípicos e uma interlíngua sintaticamente estruturada (i.e., guiada por padrões sintáticos abstraídos do insumo).

Na análise qualitativa que fizemos, observamos claramente um padrão de desenvolvimento interlingual gradual e altamente irregular, que vai do concreto ao mais abstrato, mas não é possível detectar facilmente a natureza das colocações no entorno dos verbos. Estas co-ocorrências de palavras emergentes tornam-se saliente apenas quando analisamos os dados de forma quantitativa e probabilística (i.e., em forma de redes de associação e de co-ocorrência de certas colocações de palavras).

Klein (1993:74), após análise sobre o desenvolvimento/aquisição da expressão temporal na interlíngua dos sujeitos do projeto ESF, destaca que em cada um destes aparentes ‘estágios’ de desenvolvimento<sup>71</sup> há um princípio organizacional inerente ao sistema interlingual, ou, nas palavras do próprio Klein

Em cada estágio, a língua do aprendiz não é apenas um acúmulo aleatório de formas específicas mas um *sistema* autônomo – uma variedade do aprendiz que é regida por um número de princípios organizacionais distintos. Esta é a primeira sistematicidade. O processo de aquisição é uma seqüência de variedades do aprendiz, o qual, por sua vez, segue certas regularidades. Esta é a segunda sistematicidade. A forma de expressão concreta destes dois tipos de sistematicidade depende de um número de fatores causais – princípios cognitivos gerais, natureza das

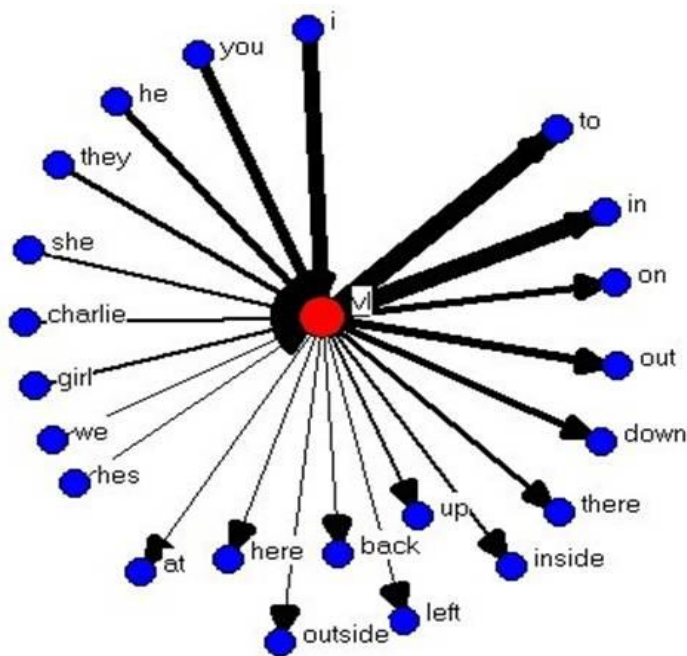
---

<sup>71</sup> É sempre importante lembrar que tais ‘estágios’ são altamente irregulares (no tocante às formas da interlíngua exibidas) e seu desenvolvimento caracteriza-se pela não linearidade.



línguas fonte e alvo, condições individuais e sociais de aprendizagem, dentre outros (grifo no original)<sup>72</sup>

Nosso foco aqui é a construção VL, por isto ilustraremos as colocações mais freqüentes no entorno do verbo prototípico *go*. Como relatado no início deste capítulo, Ke (2007) fez uma análise em rede das construções VL utilizando informações resultantes de um cálculo probabilístico de co-ocorrência de palavras, adicionando algumas restrições à rede de forma a excluir as colocações estatisticamente não significantes (ou seja, abaixo do valor  $\geq 5$  de freqüência de ocorrência). Assim, o Gráfico 23 exibe uma rede gerada para o verbo prototípico *go* levando-se em conta tais restrições estatísticas de freqüência de co-ocorrência de palavras



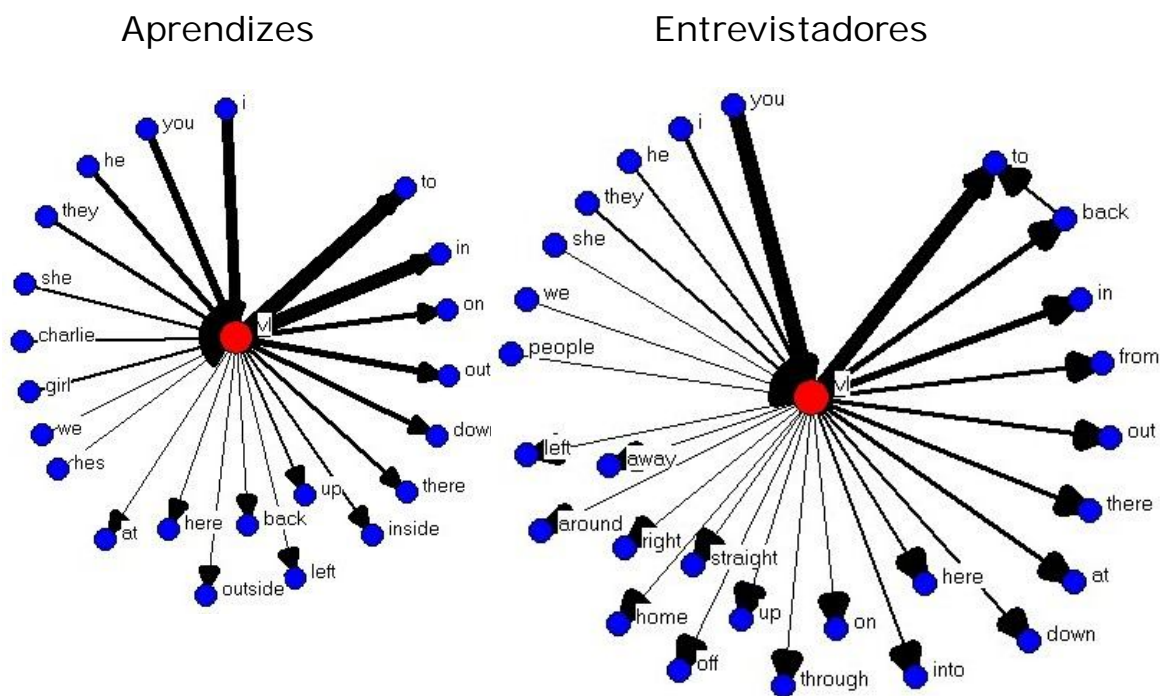
**Gráfico 23**

Rede de colocações mais freqüentes em torno do verbo *go*  
na construção VL para o grupo de aprendizes (reproduzido de Ke, 2007)

<sup>72</sup> Minha tradução de: *At each point, the learner's language is not just a random accumulation of individual forms but a system in its own right – a learner variety which is governed by a number of distinct organizational principles. This is the first systematicity. The acquisitional process is a sequence of leaner varieties, which in turn follows certain regularities. This is the second systematicity. What these two types of systematicity concretely look like depends on a number of causal factors – general cognitive principles, nature of source language and of target language, individual and social learning conditions, and others.*

Como vemos, o resultado obtido é um gráfico em rede que nos permite verificar que há, inicialmente, uma padrão observado na construção VL ancorado no verbo prototípico *go* (uma ‘ilha verbal’ tomaselliana), que se traduz no uso de um esquema agente-ação-locativo (refletido nas categorias sujeito-verbo-preposição), no qual se observa a co-ocorrência muito freqüente dos pronomes *I, you, he* e *they* no slot agente e das preposições *to, in, on, out* e *down* no slot do locativo. Como vimos na discussão de Tomasello (2003, 2000), estes esquemas pivô e ilhas verbais seriam a gênese de construções envolvendo categorias abstratas, como observado na estrutura argumental de uma língua.

Finalmente, encerrando esta ilustração dos nossos dados em forma de redes, veremos abaixo o Gráfico 24, reproduzido da análise de Ke (2007), que possibilita uma excelente comparação das redes de colocações (pares adjacentes) na construção VL no entorno do verbo *go* para os dois grupos pesquisados: aprendizes e entrevistadores



**Gráfico 24**

Comparação das redes de colocações mais freqüentes em torno do verbo *go* na construção VL para os grupos de aprendizes e entrevistadores. In: Ke (2007)

Na comparação entre as redes de colocações dos aprendizes e dos entrevistadores (falantes nativos da língua inglesa), podemos ver claramente como as redes são semelhantes, embora a ilha dos entrevistadores seja visivelmente mais rica e diversificada. A rede dos entrevistadores exhibe colocações ligeiramente mais distribuídas. Isto pode ser constatado pela espessura dos *links* entre as palavras, ou seja, quanto mais espessa menos distribuída e centrada em alguns pronomes. Entretanto, há inegavelmente uma superposição de pronomes e preposições; destas últimas, dos vinte tipos de preposições verificados nos dados dos entrevistadores, treze estão também presentes nos dados dos aprendizes.

### **7.3 Análise qualitativa do uso do verbo *put* em construções VOL**

Neste ítem será feita uma análise qualitativa das sessões nas quais o verbo *put* na construção VOL foi mais freqüentemente encontrado (sessões que envolvem as atividades do tipo *E*, que consistia em dar instruções). É neste tipo de atividade que a maior parte das construções VOL ocorrem (o equivalente a 52% dos 18 tipos de atividades realizadas no projeto ESF).

#### **7.3.1 Andrea**

Foi verificada a ocorrência de construções VOL em dez sessões de Andrea. Além da tarefa *E*, as outras tarefas de Andrea em que ocorre o uso de tal construção são as tarefas *k*, *l*, *a*, *g*, *i*, *d* & *q*. Analisaremos detalhadamente cinco sessões ao longo dos meses com o objetivo de capturar o desenvolvimento de construções VOL. Veja Quadro 11 com o código do arquivo transcrito, a data de realização da sessão, o mês do sujeito no estudo quando da sessão e verbos usados na construção com sua freqüência (em *tokens*) entre parênteses<sup>73</sup>

---

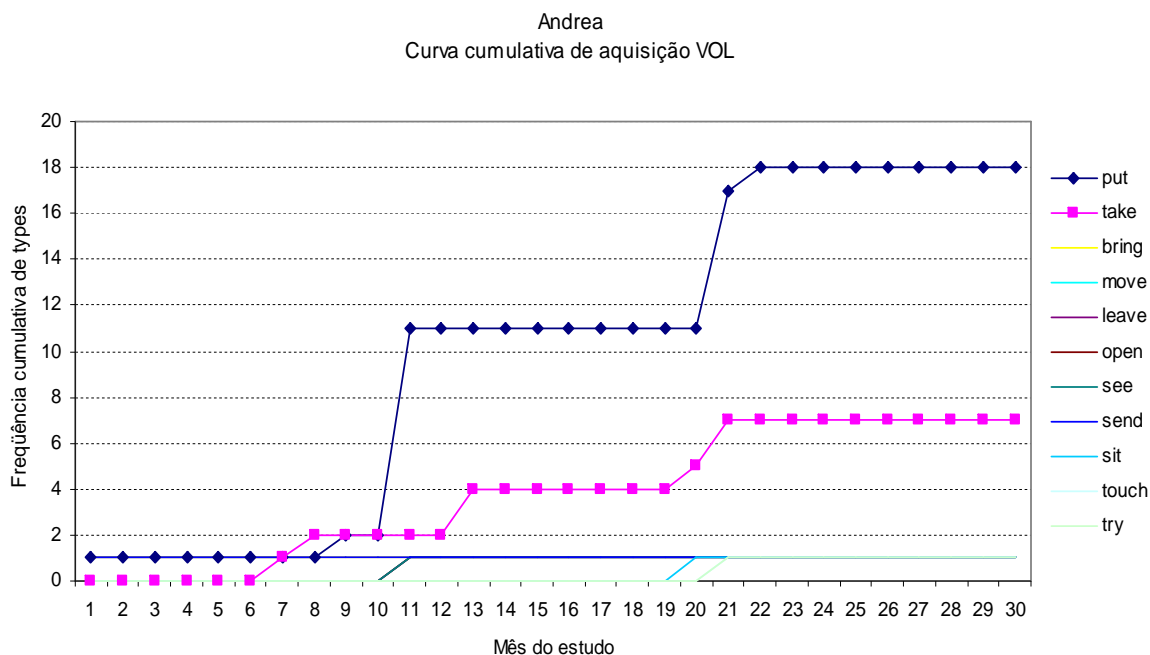
<sup>73</sup> Os exemplos aqui fornecidos são trechos inalterados das sessões transcritas nos arquivos em formato .cha. Como destacado anteriormente, o leitor encontrará TODAS as sessões em CD-ROM anexo a esta tese. O leitor também encontrará no CD o programa CLAN, que deverá ser instalado para a correta leitura dos arquivos.

| Arquivo        | Data da sessão   | Mês no estudo | Verbos usados (frequência <i>tokens</i> )       |
|----------------|------------------|---------------|---|
| liean12e.1.cha | 10 janeiro 1983  | 1             | bring (4), put (1), see (1)                     |
| liean16k.1.cha | 8 julho 1983     | 7             | take (1)  |
| liean17l.1.cha | 23 agosto 1983   | 8             | take (1)  |
| liean18a.1.cha | 28 setembro 1983 | 9             | put (1)   |
| liean22g.1.cha | 16 novembro 1983 | 11            | put (1), open (1)                               |
| liean22e.1.cha | 16 novembro 1983 | 11            | put (8), take (2), leave (1), keep (1), try (1) |
| liean24i.1.cha | 20 janeiro 1984  | 13            | take (2)  |
| liean31d.1.cha | 24 outubro 1984  | 22            | send (1), take (1)                              |
| liean32e.1.cha | 21 novembro 1984 | 23            | put (6), take (5), touch (1), sit (1)           |
| liean34q.1.cha | 21 dezembro 1984 | 24            | put (1)   |

**Quadro 11**

Sessões de Andrea nas quais ocorre a construção VOL

O Gráfico 25 ilustra o desenvolvimento de Andrea no uso de construções VOL, inicialmente ancoradas principalmente no verbo *put*, por meio de uma curva cumulativa de aquisição



**Gráfico 25**

Curva cumulativa de aquisição VOL de Andrea

***Primeira sessão, em 10 de janeiro de 1983***

Esta sessão é a primeira de Andrea no estudo ESF. Andrea parece desconhecer verbos (especialmente verbos de movimento). Seus enunciados são praticamente construídos com substantivos (i.e., são, essencialmente, o que poderíamos chamar de enunciados multi-palavras). A comunicação é bastante telegráfica e Andrea faz uso freqüente de gestos, como nestes trechos em que Andrea pede primeiro que a entrevistadora coloque a jaqueta na cadeira e depois que ela coloque (vista) a jaqueta (neste último, trecho 4, vemos que ele usa o italiano *indossare la giacca* para dizer ‘vista a jaqueta’ ou ‘put the jacket on’, em inglês)<sup>74</sup>

- 1) *er your jacket. e@s [:=T and] jacket er on the seg@s [:=T from the italian \*seggia\* chair].*
- 2) *er after get up bring the er jacket ## <i dont know> [% a is having problems expressing himself and is aware of it] er indossare@s la@s giacca@s [:=T put the jacket on] yes [% a indicates to m that she is performing the correct actions].*

Neste outro trecho, Andrea pede para colocar a caneta dentro da revista, sem utilizar verbos, apenas usando gestos e contato visual

- 3) *e@s [:=T and] er # this pens in the magazine.*

No trecho a seguir, Andrea pede que sua interlocutora coloque a bolsa dela na cadeira. Observe que Andrea não formula a frase “coloque a/sua bolsa na cadeira”, mas chama sua interlocutora com a fórmula *come near*, utilizando gestos e contato visual para se comunicar

- 4) *come near this seg@s [:= from the english \*seggia\* chair] yes er the bage [% #]# in the [/?] on the <> [% laughs].*

---

<sup>74</sup> Os comentários entre colchetes foram feitos pelo pesquisador que originalmente transcreveu as sessões.

Mais adiante no diálogo, Andrea começa a usar bastante o verbo *bring* com o objetivo de fazer com que sua interlocutora mova ou coloque algo em algum lugar, ou seja, construções VOL. Veja os exemplos:

5) *bring this seg@s* [:=T from the english \*seggia\* chair] erm # *in nother position*.

6) *bring one magazine in the bag*.

7) *bring this er pens*. Er # *this pens in the magazine*.

8) *see again in the window after bring* [/?] *bring one magazine in the bag*

Andrea emprega, pela primeira vez, o verbo *put* numa construção VOL:

9) *<put> [?] magazine in the bag*.

### ***Sessão realizada em 23 de agosto 1983***

Andrea encontra-se no oitavo mês do estudo. Uma de suas tarefas nesta sessão é dar instruções a sua interlocutora sobre como chegar a lugares específicos (*route description*). Andrea tem dificuldades em comunicar suas intenções. Usa fórmulas (*I don't understand*) e alguns poucos verbos (*take, go, turn*) começam a ser usados. Com exceção do uso de alguns verbos, sua interlíngua parece não avançar em relação à primeira sessão. Veja a única construção VOL identificada

10) *when i finish the irving road i take the left # and er i take the train for # green park*

### ***Sessão realizada quatro meses depois, em 16 de novembro 1983***

Andrea está participando do estudo há dez meses e começou a usar outros verbos na construção VOL. O verbo *bring*, que era usado quase que sempre para pedir que algo fosse movido, deslocado ou colocado em algum lugar, praticamente desapareceu nesta

construção, ao passo que o uso do verbo *put* aumentou consideravelmente (8 tokens). Andrea também faz uso agora dos verbos *take* (2 tokens), *leave* (1 token), *keep* (1 token) e *try* (1 token). Parece que Andrea está começando, de maneira bastante tímida e gradual, a desenvolver a construção *X causes Y to move Zpath/location* baseado no verbo *put*. Veja exemplos a seguir:

11) *put the silver in the ^bag.*

12) *put er on the floor.*

13) *put the bag with the silver # over the ### [% long pause] over the ##  
cuscino@s [:=T cushion].*

14) *put the box on the shelf*

15) [% long pause] *no good < > [% laughs] take the er the money ### and  
er try now in that [!] plant ### the big one # on the left. Yes and put the money #  
over ### [% long pause] yeah take the money ### and now try # under [!] the  
## the vase [% #veiz#] va [/?]*

#### ***Sessão realizada em 20 de janeiro de 1984***

Andrea encontra-se no décimo terceiro mês no projeto. Nesta sessão, a tarefa consiste em contar o filme de Charlie Chaplin (*Tempos Modernos*) que Andrea acabara de assistir. O uso de fórmulas é constante (por exemplo, *I don't know, all right, I dont understand*, etc) e o progresso de sua interlíngua é imperceptível em relação à sessão anterior. Seus enunciados exibem um padrão telegráfico. Duas construções VOL foram identificadas nesta sessão, ambas com o verbo *take*

16) ### [% long pause] *and er # take him # charlie chaplin # in this house*

17) *and after er # other policeman # with him # take him for # the police station  
for other things*

*Sessão realizada 10 meses depois, em 21 de novembro 1984*

Esta sessão foi realizada no vigésimo terceiro mês de participação de Andrea no estudo. A construção VOL permanece dominada pelo verbo *put* (6 tokens). O verbo *take* continua a ser gradativamente utilizado nesta construção (2 tokens) e surgem novos verbos *touch* (1 token) e *sit* (1 token). Uma construção esquemática mais complexa do tipo *take X and put in Zpath/location* ou *X causes Y to move Zpath/location* parece estar surgindo e sendo incorporada à interlíngua de Andrea. Veja exemplos a seguir:

18) *take the [% prolonged] the plate # under the cup. And put that over [!] the cup.*

19) *take the box and er put that near the legs of the ## seige [%  
sezEsuffix chair] # behind the table.*

20) *take -\_ the ashtray on the table and put that into the bag.*

Esta nova construção (ou sua complexificação) é uma boa evidência de que Andrea já conseguiu abstrair a construção VOL e agora já é capaz de incrementar ou expandir tal estrutura esquemática para além do uso do verbo prototípico *put*, ao fazer uso de um outro verbo, de forma a mudar, levemente, a função/significado comunicativo pretendido na interação. Veja exemplo retirado desta sessão:

4) *take the little bag and the other one and er ## go outside of the +/. And going  
# outside of the door.*

A sessão acima foi uma das últimas sessões de Andrea no projeto ESF. Como vimos, Andrea estava no seu vigésimo terceiro mês no estudo. Sua interlíngua exibe mais complexidade (se comparada à sessão primeira), e ele consegue ser bem específico em certas instruções dadas. O padrão verificado ao longo do desenvolvimento desta construção VOL guarda estreita relação com àquele visto para a construção VL. Repare que no início a interlíngua de Andrea era limitada: seus enunciados não continham verbos, o uso de palavras soltas e linguagem telegráfica eram comuns, assim como o emprego de fórmulas fixas. Também havia muito uso de gestos. Andrea começou a usar a construção VOL inicialmente com o verbo *bring* (4 ocorrências). Por volta do décimo mês, o verbo prototípico *put* começa a ser muito usado em construções VOL. A partir



deste ponto, há um progresso perceptível na interlíngua de Andrea. Um padrão interlingual mais estruturado começa a surgir e observamos que construções mais complexas emergem, o mesmo padrão verificado no desenvolvimento da construção VL verificado para os demais aprendizes.

### 7.3.2 Lavinia

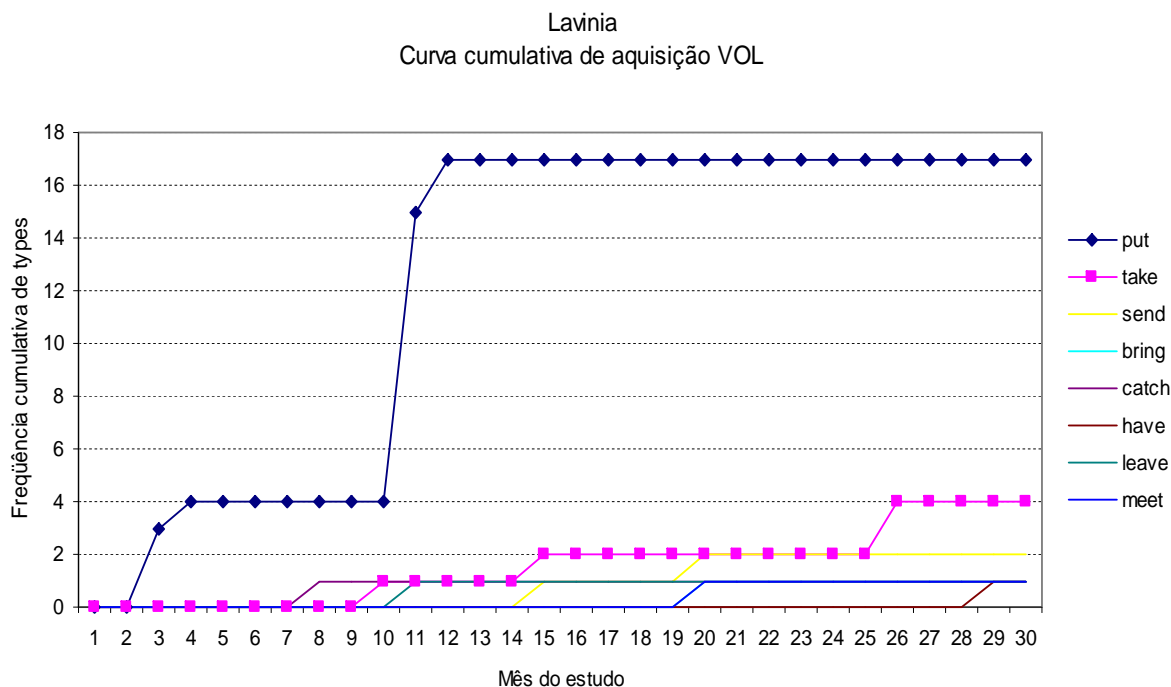
Em treze sessões de Lavinia há ocorrência da construção VOL. Além da tarefa *E*, as outras tarefas de Lavinia em que ocorre o uso de tal construção com outros verbos (incluindo *put*) são as tarefas *a*, *g*, *b*, *i*, *l*, *p* & *k*. Analisaremos algumas sessões ao longo dos meses com o objetivo de capturar o desenvolvimento de construções VOL. Veja no Quadro 12 o código do arquivo transcrito, a data de realização da sessão, mês do sujeito no estudo quando da sessão e verbos usados na construção com sua frequência (em *tokens*) entre parênteses

| Arquivo        | Data da sessão    | Mês no estudo | Verbos usados (frequência <i>tokens</i> ) |
|----------------|-------------------|---------------|---|
| liela13e.1.cha | 18 outubro 1983   | 3             | put (4 tokens)                            |
| liela14a.1.cha | 16 novembro 1983  | 4             | put (1 token)                             |
| liela18l.1.cha | 27 março 1984     | 8             | catch (1 token)                           |
| liela21p.1.cha | 10 maio 1984      | 10            | take (1 token)                            |
| liela22g.1.cha | 11 junho 1984     | 11            | put (2 tokens), leave (1 token)           |
| liela22e.1.cha | 11 junho 1984     | 11            | put (11 tokens)                           |
| liela22a.1.cha | 11 junho 1984     | 11            | put (1 token)                             |
| liela23b.1.cha | 13 julho 1984     | 12            | put (2 tokens)                            |
| liela24a.1.cha | 24 outubro 1984   | 15            | take (1 token), send (1 token)            |
| liela31a.1.cha | 11 maio 1985      | 22            | send (1 token)                            |
| liela31i.1.cha | 11 maio 1985      | 22            | bring (1 token), meet (1 token)           |
| liela25i.1.cha | 13 novembro 1985  | 28            | take (2 tokens)                           |
| liela36k.1.cha | 28 fevereiro 1986 | 31            | have (1 token)                            |

### Quadro 12

Sessões de Lavinia nas quais ocorre a construção VOL

O Gráfico 26 ilustra o desenvolvimento de Lavinia no uso de construções VOL, neste estágio de desenvolvimento ancoradas principalmente no verbo *put*, por meio de uma curva cumulativa de aquisição



**Gráfico 26**  
Curva cumulativa de aquisição VOL de Lavinia

***Sessão realizada em 18 de outubro de 1983, no terceiro mês do estudo***

Lavinia já inicia com um desenvolvimento lingüístico mais avançado do que aquele exibido por Andrea, acima discutido. De fato, Lavinia já está há 3 meses no projeto quando da realização desta sessão e possui um conhecimento mais profundo da língua quando comparada com os outros sujeitos nesta fase inicial do projeto (vide Anexo C, para uma descrição sociobiográfica dos sete sujeitos). Podemos claramente perceber que a construção VOL, ancorada sempre pelo verbo *put*, já faz parte da estruturação de sua interlíngua, i.e., a construção *X causes Y to move Zpath/location*. Também, como

observado nos exemplos de Andrea sobre a construção esquemática mais complexa do padrão *take X and put in Zpath/location* ou *X causes Y to move Zpath/location*, uma construção esquemática similar *take X and put in Zpath/location* parece já fazer parte da interlíngua de Lavinia. Veja exemplos abaixo:

- 1) *take [>2] the newspaper (...) you put the newspaper on the table*
- 2) *and n you take the [/?] the bag [/?] you take the bag. You put here*
- 3) *take m z that's that's know the name m the dish [/?] the plate for cigarettes. You er put in the bag # you put in your bag.*
- 4) *you close newspaper. Put in the bag.*

Lavinia faz uso do verbo *put* (de fato, de sua variante mais complexa *take X and put in Z*) em todas as quatro vezes em que a construção VOL foi usada nesta sessão.

***Sessão realizada oito meses depois, em 11 de junho de 1984***

Novamente, Lavinia faz uso da construção VOL utilizando o verbo *put*, juntamente com alguns exemplos da construção *take X and put in Z*, em todas as vezes que a construção VOL é usada. De forma um pouco diferente do que observamos em Andrea, observe que Lavinia faz uso deste tipo de construção como algo já congelado, fixo, como um *chunk* que se tornou uma fórmula bastante funcional e eficaz.

- 5) *## and now you put the bag <on the floor> [% laughs]*
- 6) *and ## put the bag ## in [% prolonged] like [/?] # *that's don't know the [/?] that's that's know ## there is some book - in the corner # near the cupboard**
- 7) *and [% prolonged] erm ## you have # to go # and put the purse in [% prolonged] [/?] ## near the the the [/?] the green books*
- 8) *be [>1] cause if he [/?] has to put a ladder on the wall that's that's know # and the woman say there is er a window in the kitchen*

- 9) *erm # oh no <> [% laughs] that's am right [/?] wrong no you have to put er the bag inside*
- 10) *mm #### its impossible because # the gas -\_ ## and the hot water that's put in separate that's put the money in the meter*
- 11) *put the bag on the cupboard # and [>1] there is some sheet -\_ # sheets under -\_ the sheets*

***Sessões realizadas entre 13 de julho de 1984 a 28 de fevereiro de 1986***

Lavínia estende o conhecimento da construção abastrata VOL para além do verbo prototípico *put*. Outros verbos começam a ser fortemente incorporados a esta construção, como os verbos *take*, *send*, *have*, etc. Veja exemplos

- 12) *[% laughs] yeah and they send [/?] we send them to some shops you know to collect the prize as well.*
- 13) *and when that's said to the [/?] because that's know the girl who [>1] works there and she was so disappointed as well because you know sh [/?] they send a lot of people there and then they change you know they say the the the course is full up already you know # xxx [>2] yeah [% high fall] because she phoned there as well.*
- 14) *oh you can take home them.*
- 15) *okay take me here the policeman er mm that's mean the policeman rung to the*
- 16) *[/?] to the centre that's that's know where [/?] the police station took him away you know # in police car xxx.*

A análise do desenvolvimento interlingual de Lavinia nas sessões acima é muito reveladora. Em primeiro lugar, há um padrão que se repete quando pensamos nas análises anteriores sobre a construção VL. No início, o uso da construção VOL é dominado pelo verbo *put*, mas Lavínia, por aparentemente já começar a dominar a estrutura sintática subjacente à construção VOL, consegue gradualmente cambiar outros verbos e palavras nos *slots* deste esquema. Nas sessões finais, há um processo visível de

categorização bem sucedida para esta construção, na qual outros verbos são usados para expressar sentidos comunicativos que se apresentam cada vez mais específicos, revelando, com isto, que Lavínia abstraiu um esquema sintático complexo, no qual outros exemplares lingüísticos passam a ser induzidos e usados por meio de uma co-variação coerente de atributos entre seus membros (como discutido na teorização sobre o desenvolvimento de categorias a partir de protótipos, feita no capítulo 3).

### 7.3.3 Jarnail

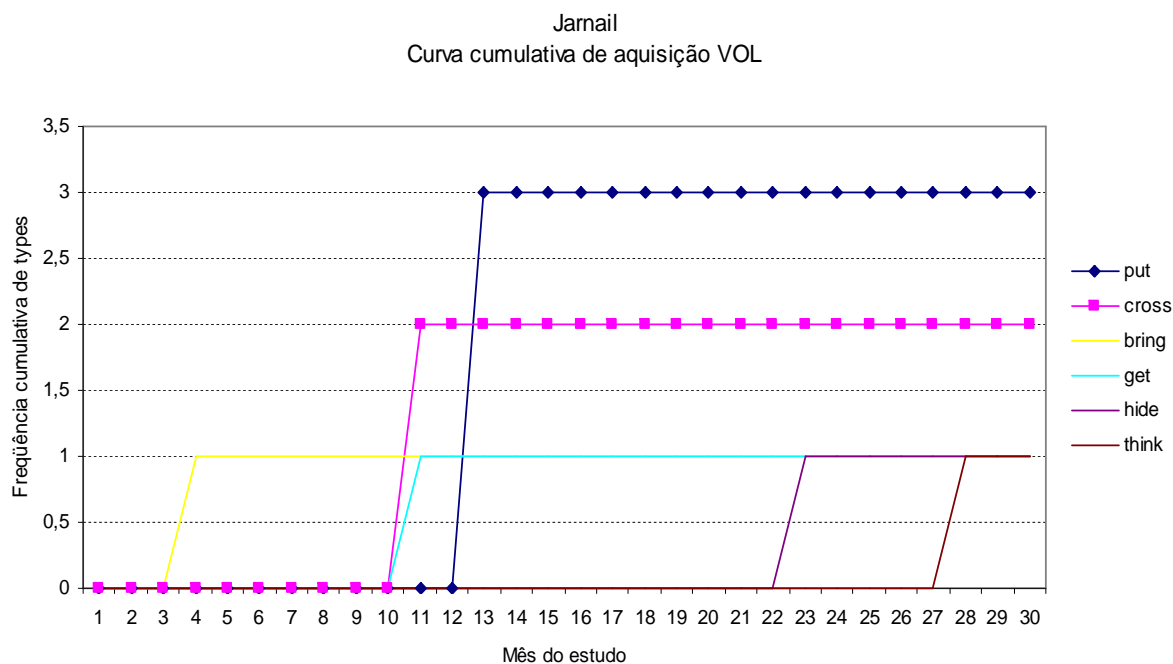
Há ocorrência da construção VOL em sete sessões de Jarnail. As construções VOL produzidas por Jarnail envolvem os verbos *put*, *cross*, *bring*, *get*, *hide* e *think* nas tarefas *e*, *a*, *l* & *r*. Analisaremos cinco sessões ao longo dos meses visando capturar o desenvolvimento de construções VOL. O Quadro 13 exhibe o código de cada arquivo transcrito, a data de realização da sessão, mês do sujeito no estudo quando da sessão e verbos usados na construção com sua frequência (em *tokens*) entre parênteses

| Arquivo        | Data da sessão    | Mês no estudo | Verbos usados (frequência <i>tokens</i> ) |
|----------------|-------------------|---------------|---|
| lpeja13e.1.cha | 12 fevereiro 1983 | 4             | bring (1 token)                           |
| lpeja17l.1.cha | 6 setembro 1983   | 11            | cross (2 tokens)                          |
| lpeja17a.1.cha | 6 setembro 1983   | 11            | get (1 token)                             |
| lpeja22e.1.cha | 11 novembro 1983  | 13            | put (3 tokens)                            |
| lpeja32e.1.cha | 30 novembro 1984  | 25            | put (2 tokens)                            |
| lpeja33e.1.cha | 30 novembro 1984  | 25            | hide (1)                                  |
| lpeja36r.1.cha | 24 abril 1985     | 30            | think (1)                                 |

### Quadro 13

Sessões de Jarnail nas quais ocorre a construção VOL

Por outro lado, o Gráfico 27 ilustra o desenvolvimento de Jarnail no uso de construções VOL, inicialmente ancoradas principalmente no verbo *put*, por meio de uma curva cumulativa de aquisição



**Gráfico 27**  
Curva cumulativa de aquisição VOL de Jarnail

*Sessão realizada em 12 de fevereiro de 1983*

Este é seu quarto mês no estudo e Jarnail basicamente se utiliza de gestos e algumas expressões de comando para conseguir comunicar suas intenções. Seu desempenho é semelhante ao de Andrea em sua primeira sessão. Sua produção na L2 nesta tarefa (i.e., tentativas de dar instruções) carece de verbos e se baseia na utilização de recursos paralingüísticos (por ex., gestos) juntamente com um substantivo e/ou locativo, como nos trechos exibidos abaixo:

- 1) *erm er < > [% pointing towards the bag] bag in the table.*
- 2) *bring bag here < > [% points towards the table top].*
- 3) *book in the table.*
- 4) *plate # in the book.*

5) *book in the bag.*

***Sessão realizada em 6 de setembro de 1983***

Jarnail encontra-se no seu décimo primeiro mês no estudo. Alguns verbos começam a ser usados, apesar de a interlíngua de Jarnail permanecer bastante telegráfica no início da sessão, baseando-se essencialmente em substantivos. Abaixo, uma única construção VOL identificada

6) *they confused me they did not tell me clearly where i could get it from*

***Sessão realizada em 11 de novembro de 1983***

Nesta sessão, no décimo terceiro mês do estudo, um novo padrão parece começar a surgir: o verbo *put* (também usado aqui na forma progressiva) é usado em todas as ocorrências da construção VOL. Jarnail também inverte a ordem canônica do inglês, possivelmente devido a uma transferência do padrão sintático do punjabi, como discutido anteriormente. De fato, este ‘sotaque sintático’ é verificado também na produção dos outros dois falantes de punjabi (Madan e Ravinder) ao longo das sessões. Os exemplos em 3 e 4 mostram a ordem LOV e VLO, respectivamente, para expressar o sentido comunicativo presente no ordenamento canônico VOL. Observe

- 1) *yes # now er # in the put in a /box blue box n s money.*
- 2) *an er n s money <putting> [>1] ## er s this is money er table*
- 3) *dust bin n s bag put in*
- 4) *yes # now er # in the put in a /box blue box n s money*

*Sessão realizada um ano depois, em 30 de novembro de 1984*

Jarnail usa o verbo *put* para expressar o sentido da construção VOL em apenas dois trechos ao longo de toda a sessão. Também usa o verbo *hide* numa construção tipicamente VOL. Sua proficiência no uso da L2 (assim como no uso da construção VOL para dar instruções) aparentemente não progrediu muito desde a última sessão, realizada há um ano. Jarnail parece apresentar um desenvolvimento muito lento de sua interlíngua, evidenciado pela sua dificuldade em dar instruções básicas.

1. no er put er in. <yeah bag> [<1]
2. okay ## er s # erns book # <putting> [>1] in er book. Er # elephant [% a little toy elephant] # elephant ### er # and s putting er book.
3. yes # erm ### [% longish pause] and the <hide # book inside> in the bag.

*Última sessão de Jarnail, realizada cinco meses depois, em 24 abril de 1985*

Para nossa surpresa, nesta última sessão Jarnail mostra-se bastante fluente no uso da língua inglesa, embora a única construção VOL identificada seja com o verbo *think*, bastante complexa por sinal, transcrita abaixo. Neste trecho, Jarnail responde a uma pergunta da entrevistadora sobre se ele se lembrava do inglês que era freqüentemente dirigido a ele durante as sessões de entrevistas do projeto (*did you remember the english you picked up from us?*)

1) *that's tried to think them over at home and later on try to use them in my english later on*

Outro bom exemplo do nível de estruturação de sua interlíngua pode ser visto nesta outra resposta dada sobre o que ele diria a uma outra pessoa a respeito do projeto do qual havia participado



2) *that's will say that it was a good project you not only did interviews but also helped people like me.*

Jarnail, assim como Andrea e Lavinia, apresenta o mesmo padrão de desenvolvimento interlíngua. Parte de gestos, fórmulas e expressões semi-fixas para o uso de estruturas mais complexas. Como Lavinia e Andrea, Jarnail parece abstrair, gradualmente, a estrutura abstrata VOL e, a partir daí, incorporar outros verbos buscando maior precisão comunicativa. Do genérico ao específico, mas inicialmente calcado em um protótipo altamente freqüente que é o verbo *put*. A utilização deste verbo prototípico, que aparece por volta do décimo terceiro mês, parece engendrar processos de generalização e categorização que aceleram sobremaneira o desenvolvimento interlíngua de Jarnail.

### 7.3.4 Ravinder

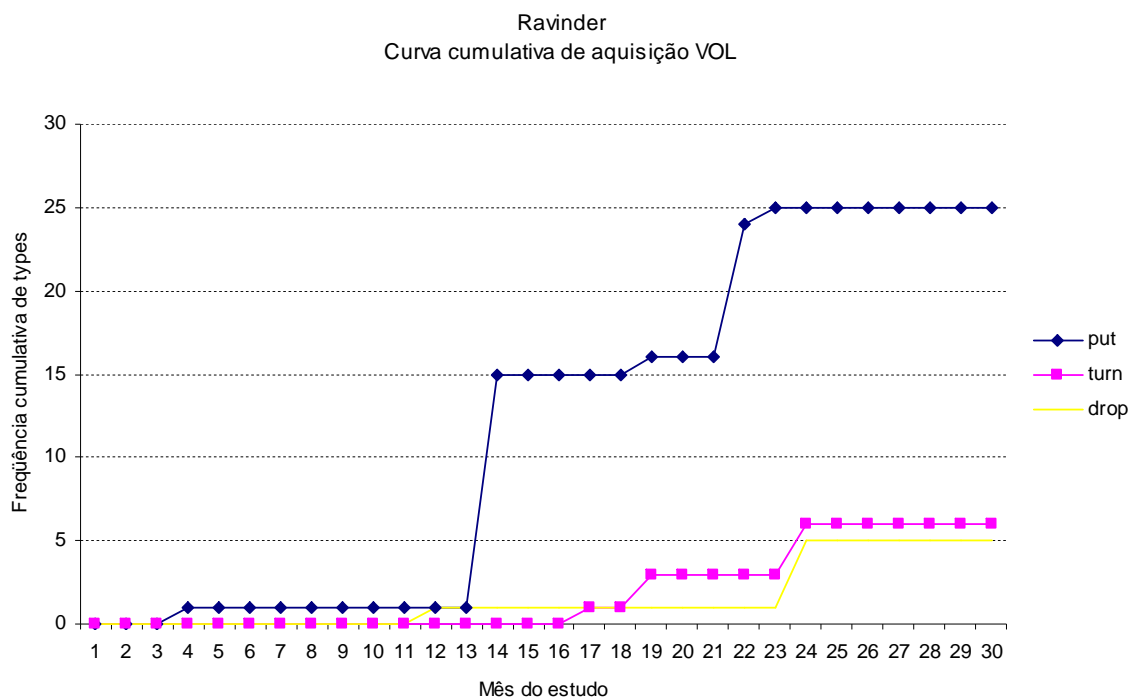
Foram detectadas dez sessões de Ravinder nas quais ocorre o uso da construção VOL, envolvendo as tarefas *e, o, m, a, l, p, i & q*. Destas, analisamos em detalhe cinco sessões nas quais a construção VOL é usada com o verbo *put*. Veja Quadro 14 com o código do arquivo transcrito, a data de realização da sessão, mês do sujeito no estudo quando da sessão e verbos usados na construção com sua freqüência (em *tokens*) entre parênteses

| Arquivo        | Data da sessão    | Mês no estudo | Verbos usados (freqüência <i>tokens</i> ) |
|----------------|-------------------|---------------|---|
| lpera13e.1.cha | 15 fevereiro 1983 | 4             | put (1 token)                             |
| lpera21o.1.cha | 19 outubro 1983   | 12            | drop (1 token)                            |
| lpera23m.1.cha | 8 dezembro 1983   | 14            | put (14 tokens)                           |
| lpera25a.1.cha | 29 março 1984     | 17            | turn (1 token)                            |
| lpera31p.1.cha | 24 julho 1984     | 21            | put (1 token)                             |
| lpera27l.1.cha | 10 julho 1984     | 21            | turn (2 tokens)                           |
| lpera32e.1.cha | 17 outubro 1984   | 24            | put (8 tokens)                            |
| lpera33q.1.cha | 5 novembro 1984   | 25            | put (1 token)                             |
| lpera34i.1.cha | 19 dezembro 1984  | 26            | drop (4 tokens)                           |
| lpera34l.1.cha | 19 dezembro 1984  | 26            | turn (3 tokens)                           |

#### Quadro 14

Sessões de Ravinder nas quais ocorre a construção VOL

O Gráfico 28 ilustra o desenvolvimento de Ravinder no uso de construções VOL, inicialmente ancoradas principalmente no verbo *put*, por meio de uma curva cumulativa de aquisição



**Gráfico 28**  
Curva cumulativa de aquisição VOL de Ravinder

*Sessão realizada em 15 de fevereiro de 1983*

Ravinder, a exemplo de Jarnail, faz muito uso de gestos e utiliza expressões multi-palavras para expressar suas intenções na L2. Ravinder está no quarto mês do estudo e seu desempenho neste estágio é semelhante ao de Jarnail e Andrea acima discutidos. As frases carecem de verbos e são compostas, essencialmente, de substantivos e locativos. Veja exemplos desta sessão, cuja tarefa consiste em dar instruções em relação à

mudança/deslocamento de objetos e pessoas num ambiente, envolvendo, portanto, referência espacial

- 1) *book <the bag in>*
- 2) *book # in the bag.*
- 3) *coat on*
- 4) *and bag the chair*
- 5) *book in the bag -\_ that's think.*

Também há a produção de alguns comandos, como estes:

- 6) *come on*
- 7) *m open the bag*

Uma única construção VOL, com o verbo *put*, é identificada.

- 8) *put one book yes ### on the read book you know # read*

### ***Sessão realizada em 19 de outubro de 1983***

Nesta sessão, Ravinder está no décimo segundo mês e já utiliza vários verbos e fórmulas em seus enunciados. Ele começa ir além de enunciados multi palavras, embora sua interlíngua seja bastante irregular, como observamos nos exemplos abaixo

- 12) *### i think must talk to this man.*
- 13) *theres two children.*
- 14) *in home # thats playing with children now.*
- 15) *sometime have cup of tea.*
- 16) *# this customer something <wanting> [?] yeah.*
- 17) *# <i dunno>*
- 18) *no <theres no time <this time of year> [?] >*
- 19) *i like it but no time.*

Nenhuma construção VOL foi identificada nesta sessão

*Sessão realizada em 8 de dezembro de 1983*

Ravinder encontra-se agora no décimo quarto mês do estudo. A tarefa consiste em dar instruções em relação à mudança/deslocamento de móveis em um ambiente, envolvendo, portanto, referência espacial. Explode o uso do verbo *put* nas construções VOL, veja exemplos

- 20) < > [*\$ laughs*] *put it in # door [/?] room.*
- 21) ## *here put it in # righthand side*
- 22) ### *mm* ## *put it in lefthand side #*
- 23) ### *put it in other room yes*
- 24) < > [*\$ laughs*] *put it in # door [/?] room*
- 25) <yeah> [*<1*] # *first # put in # blue table #* < > *the chair bring ### no*
- 26) *a horse # put in there* < > [*\$ laughs*]
- 27) *er* # *put in near telly*
- 28) *no* # *thing* ## < > *put it other side there*
- 29) *no* *put it in front # thats it # yeah*
- 30) *put in the right wall*
- 31) *put it in window*
- 32) *yeah* # *put in # lefthand side*
- 33) *yeah* *put it in # righthand side*

*Sessão realizada em 17 de outubro de 1984*

Dois meses depois da sessão discutida acima, a interlíngua de Ravinder parece ter progredido. Em seu vigésimo quarto mês no estudo, Ravinder está mais comunicativo, produz enunciados maiores e utiliza construções VOL ancoradas no verbo *put*. Seu

desenvolvimento assemelha-se aos outros seis sujeitos analisados neste subcorpus (falantes tanto de punjabi como de italiano). Alguns exemplos são fornecidos a seguir:

34) *open your bag # put on table as well*

35) *bag put in table.(OVL)*

36) *and open ## er put it out in book*

37) *and bag put in floor (OVL)*

38) *the ers smoke tray put it in book*

39) *and book put it in # bag*

40) *er bag put it in floor*

Ravinder também consegue produzir muito timidamente variações mais elaboradas da construção VOL, como a construção *take X and put in Z* (como também observado na produção dos outros sujeitos). Veja exemplo:

41) *bag that's it and pick it [/?] pick it up ## and put it in table*

#### ***Sessão realizada em 19 de dezembro de 1984***

Nesta sessão, em seu vigésimo sexto mês, Ravinder dá instruções a sua interlocutora sobre como chegar a lugares específicos (*route description*). As construções VOL identificadas utilizam o verbo *turn*

42) *<oh sorry # im turn left from mangat house>*

43) *and straight down # to ### turn left # from by sikhs temple # <its>[>1] bradford street as well*

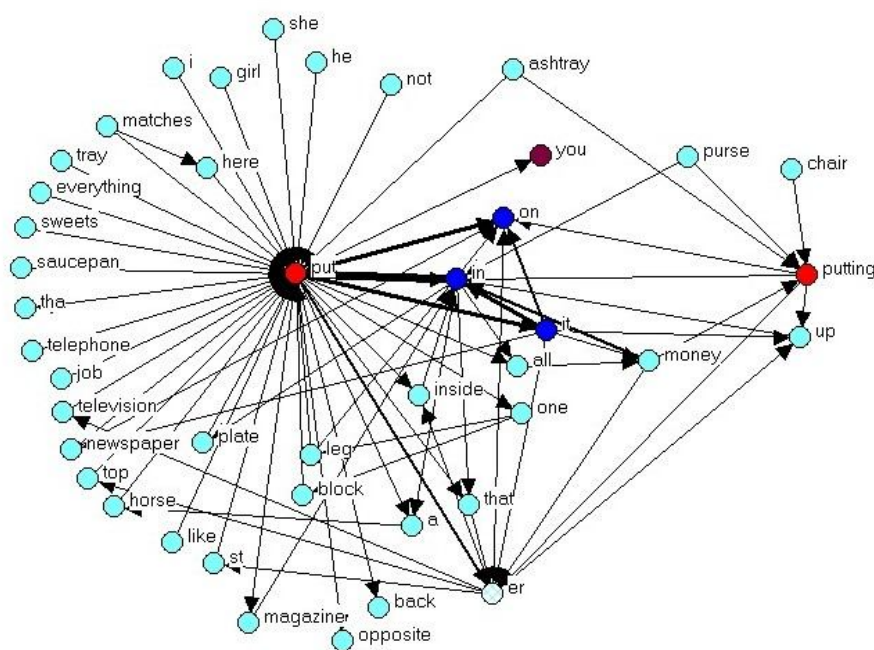
44) *turn right from there*

Apesar das observações sobre as últimas sessões acima descritas, Ravinder obteve um progresso bem mais lento do que Jarnail, Lavinia e Andrea ao longo do projeto. O desenvolvimento de sua interlíngua mostra-se bem mais 'conservador' e ele parece não

avançar para além de uma ‘variedade básica’. No entanto, o mesmo padrão de desenvolvimento observado nos demais aprendizes também é verdadeiro para Ravinder.

### 7.3.5 Discussão sobre o desenvolvimento da construção VOL

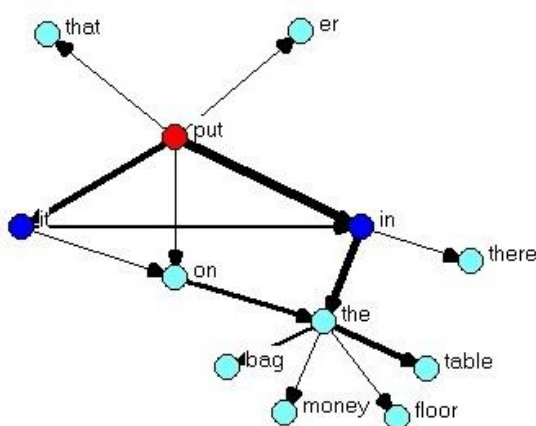
No Gráfico 29, a exemplo do que fizemos anteriormente na discussão sobre o verbo *go* na construção VL, vemos uma rede gereada para todas as ocorrências da construção VOL envolvendo o verbo prototípico *put*. Como explicado anteriormente, a espessura das linhas (conexões) entre as palavras indica o grau de frequência/ocorrência de uma dada colocação. Observe a configuração da rede



**Gráfico 29**

Rede de relações entre todas as palavras presentes nas 152 construções VOL com o verbo *put* (e ocorrências *putting*) produzidas pelos aprendizes

Na discussão das sessões no item anterior, é possível perceber nas construções dos quatro aprendizes o gradual acúmulo de palavras em torno do que poderíamos chamar de um ‘proto-esquema’ VOL, inicialmente calcado em um verbo prototípico. O Gráfico 30 refina ainda mais a rede dada acima, ao focar aquelas colocações mais frequentes ( $\geq 5$  de frequência de ocorrência) ilustrando a estrutura deste primeiro suposto esquema construcional VOL em processo de consolidação, abstraído em torno do verbo abre alas *put*



**Gráfico 30**

Rede de colocações mais frequentes na construção VOL, em torno do verbo *put*, produzidas pelo aprendizes

Uma outra característica bastante marcante observada no uso da construção VOL por parte dos aprendizes é a tendência de não nomear o objeto ou o referente em questão durante a comunicação (talvez pelo óbvio motivo de que ele já está dado, é explícito e partilhado pelos participantes do ato comunicativo).

Imagine, por exemplo, você tentando pedir (numa língua estrangeira que está começando a aprender) que sua interlocutora coloque a bolsa que ela carrega em cima de uma mesa. Provavelmente diríamos apenas ‘coloque na mesa’, ao mesmo tempo em que apontaríamos para a bolsa em questão. A menção ao referente tornar-se-ia desnecessária (e, por este mesmo motivo, o pronome objeto ‘*it*’ não seja **percebido** em

estágios iniciais de aprendizagem da língua inglesa). Tanto nossa L1 como as restrições impostas pela nossa própria interlíngua (ambos aspectos regidos por um cálculo probabilístico), o que, conseqüentemente gera limitações cognitivas que influenciam nossa atenção e percepção, fazem com que tais ‘itens’ ou ‘acessórios’ meramente lingüísticos (i.e., pronomes objeto, sufixos verbais e nominais, etc.) passem despercebidos por possuírem baixa **saliência**<sup>75</sup>. A ausência de tais marcadores gramaticais, como observado no uso inicial mais pragmático durante a aprendizagem de uma L2, é evidente nos dados aqui analisados.

Como já discutido, tal padrão (universal) refletiria uma ‘variedade básica’ observada no desenvolvimento de interlínguas e que, como destacam Perdue (1993), Klein (1993:104) e Mitchel & Myles (2004:150), é caracterizada por uma estrutura tópico-comentário, conforme a distinção de Givón dos usos pragmáticos e sintáticos que fazemos numa língua. As três fases observadas por Perdue (1993) e Klein (1993), quais sejam, a) variedades pré-básicas; b) variedades básicas; e c) para além da variedade básica (ou construções sintáticas abstratas), guardam estreita semelhança com as fases propostas por Tomasello (2003, 2000) ao longo da construção da L1 (conforme discutido anteriormente)<sup>76</sup>.

De fato, ao analisarmos qualitativamente a produção dos sete aprendizes do nosso corpus no projeto ESF, observamos um desenvolvimento de enunciados que parte do uso de substantivos concretos (um modo mais pragmático), passando pela utilização de verbos (um modo mais semântico) até uma estruturação calcada em estruturas ou esquemas abstratos (um modo mais sintático). Como vimos, este padrão também é verificado na análise do desenvolvimento da construção VL por parte dos aprendizes e corrobora a teorização discutida no capítulo 3 sobre o desenvolvimento de categorias calcado em protótipos.

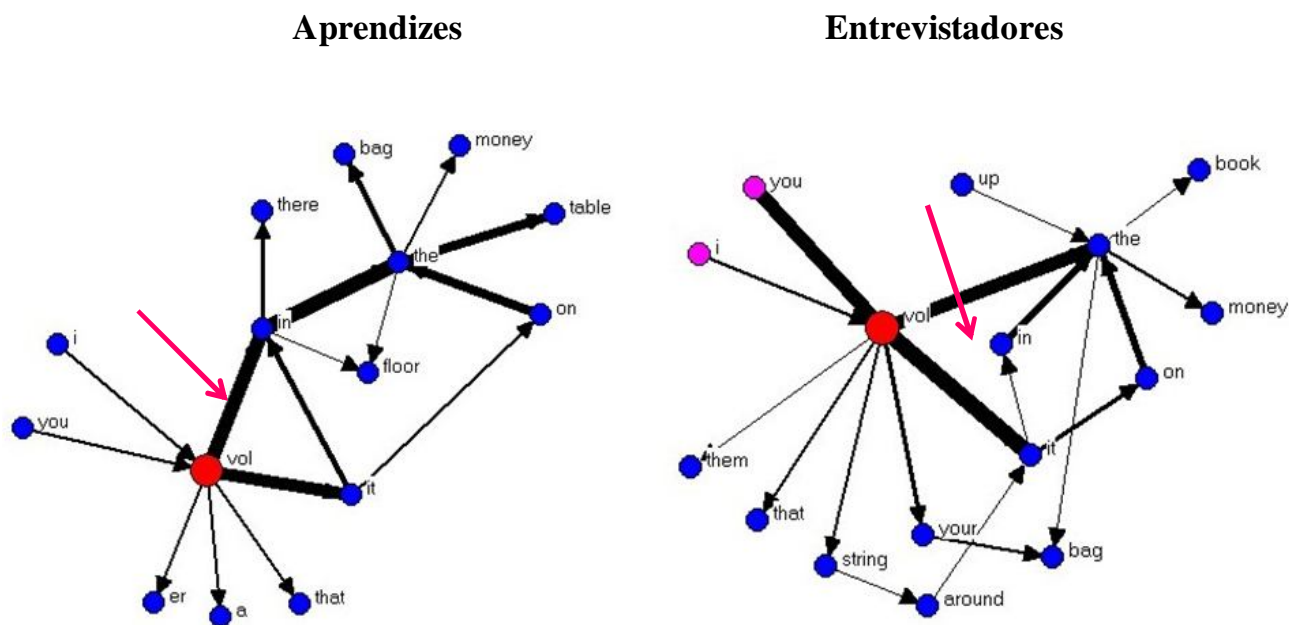
---

<sup>75</sup> Veja uma excelente discussão sobre vários aspectos e fenômenos observados no desenvolvimento de interlínguas em Nick Ellis (2006b, 2006c).

<sup>76</sup> Embora os processos cognitivos subjacentes a tal padrão, variedade básica ou interlíngua sejam universais, i.e., constituam um sistema independente da L1 e L2 envolvidas, isto não significa que o desenvolvimento interlingual não seja fortemente influenciado pelas matrizes sociolingüísticas da L1 (suas práticas sociais e normas culturais) e da L2. Tal ‘mistura’ é o que compõe as idiosincrasias e variações observadas no processo de aprendizagem de línguas, conforme nossa discussão anterior acerca do conceito selinkiano de interlíngua.



Finalmente, o não uso de referentes ou pronomes objetos por parte dos aprendizes (bem como seu uso por parte dos entrevistadores) é ilustrado a seguir no Gráfico 31, que reproduzimos a partir da análise dos nossos dados conduzida por Ke (2007). Observe as redes abaixo geradas a partir dos dados dos aprendizes e dos entrevistadores



**Gráfico 31**

Comparação em rede das colocações mais frequentes na construção VOL nos grupos de aprendizes e entrevistadores<sup>77</sup>. In: Ke (2007)

Como ilustrado no Gráfico 31, a configuração construcional VOL, que requer uma conexão entre verbo, objeto e locativo (no exemplo acima, as possibilidades são  $V + it + in\ the\dots$ , ou  $V + it + on\ the\dots$ ) encontra-se bem mais consolidada na rede dos entrevistadores do que na rede dos aprendizes, cuja principal rota é  $V + in\ the\dots$ . Mas uma segunda rota  $V + it\ in\dots$  já encontra-se em plena construção, como podemos

<sup>77</sup> A rede foi construída a partir de todas as palavras presentes nas construções VOL identificadas em nosso subcorpus, utilizando informações resultantes de um cálculo probabilístico de co-ocorrência de palavras (i.e., colocações). Foram adicionadas restrições à rede de forma a excluir os links estatisticamente não significantes. Para obter mais detalhes, veja Ke (2007).

observar pela linha menos espessa. É provável que a crescente complexificação da interlíngua, resultante do vivenciar do aprendiz na L2, resulte numa configuração semelhante àquela vista na rede dos entrevistadores acima, cujo padrão de conexão é V + artigo ou V + pronome objeto, e onde normalmente não ocorra o padrão V + preposição (em construções VOL, obviamente) em seu sistema lingüístico. Esta conexão não-canônica em construções VOL, presente na rede dos aprendizes, mas ausente na rede dos entrevistadores, é ilustrada no Gráfico 31 pelas setas em vermelho.

#### 7.4 Análise das construções VOO

Apenas 36 construções VOO foram identificadas em nosso subcorpus. Além disto, Jarnail não produziu nenhuma construção VOO ao longo de todo período do estudo. Outros, Santo e Vito, produziram apenas uma construção VOO. Resolvemos então incluir nesta análise os demais sujeitos (mesmo que seus dados do subcorpus ESF não estejam completos, como destacado anteriormente nesta tese). O Quadro 15 fornece o nome do aprendiz, o código do arquivo transcrito, a data de realização da sessão, mês do sujeito no estudo quando da sessão e verbos usados na construção com sua frequência (em *tokens*) entre parênteses

| Nome    | Arquivo        | Data da sessão | Mês | Verbos usados               |
|---------|----------------|----------------|-----|-----------------------------|
| Andrea  | liean271.1.cha | 10 Maio 1984   | 17  | explain (1)                 |
| Andrea  | liean32g.1.cha | 21 Nov 1984    | 23  | tell (1)                    |
| Lavinia | liela13n.1.cha | 18 Out 1983    | 3   | pay (1)                     |
| Lavinia | liela16a.1.cha | 20 Jan 1984    | 6   | give (2)                    |
| Lavinia | liela17a.1.cha | 29 Fev 1984    | 7   | write (1)                   |
| Lavinia | liela18a.1.cha | 27 Março 1984  | 8   | write (2)                   |
| Lavinia | liela21d.1.cha | 10 maio 1984   | 10  | give (1)                    |
| Lavinia | liela22a.1.cha | 11 Junho 1984  | 11  | give (1), ask (1), send (2) |
| Lavinia | liela22g.1.cha | 11 Junho 1984  | 11  | give (1)                    |
| Lavinia | liela23a.1.cha | 13 Julho 1984  | 12  | ask (2)                     |

|          |                |               |    |          |
|----------|----------------|---------------|----|----------|
| Lavinia  | liela23j.1.cha | 13 Julho 1984 | 12 | give (1) |
| Lavinia  | liela25a.1.cha | 13 Nov 1984   | 16 | give (2) |
| Madan    | lpema22a.1.cha | 10 Nov 1983   | 1  | buy (1)  |
| Madan    | lpema24h.1.cha | 24 Jan 1984   | 3  | buy (1)  |
| Madan    | lpema24i.1.cha | 24 Jan 1984   | 3  | give (4) |
| Madan    | lpema31d.1.cha | 27 abril 1984 | 6  | give (1) |
| Madan    | lpema32g.1.cha | 6 Nov 1984    | 13 | give (1) |
| Madan    | lpema34i.1.cha | 18 Dez 1984   | 14 | give (1) |
| Madan    | lpema35j.1.cha | 30 Jan 1985   | 15 | show (1) |
| Madan    | lpema36a.1.cha | 20 Fev 1985   | 16 | give (1) |
| Ravinder | lpera16j.1.cha | 8 Junho 1983  | 8  | give (1) |
| Ravinder | lpera17a.1.cha | 7 Set 1983    | 11 | give (1) |
| Ravinder | lpera27h.1.cha | 10 Jul 1983   | 21 | give (2) |
| Santo    | liesa16i.1.cha | 2 Fev 1984    | 8  | give (1) |
| Vito     | lievi19a.1.cha | 22 Ago 1983   | 14 | give (1) |

### Quadro 15

Todas as construções VOO identificadas em nosso subcorpus

Uma análise individualizada não se mostrou viável, pois não seria possível gerar gráficos em rede numa frequência  $\geq 5$ , padrão adotado nas nossas análises anteriores, devido ao número escasso de construções por indivíduo, como podemos ver no Quadro 15. Diante de tal impossibilidade, visualisaremos apenas os gráficos em rede com as colocações presentes no uso desta construção para os seis sujeitos investigados. Ou seja, uma análise coletiva envolvendo todos os verbos na construção VOO. Conseqüentemente, a análise não será conduzida nos mesmos moldes adotados para as construções VL e VOL em decorrência da escassez de construções VOO identificadas em nosso subcorpus (apenas 36 construções) para todos os sete sujeitos investigados. A análise será puramente baseada nas colocações observadas no entorno dos verbos na construção VOO. Mesmo com um baixo número de construções, resolvemos conduzir análise de colocações na tentativa de verificar se o padrão verificado nas duas construções anteriormente analisadas tende a se repetir aqui.

Uma observação importante. A exemplo de nossa análise do subcorpus ESF, o estudo de Goldberg, Casenhiser & Sethuraman (2004) também encontrou um baixo número de ocorrência da construção VOO no corpus de aprendizes de L1 investigado. De fato, esta construção parece ser aprendida de forma mais lenta do que as outras. Os autores explicam que isto se deve ao fato de que, aos 28 meses de vida (cerca de 2 anos de meio de idade), as crianças estão começando a aprender tal construção. Coincidentemente, os aprendizes no estudo ESF foram estudados ao longo de dois anos e meio (i.e., 30 meses), conforme discutido no capítulo 5 desta tese.

#### 7.4.1 Discussão

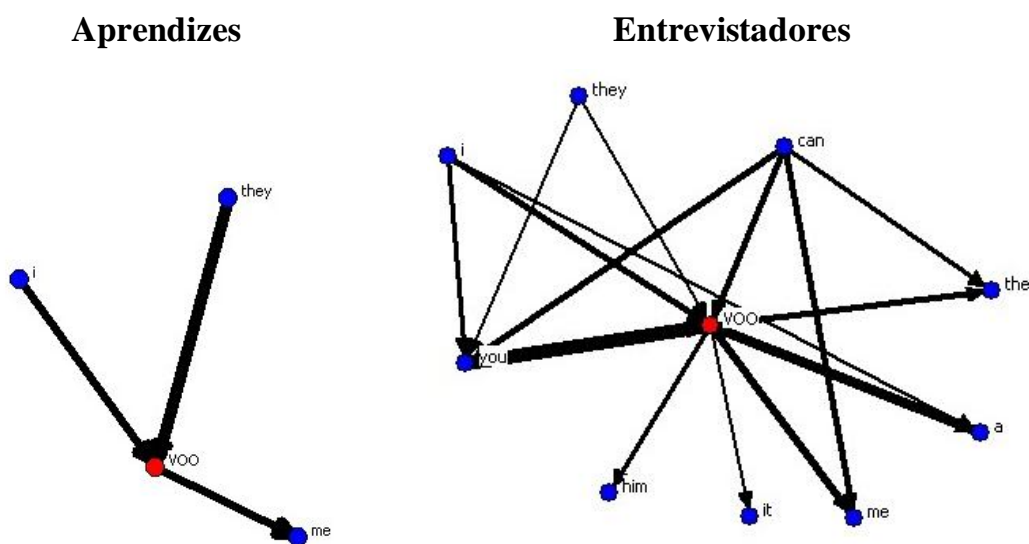
A construção VOO (ou bitransitiva) é normalmente constituída de um sujeito, um verbo e mais dois complementos, chamados de transitivo direto e transitivo indireto. Tal construção está intimamente ligada à função ou significado de transferir algo a alguém, ou seja, envolve um *agente*, um *recipiente* e um *tema*. Esta construção é conhecida na gramática tradicional como *caso dativo*, cujo termo etimologicamente deriva da palavra latina *dativus*, que significar dar. O caso dativo pode ser preposicionado ou não. Veja exemplos em português e inglês

- 1a) ela deu-me um livro (*she gave me a book*)
- 1b) ela deu um livro para mim (*she gave a book to me*)
- 2a) ela deu-lhe o troco (*she gave him the change*)
- 2b) ela deu o troco a ele (*she gave the change to him*)
- 3a) ela escreveu-me uma carta (*she wrote me a letter*)
- 3b) ela escreveu uma carta para mim (*she wrote a letter to me*)

No que se segue, listaremos todas as 36 construções VOO dos aprendizes (Quadro 16). No Gráfico 32, coluna à direita, são exibidas as redes criadas a partir de tais construções. Nestas redes são exibidas as associações entre as palavras das construções VOO (gráfico superior) e logo abaixo a rede de colocações VOO com restrição de frequência de ocorrência  $\geq 5$  (gráfico inferior).



Como podemos observar, a rede apresentada no Gráfico 32 à direita (inferior) revela a ocorrência de um **esquema verbal** muito freqüente no qual os pronomes sujeitos *I* e *they* juntamente com o pronome objeto *me* orbitam os 9 verbos que ocorrem nas 36 construções VOO produzidas pelos aprendizes. As colocações no entorno dos verbos na construção VOO dos aprendizes são comparadas à rede de colocações para a mesma construção obtida a partir das 139 construções dos entrevistadores, como revela o Gráfico 33.



**Gráfico 33**

Comparação das redes de colocações mais freqüentes na construção VOO dos aprendizes e entrevistadores. À esquerda, a rede menor dos aprendizes e, à direita, a rede mais ampla dos entrevistadores, ambas com  $\geq 5$  de freqüência de ocorrência.

Assim como na comparação das colocações entre as redes para os aprendizes e para os entrevistadores nas construções VL e VOL, vemos uma *ilha* mais rica no entorno dos 12 verbos usados na construção VOO pelos entrevistadores. O número de pronomes pessoais (*I*, *you* e *they*) e pronomes objeto (*you*, *him*, *it*, *me*) é maior na rede dos entrevistadores. Há ainda a forte presença dos artigos definido e indefinido *the* e *a*, respectivamente. O verbo modal *can* (que geralmente é usado juntamente com o verbo

principal para complementar ou alterar seu sentido) também aparece com bastante frequência na rede de colocações dos entrevistadores (27 ocorrências). Há apenas duas ocorrências de *can* nas construções VOO dos aprendizes.

Os retângulos em volta das colocações observadas nas construções VOO dos aprendizes (exibidas na lista do Quadro 16) mostram uma combinação de fórmulas (ou *chunks*) e esquemas para a construção VOO. Tal configuração parece revelar claramente os estágios iniciais e intermediários de uma seqüência que parece seguir o seguinte padrão: da fórmula para o esquema até construções mais abstratas e criativas. Veja o Quadro 17:

| Fórmula                            | Esquema                               | Construção Criativa          |
|------------------------------------|---------------------------------------|------------------------------|
| <b>They <u>give</u> me</b>         | uytr <u>give</u> gfd <b>the</b> dfdyr | <b>I drew her a heart</b>    |
| <b>I <u>give</u> you</b>           | kuyt <u>give</u> me a grters          | <b>he blew them kisses</b>   |
| <b>they <u>give</u> me</b>         | <b>they</b> jtyu <b>me the</b> jhhst  | <b>they threw me a party</b> |
| <b>shopkeeper <u>give</u> me</b>   | I yte fteret <b>the</b> tryurw        | <b>give peace a chance!</b>  |
| <b>they <u>give</u> me</b>         | <b>I send</b> gfhtr a dfdyr           | <b>I planted him an idea</b> |
| <b>they <u>give</u> you</b>        | <b>it</b> hhgut <b>you a</b> grters   | <b>she grew him a tree</b>   |
| <b>they <u>give</u> me a</b>       | gfr <u>buy</u> <b>her the</b> tryurw  | <b>she lights me a fire</b>  |
| <b>Charlie Chaplin <u>give</u></b> | <b>they</b> kitrq gfd a dfdyr         |                              |
| <b>I <u>give</u> you</b>           |                                       |                              |

### Quadro 17

Ilustração do processo de aquisição (abstração gradativa) da construção VOO a partir de fórmulas calcadas em co-ocorrência de palavras em torno do verbo prototípico *give*

Os dados dos aprendizes compreendem um período de aproximadamente 30 meses e parecem revelar, justamente, um momento ou período de transição do uso de fórmulas para o uso de esquemas. Infelizmente, caso houvesse um número maior de construções VOO, poderíamos ter uma análise mais robusta e mais próxima daquelas realizadas para as construções VL e VOL. A seguir, faço alguns comentários finais sobre as análises conduzidas neste capítulo.

## 7.5 De ilhas verbais a um arquipélago construcional

Como vimos ao longo deste capítulo, ao fazermos uma análise das sessões dos quatro sujeitos investigados (Andrea, Lavinia, Jarnail e Ravinder) com o intuito de verificar o desenvolvimento das construções VL, VOL e VOO a partir de seus respectivos verbos prototípicos (i.e., ilhas verbais), foi possível perceber que o padrão de desenvolvimento construcional observado, embora seja gradual, alicerça-se em algumas outras colocações altamente frequentes, que podemos interpretar como um **arquipélago colocacional ou construcional emergente** (i.e., ilhas verbais, ilhas pronominais, ilhas preposicionais, etc.)<sup>78</sup>. Tais ilhas e seus habitantes possibilitam um desenvolvimento que parte de fórmulas, progride para esquemas, e, posteriormente, alcança o nível de construções abstratas e criativas (como exemplificado na discussão sobre a construção VOO).

Como destacado, o registro das sessões dos sujeitos de nosso subcorpus ESF reflete um período de aproximadamente 30 meses de progresso na língua inglesa. O progresso ao longo deste período geralmente equivale a um estágio intermediário de proficiência na L2. De fato, se observarmos os gráficos de aquisição cumulativa das construções VL e VOL para os quatro sujeitos analisados (ou seja, Andrea, Lavinia, Jarnail e Ravinder) e observarmos atentamente o desenvolvimento interlingual dos quatro sujeitos constataremos, de forma clara, que, na média, os aprendizes atingiram este nível de proficiência (à exceção de Lavinia, que foi visivelmente além deste nível). Tal progresso, conforme verificado, ocorreu de forma altamente irregular e idiossincrática. Há diferenças evidentes no progresso exibido por cada um dos quatro aprendizes investigados. Dentre estas, se destaca o ritmo de aquisição das estruturas por cada um dos aprendizes. Cada aprendiz tem seu tempo. Entretanto, há também fortes semelhanças no tocante ao padrão de desenvolvimento observado nos quatro sujeitos, como buscamos mostrar ao longo de nossa análise.

De forma resumida, é claramente possível identificar tais padrões recorrentes no desenvolvimento interlingual dos sujeitos. São eles: 1) de maneira geral, nas primeiras

---

<sup>78</sup> A idéia de um arquipélago construcional emergente, resultante da análise dos dados do subcorpus ESF, é desenvolvida com mais propriedade em Ellis & Ferreira-Junior (no prelo b).



sessões do projeto, verbos são raramente usados; os enunciados normalmente não contêm verbos; os aprendizes fazem muito uso de gestos; o uso de palavras soltas e linguagem telegráfica são comuns; igualmente comum é o emprego de fórmulas fixas; 2) a comunicação inicial essencialmente telegráfica e calcada apenas em substantivos começa a incorporar alguns verbos. Expressões multi-palavras começam a surgir. Há bastante uso de fórmulas fixas e semi-fixas, mas alguns verbos leves começam a ser usados com frequência; 3) verbos prototípicos começam a aparecer e são usados nas suas respectivas construções; 4) Há uma fase de muito uso do verbo prototípico e um aumento considerável da capacidade de expressão de significados (i.e., comunicação) na L2; há um aumento considerável no uso das construções com seus respectivos verbos prototípicos; 5) Este verbo prototípico bastante usado em certas funções comunicativas propicia, gradualmente, a aprendizagem das construções VL, VOL e VOO; 6) com o passar do tempo, ou seja, à medida que o aprendiz é exposto a mais e mais insumo, tais verbos abre-alas também promovem a incorporação e uso graduais de outros verbos semanticamente próximos ao exemplar prototípico. É possível observar, então, o surgimento de estruturas mais complexas e criativas no uso da L2. A Figura 10 busca ilustrar tal seqüência de aquisição para a construção VOL (embora se aplique igualmente, como vimos, às demais construções analisadas nesta tese)

|                         |   |          |
|-------------------------|---|----------|
| 1. Put it on the table  | Put it on the table   |          |
| 2. Put it on the table  | Put it on the table   |          |
| 3. Put it on the table  | Put it on the table   |          |
| 4. Put it on the desk   | Put it on the <del>table</del>                                  |          |
| 5. Put it in the bag    | Put it <del>on</del> the <del>table</del>                       | Put it L |
| 6. Put it on the bed    | Put it <del>on</del> the <del>table</del>                       |          |
| 7. Put it in the trash  | Put it <del>on</del> the <del>table</del>                       |          |
| 8. Put it in the fridge | Put it <del>on</del> the <del>table</del>                       | V it L   |
| 9. Set it on the table  | <del>Put</del> it <del>on</del> the <del>table</del>            | V O L    |
| 10. Run it at the road  | <del>Put</del> it <del>on</del> the <del>table</del>            |          |
| 11. Get me on the way   | <del>Put</del> <del>it</del> <del>on</del> the <del>table</del> |          |
| 12. Let me at the ball  | <del>Put</del> <del>it</del> <del>on</del> the <del>table</del> |          |

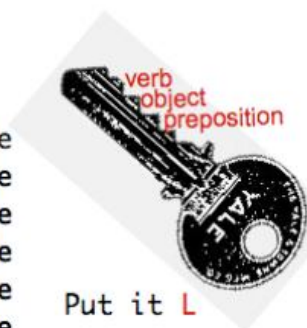


Figura 10

Ilustração da seqüência de aquisição da construção VOL: das fórmulas fixas aos esquemas às construções abstratas. In: Ellis & Ferreira-Junior (no prelo b).

Eis a chave! Como vemos na Figura 10, a experiência cumulativa com os exemplares ou fórmulas fixas VOL leva, inicialmente, ao entrincheiramento destas (um *chunk*). Um protótipo de alta frequência, neste exemplo o verbo *put*, ancora a construção como uma expressão formuláica. A experiência com outras construções VOL e com os protótipos de outras ilhas do arquipélago construcional ou colocacional (protótipos verbais, pronominais, preposicionais, etc.) propicia a generalização do esquema, com os diferentes *slots* tornando-se progressivamente definidos como atratores.

Os comentários feitos acima e resumidos nos itens 1-6 (ilustrados na Figura 10) sobre as características observadas no desenvolvimento das construções por parte dos aprendizes ao longo das sessões sintetizam, de certa forma, as análises e discussões feitas ao longo deste capítulo.

No próximo e derradeiro capítulo desta tese teço algumas considerações finais sobre o tema aqui tratado, sobre os resultados empíricos encontrados e sua importância para o aprofundamento da teorização sobre os processos envolvidos na aquisição de línguas e a possível contribuição prática destes resultados para o ensino e aprendizagem de línguas estrangeiras.

## Considerações finais

*I won't write my poem till I'm in my right mind*

Allen Ginsberg

Retomando e tentando sintetizar toda a discussão teórica<sup>79</sup> conduzida nos capítulos 1, 2, 3 e 4 desta tese sobre os supostos processos envolvidos na aprendizagem de uma língua (ao longo dos seus vários níveis, i.e., neurobiológico, simbólico-representacional, comportamental, etc.) é possível sustentar que o padrão de desenvolvimento interlingual dos aprendizes verificado em nosso subcorpus parece sustentar o arcaboço teórico delineado neste trabalho (i.e., a aprendizagem de uma L2 parece se dá, principalmente, por meio do desenvolvimento de construções emergentes, resultantes, por sua vez, de uma atividade sistêmica contínua guiada por processos probabilísticos indutivos de busca de padrões e categorização por meio de um mecanismo neurobiológico e cognitivo de *chunking*).

Tais processos neurobiológicos e cognitivos se traduzem em certos princípios organizacionais presentes na interlíngua dos aprendizes, como vimos nas análises dos capítulos 5, 6 e 7, que são expressos, neste nível comportamental, por meio da utilização de recursos pragmáticos, passando pelo emprego de expressões lexicalizadas para, finalmente, atingir o nível complexo das construções sintaticamente estruturadas. Tais princípios são destacados tanto por Perdue (1993a, 1993b), com base nos resultados de outros tipos de análises conduzidas no âmbito do projeto ESF, como em teorizações recentes de Tomasello (2000, 2003), Goldberg & Casenhiser (2008), Goldberg (2006),

---

<sup>79</sup> Discussão teórica que admito ser, de fato, “excruciantemente longa”, mas, no entanto, teve que permanecer como tal devido à abrangência e complexidade do tema aqui tratado, sob o risco, a meu ver, de se tornar algo “excruciantemente” limitado em termos intelectuais. Finalmente, admito e assumo o risco que corro ao realizar um trabalho de tal amplitude teórica, que o torna, paradoxalmente, altamente vulnerável.

Goldberg, Casenhiser & Sethuraman (2004) e Nick Ellis (2001, 2002a, 2002b, 2003 e 2006c).

As análises conduzidas neste manuscrito são limitadas e por isto não permitem amplas generalizações sobre os processos mais gerais envolvidos na aquisição de uma L2, mas seus resultados são importantes e corroboram outros estudos sobre ASL. As análises feitas no capítulo 7 desta tese (baseadas no desenvolvimento das construções VL, VOL e VOO) apontam para o gradual desenvolvimento de esquemas construcionais por parte dos aprendizes da L2 ao longo de um período de aproximadamente trinta meses de coleta de dados. O desenvolvimento ou a constante re-estruturação (re-construção) de interlínguas, como é possível observar na análise da produção dos sujeitos, apesar de idiossincrático e gradual, apresenta sistematicidades inquestionáveis, mesmo levando em conta os diferentes graus de proficiência exibidos pelos aprendizes Andrea, Lavinia, Jarnail e Ravinder.

Como destacado por Ortega & Iberri-Shea (2005), há uma escassez de estudos longitudinais na área de ASL, sendo a pesquisa que é empreendida nesta área majoritariamente composta por estudos transversais (*cross-sectional*). Embora os estudos transversais tenham a sua importância por fornecer insights valiosos sobre os processos envolvidos na aquisição de uma L2 (como atesta a contribuição de GRIES & WULFF, 2005), uma compreensão total dos processos cognitivos subjacentes à aquisição de uma língua depende, inexoravelmente, do fator *tempo*. Como observado tanto na discussão teórica sobre construções como nas análises dos quatro sujeitos no capítulo 7, o tempo é um fator crucial na construção de uma língua, ou seja, desde os *tokens* altamente frequentes, passando pelos *types* até a construção de estruturas abstratas. Construir uma língua exige tempo!<sup>80</sup>

Diante dos motivos expostos por Ortega & Iberri-Shea (2005), acredito, pois, que a contribuição maior deste trabalho se relaciona à natureza dos dados empíricos aqui analisados (i.e., uma análise longitudinal de um corpus de aprendizes de L2 ao longo de

---

<sup>80</sup> E após o lento, gradual e constante desenvolvimento de construções sobre construções, de abstrações de abstrações, é possível chegar à poesia. Ou melhor, alguns, a exemplo de Allen Ginsberg, chegam brilhantemente lá. E, pela epígrafe que fecha esta tese, extraída de *America*, não parece ser assim fácil!

um período de aproximadamente dois anos e meio). De fato, no âmbito dos estudos em ASL a partir da perspectiva construcionista não há, até onde chega meu conhecimento, uma análise de dados nos moldes feitos aqui. Finalmente, os resultados parecem confirmar, no âmbito dos estudos em ASL, a validade da teorização construcionista e os achados pioneiros relatados em Goldberg & Casenhiser (2008), Goldberg (2006), Goldberg, Casenhiser & Sethuraman (2004) no âmbito de uma L1.

Gries & Wulff (2005:196), ao discutirem a importância e impacto do enfoque construcionista para a área de ASL, destacam que

Embora os aprendizes de línguas estrangeiras estejam muito menos expostos ao insumo na L2 do que os falantes nativos na sua L1, eles ainda assim são capazes de produzir generalizações consoantes com explicações construcionistas. Portanto, argumentos contrários à validade das construções baseados no fato de o insumo ser limitado são aparentemente enganosos. Além disso, apesar das várias diferenças existentes entre a aprendizagem de uma L1 e uma L2, a natureza probabilística dos resultados e as semelhanças verificadas no tocante a falantes nativos fornecem fortes evidências acerca da plausibilidade das teorias *exemplar-based* na investigação do processo de aprendizagem de uma segunda língua/língua estrangeira, no qual a frequência de exposição e o uso de construções desempenham um papel decisivo<sup>81</sup>

Os *insights* fornecidos pelos resultados da pesquisa conduzida nesta tese podem contribuir para a área de ensino e aprendizagem de línguas estrangeiras. Tais resultados corroboram os argumentos de Gries & Wulff acima. Aplicá-los ou transformá-los instantaneamente em um método de ensino de línguas (como normalmente esperam os que acompanham de longe as teorizações sobre ASL), no entanto, é algo que exige reflexão igualmente longa. Neste sentido, sugiro a seguir alguns pontos que podem vir a

---

<sup>81</sup> Minha tradução de: *Although foreign language learners have much less input in the foreign language than native speakers have in their native language, they are still able to arrive at generalizations that lend themselves to construction-based explanations. Therefore, accounts arguing against constructions on grounds of limited input are apparently on the wrong track. In addition, in spite of the various differences between first and second/foreign language learning, the probabilistic nature of the results and their similarity to that obtained for native speakers provide strong additional support of exemplar based theories of second/foreign language acquisition in which frequency of exposure to, and use of, constructions play a vital rule (sic) (p.196).*

contribuir para a maneira de se ensinar e aprender uma língua (seja na sala-de-aula ou fora dela).

Primeiro, acredito que seja necessária uma compreensão mais concreta por parte do professor de línguas estrangeiras de como uma língua se estrutura, ou seja, uma compreensão de estrutura lingüística estreitamente relacionada às nossas funções cognitivas básicas e à forma como nos relacionamos com o mundo que nos cerca (como indicam, principalmente, as pesquisas no âmbito da Lingüística Cognitiva). Por exemplo, para Tyler (2008), fazer com que professores de línguas não familiarizados incorporem a visão de língua e formas lingüísticas como uma expressão pura de nossos processos cognitivos básicos é a principal tarefa a ser feita. Algumas tentativas de aplicação pedagógica desta nova forma de conceber uma língua são feitas em Evans & Tyler (2005), que sugerem, por exemplo, novas alternativas de explicar o significado e motivações de uso de certas preposições da língua inglesa; isto é feito por meio da tomada de consciência, por parte do aprendiz, da relação estreita entre as preposições e seus possíveis significados em relações socio-físico-espaciais.

Tyler (2008:460) explica que na simples compreensão da preposição *up* na frase *the price of gas is up* há um conhecimento maior que nos informa (inconscientemente), com base em nossa longa experiência com eventos do mundo, que quando há um aumento de quantidade há também uma elevação ou um aumento vertical de algo, quando, por exemplo, abrimos a geladeira, pegamos uma garrafa de água e enchemos gradualmente um copo na ânsia de bebê-la. Ou seja, desde criança observamos estas relações, que, por sua vez, estão refletidas nas formas lingüísticas que usamos no dia-a-dia.

Um segundo *insight* importante é a idéia de protótipos. Eles podem otimizar ou facilitar o ensino de certas categorias lingüísticas, como vemos nas três construções aqui analisadas. Como destacam Ellis & Ferreira-Junior (no prelo a), a melhor maneira de ensinar um conceito é dando um exemplo. Assim, a melhor maneira de ensinar uma categoria é fornecendo o melhor exemplo (i.e., seu protótipo). Goldberg & Casenhiser (2008) demonstram que alguns estudos indicam que, de fato, é vantajoso fornecer uma grande quantidade de exemplos prototípicos em estágios iniciais de aprendizagem de uma L2 (embora possa ser enfadonho, admitem). Portanto, tais protótipos, como vimos

nesta tese, servirão de porta de entrada para as categorias abstratas da língua, pois encerram o significado ou função de uma dada construção que esteja sendo objeto de ensino.

Um terceiro *insight* valioso pode ser a idéia de ensinar fórmulas ou fraseologismos nos estágios iniciais de aprendizagem. Como vimos em nossa análise do subcorpus ESF, os aprendizes fazem muito uso de fórmulas. No ensino de línguas estrangeiras é grande o potencial de aplicação de fórmulas no planejamento e elaboração de materiais didáticos que levem em conta tais realidades cognitivas e lingüísticas.

Muitas construções encontradas em nossos livros didáticos ainda parecem estranhas quando comparadas ao que ouvimos, de fato, em exemplos naturais (por exemplo, programas de rádio ou de TV, ou o que ouvimos em conversas com falantes nativos de uma LE). Ou seja, podemos elaborar materiais que enfatizem certos encadeamentos de palavras e estruturas lexicais levando em conta seus graus de frequência de ocorrência (i.e., colocações). O ensino destes *chunks* (ou pedaços) de língua altamente funcionais é de grande utilidade comunicativa, e muito importante nos estágios iniciais.

Para Ellis (2008), em estágios iniciais os *chunks* ou *fórmulas fixas* são aprendidos e usados de forma explícita. São, portanto, para os aprendizes iniciantes, uma maneira eficaz de comunicação, algo como o juntar de blocos em um jogo de Lego. Mas para aprendizes em níveis mais avançados da L2, por outro lado, a fraseologia torna-se um problema. Idiomaticidade implica encadeamentos socioculturalmente entrincheirados numa dada língua, e tais encadeamentos no nível da fraseologia (*chunks* bem maiores e mais complexos) são aprendidos apenas implicitamente. Portanto, é preciso muita exposição à L2.

No ensino de ESP, acredito que tais contribuições possam também ser muito valiosas no sentido de atentarmos para a importância da fraseologia, aspecto que tende a ser inconscientemente ignorado em tal metodologia devido ao enfoque quase que exclusivo em questões terminológicas (i.e., palavras ou termos específicos).

Por fim, espero que os achados empíricos aqui relatados e a discussão feita ao longo destas páginas sejam de alguma utilidade e possam contribuir, de certa forma, para o avanço e o aprofundamento teóricos na área de aquisição de segunda língua no contexto acadêmico brasileiro.



## Referências bibliográficas

BARTLETT, Frederic C. *Remembering - A Study in Experimental and Social Psychology*. Cambridge: Cambridge University Press, 1932.

BIALYSTOK, Ellen. Towards an explanation of second language acquisition. In: Gillian Brown et. al.(Eds.). *Language and Understanding*. Oxford: Oxford University Press, 1994.

BOD, Rens, HAY, Jennifer, & JANNEDY, Stephanie. Introduction. In: Rens Bod, Jennifer Hay & Stefanie Jannedy. *Probabilistic Linguistics*. Cambridge, MA: Bradford/MIT Press, p.01-10. 2003.

BROEDER, Peter & PLUNKETT, Kim. Connectionism and second language acquisition. In: N. Ellis (Ed.). *Implicit and Explicit Learning of Languages*. London: Academic Press, 1994.

BIALYSTOK, Ellen. Towards an explanation of second language acquisition. In: Gillian Brown et. al. (Eds.). *Language and Understanding*. Oxford: Oxford University Press. 1994.

BYBEE, Joan. Usage-based grammar and second language acquisition. In: Peter Robinson & Nick C. Ellis (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition*. Hillsdale, NJ: Lawrence Erlbaum, p. 216-236. 2008.

BYBEE, Joan & McCLELLAND, James. Alternatives to the combinatorial paradigm of linguistic theory based on domain general principles of human cognition. *The Linguistic Review*, Vol. 22, 2005. p. 381-410.

CHATER, Nick & MANNING, Christopher D. Probabilistic models of language processing and acquisition. *TRENDS in Cognitive Sciences*, Vol. 10, 7, 2006. p. 335-344.

CHOMSKY, Noam. On Cognitive Structures and their Development: A Reply to Piaget. In: Massimo Piattelli-Palmarini (org.) *Language and Learning (The debate between Jean Piaget and Noam Chomsky)*: London: Routledge & Kegan Paul, p.35-54. 1980

CHOMSKY, Noam. Language: Chomsky's Theory (verbete). In: Richard L. Gregory (org.) *The Oxford Companion to the Mind*: Oxford: OUP, p.419-21. 1998

CHRISTIANSEN, Morten & CHATER, Nick. Connectionist Psycholinguistics: The Very Idea. In: Morten Christiansen & Nick Chater. *Connectionist Psycholinguistics*. Ablex: Connecticut, p.01-15. 2001.

CHRISTIANSEN, Morten & CHATER, Nick. Connectionist Psycholinguistics in Perspective. In: Morten Christiansen & Nick Chater. *Connectionist Psycholinguistics*. Ablex: Connecticut, p.19-75. 2001.

CLARK, Andy. Parallel Distributed Processing. In: *Microcognition: Philosophy, Cognitive Science, and Parallel Distributed Processing*. Cambridge, MA: MIT Press, 1989. p. 83-105.

CLARK, Andy. *Associative Engines: Connectionism, Concepts and Representational Change*. Cambridge, MA: MIT Press, 1993.

CORDER, Pit. The significance of learners' errors. *International Review of Applied Linguistics*, vol. 5, 1967. p. 161-169.

CORDER, Pit. Linguistics and the language teaching syllabus. In: J.P.B Allen and P. Corder (eds). *The Edinburg Course in Applied Linguistics*. London: Oxford University Press, p. 275-284. 1973.

CROWELL, Sheila C. The neurobiology of declarative knowledge. In: John Schumann et al. *The neurobiology of learning: Perspectives from second language acquisition*. New Jersey: Lawrence Erlbaum Associates, p.75-109. 2004.

CRYSTAL, David. *The Cambridge Encyclopedia of Language (Second Edition)*. Cambridge: CUP, 2003.

*Dicionário Houaiss da Língua Portuguesa*. Rio de Janeiro: Editora Objetiva, 2001.

ELLIS, Nick C. Implicit and Explicit Language Learning - An Overview. In: N. Ellis (Ed.). *Implicit and Explicit Learning of Languages*. London: Academic Press, 1994.

ELLIS, Nick C. Emergentism, Connectionism and Language Learning. *Language Learning*, vol 48 (4), 1998. p. 631-664.

ELLIS, Nick C. Memory for language. In: Robinson, P. (Ed.). *Cognition and Second Language Instruction*. Cambridge: Cambridge University Press, p. 33-68. 2001.

ELLIS, Nick C. Frequency effects in language acquisition: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, vol. 24, 2002a, p.143-188.

ELLIS, Nick C. Reflections on frequency effects in language processing. *Studies in Second Language Acquisition*, vol. 24, 2002b. p.297-339.

ELLIS, Nick C. Constructions, Chunking, and Connectionism: The Emergence of Second Language Structure. In: C. Doughty & M. H. Long (Eds.), *The Handbook of Second Language Acquisition*. Oxford, UK: Blackwell, p. 63-103. 2003.

ELLIS, Nick C. The processes of second language acquisition. In: Van Patten et. al. (Eds.), *Form-Meaning Connections in Second Language Acquisition*. New Jersey: LEA, p. 49-76. 2004.

ELLIS, Nick C. At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, 27, 2005, Pp. 305-352.

ELLIS, Nick C. Cognitive perspectives on SLA: The Associative Cognitive CREED. *AILA Review*, vol. 19, 2006a. p. 100-121.

ELLIS, Nick C. Language Acquisition as Rational Contingency Learning. *Applied Linguistics*, vol. 27(1), 2006b. p. 1-24.

ELLIS, Nick C. Selective Attention and Transfer Phenomena in L2 Acquisition: Contingency, Cue Competition, Salience, Interference, Overshadowing, Blocking, and Perceptual Learning. *Applied Linguistics*, vol. 27(2), 2006c. p. 164-194.

ELLIS, Nick C. Phraseology: The Periphery and the Heart of Language. Prefácio em F. Meunier and S. Granger (Eds.), *Phraseology in Foreign Language Learning and Teaching*. Amsterdam: John Benjamins, p. 1-13. 2008.

ELLIS, Nick C. Optimizing the input: Frequency and Sampling in Usage-based and Form-focussed Learning. In: Michael Long & Catherine Doughty (Eds.), *The Handbook of Second and Foreign Language Teaching*. Oxford: Blackwell, no prelo.

ELLIS, Nick C. & FERREIRA-JUNIOR, Fernando. *Construction learning as a function of Frequency, Frequency Distribution, Form & Function*. Trabalho apresentado no encontro da American Association of Applied Linguistics (AAAL Conference), Washington, DC, em 29 de Março de 2008.

ELLIS, Nick C. & FERREIRA-JUNIOR, Fernando. Second language construction learning as a function of frequency, frequency distribution, and function. *The Modern Language Journal (Special Issue)*, 93(2), no prelo a.

ELLIS, Nick C. & FERREIRA-JUNIOR, Fernando. Constructions and their acquisition: Islands and the distinctiveness of their occupancy. *Annual Review of Cognitive Linguistics (special section)*, vol. 7, no prelo b.

ELLIS, Nick C. & LARSEN-FREEMAN, Diane. Language Emergence: Implications for Applied Linguistics. *Applied Linguistics*, vol. 27 (4), 2006. p. 558-589.

ELLIS, Rod. *Second Language Acquisition*. Oxford: Oxford University Press, 1997.

ELMAN, Jeffrey L. Finding structure in time. *Cognitive Science*, vol.14, 1990. p.179-211.

- ELMAN, Jeffrey L. et al. *Rethinking Innateness: a connectionist perspective on development*. Cambridge, MA: MIT Press, 1998.
- ELMAN, Jeffrey. L. An alternative view of the mental lexicon. *TRENDS in Cognitive Sciences*, vol. 8, n.7, 2004. p.301-306.
- ELMAN, Jeff, HARE, Mary e McRAE, Ken. Cues, Constraints and Competitions in Sentence Processing. In: Michael Tomasello & Dan Slobin: *Beyond Nature-Nurture: Essays in Honor of Elizabeth Bates*. New Jersey: Lawrence Erlbaum Associates, p.111-138. 2005.
- EVANS, Vyvyan & TYLER, Andrea. Applying Cognitive Linguistics to Pedagogical Grammar: The English Prepositions of Verticality. *Revista Brasileira de Linguística Aplicada*, vol.5, n.2, 2005. p.11-42.
- FERREIRA-JUNIOR, Fernando. Programas de rádio: uma alternativa de insumo em língua estrangeira. Dissertação de Mestrado publicada online em <http://libdigi.unicamp.br/document/?code=vtls000212327>. Universidade Estadual de Campinas, Unicamp, 149p. 2000.
- FERREIRA-JUNIOR, Fernando. Competências e módulos: A terminologia cognitivista na reforma curricular dos CEFETs. *Estudos Linguísticos*, vol. 33, 2004. p. 550-555.
- FERREIRA-JUNIOR, Fernando. Dos grupos de discussão às redes neurais: Reflexões sobre o desenvolvimento de um léxico mental. *Revista Brasileira de Linguística Aplicada*, vol.5, n.2, 2005. p.231-252.
- FERREIRA-JUNIOR, Fernando. Uma Interlíngua Conexionista. *Trabalhos em Linguística Aplicada*, vol. 46, n.2, 2007. p. 219-231.
- FODOR, Jerry A. Four Accounts of Mental Structure. In: *The Modularity of Mind: An Essay on Faculty Psychology*. Cambridge: MIT Press, p.1-38. 1983.
- GARDNER et al. *Inteligência: múltiplas perspectivas*. Porto Alegre: Artmed, 1998.
- GASS, Susan. *Input, Interaction, and the Second Language Learner*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1997.
- GOLDBERG, Adele. Constructions: a new theoretical approach to language. *TRENDS In Cognitive Sciences*, vol. 7, 2003. p. 219-224.
- GOLDBERG, Adele. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press, 2006.
- GOLDBERG, Adele & CASENHISER, Devin. Construction learning and second language acquisition. In: Peter Robinson & Nick C. Ellis (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition*. Hillsdale, NJ: Lawrence Erlbaum, p.197-215. 2008.

- GOLDBERG, Adele, CASENHISER, Devin & SETHURAMAN, Nitya. Learning argument structure generalizations. *Cognitive Linguistics*, vol. 15, 2004. p. 289–316.
- GÓMEZ, Rebecca & GERKEN, LouAnn. Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, vol. 70, 1999. p.109-135.
- GÓMEZ, Rebecca & GERKEN, LouAnn. Infant artificial language learning and language acquisition. *TRENDS in Cognitive Sciences*, vol.4, n. 5, 2000. p.178-186.
- GOODWIN, Charles & DURANTI, Alessandro. Rethinking Context: An Introduction. In: Charles Goodwin and Alessandro Duranti (Eds), *Rethinking Context: Language as an Interactive Phenomenon*. Cambridge: Cambridge University Press, p. 1-42. 1992.
- GRANGER, Sylviane. Prefabricated Patterns in Advanced EFL Writing: Collocations and Formulae. In: A. P. Cowie Cowie (ed.) *Phraseology: Theory, Analysis, and Applications*. Oxford: Oxford University Press, p. 145-160. 2001.
- GREGORY, Richard L. (org.) *The Oxford Companion to the Mind*: Oxford: OUP, 1998.
- GRIES, Stefan. Corpus-Based Methods in Analyses of Second Language Acquisition Data. In: Peter Robinson & Nick C. Ellis (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition*. Hillsdale, NJ: Lawrence Erlbaum, p.406-431. 2008
- GRIES, Stefan & WULFF, Stefanie. Do foreign language learners also have constructions? Evidence from priming, sorting, and corpora. *Annual Review of Cognitive Linguistics*, vol. 3, 2005. p. 182-200.
- GRIES, Stefan, HAMPE, Beate, & SCHÖNEFELD, Doris. Converging evidence: bringing together experimental and corpus data on the association of verbs and constructions. *Cognitive Linguistics*, vol. 16 (4), 2005. p. 635-676.
- HARLEY, Trevor A. *The Psychology of Language (from Data to Theory)*. Hove: Psychology Press, 1998.
- JONES, Nancy E. The neurobiology of memory consolidation. In: John Schumann et al. *The neurobiology of learning: Perspectives from second language acquisition*. New Jersey: Lawrence Erlbaum Associates, p.111-142. 2004.
- JORDAN, Michael L. Recurrent networks. In: *MITECS files on connectionism*. MIT, 1999.
- KE, Jinyun. *Network analyses for language development & language teaching*, ELI Winter Seminar Series, University of Michigan, ELI, 26 de Janeiro de 2007. ppt disponível em [http://www-personal.umich.edu/~jyke/presentELI070126\\_network.pps](http://www-personal.umich.edu/~jyke/presentELI070126_network.pps)
- KLEIN, Wolfgang. The acquisition of temporality. In: Clive Perdue (Ed.) *Adult language acquisition: Crosslinguistic perspectives* (Vol. 2- The Results). Cambridge: Cambridge University Press, p.73-118, 1993.

KRASHEN, Stephen. *Principles and Practice in Second Language Acquisition*. New York: Pergamon Press, 1982.

LANGACKER, Ronald W. Part I - Preliminaries. In: *Cognitive Grammar (A Basic Introduction)*. Oxford: Oxford University Press, p. 3-85. 2008.

LANGACKER, Ronald W. Fundamental Concepts. In: *Foundations of Cognitive Grammar*. Stanford: Stanford University Press, p.57-96. 1999.

LANTOLF, James P. & PAVLENKO, Aneta. "(S)econd (L)anguage (A)ctivity theory: understanding second language learners as people" IN: BREEN, Michael P. *Learner Contributions to Language Learning*. Harlow: Pearson Education Limited, 2001.

LARSEN-FREEMAN, Diane. Chaos/Complexity Science and Second Language Acquisition. *Applied Linguistics*, Vol. 18 (2), 1997. p.141-165.

LEE, Namhee. The Neurobiology of Procedural Knowledge. In: John Schumann et al. *The neurobiology of learning: Perspectives from second language acquisition*. New Jersey: Lawrence Erlbaum Associates, p.43-73. 2004.

LIEVEN, Elena, & TOMASELLO, Michael. Children's first language acquisition from a usage-based perspective. In: Peter Robinson & Nick C. Ellis (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition*. Hillsdale, NJ: Lawrence Erlbaum, p.168-196. 2008.

LONG, Michael & DOUGHTY, Catherine. SLA and Cognitive Science. In: Doughty, C. & Long, M. (Eds.). *The Handbook of Second Language Acquisition*. Oxford, UK: Blackwell, p. 866-870. 2003.

MARCHMAN, Virginia & THAL, Donna. Words and Grammar. In: Michael Tomasello & Dan Slobin: *Beyond Nature-Nurture: Essays in Honor of Elizabeth Bates*. New Jersey: Lawrence Erlbaum Associates, p.141-164. 2005.

MAcWHINNEY, Brian. *The CHILDES Project: Tools for Analyzing Talk*. Volume 1: Transcription format e programs. Mahwah, NJ: Lawrence Erlbaum Associates, 2000 (links disponibilizados em <http://psyling.psy.cmu.edu/papers/talkbank/> . Último acesso em 12 de Fevereiro de 2008).

MAcWHINNEY, Brian. The competition model: the input, the context, and the brain. In: Peter Robinson (org.). *Cognition and Second Language Instruction*. Cambridge: Cambridge University Press, p.69-90. 2001.

MAcWHINNEY, Brian. New Directions in the Competition Model. In: Michael Tomasello and Dan Slobin: *Beyond Nature-Nurture: Essays in Honor of Elizabeth Bates*. New Jersey: Lawrence Erlbaum Associates, p.81-110. 2005.

McCLELLAND, James, ROGERS, Timothy. The Parallel Distributed Processing Approach to Semantic Cognition. *Nature Neuroscience*, vol.4, 2003. p.310-322.

McCLELLAND, James & RUMELHART, David. An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, vol. 88, n.5, 1981. p. 375-407.

McCLELLAND, James & ELMAN, Jeffrey. The TRACE Model of Speech Perception. *Cognitive Psychology*, vol. 18, 1985. p. 1-86.

McCLELLAND, James, RUMELHART, David. & HINTON, Geoffrey. The Appeal of Parallel Distributed Processing. In: Rumelhart, McClelland and the PDP Research Group. *Parallel Distributed Processing (Explorations in the Microstructure of Cognition)*, Vol.1. MIT Press, 1986.

McENERY, Tony & WILSON, Andrew. *Corpus Linguistics*. Edinburgh: Edinburgh University Press, 1997.

McLAUGHLIN, Barry. *Theories of Second Language Learning*. London: E. Arnold, 1987.

MEARA, Paul. The dimensions of lexical competence. In: Brown, G. et al. *Performance & Competence in Second Language Acquisition*. Cambridge: Cambridge University Press, p. 35-53. 1996.

MITCHELL, Rosamond & MYLES, Florence. Functional/pragmatic perspectives on second language learning. In: R. Mitchell & F. Myles, *Second Language Learning Theories (2<sup>nd</sup> Edition)*. London: Arnold, p. 131-158. 2004.

MOTA, Mailce & ZIMMER, Márcia. Cognição e aprendizagem de L2: o que nos diz a pesquisa nos paradigmas simbólico e conexionista. *Revista Brasileira de Linguística Aplicada*, vol.5, n.2, 2005. p.155-187.

MURPHY, Gregory. Typicality and the classical view of categories. In: Gregory Murphy, *The Big Book of Concepts*. Cambridge, MA: MIT Press, p.1-40. 2002.

NEY, James & PEARSON, Bethyl. Connectionism as a model of language learning: parallels in foreign language teaching. *The Modern Language Journal*, vol. 74, n.4, 1990. p. 472-482.

OAKSFORD, Mike & CHATER, Nick. *Bayesian Rationality (the probabilistic approach to human reasoning)*. Oxford: OUP, 2007. 330p.

O'GRADY, William. The Radical Middle: Nativism Without Universal Grammar. In: Doughty, C & Long., M. (Eds.). *The Handbook of Second Language Acquisition*. Oxford, UK: Blackwell, p. 43-62. 2003.

O'REILLY, Randall & NORMAN, Kenneth. Hippocampal and neocortical contributions to memory: advances in the complementary learning systems framework. *TRENDS in Cognitive Sciences*, vol.16, n.12, 2002. p.505-510.

ORTEGA, Lourdes, & IBERRI-SHEA, Gina. Longitudinal research in second language acquisition: recent trends and future directions. *Annual Review of Applied Linguistics*, vol. 25, 2005. p. 26-45.

PERDUE, Clive. (Ed.) *Adult language acquisition: Crosslinguistic perspectives* (Vol. 1-Field Methods). Cambridge: Cambridge University Press, 1993.

PERDUE, Clive. (Ed.) *Adult language acquisition: Crosslinguistic perspectives* (Vol. 2-The Results). Cambridge: Cambridge University Press, 1993.

PIAGET, Jean. The Psychogenesis of Knowledge and Its Epistemological Significance. In: Massimo Piattelli-Palmarini (org.) *Language and Learning (The debate between Jean Piaget and Noam Chomsky)*: London: Routledge & Kegan Paul, p.23-34. 1980.

PINKER, Steven. *The Language Instinct*. London: Penguin Books, 1994.

ROGERS, Timothy & McCLELLAND, James. *Semantic Cognition: A Parallel Distributed Processing Approach*. Cambridge, MA: MIT Press, 2004.

ROSCH, Eleanor. Principles of Categorization. In: A. Collins & E. Smith (Eds.), *Readings in Cognitive Science: a perspective from psychology and artificial intelligence*. San Mateo, CA: Morgan Kauffmann Publishers, p. 312-322. 1988.

RUMELHART, David & McCLELLAND, James. On learning the past tenses of English verbs. In: D.E. Rumelhart & J. L. McClelland (orgs) In: *Parallel Distributed Processing (Explorations in the Microstructure of Cognition)*, Vol.1. MIT Press, p.216-271. 1986.

RUSSELL, Bertrand. *On Desire*. Integra do discurso disponível em <http://nobelprize.org/literature/laureates/1950/russell-lecture.html> . Acesso em agosto de 2007.

SARDINHA, Tony B. Computador, corpus e concordância no ensino do léxico-gramática da língua estrangeira. In: Wilson Leffa (org). *As palavras e sua companhia: O léxico na aprendizagem das línguas*, Pelotas: EDUCAT, UCP, 2000. Arquivo disponível em <http://www2.lael.pucsp.br/~tony/temp/publications/2000vilson.pdf> . Acesso em Fevereiro 2008.

SCHMIDT, Richard. Attention. In: Peter Robinson (org.). *Cognition and Second Language Instruction*. Cambridge: CUP, p.02-32. 2001.

SCHULTZ, Duane. *História da Psicologia Moderna*. São Paulo: Cultrix, 1981.

SCHUCHERT, Sara. The Neurobiology of Attention. In: John Schumann et al. *The neurobiology of learning: Perspectives from second language acquisition*. New Jersey: Lawrence Erlbaum Associates, p.143-173. 2004.

SCHUMANN, John et al. *The neurobiology of learning: Perspectives from second language acquisition*. New Jersey: Lawrence Erlbaum Associates. 2004.



SCHUMANN, John e WOOD, Lee. The Neurobiology of Motivation. In: In: John Schumann et al. *The neurobiology of learning: Perspectives from second language acquisition*. New Jersey: Lawrence Erlbaum Associates, p.23-42. 2004.

SEIDENBERG, Mark. Language Acquisition and Use: Learning and Applying Probabilistic Constraints. *Science*, vol. 275, 1997. p.1599-1603.

SEIDENBERG, Mark, e MacDONALD, Maryellen. Constraint Satisfaction in Language Acquisition and Processing. In: Morten Christiansen and Nick Chater. *Connectionist Psycholinguistics*. Ablex: Connecticut, p.281-318. 2001.

SELINKER, Larry. Interlanguage. *IRAL*, 10, 1972. p.209-31.

SELINKER, Larry. On the notion of 'IL competence' in early SLA research: an aid to understanding some baffling current issues. In: Gillian Brown et. al. (orgs). *Performance & Competence in Second Language Acquisition*. Cambridge: Cambridge University Press, p 92-113. 1996.

SINGLETON, David. Lexis and the lexicon: some general considerations. In: David Singleton. *Exploring the Second Language Mental Lexicon*. Cambridge: Cambridge University Press, p. 08-38, 1999.

SKEHAN, Peter. *A cognitive Approach to Language Learning*. Oxford: OUP, 1998.

TAYLOR, John R. Prototypes in Cognitive Linguistics. In: Peter Robinson & Nick C. Ellis (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition*. Hillsdale, NJ: Lawrence Erlbaum, p. 39-65. 2008.

TOMASELLO, Michael. The item-based nature of children's early syntactic development. *TRENDS in Cognitive Sciences*, Vol.4, No 4, Pp.156-163, 2000.

TOMASELLO, Michael. *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press, 2003.

TYLER, Andrea. Cognitive Linguistics and Second Language Instruction. In: Peter Robinson & Nick C. Ellis (Eds.), *Handbook of Cognitive Linguistics and Second Language Acquisition*. Hillsdale, NJ: Lawrence Erlbaum, p. 456-488. 2008.

VON GOLDAMMER, Eberhard. *Heterarchy and Hierarchy – Two Complementary Categories Of Description*. Disponível em [http://www.vordenker.de/heterarchy/het\\_intro\\_en.htm](http://www.vordenker.de/heterarchy/het_intro_en.htm) Acesso em Fevereiro 2008.

WEINREICH, Uriel. *Languages in contact*. The Hague: Mouton, 1953.

WIDDOWSON, Henry G. *Linguistics*. Oxford: Oxford University Press, 1996.

WHITE, Lydia. On the Nature of Interlanguage Representation: Universal Grammar in the Second Language. In: Doughty, C & Long., M. (Eds.). *The Handbook of Second Language Acquisition*. Oxford, UK: Blackwell, Pp. 19-42, 2003

WINTER, Bill & REBER, Arthur. Implicit Learning and the Acquisition of Natural Languages. In: N. Ellis (Ed.). *Implicit and Explicit Learning of Languages*. London: Academic Press, 1994.

ZIMMER, Márcia. Cognição e Aprendizagem de L2: uma abordagem conexionista. In: Macedo, A.C.P.; Feltes, H.; Farias, E. M. P. (Org.). *Cognição e Lingüística: Territórios, Mapeamentos e Percursos*. Porto Alegre: EDIPUCRS, EDUSC, p. 229-248, 2008.

## Anexo A

### Material utilizado nos experimentos sobre prototipicidade

| VL (verb locative)  | VOL (verb object locative)  | VOO (ditransitive)                                     |
|---|---|--|
| go, come, sit, look, get, live, put, turn, drop, move, fall, sleep, take, open, pick, run, walk, work, see, return. | put, take, turn, drop, move, bring catch, have, send, keep, open, leave, cross, sit, look, get, pick, see, push, think. | give, write, ask, send, buy, tell, show, pay, explain. |

Tabela 1 – verbos mais frequentes para cada construção

| + Best | middle |   |   | Worst - |
|--------|--------|---|---|---------|
| 1      | 2      | 3 | 4 | 5       |

Tabela 2 – Escala de cinco pontos para a classificação dos verbos apresentados



| VL   | VOL   | VOO  |
|--|---|--|
| <p>In English some verbs are used to describe the movement of something or someone (X) to a new place (Y) or in a new direction.</p>  <p>X                  Y</p> <p>e.g. The fly buzzed into the room</p> <p>Sort the verbs on following cards into five piles from those which work best in this way to those that work worst in this way</p> | <p>In English some verbs are used to describe how someone (A) causes the movement of something (X) to a new place (Y) or in a new direction.</p>  <p>A causes X                  Y</p> <p>e.g. Pat blew the feather into the drawer</p> <p>Sort the verbs on following cards into five piles from those which work best in this way to those that work worst in this way</p> | <p>In English some verbs are used to describe how someone (A) causes someone (B) to receive something (Z)</p> <p>A causes Z → B</p> <p>e.g. Pat faxed Tom the picture</p> <p>Sort the verbs on following cards into five piles from those which work best in this way to those that work worst in this way</p> |

Tabela 3 – Instruções dadas aos participantes para cada tarefa

| VL                       | Ayla | Katherine | Fernando | Jennifer | Joshua | Brian | Kelly | George | Jaimie | Maurice | Jenny | Rafael  | Silvia     | Enio       | Adriana    |
|--------------------------|------|-----------|----------|----------|--------|-------|-------|--------|--------|---------|-------|---------|------------|------------|------------|
| NS or NNS                | NS   | NS        | NS       | NS       | NS     | NS    | NS    | NS     | NS     | NS      | NS    | NNS     | NNS        | NNS        | NNS        |
| NNS's language           |      |           |          |          |        |       |       |        |        |         |       | Spanish | Portuguese | Portuguese | Portuguese |
| Age                      | 19   | 20        | 26       | 21       | 25     | 28    | 25    | 48     | 20     | 22      | 26    | 39      | 26         | 48         | 37         |
| NNS's years of residency |      |           |          |          |        |       |       |        |        |         |       | 1       | 1          | 13         | 30         |
| Gender                   | F    | F         | M        | F        | M      | M     | F     | M      | F      | M       | F     | M       | F          | M          | F          |
| go                       | 1    | 1         | 1        | 1        | 1      | 1     | 1     | 1      | 1      | 1       | 1     | 1       | 1          | 1          | 1          |
| come                     | 1    | 3         | 2        | 2        | 1      | 1     | 1     | 2      | 1      | 1       | 1     | 2       | 1          | 1          | 1          |
| sit                      | 2    | 5         | 5        | 5        | 2      | 5     | 5     | 5      | 4      | 3       | 2     | 5       | 5          | 3          | 2          |
| look                     | 3    | 5         | 5        | 4        | 5      | 5     | 5     | 5      | 3      | 3       | 5     | 4       | 5          | 5          | 4          |
| get                      | 3    | 5         | 5        | 3        | 5      | 4     | 4     | 4      | 4      | 2       | 3     | 2       | 3          | 3          | 3          |
| live                     | 3    | 4         | 5        | 4        | 5      | 5     | 5     | 4      | 2      | 5       | 5     | 5       | 5          | 5          | 1          |
| put                      | 1    | 2         | 5        | 3        | 1      | 5     | 3     | 3      | 5      | 1       | 2     | 1       | 4          | 1          | 1          |
| turn                     | 1    | 2         | 1        | 3        | 3      | 3     | 3     | 5      | 3      | 2       | 1     | 2       | 2          | 1          | 1          |
| drop                     | 1    | 2         | 3        | 2        | 1      | 1     | 2     | 3      | 5      | 2       | 2     | 2       | 4          | 1          | 1          |
| move                     | 1    | 1         | 1        | 1        | 2      | 2     | 1     | 1      | 1      | 1       | 1     | 1       | 1          | 1          | 1          |
| fall                     | 1    | 2         | 1        | 2        | 2      | 1     | 2     | 3      | 5      | 2       | 1     | 2       | 2          | 1          | 1          |
| sleep                    | 4    | 5         | 5        | 5        | 5      | 5     | 3     | 5      | 5      | 4       | 4     | 4       | 5          | 5          | 5          |
| take                     | 2    | 3         | 5        | 3        | 1      | 2     | 3     | 4      | 4      | 2       | 4     | 1       | 4          | 1          | 1          |
| open                     | 2    | 4         | 5        | 3        | 4      | 5     | 4     | 4      | 2      | 2       | 4     | 4       | 5          | 5          | 2          |
| pick                     | 5    | 4         | 5        | 4        | 4      | 5     | 5     | 5      | 4      | 3       | 4     | 5       | 5          | 4          | 1          |
| run                      | 1    | 1         | 1        | 1        | 1      | 1     | 1     | 1      | 1      | 1       | 1     | 1       | 1          | 1          | 1          |
| walk                     | 1    | 1         | 1        | 1        | 1      | 1     | 1     | 1      | 1      | 1       | 1     | 1       | 1          | 1          | 1          |
| work                     | 2    | 4         | 5        | 5        | 5      | 5     | 4     | 5      | 4      | 3       | 4     | 5       | 5          | 5          | 5          |

**Tabela 4**

Modelo de formulário usado para registrar as informações pessoais dos participantes e os resultados das tarefas por eles feitas

|                                  |                                  |
|----------------------------------|----------------------------------|
| Anita threw a hammer             | Laura got the ball into the net  |
| Audrey took the watch            | Lyn threw the box apart          |
| Barbara cut the bread            | Meg cut the ham onto the plate   |
| Beth got Liz an invitation       | Michelle got the book            |
| Chris threw Linda the pencil     | Nancy cut the tire open          |
| Dana got the mattress inflated   | Pat threw the keys onto the roof |
| Jennifer cut Terry an apple      | Paula took Sue a message         |
| Kim took the rose into the house | Rachel took the wall down        |

**Tabela 5**

Frases usadas no estudo de classificação de frases conduzido por Gries & Wulff (2005)

## **Anexo B**

### **Informação sociobiográfica dos sujeitos<sup>82</sup>**

#### **1. Informações sociobiográficas de Andrea**

Andrea era à época um italiano com trinta e poucos anos. Ele terminou o ensino médio (com aulas de francês no currículo) e então obteve qualificação profissional como técnico. Foi para a Inglaterra por motivos pessoais, após se separar de sua esposa, deixando-a juntamente com um filho na Itália. Ele trabalhava no centro de Londres como garçom e barman, trabalhando longas horas em turnos na companhia, principalmente, de colegas italianos. Com a exceção de apenas uma pessoa inglesa, seus amigos são também italianos. Ele depois conseguiu um emprego melhor numa padaria.

Quando foi entrevistado pela primeira vez ele estava tendo aulas de inglês três vezes por semana, mas pressões do trabalho (turnos de trabalho fragmentados) logo o fizeram parar o curso e não mais voltar a ter aulas de inglês. Sua esposa e filho juntaram-se a ele em Londres já perto do final das sessões do estudo ESF.

Seus poucos contatos com falantes nativos de inglês eram principalmente em instituições: companhias nas quais ele solicitava emprego, as obrigações normalmente exigidas de quem é proprietário de carro, funcionários públicos e agentes de viagem. Após a chegada de sua esposa e filho, ele passou a ter contato com escolas, médicos, dentistas, etc. Andrea era ainda iniciante em inglês quando começou o projeto ESF, apesar de frequentar aulas de inglês no período acima descrito. Por todo o período do projeto, ele permaneceu um aprendiz reservado e cauteloso mas obteve progresso considerável ao longo do estudo. Ele usava as conversas com os pesquisadores como uma oportunidade pedagógica, em especial no desenvolvimento do léxico. Entretanto, nos encontros internos e externos ao projeto ele permanecia bastante receoso/tímido no

---

<sup>82</sup> Textos traduzidos na íntegra a partir dos anexos em Perdue, 1993, vol 1, Pp.189-207 e informações disponíveis no site do projeto ESF em [http://corpus1.mpi.nl/ds/imdi\\_browser/](http://corpus1.mpi.nl/ds/imdi_browser/) ).

uso da L2. Parece ter atingido sintaxe adequada e bom vocabulário, mas distante de uma competência nativa.

## **2. Informações sociobiográficas de Lavinia**

Lavinia chegou a Inglaterra vinda de Trieste com vinte e poucos anos, na companhia de seu marido e de seu filhinho. Ela nunca havia trabalhado na Itália. Na Inglaterra, trabalhou primeiro como garçomete, mas depois que começou a participar do projeto ESF ela frequentou cursos de inglês.

Lavinia progrediu muito rapidamente no curso de inglês (passando a vários estágios). Teve dois empregos de trabalho não especializado e de vendedora informal de roupas. do curso. Seu principal contato com a língua inglesa além do curso de inglês era com as crianças (ela teve um segundo filho no final do projeto ESF), ou seja, saúde e cuidados infantis, escola, etc. No entanto, apesar de ser uma pessoa muito expansiva, ela achava extremamente difícil fazer amigos e sentia-se bastante isolada. Apesar de ter gostado de ir para o Reino Unido pelo fato de passar a trabalhar fora, as dificuldades enfrentadas por ela e seu marido com relação a emprego e moradia acabaram fazendo com que ambos retornassem a Itália.

Embora Lavinia possuísse um domínio limitado da língua inglesa ao chegar no Reino Unido, a sua aquisição da L2 foi muito rápida. Ela era obviamente bastante motivada para aprender o inglês e aproveitou ao extremo os cursos de inglês que frequentou. No final do estudo ela tornou-se bastante fluente mesmo.

## **3. Informações sociobiográficas de Santo**

Santo tinha vinte e poucos anos à época do estudo ESF. Nascido na província de Nápoles, sua L1 era o dialeto napolitano do italiano falado em sua cidade natal. Com oito anos de educação formal, chegou na Inglaterra em janeiro de 1983 para trabalhar, assim como por motivos pessoais, e esperava ficar indefinidamente. Sua primeira entrevista no projeto ocorreu seis meses após sua chegada. Na época, alugava um quarto numa casa com pessoas de várias nacionalidades, trabalhava num restaurante italiano e tinha uma namorada falante de italiano. Relatava falar italiano tanto no trabalho como

com sua namorada e família. Entretanto, no início do estudo suas habilidades de produção e compreensão do inglês eram consideradas muito boa. A habilidade de leitura, por outro lado, era fraca e ele afirmava não saber escrever em inglês.

No início, o repertório de Santo se assemelhava bastante ao de Andrea quando comparados dentro do mesmo período de tempo. Mas, num forte contraste em relação ao desenvolvimento de Andrea, a interlíngua de Santo mudou pouco ao longo do período de observação. Santo se tornou mais proficiente no uso do mesmo repertório.

#### **4. Informações sociobiográficas de Vito**

Vito chegou a Inglaterra para trabalhar aos 23 anos, vindo da Itália. Possuía ensino médio incompleto.

#### **5. Informações sociobiográficas de Madan**

Madan chegou na Inglaterra em fevereiro de 1981, aos 25 anos, para um casamento arranjado com uma mulher punjabi de cidadania britânica residente no Reino Unido há dezoito anos. Madan nasceu na Índia e cresceu falando o dialeto mawlai do punjabi. Ele havia completado seis anos de ensino primário na sua cidade nativa e, após o período escolar, ali trabalhado na fazenda da família. Estou inglês como LE durante um ano e hindi durante três.

Antes de chegar a Inglaterra, Madan passou um período de tempo trabalhando no Afeganistão, Jordânia e Síria, mas afirmava não ter aprendido nenhuma outra língua durante este tempo. Na época do projeto ESF, que teve início vinte meses após sua chegada na Inglaterra, ela trabalhava como operador de prensa em uma fábrica, onde usava tanto o inglês como o punjabi, embora nos intervalos para refeição no trabalho ele ficava com os amigos falantes de punjabi. Ele dividia sua casa com a família da irmã de sua esposa. Embora sua esposa falasse inglês muito bem, ele relatava pouca exposição ao inglês: a maior parte no trabalho, à noite na loja do irmão, ou por meio da TV. A esposa de Madan e sua irmã se comunicavam em inglês, mas Madan ficava obviamente inibido pela proficiência das duas em inglês e assim só falava punjabi na presença delas.



Apesar de trabalhar muito, Madan era extremamente cooperativo e parecia adorar as sessões. No início do projeto sua competência na L2 era muito limitada, mas ele era muito cooperativo e buscava interação. Ele se esforçava ao extremo para ser compreendido na L2 e era incansável em se fazer compreender. Ao final do projeto, sua confiança no uso da L2, principalmente com os pesquisadores, melhorou consideravelmente. Por exemplo, ele discordava de comentários dos falantes nativos mas mantinha-se interativamente cooperativo. Sua competência em sintaxe e léxico ainda era de certa forma limitada.

## **6. Informações sociobiográficas de Ravinder**

Ravinder chegou no Reino Unido aos vinte anos, procedente de uma vila punjab onde fez o primário. Logo após chegar ele casou com uma garota que cresceu e estudou no Reino Unido e, portanto, era considerada falante nativa do inglês. Entretanto, ela e Ravinder se comunicavam principalemnte em punjabi.

Inicialmente, o cunhado de Ravinder apoiou a idéia de que ele deveria freqüentar aulas de inglês, mas sua motivação acabou a partir da hora em que ele percebeu que poderia ser útil nos negócios de peixe e batata-fritas da família e material de construção com ou sem o conhecimento da língua inglesa. Assim, o principal contato de Ravinder com o inglês era por meio de clientes nos dois comércios. Tal contato foi também suplementado com contatos sociais em bares, embora muitos bares na área na qual ele vivia eram de proprietários falantes de punjabi.

Ravinder era tímido e às vezes bastante embaraçado durante as sessões devido a sua baixa competência em inglês, embora tenha dito certa vez que queria aprender e gostava da língua inglesa. Ao final do projeto, o contato mínimo com o inglês deu a ele confiança para ser aprovado no exame de motorista e passar a fazer entregas para o seu cunhado, o que dava a ele a impressão de que seu inglês estava adequado às suas necessidades.

## **7. Informações sociobiográficas de Jarnail**

Jarnail chegou na Inglaterra em Abril de 1982 aos 27 anos, procedente da província de Punjab, na Índia, para um casamento arranjado com uma mulher oriunda de uma família punjabi, mas nascida no Reino Unido. Sua qualificação profissional era como motorista de cargas pesadas. Jarnail só falava Punjabi com a esposa e parentes.