

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

INSTITUTO DE LETRAS

CURSO DE PÓS-GRADUAÇÃO EM LETRAS

**A PRODUÇÃO DE PERGUNTAS EM INGLÊS COMO LÍNGUA
ESTRANGEIRA POR APRENDIZES BRASILEIROS**

UMA PERSPECTIVA CONEXIONISTA

CLÁUDIO PASSOS DE OLIVEIRA

Dissertação apresentada à Banca
Examinadora como requisito parcial à
obtenção do título de Mestre em Letras:
Estudos da Linguagem

Orientadora: Professora Dra. Luciene Juliano Simões

Porto Alegre, novembro de 2001

*I will put Chaos into fourteen lines
And keep him there; and let him thence escape
If he be lucky; let him twist and ape
Flood, fire, and demon - his adroit designs
Will strain to nothing in the strict confines
Of this sweet Order, where, in pious rape,
I hold his essence and amorphous shape,
Till he with Order mingles and combines.
Past are the hours, the years, of our duress,
His arrogance, our awful servitude:
I have him. He is nothing more nor less
Than something simple not yet understood;
I shall not even force him to confess;
Or answer. I will only make him good.*

Edna St. Vincent Millay

AGRADECIMENTOS

Olho para trás e me dou conta: esta dissertação tinha tudo para não se materializar. Ela dependeu demasiadamente da fortuna, existiu sob ameaça permanente. Se não fosse a paciência de Jô de minha amada esposa, Luciane, seus cafunés e palavras de compreensão, se não fossem os sorrisos de incentivo de meus pais, Paulo e Suely, e os fluidos positivos, via satélite, de meu irmão, Cássio, e da cunhadinha, Giovana, sei que este trabalho simplesmente não haveria vindo ao mundo. E se meus fiéis amigos terapêuticos, Marco e Mário, não houvessem me arrancado de casa aos sábados pra tomar cerveja e jogar conversa fora? Certamente esta obra haveria morrido já na Introdução. Sim, sua confecção esteve sempre por um fio. Sei que, se não fosse a heróica resistência da UFRGS ao desmanche federal, a preciosa ajuda e os instigantes encontros com a minha querida orientadora, Luciene, o apoio e as excelentes aulas dos meus dedicados professores do mestrado, nada haveria sido. O fracasso e o desânimo haveriam prevalecido, tenho certeza. E ainda foram necessários os cafés com salgadinho no Bar do Antônio, na agradável companhia dos colegas mestrandos. Sem tudo isso... nem pensar, não haveria dado em nada. Aliás, sem café não haveria dado, e muito menos sem aquele cálice inspirador de vinho tinto. Foi assim, no fio da navalha, todo o tempo. Se não fossem os recursos bibliográficos e as dicas das amigas Tirza, Ingrid e Márcia (essa última culpada direta pelo meu cacoete conexcionista), e a flexibilidade brasileira em relação ao xerox... Eu não teria por que agradecer, pois simplesmente nada haveria sido. E se o meu solidário colega de Cultural, Clóvis, não houvesse se interessado pelo assunto e me ajudado a pensar o problema? Sem chances, esta dissertação haveria morrido no parto. E ainda houve o chocolate. Aquele dos indispensáveis bombons que utilizei para subornar meus queridos colegas do Cultural e seus alunos a serem minhas cobaias voluntárias. Mas de nada haveria adiantado ter cobaias se não houvesse existido a boa vontade e o incentivo do João, Gerente de Cursos do Cultural, ou os apoios da Vânia, Diretora de Cursos, e do Dr. Amaral, nosso Presidente. Suas generosas autorizações para o bem da ciência e da minha saúde mental foram essenciais e significativas. Sem tudo isso, não tenho dúvidas, não haveria como, não seria, não daria.

Sim, tenho muito a agradecer à enorme rede humana de afeto e apoio que se costurou ao meu redor ao longo desses mais de dois anos e meio de trabalho. A ela dedico as páginas a seguir.

SUMÁRIO

LISTA DE ABREVIATURAS	VI
LISTA DE FIGURAS	VII
LISTA DE QUADROS	VIII
LISTA DE TABELAS	IX
RESUMO	X
ABSTRACT	XI
INTRODUÇÃO	001
1 FUNDAMENTOS TEÓRICOS	006
1.1 Um Breve Histórico do Conexionismo	007
1.2 As Redes Neurais	009
1.3 A Abordagem Conexionista da Aquisição da Linguagem	017
1.4 Os Modelos Conexionistas	022
1.5 O Aprendizado nas Redes PDP	026
1.6 As Funções Harmônicas	031
1.7 Críticas ao Conexionismo	035
1.8 O Conexionismo e a Aquisição de L2	043
2 PRODUÇÃO DE PERGUNTAS POR APRENDIZES DE INGLÊS COMO L2	053
2.1 Os Padrões Interrogativos do Inglês	053
2.1.1 As Perguntas Diretas	054
2.2 Metodologia	058
2.2.1 Os Sujeitos	058
2.2.2 O Instrumento	059
2.2.3 A Tabulação dos Dados	061
2.3 Os Resultados	067
2.3.1 O Questionário	068

2.3.2 O Formulário	069
2.3.3 A Análise Estatística	072
3 A SIMULAÇÃO	079
3.1 A Codificação dos Padrões Interrogativos do Inglês	079
3.1.1 As Entradas	079
3.1.2 As Saídas	081
3.2 A Rede	083
3.3 O Treinamento	085
3.4 A Pontuação das Saídas	091
3.5 A Análise de Agrupamento	092
3.6 As Frequências de Ocorrência	092
3.7 A Harmonia	093
3.8 Os Resultados	094
3.8.1 As Saídas e suas Pontuações	094
3.8.2 A Análise de Agrupamento	098
3.8.3 As Frequências Relativas de Ocorrência	100
3.8.4 A Harmonia	104
4 DISCUSSÃO	106
4.1 A Produção dos Sujeitos	106
4.2 A Simulação	112
4.2.1 As Estruturas Inatas da Rede	112
4.2.1.1 A Arquitetura	113
4.2.1.2 A Codificação	113
4.2.1.3 O Modo de Aprender	115
4.2.2 O <i>Input</i> da Rede	116
4.2.3 As Saídas	119
CONCLUSÃO	130
BIBLIOGRAFIA CONSULTADA	136
ANEXOS	141
Anexo A: Instruções Fornecidas aos Sujeitos	142
Anexo B: Formulário de Coleta de Dados	143
Anexo C: Questionário Aplicado aos Sujeitos	144
Anexo D: Codificação de Padrões de Entrada	145
Anexo E: Codificação de Padrões de Saída	147
Anexo F: Saídas da Rede	149
Anexo G: Análise de Agrupamento	160

LISTA DE ABREVIATURAS

Csuj-aux	Concordância Sujeito – Verbo Auxiliar
Caux-main	Concordância Verbo Auxiliar – Verbo Principal
CA	Análise de Agrupamento
H	Harmonia
IL	Interlíngua
L1	Língua Materna
L2	Segunda Língua
MAL	Linguagem Artificial em Miniatura
MM	Modelo Multidimensional
MP	<i>Multilayer Perceptron</i>
N1	Nível de Proficiência 1
N2	Nível de Proficiência 2
N3	Nível de Proficiência 3
N4	Nível de Proficiência 4
NPsubj	Sintagma Nominal - Sujeito
PCA	<i>Análise de Componentes Principais</i>
PDP	<i>Parallel Distributed Processing</i>
PB	Português Brasileiro
S	Sujeito
SRN	<i>Simple Recurrent Network</i>
TH	Teoria da Harmonia
TO	Teoria da Otimização
U	Unidade Neural
VSG	<i>Visitation Set Gravitation</i>
VPaux	Sintagma Verbal – Auxiliar
VPmain	Sintagma Verbal – Principal
XP	Sintagma Qualquer

LISTA DE FIGURAS

FIGURA 1	Constituintes da Célula Nervosa	010
FIGURA 2	<i>Unidade Neural Artificial</i>	011
FIGURA 3	Tipos de Funções de Ativação	011
FIGURA 4	Rede Neural do Tipo MP	012
FIGURA 5	Taxa de Padrões por Nível de Proficiência	070
FIGURA 6	Progressão das Médias de N1-4	073
FIGURA 7	Esquema da Rede	084
FIGURA 8	Nível de Dificuldade de Aprendizado por Etapa	097
FIGURA 9	Precisão Média dos Padrões de Saída	097
FIGURA 10	<i>Desvio Padrão Médio das Saídas</i>	098
FIGURA 11	Frequências Médias de Ocorrência	100
FIGURA 12	Variâncias Médias	101
FIGURA 13	Frequências Relativas do Padrão A8	102
FIGURA 14	Variâncias de A8	103
FIGURA 15	Progressão de H para A8	104

LISTA DE QUADROS

QUADRO 1	Currículo dos Sujeitos	059
QUADRO 2	Codificação dos Dados	063
QUADRO 3	Configuração da Rede	083
QUADRO 4	Padrões de Treinamento	085
QUADRO 5	Parâmetros de Treinamento	088
QUADRO 6	Arquivo de Mapeamento	089
QUADRO 7	Codificação Utilizada na Tradução das Saídas	090

LISTA DE TABELAS

TABELA 1	Perfil Geral dos Sujeitos	068
TABELA 2	Resultados Numéricos Gerais Obtidos Através do Formulário	069
TABELA 3	Resultados por Padrão Interrogativo	070
TABELA 4	Resultados da ANOVA	073
TABELA 5	Resultados dos Testes <i>Post Hoc</i>	075
TABELA 6	Especificações do Treinamento	095
TABELA 7	Resultados Gerais de Ordenamento para A8	103

RESUMO

Este trabalho apresenta um estudo sintático comparativo entre uma simulação conexionista da produção de perguntas em inglês como L2 e dados empíricos relativos ao mesmo fenômeno, coletados para esse fim.

Os dados empíricos foram obtidos através de uma investigação transversal da produção escrita de perguntas em inglês como L2 por alunos brasileiros de uma escola de idiomas de Porto Alegre, divididos em quatro faixas de proficiência. As perguntas produzidas por esses aprendizes foram tabuladas com base em seu nível de precisão estrutural e analisadas estatisticamente.

A simulação conexionista foi realizada através de uma rede neural do tipo *multilayer perceptron*, construída a partir de um simulador. Essa rede foi configurada de modo a operar no âmbito da sintaxe de perguntas em inglês e treinada com base na progressão curricular dos aprendizes estudados. O desempenho da rede foi avaliado nos diferentes estágios do seu treinamento, através da análise tanto das saídas produzidas como das configurações internas construídas.

A partir dos resultados obtidos em ambas as frentes de investigação, a saber, a coleta junto aos aprendizes e a simulação conexionista, foi desenvolvida uma discussão comparativa. Essa discussão salientou algumas potencialidades importantes da utilização da ferramenta neural na formulação de hipóteses de teste em relação ao fenômeno investigado, notadamente no que tange à relação entre a frequência e a variância de componentes sintáticos no *input* em L2 com o qual os aprendizes interagem.

ABSTRACT

This work presents a comparative syntactic study between a connectionist simulation of question production in L2 English and empirical data related to the same phenomenon, collected for this purpose.

The empirical data were obtained in a cross-sectional investigation of the written production of questions in L2 English by Brazilian learners at a language school in Porto Alegre, divided into four proficiency levels. Based on their level of structural accuracy, the questions produced by these learners were tabulated and statistically analyzed.

The connectionist simulation was carried out using a multilayer perceptron neural network devised through a simulator. This network was designed to operate within the syntax of questions in English and trained based on the curricular sequence of the learners studied. Network performance was assessed at different stages of its training through the analysis of both the output produced and the internal configurations constructed.

From the data obtained in both investigation efforts, namely the data collection with the learners and the connectionist simulation, a comparative discussion was developed. This discussion pointed out to some important potentialities of the use of the neural tool in the formulation of test hypothesis as far as the investigated phenomenon is concerned, mainly in aspects regarding the relation between the frequency and the variance of syntactical components in the input in L2 with which learners interact.

INTRODUÇÃO

Esta dissertação é culpa de meu interesse pelo modo como as pessoas aprendem uma língua estrangeira. Uma inquietação que surgiu em sala de aula mesmo, enquanto eu observava aquele monte de gente tentando aprender inglês, encarando uma tarefa que, sabem todos os que já tentaram, é das mais árduas. Observei-os atentamente durante cinco anos e minha mente ficou, pouco a pouco, transbordando de perguntas. Da maturação informal das perguntas nasceram o mestrado e, conseqüentemente, as leituras. Lembro de meu assombro ao ler as primeiras páginas sobre gerativismo: elegante, explicativo, maravilhoso. Nele depositei, inicialmente, minhas esperanças de desenvolver um trabalho consistente em aquisição de L2. Também já tinha ouvido falar de conexionismo, mas não tinha muito clara a noção exata do que se tratava. Eu estava muito ocupado lendo Haegeman, imaginando se um dia eu iria poder falar cara a cara com o Chomsky.

Pois eis que um belo dia... Peguei um artigo da Larsen-Freeman para ler: *Chaos / Complexity Science and Second Language Acquisition* (A Ciência do Caos / da Complexidade e a Aquisição de Segunda Língua - Larsen-Freeman 1997). Assim, desavisado, para quebrar um pouco a minha rotina de gerativista amador. Aquelas coisas da vida: o artigo era parte das leituras obrigatórias de uma das cadeiras do mestrado, que acabei não lendo por ocasião de uma viagem a trabalho. Uma colega também o tinha mencionado elogiosamente, em uma outra ocasião, lembro, mas nunca cumpri a promessa que fiz a mim mesmo de lê-lo. Mas naquele dia eu o li, meio distraído no começo, depois, avidamente. Terminada a leitura, deitei-me na

cama e pensei, pensei, até ter dor de cabeça, sob o olhar de compaixão de minha esposa. Recomecei tudo do zero, adeus gerativismo.

Essa dissertação também é culpa daquela leitura, da inquietação que ela gerou em mim. Isso, uma inquietação, um processo de mudança. Hoje vejo que as leituras gerativistas, de certa forma, me remetiam a uma planície de certa tranqüilidade, a um estado de conhecedor das regras da linguagem. O equilíbrio imbuído na noção de competência lingüística, propositalmente alheia aos gaguejos e erros da vida em sala de aula, me agradou muito por algum tempo. Tive, durante aquele período, a impressão de que uma língua era aprendida pulando-se de um estado de estabilidade para outro, rumo ao nirvana. Ao ler aquele artigo da Larsen-Freeman, contudo, me lembrei dos meus alunos, das variações que eles produziam em suas falas de uma aula para outra, das instabilidades das suas linguagens. Pensei em mundo mental inquieto, pulsante, caótico, em bilhões de descargas elétricas cruzando redes neurais de complexidade espantosa. Resolvi, então, valer-me de uma ferramenta relativamente nova à área de aquisição de L2, o conexionismo, em uma tentativa de desenvolver um trabalho sob a ótica dos sistemas não-lineares, os quais, creio, têm muito a dizer sobre a aquisição de L2.

Este trabalho se propõe a pensar os sistemas de interlíngua (IL) dos aprendizes de L2 como sendo “dinâmicos, complexos, não-lineares, caóticos, imprevisíveis, sensíveis às condições iniciais, abertos, auto-organizáveis, sensíveis ao *feedback* e adaptáveis” (Larsen-Freeman 1997, p. 142), ou seja, fascinantes de se estudar. Em aquisição de L2 se sabe que há uma vasta gama de fatores que interagem entre si e determinam o percurso de desenvolvimento da IL, desde a influência da L1, o tipo de *input* e a marcação das estruturas à idade, os fatores sócio-psicológicos e os motivacionais, só para citar alguns. A utilização, neste trabalho, da ferramenta fundamental do conexionismo, a rede neural, é uma tentativa de respeitar a natureza múltipla que, acredito, seja inerente ao processo de aquisição de L2.

Mais especificamente, o presente trabalho teve por objetivo investigar, sob o ponto de vista da sintaxe, a produção de perguntas em inglês por falantes brasileiros

jovens em um ambiente instrucional formal. Escolhi esse tipo de ambiente por ele ser meu *habitat* natural, por ter sido o berço de minhas inquietações em relação ao processo do aprendizado. Ainda cruzo pelos corredores, quase que diariamente, com muitos dos meus valorosos sujeitos. Além disso, creio ser a sala de aula o grande desafio da aquisição de L2. Estou ciente, contudo, do quanto a vinculação da pesquisa em L2 com a prática pedagógica e as cobranças que essa combinação enseja podem, por vezes, atrapalhar uma frente de estudos enquanto empresa científica, mas não necessariamente tecnológica. Quanto ao fenômeno investigado, o ordenamento sintático verbo / sujeito na produção de perguntas em L2, ele foi escolhido por causa da dificuldade geral de posposição do sintagma verbal que sempre observei em meus alunos. Como veremos no capítulo teórico, a sintaxe é uma importante frente de investigação conexionista em sua queda de braço com o gerativismo.

A escolha de um fenômeno isolado como foco de investigação, nomeadamente a formação de perguntas em L2, contudo, não é de todo coerente com a visão sistêmica de linguagem proposta pelo próprio conexionismo. Sob a ótica dos sistemas complexos, muito provavelmente nenhum fator lingüístico é, em si só, determinante no processo de aquisição de L2. O estudo de qualquer um desses fatores em isolamento faz perder, em certa medida, a visão do todo. Por outro lado, estudar o todo em aquisição de L2 é, sem dúvida, algo muito mais fácil de se dizer do que se fazer. Minha opção foi pelo recorte, sim, pela investigação dentro do domínio da sintaxe das perguntas em L2, mas também optei por me valer da abordagem conexionista justamente para dar ênfase aos aspectos sistêmicos das subpartes desse domínio. Creio que estudar as partes do todo através de um instrumento que, pelo menos teoricamente, possa vir a ser capaz de progressivamente integrá-las e, quiçá, finalmente revelar a natureza do todo que as transcende, seja uma opção razoável. É justamente a possibilidade de interconexão que faz das redes neurais um instrumento de investigação tão fascinante. A escolha deste trabalho mistura *langue e parole*, abdica dos algoritmos elegantes do gerativismo em prol do processamento paralelo de uma rede neural, no qual tudo acontece ao mesmo tempo, com fatores se fundindo na produção de um todo que é maior do que a soma das partes. Uma escolha, com os prós e contras que essa palavra traz em si mesma.

Minha principal fonte de consulta e inspiração foi o trabalho de pesquisadores como Jeffrey L. Elman, Nick C. Ellis e Kim Plunkett, cujas criteriosas simulações conexionistas ajudaram a consolidar o uso da rede neural como recurso legítimo e promissor na investigação do fenômeno da aquisição da linguagem. Como veremos no capítulo teórico, as conquistas desses e de outros pesquisadores conexionistas lançaram as bases do *proficuo embate de idéias que hoje se desenrola no estudo da aquisição da linguagem*, em cujo escopo se insere essa dissertação. Nele fui convidado a entrar através das excelentes aulas e conversas informais no mestrado da UFRGS, com seus professores e alunos instigantes e competentes, sem os quais nada haveria sido.

Esta dissertação se divide, basicamente, em quatro partes principais. A primeira é o capítulo teórico, no qual serão explicados os fundamentos do conexionismo, mencionados seus principais críticos e descritas as principais tentativas de aplicá-lo à aquisição de L2. O segundo capítulo descreve o experimento realizado com aprendizes de inglês como L2, divididos em faixas de proficiência, que produziram perguntas escritas em L2. O Capítulo 3 apresenta os detalhes relativos à simulação conexionista, elaborada com o objetivo de ser um paralelo à coleta realizada junto aos sujeitos. Os resultados obtidos em ambas as etapas (o experimento e a simulação) são discutidos no Capítulo 4. Todas as traduções aqui realizadas, salvo se indicado o contrário, são de minha autoria e responsabilidade.

Espero que a leitura deste trabalho seja uma experiência prazerosa e elucidativa para os interessados na investigação de fenômenos relacionados à aquisição de L2 sob uma perspectiva diferente das geralmente utilizadas na literatura. Acredito que os leitores que apresentarem a resistência necessária aos por vezes enigmáticos modos de operar das redes neurais serão recompensados com a possibilidade de vislumbrar novos caminhos a serem desbravados em aquisição de L2. Como salienta Gleick (1987), ao comentar a visão de alguns físicos em relação à teoria do caos, pode valer a pena abordar a aquisição de L2 como “uma ciência do processo, não do estado, do tornar-se, não do ser”(p.5).

Como afirmei no início desta introdução, esta dissertação nasceu de perguntas não resolvidas, cuja formação minha experiência em sala de aula ensinou. Aos leitores das páginas que se seguem, um comentário final: minhas perguntas continuam sem resposta. Gosto de crer que elas estão, contudo, melhor elaboradas, e espero que elas possam auxiliar na formação de perguntas ainda melhores nos leitores deste trabalho.

CAPÍTULO 1 - FUNDAMENTOS TEÓRICOS

O conexionismo é uma abordagem bastante recente no estudo da aquisição da linguagem, atualmente a disputar seu espaço com o gerativismo. Na revisão teórica que se segue, essa condição ficará evidente, tanto pela relativa novidade da utilização das redes neurais como instrumento fundamental na análise dos fenômenos relativos à aquisição, como pela freqüente menção ao gerativismo como parâmetro de referência (geralmente antagônica) na definição dos pressupostos conexionistas.

O paradigma conexionista questiona pressupostos inatistas; um questionamento que, à primeira vista, pode parecer (e por vezes de fato é) uma negação dos fundamentos do gerativismo. Deve-se ter em mente, contudo, que o momento histórico conexionista é o do esticar do pêndulo para o lado oposto, de um rompimento com uma maneira de se pensar o modo como os seres humanos adquirem a linguagem. Um olhar mais informado à totalidade da proposta conexionista, contudo, deixa claro que não se trata exatamente de uma mudança radical de perspectiva, mas sim de mais uma oportunidade de, como já aparece na literatura (Pinker 1998, Prince & Smolensky 1997, Smolensky 1999, Steedman 1999), haver uma reacomodação de modelos e maneiras de se conceber a aquisição da linguagem, para que a entendamos de modo mais amplo. Como salientam Christiansen & Chater (1999):

“Se modelos conexionistas realistas do processamento da linguagem puderem ser construídos, então poderá haver a possibilidade de um repensar radical não somente em relação

à natureza desse processamento, mas também da própria estrutura da linguagem. Talvez a melhor descrição da linguagem resida na estrutura de redes complexas, podendo ser apenas aproximada através das regras da gramática. Ou talvez os métodos de aprendizado conexionistas não se desenvolverão, e o conexionismo só poderá ter sucesso na reimplantação dos modelos simbólicos tradicionais” (p. 433).

No que respeita à aquisição de L2, a empresa conexionista encontra-se em estágio ainda mais incipiente. Como conseqüência desse fato, há pouquíssima literatura a esse respeito. Neste capítulo serão revisados alguns estudos importantes que tratam do assunto. Antes disso, porém, será necessária uma revisão geral da proposta conexionista para a aquisição da linguagem (L1, via de regra), para que se possa *situar o presente estudo dentro dos parâmetros de um debate mais amplo, que chega bastante intenso ao campo da aquisição de L2.*

1.1 UM BREVE HISTÓRICO DO CONEXIONISMO

O conexionismo traz a máquina computacional neural como metáfora base do *pensar-se a aquisição da linguagem. Impulsiona a empresa conexionista o credo de que essa máquina amplia nossa capacidade qualitativa de análise por via quantitativa, dado o assombroso volume de fatores que ela pode levar em consideração simultaneamente e dadas as complexas relações que esses fatores estabelecem entre si.* Como salientam Christiansen & Chater (1999), historicamente as teorias do pensamento humano têm inspirado as tentativas de construção de máquinas de processamento de informação. Atualmente, o paradigma dominante nessa área baseia-se na visão de que a cognição humana tem seus fundamentos no *processamento de símbolos. Essa tradição tem suas origens na sugestão de Boole, ainda no século XIX, de que a teoria da lógica e da probabilidade descreveria as “Leis do Pensamento” e de que o raciocínio consoante a elas poderia ser executado através regras simbólicas.* Nos anos trinta, Turing já argumentava que todo e qualquer pensamento poderia ser demonstrado através de operações simbólicas (Máquina de Turing). Seguiu-se, daí, o moderno computador digital idealizado por

von Neuman, a moderna ciência da computação, a inteligência artificial, o gerativismo e a ciência cognitiva simbólica.

O conexionismo possui uma origem distinta, pois se inspira no processamento do cérebro humano. Nos anos quarenta, McCulloch e Pitts criaram uma incipiente, mas influente, idealização do funcionamento neuronal. Nas décadas de cinquenta e sessenta, alguns pesquisadores construíram sistemas computacionais com base em idealizações semelhantes. Esses modelos apresentavam características interessantes, pois podiam aprender a partir da experiência, ao invés de serem previamente concebidos, o que fazia com que fossem tidos como modelos bastante plausíveis da cognição humana. Até então, as abordagens simbólica e conexionista apresentavam-se como modelos alternativos de representação da inteligência humana, sem que houvesse uma predominância clara de nenhum deles. Na década de sessenta, contudo, Chomsky desafiou a noção conexionista de aprendizado lingüístico através da defesa do inatismo, em seu elegante modelo teórico, o que, conjugado ao pouco poder computacional das máquinas de então, causou o abandono do conexionismo e o conseqüente estabelecimento do *paradigma simbólico como dominante*.

O recente aumento do poder dos computadores e os avanços dos estudos do cérebro causaram o *renascimento do conexionismo como alternativa ao modelo computacional simbólico da cognição humana*. Propondo-se a reavaliar a questão do inatismo e do paradigma simbólico da aquisição da linguagem, o conexionismo lança mão de novos avanços na neurologia e na ciência da computação com o intuito de explicar comportamentos para os quais tradicionalmente foram buscadas explicações dentro do paradigma simbólico. Os pesquisadores do conexionismo buscam um modelo unificado de cognição, que permita o abandono da distinção chomskiana entre competência e desempenho. Essa nova metáfora da faculdade humana da *linguagem tem questionado pilares do pensamento inatista como, por exemplo, a questão em torno da pobreza do estímulo, um dos principais argumentos em favor dos mecanismos inatos que justificariam a aquisição das complexas estruturas das línguas por parte dos seres humanos a partir de um estímulo degenerado*. As abordagens simbólicas descrevem um tal sistema mental através de símbolos e regras

definidas. Já o conexionismo, como veremos em seguida, minimiza (mas não elimina) o aparato inato, priorizando a capacidade humana de desvendar sistemas complexos através do estímulo lingüístico que recebemos do ambiente, degenerado, mas suficiente.

Para que se entenda melhor a proposta conexionista, é necessário que se compreenda o funcionamento de sua metáfora básica, a rede neural (tema da próxima seção). Afastando-se do behaviorismo, como salientam Elman *et alii* (1996), o conexionismo não prega uma volta à concepção da criança como *tabula rasa*, pois as redes conexionistas constroem representações internas capazes de generalização e criação. Não há dúvida, contudo, que o estímulo recebido do ambiente volta ao centro do palco, mas agora em oposição a um aparelho criador inato que talvez tenha, com o passar do tempo, se tornado poderoso demais.

1.2 AS REDES NEURAIS

Redes neurais artificiais são redes computacionais cujo modelo matemático de funcionamento se inspira na composição do cérebro do ser humano, organismo inteligente que tem a capacidade de aprender a partir da experiência. Nosso cérebro é formado pelos neurônios: células compostas por dendrites, que são, basicamente, um conjunto de terminais de entrada, por um corpo central e pelos axônios, que são longos terminais de saída (Figura 1). A comunicação entre essas células se dá através de regiões de contato, denominadas sinapses, por onde passam impulsos nervosos. Os impulsos recebidos por um neurônio X são constantemente processados e, ao atingir um dado limiar de ação, fazem com que ele dispare. Esse disparo produz uma substância neurotransmissora que flui do corpo do neurônio para os axônios, e daí para outros neurônios. Essa substância pode diminuir ou aumentar a polaridade da membrana pós-sináptica, inibindo ou excitando a geração dos pulsos de X para os outros neurônios, dependendo de fatores como a geometria da sinapse e o tipo de substância neurotransmissora (Tatibana & Kaetsu).

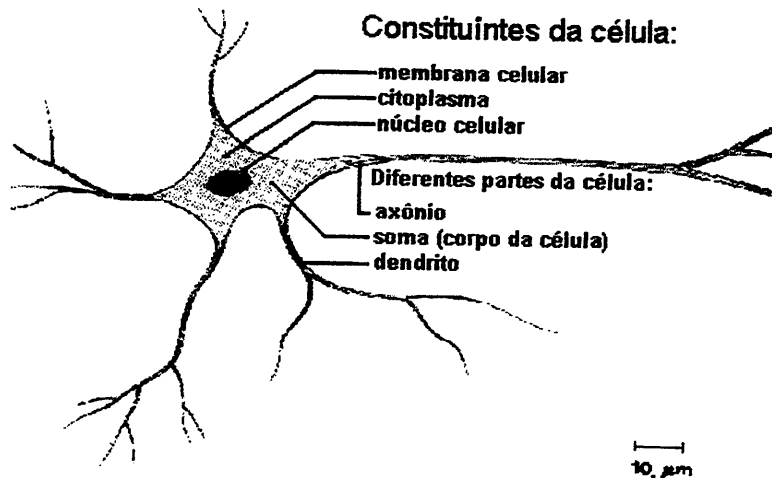


Figura 1: Constituintes da Célula Nervosa

Uma rede neural artificial consiste tipicamente de um conjunto de elementos densamente interconectados (obviamente não tão densa e nem tão complexamente conectados como os neurônios do cérebro humano), sendo esses elementos denominados unidades ou nódulos (Figura 2). A estrutura de uma unidade neural tem como objetivo imitar o funcionamento básico do neurônio, representado por suas unidades, e se dá, basicamente, da seguinte maneira:

- Sinais são apresentados à entrada da unidade;
- Cada sinal é multiplicado por um número, ou peso, que indica a sua influência na saída da unidade;
- É realizada a soma ponderada dos sinais, o que produz um nível de atividade;
- Se esse nível de atividade exceder um certo limite (relativo ao tipo de função de ativação da unidade em questão), a unidade produzirá uma determinada resposta de saída.

Tantos as entradas como as saídas são representadas como grandezas vetoriais, da seguinte forma:

- Vetor de Entrada: $X = (x_1, x_2, \dots, x_n)$
- Vetor de Peso: $W = (w_1, w_2, \dots, w_n)$
- Soma Ponderada dos Sinais: $A = \sum(w_i \cdot x_i) = W \cdot X$

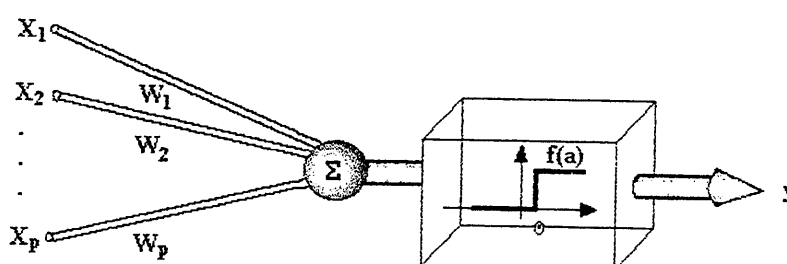


Figura 2: Unidade Neural Artificial

Uma função ativação (f) determina o valor de saída (Y) de uma unidade (U) a partir de A :

$$Y_u = f(A)$$

Algumas das funções mais utilizadas nas redes são a *hard limiter*, a *threshold logic* e a *sigmoid* (Figura 3).

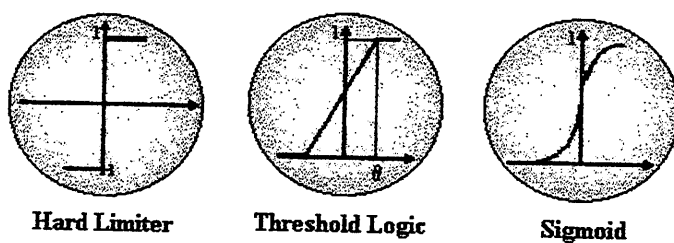


Figura 3: Tipos de funções de ativação

As unidades de uma rede neural são conectadas em camadas. Essas camadas são, nas redes do tipo MP (*Multilayer Perceptron*: redes de mais de uma camada), de três tipos (Figura 4):

- Camada de Entrada: onde os padrões vetoriais são apresentados à rede;
- Camadas Intermediárias ou Escondidas: onde é realizado, através das conexões ponderadas, a maior parte do processamento. Essas camadas podem ser consideradas extratoras de características das entradas;
- Camada de Saída: onde o resultado final é concluído e apresentado.

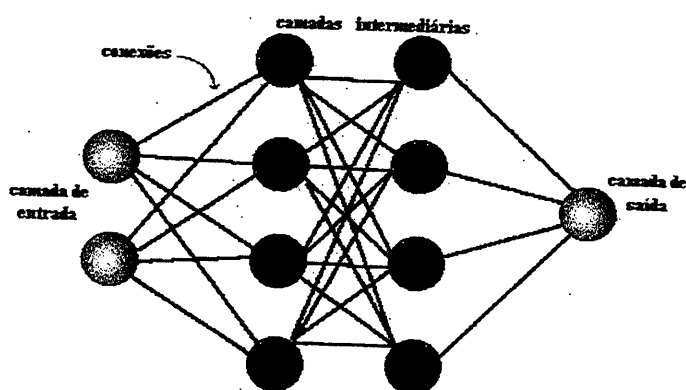


Figura 4: Rede Neural do Tipo MP

Uma das características mais importantes das redes neurais é o fato de elas serem capazes de aprender a partir do estímulo a elas apresentado. Esse aprendizado pode se dar de três maneiras:

- Supervisionado: quando é utilizado um agente externo que indica à rede a resposta desejada para o padrão de entrada;
- Não Supervisionado (auto-organização): quando não existe um agente externo indicando a resposta desejada para os padrões de entrada;

- Reforço: quando um crítico externo avalia a resposta fornecida pela rede.

Para que aprenda, a rede é submetida a um regime de treinamento; nele, os pesos das conexões da rede são modificados através de um algoritmo, visando a melhorar o desempenho da rede. A cada ciclo, todos os N pares do conjunto de treinamento (entradas e saídas) lhe são apresentados. O ajuste dos pesos das conexões da rede pode ser feito, via de regra, de duas maneiras:

- Modo Padrão: A correção dos pesos acontece a cada apresentação à rede de um exemplo do conjunto de treinamento. Cada correção de pesos baseia-se somente no erro do exemplo apresentado naquela iteração. Assim, em cada ciclo (apresentação de todos os padrões do conjunto de entrada) ocorrem N correções;

- Modo Batch: Apenas uma correção é feita por ciclo. Todos os exemplos do conjunto de treinamento são apresentados à rede, seu erro médio é calculado e a partir desse erro fazem-se as correções dos pesos.

Um algoritmo de aprendizado é um conjunto de regras bem definidas para a solução de um problema de aprendizado. Existem muitos tipos de algoritmos específicos para determinados modelos de redes neurais, que diferem entre si principalmente pelo modo como os pesos são modificados. Entre os mais utilizados para as modificações está a regra delta; nela, o ajuste (Δw_{ij}) do valor do peso da conexão entre uma unidade emissora (U_i) e uma receptora (U_j) é calculado da seguinte maneira:

$$\Delta w_{ij} = \eta \cdot \delta_i \cdot y_j$$

Onde: (η) é de taxa de aprendizado; $\delta_i = f'(A) \cdot (t_j - y_j)$ é o erro; (f') é a derivada primeira da função de ativação $f(A)$; (t_j) é a saída desejada e (y_j) é a saída real.

Esse algoritmo, porém, só pode ser utilizado em redes de apenas duas camadas. Por essa razão, as redes MP utilizadas nas simulações conexionistas usam o

backpropagation (propagação reversa), também denominado regra delta generalizada. Esse algoritmo foi desenvolvido por Rumelhart, Hinton & Williams (1986) para redes de camadas múltiplas. Nele, o aprendizado se dá, na primeira iteração, pela comparação, na camada de saída, do vetor de saída real com o vetor de saída esperado. Caso essa diferença seja zero, não ocorre nenhuma mudança nos pesos das conexões. Caso ela não seja zero, o erro é calculado através da regra delta e propagado reversamente ao longo da rede, causando um ajuste geral das conexões.

A lógica fundamental desses algoritmos, portanto, é a de ajustar os pesos das conexões da rede neural de modo a minimizar a diferença entre a saída real e a saída esperada. Durante um regime de treinamento supervisionado, tipicamente são atribuídos, às conexões da rede, pesos iniciais aleatórios de distribuição uniforme, com média de 0.0 e com limites inferiores e superiores de $+e - 1$ (Plunkett & Elman 1997). Quando o primeiro padrão vetorial de entrada é apresentado à rede, ela produz uma determinada saída vetorial. Calcula-se, então, a diferença entre a saída obtida e a desejada, realizando-se os devidos ajustes de peso com vistas a reduzir esse erro. Esse procedimento é então repetido, com uma velocidade definida pela taxa de aprendizado (η) do algoritmo, até que se tenha um erro (δ) suficientemente pequeno.

Elman *et alii* (1996) e Plunkett & Elman (1997) apresentam em detalhe o mecanismo de aprendizado das redes neurais. Ao ser treinada para resolver uma dada tarefa de aprendizado, uma rede neural desenvolve um complexo sistema de representações internas em suas unidades intermediárias. São essas representações internas que a habilitam tanto a responder adequadamente às entradas para as quais foi treinada como a serem capazes de generalizar em relação a novas entradas. Mais importante do que o fato de uma rede conseguir obter sucesso em seu aprendizado é a maneira pela qual tal sucesso é obtido. Uma rede pode, por exemplo, ser confrontada com um conjunto de entradas simples e memorizar as respostas adequadas ao invés de desenvolver representações internas com poder de generalização, isto é, só funcionar para o conjunto de entrada utilizado em seu treinamento.

Para se ter acesso às configurações internas de uma rede treinada, deve-se utilizar algumas ferramentas de análise que permitam ver o que de fato aconteceu durante seu processo de aprendizado. Plunkett & Elman (1997) fazem um levantamento das ferramentas de análise mais comuns:

- Erro: cálculo do nível de erro, por número de ciclos de treinamento, das saídas produzidas pela rede em relação às saídas esperadas. Pode-se calcular tanto o erro global como o individual de alguns dos padrões de entrada. Ao optar-se pelo primeiro, contudo, deve-se ter em mente que muitas vezes o valor global do erro oculta diferenças importantes no que tange à variação entre padrões distintos de entrada;
- Análise de Agrupamento (*Cluster Analysis - CA*): Após o treinamento, apresentam-se alguns vetores de teste à rede. Esses padrões produzem ativações nas unidades intermediárias, que são gravadas e rotuladas. Os padrões das unidades intermediárias são vetores em um espaço multidimensional (cada unidade intermediária representa uma dimensão), e a CA mostra níveis de similaridade entre esses vetores, calculados com base na distância euclidiana entre eles. Os agrupamentos são feitos de modo hierárquico, na forma de um diagrama arbóreo;
- Análise de Componentes Principais (*Principal Component Analysis - PCA*): Essa ferramenta de análise permite visualizar seções de interesse do espaço vetorial, pois se pode mover a perspectiva de visualização ao longo de vários planos diferentes. A PCA é um procedimento que permite descobrir eixos alternativos para a representação dos vetores sendo analisados, de modo que as variações em distribuição espacial possam ser mais bem compreendidas. Aqui, também, cada unidade intermediária da rede corresponde a uma dimensão do espaço vetorial (um componente principal). A PCA organiza esses componentes em ordem decrescente de variação explicada, ou seja, permite que se saiba onde houve concentração de representações internas;

- Diagramas de Hinton: os pesos das conexões da rede são mostrados na forma de quadros coloridos; o tamanho e a cor de um quadro representam, respectivamente, a magnitude e o sinal da conexão: excitatório (+) ou inibitório (-). Essa ferramenta permite que se veja os pesos das conexões da rede, o que serve de indicação de como as transformações nas conexões estão acontecendo.

Há vários tipos de configurações possíveis de redes neurais. Essas configurações, conjugadas a algoritmos como os vistos acima, são capazes de resolver uma gama de problemas de aprendizado, sejam eles lingüísticos ou não. Para os propósitos deste trabalho, é essencial que se compreenda o que de fato representa uma rede treinada, ou seja, que tipo de explicação ela pode nos fornecer quanto aos problemas que consegue resolver. Ao término de um período de treinamento, uma rede neural adaptou seus parâmetros internos, a saber, as ativações de suas unidades e os pesos das conexões entre elas, de modo a produzir a melhor combinação possível desses parâmetros para fornecer a melhor resposta possível (i.e., a mais próxima possível da resposta esperada) para as entradas que lhe são fornecidas durante o treinamento. Isso é alcançado porque a rede condensa as características das entradas que recebe, tira delas sua essência, construindo uma lógica interna travestida em ativações e pesos. Essa lógica interna responderá, via de regra, previsivelmente em relação ao que já é conhecido (i.e., ao que fez parte de seu conjunto de treinamento) e construirá generalizações em relação a dados novos. Um sistema desse tipo representa uma metáfora muito interessante tanto em relação ao cérebro humano, pois possui elementos de inspiração biológica, a saber, unidades e conexões, como em relação à mente, com seu algoritmo de aprendizado e suas representações internas maleáveis. Veremos ao longo deste capítulo que as representações internas das redes estão sendo utilizadas como base para a construção de gramáticas de base conexionista, como no caso da recente Teoria da Otimização.

Analisaremos, na próxima seção, de que forma o conexionismo se vale das redes neurais para tentar resolver problemas da aquisição da linguagem.

1.3 A ABORDAGEM CONEXIONISTA DA AQUISIÇÃO DA LINGUAGEM

Foi salientado acima o fato de a abordagem conexionista ter reformulado a maneira de muitos estudiosos pensarem a aquisição da linguagem, i.e., o modo como muitos cientistas cognitivos vêem o aprendizado e as representações e processos mentais relativos à faculdade humana da linguagem. Como decorrência desse fato, o inatismo e seus fundamentos simbólicos têm sido bastante questionados pelos advogados dessa nova visão, empenhados na procura de espaço na arena da discussão teórica. Gasser (1990) nos fornece um retrato bastante preciso de como o conexionismo se insere na atual discussão em torno dos méritos e deméritos do inatismo como explicação plausível da aquisição da linguagem:

“Apesar do apelo poderoso dos símbolos, das regras e da lógica, as visões tradicionais possuem uma fragilidade muito pouco humana. Entidades conceituais e lingüísticas são relacionadas a categorias na base do tudo ou nada, regras tipicamente se aplicam a seqüências fixas, e lida-se precariamente com desvios em relação a padrões esperados. Nos modelos conexionistas a fragilidade é evitada porque não há símbolos discretos e regras como tais; as entidades que um sistema conexionista utiliza para caracterizar o mundo são padrões fluidos de ativação ao longo de partes de uma rede. Além disso, o conexionismo recupera a ênfase no aprendizado na ciência cognitiva. Nos modelos simbólicos geralmente se pressupõe que é suficiente caracterizar-se um ponto em particular no processo de aquisição. A maioria dos conexionistas não concorda com isso: o mais interessante é como o sistema progride de um estágio para outro, e os conexionistas têm desenvolvido vários algoritmos de aprendizado novos em rede para serem estudados e aplicados a certos domínios-problema” (p. 179).

Esse trecho deixa claro que o conexionismo apresenta, portanto, uma alternativa ao paradigma simbólico, visão dominante até o presente nos estudos em aquisição da linguagem. Como ressaltamos acima, a maioria dos pressupostos conexionistas são apresentados e definidos, na literatura, em relação ou em oposição ao gerativismo. Não obstante, Gasser (1990) ressalta que, no que tange às estruturas

consideradas inatas pelos modelos gerativos tradicionais, o conexionismo não as nega totalmente, pois a própria arquitetura da rede tipifica algo biologicamente inato, mas propõe uma reanálise do escopo de ação dessas estruturas, tendendo a minimizá-lo. Conexionistas geralmente acreditam, por exemplo, na interacionalidade, e não na modularidade, como fator principal na aquisição do conhecimento lingüístico, pois *ênfatizam o empiricismo em detrimento do inatismo*.

Elman *et alii* (1996) listam as características principais da abordagem conexionista:

“Há que se distinguir os mecanismos do inatismo de seu conteúdo. Não é necessário que haja um mapeamento direto entre eles. O conexionismo sugere que, para os comportamentos cognitivos de ordem superior, a maioria dos resultados específicos a certos domínios são provavelmente obtidos através de meios independentes a esses domínios;

“A relação entre mecanismos e comportamentos é frequentemente não-linear. Efeitos drásticos podem ser causados por pequenas mudanças;

“O que aparenta ser um evento ou comportamento isolado pode ter várias causas subjacentes, algumas delas distantes no tempo;

“O conhecimento, em última análise, refere-se a um padrão específico de conexões sinápticas no cérebro. Nesse sentido específico, o conexionismo alega não existir, provavelmente, conhecimento de ordem superior que seja inato;

“O processo de desenvolvimento em si está no âmago da aquisição do conhecimento. Alguns comportamentos complexos podem não ser adquiríveis salvo se através de um percurso desenvolvimentista [veja também Elman 1993];

“O conexionismo nos fornece um aparato conceitual útil para que entendamos as formas emergentes e as interações de restrições em níveis múltiplos, mas definitivamente não se deve pensar que ele seja um empiricismo radical. O conexionismo ainda está em sua infância, e o desenvolvimento fornece uma rica gama de fenômenos que desafiam a atual tecnologia;

“O desenvolvimento é um processo de emergência. O conexionismo nos provê um meio conceitual para o estudo das condições sob as quais as formas emergentes surgem e como essas formas podem ser restringidas” (p. 359).

Os pesquisadores que trabalham dentro da visão simbólica defendem a possibilidade da distinção entre *hardware* e *software*, ou seja, que as atividades mentais relativas à linguagem podem ser descritas sem que haja a necessidade de se fazer referência a estruturas ou processos neurais. A cognição, dentro dessa visão, consiste fundamentalmente da aplicação de regras. Comportamentos não regrados, i.e., as exceções, são explicados através da ação da memória. Existe, portanto, um mecanismo dual para dar conta, por exemplo, da aquisição do frequentemente citado sistema de verbos regulares e irregulares do inglês (Pinker 1998).

Segundo Gasser (1990), um sistema regido por regras requer três elementos básicos: “... (a) um programa central que dirija o sistema durante os processos nele envolvidos, sejam eles a compreensão, a produção ou a ‘derivação’ (sic); (b) um mecanismo sofisticado de identificação de padrões visando à seleção de regras apropriadas a dados contextos; (c) *símbolos*, ou seja, sinais que denotem outros sinais ou estruturas plenamente desenvolvidas na memória” (p. 184).

Os pesquisadores conexionistas, como já foi visto, rejeitam algumas das premissas básicas da ciência cognitiva simbólica, principalmente no que concerne à que considera o comportamento neural como irrelevante à explicação da cognição. Eles acreditam que o cérebro utiliza-se de vastas quantidades de processamento paralelo, o que faz com que modelos de inspiração neural tenham mais poder explicativo do que os de base simbólica. Para os conexionistas, portanto, “... regras e símbolos são fenômenos emergentes, i.e., surgem a partir de interações complexas com elementos e processos primitivos” (Gasser 1990, p. 184). Essa nova maneira de se pensar os processos cognitivos humanos, entre eles a aquisição de L2, é comumente definido como o *paradigma sub-simbólico*.

Os estudiosos do conexionismo alegam que as explicações do paradigma simbólico são aproximações descritivas úteis. Um dos exemplos mais contundentes em favor dessa forte alegação é o sucesso na simulação da curva de aprendizado do passado em inglês (Rumelhart & McClelland 1987, Plunkett & Marchman 1991). Os modelos utilizados nessas simulações foram redes conexionistas simples capazes de gerar, através de treinamentos utilizando representações de *input* provável, tanto formas regulares como irregulares do passado em inglês, sem que lhes fosse fornecida nenhuma informação referente a quaisquer regras. Ou seja, as redes apresentaram um comportamento regrado, mas não gerado por regras, assim como produziram irregularidades com sucesso, o que contraria a necessidade do mecanismo dual regras / memória do inatismo. As regras aparentes apresentam-se, nesse contexto, de forma distribuída ao longo das associações entre as unidades da rede, de modo que cada regra envolve muitos pesos de conexões, pois eles participam da formação de inúmeros padrões. De forma análoga, um mesmo peso participa de várias regras, i.e., participa da construção de vários padrões de resposta correlacionados. Esse tipo de sistema trabalha, portanto, com um modo distribuído de representação: vários processos ocorrem paralelamente na rede para que ela possa produzir suas saídas.

Gasser (1990) analisa, a partir dessas características, o que substitui os símbolos nos modelos conexionistas:

“O que ocupa o lugar dos símbolos dos modelos tradicionais? Não haveria a necessidade de existir um lugar ou um padrão invariável na memória para conceitos como PERNA, FRITAR e VELAR? A resposta é que em um sistema conexionista não se pode dizer que uma coisa é uma PERNA ou não é uma PERNA; ao contrário, as coisas são mais ou menos como pernas. Poderá existir um grupo de unidades que tenda a ser ativada mais fortemente quando uma perna é percebida ou imaginada, mas não seria possível traçar-se uma linha divisória nítida em torno desse grupo de unidades, e cada uma delas também participaria da representação de outros conceitos. Dessa forma, quando (mais ou menos) PERNA é ativado, um número indefinido de outros conceitos também são ativados parcialmente. Note que essa abordagem permite que a metáfora e a analogia sejam

tratadas como processos naturais, e não como fenômenos periféricos não passíveis de formalização” (p.185).

Há, portanto, nos sistemas conexionistas, tanto uma flexibilidade de resposta a estímulos externos, uma vez que a rede constrói suas representações, em larga medida, a partir do ambiente, como uma flexibilidade na delimitação de conceitos, pois eles encontram-se difusos ao longo de um sistema complexo de múltiplas conexões paralelas. Isso permite, como salienta o autor, que se trabalhe com aproximações que talvez tenham uma realidade mental maior do que as abordagens que lidam com a dicotomia pertence ou não-pertence a uma dada categoria conceitual.

O que está em jogo aqui é o próprio conceito de conceito. Como enfatiza Thornton (1996), *conceitos fazem o mundo mais simples e fácil de entender: não percebemos, lembramos ou falamos sobre cada objeto ou evento como se ele fosse único. Nos referimos a ele como exemplo de uma classe ou conceito sobre o qual já sabemos algo. Geralmente um conceito é implementado através de um mecanismo que divide um conjunto de dados em dois subconjuntos: os que estão dentro e os que estão fora do escopo desse conceito. A delimitação dessas duas áreas, ou seja, a implementação de um conceito, no conexionismo, é realizada através de detectores de facetas, tais como os mecanismos de agrupamento e de componência principal (seção 1.2). Esses mecanismos produzem altos níveis de saída para certas classes de entradas, podendo, portanto, ser entendidos como conceitos de natureza probabilística que incluem as classes em questão. Isso é feito através de um algoritmo de aprendizado que opera com base na semelhança, pois a premissa fundamental desse tipo de algoritmo é a de que exemplos que se encontram dentro do escopo do conceito serão mais similares entre si do que os que se encontram fora. Esse mecanismo de aprendizado tem como objetivo produzir conceitos que maximizam a semelhança intraclasses e minimizam a semelhança interclasses. Essa similaridade, no caso de uma rede conexionista, é avaliada através de distâncias em um espaço vetorial.*

Essa distinção conceitual básica nos conduz a outras distinções importantes para os modelos conexionistas: as que se estabelecem, dentro dos modelos cognitivos, entre exemplos (instâncias) e classes (tipos), e dentro do modelo lingüístico, entre gramática (geral) e léxico (específico), e processos regulares (gerais) e irregulares (específicos). Nos modelos conexionistas, essas distinções são uma questão de grau, não de tipo (Ellis & Schmidt 1998). Ou seja, um padrão de ativação pode tanto representar algo geral como específico, isso dependendo apenas do número de unidades envolvidas, ou seja, do quanto a entidade em pauta é especificada.

Especificamente no que tange à sintaxe, objeto deste estudo, o desafio conexionista está em estender a observação de Gasser (1990), de algo ser “mais ou menos uma perna”, para algo ser mais ou menos um sujeito ou um objeto, por exemplo, pois as restrições interpretativas da sintaxe são fortemente categóricas. Ou seja, modelos conexionistas bem sucedidos de processamento sintático como o de Elman (1993), que simulou o processamento de sentenças complexas envolvendo subordinação, concordância e estruturas argumentais, devem ser refinados. Uma das tentativas atuais mais promissoras de incrementar o modelo SRN (*Simple Recurrent Network* – Rede de Recorrência Simples) de Elman é o modelo VSG (*Visitation Set Gravitation* – Gravitação do Conjunto de Visitação) de Tabor, Juliano & Tanenhaus (1997), que desenvolve suas representações sintáticas através de um módulo gravitacional: um processador dinâmico que transforma as representações produzidas por uma rede SRN em hipóteses interpretativas.

Será realizada, na seção que se segue, uma análise mais específica de aspectos relativos aos modelos conexionistas.

1.4 OS MODELOS CONEXIONISTAS

Os modelos conexionistas que têm norteado a discussão deste trabalho são os chamados modelos de processamento distribuído, ou PDP (*Parallel Distributed Processing*), que têm sido os mais discutidos por suas características inovadoras. Há,

contudo, outros tipos de modelos conexionistas, como os de base localista (Gasser 1990), por exemplo, nos quais as unidades da rede representam conceitos específicos (*azul, João, animado, etc.*), que estão fora do escopo deste trabalho.

Como foi visto acima, nos modelos distribuídos os conceitos encontram-se espalhados por várias unidades, cada unidade participando da representação de vários conceitos. Seus componentes individuais são altamente interativos: o conhecimento de eventos, conceitos e linguagens é representado de maneira difusa, em consonância com o que se acredita ser a base do sistema cognitivo humano. Os modelos PDP possuem inspiração neural: o comportamento das unidades pode ser comparado com o dos neurônios (ativação = excitação), e as conexões entre as unidades lembram seus axônios e dendrites (Broeder & Plunkett 1994).

Christiansen & Chater (1999) salientam que, grosso modo, o cérebro humano é formado por um número elevado de unidades simples de processamento, os neurônios, que se encontram densamente interconectados de modo complexo. A visão conexionista de processamento de informações acredita que esse processamento se dê, dentro desse contexto, através de operações simultâneas e cooperativas dessas unidades simples. Essa visão também crê que os neurônios se comuniquem através de valores numéricos, representados pelas suas taxas de descarga elétrica, e não através de mensagens simbólicas. Os neurônios, portanto, realizam, sob essa ótica, o mapeamento de entradas numéricas, recebidas de outros neurônios, para saídas também numéricas, que por sua vez são transmitidas a outros neurônios. Os modelos conexionistas simulam tais arquitetura e propriedades, pois consistem de unidades de processamento densamente interconectadas de modo complexo, que funcionam simultânea e intrincadamente, nas quais as saídas são funções das entradas.

Obviamente os modelos PDP não são de fato realistas, pois consistem de drásticas simplificações, tanto da real arquitetura do cérebro como das complexidades de seu funcionamento. Isso deixa, para Christiansen & Chater (1999), duas opções claras de pesquisa: uma no sentido de aumentar o realismo biológico das

redes (o que significaria, pelo menos em tese, operar no âmbito dos bilhões de unidades) e a outra no sentido de simular o comportamento humano, utilizando dados oriundos da psicologia cognitiva, da lingüística e neuropsicologia cognitiva, ao invés de informações obtidas via neurobiologia. A segunda opção tem sido a mais freqüente no contexto da pesquisa conexionista, significando uma competição direta com as máquinas simbólicas de processamento de informação.

Christiansen & Chater (1999, p. 420) também listam os temas mais recorrentes no debate acerca da validade dos modelos conexionistas, que são as seguintes:

- a) Aprendizado: Os modelos conexionistas tipicamente aprendem a partir da experiência, enquanto que os modelos simbólicos de processamento da linguagem são tipicamente pré-especificados e não aprendem;
- b) Generalização: Poucos aspectos da linguagem são simples o suficiente para serem aprendidos através da memorização (*rote learning*). Ter a habilidade de generalizar é crucial para qualquer modelo conexionista;
- c) Representação: As redes constroem códigos internos, durante o aprendizado, para se adequar à tarefa que lhes é imposta. É crucial que mecanismos para a compreensão desses códigos sejam desenvolvidos, pois só assim se saberá o que a rede de fato fez. Além disso, o modo de codificação das entradas e saídas, uma escolha do pesquisador, afeta diretamente o desempenho da rede, além de pôr em cheque a validade da simulação;
- d) Regras Versus Exceções: Muitos aspectos da linguagem exibem semi-regularidades, i.e., regularidades que, no geral, funcionam, mas que admitem exceções. Os modelos simbólicos dão conta desse problema através da utilização de dois sistemas distintos: regras e listas de exceções (memorizadas). Ao contrário, muitos dos modelos conexionistas dão conta tanto das regularidades como das exceções através de um mecanismo único.

Uma premissa básica desses modelos, como salientam Elman *et alii* (1996), é a de que comportamentos complexos podem emergir a partir da interação entre constituintes simples e seu meio. Como já se salientou, outro fator importante em relação ao paradigma conexionista é sua inspiração neuronal. Computações em paralelo realizadas por unidades conectadas em uma rede de linhas balanceadas nos fornecem uma maneira de abstrairmos um sistema nervoso central. Os modelos conexionistas pretendem explicar o processamento cognitivo através desse tipo de computação, ou seja, através de estruturas de dados e dos processos que operam sobre esses dados, relacionando entradas e saídas. Ao contrário dos modelos gerativistas, essa linha de análise não traça nenhuma distinção fundamental entre competência e desempenho. O que de fato importa em tais modelos computacionais da linguagem é a compreensão, a produção e a representação, seja ela lingüística ou não, que possibilitam esses processos. O computador figura, nesse contexto, como uma ferramenta indispensável para testar-se a validade desses complexos modelos. Já vimos que modelos PDP, na verdade, não são novidade, pois os primeiros datam da década de quarenta. Contudo, os novos algoritmos de aprendizado e o relativo desgaste dos modelos simbólicos de cognição têm dado novo impulso ao estudo da aquisição da linguagem nesses moldes. Segundo Gasser (1990, p. 181), a maioria desses novos modelos computacionais têm as seguintes características gerais:

- A memória do sistema consiste de uma rede de unidades de processamento ligadas entre si através de conexões balanceadas. O peso de cada conexão equivale à quantidade que determina o quanto a unidade na entrada ativa ou inibe a unidade na saída;
- O comportamento das unidades imita, de maneira aproximada, o comportamento dos neurônios. Elas somam o insumo que recebem nas conexões e computam uma ativação, que é uma função do insumo total, e um valor de saída, que é uma função da ativação. O valor de saída de uma unidade é passado para outras unidades;

- Os efeitos de memória a aprendizado podem ser representados através de padrões de ativação e de ajuste de pesos;
- O processamento é paralelo, várias decisões podem ser tomadas ao mesmo tempo;
- O controle é distribuído, e não há regras a serem executadas.

Será descrita, na seção a seguir, a maneira como esses modelos aprendem.

1.5 O APRENDIZADO NAS REDES PDP

Já se ressaltou que os modelos conexionistas estão fortemente voltados aos processos (em contraste aos produtos) cognitivos, enfatizando sua sensibilidade aos contextos. Os modelos PDP são, segundo Gasser (1990, p. 181), os mais adaptados ao estudo do aprendizado lingüístico, sendo suas propriedades básicas as seguintes:

- Robustez e Degradação Suave: *Input* incompleto e errôneo não causa o colapso do sistema;
- Representações Graduadas: Os conceitos adquiridos pelo sistema não são categorias discretas. Elementos pertencem a categorias conexionistas em graus diversos; as representações evoluem continuamente ao longo do processo de aprendizado e os conceitos podem se imbricar livremente;
- Tamanho de Memória Fixo: Como o conhecimento é distribuído pelas conexões do sistema, a adição de conhecimento não necessita, necessariamente, de novas unidades e conexões;
- Generalização Automática e Comportamento Regrado: Ao aprenderem padrões específicos, os sistemas conexionistas constroem o conhecimento que os

habilita a lidar com padrões semelhantes. Eles realizam generalizações, mas elas não aparecem explicitamente na rede;

- Interação de Fontes Múltiplas de Conhecimento: Os sistemas conexionistas operam com base na ativação em paralelo de várias partes da rede, o que possibilita a investigação em domínios caracterizados por decisões nas quais várias fontes de conhecimento são acionadas simultaneamente.

O processo de aprendizado geralmente se dá, nas redes PDP, através da *associação de padrões de relação entrada / saída*. Cada saída representa um padrão de ativação relacionado a uma dada entrada. A rede constrói gradualmente representações internas sobre as regularidades das entradas; o treinamento da rede resulta em saídas (aproximadamente) corretas em relação às entradas fornecidas. Além disso, uma rede conexionista pode vir a gerar saídas apropriadas para entradas para as quais ela não foi treinada diretamente, através da generalização via representação interna de regularidades. Ou seja, as redes extraem as regularidades das entradas e executam generalizações baseadas nessas descobertas *a posteriori*, i.e., nenhuma informação nesse sentido é fornecida *a priori*. As categorias representadas na rede após o treinamento são, portanto, fruto de propriedades intrínsecas às entradas.

Elman (1993, p. 86-93) salienta quatro propriedades básicas do processo de aprendizado das redes PDP:

- A Estatística Como Base do Aprendizado: A mola propulsora da inferência nos algoritmos de aprendizado da maioria das redes neurais é a estatística. As redes PDP baseiam-se na representatividade de seus conjuntos de dados. Se eles forem muito reduzidos, a rede pode vir a não ser capaz de descobrir generalizações que caracterizam as populações maiores. É importante frisar, contudo, que a estatística não é o resultado do aprendizado, pois se o fosse o trabalho das redes PDP não representaria mais do que uma tabela de coocorrências, simplesmente reeditando credos behavioristas. As redes PDP são aproximadores de funções, não compiladores

ou tabelas de referência. Aprender significa descobrir a função que subjaz aos dados de treinamento. Apesar de o algoritmo de aprendizado ser altamente sensível à estatística, o resultado final está mais próximo das regras do gerativismo do que de uma tabela de referência. Em um estudo de morfossintaxe comparando o desempenho de uma rede PDP e seres humanos, Ellis & Schmidt (1998) concluem que, *sob uma perspectiva associativa, regularidade e freqüência são, em sua essência, o mesmo fenômeno, interagindo durante o processo de aprendizado. Segundo esses autores, a regularidade surge a partir da consistência exibida pelas associações mais freqüentes da língua;*

- A Representação da Experiência: Ao entrar em contato com um padrão durante o treinamento, a rede PDP ajusta, através de seu algoritmo de aprendizado, os pesos de suas conexões. Esses novos pesos representam a função aprendida pela rede até aquele momento. Uma vez processado, o padrão em si é descartado, deixando como legado apenas a atualização das representações da rede. Ou seja, o padrão em si não pode ser mais utilizado pelo algoritmo no processo de aprendizado;
- Restrições na Formulação de Novas Hipóteses: Ao contrário dos modelos gerativos, de base simbólica, nos quais o espaço criado por todas as hipóteses possíveis para um dado sistema corresponde a todas as regras (proposições simbólicas) consoantes à gramática desse sistema, nos modelos conexionistas as hipóteses são implementadas na forma dos valores dos pesos das conexões entre as unidades da rede. Isso remete a uma importante diferença na seqüência de formulações de hipóteses, pois sob a perspectiva gerativista uma hipótese pode ser descartada e sua sucessora diferir completamente dela, uma vez que não há restrições advindas do histórico temporal do aprendizado. O espaço de formulação de hipóteses nas redes PDP, no entanto, deve ser necessariamente contínuo, uma vez que essas redes utilizam o gradiente descendente para essa reformulação. O gradiente descendente é um expediente matemático utilizado pelo algoritmo de aprendizado que permite a exploração da superfície tri-dimensional do gráfico do erro de maneira eficiente, i.e., visa a determinar a combinação de pesos que resulta no erro mínimo. Essa descida ao ponto mínimo absoluto da superfície do gráfico do erro é de natureza

contínua, pois inicia com alguma hipótese escolhida aleatoriamente e só lhe são permitidas pequenas mudanças incrementais durante o percurso. Novas hipóteses têm que ter um certo grau de similaridade em relação às suas predecessoras. Esse mesmo processo pode, contudo, levar a investigação do algoritmo a um mínimo local, de onde ele não consiga sair;

- Mudança na Habilidade de Aprender ao Longo do Tempo: As redes PDP geralmente utilizam o *backpropagation* como algoritmo de aprendizado. É assim que a rede ajusta os pesos das suas conexões em resposta aos erros que são produzidos durante o treinamento. A função de ativação da maioria das redes PDP é a *sigmoid* (Figura 3), que possui uma série de propriedades importantes:

a) As entradas são condensadas, de modo que a saída sempre variará entre 0.0 e 1.0;

b) Uma entrada líquida de 0.0 significa uma ativação de 0.5 (ponto intermediário do escopo de ativação da unidade). Entradas positivas resultam em ativações superiores a 0.5, enquanto que entradas negativas resultam em ativações inferiores a 0.5;

c) A função ativação é monotônica (não apresenta subidas e descidas dentro de um mesmo intervalo), mas é não linear;

d) 0.0 é o nível de maior sensibilidade da função ativação *sigmoid*, havendo saturação em resposta a entradas de grande magnitude, sejam elas positivas ou negativas.

Essas características têm conseqüências interessantes para o processo de aprendizado das redes PDP. A inicialização do processo, por exemplo, é geralmente feita através do ajuste dos pesos com valores aleatórios próximos de zero, de maneira que a entrada líquida nas unidades da rede seja também próxima de zero. Isso coloca a rede em um ponto de sensibilidade máxima.

Como a função ativação das unidades da rede é geralmente a *sigmoid*, sua derivada (taxa de variação) é máxima em 0.0, caindo assintoticamente quando se aproxima dos extremos. Isso faz com que o ajuste dos pesos, que é calculado por:

$$\Delta w_{ij} = \eta \cdot \delta_i \cdot y_j, \text{ onde } \delta_i = f'(A) \cdot \sum_k \delta_k \cdot w_{ki}$$

tenda a ser menor quando a unidade começa a se saturar, pois δ_i tende a diminuir, uma vez que (f') representa a inclinação da função *sigmoid*.

A existência dessas características significa que, no início do treinamento, as unidades estão mais sensíveis às entradas, e os pesos podem ser modificados mais facilmente. Com o desenvolver do treinamento, as unidades aproximam-se de seus pontos de saturação, o que faz com que a rede assuma um comportamento mais categórico em relação às diferenças entre as entradas, pois a modificação dos pesos torna-se mais difícil. Ou seja, a *maleabilidade inicial da rede dá lugar a uma crescente rigidez*. Há, portanto, um processo de fossilização inerente ao processo de aprendizado. Quanto mais a rede sabe, mais difícil se torna para ela aprender algo novo.

Nota-se, então, que a maneira como as redes PDP aprendem está longe da perfeição. Elas tanto se comprometem com falsas hipóteses no começo do processo, e têm que se livrar delas de modo lento e gradual, correndo o risco de ficarem presas em um mínimo local, ao invés de atingirem o menor erro possível (mínimo absoluto), como também aprendem menos com o passar do tempo. Essas características aparentemente negativas, no entanto, fazem com que as redes PDP sejam tão interessantes do ponto de vista dos estudos da linguagem, pois simulam aspectos do aprendizado humano ao longo do tempo. Entre eles estão os efeitos, no aprendizado, da idade, do desenvolvimento da memória e da transferência interlingüística (Sokolik 1990, Elman 1993, Gasser 1990, respectivamente), só para citar alguns.

Recentemente, o conexionismo tem procurado um refinamento ainda maior de suas propostas no campo da aquisição da linguagem. Isso implica, necessariamente, em um melhor entendimento do funcionamento dos mecanismos internos das redes neurais. Um dos produtos desse esforço foi o desenvolvimento matemático do conceito de grau de harmonia de uma rede, utilizado em abordagens como a da Teoria da Otimização, assunto da próxima seção.

1.6 AS FUNÇÕES HARMÔNICAS

Para que o conexionismo possa formular postulados consistentes e precisos a respeito do nível de abstração e organização mental que uma teoria da gramática exige, é preciso que se estabeleçam relações entre as interações de nível inferior com o *input* e princípios cognitivos mais gerais, de nível superior. O credo fundamental desse tipo de tentativa é o de que descrições de nível superior da computação conexionista possam servir de base para uma teoria mais abrangente do funcionamento da mente humana, mais especificamente no que tange à faculdade da linguagem. Smolensky (1999) apresenta uma possibilidade desse tipo de casamento, denominando-a Aproximação Simbólica:

“No domínio da linguagem, os padrões de ativação que constituem as representações mentais admitem descrições de nível superior, abstratas, muito próximas dos tipos de estruturas discretas, abstratas, que servem de base à teoria lingüística simbólica” (p. 159).

Essa união está sendo realizada com inspiração na maneira como as ciências físicas relacionam seus sistemas de análise de nível superior e inferior: explica-se os princípios fundamentais da organização mental em nível gramatical a partir de conseqüências de larga escala de interações em um nível inferior.

Lembrando-se que nos modelos PDP tipicamente:

- Representações mentais são padrões distribuídos de atividade numérica;

- Processos mentais são transformações maciças em paralelo de padrões de atividade por padrões de conexões numéricas;
- A aquisição do conhecimento é o resultado da interação de características arquitetônicas e regras de aprendizado inatas (i.e., inerentes ao modelo) e a modificação dos pesos das conexões através da experiência.

tem-se aí a descrição de sistemas que se caracterizam pela complexidade e não-linearidade, para os quais as ferramentas de análise ainda são escassas. Uma dessas raras ferramentas são as funções de Lyapunov. Essas funções atribuem um valor numérico a cada estado global possível de um dado sistema dinâmico de modo que, com a mudança de estado do sistema ao longo do tempo, esse valor aumenta. Uma das denominações comumente citadas na literatura para esse tipo de função é Função de Energia. No caso do conexionismo lingüístico, ela tem sido chamada de Função Harmônica (Prince & Smolensky 1997). A harmonia (H) de um dado padrão de entrada mede seu grau de conformidade com as restrições implícitas às conexões da rede. Ou seja, a harmonia do padrão medirá sua boa-formação em relação ao sistema de restrições intrínseco aos pesos das conexões do modelo conexionista em questão.

A função harmônica básica de uma rede conexionista é dada por:

$$H = \sum_{ij} a_i \cdot w_{ij} \cdot a_j$$

onde a_i é o valor de ativação da unidade i , a_j é o valor da ativação da unidade j e w_{ij} é o valor do peso da conexão entre essas unidades.

Entre duas unidades quaisquer da rede, se o valor de H for negativo, isso significará que houve violação da restrição imposta pela conexão que une as duas unidades. Se, por outro lado, H for positivo, isso significará que a restrição foi satisfeita, o que, portanto, maximiza H . Em uma rede contendo várias unidades, a

harmonia de um dado padrão de entrada é a soma de todas as micro-harmonias computadas para todos os pares de unidades conectadas (Prince & Smolensky 1997).

Duas correntes importantes desse tipo de união entre conexionismo e teoria da gramática são a Teoria da Harmonia (Smolensky 1986) e a Teoria da Otimização (Prince & Smolensky 1991). Segundo Tesar & Smolensky (1996), a Teoria da Harmonia (TH) postula que a boa formação de uma representação em uma rede neural é medida numericamente pelo valor de sua H , e que a probabilidade (P) de ocorrência de uma representação é também governada por H , onde $P \propto e^H$. Segundo a TH, uma representação é composta de uma parte explícita e uma implícita, e a função harmônica fornece o modelo que relaciona essas duas partes: dada uma certa estrutura explícita, associá-la a diferentes estruturas implícitas nos fornece valores diferentes de H e, conseqüentemente, diferentes valores de P . Quanto maior o valor de H , mais bem formada será a estrutura. A Gramática Harmônica (Legendre, Miyata & Smolensky 1990), uma aplicação da TH à lingüística, pressupõe que as estruturas implícita e explícita façam parte das descrições estruturais lingüísticas e que o modelo que as relaciona seja uma gramática.

A Teoria da Otimização (TO) também utiliza o conceito de harmonia na formulação de seus princípios, mas com a diferença de que a harmonia de suas descrições estruturais não é computada numericamente. Um dos princípios fundamentais da TO é o de que as gramáticas possuem restrições em relação à boa formação das estruturas lingüísticas, e que essas restrições encontram-se em conflito intenso. Um fator importante na TO é o fato de a disputa entre essas restrições ser satisfeita otimizada, não perfeitamente. Prince & Smolensky (1997) apresentam um exemplo bastante ilustrativo desse princípio:

“O inglês opera sob restrições que incluem o fato de sua ordem sintática básica ser sujeito-verbo-objeto; contudo, em uma frase como *What did John see?* [O que João viu?] é o objeto que figura primeiro. Isso demonstra a força superior de uma restrição que exige que palavras interrogativas como *what* se posicionem no começo da frase. Não obstante, até mesmo essa restrição não é absoluta: devemos dizer *Who saw what?* [Quem viu o que?] com a palavra interrogativa no

papel de objeto aparecendo em sua posição canônica; a alternativa potencial *Who what saw?* [Quem o que viu?], com todas as palavras interrogativas amontoadas na posição inicial, certamente gramatical em outras línguas, entra em conflito com um outro princípio de estrutura de oração que é, em inglês, ainda mais forte do que a necessidade de palavras interrogativas ocuparem a posição inicial. Portanto, *Who saw what?* é a estrutura gramatical, satisfazendo as restrições da gramática não de maneira perfeita, mas otimizada: nenhuma alternativa se sai melhor se levamos em consideração as forças relativas das restrições da gramática do inglês” (p. 1604).

Une-se a esse princípio outro elemento importante para a formalização de uma teoria gramatical desse tipo: não há sucesso na satisfação de uma restrição mais fraca que compense o fracasso em uma restrição mais forte. A TO opera, portanto, com base na noção de Dominância Hierárquica Estrita, segundo a qual cada restrição mais forte é absolutamente prioritária em relação a todas as outras restrições mais fracas. Esse tipo de dominância limita severamente as possibilidades algébricas de combinação de restrições dentro da gramática, possibilitando que restrições simples de boa formação dêem origem a conseqüências gramaticais extremamente complexas, o que satisfaz requisitos básicos da teoria da linguagem humana. A TO hipotetiza que esse conjunto de restrições seja universal: não como possibilidades de escolha, mas como presença obrigatória. A gramática de uma língua, sob essa ótica, consiste da imposição de uma hierarquia de dominância estrita a esse conjunto universal de restrições.

A TO também universaliza a função que determina, dada uma entrada qualquer, o conjunto de possibilidades de saída que competem pela otimização: todas as línguas levam em consideração as mesmas possibilidades no processamento de uma dada entrada. Restrições podem, nesse contexto, tanto não serem nunca violadas como serem permanentemente violadas. Para a TO, portanto, “gramaticalidade significa a satisfação otimizada de exigências conflitantes de restrições violáveis” (Prince & Smolensky 1997, p. 1605).

Smolensky (1999, p. 596) resume os princípios gerais da TO da seguinte forma:

- A gramática de uma língua produz, a partir de uma dada entrada, uma estrutura lingüística de saída que maximiza a Harmonia;
- A Harmonia de uma saída em potencial equivale à medida que ela satisfaz, simultaneamente, um conjunto de restrições violáveis da boa-formação lingüística;
- Essas restrições possuem forças distintas, o que determina qual delas prevalecerá quando houver conflito de restrições;
- A gramática de uma língua consiste de uma hierarquia de restrições, da mais forte para a mais fraca; uma restrição mais forte tem supremacia absoluta sobre as que se encontram abaixo;
- O conjunto das saídas possíveis e das restrições é o mesmo para todas as línguas; as gramáticas das línguas diferem somente em como as restrições são hierarquizadas.

Apesar de avanços significativos de formalização como a TO, o conexionismo também apresenta uma série de limitações e desafios a serem vencidos. Alguns deles serão analisados a seguir.

1.7 CRÍTICAS AO CONEXIONISMO

Com já foi mencionado neste capítulo, a empresa conexionista ainda encontra-se em um estado relativamente incipiente, portanto há muita controvérsia em torno de suas potencialidades e de seu escopo de ação. Muito já foi discutido desde a publicação da pioneira simulação da aquisição das formas verbais que compõem o passado do inglês por Rumelhart & McClelland (1987) e da subsequente reação de pesquisadores gerativistas (Pinker & Prince 1988, Fodor & Pylyshyn

1988), que representou a entrada oficial, por assim dizer, do conexionismo na arena da aquisição da linguagem. Mais recentemente, em uma comparação do inglês com outras línguas, Pinker (1998) desenvolve uma crítica ao conexionismo através da defesa da mente simbólica e do mecanismo dual da aquisição da linguagem, nomeadamente a memória (léxico) e as regras (gramática).

O primeiro princípio inclui a noção de um dicionário mental, uma lista finita de palavras memorizadas, sendo essas palavras símbolos arbitrários (uma conexão entre um sinal e uma idéia compartilhada por todos os membros de uma comunidade). A utilização desse mecanismo supõe que tanto o falante quanto o ouvinte possuam um registro de entrada de memória para a associação de um dado símbolo. Esse registro terá três vias associativas: o som, o significado e a categoria gramatical do item. Para um falante adulto de inglês, o número de entradas desse tipo oscila entre 60 e 100 mil (Pinker 1994). Esse número permite a esse falante expressar-se de maneira bastante eficiente.

O segundo princípio, a gramática, parte do pressuposto que os seres humanos, além de aprenderem itens lexicais isolados, os combinam em seqüências durante a fala. Cada falante possui um algoritmo mental finito que o capacita a gerar um número infinito de sentenças possíveis, cada uma representando um pensamento distinto. O sentido de uma frase, nesse contexto, se define pela combinação dos sentidos individuais dos itens lexicais com a maneira como eles são combinados. Essas combinações são de caráter recursivo, ou seja:

$$S \rightarrow NP + VP, \text{ onde } VP \rightarrow V + (NP) + (S)$$

o que possibilita associações em ambos os sentidos (da esquerda para a direita e vice-versa), criando um *loop* capaz de gerar sentenças de qualquer tamanho. A gramática de uma língua, portanto, tem caráter combinatório, e o número de mensagens possíveis cresce exponencialmente com o tamanho da frase, fazendo com que o falante possa expressar um número infinito de sentidos.

Esse mecanismo dual, contestado pelos conexionistas (Plunkett & Juola 1999), é basilar ao pensamento dos gerativistas. Nas palavras de Pinker (1998):

“A linguagem maximiza as vantagens distintas das palavras e das regras na combinação das duas, cada uma delas sendo manipulada por um sistema psicológico distinto. Há um léxico de palavras para entidades comuns e idiossincráticas; o mecanismo psicológico que dele se ocupa é simplesmente um tipo de memória. Existe, também, um sistema separado de regras gramaticais combinatórias para novas combinações de entidades; o mecanismo psicológico que dele se ocupa é a computação simbólica” (p. 221).

O autor analisa, sob essa ótica, questões envolvendo flexões regulares e irregulares dos verbos do inglês em comparação com outras línguas, alertando para o fato de que a noção de memória descrita acima apresenta alguns problemas, pois as formas flexionais irregulares não são apenas um conjunto arbitrário de exceções, pois, como ressaltam os conexionistas, há padrões entre essas formas. Isso faz com que elas não possam ser simplesmente explicadas através de um léxico de itens armazenados. Pinker (1998) propõe, então, uma reavaliação da noção de memória apresentada acima, advogando que ela não seja apenas uma lista não-estruturada de itens, mas sim parcialmente associativa, como nos modelos conexionistas, na qual “características estão associadas a características, e palavras a palavras (p. 225)”.

Não obstante, quanto às formas regulares, Pinker defende a idéia de que elas sejam geradas por uma regra simples de concatenação simbólica, podendo ser associadas a qualquer palavra, independentemente de seu *status* na memória. Para tanto, o autor analisa a correlação existente, tanto no inglês como na maioria das línguas, entre frequência e irregularidade, alegando que só os verbos de alta frequência mantêm, ao longo dos anos, sua irregularidade, pois são facilmente memorizados de geração para geração. Por outro lado, verbos inicialmente irregulares que caem em desuso são regularizados (Bybee 1985), pois a regra de regularização simbólica pode ser aplicada a um verbo independentemente da sua frequência.

Pinker também analisa o fenômeno da regularização sintática, casos nos quais algumas formas irregulares exigem, misteriosamente, flexões regulares em alguns contextos, como em:

(1) Ex.: *All my daughter's friends are low-lives (*low-lives)*. (Todos os amigos da minha filha são marginais.) (p. 229)

O autor conclui que, ao contrário do que alegam os conexionistas (Rumelhart & McClelland 1987), as formas flexionadas regulares não são operações *default* que se aplicam sempre que a memória não consegue recuperar uma forma flexionada. A flexão regular ocorre “de maneira livre sempre que a memória fracassa, porque a flexão regular é computada através de uma operação mental que não *precisa* ter acesso aos conteúdos da memória, ou seja, uma regra de processamento simbólico” (p. 240).

As evidências citadas no estudo de Pinker corroboram, apesar da incorporação de noções conexionistas em relação à estrutura da memória, a hipótese básica gerativista de que a estruturação da linguagem humana lança mão de dois mecanismos mentais, a saber, “a memória, para os signos arbitrários que subjazem às palavras, e a computação simbólica, para a utilização infinita dos meios finitos que subjazem à gramática” (p.240).

O mecanismo dual não é, contudo, como já foi mencionado neste capítulo, a única frente de disputa entre as duas teorias. Há outros grandes desafios com os quais o conexionismo tem que se defrontar, entre eles estando o debate em torno de questões fundamentais como a pobreza do estímulo (Seidenberg 1994), o grau de inatismo necessário para que se explique a aquisição da linguagem e a capacidade de geração de combinações inéditas similares às dos seres humanos. No que tange particularmente a esse último item, as redes conexionistas têm de ser capazes, sem que a elas sejam fornecidas regras explícitas de manipulação simbólica, de produzir estruturas às quais jamais tenham sido expostas, assim como reconhecer como anômalas estruturas com as quais ainda não tenham entrado em contato.

Marcus (1998), em sua análise dos modelos conexionistas mencionados em Elman *et alii* (1996), aborda essas questões referentes ao inatismo e à capacidade de generalização das redes conexionistas. Segundo o autor, Elman *et alii* (1996) tenta tanto fornecer mecanismos ao construtivismo como negar que as representações, os módulos e o maquinário de manipulação simbólica sejam inatos. Alinhando a visão conexionista à linha construtivista de Piaget, Marcus alega que esse tipo abordagem, que pressupõe a emergência de representações qualitativamente novas durante o processo de aquisição da linguagem, não se sustenta, até o presente, por uma explicação computacional concreta de como esse mecanismo pode de fato conduzir a redescrições representacionais. As alegações principais de Elman *et alii* (1996, p. 160), segundo o autor, são as de que:

- Não há representações inatas, mas sim representações aprendidas;
- Há poucos módulos inatos, mas há módulos aprendidos.

Para Marcus, portanto, a plausibilidade dessas afirmações dependerá da construção de modelos que desenvolvam novas representações e módulos sem que lhes sejam fornecidos representações e módulos inatos.

Quanto às representações, o autor afirma que as codificações vetoriais das entradas e saídas das redes utilizadas em Elman *et alii* (1996), das quais depende diretamente o sucesso ou fracasso desses modelos, são, necessariamente, representações fornecidas *a priori* a eles, portanto inatas. O que as redes aprendem não são, sob essa ótica, as representações em si mesmas, mas correspondências entre conjuntos de representações.

No que diz respeito aos módulos aprendidos, Marcus (1998) alega que, semelhante à questão das representações, os modelos utilizados em Elman *et alii* (1996) são módulos, uma vez que cada problema é resolvido através de um modelo

encapsulado em termos de informação, especializado naquela tarefa específica. Segundo o autor:

“Essas arquiteturas são sempre inteiramente pressupostas; nenhum desses modelos aprende a se dividir em módulos novos. Apesar dos autores de RI [Rethinking Innateness, Elman *et alii* (1996)] repetidamente defenderem uma posição em que módulos ‘emergem’, eles nunca de fato apresentam um modelo no qual módulos emergjam” (p. 163).

Além disso, o autor critica o fato de os modelos conexionistas diferirem tanto um do outro em termos de representação de entradas, de saídas e de regimes de treinamento, questionando a possibilidade de reuni-los em um único modelo explicativo.

Marcus (1998) também analisa as representações internas das unidades intermediárias das redes apresentadas em Elman *et alii* (1996), especialmente as representações desenvolvidas pelas redes construídas por Jeffrey Elman (1990, 1993), alegando não serem elas comparáveis a categorias gramaticais, ao contrário do que os autores afirmam. Os padrões comuns nas unidades intermediárias, relativos às entradas que representavam a categoria substantivo em Elman (1990), por exemplo, refletem, segundo Marcus, características distribucionais do corpo de treinamento, e não das entradas em si. O autor afirma que “padrões semelhantes de atividade nas unidades intermediárias refletem *palavras que tenham ocorrido em circunstâncias semelhantes, não palavras que pertençam à mesma categoria gramatical*. Não se pode afirmar, portanto, que as unidades intermediárias tenham desenvolvido representações abstratas da categoria *substantivo*” (p. 164). As unidades intermediárias formam, portanto, categorias que se baseiam em representações pré-existentes. Ou seja, essas categorias não são representações genuinamente novas. Corroborando esse fato estaria a incapacidade das redes de, ao receberem entradas desconhecidas, generalizar apropriadamente, da maneira como fazem os seres humanos, apesar da sua grande capacidade de, a partir de um treinamento adequado, aprender basicamente qualquer conjunto de treinamento. Sua capacidade de generalização de relações para novos elementos depende, segundo o

autor, da familiaridade da rede com esses elementos, enquanto que, dentro do paradigma simbólico, as regras são facilmente estendidas a elementos desconhecidos, pois elas constituem-se de relações abstratas entre variáveis. Segundo Marcus, isso se deve, em larga medida, a algoritmos como o *backpropagation*, baseados na regra delta e largamente utilizados nas simulações conexionistas, pois o ajuste dos pesos das conexões entre unidades é feito com base em informações disponíveis localmente para essas unidades, ao invés de utilizar informações globais sobre o desempenho geral da rede.

O que na visão de críticos do conexionismo como Pinker e Marcus são fraquezas, seus advogados analisam diferentemente. Em Plunkett (2000), a característica fortemente estatística e diversificada da empresa conexionista é vista como um avanço importante nos estudos da linguagem:

“O sucesso das redes conexionistas em imitar o desenvolvimento comportamental, cognitivo e lingüístico reside em sua sensibilidade às regularidades estatísticas inerentes ao ambiente. É importante escolher o tipo certo de rede para chegar a essas estatísticas. Uma vez escolhidos, a rede (ou sistema de redes) é capaz de integrar informações oriundas de múltiplas fontes e modalidades para construir representações cognitivas que não poderiam ter emergido de domínios isolados. O todo é maior do que a soma das partes” (p.7).

Também indo de encontro às críticas em relação à natureza das representações conexionistas está o já mencionado estudo de Ellis & Schmidt (1998) envolvendo o aprendizado, por parte de falantes adultos de inglês, de aspectos morfossintáticos de uma MAL (*Miniature Artificial Language* - Linguagem Artificial em Miniatura), em que frequência e regularidade encontram-se combinadas fatorialmente. Os autores conseguiram, em uma simulação conexionista, resultados intimamente relacionados aos obtidos nos dados coletados com os sujeitos. Ellis & Schmidt alegam, portanto, que a interação ocorrida entre regularidade e frequência é uma consequência natural da Lei Exponencial do Aprendizado (Newell 1990) e, por conseguinte, inteiramente consistente com processos de aprendizado associativos. Os

autores afirmam, então, que “sob uma perspectiva associativa, regularidade e frequência são essencialmente o mesmo fator com nomes diferentes” (p.329).

Conclusões como essa refletem o fato de esses pesquisadores verem no ambiente, e não em estruturas inatas, a grande fonte de explicação para o processo de aquisição da linguagem. Ao mencionarem os cuidados a serem tomados no treinamento da rede, Ellis & Schmidt ressaltam a escolha apropriada do conjunto de treinamento utilizado para expor a rede a evidências lingüísticas, pois ela deve “refletir de maneira adequada tipos e ocorrências (*types and tokens*), em razões representativas de formas regulares e irregulares, em uma seqüência que reflita de maneira plausível a exposição dos aprendizes em diferentes estágios de desenvolvimento” (p. 318). Embora esse tipo de maleabilidade de treinamento possa vir a melhorar o resultado de um experimento como esse, é difícil sustentá-la contra críticas de manipulação do regime de treinamento como as de Marcus (1998), Prince & Pinker (1988) e Fodor & Psylun (1988).

Hoje em dia há a aceitação, por parte dos críticos do conexionismo, de que modelos do tipo PDP, quando apropriadamente conectados, possam vir a ser capazes de implementar os fundamentos de um sistema cognitivo de maneira semelhante a como o *hardware* de um computador fornece o ambiente necessário à programação simbólica (Broeder & Plunkett 1994). Além disso, é de reconhecimento geral o fato de que algo similar aos modelos PDP tenha que constituir a base neurológica da mente aparentemente simbólica. Essa ótica, contudo, relega o papel dos modelos PDP ao nível neurológico, desqualificando-o no que tange aos processos cognitivos de nível superior, os quais seriam mais bem explorados através da computação de base simbólica. Como ressaltam Broeder & Plunkett (1994), a comunidade conexionista refuta esse papel, pois “um dos motivos principais para a construção de modelos PDP de processamento cognitivo foi o fato de a abordagem simbólica parecer carecer de certas características necessárias justamente no nível cognitivo de funcionamento” (p. 448).

Apesar de já haver, como é o caso da TO, tentativas de conciliar as duas abordagens (a simbólica e a conexionista), Broeder & Plunkett (1994) salientam que:

“... há um perigo nessa abordagem [a conciliatória]. Uma das vantagens dos sistemas PDP é que eles podem aprender. Ao contrário, os sistemas simbólicos são notoriamente impenetráveis e modulares. Eles tipicamente incorporam princípios universais que não podem ser extraídos sem ajuda do estímulo ao qual são expostos. Ao aceitarmos uma arquitetura que conjuga princípios simbólicos e de PDP, estamos também aceitando uma abordagem não desenvolvimentista para a cognição humana. O aprendiz de uma língua certamente já vem equipado com um conjunto sofisticado de restrições para o processamento de seu meio. Não obstante, devemos ter cuidado no sentido de nos certificarmos de que a competência atribuída à mente do aprendiz por um mecanismo de processamento simbólico não exceda os fatos relativos ao desenvolvimento humano” (p. 448).

Considerando-se terminada, para os propósitos deste trabalho, a descrição do panorama geral da empresa conexionista no estudo da aquisição da linguagem, será a seguir abordada a área de aquisição de L2 propriamente dita.

1.8 O CONEXIONISMO E A AQUISIÇÃO DE L2

Território ainda largamente inexplorado pelo conexionismo, a aquisição de L2 apresenta desafios ainda maiores para essa nova linha de pesquisa. Como enfatizam Broeder & Plunkett (1994), ainda há muita controvérsia em torno da medida da influência tanto da qualidade quanto da quantidade do *input* no modo como a L2 é percebida e processada pelos aprendizes, tornando-se ou não *intake* (Ellis 1994, Gass & Selinker 1994). Também não se encontra bem definida, na literatura, a relação entre o *input* e a proficiência dos aprendizes. Ou seja, a aquisição de L2 possui uma ampla variedade de contextos de aprendizado (não quase que apenas o naturalístico, como é o caso da L1), principalmente no que tange ao ambiente instrucional de língua estrangeira (caso do presente estudo). Sabe-se que esses aspectos contextuais

relativos ao ambiente de aquisição, assim como demandas comunicativas impostas ao aprendiz, podem afetar a aquisição de L2.

Os fatores descritos acima complexificam o estudo de base conexionista na área de L2, pois as redes neurais devem ser treinadas para que possam resolver os problemas a elas apresentados, e os treinos sempre vêm imbuídos de pressuposições sobre o tipo e a quantidade do estímulo recebido por um aprendiz. Gasser (1990, p. 189) lista alguns outros fatores que diferenciam as aquisições de L1 e L2, tendo, portanto, impacto potencial na realização de simulações nessa área:

- Há influência interlingüística entre L1 e L2;
- Mudanças de natureza neurológica ou de desenvolvimento cognitivo que não estão diretamente relacionadas à linguagem podem restringir a habilidade do indivíduo de adquiri-la, ou podem predispor-lo a utilizar estratégias de aquisição específicas.

Apesar das dificuldades inerentes à tarefa, já houve alguns estudos substantivos na área de aquisição de L2. Blackwell & Broeder (1992, *apud* Broeder & Plunkett 1994) realizaram uma série de simulações visando a explorar o efeito da frequência do *input* e da interação entre L1/L2 em processos de aquisição de L2. Os autores treinaram uma rede MP convencional de três camadas, utilizando *backpropagation*, a produzir formas pronominais em línguas diferentes (turco, árabe e holandês), a partir de especificações funcionais de uso pronominal.

A camada de entrada da rede especificou as diferentes funções codificadas nos sistemas pronominais do turco, do árabe e do holandês, a saber, pessoa, número, *status*, gênero, ou caso, além da língua em questão. As saídas foram configuradas em dois blocos: o primeiro representou os pronomes do turco ou árabe, com tamanho variável dependendo da língua em questão; o segundo, os pronomes do holandês. A tarefa da rede foi a de mapear o grupo de funções pronominais da entrada a formas pronominais específicas nas saídas. A rede foi configurada com 17 unidades de

entrada, 10 intermediárias e 36 ou 8 de saída, para os pronomes do turco ou o árabe (nessa ordem), além das 15 unidades relativas às formas pronominais do holandês.

O treinamento foi realizado com a entrada de todas as combinações forma / função possíveis para o turco ou árabe (144 tipos), seguido da entrada de todas as combinações para os pronomes do holandês, de acordo com as frequências relativas das suas relações forma / função (obtidas a partir dos dados em Broeder 1991¹, num total de 8644 instâncias). A tarefa da rede consistiu, então, em aprender *primeiramente o mapeamento pronominal para o turco ou o árabe, e depois para o holandês*.

As simulações produziram os seguintes resultados:

- Nos estágios iniciais de aprendizado, um grupo de pronomes é adquirido relativamente rápido, enquanto a velocidade de aquisição dos demais se concentra em torno de um mesmo valor. Esses últimos são adquiridos de forma sucessiva;
- As formas adquiridas por último diferem no modo como elas são afetadas pela aquisição das formas adquiridas anteriormente;
- A ordem de aquisição dos pronomes do holandês observada nas redes, tanto pelos falantes de turco como de árabe, se assemelha à ordem encontrada em Broeder (1991);
- Tanto as simulações como o estudo com os sujeitos apresentaram os *mesmos princípios gerais de aquisição*:
 - a) Os pronomes de 1ª pessoa foram adquiridos antes dos pronomes de 2ª pessoa;

¹ Anteriormente a essas simulações, Broeder realizou um estudo longitudinal da aquisição de holandês por quatro falantes adultos de turco e quatro de árabe.

- b) As formas singulares foram aprendidas antes das plurais;
- c) Formas subjetivas foram adquiridas antes das formas objetivas e possessivas;
- d) As formas masculinas foram adquiridas antes das femininas.

No geral, os pronomes demonstraram ser um grupo bastante homogêneo da interlíngua do holandês, com uma alta impenetrabilidade às propriedades da L1.

Outro estudo conexionista importante em L2 na área da transferência lingüística foi o conduzido por Gasser (1990). O autor partiu do princípio de que o processamento lingüístico equivale a um processo de preenchimento de padrões. Um padrão inclui elementos de todos os tipos a partir dos quais um falante pode tecer generalizações, ou seja, tanto elementos formais quanto contextuais. Segundo o autor, as redes auto-associativas são capazes de implementar esse tipo de processo.

Para Gasser, os modelos conexionistas são bastante apropriados ao estudo da transferência, pois quando “uma rede aprende a associar um padrão P1 a um padrão P2, ao ser apresentada a um padrão P3, ela tenderá a ativar um padrão que se assemelhe a P2 somente na medida em que P3 se assemelhe a P1” (p. 189). Além disso, o autor também utilizou, como pressuposto para as suas simulações, a idéia de que os primitivos de forma e conteúdo sejam idênticos em todas as línguas. Ou seja, as unidades básicas de rede através das quais a forma e o conteúdo são representados são as mesmas para L1 e L2. A transferência, dentro desse contexto, se daria sempre que houvesse uma sobreposição de qualquer tipo entre L1 e L2.

Gasser (1990) utilizou uma rede MP auto-associativa na qual padrões de entrada foram mapeados a padrões idênticos de saída através de uma camada de unidades intermediárias. Tanto as camadas de entrada como as de saída possuíam três blocos distintos: o primeiro, representando a linguagem sendo aprendida (L1 ou L2), o segundo, a forma das palavras e suas posições na oração, e o terceiro, o

conteúdo semântico das palavras e seu papel na proposição da oração. As entradas foram construídas com seqüências de duas palavras, e os papéis semânticos utilizados foram agente (sujeito) e processo (verbo). O sistema foi treinado com *backpropagation*. A rede foi configurada com 51 unidades de entrada, 25 intermediárias, 51 de saída e treinada com um grupo pequeno de palavras e significados (6 verbos e seus 6 significados, 6 substantivos e seus 6 significados).

Após a primeira fase de treinamento, durante a qual a rede aprendeu a mapear entradas e saídas idênticas com alta precisão, ela se demonstrou capaz de preencher eficazmente padrões parciais de entrada tanto com elementos conhecidos (parte do conjunto de treinamento), como com elementos desconhecidos, com um grau de precisão um pouco menor para esses últimos.

A segunda fase de treinamento da rede foi concebida com vistas ao teste da transferência lingüística ($L1 \rightarrow L2$, $L2 \rightarrow L1$), portanto foram utilizadas entradas que diferiam das entradas da primeira fase em diferentes graus. Alguns resultados importantes foram então obtidos:

- Os padrões de L1 claramente sofreram interferência dos padrões de L2, perdendo um pouco de sua precisão de mapeamento;
- Os padrões de L2 sempre apresentaram precisão menor em comparação aos da L1;
- Os padrões de L2 nunca foram tão difíceis de mapear quanto os da L1 na primeira fase de treinamento, mesmo quando havia um alto grau de diferença entre os padrões de L1 e L2;
- O mapeamento sintático foi mais difícil quando os padrões de L1 e L2 para tal eram mais díspares, mas a dificuldade tendeu a diminuir ao longo do treinamento;

- O mapeamento sintático foi mais fácil quando os padrões de L1 e L2 para tal eram mais semelhantes, mas a facilidade também tendeu a diminuir ao longo do treinamento.

Esses resultados se assemelham a outros já encontrados na literatura em relação ao fenômeno da transferência. A influência da L2 na L1, por exemplo, consta em Sharwood Smith (1983). Kellerman (1978) aponta a percepção, por parte do aprendiz, da relação entre L1 e L2 como um fator determinante do grau de transferência, o que se assemelha à diferença de dificuldade de aprendizado apresentada pela rede quando os padrões sintáticos de L1 e L2 eram díspares.

A diferença entre a aprendizagem de L2 por parte de crianças e de adultos foi investigada por Sokolik (1990), que realizou uma simulação conexionista em que houve a manipulação da taxa de aprendizado da rede para contemplar diferenças de ordem neurológica tidas como intrínsecas ao desenvolvimento humano. Essa taxa foi alterada com base no fator NGF (*Nerve Growth Factor* – Fator de Crescimento de Nervo, Purves 1986, *apud* Sokolik 1990).

Sokolik utilizou, nessa simulação, uma rede de duas camadas e um algoritmo de ajuste de pesos baseado na regra delta. O autor comparou a aquisição de L2 por uma criança de três anos de idade e por um adulto de 25, atribuindo-lhes taxas de aprendizado de 0.25 e 0.05, respectivamente, com o objetivo de refletir a diminuição da taxa de aprendizado com a idade. Essa pressuposição é justificada, em Sokolik (1990), por haver “ampla aceitação do fato de as crianças adquirirem uma segunda língua mais prontamente do que os adultos” (p. 690). O resultado da simulação mostrou que a criança adquiriu (i.e., obteve 95% de precisão) uma estrutura hipotética X após 10 instâncias de treinamento, enquanto que o adulto só adquiriu a mesma estrutura X após 55 instâncias.

Sokolik & Smith (1992) realizaram outro experimento em L2, dessa vez visando a investigar a aquisição do gênero no francês como L2, considerado um fenômeno adquirido cedo em L1 e tarde em L2 (Andersen 1984, *apud* Sokolik &

Smith 1992). Foi investigada a possibilidade de a informação inerente à estrutura de substantivos isolados do francês ser suficiente para que a atribuição de gênero pudesse ocorrer eficientemente, sem que fosse necessário o acesso a outras fontes de informação (regras, por exemplo).

Segundo os autores, a aquisição do gênero do francês poderia se dar, sob o ponto de vista simbólico, por três vias distintas. O aprendiz poderia:

- Aprender a reconhecer que certos agrupamentos ortográficos ou fonéticos servem como indicações para a atribuição de gênero;
- Basear-se em informações contextuais que especificam os gêneros dos substantivos, a saber, a concordância de artigos e adjetivos;
- Basear-se na memorização de substantivos associados a artigos marcados para o gênero.

Sokolik & Smith utilizaram um associador de padrões com duas camadas, contendo 224 unidades de entrada para a identificação das letras do alfabeto e suas posições nos substantivos franceses, e duas unidades de saída, para a indicação do gênero (masculino ou feminino). O *corpus* de treinamento foi construído a partir de palavras de 3 a 5 letras, retiradas de um livro-texto de francês básico, somando 406 nomes masculinos e 317 femininos, de onde foram escolhidos, aleatoriamente, 300 de cada, totalizando 600 substantivos. Desses foram escolhidos, também balanceada e aleatoriamente, 450 itens para o treinamento e 150 para o teste de generalização. Foi utilizada a regra delta como algoritmo de aprendizado, e a classificação das saídas em masculino ou feminino foi estabelecida a partir do maior valor de ativação da unidade correspondente.

Houve três fases de treinamento: a primeira, simulando o aprendizado a partir de uma situação de *tabula rasa*, i.e., com os pesos iniciais das conexões iguais a zero; a segunda, simulando a existência de conhecimento prévio, com atribuição

aleatória de pesos iniciais; e a terceira, simulando a queda na capacidade de aquisição ao longo da idade, com base em Sokolik (1990). Depois de treinada, a rede demonstrou a capacidade de associar com precisão as características ortográficas dos substantivos franceses do conjunto de treinamento ao seu gênero, sem que lhe fosse fornecida nenhuma informação adicional (regras ou contextos). Além disso, a rede também demonstrou ter capacidade de atribuir gênero a substantivos novos, também com elevado grau de precisão.

Com base nesses resultados, os autores sugerem que:

- A detecção de padrões é possivelmente uma base mais confiável de atribuição de gênero no aprendizado de francês como L2 do que o contexto ou as regras;
- Aprendizes adultos de L2 aprendem a atribuir gênero de forma mais lenta do que as crianças, tanto por causa da maturação neural (taxa de aprendizado menor) como por causa da influência da L1 (presença de pesos iniciais treinados para a L1 e não treinados para a L2);
- A precisão em atribuir gênero a substantivos desconhecidos cresce com a precisão em atribuir gênero aos membros do grupo de treinamento, ou seja, o aprendiz melhora sua intuição gramatical ao longo do aprendizado sem receber evidência explícita de nenhuma regra.

Os estudos citados até aqui já representam avanços significativos do conexionismo na pesquisa da aquisição de L2, mas há muito a ser feito ainda. Broeder & Plunkett (1994) ressaltam as novas possibilidades que a abordagem conexionista faz surgir para o estudo em L2. Entre elas está a ênfase no processo (aprendizado) e seus mecanismos de mudança, ao invés de no produto (estágio), e no funcionalismo, ao invés de na competência, situação geralmente invertida nos estudos de inspiração simbólica. Nas palavras de Schmidt (1994):

“Os modelos conexionistas são especialmente bons exatamente na simulação daqueles aspectos da linguagem e do aprendizado da linguagem nos quais os sistemas baseados em regras se demonstram pouco apropriados: o aprendizado gradual de áreas da gramática para as quais nenhuma distinção clara entre casos regulares e excepcionais pode ser traçada” (p. 192).

Broeder & Plunkett (1994) também salientam o fato de o conexionismo trazer consigo a possibilidade de uma investigação mais abrangente e precisa por modelos funcionalistas como o Modelo da Competição (Bates & MacWhinney 1989), pois ele possibilita a análise de como vários tipos de marcadores (*cue-types*) interagem durante o aprendizado. Além disso, Broeder & Plunkett enfatizam a propriedade dos modelos conexionistas para o estudo do desenvolvimento dos sistemas de interlíngua dos aprendizes, extremamente dinâmicos tanto sob uma perspectiva transversal como longitudinal. Um dos reflexos dessa nova perspectiva foi a extensão do Modelo de Pidginização de Schumann, de modo que ele incluísse processos cognitivos (Schumann 1990), num reconhecimento de que uma teoria completa da aquisição de L2 tem de conter uma explicação consistente desses processos.

A máquina computacional de base conexionista fornece ao pesquisador em L2 um instrumento versátil e poderoso de teste e adaptação de hipóteses e algoritmos de aprendizagem. Como apontam Hatch, Shirai & Fantuzzi (1990), o conexionismo pode vir a ser uma maneira de integrar os módulos isolados do vasto escopo de investigação dos estudos em aquisição de L2. Esses autores vislumbram o futuro da empresa conexionista através da interessante metáfora do retroprojetor, na qual várias lâminas encontram-se sobrepostas, representando as diferentes frentes de investigação da aquisição da L2 no presente, e através das quais passa uma luz que, de acordo com o movimentar dessas lâminas, salienta um ou outro aspecto nelas contido. Ou seja, apesar da aparente modularidade das lâminas, nesse contexto elas encontram-se multiplamente interconectadas em graus que mudam permanentemente. Tais dinamismo e complexidade requerem, segundo os autores, uma ferramenta de análise compatível:

“Com os supercomputadores, podemos vislumbrar o processamento em paralelo de componentes de maneira muito semelhante à nossa metáfora do retroprojektor. Com computadores em rede, podemos ver como as exigências da comunicação entre mais de duas pessoas pode ser modelada. Com o processamento em paralelo, um modelo não seqüencial, no qual as conexões são ativadas e trabalham simultaneamente, com a possibilidade de se suprirem mutuamente, caso necessário, torna-se um modelo de processamento lingüístico possível. As ‘partes’ não precisam ser encaradas como módulos ou transparências fixas e separadas, mas sim como um grande número de conexões que possam vir a ser ligadas ou ativadas em conjunto, de maneira a formar padrões conectados por ligações” (p. 704).

Como foi mencionado na Introdução, o presente estudo abordará, sob a luz de *uma simulação conexionista, a produção de perguntas em inglês como língua estrangeira por falantes de português de quatro níveis de proficiência*. O próximo capítulo apresentará o experimento realizado junto aos aprendizes com essa finalidade. Primeiramente, será traçado um panorama geral dos tipos de padrões interrogativos da sintaxe do inglês moderno e de seu processo geral de produção. Depois disso, serão detalhados os procedimentos utilizados e os resultados obtidos através deles.

CAPÍTULO 2 - PRODUÇÃO DE PERGUNTAS POR APRENDIZES DE INGLÊS COMO L2

O presente estudo tem por objetivo estabelecer relações entre uma simulação conexionista da produção de estruturas em inglês como L2 e dados empíricos relativos ao mesmo fenômeno, coletados para esse fim. Dessa forma, foi realizada uma investigação transversal da produção de estruturas interrogativas em inglês como L2 por alunos brasileiros de uma escola de idiomas de Porto Alegre, divididos em quatro faixas de proficiência. Concomitantemente, foi simulada a produção dessas estruturas através de uma rede conexionista construída no simulador *T-learn* (Plunkett & Elman 1997), cujo desempenho foi avaliado nos diferentes estágios do treinamento ao qual ela foi submetida. Pôde-se, dessa forma, realizar uma análise comparativa através do cruzamento dos resultados obtidos na simulação com os relativos à análise dos dados dos sujeitos.

A seguir será traçado um panorama geral das estruturas interrogativas do inglês. Isso feito, serão apresentados os procedimentos adotados no experimento com os sujeitos e os resultados obtidos. A simulação conexionista será detalhada no Capítulo 3, no qual também serão apresentados os seus resultados. A análise dos resultados de ambas as etapas e suas inter-relações serão discutidas no Capítulo 4.

2.1 OS PADRÕES INTERROGATIVOS DO INGLÊS

As estruturas interrogativas do inglês moderno dividem-se em dois grandes grupos: as perguntas diretas e as indiretas. Um exemplo de pergunta indireta do inglês é:

- (1) Ex: *I would like to know if you are going to be at the party* (Eu gostaria de saber se você estará na festa).

Esse tipo de interrogação não possui nenhuma diferença de ordenamento de itens lexicais em relação às estruturas declarativas do inglês. O presente estudo se ocupará das perguntas diretas, que apresentam modificações de ordenamento sintático em relação às estruturas declarativas.

2.1.1 As Perguntas Diretas

Baker (1996) classifica as interrogações diretas do inglês nos seguintes tipos:

a) Perguntas Sim-Não

Esse tipo de pergunta apresenta estrutura frasal verbal denominada finita invertida, a qual implica em uma mudança na ordem das palavras em relação à estrutura declarativa de uma estrutura frasal verbal finita de propósitos específicos. Essa última possui um papel especial na sintaxe do inglês, sendo encontrada em frases enfáticas do tipo:

- (2) Ex.: *I DID go to the post office.* (Eu de fato fui ao correio).

que contrastam com frases do tipo:

- (3) Ex.: *I went to the post office.* (Eu fui ao correio.)

denominadas finitas de propósito geral. A diferença fundamental entre esses dois tipos de estrutura frasal é o fato de que enquanto (3) contém uma frase cujo núcleo é

um verbo temporalizado de modo comum, (2) contém uma frase correspondente cujo núcleo é uma raiz verbal, frase essa que serve de complemento a uma forma temporalizada de DO.

Para alguns outros verbos de caráter excepcional, a saber, as formas temporalizadas de BE, as formas temporalizadas do HAVE perfeito, assim como todos os verbos modais, a estrutura enfática e a normal coincidem, ou seja, frases como:

(4) Ex.: **Jack DOES can play the guitar.* (Jack toca violão.)

não são frases do inglês.

Como as perguntas diretas do inglês, como já foi salientado acima, formam-se a partir de uma frase verbal finita de propósito específico, sua composição pode ser descrita através dos seguintes passos:

1. Utilize uma estrutura frasal verbal finita de propósitos específicos;

2. Mova o verbo nuclear dessa estrutura para a esquerda da frase nominal no papel de sujeito.

(5) Ex.: *Did John arrive late last night?* (O João chegou tarde ontem?)

b) Perguntas Alternativas

Esse tipo de interrogação é composto por duas (ou mais) estruturas finitas invertidas unidas pela conjunção OR (ou).

(6) Ex.: *Did John sue Karen or did Karen sue John?* (O João processou a Karen ou a Karen processou o João?)

c) Perguntas Negativas Sim-Não

A seqüência formada por um núcleo verbal de propósito específico e uma partícula negativa NOT contraída funciona sintaticamente, no inglês, como um núcleo verbal de propósito específico por si só. Por essa razão, essas perguntas seguem as regras de formação das estruturas finitas invertidas:

1. Utilize uma estrutura frasal verbal finita de propósitos específicos na negativa;
2. Mova o verbo nuclear e o NOT contraído dessa estrutura para a esquerda da frase nominal no papel de sujeito.

(7) Ex.: *Hasn't George written this composition?* (O George não escreveu esta redação?)

Esse tipo de pergunta é, tipicamente, utilizado em confirmações.

d) Perguntas Tag

Geralmente adicionadas ao final de uma frase declarativa, as perguntas *tag* seguem a seguinte regra de formação:

1. Utilize como sujeito o pronome que concorde com o sujeito da frase declarativa;
2. Coloque, à esquerda desse sujeito, um verbo de propósito específico que se encaixe com o verbo da frase declarativa;
3. Se a sentença declarativa for afirmativa, faça com que o verbo de propósito específico seja negativo, e vice-versa.

(8) Ex.: *Martin lives in Denver, doesn't he?* (O Martin mora em Denver, não mora?)

Essas perguntas também são tipicamente utilizadas em confirmações.

e) Perguntas Frasais

Também baseadas em estruturas finitas invertidas, as perguntas frasais são formadas da seguinte maneira:

- Se a frase interrogativa for o sujeito da pergunta, una a frase interrogativa a uma sentença finita comum com um sintagma nominal subjetivo faltando.

(9) Ex.: *Whose dog bit the postman?* (O cachorro de quem mordeu o carteiro?)

- Caso contrário, una o sintagma interrogativo a uma estrutura finita invertida com um sintagma do mesmo tipo do sintagma interrogativo faltando.

(10) Ex.: *What did Fred do with the rug?* (O que o Fred fez com o tapete?)

Haegeman (1994) menciona o fato de algumas perguntas frasais serem formadas a partir de sentenças finitas comuns, através da substituição de um constituinte por uma palavra interrogativa. A autora as denomina Perguntas Eco:

(11) Ex.: *Peter will invite whom?* (O Peter vai convidar quem?)

Nessa mesma linha, mas com a ausência da substituição, estão algumas confirmações simples, que não implicam em nenhum tipo de transformação sintática em relação à forma declarativa, havendo apenas mudança de entonação.

(12) Ex.: *He saw her yesterday?* (Ele a viu ontem?)

Esse é o quadro geral das estruturas interrogativas diretas do inglês. A próxima seção apresentará o experimento realizado com os alunos da escola de idiomas de Porto Alegre.

2.2 METODOLOGIA

O experimento realizado junto ao Instituto Cultural Brasileiro Norte-Americano, em Porto Alegre, teve como objetivo principal a análise da produção tanto de interrogações diretas do tipo sim / não como frases. A coleta dos dados foi transversal, envolvendo quatro níveis distintos de proficiência. Como se verá a seguir, optou-se por trabalhar com base na noção de um sujeito geral de nível de proficiência X, ou seja, um protótipo concebido a partir da soma das produções dos sujeitos pertencentes a cada um dos níveis de proficiência estudados. A análise estatística dos dados foi, dessa forma, elaborada com base no total das interrogações classificáveis produzidas por cada nível de proficiência estudado.

2.2.1 Os Sujeitos

Foram pesquisados alunos brasileiros cursando inglês como língua estrangeira na escola de inglês supracitada, matriculados em cursos regulares (aulas duas vezes por semana, com duração de 1h20min cada, em um total de 48 horas de instrução por semestre). Esses alunos possuíam, na data da coleta, entre 14 a 19 anos de idade e pertenciam a quatro níveis diferentes de proficiência, cada nível correspondendo a 96 horas de instrução formal (dois semestres). O critério de proficiência utilizado foi a classificação inicial e os critérios de avaliação do curso estudado, ou seja, os sujeitos não foram submetidos a nenhum teste específico para esse fim.

A faixa de idade utilizada foi considerada suficientemente homogênea, de modo que não foi necessário levar o fator idade (Sokolik 1990) em consideração na

análise da produção desses sujeitos². Foram estudados, ao todo, 78 indivíduos, distribuídos da seguinte forma:

- Nível 1 – 22 alunos (de 14 a 16 anos);
- Nível 2 – 22 alunos (de 14 a 18 anos);
- Nível 3 – 17 alunos (de 14 a 19 anos);
- Nível 4 – 17 alunos (de 16 a 19 anos).

No que concerne ao material didático com o qual os sujeitos entraram em contato, se sabe precisamente qual o currículo do curso pesquisado (excetuadas as singularidades inerentes à instrução recebida por cada um dos grupos, não consideradas para os propósitos deste estudo), pois ele é baseado nos livros textos adotados. Por outro lado, não se dispõe de informações no que respeita aos currículos das escolas regulares freqüentadas por cada um dos sujeitos, pois não houve controle em relação a esse aspecto.

Os livros-texto utilizados pelos sujeitos (Maurer & Schoenberg 1998, Richards & Sandy 1998) são estruturados com base no *input gramaticalmente seqüenciado*, em sua maior parte, a partir dos tempos verbais do inglês. Abaixo se encontram listados os tempos verbais que constam do programa dos quatro níveis de proficiência pesquisados:

Nível	Unidade	Tempos verbais apresentados
1	1	<i>Present Simple (to be)</i>
1	2	<i>Present Simple (there + to be)</i>
1	3	<i>Imperative</i>
1	4	<i>Present Continuous</i>
1	5	<i>Present Simple</i>
1	6	<i>Future (going to)</i>

² Apesar da falta de consenso sobre os pormenores do assunto, a idade é amplamente reconhecida como um fator influente no processo de aquisição de L2 (Ellis 1994), dadas as transformações neurofisiológicas que ocorrem nos seres humanos ao longo de seu desenvolvimento.

1	7	<i>Modals (can) + Have to</i>
1	8	Nenhuma forma verbal específica
1	9	<i>Past Simple (to be)</i>
1	10	<i>Past Simple</i>
2	1	<i>Past Continuous</i>
2	2	<i>Modals (will)</i>
2	3	Nenhuma forma verbal específica
2	4	<i>Present Perfect</i>
2	5	<i>Modals (should, could)</i>
2	6	Nenhuma forma verbal específica
2	7	<i>Had Better</i>
2	8	<i>Would Rather</i>
2	9	<i>Gerunds</i>
2	10	<i>Modals (will be able to, will have to)</i>
3	1	<i>Present Perfect Continuous</i>
3	2	<i>Passive voice (Present Simple, Past Simple, Will)</i>
3	3	<i>Modals (may, might) + Tag Questions</i>
3	4	<i>Used to</i>
3	5	Nenhuma forma verbal específica
3	6	Nenhuma forma verbal específica
3	7	<i>Modals (can, must, might, should)</i>
3	8	Nenhuma forma verbal específica
3	9	<i>Was going to, Modals (would)</i>
3	10	<i>Infinitives, Emphatic Forms</i>
4	1	<i>Gerunds</i>
4	2	<i>Modals (should, ought to) + Infinitives, Gerunds</i>
4	3	Nenhuma forma verbal específica
4	4	Nenhuma forma verbal específica
4	5	Nenhuma forma verbal específica
4	6	Nenhuma forma verbal específica
4	7	<i>Pres. Perfect, Simple Past, Past Contin. & Perfect</i>
4	8	<i>Modals (should have done, ought to have done)</i>
4	9	Nenhuma forma verbal específica
4	10	Nenhuma forma verbal específica
4	11	<i>Future Perfect, Future Perfect Continuous</i>
4	12	<i>Passive Voice (Present Continuous)</i>

Quadro 1: Currículo dos Sujeitos

2.2.2 O Instrumento

Os dados foram coletados a partir de dois instrumentos diferentes: um formulário, no qual pediu-se que os sujeitos escrevessem 10 perguntas em inglês,

direcionadas a um entrevistado hipotético, e um questionário, através do qual se pediu que os sujeitos fornecessem informações sobre seus históricos no que tange ao aprendizado de línguas estrangeiras.

O formulário foi construído e contextualizado a partir de um evento ocorrendo simultaneamente à coleta, os *Jogos Olímpicos de Sydney*. A coleta foi realizada durante um dos períodos de aula de cada uma das turmas selecionadas, quando foi requisitado que os alunos imaginassem estarem na Austrália, na condição de repórter esportivo, e que preparassem 10 perguntas, em inglês, a serem utilizadas em uma entrevista com o nadador Australiano Ian Thorpe. Essa tarefa foi escolhida por três motivos básicos: pela motivação que um evento esportivo daquele porte causaria na população alvo; pelo pressuposto conhecimento prévio, por grande parte dos sujeitos, em relação ao nadador a ser entrevistado, o qual recebeu especial atenção da mídia durante os *Jogos*; e pela familiaridade que os alunos de cursos de idiomas têm com o tipo de situação proposta, na qual eles devem se imaginar em certos papéis sociais e se comunicar na língua alvo, pois essa é uma prática bastante comum nas aulas comunicativas de L2.

Os sujeitos receberam três páginas impressas: a primeira, contendo instruções gerais, em português (Anexo A); a segunda, contendo o formulário, com estímulo visual (fotos do nadador e logotipo olímpico), 12 sugestões (opcionais) de tópicos para as perguntas em inglês e 10 linhas numeradas onde as escrevessem (Anexo B); a terceira, contendo o questionário a respeito de seus históricos de aprendizado em língua estrangeira (Anexo C).

Não foi estabelecido um limite de tempo para que os sujeitos escrevessem as perguntas e completassem o questionário. Não obstante, como as instruções, explicadas em voz alta, em português, solicitavam que eles executassem a tarefa no tempo mais curto possível e não reformulassem perguntas já escritas, nenhum dos grupos demorou mais de 20 minutos para concluir a tarefa.

2.2.3 A Tabulação dos Dados

Primeiramente, foram tabuladas as respostas obtidas a partir do questionário, para que se pudesse traçar um perfil geral dos sujeitos investigados no que concerne ao seu tempo de estudo de inglês na escola pesquisada (com ou sem interrupções), fora dela e na escola regular, assim como seu contato com línguas que não fossem o inglês. Os itens tabulados foram:

1. Tempo de estudo na escola pesquisada (em semestres);
2. Interrupção do curso na escola pesquisada (sim / não);
3. Tempo de inglês na escola regular (em anos);
4. Estudo em outros cursos de inglês (sim / não);
5. Estudo de outras línguas (sim / não).

Para os itens 1 e 3 foram calculados, em semestres e anos, respectivamente, a média aritmética (μ) e o desvio padrão médio (s) do tempo de estudo dos alunos na escola pesquisada, ou seja, o tempo de exposição dos alunos ao currículo escolhido por aquela instituição. Já para os itens 2, 4 e 5, calculou-se a frequência relativa (fr) de ocorrência dos eventos citados, igual ao número de sujeitos que responderam positivamente à pergunta dividido pelo número total de sujeitos da amostra para aquele nível de proficiência, sem que fosse levado em consideração, para os propósitos desse perfil geral dos sujeitos, o tempo específico de duração de cada uma das ocorrências.

No que tange ao formulário, a tabulação das interrogações produzidas pelos sujeitos baseou-se em critérios sintático-morfológicos relativos a três categorias, a saber: VPaux (sintagma verbal - auxiliar), NPsuj (sintagma nominal - sujeito), VPmain (sintagma verbal - principal), nessa ordem. Os elementos utilizados nas

perguntas foram, primeiramente, codificados segundo o Quadro 2, através do qual foram reduzidos a certas características dos núcleos de seus XPs:

VPaux		NPsuj		VPmain	
Código	Categoria	Código	Categoria	Código	Categoria
am ar is	am are is	1ps	1ª pessoa do singular	st	infinitivo (do inglês, <i>stem</i>)
ws wr	was were	2ps	2ª pessoa do singular	ng	particípio presente
do ds	do does	3ps	3ª pessoa do singular	ir	verbo conjugado irregularmente no passado ou particípio passado
dd	did	1pp	1ª pessoa do plural	ed	verbo conjugado regularmente no passado ou particípio passado
hv	have	2pp	2ª pessoa do plural	3s	verbo conjugado na 3ª pessoa do singular do <i>Present Simple</i>
hs	has	3pp	3ª pessoa do plural		
hd	had				
md	modal				

Quadro 2: Codificação dos Dados

Somente as perguntas que se enquadraram na classificação acima (ou seja, seguiram o padrão VPaux – NPsuj – VPmain) e apresentaram pelo menos um item lexical explícito na posição NPsuj foram codificadas, as demais foram agrupadas em uma categoria denominada “Não classificável³”:

- (13) Ex.: **Are you have a lot of friends?* = Classificável [Sujeito (S) 13, Nível 1 (N1)]

³ Como o objetivo do estudo foi investigar padrões interrogativos que apresentassem movimento do verbo auxiliar, padrões que não apresentaram esse movimento foram excluídos. A presença

- (14) Ex.: *What about your hobbies?* = Não classificável [S8, N3]

A codificação se deu com base na posição NPsuj, a partir da qual foram selecionados os elementos que ocuparam a posição VPaux, à esquerda, e a posição VPmain, à direita. Somente essas três posições foram utilizadas nas codificações.

- (15) Ex.: **How long are you training swimming?* = AR (VPaux) + 2PS (NPsuj) + NG (VPmain) [S6, N4]

Com exceção da posição NPsuj, as outras duas posições puderam ser ocupadas por elementos que, tradicionalmente, ocupariam as posições à direita ou à esquerda, como em:

- (16) Ex.: **Where live your family?* = ST (VPaux) + 3PS (NPsuj) [S21, N1]

Erros de ortografia não foram levados em consideração, salvo quando causaram ambigüidade (como *thing* x *think*, por exemplo), situação na qual ficou ao critério do codificador decidir a qual categoria os itens lexicais em questão pertenciam, com base no contexto sintático tido como o mais provável.

O item lexical *family*, que apareceu na posição NPsuj em várias das construções dos sujeitos pesquisados, foi considerado ou 3PP, ou 3PS, dependendo das indicações contextuais sintáticas fornecidas por VPaux e VPmain.

- (17) Ex.: *How is your family?* = IS (VPaux) + 3PS (NPsuj) [S2, N1]

- (18) Ex.: *How are family?* = AR (VPaux) + 3PP (NPsuj) [S9, N1]

Os seguintes casos foram considerados evasão de tarefa:

obrigatória de um elemento em NPsuj também serviu ao mesmo propósito, pois possibilitou uma posição de referência para esse movimento.

- Perguntas não concluídas.

(19) Ex.: **Do you have...?* = Evasão de tarefa [S11, N1]

- Mais de duas perguntas idênticas até a posição VPmain, em seqüência (nesse caso só foram consideradas válidas as duas primeiras perguntas).

(20) Ex.: *Do you have a brother?* = DO (VPaux) + 2PS (Npsuj) + ST (VPmain);

Do you have a sister? = DO (VPaux) + 2PS (Npsuj) + ST (VPmain);

**Do you have a girl-friend?* = Evasão de tarefa [S20, N2].

Com vistas à análise estatística dos dados, foi utilizado um sistema de pontuação e foram adicionadas duas categorias classificatórias, nomeadamente Csubj-aux (concordância Npsuj / VPaux) e Caux-main (concordância VPaux / VPmain), nessa ordem. Os critérios utilizados na pontuação levaram em consideração a presença ou não do elemento pertencente à categoria na posição a ele reservada (categorias VPaux, Npsuj e VPmain) e a concordância entre essas categorias (Csubj-aux e Caux-main). Critérios atendidos receberam pontuação 1 e não atendidos, pontuação 0. A pontuação mínima ficou, desse modo, em 1 ponto, e a máxima, em 5 pontos⁴.

(21) Ex.: *Where do you live?* = DO (VPaux) + 2PS (Npsuj) + ST (VPmain) [S14, N1]. Pontuação: 1 (VPaux) + 1 (Npsuj) + 1 (VPmain) + 1 (Csubj-aux) + 1 (Caux-main) = 5

⁴ Inicialmente, a pontuação foi realizada com as categorias VPaux, Npsuj e VPmain, mas a análise estatística dos dados não acusou nenhuma diferença significativa entre os quatro níveis de proficiência estudados, o que exigiu um refinamento maior das categorias de pontuação, obtido com a adição de Csubj-aux e Caux-main.

(22) Ex.: **When do you studied?* = DO (VPaux) + 2PS (NPsuj) + ED (VPmain) [S14, N1]. Pontuação: 1 (VPaux) + 1 (NPsuj) + 1 (VPmain) + 1 (Csubj-aux) + 0 (Caux-main) = 4

As categorias posição e concordância foram tratadas de forma interdependente, ou seja, perguntas do tipo:

(23) Ex.: **Is you training at Olympics games in the future?* [S22, N2]. Pontuação: 1 (VPaux) + 1 (NPsuj) + 1 (VPmain) + 0 (Csubj-aux) + 1 (Caux-main) = 4

receberam pontuação 1 para a posição VPaux (*is* funciona como auxiliar nessa posição), 1 para a posição NPsuj (pontuação obrigatória, uma vez que não foram tabuladas perguntas sem um elemento lexical explícito em NPsuj), 1 para VPmain (*training* é verbo principal no sintagma), 0 para a concordância Csubj-aux (*is* não concorda com o sujeito) e 1 para concordância Caux-main (*is* está em sintonia com a conjugação do verbo principal).

Já a ausência de elemento em VPaux, como em:

(24) Ex.: **What you do in your free time?* [S2, N3]. Pontuação: 0 (VPaux) + 1 (NPsuj) + 1 (VPmain) + 0 (Csubj-aux) + 0 (Caux-main) = 2

recebeu pontuação 0 tanto em VPaux como em Csubj-aux e Caux-main, pois a inexistência de um elemento em VPaux inviabilizou qualquer tipo de concordância.

Foi atribuída pontuação 1 a posições canonicamente vazias em VPmain, com em:

(25) Ex.: *How old are you?* [S1, N1]. Pontuação: 1 (VPaux) + 1 (NPsuj) + 1 (VPmain) + 1 (Csubj-aux) + 1 (Caux-main) = 5

pois esse é o padrão esperado nesses casos. Isso, por sua vez, causou pontuação 1 para a categoria Caux-main, pois a ausência de um elemento em VPmain inviabilizou qualquer tipo de discordância entre o verbo auxiliar e o elemento nulo que ocupa a posição VPmain. Em frases desse tipo, portanto, o verbo *to be* foi sempre considerado, para efeitos de tabulação, um verbo auxiliar, pelo fato de ele executar o mesmo tipo de movimento dos demais verbos considerados auxiliares.

Não houve, de modo geral, julgamento em relação a aspectos contextuais, ou seja, as perguntas não foram julgadas em relação à sua coerência pragmática (salvo nos casos de ambigüidade ortográfica supracitados).

Elementos não nucleares às categorias de posição como, por exemplo, artigos e advérbios, não influíram nas categorizações de posição e concordância. Portanto, em:

(26) Ex.: **Have you ever take drugs to be a better swimmer?* [S13, N3] = HV (VPaux) + 2PS (NPsuj) + ST (VPmain). Pontuação: 1 (VPaux) + 1 (NPsuj) + 1 (VPmain) + 1 (Csuj-aux) + 0 (Caux-main) = 4

(27) Ex.: **Wasn't you traveling last week?* [S10, N2] = WS (VPaux) + 2PS (NPsuj) + NG (VPmain). Pontuação: 1 (VPaux) + 1 (NPsuj) + 1 (VPmain) + 0 (Csuj-aux) + 1 (Caux-main) = 4

nem *ever* (26) nem *n't (not)* (27) foram levados em consideração para efeitos de pontuação.

A seguir serão apresentados os resultados obtidos através das tabulações descritas nesta seção.

2.3 OS RESULTADOS

Primeiramente, serão apresentados os resultados obtidos através do questionário, que visou a coletar informações gerais sobre o perfil dos sujeitos no

que tange aos seus históricos de aprendizado de línguas estrangeiras. Subseqüentemente, serão detalhados os resultados obtidos através da tabulação dos formulários, nos quais os sujeitos escreveram as perguntas em inglês, tanto no que tange aos resultados quantitativos gerais obtidos como em relação aos tipos de padrões produzidos e suas respectivas pontuações. Finalmente, serão apresentados os resultados gerados pela análise estatística dessas produções⁵.

2.3.1 O Questionário

Os resultados obtidos através do questionário foram os seguintes:

Nível de proficiência	1 Tempo estudo no curso pesquisado (sem.)	2 Interrupção no curso pesquisado	3 Tempo inglês na escola regular (anos)	4 Inglês em outros cursos	5 Estudo de outras L2s
N1 (2º semestre)	$\mu = 1,77$ $s = 0,35$	fr = 0,00	$\mu = 4,09$ $s = 1,73$	fr = 0,14	fr = 0,23
N2 (4º semestre)	$\mu = 3,50$ $s = 0,77$	fr = 0,00	$\mu = 4,95$ $s = 1,78$	fr = 0,27	fr = 0,41
N3 (6º semestre)	$\mu = 4,05$ $s = 1,58$	fr = 0,00	$\mu = 5,35$ $s = 1,27$	fr = 0,23	fr = 0,18
N4 (8º semestre)	$\mu = 5,64$ $s = 1,83$	fr = 0,12	$\mu = 5,94$ $s = 1,25$	fr = 0,35	fr = 0,41

Como se pode observar, os valores da seção 1 indicam um aumento gradual seguido de leve queda, ao longo das faixas de proficiência N1, N2, N3 e N4 (com número de sujeitos pesquisados equivalente a 22, 22, 17 e 17, respectivamente), da diferença entre o tempo esperado de estudo dos sujeitos na escola pesquisada à data da coleta (2, 4, 6 e 8 semestres, respectivamente, correspondendo aos níveis de proficiência 1, 2, 3 e 4) e a média real (μ) de estudo dos sujeitos [1,77 (13%); 3,50

⁵ Realizada com o auxílio do Departamento de Estatística da UFRGS, ao qual eu gostaria de expressar meu grande apreço e agradecimento.

(14%); 4,05 (48%); 5,64 (41%); nessa ordem). Essa progressão ocorre conjugada a um aumento progressivo do desvio padrão médio (s) para N1-4, ou seja, aumenta a variação intragrupo dos tempos de estudo dos sujeitos com o aumento da proficiência. Observa-se, na seção 2, que as interrupções de curso praticamente inexistem, pois apenas N4 apresenta $fr \neq 0,00$, ainda assim com um valor bastante baixo ($fr = 0,12$). Na seção 3, observa-se o esperado crescimento gradual do tempo médio de aulas de inglês na escola regular, dado o aumento da faixa etária dos sujeitos ao longo de seus estudos, que varia de 4,09 a 5,94 anos. Já a variação intragrupo, nesse caso, diminui com o aumento da proficiência dos sujeitos, pois o valor de (s) decresce. A seção 4 mostra que há graus de incidência de estudo em outros cursos de inglês diferentes de zero ($fr \neq 0,00$ para N1-4) em todas as faixas de proficiência, o que é indesejável no que respeita à homogeneidade curricular dos sujeitos. Similarmente, os valores de (fr), na seção 5, revelam que, em todos os quatro níveis de proficiência, houve sujeitos que tiveram contato com outras L2s que não o inglês, fato que também é indesejável no que tange à homogeneidade da amostra.

2.3.2 O Formulário

A Tabela 2 exhibe os resultados quantitativos gerais da coleta de dados realizada a partir do formulário.

Nível de proficiência	Sujeitos pesquisados	Perguntas classificáveis produzidas	Tipos de padrão interrogativo produzidos	Razão número de tipos / número de perguntas
N1	22	201	27	0,134328
N2	22	212	41	0,193396
N3	17	161	32	0,198757
N4	17	171	36	0,210526
Total	78	745	68*	

(*) Houve, como seria de se esperar, repetição de padrões entre os níveis de proficiência.

Os 78 sujeitos pesquisados produziram, no total, 745 perguntas e 68 tipos de padrões interrogativos classificáveis. O cálculo da proporção entre o número de tipos de padrão e o número de perguntas produzidas por nível revelou um aumento dessa razão com o aumento da proficiência (Figura 5). Ou seja, os sujeitos, nesse sentido, se mostraram gradualmente mais criativos de N1 a N4 (mas não necessariamente mais precisos, como revelará a análise estatística):

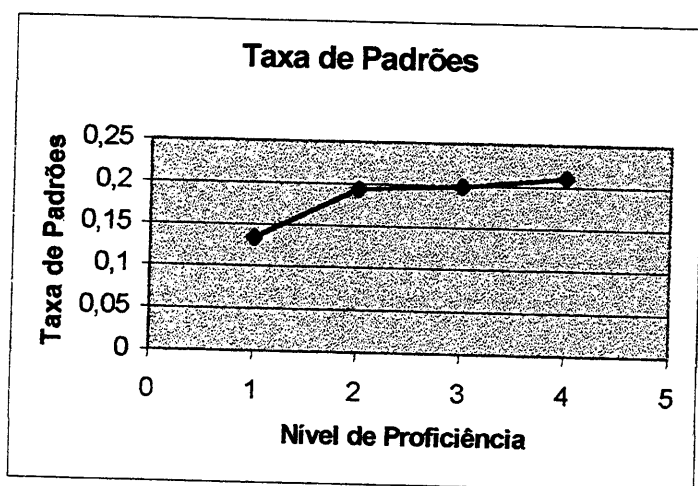


Figura 5: Taxa de Padrões por Nível de Proficiência

A Tabela 3 apresenta, em detalhe, os 68 padrões interrogativos produzidos pelos sujeitos de N1-4. Esses padrões foram ordenados por ordem de aparição na tabulação e por ordem crescente de pessoa, sendo pontuados, como foi salientado na seção 2.2.3, com valores variando de 1 a 5, segundo o seu grau de consonância com os padrões canônicos do inglês. Foi utilizada, também, a codificação do Quadro 2 para os elementos ocupando as posições VPaux, NPsubj e VPmain.

TABELA 3
RESULTADOS POR PADRÃO INTERROGATIVO

Tipo	Pontos	Padrões			Ocorrências			
		VPaux	NPsubj	VPmain	N1	N2	N3	N4
1*	5	do	2ps	st	85	86	64	80
2*	5	ar	2ps		15	5	8	7
3	2		2ps	ir	2	1	1	4
4	2		2ps	md	1	1	0	0
5	1		2ps	3s	1	0	0	0
6	2		2ps	st	23	36	19	9

7*	5	md	2ps	st	4	1	3	2
8	2		2ps	ng	1	3	0	0
9	4	ar	2ps	st	6	3	2	1
10	4	do	2ps	3s	2	0	0	0
11	3	do	2ps	ar	1	0	0	0
12	3	do	2ps		1	1	0	0
13	1	st	2ps		2	1	1	0
14*	5	dd	2ps	st	0	5	6	8
15*	5	ar	2ps	ng	0	1	2	2
16	3	do	2ps	do	0	1	0	0
17*	5	wr	2ps		0	2	0	0
18	4	ws	2ps		0	1	1	0
19	4	do	2ps	ir	0	1	0	2
20	4	ws	2ps	ng	0	1	0	0
21	2		2ps	ed	0	1	0	3
22	1		2ps		0	2	0	1
23	4		2ps	ar	0	1	0	0
24	3	do	2ps	is	0	1	0	0
25	2		2ps	do	0	1	0	0
26	3		2ps	is	0	1	0	0
27	3	dd	2ps		0	1	0	1
28	4	is	2ps	ng	0	1	0	0
29	3	do	2ps	hv	0	0	1	0
30*	5	hv	2ps	ed	0	0	2	1
31	3		2ps	wr	0	0	1	0
32*	5	hv	2ps	ir	0	0	3	2
33	4	is	2ps		0	0	1	0
34	4	dd	2ps	ed	0	0	1	3
35	3		2ps	ws	0	0	1	1
36	4	hv	2ps	st	0	0	3	0
37	4	dd	2ps	ir	0	0	1	1
38	4	do	2ps	ng	0	0	2	0
39	4	do	2ps	ed	0	0	1	0
40*	4	hd	2ps	ed	0	0	0	1
41*	5	is	3ps		22	17	15	12
42	4	ar	3ps		3	4	1	0
43	1	st	3ps		2	1	0	0
44	1		3ps		6	3	1	4
45	3		3ps	is	2	4	0	5
46	3		3ps	ws	1	0	0	1
47	2		3ps	st	2	2	0	1
48*	5	ws	3ps		0	1	1	5
49	1		3ps	3s	0	1	2	0
50*	5	hs	3ps	ir	0	1	0	0
51*	5	ds	3ps	st	0	0	2	0
52	4	wr	3ps		0	0	0	1

53	2		3ps	hd	0	0	0	1
54	2		3ps	ir	0	0	0	1
55	2		2pp	st	0	0	0	1
56	4	is	3pp		4	8	6	2
57	3	do	3pp		1	0	0	0
58	1		3pp		7	1	0	1
59	3		3pp	ar	1	1	0	0
60	2		3pp	st	1	1	1	1
61*	5	ar	3pp		3	5	6	2
62	3		3pp	is	2	1	1	0
63	2	is	3pp	is	0	2	0	0
64	3		3pp	ws	0	0	1	0
65*	5	dd	3pp	st	0	0	0	1
66*	5	hd	3pp	ed	0	0	0	1
67*	5	do	3pp	st	0	0	0	1
68	2		3pp	ng	0	0	0	1
NC**					6	12	9	5
ET**					9	1	0	1
*								
Total					216	225	170	177

(*) Padrão Canônico; (**) Não classificável; (***) Evasão de tarefa

Dos 68 tipos de padrões classificáveis produzidos, nota-se que apenas 17 são canônicos (tipos 1, 2, 7, 14, 15, 17, 30, 32, 40, 41, 48, 50, 51, 61, 65, 66 e 67). Há um predomínio dos padrões na segunda pessoa do singular (2ps), com 40 tipos, seguido dos padrões na terceira pessoa do singular (3ps), com 14, na terceira pessoa do plural (3pp), com 13, e na segunda do plural (2pp), com 1. Desses, os padrões mais frequentemente utilizados pelos sujeitos foram os dos tipos 1 (do + 2ps + st, 315 ocorrências), 6 (2ps + st, 87 ocorrências), 41 (is + 3ps, 66 ocorrências), e 2 (ar + 2ps, 35 ocorrências), ou seja, apesar de a grande maioria dos tipos produzidos não serem canônicos, quantitativamente as perguntas canônicas predominaram em todos os níveis de proficiência. Houve, também, uma baixa incidência tanto de perguntas não classificáveis (NC), com 32 ocorrências, como de evasões de tarefa (ET), com 11 ocorrências.

2.3.3 A Análise Estatística

A Figura 6 apresenta a progressão das médias das pontuações obtidas para os quatro níveis de proficiência investigados:

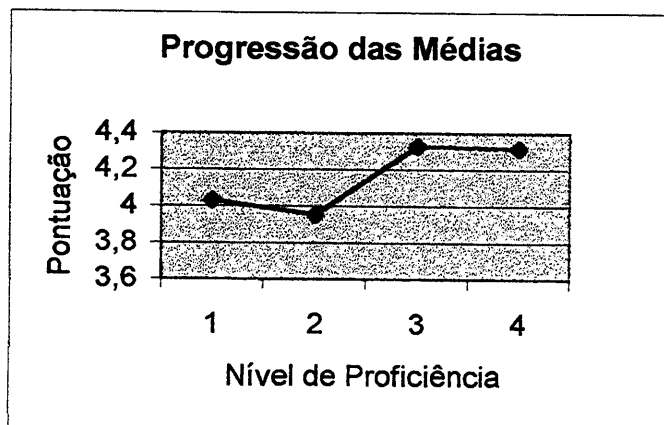


Figura 6: Progressão das Médias de N1-4

Com o objetivo de se saber se houve variações significativas entre as médias (μ) das pontuações (que equivalem à consonância das perguntas aos padrões interrogativos canônicos do inglês) observadas em N1-4, foi realizada, primeiramente, uma análise de variância (ANOVA) de fator único para um nível de significância $\alpha = 0,01$. A hipótese de teste (H_0) foi a de que não havia diferença significativa entre as distribuições de pontuação dos níveis de proficiência N1, N2, N3 e N4, ou seja, que a média de pontos obtidos nas perguntas produzidas pelos sujeitos de cada nível era estatisticamente igual às dos outros níveis ($H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$). Dado o valor de α , a confirmação de H_0 permitiria que se afirmasse com 99% de certeza que essa hipótese era verdadeira. A hipótese alternativa H_1 , conseqüentemente, foi a de que havia diferenças significativas entre as distribuições de pontuação de N1-4.

Os resultados da análise foram os seguintes:

TABELA 4				
RESULTADOS DA ANOVA				
RESUMO				
Grupo	Contagem	Soma	Média	Variância
Nível 1	201	810	4,029851	2,099104

Nível 2	212	838	3,95283	1,931414
Nível 3	161	697	4,329193	1,372205
Nível 4	171	738	4,315789	1,511455

VARIACÃO

Tipo	SQ	gl	MQ	F	valor-P	F crítico
Entre grupos	20,91037	3	6,970124	3,961241	0,00811	3,80814
Dentro dos grupos	1.303,849	741	1,759581			
Total	1.324,76	744				

Na primeira seção da Tabela 4, *Grupo* significa o nível de proficiência dos sujeitos, *Contagem*, o número de perguntas produzidas pelos sujeitos, *Soma*, a soma dos pontos obtidos por todas as perguntas produzidas, *Média*, a média aritmética dos pontos obtidos e *Variância*, a dispersão desses pontos. Nota-se que, a princípio, as médias agruparam-se em dois blocos de valores próximos, a saber, o composto por N1 e N2 e o composto por N3 e N4. A ANOVA realizada não nos permite precisar, contudo, se esses agrupamentos específicos se justificam estatisticamente, pois ela é um teste genérico da diferença entre as médias de um número k qualquer de grupos.

A segunda seção apresenta as variações intra e intergrupo analisadas pela ANOVA. Nela, *SQ* é a soma dos quadrados, *gl*, os graus de liberdade e *MQ*, as médias dos quadrados. Para que se possa confiar que H_0 seja verdadeira, o valor da estatística F (calculada a partir da razão entre o erro médio quadrado intra e interamostral) deve ser menor ou igual a ao valor crítico de F para $\alpha = 0,01$, ou seja, menor ou igual a $F = 3,80814$. Como o valor de F , nesse caso, é igual a 3,961241, o valor da probabilidade P é 0,00811. Esse valor de P permite que H_0 seja rejeitada, pois P é menor do que o nível de significância α . Há, portanto, variação significativa entre os níveis N1, N2, N3 e N4, ou seja, H_1 é verdadeira.

Como já se sabe que há diferenças significativas entre as médias das pontuações de N1-4, há que se determinar onde exatamente residem essas diferenças. Foram então realizados, para esse fim, uma série de testes *post hoc* de comparações múltiplas entre as médias de N1-4, a saber:

- Fisher's LSD – *Least Significant Difference* (Diferença Menos Significativa);
- Scheffé;
- Newman-Keuls;
- Duncan;
- Tukey HSD – *Honestly Significantly Different* (Diferença Honestamente Significativa);
- Tukey HSD para N desiguais.

Os resultados obtidos para cada um dos testes foram os seguintes:

Fisher's LSD		{1}	{2}	{3}	{4}
		4,029851	3,952830	4,329193	4,315790
N1	{1}		0,555515	0,033201	0,038609
N2	{2}	0,555515		0,006800	0,007934
N3	{3}	0,033201	0,006800		0,926714
N4	{4}	0,038609	0,007934	0,926714	
Scheffé		{1}	{2}	{3}	{4}
		4,029851	3,952830	4,329193	4,315790
N1	{1}		0,950762	0,208617	0,232410
N2	{2}	0,950762		0,061954	0,070081
N3	{3}	0,208617	0,061954		0,999793
N4	{4}	0,232410	0,070081	0,999793	
Newman-Keuls		{1}	{2}	{3}	{4}
		4,029851	3,952830	4,329193	4,315790
N1	{1}		0,555340	0,083153	0,038269
N2	{2}	0,555340		0,033626	0,021210
N3	{3}	0,083153	0,033626		0,926697
N4	{4}	0,038269	0,021210	0,926697	
Duncan		{1}	{2}	{3}	{4}
		4,029851	3,952830	4,329193	4,315790

N1	{1}		0,555340	0,159392	0,038269
N2	{2}	0,555340		0,097523	0,041969
N3	{3}	0,159392	0,097523		0,926697
N4	{4}	0,038269	0,041969	0,926697	
Tukey HSD	{1}		{2}	{3}	{4}
		4,029851	3,952830	4,329193	4,315790
N1	{1}		0,935215	0,142467	0,162294
N2	{2}	0,935215		0,033626	0,038874
N3	{3}	0,142467	0,033626		0,999721
N4	{4}	0,162294	0,038874	0,999721	
Tukey HSD (N desiguais)	{1}		{2}	{3}	{4}
		4,029851	3,952830	4,329193	4,315790
N1	{1}		0,937515	0,178826	0,190472
N2	{2}	0,937515		0,053189	0,055402
N3	{3}	0,178826	0,053189		0,999733
N4	{4}	0,190472	0,055402	0,999733	

A Tabela 5 apresenta os valores de P para todos os testes *post hoc* realizados. Cada seção mostra, no canto superior esquerdo, o nome do teste em questão. Abaixo dele figuram os quatro níveis de proficiência dos sujeitos, enquanto que à sua direita estão listadas as médias de pontuação de cada um desses níveis. Os valores de P, que ocupam a parte central de cada uma das seções, indicam o nível de significância para cada par de médias comparadas. Esses valores são comparados ao nível de significância de referência (α), para que se possa rejeitar ou não a hipótese de teste (H_0 = Não há diferença estatística entre as duas médias sendo comparadas). Nos casos em que P tenha valor inferior ao de α , H_0 será rejeitada, sendo H_1 , conseqüentemente, aceita.

(28) Ex.: Tukey HSD: $\mu_2 \neq \mu_3$, pois $P = 0,033626$, i.e., $P < \alpha$ (0,05) na comparação da média de N2 (3,952830) com a de N3 (4,329193).

Considerando-se $\alpha = 0,01$, tem-se os seguintes conclusões para os testes *post hoc*:

- Fisher's LSD: $\mu_2 \neq \mu_3$; $\mu_2 \neq \mu_4$;

- Scheffé: $\mu_1 = \mu_2 = \mu_3 = \mu_4$;
- Newman-Keuls: $\mu_1 = \mu_2 = \mu_3 = \mu_4$;
- Duncan: $\mu_1 = \mu_2 = \mu_3 = \mu_4$;
- Tukey HSD: $\mu_1 = \mu_2 = \mu_3 = \mu_4$;
- Tukey HSD ($\neq N$): $\mu_1 = \mu_2 = \mu_3 = \mu_4$.

Considerando-se $\alpha = 0,05$:

- Fisher's LSD: $\mu_1 \neq \mu_3$ e μ_4 ; $\mu_2 \neq \mu_3$ e μ_4 ;
- Scheffé: $\mu_1 = \mu_2 = \mu_3 = \mu_4$;
- Newman-Keuls: $\mu_1 \neq \mu_4$; $\mu_2 \neq \mu_3$ e μ_4 ;
- Duncan: $\mu_1 \neq \mu_4$; $\mu_2 \neq \mu_4$;
- Tukey HSD: $\mu_2 \neq \mu_3$ e μ_4 ;
- Tukey HSD ($\neq N$): $\mu_1 = \mu_2 = \mu_3 = \mu_4$.

Esses resultados gerais apontam para a existência de uma diferença significativa entre a média de N2 em relação tanto à de N3 como à de N4, pois essa diferença aparece no teste Fisher's LSD para $\alpha = 0,01$ e nos testes Fisher's LSD, Newman-Keuls, Duncan e Tukey HSD para $\alpha = 0,05$, suficientemente confirmando H_1 . Por outro lado, H_0 se confirma na comparação entre a média de N1 e a de N2. Tem-se, portanto, uma progressão de médias do tipo platô - rampa - platô, na qual $\mu_1 = \mu_2 \neq \mu_3 = \mu_4$.

Com isso conclui-se a exposição dos resultados obtidos a partir da coleta de dados junto aos sujeitos. No próximo capítulo serão descritos os procedimentos e os resultados da simulação conexionista.

CAPÍTULO 3 - A SIMULAÇÃO

A segunda parte do experimento consistiu de uma simulação conexionista visando a analisar, através do modelo construído e do treinamento a que esse modelo foi submetido, aspectos relativos à produção de perguntas em L2. Serão apresentados, nas seções a seguir, os detalhes da construção e do treinamento do modelo elaborado para os propósitos dessa investigação, assim como os resultados obtidos através desses procedimentos.

3.1 A CODIFICAÇÃO DOS PADRÕES INTERROGATIVOS DO INGLÊS

Para que a rede conexionista utilizada na simulação da produção de perguntas em L2 pudesse ser treinada e aprender, foi necessário que suas entradas e saídas esperadas fossem codificadas vetorialmente. Foram utilizadas bases de codificação distintas para as entradas e para as saídas da rede.

3.1.1 As Entradas

A codificação das entradas partiu de uma estrutura de frase do tipo:

$S \rightarrow NP + VP$

Onde:

$VP \rightarrow VPaux + VPmain$

Portanto:

$S \rightarrow NP_{suj} + VPaux + VPmain$

Esses sintagmas foram então codificados de acordo com as características dos seus núcleos:

- NP_{suj} foi codificado com base nas categorias Pessoa (subcategorias 1^a, 2^a, 3^a) e Número (subcategorias Singular, Plural);
- $VPaux$ foi codificado com base nas categorias utilizadas em NP_{suj} , de modo que a concordância entre esses sintagmas pudesse ser representada, acrescidas da categoria Tempo (subcategorias Presente, Passado, Futuro);
- $VPmain$ foi codificado a partir da categoria Tempo / Concordância com $VPaux$ (subcategorias Presente, Passado, Particípio Presente, Particípio Passado e Infinitivo).

Assim, uma frase do tipo '*He is talking*' (Ele está conversando), por exemplo, foi representada como:

$S \rightarrow 3^a, \text{ sing. } (NP_{suj}) + 3^a, \text{ sing., pres. } (VPaux) + \text{ part. pres. } (VPmain)$

Foram construídas, a partir daí, representações vetoriais binárias para cada padrão de entrada da rede, nas quais 1 representou a presença de uma dada subcategoria e 0, a ausência da mesma. Todas as codificações estão listadas no Anexo D. Dado o tipo de representação utilizada, a entrada incluiu três módulos de unidades de entrada (NP_{suj} , $VPaux$ e $VPmain$), somando um total de 18 unidades, com a seguinte composição:

- 5 unidades representando NP_{subj} (3 para Pessoa e 2 para Número);
- 8 unidades representando VP_{aux} (3 para Pessoa, 2 para Número e 3 para Tempo);
- 5 unidades representando VP_{main} (5 para Tempo / Concordância).

A rede recebeu, portanto, vetores de tamanho 18 como entradas, compostos da seqüência 5 – 8 – 5, correspondendo a NP_{subj} – VP_{aux} – VP_{main}.

Abaixo estão listadas algumas observações importantes em relação à codificação utilizada para as entradas da rede:

- Quando as entradas da rede representaram frases com ausência de elementos na posição VP_{aux}, como, por exemplo, ‘*You like _____*’ (Você gosta de _____), foram atribuídos valores 0 apenas às três unidades finais, referentes à subcategoria Tempo, da seqüência de 8 unidades de VP_{aux}, de modo que se preservasse o fator concordância de NP com VP (5 primeiras unidades da seqüência de VP_{aux}), independentemente de VP ser do tipo VP_{aux} ou VP_{main};
- As representações da categoria Tempo foram baseadas na morfologia da flexão do verbo auxiliar em inglês, ou seja, a frase ‘*They are playing _____*’ (Eles estão jogando _____) foi marcada, em VP_{aux}, como presente;
- Quando as entradas representaram frases modais como, por exemplo, ‘*You can dance*’ (Você pode dançar), foi utilizada uma marcação temporal múltipla: presente, passado e futuro receberam valor 1⁶.

3.1.2 As Saídas

⁶ Optou-se por esse tipo de representação uma vez que foi utilizada uma categoria geral de saída (*mod*) para as formas modais do inglês e elas, no conjunto, englobam essas três categorias temporais.

A codificação das saídas da rede utilizou uma estrutura frasal representando o deslocamento, no inglês, do verbo auxiliar para a esquerda de NP_{suj}:

$$S \rightarrow \text{VPaux} + \text{NP}_{\text{suj}} + \text{VPmain}$$

As categorias utilizadas para a representação dos padrões interrogativos, detalhadas no Anexo E, foram as seguintes:

- VPaux → Foram utilizadas as formas verbais *am, are, is, was, were, do, does, did, have, has, had* e uma categoria modal geral, *mod*;
- NP_{suj} → Foram utilizadas as formas pronominais *I, you, he* (englobando *she e it*), *we, they*, representando tanto elas mesmas quanto o universo de NPs que elas podem substituir;
- VPmain → Empregou-se as representações *st (stem), st + ed, irr (irregular) e st + ing*, significando, respectivamente, um radical verbal qualquer, um radical verbal qualquer acrescido da desinência *ed* do inglês (marca do passado e do particípio passado regulares), todas as formas verbais irregulares do inglês, tanto do passado como do particípio passado, e um particípio presente qualquer.

Esses três grupos foram representados por três conjuntos de vetores binários de tamanho 12, 5 e 4, respectivamente, totalizando um vetor de 21 unidades. Foi utilizado o mesmo sistema de atribuição de valores das entradas, ou seja, as características presentes nas saídas receberam valor 1 e as ausentes, 0. A rede produziu, portanto, vetores de tamanho 21 em suas saídas, compostas da seqüência 12 – 5 – 4, correspondendo a VPaux – NP_{suj} – VPmain.

Abaixo encontram-se algumas observações importantes em relação à codificação utilizada para as saídas da rede:

- Foi utilizada uma categoria geral *mod* para todas as saídas que envolviam os verbos modais do inglês, ou seja, a rede contemplou o mapeamento sintático-morfológico de saídas e entradas no que tange à concordância *mod* – VPmain, mas ignorou diferenças de Tempo entre os diferentes integrantes da classe *mod*⁷;
- Como as saídas operaram ao nível da forma, foi utilizada apenas uma unidade na representação de *you* (2ª pessoa ou do singular ou do plural no inglês);
- Seguindo o tipo de orientação adotada em relação aos verbos modais, foi utilizada uma categoria geral *irr* para todos os participios passados irregulares do inglês⁸. Foi contemplada, portanto, a concordância VPaux – VPmain, mas foram ignoradas diferenças entre as diversas formas irregulares dos participios passados do inglês.

3.2 A REDE

Todas as simulações do presente estudo utilizaram o simulador de redes neurais *T-learn*, software desenvolvido por Kim Plunkett e Jeffrey Elman para a realização de simulações conexionistas. Foi criado um projeto base denominado *question.cf*, com a seguinte configuração:

CONFIGURAÇÃO DE QUESTION.CF
<p><i>NODES:</i> <i>nodes = 31</i> <i>inputs = 18</i> <i>outputs = 21</i> <i>output nodes are 11-31</i></p> <p><i>CONNECTIONS:</i> <i>groups = 0</i> <i>1-10 from i1-i18</i> <i>11-31 from 1-10</i> <i>1-31 from 0</i></p>

⁷ Essa opção foi adotada por causa de restrições no tamanho de rede comportado por *T-learn*.

⁸ Essa opção, assim como no caso dos modais, foi adotada por causa de restrições no tamanho de rede comportado por *T-learn*.

SPECIAL:
selected = 1-10
weight_limit = 1.00

Quadro 3: Configuração da Rede

Essa rede é um auto-associador que possui 18 unidades de entrada, 21 unidades de saída, uma unidade de tendência (*bias*) e 10 unidades intermediárias, todas densamente interconectadas (Figura 7).

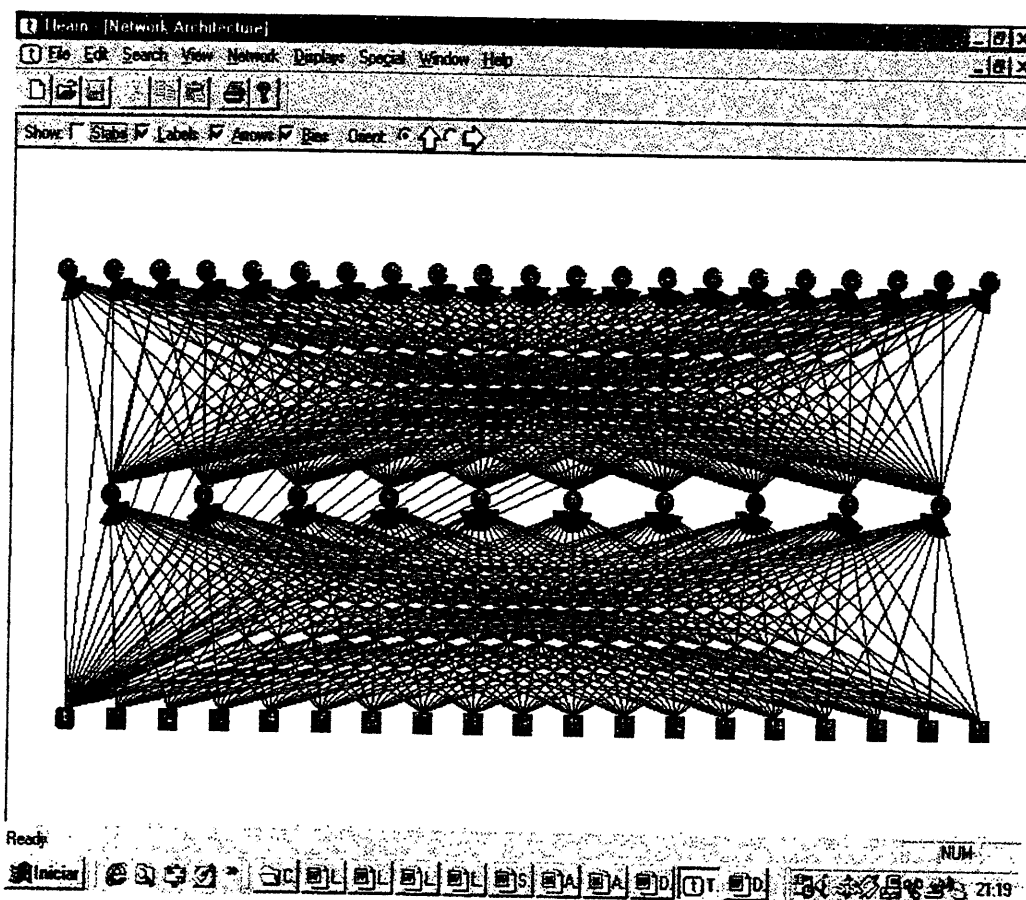


Figura 7: Esquema da rede

Esse tipo de configuração foi inspirada na rede utilizada por Plunkett & Juola (1999) em sua simulação conexionista da aquisição da morfologia do passado e do plural no inglês como língua materna. As unidades de entrada da rede estão

conectadas às unidades intermediárias e essas, às unidades de saída. A unidade *bias*⁹ está conectada a todas as unidades intermediárias e de saída da rede.

3.3 O TREINAMENTO

Para que fosse possível comparar os resultados obtidos a partir dos sujeitos com os resultados obtidos pelo simulador *conexionista*, idealizou-se um regime de treinamento baseado nas características do tipo de instrução formal que os sujeitos receberam em sala de aula. Treinou-se a rede de modo incremental, ou seja, grupos cumulativos de padrões interrogativos em seqüência foram a ela apresentados, cada grupo sendo treinado individualmente. Com isso teve-se por objetivo simular a instrução por unidades na qual baseia-se o curso de inglês no qual os sujeitos estudavam (seção 2.2.1). Foi construído, a partir dessa seqüência, um conjunto de treinamento formado por 7 tempos verbais base, seqüenciados de acordo com a progressão curricular geral dos sujeitos, totalizando 54 padrões interrogativos para o treinamento da rede. Esses padrões encontram-se listados no Quadro 4.

PADRÕES DE TREINAMENTO						
#	Afirmativo			Interrogativo		
	NPsubj	VPaux	VPmain	VPaux	NPsubj	VPmain
(A) PRESENT SIMPLE						
1	I	am	-	am	I	-
2	you	are	-	are	you	-
3	he, she, it	is	-	is	he, she, it	-
4	we	are	-	are	we	-
5	you	are	-	are	you	-
6	they	are	-	are	they	-
7	I	-	talk*	do	I	talk
8	you	-	talk	do	you	talk
9	he, she, it	-	talks	does	he, she, it	talk
10	we	-	talk	do	we	talk
11	you	-	talk	do	you	talk
12	they	-	talk	do	they	talk
(B) PRESENT CONTINUOUS						

⁹ A presença da unidade *bias* na configuração da rede neural faz com que suas unidades tenham ativações *default* individuais para iniciar o treinamento, ou seja, assegura que cada unidade possua um certo patamar de ativação. Essa unidade tem ativação constante igual a 1.

13	I	am	talking	am	I	talking
14	you	are	talking	are	you	talking
15	he, she, it	is	talking	is	he, she, it	talking
16	we	are	talking	are	we	talking
17	you	are	talking	are	you	talking
18	they	are	talking	are	they	talking
<i>(C) MODAL</i>						
19	I	can**	talk	can	I	talk
20	you	can	talk	can	you	talk
21	he, she, it	can	talk	can	he, she, it	talk
22	we	can	talk	can	we	talk
23	you	can	talk	can	you	talk
24	they	can	talk	can	they	talk
<i>(D) PAST SIMPLE</i>						
25	I	was	-	was	I	-
26	you	were	-	were	you	-
27	he, she, it	was	-	was	he, she, it	-
28	we	were	-	were	we	-
29	you	were	-	were	you	-
30	they	were	-	were	they	-
31	I	-	talked / spoke***	did	I	talk
32	you	-	talked / spoke	did	you	talk
33	he, she, it	-	talked / spoke	did	he, she, it	talk
34	we	-	talked / spoke	did	we	talk
35	you	-	talked / spoke	did	you	talk
36	they	-	talked / spoke	did	they	talk
<i>(E) PAST CONTINUOUS</i>						
37	I	was	talking	was	I	talking
38	you	were	talking	were	you	talking
39	he, she, it	was	talking	was	he, she, it	talking
40	we	were	talking	were	we	talking
41	you	were	talking	were	you	talking
42	they	were	talking	were	they	talking
<i>(F) PRESENT PERFECT</i>						
43	I	have	talked / spoken	have	I	talked / spoken
44	you	have	talked / spoken	have	you	talked / spoken
45	he, she, it	has	talked / spoken	has	he, she, it	talked / spoken
46	we	have	talked / spoken	have	we	talked / spoken
47	you	have	talked / spoken	have	you	talked / spoken
48	they	have	talked / spoken	have	they	talked / spoken
<i>(G) PAST PERFECT</i>						

49	I	had	talked / spoken	had	I	talked / spoken
50	you	had	talked / spoken	had	you	talked / spoken
51	he, she, it	had	talked / spoken	had	he, she, it	talked / spoken
52	we	had	talked / spoken	had	we	talked / spoken
53	you	had	talked / spoken	had	you	talked / spoken
54	they	had	talked / spoken	had	they	talked / spoken

(*) *talk* representa um verbo regular ou irregular não modal e não *to be* qualquer do inglês; (**) *can* representa um verbo modal qualquer do inglês; (***) *spoke* representa um verbo irregular qualquer, não modal, do inglês.

Quadro 4: Padrões de Treinamento

Esses padrões foram codificados na forma de vetores binários, de modo que pudessem ser introduzidos na rede construída no simulador. Para isso, foi utilizado o sistema de codificação descrito na seção 3.1. Foram realizados 7 estágios de treinamento e teste, correspondentes aos 7 grupos de tempos verbais relacionados acima. O treinamento foi realizado de forma incremental, ou seja, a rede foi treinada com um número crescente de grupos de padrões de entrada. O número de ciclos utilizados para cada estágio de treinamento dependeu da quantidade necessária para que se atingisse 100% de precisão nas saídas traduzidas com base na distância euclidiana entre os vetores. Para cada um dos estágios, quando esse ponto foi atingido, os pesos das conexões das unidades intermediárias eram salvos e utilizados como ponto de partida para o treinamento do grupo seguinte de padrões.

Primeiramente, foram criados 7 projetos para treinamento, denominados *t-groupX.cf*, nos quais $X = A, B, \dots, G$. Todos esses arquivos foram idênticos a *question.cf*, uma vez que eles representaram, metaforicamente, o mecanismo de processamento de um sujeito qualquer, considerado uma constante ao longo das fases de treinamento. Foram então criados 7 arquivos de treinamento, contendo as representações vetoriais de tamanho 18 dos arquivos de entrada, relativos aos sete grupos de padrões base (afirmativos) dos tempos verbais listados no Quadro 4. Esses arquivos foram denominados *t-groupX.data*. Respectivamente, foram criados 7 arquivos denominados *t-groupX.teach*, contendo os padrões base interrogativos correspondentes, i.e., as saídas esperadas de tamanho vetorial 21. Além desses, foram construídos, também, 7 arquivos de teste, denominados *groupY* ($Y = A, B, \dots$,

G), de extensão *.data* e *.teach*, contendo apenas os padrões relativos a cada um dos tempos verbais do Quadro 4.

Os arquivos de treinamento tiveram seu conteúdo construído em caráter cumulativo. Dessa forma, para $X = A$, $t\text{-group}A = \text{group}A$; para $X = B$, $t\text{-group}B = \text{group}A + \text{group}B$; para $X = C$, $t\text{-group}C = \text{group}A + \text{group}B + \text{group}C$, e assim sucessivamente.

O treinamento deu-se dentro dos parâmetros gerais exibidos no Quadro 5.

PARÂMETROS DE TREINAMENTO PARA T-GROUPA-G	
Número de Instâncias: Número de vezes que vetores quaisquer são apresentados às entradas da rede.	Esse número foi estabelecido como sendo 10 ou 15 vezes o número de padrões de entrada de <i>t-groupY.data</i> , dependendo no número necessário para se obter uma precisão de 100% nas respostas.
Semente Aleatória: Indica os valores iniciais dos pesos das conexões da rede. A utilização de uma mesma semente (o valor pode ser qualquer um) para um grupo de simulações garante que os valores iniciais dos pesos serão sempre idênticos e que os padrões de entrada serão sempre ativados na mesma ordem aleatória. Obs: A escolha dos valores dos pesos iniciais, apesar de aleatória, respeita sempre os limites estabelecidos em <i>question.cf</i> .	Foi escolhido o valor 1 para a primeira simulação (<i>t-groupA</i>). As demais (<i>t-groupB-G</i>) utilizaram valores dos treinamentos imediatamente anteriores, determinados pelo arquivo de pesos que tenha conseguido 100% de precisão na tradução das saídas (utilizando-se a distância euclidiana entre os vetores).
Regime: Periodicidade do ajuste dos pesos da rede.	Padrão (ajuste a cada instância).
Modo: Ordem de apresentação das entradas à rede.	Aleatório (com reposição, i.e., entradas utilizadas sempre estão disponíveis para serem reutilizadas).
Tipo: Maneira como se dá o aprendizado	Supervisionado (presença das saídas esperadas em <i>t-groupY.teach</i>).
Algoritmo: Programa de aprendizado utilizado pela rede.	<i>Backpropagation</i> .
Erro: Tipo de cálculo de erro utilizado pelo algoritmo de aprendizado	Entropia-X, registrada a cada instância de treinamento.
Taxa de Aprendizado: Taxa de ajuste, pelo algoritmo de aprendizado, dos	Utilização do valor 0.1 (<i>T-learn</i> oferece o intervalo de 0 – 10 para esse valor),

pesos das conexões da rede, em direção à saída desejada.	por permitir uma melhor visualização da progressão do aprendizado ao longo do treinamento.
<u>Momento</u> : Parâmetro que indica a proporção da mudança de pesos, efetuada na instância imediatamente anterior, a ser utilizada no ajuste corrente.	Utilizado o valor 0 (<i>T-learn</i> oferece valores entre 0 – 1 para esse parâmetro).
<u>Registro dos Pesos</u> : Arquivos criados por <i>T-learn</i> com os valores dos pesos das conexões em um dado momento do treinamento.	Arquivos criados a cada N instâncias, sendo N = número de padrões de entrada de <i>t-groupY.data</i>

Quadro 5: Parâmetros de Treinamento

As saídas produzidas por esse treinamento foram de natureza vetorial e registradas em intervalos regulares. Para que se pudesse interpretar linguisticamente esses registros, os vetores correspondentes às saídas obtidas foram codificados através do arquivo de mapeamento *mapp.out*, configurado da seguinte forma:

CONFIGURAÇÃO DE MAPP.OUT												
MAPPINGS:												
1-12 from VA												
13-17 from NP												
18-21 from VM												
VA:												
a	1	0	0	0	0	0	0	0	0	0	0	0
b	0	1	0	0	0	0	0	0	0	0	0	0
c	0	0	1	0	0	0	0	0	0	0	0	0
d	0	0	0	1	0	0	0	0	0	0	0	0
e	0	0	0	0	1	0	0	0	0	0	0	0
f	0	0	0	0	0	1	0	0	0	0	0	0
g	0	0	0	0	0	0	1	0	0	0	0	0
h	0	0	0	0	0	0	0	1	0	0	0	0
i	0	0	0	0	0	0	0	0	1	0	0	0
j	0	0	0	0	0	0	0	0	0	1	0	0
k	0	0	0	0	0	0	0	0	0	0	1	0
l	0	0	0	0	0	0	0	0	0	0	0	1
m	0	0	0	0	0	0	0	0	0	0	0	0
NP:												
n	1	0	0	0	0							
o	0	1	0	0	0							

p	0	0	1	0	0
q	0	0	0	1	0
r	0	0	0	0	1
s	0	0	0	0	0
VM:					
t	1	0	0	0	
u	0	1	0	0	
v	0	0	1	0	
w	0	0	0	1	
x	0	0	0	0	

Quadro 6: Arquivo de Mapeamento

O Quadro 6 apresenta, primeiramente, a seção *MAPPING* (mapeamento), contendo os três subgrupos utilizados para a codificação das saídas, a saber, VA (VPaux), com 12 algarismos, NP (NPsuj), com 5, e VM (VPmain), com 4, totalizando os 21 algarismos a serem mapeados. A seguir, tem-se a seqüência vetorial correspondente a cada saída possível (segundo o Anexo E) para cada um dos subgrupos, codificadas por um caractere ASC. Finalmente, para que as saídas codificadas por *mapp.out* pudessem ser visualizadas em sua forma lingüística, uma vez que *T-learn* só permite a utilização de um caractere ASC para esse tipo de mapeamento, foi criado um arquivo de recodificação denominado *tradu*. Esse arquivo permitiu a recodificação dessas saídas com até dois caracteres ASC:

TRADUÇÃO DAS SAÍDAS POR TRADU					
VPaux		NPsuj		VPmain	
ASC	Saída	ASC	Saída	ASC	Saída
a	am (<i>am</i>)	n	I (<i>I</i>)	t	st (<i>talk</i>)
b	ar (<i>are</i>)	o	yu (<i>you</i>)	u	ed (<i>talked</i>)
c	is (<i>is</i>)	p	he (<i>he, she, it</i>)	v	ir (<i>spoken</i>)
d	ws (<i>was</i>)	q	we (<i>we</i>)	w	ng (<i>talking</i>)
e	wr (<i>were</i>)	r	th (<i>they</i>)	x	-
f	do (<i>do</i>)	s	-		
g	ds (<i>does</i>)				
h	dd (<i>did</i>)				
i	hv (<i>have</i>)				
j	hs (<i>has</i>)				
k	hd (<i>had</i>)				

l	md (<i>can</i>)
m	-

Quadro 7: Codificação Utilizada na Tradução das Saídas

O Quadro 7 apresenta os três subgrupos de codificação das saídas (VPaux, NPsubj e VPmain) em seqüência, da esquerda para a direita. Cada subgrupo está subdividido em duas colunas: a da esquerda, contendo os caracteres produzidos por *mapp.out*, e a da direita, contendo a recodificação, em até dois algarismos, desses caracteres, seguida do elemento lingüístico que elas representam (entre parênteses, de acordo com a representação do Quadro 4). Dessa forma, seqüências binárias que foram mapeadas a partir de *mapp.out* foram traduzidas, através de *tradu*, para seu padrão interrogativo de referência:

(1) Ex.: (0 0 1 0 0 0 0 0 0 0 0) + (0 0 1 0 0) + (1 0 0 0) = cpt, onde cpt = IS + HE + ST (c = is, p = he, t = st) (Grupo A).

3.4 A PONTUAÇÃO DAS SAÍDAS

Traduzidas as saídas, elas foram pontuadas em relação à sua conformidade aos padrões de treinamento, com valores oscilando entre 0 e 3. Cada ponto atribuído correspondeu a um elemento correto na posição de saída em relação ao padrão esperado. Dessa maneira:

(2) Ex.: Saída Real (DO + WE + ST) x Saída Esperada (MD + WE + ST)
0 + 1 + 1 = 2 pontos (Grupo C);

(3) Ex.: Saída Real (DO + ____ + ST) x Saída Esperada (DO + I + ST)
1 + 0 + 1 = 2 pontos (Grupo A);

(4) Ex.: Saída Real (____ + WE + ____) x Saída Esperada (AM + I + ____)
0 + 0 + 1 = 1 ponto (Grupo A);

(5) Ex.: Saída Real ($HV + YU + ED$) x Saída Esperada ($HV + YU + ED$)
 $1 + 1 + 1 = 3$ pontos (Grupo F);

(6) Ex.: Saída Real ($AR + WE + ___$) x Saída Esperada ($DO + YU + ST$)
 $0 + 0 + 0 = 0$ ponto (Grupo A).

3.5 A ANÁLISE DE AGRUPAMENTO

A segunda ferramenta de avaliação de resultados utilizada neste estudo foi a *Análise de Agrupamento (CA)*. Como foi salientado na seção 1.2, essa é uma medição realizada após o treinamento, quando se apresentam alguns vetores de teste à rede. Esses padrões produzem ativações nas unidades intermediárias, que são gravadas e rotuladas. Sendo os padrões das unidades intermediárias vetores em um espaço multi-dimensional, a CA calcula as distâncias euclidianas entre eles e apresenta, sob a forma de um diagrama arbóreo (Anexo G), os agrupamentos hierárquicos que se formam ao longo do treinamento. Dessa forma, pôde-se observar tanto a aproximação como o distanciamento entre si dos vetores compostos pelos valores das ativações das unidades intermediárias após novos padrões interrogativos do conjunto de treinamento haverem sido introduzidos em regime incremental.

A progressão do agrupamento dos vetores ativação das unidades intermediárias auxilia na investigação de como a rede resolveu o problema de aprendizado a ela proposto, pois revela que características essenciais das entradas ela utilizou para a produção de saídas corretas. Ao hierarquizar suas configurações internas, a rede revela seu critério para tal.

3.6 AS FREQUÊNCIAS DE OCORRÊNCIA

Com o objetivo de investigar a influência dos fatores frequência relativa de ocorrência e variância na formação das configurações internas da rede, foram calculadas as frequências relativas (fr) médias de ocorrência e as variâncias para todos os elementos pertencentes às categorias VPaux, NPsubj e VPmain dos padrões

que compuseram o conjunto de treinamento em cada um dos seus sete estágios. Como o treinamento da rede foi realizado em regime incremental, para cada estágio de treinamento foram, portanto, calculados os valores de fr em relação aos grupos de padrões que o compuseram.

(7) Ex.: Estágio C, fr (ARE) = 8 ocorrências / 24 padrões = 0,33333.

Os valores de fr de ocorrência assim calculados equivaleram, não obstante, a frequências prováveis, pois, no regime de treinamento adotado, os padrões foram apresentados à rede em ordem aleatória.

3.7 A HARMONIA

Com o objetivo de dispor de mais um parâmetro de análise, foi calculada a harmonia (H) da rede para o padrão A8 [DO + YOU + TALK] – o mais recorrente nas produções dos sujeitos – em cada uma das sete fases de treinamento da rede. A análise da progressão desse valor ao longo do treinamento fornece uma medida do grau de consonância das ativações das unidades intermediárias da rede em relação a A8 com as restrições impostas pelos valores dos pesos das conexões ao final do treinamento. Nesse caso, quanto maior o valor de H, maior a consonância.

Essas restrições, segundo Prince & Smolensky (1997), equivalem a restrições de gramaticalidade no universo das estruturas com as quais a rede entrou em contato. Ou seja, é no ajuste dos pesos da rede que ficam codificadas as restrições de boa formação que subjazem ao conjunto de treinamento.

Como já foi mencionado na seção 1.6, o cálculo da harmonia de uma rede para um dado padrão de entrada é dado por:

$$H = \sum_{ij} a_i \cdot w_{ij} \cdot a_j$$

Onde a_i é o valor de ativação da unidade i , a_j é o valor da ativação da unidade j e w_{ij} é o valor do peso da conexão entre essas unidades. Em uma rede contendo várias unidades, o valor de H para um dado padrão é a soma (Σ) de todas as micro-harmonias computadas para todos os pares de unidades conectadas. Foram, portanto, calculadas e somadas, para o padrão A8, as micro-harmonias das 421 conexões da rede construída em *T-learn*.

O cálculo foi executado da seguinte maneira:

- Em cada estágio do treinamento, foi salvo o vetor dos valores das ativações relativas ao padrão A8 nas 10 unidades intermediárias da rede;
- Para a camada de entrada, foi utilizado, para cada um dos sete conjuntos de treinamento, o valor binário do vetor de codificação de entrada do padrão A8, que equivale ao valor das ativações para esse padrão, uma constante;
- Para a camada de saída, foi utilizado, para cada um dos sete conjuntos de treinamento, o vetor de saída relativo a A8 obtido ao término do total de 55 ciclos do treinamento (2.249 instâncias), também uma constante.

Dessa forma, a única variável no cálculo de H foi o valor das ativações das unidades intermediárias, pois tanto os valores das ativações de entrada como de saída foram mantidos constantes.

3.8 OS RESULTADOS

Nesta seção serão apresentados os resultados obtidos através dos registros realizados nas saídas da rede, na progressão dos diagramas da análise de agrupamento e na medição dos níveis de harmonia da rede ao longo do treinamento.

3.8.1 As Saídas e suas Pontuações

O treinamento do primeiro grupo de vetores (*Present Simple*) foi realizado com 15 ciclos (180 instâncias, i.e., 15 vezes os 12 padrões contidos em *t-groupA.data*), pesos iniciais aleatórios (semente =1) e registro de pesos a cada 12 instâncias. Ao término do treino, a rede foi testada com os padrões contidos em *groupA* (nesse caso, como esse foi o primeiro teste, $groupA = t-groupA$) e com os arquivos de pesos registrados durante o treinamento (um registro por ciclo), até que fosse obtido 100% de precisão nas saídas. Isso foi alcançado às 145 instâncias.

O treinamento do segundo grupo de vetores (*Present Continuous*) foi realizado com 10 ciclos (180 instâncias, i.e., 10 vezes os 18 padrões contidos em *t-groupB.data*, com $t-groupB = groupA + groupB$), pesos iniciais contidos em *t-groupA-145.wts* e registro de pesos a cada 18 instâncias. Ao término do treino, a rede foi testada com os padrões contidos em *groupB*, utilizando os arquivos de pesos registrados durante o treinamento (um por ciclo, i.e., 18 instâncias), até que fosse obtido 100% de precisão nas saídas. Isso foi alcançado às 218 instâncias.

Essa lógica de procedimentos foi, então, repetida para os demais grupos de treinamento, de acordo com a tabela abaixo:

Etapas de treinamento		Número de ciclos total do treinamento	Número de instâncias para cada registro de pesos	Número de instâncias para precisão de 100%
A	<i>Present Simple</i> (12 padrões)	15	12	145
B	<i>Present Continuous</i> (6 padrões)	10	18	218
C	<i>Modal</i> (6 padrões)	10	24	291
D	<i>Past Simple</i> (12 padrões)	10	36	651
E	<i>Past Continuous</i> (6 padrões)	10	42	862
F	<i>Present Perfect</i> (6 padrões)	15	48	1.439
G	<i>Past Perfect</i> (6 padrões)	10	54	1.764

Na Tabela 6 estão listados, na coluna mais à esquerda, os sete estágios de treinamento, os grupos de padrões interrogativos do inglês que foram adicionados a

cada uma dessas etapas, seguidos do tempo verbal a que se referem, e o número de padrões diferentes (6 ou 12) de cada um desses grupos. A segunda coluna apresenta o número total de ciclos de treinamento (1 ciclo = número de padrões utilizados nesse estágio) ao qual cada um dos grupos foi submetido. Nota-se que os grupos A (*Present Simple*) e F (*Present Perfect*) foram os únicos que necessitaram de 15 ciclos para que a precisão de 100% fosse atingida. A terceira coluna mostra o número de padrões utilizados em cada uma das sete etapas, de acordo com o regime acumulativo descrito na seção 3.3. O registro das saídas se deu, para cada um dos estágios, em intervalos iguais a esse número de padrões. A quarta e última coluna registra o número de instâncias no qual congelaram-se os valores dos pesos das conexões da rede, pois a precisão de 100% fora atingida. Esses pesos foram utilizados como pontos de partida para os estágios seguintes de treinamento.

Pode-se calcular, a partir dos valores das duas últimas colunas da Tabela 6, o nível de dificuldade para se atingir 100% de precisão nas saídas ao longo do treinamento, dividindo-se a diferença dos valores da quarta coluna em relação aos seus antecedentes (número de instâncias necessárias, naquela etapa, para a obtenção de 100% de precisão) pelos valores da terceira coluna (número de padrões utilizados na etapa).

(8) Ex.: Grupo E ($862 - 651 = 211 / 42 = 5,02$)

A Figura 8 exibe a progressão dessa dificuldade. Observa-se um declínio acentuado de A para B e um declínio mais ameno de B para C, havendo, a partir daí, uma oscilação de valores até a etapa final, G. O gráfico também mostra que há uma elevação do grau de dificuldade nas transições de C para D (*Modal* → *Past Simple*) e de E para F (*Past Continuous* → *Present Perfect*).

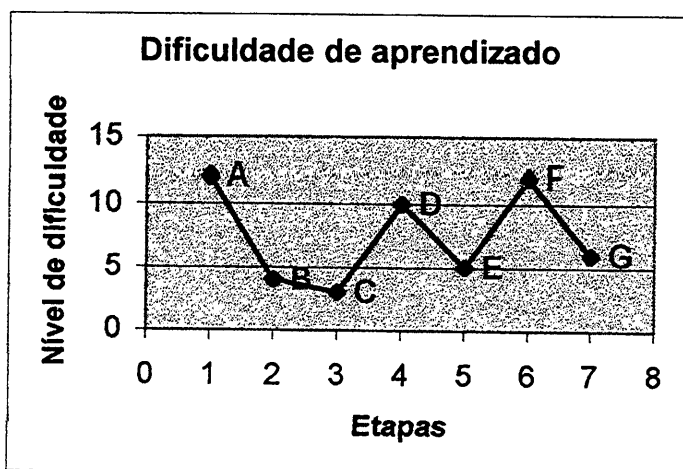


Figura 8: Nível de Dificuldade de Aprendizado por Etapa

O Anexo F apresenta, em detalhe, as saídas produzidas pela rede durante o treinamento, as quais foram registradas nos intervalos definidos na Tabela 6. Essas saídas foram pontuadas segundo os critérios apresentados na seção 3.3 (pontuação variando de 0 a 3). As médias das pontuações dessas saídas, por grupo de treinamento, estão representadas nas Figura 9. A Figura 10 apresenta os valores dos seus desvios padrões.

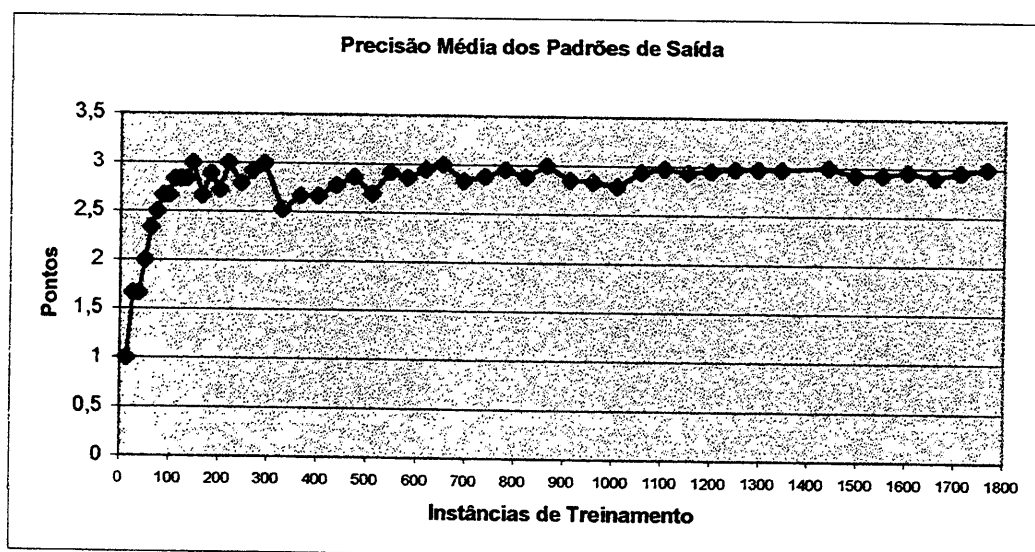


Figura 9: Precisão Média dos Padrões de Saída

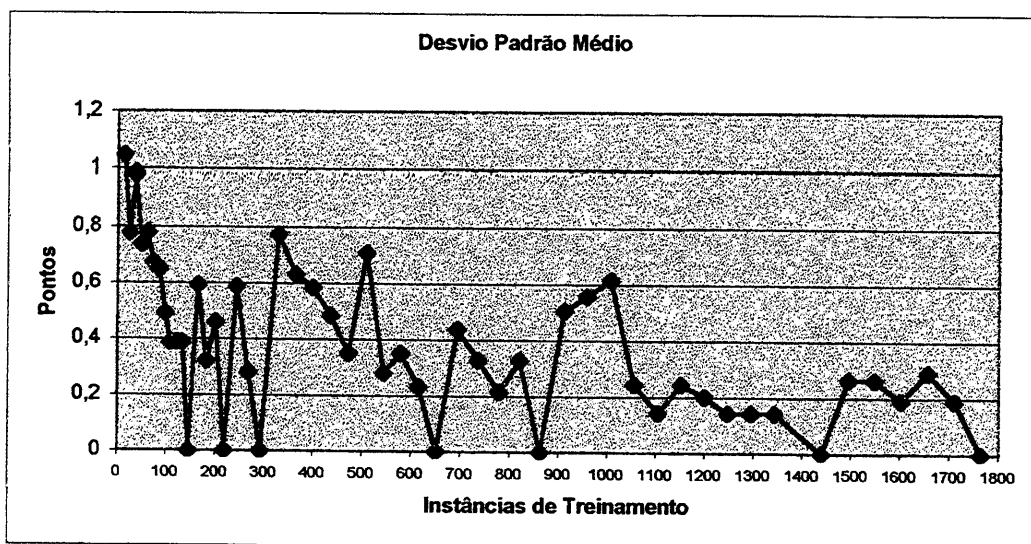


Figura 10: Desvio Padrão Médio das Saídas

De um modo geral, nota-se que há um aumento não linear da precisão média das produções da rede ao longo do treinamento, acompanhado de uma diminuição gradual, mas bastante oscilante, do desvio padrão. A precisão também apresenta oscilações, porém menos acentuadas. Observa-se, portanto, que há a presença do fenômeno do desaprendizado, ou seja, há quedas na média de precisão durante o treinamento, indicando que a rede, muitas vezes, acerta, erra e acerta novamente. Em relação às transições entre os estágios de treinamento, nota-se uma maior oscilação da precisão durante as fases iniciais de treinamento: Grupo A (145 instâncias) → Grupo B (218 instâncias), Grupo B (218 instâncias) → Grupo C (291 instâncias) e Grupo C (291 instâncias) → Grupo D (651 instâncias).

3.8.2 A Análise de Agrupamento

Foi realizada, também, a análise de agrupamento (CA) das configurações internas das unidades intermediárias, relativa a todos os padrões do conjunto de treinamento. Como já foi salientado na seção 1.2, a CA é um agrupamento hierárquico baseado nas distâncias euclidianas entre os vetores ativação das unidades intermediárias da rede. Os padrões que figuram mais próximos no diagrama da CA

são os com menor distância entre si, ou seja, são os que têm maior grau de similaridade entre si no que tange às representações internas (ativações) da rede.

Os diagramas arbóreos finais relativos aos 7 estágios de treinamento estão listados no Anexo G. Observa-se, na progressão espacial apresentada, que o principal fator de agrupamento dos padrões vetoriais, de um modo geral, foi o tipo de elemento na posição NPsuj. Portanto, ao longo das sete fases de treinamento, os grupos de ramos adjacentes da CA se formaram, majoritariamente, em torno das seis pessoas do inglês que ocuparam a posição NPsuj. Somente no estágio final (G), houve uma aglomeração cujo fator principal de união dos padrões foi VPmain (particípio passado, ED ou IR).

Especial atenção foi dedicada, no Anexo G, à progressão de A8 no espaço vetorial multidimensional. Para uma melhor visualização do processo, ao lado de cada um dos diagramas (Grupos A, B, C, D, E, F e G) encontram-se listados, na mesma ordem vertical em que figuram no diagrama, os padrões espacialmente mais próximos de A8, ou seja, pertencentes a ramos adjacentes. Seguindo a lógica geral, A8 apareceu, nos diagramas, circundado pelos outros padrões NPsuj na segunda pessoa (tanto do plural como do singular, pois nesse caso há coincidência de forma, YOU). Refinando-se a análise através da inclusão de fatores secundários de agrupamento, se observa, a partir da progressão apresentada no Anexo G, a seguinte ordem de influência na hierarquização das distâncias euclidianas no espaço vetorial (em ordem decrescente):

- NPsuj → VPaux → VPmain

NPsuj, portanto, é o fator de aglomeração mais influente ao longo do treinamento, seguido de VPaux, em segundo, e VPmain, em terceiro. No estágio G, contudo, houve, como já se ressaltou, o afastamento dos padrões contendo ED e IR (fases F e G) da aglomeração em torno de A8. Esses padrões formaram um grupo cuja hierarquia interna foi invertida para a ordem VPmain → VPaux → NPsuj.

3.8.3 As Frequências Relativas de Ocorrência

A Figura 11 apresenta a progressão dos valores das frequências relativas médias de ocorrência para as categorias NP_{suj}, VP_{aux} e VP_{main}, ao longo das sete fases de treinamento às quais a rede foi submetida.

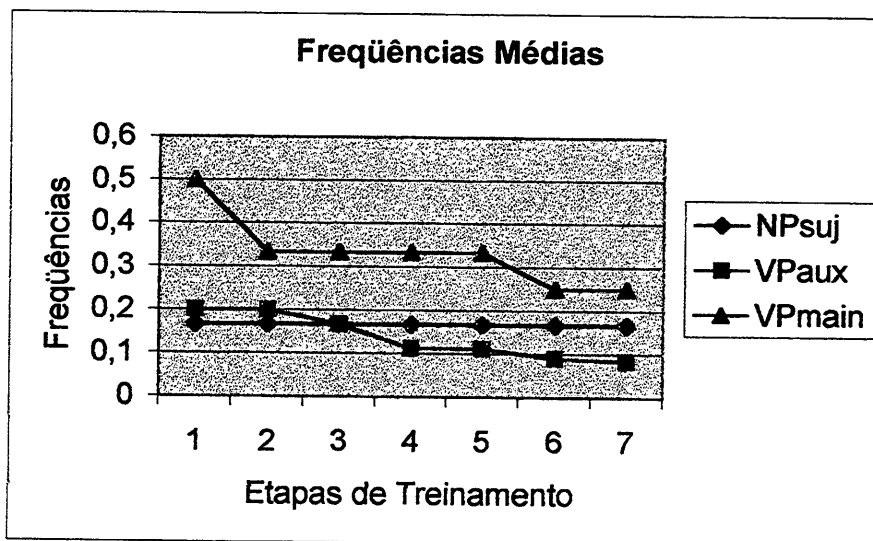


Figura 11: Frequências Médias de Ocorrência

Observa-se, na Figura 11, que VP_{main} apresenta os valores mais elevados de fr ao longo de todo o treinamento. Esse padrão também apresenta a menor regularidade de frequência, pois claramente sofre as oscilações mais bruscas. Já NP_{suj} apresenta oscilação nula em seus valores, ocupando o último lugar em grandeza até o estágio 3 (fase C), quando passa a ocupar o segundo lugar. As frequências de VP_{aux} oscilam moderadamente e, a partir do estágio 3, assumem a última posição em termos de grandeza de valores de frequência.

Em relação aos resultados obtidos na CA (seção 3.5), observa-se que os valores médios de fr de ocorrência para as categorias NP_{suj}, VP_{aux} e VP_{main} não correspondem fielmente à hierarquia geral de aglomeração revelada pela CA. Nota-se que a maior valor de frequência média de ocorrência é sempre a de VP_{main}, havendo a inversão de superioridade de valor entre NP_{suj} e VP_{aux} a partir da etapa 3 (fase C).

A Figura 12 mostra o valor das variâncias médias para cada uma das sete fases de treinamento. Observa-se que NPsuj teve variância média nula ao longo de todo o treinamento, sendo, portanto, sempre o menor valor. Novamente, a etapa 3 (fase C) representou uma inversão de predominância de valores, na qual os valores da variância média de VPmain ultrapassaram os de VPaux. Nota-se também, a partir desse ponto, uma relação exatamente inversa à escala de nível de influência de categorias que resultou da CA. No geral, as menores variâncias foram as dos elementos pertencentes à categoria NPsuj (excetuadas as fases A e B, nas quais esses valores foram iguais aos valores dos elementos de VPmain). As maiores variâncias, por sua vez, encontram-se na categoria VPmain, ficando os valores dos elementos de VPaux no meio termo.

A partir da fase C, portanto, a variância média parece ter relação inversa com a escala de nível de influência de agrupamento revelada pela CA. Observa-se, também, que na fase G os valores de variância média de VPmain, VPaux e NPsuj são bastante próximos, o que coincide com a formação, na CA, de um grupo isolado pelos padrões do tipo *Perfect*.

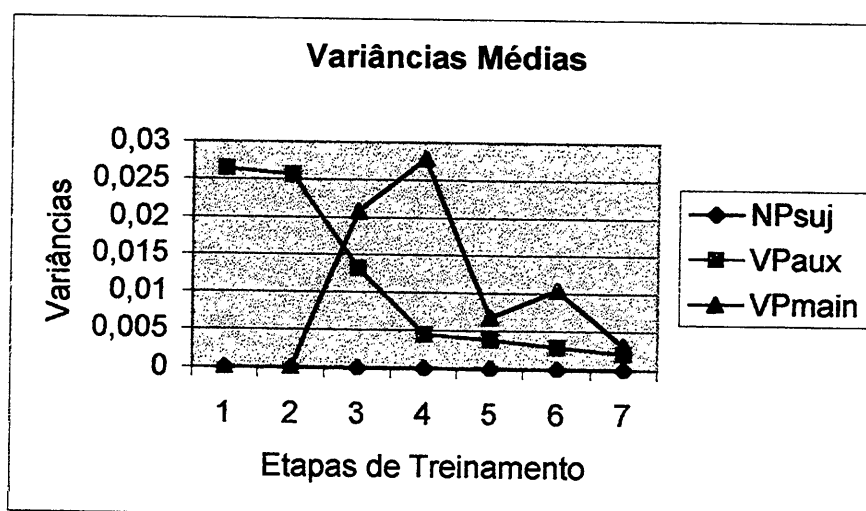


Figura 12: Variâncias Médias

A Figura 13 apresenta a evolução dos valores da frequência relativa de ocorrência, no conjunto de treinamento, para os elementos que compõem o padrão

A8 [DO + YOU(2PS) + TALK]. Observa-se que os valores relativos a TALK são, ao longo de todo o treinamento, os mais elevados. Nota-se, também, que não há variação nas freqüências de ocorrência de YOU(2PS), que aparece como o menor valor até a etapa 4 (D), na qual o valor de DO, que vinha decrescendo, passa a ser o menor de todos. Há, portanto, uma maior regularidade, i.e., uma menor oscilação dos valores de freqüência de YOU(2PS) ao longo do treinamento. Quanto a DO, se observa que, apesar de esse elemento apresentar a maior variação total (do estágio 1 para o estágio 7), suas taxas de variação (inclinação da curva) de um estágio para o outro são ou iguais ou menores do que as de TALK, com exceção da passagem do estágio 3 (fase C) para o estágio 4 (fase D), na qual TALK não varia. TALK é, claramente, o elemento que sofre as variações mais bruscas de freqüência ao longo do treinamento, ou seja, apresenta menor regularidade geral de freqüência.

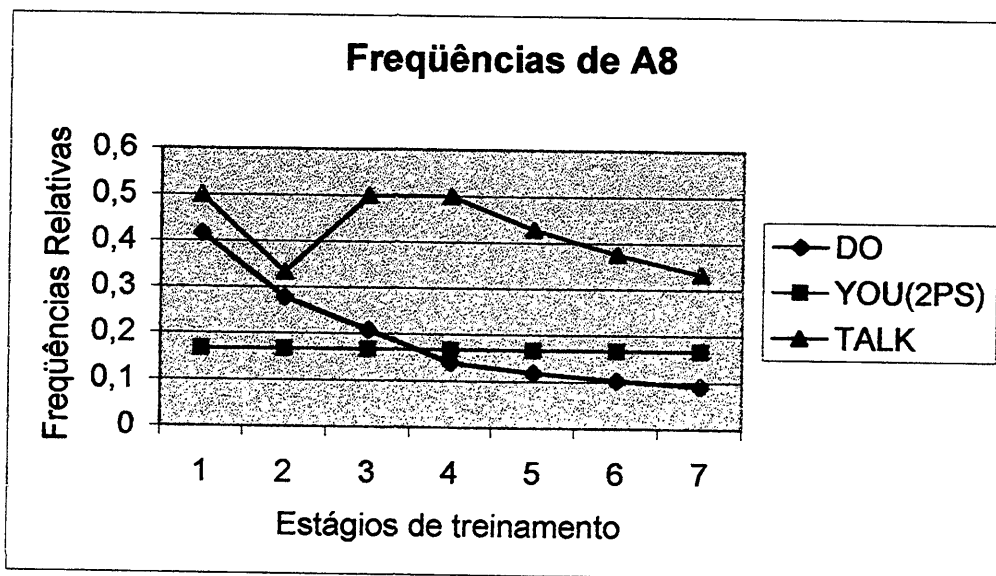


Figura 13: Freqüências Relativas do Padrão A8

A Figura 14 apresenta a progressão das variâncias de A8 ao longo do treinamento. Nota-se a completa estabilidade do valor da variância de YOU(2PS) em 0, enquanto que esse valor, para DO, começa bastante elevado, caindo acentuadamente do estágio 1 para o 2 (A → B) e continuando a decrescer até praticamente anular-se no estágio 5 (E), se igualando a YOU(2PS). TALK apresenta variância nula nos estágios 1 e 2, assumindo, a partir daí, os valores mais elevados de

variância e apresentando as maiores oscilações. O estágio 3 (fase C) estabelece uma mudança na ordem de grandeza entre YOU(2PS), DO e TALK. A partir da fase C, TALK passa a ser o elemento a exibir os valores mais elevados, substituindo DO.

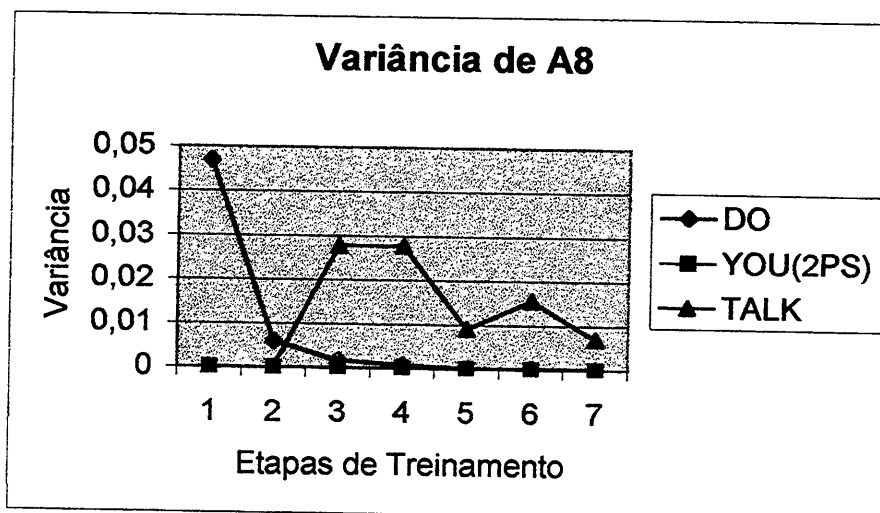


Figura 14: Variância de A8

O resultado geral dos ordenamentos (hierárquicos e de frequência relativa de ocorrência) entre os elementos pertencentes a NP_{suj}, VP_{aux} e VP_{main}, em relação a A8 encontram-se resumidos na Tabela 7.

TABELA 7

RESULTADOS GERAIS DE ORDENAMENTO PARA A8

Fases	Ordem decrescente de frequência	Ordem decrescente de variância	Ordem decrescente de nível de influência na CA
A	VP _{main} → VP _{aux} → NP _{suj}	VP _{aux} → VP _{main} = NP _{suj}	NP _{suj} → VP _{aux} → VP _{main}
B	VP _{main} → VP _{aux} → NP _{suj}	VP _{aux} → VP _{main} = NP _{suj}	NP _{suj} → VP _{aux} → VP _{main}
C	VP _{main} → VP _{aux} → NP _{suj}	VP _{aux} → VP _{main} = NP _{suj}	NP _{suj} → VP _{aux} → VP _{main}
D	VP _{main} → VP _{aux} → NP _{suj}	VP _{main} → VP _{aux} → NP _{suj}	NP _{suj} → VP _{aux} → VP _{main}
E	VP _{main} → NP _{suj} → VP _{aux}	VP _{main} → VP _{aux} → NP _{suj}	NP _{suj} → VP _{aux} → VP _{main}
F	VP _{main} → NP _{suj} → VP _{aux}	VP _{main} → VP _{aux} → NP _{suj}	NP _{suj} → VP _{aux} → VP _{main}
G	VP _{main} → NP _{suj} → VP _{aux}	VP _{main} → VP _{aux} → NP _{suj}	NP _{suj} → VP _{aux} → VP _{main}

	VPaux	NPsuj	VPmain
--	-------	-------	--------

A Tabela 7 apresenta, por fase de treinamento (coluna mais à esquerda), os ordenamentos decrescentes de frequência e variância para A8 (colunas centrais), seguidos do ordenamento hierárquico revelado pela CA desse padrão. Observa-se que VPmain sempre apresentou a frequência mais elevada (e também as maiores oscilações, como foi salientado acima), enquanto que NPsuj sempre apresentou a menor variância (e também as menores oscilações de frequência). Para o padrão A8, portanto, aparentemente a menor variância (NPsuj) equiparou-se a uma maior influência na CA, enquanto que a maior frequência (VPmain) não garantiu nenhum nível de influência na hierarquização desse agrupamento.

3.8.4 A Harmonia

Os resultados obtidos a partir do cálculo da harmonia para o padrão A8 encontram-se registrados na Figura 15, cujo gráfico ilustra o modo de progressão do valor de H para A8 durante o percurso do estágio 1 ao 7 (fases A – G).

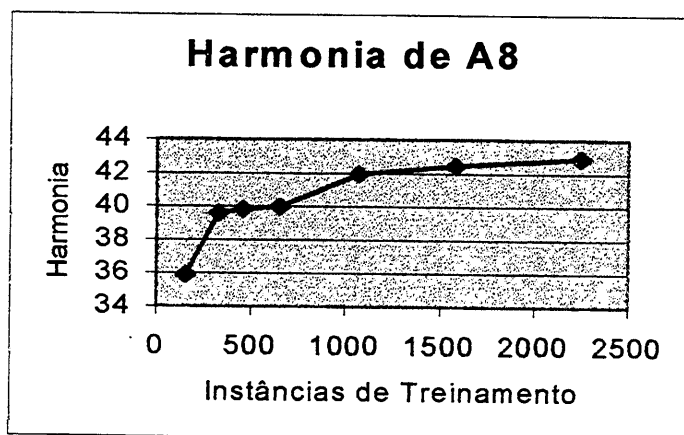


Figura 15: Progressão de H para A8

Observa-se que o valor de H para A8 sofre um aumento gradual, de caráter aproximadamente exponencial, ao longo do treinamento. Nota-se uma elevada taxa de crescimento de H das 150 às 325 instâncias de treinamento (estágio A), havendo, após esse ponto, uma queda acentuada dessa taxa, a qual se estende até as 651

instâncias (progressão B - C - D). A seguir, observa-se uma retomada do aumento da taxa de crescimento de H (embora não tão acentuada como no percurso A - B), até se alcançarem as 1.071 instâncias (estágio E). A partir daí, o crescimento de H continua, porém de modo progressivamente mais ameno (estágios F, G).

O grau de conformidade de A8 com os parâmetros internos da rede, portanto, aumenta, como era de se esperar, ao longo do treinamento. Seu grau de adaptação às restrições construídas pela rede, todavia, apresenta oscilações, principalmente no que tange ao percurso B - C - D, no qual há uma aparente dificuldade nesse sentido.

No Capítulo 4 serão discutidos, em paralelo, os resultados obtidos na coleta junto aos sujeitos e na simulação conexionista. Como já foi enfatizado, essa discussão terá como objetivo o estabelecimento de conexões que contribuam para a avaliação tanto da natureza do fenômeno em estudo (a formação de interrogações por aprendizes brasileiros de inglês como língua estrangeira) como das perspectivas de se utilizar modelos conexionistas para o estudo da aquisição de L2.

CAPÍTULO 4 - DISCUSSÃO

Neste capítulo serão discutidos os resultados obtidos nas duas fases do experimento descritas nos Capítulos 2 e 3, a saber, a produção de perguntas pelos aprendizes brasileiros e pela rede neural, respectivamente. Primeiramente, serão discutidos os resultados obtidos através da análise estatística da progressão da precisão na produção, por parte dos sujeitos estudados, de perguntas em inglês como L2. Subseqüentemente, serão analisados os dados obtidos a partir da simulação conexionista dessa produção, com especial ênfase na discussão dos resultados relativos ao mais freqüente dos padrões produzidos, o padrão A8 [DO + YOU(2PS) + TALK]. Será traçado, então, um paralelo com os resultados gerais obtidos a partir dos sujeitos. Finalmente, será desenvolvida uma reflexão crítica sobre a utilização de modelos conexionistas para este tipo estudo sobre a aquisição de L2.

4.1 A PRODUÇÃO DOS SUJEITOS

O perfil geral dos sujeitos pesquisados, apresentado na Tabela 1 (seção 2.3.1), indica que os níveis de proficiência N1-4 encontram-se dispostos, de modo geral, de forma crescente, tanto em relação ao tempo de estudo dentro do currículo do curso de inglês como no que tange ao estudo do inglês em suas escolas regulares. Isso reflete a boa homogeneidade da amostra, pois indica que houve um nível crescente de contato com o currículo ao longo da progressão N1-4. Há, contudo, além do aumento do tempo real de contato com o currículo do curso estudado, um aumento, ao longo de N1-4, da variação intragrupo em relação a esse item. No que respeita ao contato

com os currículos das escolas regulares há, além do aumento progressivo do tempo de contato, uma diminuição da variação intragrupo. Esses fatos também refletem a boa homogeneidade da amostra, apesar do grau de imprevisibilidade que o contato com o histórico escolar regular traz, dentro da ótica adotada aqui, em relação ao *input* em L2 com o qual os sujeitos interagiram ao longo de seu aprendizado, dada a suposta heterogeneidade dos ambientes escolares.

A codificação utilizada, ancorada na posição ocupada pelo sujeito (NP_{suj}), a partir da qual foram posicionadas as posições verbo auxiliar (VP_{aux}) e verbo principal (VP_{main}), adequou-se bem à tarefa, dados os baixos índices de perguntas enquadradas na categoria NC (Tabela 3, seção 2.3.2). Essa classificação, contudo, lançou mão de categorias sintático-morfológicas de escopo variável, tais como NP_{suj}, *modal*, *irregular* e *was*, portanto mais ou menos sensíveis a variações lexicais ao nível da forma. A pontuação das interrogações produzidas por cada um dos quatro níveis investigados, que levou em conta os critérios gerais posição e concordância, operou eficazmente como um medidor geral da precisão das produções dos sujeitos em relação às formas canônicas do inglês, permitindo a tabulação precisa exibida pela Tabela 3.

De um modo geral, houve um aumento gradual da taxa de padrões diferentes produzidos pelos sujeitos ao longo do eixo da proficiência (Tabela 2 e Figura 5, seção 2.3.2), sendo a maioria deles não canônicos. Quanto aos níveis de precisão alcançados, a análise estatística realizada na seção 2.3.3 revelou a ausência de uma diferença significativa entre N1 e N2. De N2 para N3, contudo, houve um aumento significativo na precisão da formação das interrogações. Após esse salto quantitativo, porém, observa-se um novo platô, pois não há diferença significativa entre as médias de pontuação de N3 e N4. Como os intervalos entre N1, N2, N3 e N4 são regulares, observa-se que os sujeitos são capazes de criar mais tipos de padrões ao longo de todo o seu aprendizado, mas não necessariamente de serem mais precisos, pois não há diferença significativa neste sentido nos intervalos N1-2 e N3-4.

A progressão da precisão dos sujeitos ao longo do eixo da proficiência foi, portanto, não linear, do tipo platô – rampa – platô (Figura 6, seção 2.3.3). Esse modo de desenvolvimento não linear não é novidade em aquisição de L2. Um dos modelos teóricos mais comumente citados na literatura em L2 a respeito do processo de produção de estruturas de IL em inglês, a escala de estágios do Modelo Multidimensional (MM), opera dentro de parâmetros não lineares de desenvolvimento de IL. O MM foi desenvolvido pelos pesquisadores envolvidos no projeto ZISA (Meisel, Clahsen & Pienemann 1981). Esses pesquisadores estudaram, tanto longitudinal como transversalmente, a aquisição do alemão como L2 por parte de adultos e crianças falantes nativos de línguas românicas. Esse modelo foi elaborado com base em estudos de dados naturalísticos (em um ambiente de aprendizado diferente, portanto, da situação dos sujeitos deste estudo), relacionando processos cognitivos subjacentes a estágios de desenvolvimento e fatores variacionais. Posteriormente, Pienemann & Johnston (1987, *apud* Pienemann, Johnston & Brindley 1988) estenderam os resultados do ZISA à aquisição do inglês como L2. O MM já alertava que, em aquisição de L2 “a precisão varia com o desenvolvimento e com a variação. (...) Ganhos em precisão não são fenômenos lineares. Ao contrário, a correção aumenta em ambas as dimensões. Essa é a razão pela qual não se pode igualar correção com desenvolvimento” (Pienemann, Johnston & Brindley 1988, p. 222).

No que tange ao ordenamento sintático de IL na produção das interrogações do inglês, o MM prevê que os seguintes estágios gerais de desenvolvimento estejam presentes na IL dos aprendizes (Lightbown & Spada 1993):

1. Palavras isoladas e fórmulas: *How are you?*
2. Ordem declarativa: *He like eat pizza?*
3. Pré-posição:
 - a) *WH* *What you do on weekends?*

- b) *DO* *Do you have car?*
- c) Outros *Is the boy like the picture?*
4. Inversões *WH* e Sim-Não:
- a) Cópula *Where is the post office?*
- b) Auxiliares (*DO* excluído) *Are you going out tonight?*
5. Inversões *WH*:
- a) *DO* *What do you plan to do?*
- b) Outros auxiliares *How is he arriving?*
6. Perguntas Complexas:
- a) *Tags* *It's better, isn't it?*
- b) Negativas *Why can't you go?*
- c) Encaixadas *Can you tell me where the bank is?*

Essas referências estruturais são, sob a ótica do MM, referências em nível de emergência de dada estrutura, não de domínio estatístico da mesma. Como vimos na Tabela 3 (seção 2.3.2), os sujeitos pesquisados neste estudo produziram uma vasta gama de tipos de estrutura interrogativas em inglês como L2, canônicos ou não, os quais surgiram em diferentes faixas de proficiência. Essas aparições podem, segundo o MM, servir de indicador em relação aos estágios listados acima. O modelo transversal de investigação utilizado neste estudo, contudo, não possibilita a

verificação da progressão desses marcadores nas IL dos sujeitos pesquisados. Interessantemente, houve, contudo, o aumento do nível de criatividade dos sujeitos ao longo de seu aprendizado (Tabela 2, seção 2.3.2), o que não significou, necessariamente, como já foi mencionado, um aumento nos níveis de consonância com os padrões canônicos do inglês. Uma hipótese de investigação possível nesse sentido seria a comparação da criatividade de IL dos sujeitos e a escala de estágios do MM. A análise do nível de variedade da produção de padrões pode ser muito mais ampla e relevante ao estudo da evolução da IL dos sujeitos do que as referências do MM, em sua maioria referências em direção à precisão estrutural.

O percurso da aquisição de L2 é, segundo abordagens como o MM, não linear e sistêmico, no qual vários fatores entram em combinação na produção de certos resultados. Curvas de desenvolvimento em U, por exemplo, correspondem ao fenômeno denominado, em aquisição da linguagem, de reestruturação. McLaughlin (1990) define reestruturação, em aquisição de língua materna, como “caracterizada por uma mudança descontínua, i.e., *qualitativa*, que ocorre enquanto a criança se desloca de um estágio de desenvolvimento para outro. Cada novo estágio corresponde a uma nova organização interna, não meramente à adição de novos elementos estruturais” (p. 117). Segundo Lightbown (1985), isso ocorre “porque a linguagem é um sistema hierárquico complexo, cujos componentes interagem de maneiras não-lineares. Visto sob essa ótica, um aumento na taxa de erro em uma área pode se refletir em um aumento da complexidade ou precisão em outra, seguido da supergeneralização de uma estrutura recentemente adquirida, ou simplesmente de um tipo de sobrecarga de complexidade que força a uma reestruturação, ou pelo menos a uma simplificação, em outra parte do sistema” (p. 177). Nos dados coletados neste estudo, portanto, houve uma variação, ao longo do eixo da proficiência, tanto qualitativa, pois variaram os tipos de padrões produzidos pelos sujeitos de cada um dos níveis (Tabela 3), quanto quantitativa, pois variaram os níveis de pontuação médios (i.e., de consonância a padrões canônicos) computados para os diferentes níveis.

O mero registro do comportamento não-linear do percurso interlingüístico é, contudo, em si só uma ferramenta de análise de nível descritivo. No MM, por exemplo, o nível explicativo é alcançado através da argumentação de que a complexidade psicológica de uma estrutura depende do grau necessário de reordenamento e rearranjo do material lingüístico no processo de mapeamento da *semântica subjacente às formas superficiais*. Sob essa ótica, esses níveis de complexidade psicológica predirão a ordem de aquisição das estruturas, pois eles exigirão estratégias de processamento condizentes em termos de complexidade (Ellis 1994). Os dados apresentados por este estudo indicaram, como já foi mencionado, um aumento no nível de precisão de N2 para N3, comportamento não observado nos intervalos N1-N2 e N3-N4. Como já foi salientado, uma vez que esses são intervalos regulares, com diferença de 96 horas de instrução entre cada um deles, tem-se, na seqüência platô – rampa – platô, um comportamento de característica não linear. Não houve, portanto, progresso significativo dos sujeitos de N1 para N2 no que concerne à precisão na produção das estruturas interrogativas do inglês, apesar de as horas de instrução, ou seja, o volume pressuposto de interação com o *input*, ter progressão linear. O mesmo ocorreu em relação ao percurso N3-4. Ao mesmo tempo, contudo, a variedade de estruturas produzidas (Tabela 2) aumentou no percurso N1-4, sendo a maioria dessas estruturas não canônicas. Do ponto de vista do instrutor, portanto, o que se observaria em sala de aula, sob essas condições, seria um aparente retrocesso no processo de aprendizado, pois cada vez mais as estruturas não canônicas e as oscilações estariam presentes na fala dos alunos. O que estaria ocorrendo, não obstante, seria uma preparação mental, por assim dizer, por parte dos aprendizes, para o salto de precisão observado em N2-3.

Este estudo supôs, com base em uma visão de cunho conexionista do fenômeno da aquisição de L2, que o cerne da questão está na interação entre o tipo de *input* com o qual os sujeitos tiveram contato durante seus aprendizados e os mecanismos mentais, inatos ou formados via aquisição de L1, que esses sujeitos trazem consigo às tarefas de aprendizado com as quais se defrontam. A simulação descrita no Capítulo 3 teve por objetivo, como já foi mencionado, a investigação da produção de perguntas em L2 dentro dessa linha de raciocínio. A seguir serão discutidos os

resultados dessa simulação, cujo principal objetivo foi o de tentar propor alguns caminhos explicativos sobre o tipo de processo interlingüístico que possivelmente leva à produção, em aquisição em L2, de comportamentos não lineares do tipo observado nos dados coletados aqui.

4.2 A SIMULAÇÃO

Modelos conexionistas da aquisição da linguagem trazem consigo pressuposições sobre esse processo. Tenta-se, via de regra, simular o que se julga ser a tarefa cognitiva imposta ao aprendiz durante o aprendizado (de uma L2, no caso do presente estudo). Além disso, esses modelos, pelo menos até o presente momento, propõem-se a elucidar questões específicas em relação à aquisição da linguagem, realizando escolhas que sempre deixam de fora alguns níveis de análise como o pragmático, por exemplo. O modelo aqui utilizado também se caracterizou, portanto, tanto por fazer pressuposições a respeito da tarefa do aprendiz como por possuir limitações de escopo. Essas questões serão analisadas em detalhe nas seções a seguir.

Como já foi mencionado no Capítulo 3, o modelo utilizado neste estudo teve por objetivo simular a produção de estruturas interrogativas do inglês como L2 por falantes de português brasileiro (PB), sob a ótica da relação *input / output*. Para que se possa traçar alguns paralelos entre o comportamento da rede e o dos sujeitos pesquisados de modo mais elucidativo, serão analisadas, primeiramente, as características intrínsecas (inatas) ao modelo. Subseqüentemente, será analisada a maneira como o *input* foi processado pela rede durante o treinamento. Finalmente, serão interpretados os resultados obtidos (*output*) na simulação sob a ótica da precisão estrutural (pontuação das saídas). Será, também, investigado, através da utilização da análise de agrupamento e pelo cálculo da harmonia, o processo de construção de configurações internas da rede.

4.2.1 As Estruturas Inatas da Rede

Em um modelo conexionista, o inato é representado pelo que é intrínseco à arquitetura do mesmo, à codificação dos dados e ao seu modo de aprender. Ou seja, inato é aquilo que já existe antes de ele ser treinado, não se alterando ao longo do processo de aprendizado. Como já foi ressaltado, o modelo construído para os fins deste estudo tanto faz pressuposições importantes a respeito do conhecimento prévio que o falante traz para a execução da tarefa de produção de perguntas em L2 como também simplifica alguns aspectos dessa produção.

Serão agora analisadas, separadamente, as características intrínsecas ao modelo utilizado neste estudo, a saber, sua arquitetura, a codificação utilizada e o seu modo de aprender.

4.2.1.1 A Arquitetura

A arquitetura da rede utilizada neste trabalho seguiu o modelo MP utilizado por Gasser (1990), composto de uma camada de entrada, uma camada intermediária e uma de saída (Figura 7, seção 3.2). Esse tipo de arquitetura permite que se formem configurações internas nas unidades intermediárias, aumentando a capacidade de aprendizado da rede¹⁰. No caso da rede utilizada aqui, o número de unidades intermediárias é inferior tanto ao número de unidades de entrada quanto ao de saída, o que garante um afunilamento de informações, i.e., ajuda a evitar que a rede simplesmente memorize os pares de entrada e saída. Esse tipo de arquitetura só é possível, contudo, quando o tamanho do conjunto de treinamento não excede a capacidade da generalização da camada intermediária projetada. Em sua simulação da aquisição de morfemas do passado e do plural do inglês, por exemplo, Plunkett & Juola (1999) utilizaram mais unidades na camada intermediária do que nas outras duas.

4.2.1.2 A Codificação

¹⁰ Esse tipo de arquitetura permitiu que os problemas ocorridos na simulação da aquisição dos morfemas do passado do inglês em Rumelhart & McClelland (1986), que utilizaram uma rede de apenas duas camadas, fossem resolvidos em Plunkett & Marchman (1991).

Um aspecto importante do sistema de codificação utilizado foi a pressuposição da existência de algumas categorias gramaticais plenamente funcionais na mente do aprendiz de L2. Ou seja, não coube à rede descobri-las, i.e., desenvolver, através de treinamento, representações internas que a capacitasse a decodificá-las a partir do estímulo externo. Essas categorias anteriores à simulação foram Sintagma, Pessoa, Número, Tempo e Concordância, com suas respectivas subcategorias. Como a tarefa do experimento com os sujeitos foi a produção escrita de perguntas, considerou-se que ela consistiu, minimamente, da realização do mapeamento entre essas categorias inatas ao projeto (entradas da rede) e as formas lingüísticas do inglês que expressam perguntas a elas relacionadas (saídas).

Outro fator importante, também intrínseco ao sistema de codificação, foi o ordenamento das entradas, construídas na ordem NPsuj – VPaux – VPmain, seguindo o ordenamento canônico, para as interrogações, do PB, L1 dos sujeitos. Como salienta Rossi (1996), em sua análise diacrônica abrangendo as interrogações diretas, no PB há amplo domínio do NPsuj pré-verbal, praticamente não havendo, portanto, o deslocamento para a esquerda do auxiliar em VPaux. Essa diferença de ordenamento sintático foi contemplada na codificação de saída do modelo, ou seja, pressupôs-se que essa reestruturação seja verdadeira na formação das perguntas em inglês por falantes de PB. A combinação da codificação das entradas e saídas esperadas da rede, portanto, considerou o ordenamento do tipo NPsuj – VPaux – VPmain como ponto de partida mental dos sujeitos pesquisados para o domínio das perguntas, a partir do qual foi construído o ordenamento VPaux – NPsuj – VPmain (Anexos D e E).

Em termos de abrangência, a codificação do modelo foi criada de forma a poder englobar os principais padrões interrogativos com os quais os sujeitos estudados entraram em contato durante seus estudos no curso de inglês que frequentaram. É importante que se leve em consideração, todavia, algumas limitações de escopo desse modelo, pois ele não contempla certas diferenças individuais nem entre diferentes verbos modais nem entre diferentes participios passados irregulares do inglês. Além disso, ele também simplifica as possibilidades

de expressão de relações temporais das formas utilizadas pelas saídas da rede. O modelo realiza, para os padrões verbais interrogativos do conjunto de treinamento, um mapeamento unívoco entre estrutura profunda (entrada) e superficial (saída), não levando em consideração, por exemplo, as possibilidades de expressão temporal futura do *Present Simple* e do *Present Continuous*, ambos assinalados apenas como tempo presente na codificação.

4.2.1.3 O Modo de Aprender

O algoritmo de aprendizado utilizado no treinamento da rede foi, como demonstra o Quadro 5 (seção 3.3), o *backpropagation*, de maneira supervisionada, cujo objetivo, segundo Plunkett & Elman (1997), é o de:

“... aprender um conjunto único de pesos de maneira tal que qualquer padrão de entrada produzirá o padrão de saída correto. De um modo geral, também se deseja que esses pesos permitam que a rede seja capaz de generalizar em relação a padrões não pertencentes ao grupo de treinamento”(p. 10).

O algoritmo de aprendizado da rede, como já foi salientado no Capítulo 1 (seção 1.5), é responsável por um processo gradual e cuidadoso nessa direção, reduzindo o erro obtido nas saídas através da propagação reversa do ajuste dos pesos das conexões, aprendendo cada vez menos ao longo do treinamento e sempre correndo o risco de ficar preso em um mínimo local.

A rede utilizada para a simulação da produção dos padrões interrogativos do inglês utilizou uma taxa de aprendizado de 0.1, valor bastante reduzido, que representou um aprendiz que se desenvolveu de forma lenta e gradual, sendo 100% bem sucedido ao longo do processo (todos os padrões do grupo de treinamento foram aprendidos). A realidade, contudo, poderia ter sido outra, caso houvesse sido utilizada uma taxa de aprendizado maior, cujos ajustes maiores nos pesos da rede poderiam ter tanto aprendido todos os padrões de forma mais rápida como também

ter enrijecido as conexões da rede prematuramente, causando a prisão a um mínimo local.

Em termos de investigação de fenômenos relativos à aquisição de L2, essa flexibilidade do modelo em relação a taxas de aprendizado é bastante interessante, pois condiz, a princípio, com as diferenças individuais em termos de sucesso de aprendizado, sem que se tenha que pensar nem em mudanças na arquitetura da rede (biologia) nem no regime de treinamento (estímulo externo).

Todas as características citadas nas três áreas inatas analisadas acima – arquitetura, codificação e modo de aprender – são inerentes ao projeto montado para os fins deste estudo, i.e., são basilares a qualquer análise que se venha a fazer a partir dos resultados obtidos após o treinamento da rede. É na estrutura da rede que se iniciam as opções que delinham a ótica da investigação, pois lá estão seus pontos de partida. Por sua vez, é no treinamento que são feitas suposições importantes sobre o *input*. A rede e a codificação elaboradas para este estudo demonstram, sob essa ótica, duas intenções fundamentais:

- Supor que o aprendiz adolescente inicia a tarefa de aprender uma L2 já dotado de um sistema semântico e sintático razoavelmente desenvolvido através de sua L1;
- Trabalhar dentro de um domínio extremamente específico, a saber, o dos padrões verbais interrogativos do inglês do tipo VPaux – NPsuj – VPmain, em seus mapeamentos forma / função temporal mais comuns.

Serão analisados, agora, os fatores externos à rede e o resultados obtidos na simulação.

4.2.2 O *Input* da Rede

Viu-se, na seção 2.2.1, que o treinamento da rede foi baseado, de forma aproximada, na progressão curricular da escola pesquisada no que se refere às formas verbais, de acordo com os livros-texto adotados para os níveis pesquisados. Essa ordem de estruturas foi representada, na simulação, pela seqüência de sete grupos de estruturas verbais, nomeadamente A, B, C, D, E e F, apresentados à rede de modo seqüenciado (foi incluído um grupo de padrões novos por vez), cumulativo (a rede teve que aprender sempre o que já sabia acrescido dos padrões novos) e supervisionado (houve sempre a presença de saídas esperadas para as entradas).

Em termos de escopo, há que se salientar que houve uma gama de padrões interrogativos eventualmente presentes no *input* dos sujeitos para os quais a rede não foi treinada, como, por exemplo, os padrões da voz passiva, das perguntas do tipo *tag*, das perguntas com *there + be*, etc., além dos padrões do tipo *Perfect Continuous* (*He has been eating a lot recently* – Ele tem comido muito ultimamente), apenas parcialmente contemplados através do treinamento com padrões do tipo *Perfect*.

O processo de aprendizado da rede foi baseado no critério da precisão total das saídas, ou seja, a rede somente entrou em contato com um grupo novo de padrões após ter produzido 100% de respostas corretas para os padrões anteriores. A simulação, portanto, constituiu-se em uma situação razoavelmente idealizada, pois a rede pôde de fato aprender um grupo de padrões antes de confrontar-se com um grupo novo. As respostas produzidas pela rede foram o produto da tradução alfanumérica dos vetores numéricos produzidos por *T-learn*. Como foi ressaltado na seção 3.3, essa tradução foi feita com base na distância euclidiana entre esses vetores e nas respostas possíveis dentro do universo composto pelos padrões de saída do grupo de treinamento. Além disso, a simulação, executada sob regime supervisionado, supôs a presença constante de *feedback* corretivo.

O modo de treinamento escolhido também influenciou o tipo de interação dos padrões com a rede. Apesar do seqüenciamento do *input* ter sido composto da progressão previamente ordenada dos sete grupos de padrões escolhidos, houve, no que concerne ao seqüenciamento interno dos padrões durante cada uma das sete fases

do treinamento, uma exposição randômica dos mesmos. Isso foi feito com o objetivo de se aproximar um pouco mais o regime de treinamento da provável variedade de frequência de exposição ao *input* do ambiente de sala de aula.

Quanto ao regime de treinamento, optou-se pelo regime padrão, ou seja, houve o reajuste dos pesos das conexões da rede a cada instância interativa. Esse tipo de regime iguala *input* a *intake* (Corder 1967), pois toda e qualquer exposição da rede a uma instância de um dado padrão causará algum nível de reconfiguração interna de seus pesos, mesmo que essa reconfiguração não chegue a alterar a saída correspondente àquele padrão.

A taxa de aprendizado da rede foi mantida constante ao longo de todo treinamento, portanto o modelo de aquisição proposto não levou em consideração os efeitos da idade na velocidade de aprendizado dos seres humanos (Sokolik 1990). A homogeneidade da amostra utilizada no experimento com os sujeitos permitiu que esse elemento fosse desconsiderado.

A escolha do conjunto de treinamento se deu dentro de um domínio bastante definido, a saber, os sete tipos de padrões interrogativos contidos no grupo de treinamento. Isso acarreta um isolamento do domínio escolhido, o que conseqüentemente não permite, apesar da visão sistêmica, de reestruturação interna, deste experimento, a possibilidade de qualquer influência externa ao domínio escolhido nas configurações internas da rede. Esse fato implica em um recorte de *input* que não contempla a influência que outras estruturas gramaticais do inglês possam ter no domínio das estruturas interrogativas. Esse tipo de influência, contudo, já foi observada tanto em experimentos envolvendo gramáticas artificiais (Manza & Reber 1997, *apud* Dienes, Altmann & Gao 1999) como em aquisição de L2 (Hyltenstam 1977).

De um modo geral, a interação da rede construída para os propósitos deste estudo com o *input* a ela fornecido, enquanto modelo humano de aprendizado, tem, minimamente, as seguintes características:

- A rede-aprendiz interagiu com um *input* seqüenciado em nível macrocurricular e randômico em nível microcurricular;
- Ela recebeu *feedback* corretivo constante e 100% eficiente, pois todo e qualquer *input* tornou-se *intake* e os padrões oferecidos como modelos foram sempre os corretos;
- A rede-aprendiz não sofreu nenhuma alteração na sua velocidade de aprendizado ao longo de sua exposição ao *input*;
- Ela só operou no domínio das estruturas interrogativas do tipo VPaux – NPsub – VPmain.

4.2.3 As Saídas

Como já foi salientado, as entradas da rede resultaram da codificação binária dos elementos lingüísticos que fizeram parte do seu conjunto de treinamento. Esse treinamento foi, dada a natureza do simulador neural *T-learn*, essencialmente matemático, ou seja, de natureza numérica. As saídas da rede durante o treinamento foram, portanto, originalmente, saídas variando de 0 a 1. Para que fosse possível se visualizar o que elas representavam em termos lingüísticos, foi necessário que elas fossem traduzidas, i.e., que o processo inverso ao das entradas ocorresse. Essa tradução, ou recodificação, foi realizada através de um expediente matemático, a distância euclidiana, cuja característica principal é sempre “ter o que dizer”, i.e., sempre decidir qual o vetor de referência (ou seja, pertencente ao universo das saídas possíveis, que é composto da adição do conjunto de treinamento com a saída nula) que mais se aproxima da saída real obtida. Diferentemente de outro recurso de recodificação, a saída por arredondamento (*threshold output*), através do qual a rede pode gerar saídas do tipo “?”, a distância euclidiana força a rede a produzir a melhor saída possível dentro do seu universo de operação.

No Capítulo 3 foram realizadas algumas medições relativas às saídas obtidas ao longo do treinamento da rede. Inicialmente, se mediu o nível de dificuldade enfrentado pela rede durante a tarefa. A Figura 8 (seção 3.8.2) mostra o nível de dificuldade oscilante demonstrado pela rede para produzir as respostas exigidas pelo treinamento, apontando para a existência de uma elevação do grau de dificuldade nos percursos $C \rightarrow D$ (*Modal* \rightarrow *Past Simple*) e $E \rightarrow F$ (*Past Continuous* \rightarrow *Present Perfect*). A seguir, as saídas relativas aos 54 padrões de treinamento foram pontuadas com valores variando de 1 a 3 (seção 3.8.2). Observou-se um percurso não linear ascendente de oscilação desses valores ao longo do treinamento, ficando evidenciada a presença do fenômeno da reestruturação (desaprendizado), particularmente claro nas fases iniciais do treinamento ($A \rightarrow D$). A partir desse ponto, apesar de o nível de dificuldade de aprendizado voltar a subir no percurso $E \rightarrow F$, o nível geral de instabilidade do sistema mantém-se baixo. Observou-se, também, uma grande oscilação dos níveis de variação intragrupo ao longo de todo o treinamento (Figura 10), um pouco menos acentuada nas fases finais do treinamento. Ou seja, as conexões da rede vão se solidificando (Elman 1993), realizando mudanças mais conservadoras durante o aprendizado. As Figuras 9 e 10 demonstram que por parte da rede houve, como nos resultados obtidos a partir dos sujeitos, um percurso de caráter geral ascendente no que respeita ao nível de precisão nas perguntas produzidas. A rede, contudo, apresentou oscilações em U ($A \rightarrow D$) que não foram constatadas nos resultados obtidos a partir dos sujeitos. A rede se apresenta, portanto, mais instável que os sujeitos ao longo do processo de aprendizado. A presença dessa instabilidade é corroborada pelo fato de que, em se tratando de uma simulação conexionista, a reestruturação explícita nas saídas produzidas pela rede reflete uma diferença de nível, i.e., quantitativa, que, dados os critérios utilizados pela distância euclidiana entre os vetores do grupo de treinamento, acaba por se refletir em uma diferença qualitativa de produção verbal. Ou seja, há uma reestruturação numérica constante ao longo do treinamento que somente torna-se relevante em termos de saída ao ultrapassar o patamar qualitativo euclidiano da proximidade vetorial. Esse processo, se verdadeiro no caso dos sujeitos, foi bem menos aparente, manifestando-se apenas no percurso N2-3.

Ao se comparar esse comportamento da rede com o tipo de progressão de precisão obtida a partir da coleta de dados junto aos sujeitos, que revelou uma seqüência não linear do tipo platô – rampa – platô, com uma elevação significativa do nível de precisão de N2 para N3 (seção 2.3.3), pode-se pensar em um caminho de investigação desse tipo de fenômeno com base na natureza dinâmica através da qual a simulação busca representá-lo. É curioso pensar que, de N1 para N2, intervalo que representa 96 horas de instrução formal em inglês, não tenha ocorrido nenhuma mudança significativa nos níveis de precisão das estruturas interrogativas de IL dos sujeitos, enquanto que, de N2 para N3, intervalo que implica nas mesmas 96 horas de instrução, essas estruturas tenham sofrido mudanças significativas. Isso leva a pensar em um processo constante e subjacente no qual micromudanças se combinam até que ultrapassem o patamar da visibilidade, ou seja, sejam explicitadas nas produções verbais dos sujeitos. Sob essa ótica, a rede, dada a sua condição de metáfora neurológica (seção 1.2), pode ser uma maneira bastante interessante de se tentar um acesso aos pormenores desse processo invisível de mudança. Ao final deste capítulo serão traçadas algumas considerações sobre possíveis frentes de investigação nesse sentido.

A análise de agrupamento (CA), realizada subseqüentemente à pontuação das saídas, teve como preocupação central o agrupamento em torno do padrão A8, revelando a sua movimentação ao longo do espaço vetorial multidimensional. Sendo a CA um agrupamento hierárquico baseado nas distâncias euclidianas entre os vetores, os padrões que figuraram mais próximos entre si, no diagrama arbóreo, foram os com o maior grau de similaridade no que tange às representações internas (ativações) das unidades intermediárias da rede. Viu-se, também, que a similaridade na categoria NPsuj foi, de modo geral, o principal fator geral de aglomeração.

Como já foi salientado na seção 3.8.2, a análise do caminho traçado por A8 nos sete estágios de treinamento (excetuado o estágio G, no qual VPmain foi o fator mais influente para uma das aglomerações), revelou que os elementos se agruparam segundo a seguinte ordem decrescente de influência de categoria:

- $NPsuj \rightarrow VPaux \rightarrow VPmain$

A investigação da relação entre a frequência dos elementos pertencentes a cada uma dessas categorias (Tabela 7, seção 3.8.3) e a ordem de influência de agrupamento revelada pela CA demonstrou que, ao longo do treinamento:

- As frequências relativas médias de ocorrência dos padrões do grupo de treinamento seguem a ordem decrescente $VPmain \rightarrow VPaux \rightarrow NPsuj$ até a fase C, mas $VPmain \rightarrow NPsuj \rightarrow VPaux$ após esse ponto;
- O valor das variâncias médias desses padrões seguiu a ordem decrescente $VPaux \rightarrow NPsuj = VPmain$ até a fase B, mas $VPmain \rightarrow VPaux \rightarrow NPsuj$ após esse ponto;
- Os valores das frequências relativas de ocorrência para os elementos pertencentes ao padrão A8 seguem a ordem $VPmain \rightarrow VPaux \rightarrow NPsuj$ até a fase C, a partir da qual estabeleceu-se a ordem de grandeza $VPmain \rightarrow NPsuj \rightarrow VPaux$;
- As variâncias dos elementos constituintes do padrão A8 apresentaram ordenamento do tipo $VPaux \rightarrow VPmain = NPsuj$ até a fase B, a partir da qual o ordenamento de categorias passou a ser $VPmain \rightarrow VPaux \rightarrow NPsuj$.

Esses resultados contrastam com a ordem hierárquica apresentada pela CA. *NPsuj* não foi, em nenhum momento, o padrão mais frequente nas sete fases de treinamento da rede. Ou seja, na CA houve um agrupamento cujo ordenamento de influência não foi determinado diretamente pelo número de vezes no qual a rede entrou em contato com determinado padrão, i.e., suas frequências relativas dentro do universo dos grupos de treinamento. Há, portanto, indícios de que as aglomerações das representações internas das unidades intermediárias da rede foram desenvolvidas, majoritariamente, a partir da regularidade de frequência de exposição ao *input* em relação a uma dada categoria inata (*NPsuj*, *VPaux* ou *VPmain*), não da quantidade

dessa exposição. Os valores mais estáveis de frequência e variância, com menos oscilações, predominaram.

Houve, contudo, casos nos quais a predominância da regularidade como fator de hierarquização não foi clara, como nas fases A e B de treinamento, nas quais o valor médio de variância para os elementos de NPsuj foram iguais aos de VPmain. Além disso, na fase G a menor variância média é a de NPsuj, enquanto que, na CA, NPsuj perde essa predominância em relação aos padrões do *Present Perfect* (F) e do *Past Perfect* (G), que se afastam da aglomeração em torno de YOU, formando o único grupo com ordenamento hierárquico do tipo VPmain → VPaux → NPsuj.

Exceções como essas sugerem que, na verdade, os fatores frequência e variância interagiram constantemente ao longo do processo de aprendizado da rede, oscilando em nível de influência na hierarquização dos elementos pertencentes às categorias NPsuj, VPaux e VPmain. Ao longo do treinamento supervisionado, a rede parece ter elaborado, a partir da interação com cada um dos padrões do conjunto de treinamento, configurações internas capazes de formar perguntas a partir da relação entre frequência de ocorrência e variância dos elementos que compõem as categorias NPsuj, VPaux e VPmain. A hierarquização dessas categorias inatas à rede teria sido construída, sob essa ótica, a partir de grandezas intrínsecas ao regime de treinamento, nomeadamente frequência e variância, que estabeleceram, ao longo do processo de aprendizado, ranqueamentos sintáticos.

Para o padrão A8, as restrições de boa formação construídas pela rede estão, segundo os resultados gerais da CA, dispostas na seguinte ordem hierárquica geral:

1. NPsuj
2. VPaux
3. VPmain

Como foi visto na seção 1.6, um dos princípios fundamentais da TO é o de que as gramáticas possuem restrições em relação à boa formação das estruturas lingüísticas, e que essas restrições encontram-se em conflito intenso. Viu-se, também, que um fator importante na TO é o fato de a disputa entre essas restrições ser satisfeita otimizada, não perfeitamente. Sob essa ótica, pode-se pensar na ordem acima como uma hierarquia de restrições em conflito a serem satisfeitas. Nesse processo, as perguntas seriam formuladas, pela rede, a partir do elemento em NP_{suj}, ao qual é anexado um VP_{aux} que concorde com esse elemento e, por último, um elemento em VP_{main} que concorde com VP_{aux}, sendo esses três passos obrigatórios tanto no que tange à presença dos elementos como no que diz respeito ao seu ordenamento. Esse ordenamento hierárquico, na verdade, não seria um ordenamento temporal em termos de produção, pois o processamento da rede se dá de modo paralelo, mas sim uma maneira organizada de se conceber o modo como a rede obedece, de modo simultâneo, suas restrições internas.

Estendendo-se os resultados em relação a A8 aos outros padrões de treinamento, tem-se a rede descobrindo como formular perguntas em inglês a partir da sua capacidade de hierarquização de suas configurações internas, desenvolvidas na interação supervisionada com o *input*. Essas configurações seriam, então, capazes de executar as várias operações em paralelo necessárias para a produção de qualquer um dos padrões do conjunto de treinamento, respeitando a hierarquia estabelecida por essas configurações. Além disso, essas hierarquias poderiam vir a sofrer alterações ao longo do aprendizado da rede, apresentando inversões de nível de influência como se observou na exceção relativa à fase G de treinamento, dependendo de como se desenvolvesse as relações entre frequência e variância durante o processo.

Há, portanto, dada a maneira como a simulação foi elaborada, rigidez no que tange às categorias inatas NP_{suj}, VP_{aux} e VP_{main}, mas haveria flexibilidade em relação à hierarquização entre os elementos que ocupam essas categorias. Essa combinação de rigidez com flexibilidade pode ser uma maneira interessante de se pensar a evolução da IL dos aprendizes de L2, admitindo-se que eles já possuam

categorias lingüísticas desenvolvidas durante (ou antes, no caso de estruturas inatas) a aquisição de L1 e que o mapeamento da L2 é, em grande medida, constantemente determinado pela interação com o *input*. Com o passar do tempo, contudo, a influência do *input* vai diminuindo em importância, pois as conexões da rede vão, como aponta Elman (1993), enrijecendo ao longo do treinamento e conseqüentemente, solidificando suas configurações internas, ou seja, suas hierarquizações sintáticas.

Esse modo de interação dinâmica entre frequência e variância foi abordado, sob a ótica da regularidade morfológica, em Ellis & Schmidt (1998), em seu estudo da aquisição da morfossintaxe de flexões regulares e irregulares em MALs. Os autores afirmam, baseados na Lei Exponencial do Aprendizado (*Power Law of Learning*), que nos resultados de seu estudo “frequência e regularidade estão combinadas fatorialmente” (p. 307), portanto não se constituem em um mecanismo dual (regras *versus* memória), pois “sob uma perspectiva associacionista [frequência e regularidade] são, essencialmente, o mesmo fator com nomes diferentes” (p. 329). A Lei Exponencial do Aprendizado (Newell 1990) interpreta o aprendizado humano como sendo a resultante de mecanismos associativos básicos envolvidos na formação de novos *chunks* (seqüências lingüísticas não analisadas) e dos efeitos da frequência na acessibilidade dessas representações. A memória humana exhibe propriedades como a Lei Exponencial do Aprendizado porque essas propriedades constituem-se em uma resposta otimizada a um ambiente no qual a probabilidade de ocorrência de um elemento em um dado momento é uma função exponencial de sua frequência de ocorrência no passado (Anderson & Schooler 1991). A ativação neural, que controla o aprendizado, reflete a probabilidade de um elemento ocorrer no ambiente. Dessa forma, os processos neurais são construídos de forma a adaptar seu comportamento às propriedades estatísticas do ambiente (Anderson 1993). Os sistemas conexionistas, por sua vez, são construídos de forma a fazer a mesma coisa (Chater 1995, *apud* Ellis & Schmidt 1998).

Tanto experimentos utilizando MALs como simulações conexionistas permitem o acesso ao histórico de frequências de interação (os primeiros, em relação

aos sujeitos, e as segundas, à rede neural artificial), o que geralmente não encontra-se disponível para estudos da aquisição de L1 ou L2 em ambiente natural. Nesses, o histórico de interação com certos elementos do *input* é, normalmente, algo sobre o qual se tem muito pouco conhecimento ou controle. No caso dos sujeitos pesquisados neste estudo, pode-se pensar na ausência de alteração significativa no nível de precisão nos percursos N1-2 e N3-4 como resultado da inconstância da frequência de exposição aos elementos presentes no *input*, que, em um regime de instrução formal seqüenciado em unidades que alternam tópicos estruturais, apresentam um nível baixo de reapresentação. Esse esparsamento, segundo abordagens como a Lei Exponencial do Aprendizado, dificultaria a acessibilidade desses elementos na memória dos aprendizes. Novamente, a rede neural parece ser uma ferramenta bastante interessante na investigação da propriedade, em aquisição de L2, desse tipo de hipótese, pois segue, como já foi mencionado, esse tipo de lógica de aprendizado.

Na simulação realizada neste estudo, a regularidade de ocorrência (pouca oscilação de frequência, baixa variância) de elementos na posição NPsuj aponta, em um paralelo com a produção dos sujeitos, para uma importância desses elementos no aprendizado da formação de perguntas de inglês como L2. Nos dados coletados e tabulados na seção 2.3.2 (Tabela 3), observa-se que 32 das 788 perguntas coletadas não puderam ser enquadradas no sistema de classificação proposto. Nesses 32 casos, apenas 7 são de omissão de elemento obrigatório em NPsuj. Em contraste, há a omissão de elemento obrigatório em VPaux em 26 dos 68 tipos de padrão produzidos pelos sujeitos de N1-4. Existe, portanto, uma tendência de presença obrigatória de elemento em NPsuj. Sob a ótica da simulação conexionista, esse fato pode ter origem na alta frequência e regularidade de referências a sujeitos em perguntas no mundo real, seja qual for o contexto temporal. Ou seja, no domínio das perguntas haveria maior oscilação de frequências em relação aos elementos pertencentes a VPaux e VPmain, o que dificultaria mais o mapeamento nessas duas categorias. Uma maior investigação em torno dessa hipótese exigiria, contudo, uma rede muito mais complexa do que a utilizada para os fins deste estudo. Nessa nova rede, o mapeamento deveria incluir diferenciações morfossintáticas mais refinadas em relação a itens abrangentes como THEY (e todos os outros cinco elementos de

NPsuj), por exemplo, devendo ser capaz de reconhecer itens lexicais como *friends* (amigos) como pertencentes a essa subcategoria. O mesmo procedimento poderia, também, ser adotado em relação à categoria VPmain, pois ST, ED, IR, etc., teriam que ser subdivididos em várias subcategorias morfológicas verbais. Tais categorias, inclusive, já foram simuladas com sucesso em Plunkett & Juola (1999). Não há, a princípio, nenhum impedimento de ordem conceitual em relação a uma expansão de escopo desse tipo de rede, pois técnicas como a utilização de produtos tensores (Prince & Smolensky 1997) permitem a utilização de números bastante elevados de tipos de padrões de treinamento, o que viabiliza a inclusão de nuances morfológicas e semânticas na representação dos dados. As redes neurais podem, pelo menos em princípio, trabalhar em vários níveis de análise linguística (sintático, semântico, fonológico, etc.) simultaneamente, bastando para isso que se construam codificações vetoriais que incluam esses fatores.

Voltando-se à análise dos procedimentos adotados neste estudo, viu-se que, após a CA, foi medida a harmonia (H) da rede para o padrão A8 nos sete estágios de treinamento. O resultado foi, como apresenta a Figura 15 (seção 3.8.4), uma curva de caráter exponencial com uma leve depressão em termos de taxa de crescimento no percurso entre as fases B, C e D. Observou-se, portanto, um comportamento não-linear caracterizado por um aumento do grau de harmonia da rede, i.e., um aumento, ao longo do treinamento, do grau de conformidade do padrão A8 às restrições sintáticas representadas pelos pesos das conexões da rede. Comparando-se a curva resultante da medição de H para A8 às Figuras 13 e 14, nas quais se encontram representados os valores das frequências relativas de ocorrência e de variância dos três elementos que o constituem, se observa que a depressão de H coincide com mudanças de orientação das curvas de TALK, que de descendentes passam a ascendentes. Há, então, uma coincidência entre o aumento da variação de frequência e da variância de um dos elementos constituintes de A8, o que aumenta a instabilidade total do padrão, com a queda da taxa de aumento do nível de harmonia da rede para A8. Esse fato corrobora a suspeita de a relação entre frequência e variância ser um fator determinante na elaboração das configurações internas da rede, tanto sob a ótica das ativações das unidades intermediárias (CA) como das

restrições impostas pelos pesos de suas conexões (H). Por conseguinte, sob a ótica da simulação conexionista descrita neste estudo, os fatores que dão origem às restrições de gramaticalidade, de boa formação sintática, aparentemente são, no que tange à relação *input / output*, originados na relação entre frequência e variância. Como já foi enfatizado, a rede constrói, a partir de suas estruturas inatas (seção 4.2.1), os refinamentos sintáticos para cada um dos padrões do conjunto de treinamento através de suas configurações internas. Dentro do âmbito da hipótese proposta por este trabalho, esses refinamentos seriam tão gerais ou tão específicos quanto a similaridade vetorial entre esses padrões e a frequência em que se encontram no ambiente permitirem. Essas características do ambiente podem, portanto, reconfigurar parcialmente as noções de gramaticalidade da rede, mas têm de respeitar as bases estruturais e de codificação da mesma, ou seja, suas estruturas inatas.

Dessa forma, se for admitido que o tipo de rede, de regime de treinamento e de codificação utilizados neste estudo são aproximações válidas da tarefa imposta aos sujeitos pesquisados, pode-se analisar o salto quantitativo (que, como vimos na seção 4.1, pode também ser considerado um salto qualitativo) de precisão de N2 para N3 (Figura 6, seção 2.3.3) como tendo origem nas relações entre frequência e variância no treinamento recebido pelos sujeitos. Pode-se cogitar, por conseguinte, que esses sujeitos, munidos de estruturas neurais inatas que contemplam a estrutura sintagmática básica das línguas humanas, entrariam em interação supervisionada (via instrução formal) com o *input* lingüístico em L2 de modo a construir os refinamentos sintáticos dentro desses sintagmas. Esses refinamentos teriam origem no desenvolvimento de restrições impostas às combinações possíveis de elementos ao longo do processo associativo de aprendizado, a partir do *feedback* corretivo disponível no ambiente (implícito ou explícito) e da relação entre frequência e variância dos elementos que compõem o *input* disponível. Essas restrições seriam formadas gradualmente, produzindo saídas que respeitam o estado de coisas sincrônico a essa produção, i.e., as hierarquias sintáticas representadas pelas configurações mentais dos sujeitos naquele estágio de aprendizado. Dentro desse universo restritivo, padrões de *input* seriam processados contra limiares paramétricos de atividade neural (no caso deste estudo, representados pela distância euclidiana),

cuja ultrapassagem determinaria o tipo de elemento produzido. Sendo esse um processo do tipo PDP, as saídas relativas a qualquer domínio da linguagem teriam um comportamento não linear, sensível a variações do *input*, cujas frequência e variância seriam determinantes na formação das representações mentais e, conseqüentemente, nas produções lingüísticas dos sujeitos. Cada combinação de itens produzida representaria, dentro desse contexto, uma resposta otimizada às restrições impostas pelas configurações mentais de cada sujeito, resultantes de seu histórico interacional particular, construído sob a tutela de suas bases inatas, único em sua especificidade e com graus variáveis de relação com a gama de outras estruturas em nível intra e inter dominial.

CONCLUSÃO

A comprovação da validade do tipo de abordagem da aquisição de L2 proposta por este trabalho requer, *minimamente*, duas condições básicas, a saber, o controle preciso em relação ao histórico de *input* dos sujeitos e uma configuração de rede e codificação de *input* que representem sistemas mentais inatos aos sujeitos. Além desses, há que se utilizar algoritmos de aprendizado que se aproximem das capacidades indutivas humanas, assim como submeter a rede a um regime de treinamento compatível com o histórico de *input* dos sujeitos, tanto quantitativa como qualitativamente. Como foi salientado no Capítulo 1, já existe, na literatura, um *corpus* de evidência bastante robusto em relação à validade da utilização das redes neurais na simulação de fenômenos relativos à aquisição da linguagem (principalmente em L1). Há, não obstante, uma variedade de configurações de rede sendo utilizadas na resolução de diferentes problemas de aquisição da linguagem, como, por exemplo, as redes SRN, para a investigação de efeitos maturacionais (Elman 1993), as redes MP, do tipo que utilizamos neste estudo, para a investigação da morfologia (Plunkett & Juola 1999) e as redes VSG, para o estudo do processamento sintático (Tabor, Juliano & Tanenhaus 1997). Apesar da diversidade de modelos, há uniformidade no credo de que a aquisição da linguagem, seja ela de L1 ou L2, é, essencialmente, um fenômeno de natureza distribuída, no qual as categorias lingüísticas são construídas a partir da interação das estruturas inatas do indivíduo com o estímulo externo. O papel dessas estruturas inatas nas simulações, contudo, é, geralmente, reduzido a um mínimo necessário, de maneira que a rede possa executar sua tarefa com sucesso, i.e., comportar-se de maneira similar aos

seres humanos. Esse sucesso, quando atingido, leva a crer que as pressuposições adotadas foram acertadas.

O experimento realizado aqui demonstrou que os sujeitos testados exibiram um comportamento não linear em relação à produção de perguntas em L2, caracterizado por uma elevação de N2 para N3 no que tange à precisão das perguntas produzidas em relação às formas canônicas do inglês. A segunda fase do experimento, a simulação conexionista, procurou recriar as prováveis condições de aprendizado desses sujeitos através de um treinamento baseado na progressão de estruturas gramaticais interrogativas que compuseram o currículo dos indivíduos pesquisados e de um regime de treinamento acumulativo. O comportamento exibido pela rede foi, como no caso dos sujeitos, não linear sob uma série de aspectos. Primeiramente, observou-se que houve níveis de dificuldade oscilantes em relação aos estágios de teste, elevação do grau de dificuldade nos percursos $C \rightarrow D$ (*Modal* \rightarrow *Past Simple*) e $E \rightarrow F$ (*Past Continuous* \rightarrow *Present Perfect*). A seguir, a pontuação das saídas da rede apontou para a presença de reestruturação durante o processo de aprendizado, com a ocorrência de padrões de desenvolvimento em U (Figuras 9 e 10) ao longo do treinamento, ou seja, com presença do processo de desaprendizado. Posteriormente, a análise de agrupamento revelou uma relação importante entre a frequência e a variância ao longo do processo de aprendizado, com a coincidência do percurso $B \rightarrow C \rightarrow D$ ser, para o padrão A8, novamente um período de oscilações (Figuras 13 e 14), havendo a elevação acentuada do nível de frequência para o elemento TALK (VPmain). Finalmente, ao se medir o nível de harmonia da rede para A8 nos sete estágios de treinamento (Figura 15), observou-se uma progressão de caráter exponencial, também com oscilações no percurso $B \rightarrow C \rightarrow D$, dessa vez na taxa de crescimento do nível de harmonia da rede para A8, havendo uma coincidência com o aumento da variação de frequência e da variância de um dos elementos constituintes de A8 (TALK – VPmain). Como foi salientado, isso aumenta a instabilidade total do padrão, com conseqüente queda da taxa de aumento do nível de harmonia da rede para A8. A plotagem de H na verdade representa, para a rede, uma curva de aprendizado, pois H cresce com o aumento do

nível de conformidade das saídas produzidas com as restrições gramaticais impostas pelas suas configurações internas.

Admitindo-se a validade do modelo conexionista construído em *T-Learn* como ferramenta de investigação, o comportamento exibido pela rede prevê, como causa provável para a elevação do patamar de precisão de N2 para N3, a oscilação do nível de incorporação de elementos ao sistema de IL dos sujeitos, ou seja, a variação do número de instâncias de contato, por parte dos sujeitos, com as estruturas *interrogativas do inglês, conjugado a uma maior ou menor constância desse contato*. Essa oscilação causaria, ao longo do processo de aprendizado, a alteração do jogo de forças que ocorre no processo de hierarquização das configurações mentais internas desses sujeitos, o que afetaria suas produções de IL.

Essa é, claramente, uma hipótese de teste. A similaridade encontrada entre o comportamento dos sujeitos e da rede neural elaborada para os fins deste estudo limita-se ao fato de ambos apresentarem um caráter ascendente e não linear, e de ambos terem sido fruto de um treinamento subdividido em fases (unidades dos livros-texto, no caso dos sujeitos, e estágios de treinamento, no caso da rede). O teste da hipótese de que a relação entre frequência e variância está no cerne do aprendizado formal de L2 requer uma investigação dentro de parâmetros específicos. Pode-se pensar, por exemplo, em um estudo longitudinal que ou lance mão do uso de uma MAL, como no caso de Ellis & Schmidt (1998), ou que consiga um nível de controle em relação a todo e qualquer *input* com o qual os sujeitos entrem em contato que permita que se estabeleça um regime de treinamento próximo do histórico real de contato desses sujeitos com as estruturas de L2. No primeiro, perde-se em validade ecológica, no segundo, perde-se em controle experimental (Jung 1971). De qualquer forma, avanços como a utilização de representações vetoriais matriciais e de produtos tensores na codificação dos dados a serem processados pela rede (Gasser 1990, Prince & Smolensky 1997), o aumento constante do poder de processamento dos simuladores neurais e novas configurações de rede como o VSG (Tabor & Tanenhaus 1999), parecem apontar para um campo potencialmente prolífico no estudo da aquisição de L2. Como foi mencionado no Capítulo 1, pesquisas nessa área

têm auxiliado a refinar as concepções inatistas a respeito das bases biológicas sobre as quais construímos nossos conhecimentos lingüísticos ao longo do tempo. Os resultados futuros desses estudos fornecerão, espera-se, informações mais claras no sentido de se poder elaborar modelos mais realistas de aquisição de L2, que possam não somente simular o comportamento humano com precisão, mas também, e principalmente, dada sua maleabilidade, ser uma fonte rica de formulação de novas hipóteses de teste. Mais especificamente no que tange ao experimento descrito aqui, uma reelaboração e expansão de escopo do conjunto de treinamento, apoiada em dados longitudinais criteriosos, pode vir a lançar luzes sobre maneiras de otimizar a elaboração dos currículos e a instrução de L2, de modo a criar melhores condições de os alunos interagirem com o *input* de forma a acomodar as reestruturações de suas ILs da maneira mais eficaz possível.

O que o conexionismo terá ou não de fato a oferecer ao estudo da aquisição de L2 dependerá, basicamente, do grau de sucesso das simulações a serem realizadas na área, entendendo-se aqui sucesso como abrangendo o conjunto coerente das pressuposições dos modelos neurais com os resultados por eles atingidos. Os pressupostos devem ser validados pela comunidade científica e os resultados devem, obrigatoriamente, espelhar os fatos do processo de aprendizado humano. O quão simplistas ou não os modelos conexionistas em aquisição de L2 serão dependerá, acima de tudo, da compreensão dos pesquisadores em relação à natureza dos fatores em jogo nesse processo. Como implementar noções como a do Filtro Afetivo (*Affective Filter* - Krashen 1982), por exemplo? Relacionar estados psicológicos e processos vetoriais pode parecer, a princípio, uma idéia não implementável. Admitindo-se, contudo, que os primeiros possam ser classificados e, portanto, codificados vetorialmente, e que os segundos possam processar interrelações entre esses estados e outros fatores relevantes ao processo de aquisição de L2, tem-se um cenário mais favorável. Modelos cuidadosamente concebidos podem, gradualmente, mostrar-se capazes de simular as transformações que normalmente ocorrem em nosso cérebro, possibilitando que esses processos possam ser analisados mais de perto. As funções matemáticas não lineares podem vir a se demonstrar mais versáteis e explicativas do que conceitos como o de estágio, por exemplo, revelando-se ferramentas que permitam a manipulação e

potencializem a compreensão das densas relações interfatoriais em jogo no processo de aquisição de L2.

Outro fator relevante ao desenvolvimento dessa frente de estudos é a nova realidade que ela apresenta ao pesquisador, que tem nas mãos um instrumento com um grau de maleabilidade criativa que lhe permite codificar vetorialmente suas convicções e suspeitas e pô-las imediatamente à prova, obtendo resultados e gráficos de análise em um espaço bastante curto de tempo. No caso do presente estudo, não foram poucas as codificações e pré-simulações realizadas e descartadas. Se a opção feita aqui teve cunho fortemente estrutural e, em grande medida, simplificador, essa foi precisamente uma opção, inserida no contexto de pesquisa proposto, não um caminho ditado pelo connexionismo. A ferramenta em si é potencialmente integradora e multiuso. Se reduzida ao seu núcleo mínimo, tem-se um princípio operacional bem definido, i.e., uma maneira de pensar algorítmica, lógica, indutiva, aliada a uma base biológica sincronicamente fixa, restritiva, representada pela arquitetura da rede, e a uma maneira de conceber o mundo da linguagem construída pelo pesquisador, representada pelo sistema de codificação utilizado. Esses três fatores são a base de funcionamento connexionista. Se, ao longo do tempo, essa base não se provar suficientemente explicativa como metáfora do processo da cognição humana, será inevitavelmente descartada, não tendo, conseqüentemente, futuro na aquisição de L2. Optou-se, aqui, por partir-se do sintagma, uma noção gerativa, por entender-se que ela tenha realidade psicológica para o aprendiz adolescente de L2. Esse credo norteou a construção da rede e a codificação dos dados. Poderia ter-se partido de noções mais basilares, como os morfemas do inglês, percorrendo-se o trajeto até o nível do enunciado, bastando para isso se elaborar uma rede com uma amplitude compatível com uma tal ampliação de escopo. O tamanho dos vetores de codificação, nesse caso, seria enormemente multiplicado, assim como seriam multiplicados os cuidados com os pressupostos de codificação e as dificuldades na análise dos resultados. Teoricamente, pelo menos, pode-se traçar o percurso da fonologia à pragmática, tanto via codificação como via interconexão de redes – as quais, nesse caso, funcionariam como módulos lingüísticos integrados – mas há que se estar preparado para tal. Há, nesse caso, uma demanda muito maior no que tange à capacidade de interpretação matemática, mas a ferramenta em si se presta, a princípio, a essa linha de investigação – tem flexibilidade implementativa e lógica interna para tal. O quão humana será essa lógica interna é o cerne da questão – o grande desafio do connexionismo em L2. Como na

metáfora do retroprojeter (seção 1.8), a adição de lâminas agrega potencial combinatório: mas também traz à tona o limite da discernibilidade, ou seja, o pesquisador terá de ser capaz de interpretar consistentemente as intrincadas relações entre esses subconjuntos de fatores, assim como escolher criteriosamente as lâminas a serem exibidas. Caberá a ele, portanto, integrar-se criativa e conscientemente ao mundo pulsante dos sistemas não lineares, sabendo analisar a soma das partes que transcende o todo. No universo da miríade de fatores envolvidos na aquisição de L2, essa parece ser uma opção bastante desafiadora e atraente.

BIBLIOGRAFIA CONSULTADA

- ANDERSEN, R. What's gender good for anyway?. In: ANDERSEN, R. (org.). *Second languages: A cross-linguistic perspective*. Rowley, MA: Newbury House, 1984.
- ANDERSON, J. *Rules of the mind*. Hillsdale, NJ: Lawrence Earlbaum, 1993.
- ANDERSON, J.; SCHOOLER, L. Reflections of the environment in memory. *Psychological Science*, v. 2, p. 396-408, 1991.
- BAKER, C. *English syntax*. Cambridge, MA: MIT Press, 1996.
- BATES, E.; MACWHINNEY, B. Competition, variation and language learning. In: MACWHINNEY, B. (org.). *Mechanisms of language acquisition*. London: Lawrence Earlbaum, 1987.
- BLACKWELL, A.; BROEDER, P. Interference and facilitation in SLA: a connectionist perspective. *Seminar on Parallel Distributed Processing and Natural Language Processing*. San Diego: UCSD, maio, 1992.
- BROEDER, P. *Talking about people: a multiple case study on adult language acquisition*. Amsterdam: Swets & Zeitlinger, 1991.
- BROEDER, P.; PLUNKETT, K. Connectionism and second language acquisition. In: Ellis, N. (org.). *Implicit and explicit learning of languages*. San Diego, CA: Academic Press, 1994.
- BYBEE, J. *Morphology: A study of the relation between meaning and form*. Philadelphia, PA: Benjamins, 1985.
- CHATER, N. Neural networks: the new statistical models of mind. In: LEVY, J.; BAIRAKTARIS, D.; BULLINARIA, J.; CAIRNS, P. (orgs.). *Connectionist models of memory and language*. London: UCL Press, 1995.
- CHRISTIANSEN, M.; CHATER, N. Connectionist natural language processing: the state of the art. *Cognitive Science*, v. 23, n. 4, p. 417-37, 1999.
- CORDER, S. The significance of learner's errors. *International Review of Applied Linguistics*, v. 5, p. 161-170, 1967.
- DIENES, Z.; ALTMANN, G.; GAO, S. Mapping across domains without feedback: a neural network model of transfer of implicit knowledge. *Cognitive Science*, v. 23, n. 1, p. 53-82, 1999.

- ELLIS, N.; SCHIMIDT, R. Rules or associations in the acquisition of morphology? The frequency by regularity interaction in human and PDP learning of morphosyntax. *Language and Cognitive Processes*, v.13, n.2-3, p. 307-36, 1998.
- ELLIS, R. *The study of second language acquisition*. Oxford: Oxford University Press, 1994.
- ELMAN, J. Finding structure in time. *Cognitive Science*, v. 14, p. 179-211, 1990.
- ELMAN, J. Learning and development in neural networks: the importance of starting small. *Cognition*, v. 48, p. 71-99, 1993.
- ELMAN, J.; BATES, E.; JOHNSON, M.; KARMILOFF-SMITH, A.; PARISI, D.; PLUNKETT, K. *Rethinking innateness: a connectionist perspective on development*. Cambridge, MA: MIT Press, 1996.
- FODOR, J.; PYLYSHYN, Z. Connectionism and cognitive architecture. *Cognition*, v. 28, p. 3-71, 1988.
- GASS, S.; SELINKER, L. *Second language acquisition: an introductory course*. Hillsdale, NJ: Lawrence Earlbaum, 1994.
- GASSER, M. Connectionism and universals of second language acquisition. *Studies in Second Language Acquisition*, v. 12, p. 179-99, 1990.
- GLEICK, J. *Chaos: Making a new science*. New York: Penguin Books, 1987.
- HAEGEMAN, L. *Introduction to government and binding theory*. 2.ed. Cambridge, MA: Blackwell, 1994.
- HATCH, E.; SHIRAI, Y.; FANTUZZI, C. The need for an integrated theory: connecting modules. *TESOL Quarterly*, v. 24, n. 4, p. 697-716, 1990.
- HYLTENSTAM, K. Implicational patterns in interlanguage syntax variation. *Language Learning*, v. 27, p. 383-411, 1977.
- JUNG, J. *The experimenter's dilemma*. New York: Harper & Row, 1971.
- KELLERMAN, E. Giving learners a break: native speaker intuitions as a source of predictions about transferability. *Working Papers in Bilingualism*, v. 15, p. 59-92, 1978.
- KRASHEN, S. *Principles and practice in second language acquisition*. London: Pergamon, 1982.
- LARSEN-FREEMAN, D. Chaos/Complexity science and second language acquisition. *Applied Linguistics*, v. 18, n. 2, p. 141-165, 1997.

- LEGENDRE, G.; MIYATA, Y.; SMOLENSKY, P. Harmonic grammar - a formal multi-level connectionist theory of linguistic well-formedness: theoretical foundations. In: 12th Annual Conference of the Cognitive Science Society, jul., 1990, Cambridge, MA. *Anais*. Cambridge, MA, 1990. p. 388-95.
- LIGHTBOWN, P. Great expectations: second language acquisition research and classroom teaching. *Applied Linguistics*, v. 6, p. 173-89, 1985.
- LIGHTBOWN, P.; Spada, N. *How languages are learned*. Oxford: Oxford University Press, 1993.
- MANZA, L.; REBER, A. Representation of tacit knowledge: transfer across stimulus forms and modalities. In: BERRY, D. (org.). *How implicit is implicit learning?*. Oxford: Oxford University Press, 1997.
- MARCUS, G. Can connectionism save constructivism?. *Cognition*, v. 66, p. 153-82, 1998.
- MAURER, J.; SCHOENBERG, I. *True Colors: an EFL course for real communication*. White Plains, NY: Longman, 1998. Vol. 1-3
- MCLAUGHLIN, B. Restructuring. *Applied Linguistics*, v. 11, p. 113-28, 1990.
- MEISEL, J.; CLAHSSEN, H.; PIENEMANN, M. On determining developmental stages in natural second language acquisition. *Studies in Second Language Acquisition*, v. 2-3, p. 109-35, 1981.
- NEWELL, A. *Unified theories of cognition*. Cambridge, MA: Harvard University Press, 1990.
- PIENEMANN, M.; JOHNSTON, M. A predictive framework of second language acquisition. Manuscript Original, 1987.
- PIENEMANN, M.; JOHNSTON, M.; BRINDLEY, G. Constructing acquisition-based procedure for language assessment. *Studies in Second Language Acquisition*, v. 10, p. 217-43, 1988.
- PINKER, S. *The language instinct: how the mind creates language*. New York: Harper Perennial, 1994.
- PINKER, S. Words and rules. *Lingua*, v. 106, 219-42, 1998.
- PINKER, S.; PRINCE, A. On language and connectionism: analysis of a parallel distributed processing model of language acquisition. *Cognition*, v. 28, p. 73-193, 1988.
- PLUNKETT, K. Connectionism today. Manuscript Original, 2000.

- PLUNKETT, K.; ELMAN, J. *Exercises in rethinking innateness: a handbook for connectionist simulations*. Cambridge, MA: MIT Press, 1997.
- PLUNKETT, K.; JUOLA, P. A connectionist model of English past tense and plural morphology. *Cognitive Science*, v. 23, n. 4, p. 463-90, 1999.
- PLUNKETT, K.; MARCHMAN, V. U-shaped learning and frequency effects in a multi-layered perceptron: implications for child language acquisition. *Cognition*, v. 38, p. 43-102, 1991.
- PRINCE, A.; SMOLENSKY, P. Notes on connectionism and Harmony Theory in linguistics. Relatório Técnico. Departamento de Ciências de Computação, Universidade do Colorado, Bolder, CO, 1991.
- PRINCE, A.; SMOLENSKY, P. Optimality: from neural networks to universal grammar. *Science*, v. 275, p. 1604-10, 1997.
- PURVES, D. The trophic theory of neural connections. *Trends in Neuroscience*, v. 9, p. 486-89, 1986.
- RICHARDS, J.; SANDY, C. *Passages: an upper-level multi-skills course*. Cambridge: Cambridge University Press, 1998. Vol. 1.
- ROSSI, M. Estudo diacrônico sobre as interrogativas do português do Brasil. In: ROBERTS, I.; KATO, M. *Português brasileiro: uma viagem diacrônica*. Campinas: Unicamp, 1996.
- RUMELHART, D., HINTON, G.; WILLIAMS, R. Learning internal representations by error propagation. In: RUMELHART, D.; MCCLELLAND, J. (org.). *Parallel Distributed Processing: Explorations in the microstructure of cognition*. Cambridge, MA: MIT Press, 1986. Vol. 1.
- RUMELHART, D.; MCCLELLAND, J. Learning the past tenses of English verbs: implicit rules or Parallel Distributed Processing?. In: MACWHINNEY, B. *Mechanisms of language acquisition*. London: Lawrence Earlbaum, 1987.
- SCHMIDT, R. Implicit learning and the cognitive unconscious: of artificial grammars and SLA. In: Ellis, N. (org.). *Implicit and explicit learning of languages*. San Diego, CA: Academic Press, 1994.
- SCHUMANN, J. Extending the scope of the acculturation / pidginization model to include cognition. *TESOL Quarterly*, v. 24, n. 4, p. 667-83, 1990.
- SEIDENBERG, M. Language and connectionism: the developing interface. *Cognition*, v. 50, p. 385-401, 1994.

- SHARWOOD SMITH, M. On first language loss in the second language acquirer: problems of transfer. In: GASS, S. & SELINKER, L. (org.). *Language transfer in language learning*. Rowley, MA: Newbury House, 1983.
- SMOLENSKY, P. Information processing in dynamical systems: foundations of Harmony Theory. In: RUMELHART, D.; J. MCCLEALLAND (org.). *Parallel Distributed Processing: explorations in the microstructure of cognition*. Cambridge, MA: MIT Press, 1986. Vol. 1.
- SMOLENSKY, P. Grammar-based connectionist approaches to language. *Cognitive Science*, v. 23, n. 4, p. 589-613, 1999.
- SOKOLIK, M. Learning without rules: PDP and the resolution of the adult language learning paradox. *TESOL Quarterly*, v. 24, n. 4, p. 685-96, 1990.
- SOKOLIK, M.; SMITH, M. Assignment of gender to French nouns in primary and secondary language: a connectionist model. *Second Language Research*, v. 8, n. 1, p. 39-58, 1992.
- STEEDMAN, M. Connectionist processing in perspective. *Cognitive Science*, v. 23, n. 4, p. 615-34, 1999.
- TABOR, W.; JULIANO, C.; TANENHAUS, M. Parsing in a dynamical system: an attractor-based account of the interaction of lexical and structural constraints in sentence processing. *Language and Cognitive Processes*, v. 12, p. 211-71, 1997.
- TATIBANA, C. & KAETSU, D. *Redes Neurais*. Disponível em: <<http://www.din.uem.br/ia/neurais>> Acesso em: 9 mar. 2001.
- THORNTON, C. Why concept learning is a good idea. In: CLARK, A.; MILLICAN, P. (org.). *Connectionism, concepts and folk psychology: the legacy of Allan Turing*. Oxford: Clarendon Press, 1996. Vol. 2.

ANEXOS

ANEXO A – INSTRUÇÕES FORNECIDAS AOS SUJEITOS**Instruções**

Você é um repórter esportivo.

Você tem que entrevistar o Ian Thorpe, um famoso nadador australiano.

Escreva, no formulário fornecido, dez perguntas EM INGLÊS para utilizar durante a entrevista.

Você pode utilizar os assuntos sugeridos na tabela, como também pode perguntar sobre outras coisas que quiser.

Termine as dez perguntas o mais rápido possível. NÃO volte a questões já escritas para reformulá-las.

Utilize suas próprias idéias e o seu inglês: NÃO copie dos colegas, NÃO consulte nenhum livro ou caderno e NÃO me pergunte nada.

Isto NÃO é um teste, é um experimento, portanto erros ou acertos NÃO fazem nenhuma diferença.

Após terminar as perguntas, preencha o questionário em português sobre você.

Tudo terminado, você ganha um bombom como prêmio de participação!

Thank you very much!!!

ANEXO B – FORMULÁRIO DE COLETA DE DADOS



Ian Thorpe

habits	girls
future	family
hometown	food
hobbies	Olympics
clothes	traveling
friends	school



Questions for the interview:

1. _____
2. _____
3. _____
4. _____
5. _____
6. _____
7. _____
8. _____
9. _____
10. _____

ANEXO C – QUESTIONÁRIO APLICADO AOS SUJEITOS

Questionário

Nome: _____

Idade: _____

Nível: _____

1. Há quantos semestres você estuda no Cultural? (inclua este semestre)

Há _____ semestre(s)

2. Você já interrompeu este curso alguma vez?

(a) Não.

(b) Sim. Especifique: _____

3. Você estuda (ou estudou) inglês na escola?

(a) Não.

(b) Sim. Comecei na _____ série do ensino (fundamental / médio) e tive inglês em todos os anos. Hoje estou na _____ série.

(c) Sim. Comecei na _____ série do ensino (fundamental / médio) mas não tive inglês por _____ anos no total. Hoje estou na _____ série.

Você já está na faculdade? () Não () Sim, estou cursando _____.

4. Você já estudou inglês fora do Cultural com professor particular ou em outros cursos?

(a) Não.

(b) Sim. Especifique: _____

Por quanto tempo e com que frequência? _____

Com que idade iniciou? _____

5. Você já estudou outras línguas além do português e do inglês?

(a) Não.

(b) Sim. Especifique: _____

Por quanto tempo e com que frequência? _____

Com que idade iniciou? _____

			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
			NPsubj					VPaux						VPmain							
			Per		Num			Per			Num			Ten			Ten/Agr				
			1st	2nd	3rd	Sing	Plur	1st	2nd	3rd	Sing	Plur	Pres	Pas	Fut	Pres	Pas	Pr Part	Pa Part	Stem	
A Present Simple																					
1	I	am	1	0	0	1	0	1	0	0	1	0	1	0	0	0	0	0	0	0	
2	you	are	0	1	0	1	0	0	1	0	1	0	1	0	0	0	0	0	0	0	
3	he	is	0	0	1	1	0	0	0	1	1	0	1	0	0	0	0	0	0	0	
4	we	are	1	0	0	0	1	1	0	0	0	1	1	0	0	0	0	0	0	0	
5	you	are	0	1	0	0	1	0	1	0	0	1	1	0	0	0	0	0	0	0	
6	they	are	0	0	1	0	1	0	0	1	0	1	1	0	0	0	0	0	0	0	
7	I	talk	1	0	0	1	0	1	0	0	1	0	0	0	0	1	0	0	0	0	
8	you	talk	0	1	0	1	0	0	1	0	1	0	0	0	0	1	0	0	0	0	
9	he	talks	0	0	1	1	0	0	0	1	1	0	0	0	0	1	0	0	0	0	
10	we	talk	1	0	0	0	1	1	0	0	0	1	0	0	0	1	0	0	0	0	
11	you	talk	0	1	0	0	1	0	1	0	0	1	0	0	0	1	0	0	0	0	
12	they	talk	0	0	1	0	1	0	0	1	0	1	0	0	0	1	0	0	0	0	
B Present Continuous																					
13	I	am talking	1	0	0	1	0	1	0	0	1	0	1	0	0	0	0	1	0	0	
14	you	are talking	0	1	0	1	0	0	1	0	1	0	1	0	0	0	0	1	0	0	
15	he	is talking	0	0	1	1	0	0	0	1	1	0	1	0	0	0	0	1	0	0	
16	we	are talking	1	0	0	0	1	1	0	0	0	1	1	0	0	0	0	1	0	0	
17	you	are talking	0	1	0	0	1	0	1	0	0	1	1	0	0	0	0	1	0	0	
18	they	are talking	0	0	1	0	1	0	0	1	0	1	1	0	0	0	0	1	0	0	
C Modal																					
19	I	can talk	1	0	0	1	0	1	0	0	1	0	1	1	1	0	0	0	0	1	
20	you	can talk	0	1	0	1	0	0	1	0	1	0	1	1	1	0	0	0	0	1	
21	he	can talk	0	0	1	1	0	0	0	1	1	0	1	1	1	0	0	0	0	1	
22	we	can talk	1	0	0	0	1	1	0	0	0	1	1	1	1	0	0	0	0	1	
23	you	can talk	0	1	0	0	1	0	1	0	0	1	1	1	1	0	0	0	0	1	
24	they	can talk	0	0	1	0	1	0	0	1	0	1	1	1	1	0	0	0	0	1	

D Past Simple

			NPsubj			VPaux									VPmain				
			Per		Num	Per			Num		Ten			Ten/Agr					
1st	2nd	3rd	Sing	Plur	1st	2nd	3rd	Sing	Plur	Pres	Pas	Fut	Pres	Pas	Pr-Part	Pa Part	Stem		
25	I	was	1	0	0	1	0	1	0	0	1	0	0	0	0	0	0	0	
26	you	were	0	1	0	1	0	0	1	0	0	1	0	0	0	0	0	0	
27	he	was	0	0	1	1	0	0	0	1	1	0	0	0	0	0	0	0	
28	we	were	1	0	0	0	1	1	0	0	0	1	0	0	0	0	0	0	
29	you	were	0	1	0	0	1	0	1	0	0	1	0	0	0	0	0	0	
30	they	were	0	0	1	0	1	0	0	1	0	1	0	0	0	0	0	0	
31	I	talked	1	0	0	1	0	1	0	0	0	0	0	0	1	0	0	0	
32	you	talked	0	1	0	1	0	0	1	0	0	0	0	0	1	0	0	0	
33	he	talked	0	0	1	1	0	0	0	1	1	0	0	0	1	0	0	0	
34	we	talked	1	0	0	0	1	1	0	0	0	1	0	0	1	0	0	0	
35	you	talked	0	1	0	0	1	0	1	0	0	0	0	1	0	0	0	0	
36	they	talked	0	0	1	0	1	0	0	1	0	0	0	1	0	0	0	0	

E Past Continuous

37	I	was	1	0	0	1	0	1	0	0	1	0	0	0	0	1	0	0
38	you	were	0	1	0	1	0	0	1	0	0	1	0	0	0	1	0	0
39	he	was	0	0	1	1	0	0	0	1	1	0	0	0	1	0	0	0
40	we	were	1	0	0	0	1	1	0	0	0	1	0	0	1	0	0	0
41	you	were	0	1	0	0	1	0	1	0	0	1	0	0	1	0	0	0
42	they	were	0	0	1	0	1	0	0	1	0	1	0	0	1	0	0	0

F Present Perfect

43	I	have	1	0	0	1	0	1	0	0	1	0	0	0	0	0	1	0
44	you	have	0	1	0	1	0	0	1	0	1	0	0	0	0	0	1	0
45	he	has	0	0	1	1	0	0	0	1	1	0	0	0	0	0	1	0
46	we	have	1	0	0	0	1	1	0	0	0	1	0	0	0	0	1	0
47	you	have	0	1	0	0	1	0	1	0	1	0	0	0	0	0	1	0
48	they	have	0	0	1	0	1	0	0	1	1	0	0	0	0	0	1	0

G Past Perfect

49	I	had	1	0	0	1	0	1	0	0	0	1	0	0	0	0	1	0
50	you	had	0	1	0	1	0	0	1	0	0	1	0	0	0	0	1	0
51	he	had	0	0	1	1	0	0	0	1	1	0	0	0	0	0	1	0
52	we	had	1	0	0	0	1	1	0	0	0	1	0	0	0	0	1	0
53	you	had	0	1	0	0	1	0	1	0	0	1	0	0	0	0	1	0
54	they	had	0	0	1	0	1	0	0	1	0	1	0	0	0	0	1	0

			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
			VPaux										NPsuj					VPmain					
			Forms										Forms					Forms					
			am	are	is	was	were	do	does	dId	have	has	had	mod	I	you	he	we	they	st	st + ed	lrr	st + Ing
A Present Simple																							
1	am	I	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
2	are	you	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
3	is	he	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
4	are	we	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
5	are	you	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
6	are	they	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
7	do	I	talk	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	1	0	0	0
8	do	you	talk	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0
9	does	he	talk	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	1	0	0	0
10	do	we	talk	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	1	0	0	0
11	do	you	talk	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0
12	do	they	talk	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	1	0	0	0
				0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0
B Present Continuous																							
13	am	I	talk	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
14	are	you	talk	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
15	is	he	talk	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
16	are	we	talk	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
17	are	you	talk	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
18	are	they	talk	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
				0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
C Modal																							
19	can	I	talk	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	1	0	0	0
20	can	you	talk	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1	0	0	0
21	can	he	talk	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0
22	can	we	talk	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	1	0	0	0
23	can	you	talk	0	0	0	0	0	0	0	0	0	0	1	0	1	0	1	0	1	0	0	0
24	can	they	talk	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0
				0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	1	0	0	0

				VPaux									NPsuj					VPmain						
				Forms									Forms					Forms						
				am	are	is	was	were	do	does	did	have	has	had	mod	I	you	he	we	they	st	st + ed	irr	st + Ing
D Past Simple																								
25	was	I		0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	
26	were	you		0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	
27	was	he		0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	
28	were	we		0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	
29	were	you		0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	
30	were	they		0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	
31	did	I	talk	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0	0	
32	did	you	talk	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	1	0	0	0	
33	did	he	talk	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0	0	
34	did	we	talk	0	0	0	0	0	0	0	1	0	0	0	0	0	1	0	0	1	0	0	0	
35	did	you	talk	0	0	0	0	0	0	0	1	0	0	0	0	1	0	1	0	1	0	0	0	
36	did	they	talk	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0	
E Past Continuous																								
37	was	I	talking	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	
38	were	you	talking	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	
39	was	he	talking	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	
40	were	we	talking	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	
41	were	you	talking	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	
42	were	they	talking	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	
F Present Perfect																								
43	have	I	talked	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	1	0	
44	have	you	talked	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1	1	0	
45	has	he	talked	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	1	0	
46	have	we	talked	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	1	1	1	0	
47	have	you	talked	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	1	1	0	
48	have	they	talked	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	1	1	0	
G Past Perfect																								
49	had	I	talked	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	1	1	1	0	
50	had	you	talked	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	1	1	0	
51	had	he	talked	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	1	1	0	
52	had	we	talked	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	1	1	1	0	
53	had	you	talked	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1	1	1	0	
54	had	they	talked	0	0	0	0	0	0	0	0	0	1	0	0	1	0	0	0	1	1	1	0	

ANEXO F - SAÍDAS DA REDE

- Group A (Present Simple)

Códigos	Saídas			Saídas esperadas		
<i>Output activations using t-groupA-13.wts and groupA.data (Training Set) using Output Translation mapp.out (Euclidean Distance)</i>						
m q x (target: a n x)		we		am	I	
b s x (target: b o x)	ar			ar	yu	
m s x (target: c p x)				is	he	
b q x (target: b q x)	ar	we		ar	we	
b q x (target: b o x)	ar	we		ar	yu	
b q x (target: b r x)	ar	we		ar	th	
m q x (target: f n t)		we		do	I	st
m s x (target: f o t)				do	yu	st
m s x (target: g p t)				ds	he	st
m q x (target: f q t)		we		do	we	st
b q x (target: f o t)	ar	we		do	yu	st
m q x (target: f r t)		we		do	th	st
<i>Output activations using t-groupA-25.wts and groupA.data (Training Set) using Output Translation mapp.out (Euclidean Distance)</i>						
m s x (target: a n x)				am	I	
c s x (target: b o x)	is			ar	yu	
c p x (target: c p x)	is	he		is	he	
f q x (target: b q x)	do	we		ar	we	
m s x (target: b o x)				ar	yu	
c s x (target: b r x)	is			ar	th	
f s t (target: f n t)	do		st	do	I	st
f s x (target: f o t)	do			do	yu	st
c p t (target: g p t)	is	he	st	ds	he	st
f q t (target: f q t)	do	we	st	do	we	st
f q x (target: f o t)	do	we		do	yu	st
f s t (target: f r t)	do		st	do	th	st
<i>Output activations using t-groupA-37.wts and groupA.data (Training Set) using Output Translation mapp.out (Euclidean Distance)</i>						
b s x (target: a n x)	ar			am	I	
b o x (target: b o x)	ar	yu		ar	yu	
b o x (target: c p x)	ar	yu		is	he	
b q x (target: b q x)	ar	we		ar	we	
b q x (target: b o x)	ar	we		ar	yu	
b s x (target: b r x)	ar			ar	th	
f s t (target: f n t)	do		st	do	I	st
b o x (target: f o t)	ar	yu		do	yu	st
m s t (target: g p t)			st	ds	he	st
f q t (target: f q t)	do	we	st	do	we	st
b q x (target: f o t)	ar	we		do	yu	st
m q t (target: f r t)		we	st	do	th	st
<i>Output activations using t-groupA-49.wts and groupA.data (Training Set) using Output Translation mapp.out (Euclidean Distance)</i>						
m o x (target: a n x)		yu		am	I	
m o x (target: b o x)		yu		ar	yu	
m o x (target: c p x)		yu		is	he	

box (target: b q x)	ar	yu		ar	we	
box (target: b o x)	ar	yu		ar	yu	
mo x (target: b r x)		yu		ar	th	
f s t (target: f n t)	do		st	do	I	st
f o t (target: f o t)	do	yu	st	do	yu	st
m p t (target: g p t)		he	st	ds	he	st
f s t (target: f q t)	do		st	do	we	st
f o t (target: f o t)	do	yu	st	do	yu	st
f s t (target: f r t)	do		st	do	th	st
<i>Output activations using t-groupA-61.wts and groupA.data (Training Set) using Output Translation mapp.out (Euclidean Distance)</i>						
box (target: a n x)	ar	yu		am	I	
box (target: b o x)	ar	yu		ar	yu	
mo x (target: c p x)		yu		is	he	
b q x (target: b q x)	ar	we		ar	we	
box (target: b o x)	ar	yu		ar	yu	
b s x (target: b r x)	ar			ar	th	
f s t (target: f n t)	do		st	do	I	st
f o t (target: f o t)	do	yu	st	do	yu	st
m p t (target: g p t)		he	st	ds	he	st
f q t (target: f q t)	do	we	st	do	we	st
f o t (target: f o t)	do	yu	st	do	yu	st
f s t (target: f r t)	do		st	do	th	st
<i>Output activations using t-groupA-73.wts and groupA.data (Training Set) using Output Translation mapp.out (Euclidean Distance)</i>						
b s x (target: a n x)	ar			am	I	
box (target: b o x)	ar	yu		ar	yu	
c p x (target: c p x)	is	he		is	he	
b q x (target: b q x)	ar	we		ar	we	
box (target: b o x)	ar	yu		ar	yu	
b s x (target: b r x)	ar			ar	th	
f s t (target: f n t)	do		st	do	I	st
f o t (target: f o t)	do	yu	st	do	yu	st
m p t (target: g p t)		he	st	ds	he	st
f q t (target: f q t)	do	we	st	do	we	st
f o t (target: f o t)	do	yu	st	do	yu	st
f s t (target: f r t)	do		st	do	th	st
<i>Output activations using t-groupA-85.wts and groupA.data (Training Set) using Output Translation mapp.out (Euclidean Distance)</i>						
m s x (target: a n x)				am	I	
box (target: b o x)	ar	yu		ar	yu	
c p x (target: c p x)	is	he		is	he	
b q x (target: b q x)	ar	we		ar	we	
box (target: b o x)	ar	yu		ar	yu	
b r x (target: b r x)	ar	th		ar	th	
f s t (target: f n t)	do		st	do	I	st
f o t (target: f o t)	do	yu	st	do	yu	st
g p t (target: g p t)	ds	he	st	ds	he	st
f q t (target: f q t)	do	we	st	do	we	st
f o t (target: f o t)	do	yu	st	do	yu	st
f s t (target: f r t)	do		st	do	th	st
<i>Output activations using t-groupA-97.wts and groupA.data (Training Set) using Output Translation mapp.out (Euclidean Distance)</i>						
m n x (target: a n x)		I		am	I	
box (target: b o x)	ar	yu		ar	yu	
c p x (target: c p x)	is	he		is	he	

b q x (target: b q x)	ar	we		ar	we	
b o x (target: b o x)	ar	yu		ar	yu	
b s x (target: b r x)	ar			ar	th	
f s t (target: f n t)	do		st	do	I	st
f o t (target: f o t)	do	yu	st	do	yu	st
g p t (target: g p t)	ds	he	st	ds	he	st
f q t (target: f q t)	do	we	st	do	we	st
f o t (target: f o t)	do	yu	st	do	yu	st
f s t (target: f r t)	do		st	do	th	st

Output activations using t-groupA-109.wts and groupA.data (Training Set) using Output Translation mapp.out (Euclidean Distance)

m n x (target: a n x)		I		am	I	
b o x (target: b o x)	ar	yu		ar	yu	
m p x (target: c p x)		he		is	he	
b q x (target: b q x)	ar	we		ar	we	
b o x (target: b o x)	ar	yu		ar	yu	
b r x (target: b r x)	ar	th		ar	th	
f n t (target: f n t)	do	I	st	do	I	st
f o t (target: f o t)	do	yu	st	do	yu	st
g p t (target: g p t)	ds	he	st	ds	he	st
f q t (target: f q t)	do	we	st	do	we	st
f o t (target: f o t)	do	yu	st	do	yu	st
f r t (target: f r t)	do	th	st	do	th	st

Output activations using t-groupA-120.wts and groupA.data (Training Set) using Output Translation mapp.out (Euclidean Distance)

a n x (target: a n x)	am	I		am	I	
b o x (target: b o x)	ar	yu		ar	yu	
m p x (target: c p x)		he		is	he	
b q x (target: b q x)	ar	we		ar	we	
b o x (target: b o x)	ar	yu		ar	yu	
b r x (target: b r x)	ar	th		ar	th	
f n t (target: f n t)	do	I	st	do	I	st
f o t (target: f o t)	do	yu	st	do	yu	st
f p t (target: g p t)	do	he	st	ds	he	st
f q t (target: f q t)	do	we	st	do	we	st
f o t (target: f o t)	do	yu	st	do	yu	st
f r t (target: f r t)	do	th	st	do	th	st

Output activations using t-groupA-133.wts and groupA.data (Training Set) using Output Translation mapp.out (Euclidean Distance)

m n x (target: a n x)		I		am	I	
b o x (target: b o x)	ar	yu		ar	yu	
c p x (target: c p x)	is	he		is	he	
b q x (target: b q x)	ar	we		ar	we	
b o x (target: b o x)	ar	yu		ar	yu	
b r x (target: b r x)	ar	th		ar	th	
f n t (target: f n t)	do	I	st	do	I	st
f o t (target: f o t)	do	yu	st	do	yu	st
f p t (target: g p t)	do	he	st	ds	he	st
f q t (target: f q t)	do	we	st	do	we	st
f o t (target: f o t)	do	yu	st	do	yu	st
f r t (target: f r t)	do	th	st	do	th	st

Output activations using t-groupA-145.wts and groupA.data (Training Set) using Output Translation mapp.out (Euclidean Distance)

a n x (target: a n x)	am	I		am	I	
b o x (target: b o x)	ar	yu		ar	yu	

c p x (target: c p x)	is	he		is	he	
b q x (target: b q x)	ar	we		ar	we	
b o x (target: b o x)	ar	yu		ar	yu	
b r x (target: b r x)	ar	th		ar	th	
f n t (target: f n t)	do	I	st	do	I	st
f o t (target: f o t)	do	yu	st	do	yu	st
g p t (target: g p t)	ds	he	st	ds	he	st
f q t (target: f q t)	do	we	st	do	we	st
f o t (target: f o t)	do	yu	st	do	yu	st
f r t (target: f r t)	do	th	st	do	th	st

- **Group B (Present Continuous)**

Códigos	Saídas			Saídas esperadas		
<i>Output activations using t-groupB-164.wts and groupB.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
m n x (target: a n w)		I		am	I	ng
b o x (target: b o w)	ar	yu		ar	yu	ng
c p w (target: c p w)	is	he	ng	is	he	ng
b q x (target: b q w)	ar	we		ar	we	ng
b o x (target: b o w)	ar	yu		ar	yu	ng
b r x (target: b r w)	ar	th		ar	th	ng
<i>Output activations using t-groupB-182.wts and groupB.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
a n x (target: a n w)	am	I		am	I	ng
b o w (target: b o w)	ar	yu	ng	ar	yu	ng
c p w (target: c p w)	is	he	ng	is	he	ng
b q x (target: b q w)	ar	we		ar	we	ng
b o w (target: b o w)	ar	yu	ng	ar	yu	ng
b r w (target: b r w)	ar	th	ng	ar	th	ng
<i>Output activations using t-groupB-200.wts and groupB.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
a n x (target: a n w)	am	I		am	I	ng
b o x (target: b o w)	ar	yu		ar	yu	ng
c p w (target: c p w)	is	he	ng	is	he	ng
b q x (target: b q w)	ar	we		ar	we	ng
b o x (target: b o w)	ar	yu		ar	yu	ng
b r x (target: b r w)	ar	th		ar	th	ng
<i>Output activations using t-groupB-218.wts and groupB.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
a n w (target: a n w)	am	I	ng	am	I	ng
b o w (target: b o w)	ar	yu	ng	ar	yu	ng
c p w (target: c p w)	is	he	ng	is	he	ng
b q w (target: b q w)	ar	we	ng	ar	we	ng
b o w (target: b o w)	ar	yu	ng	ar	yu	ng
b r w (target: b r w)	ar	th	ng	ar	th	ng

- **Group C (Modal)**

Códigos	Saídas			Saídas esperadas		
<i>Output activations using t-groupC-243.wts and groupC.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
l n t (target: l n t)	md	I	st	md	I	st
l o t (target: l o t)	md	yu	st	md	yu	st
m p x (target: l p t)		he		md	he	st
m q t (target: l q t)		we	st	md	we	st
l o t (target: l o t)	md	yu	st	md	yu	st
m r x (target: l r t)		th		md	th	st
<i>Output activations using t-groupC-267.wts and groupC.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
l n t (target: l n t)	md	I	st	md	I	st
l o t (target: l o t)	md	yu	st	md	yu	st
l p t (target: l p t)	md	he	st	md	he	st
f q t (target: l q t)	do	we	st	md	we	st
l o t (target: l o t)	md	yu	st	md	yu	st
m r t (target: l r t)		th	st	md	th	st
<i>Output activations using t-groupC-291.wts and groupC.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
l n t (target: l n t)	md	I	st	md	I	st
l o t (target: l o t)	md	yu	st	md	yu	st
l p t (target: l p t)	md	he	st	md	he	st
l q t (target: l q t)	md	we	st	md	we	st
l o t (target: l o t)	md	yu	st	md	yu	st
l r t (target: l r t)	md	th	st	md	th	st

- **Group D (Past Simple)**

Códigos	Saídas			Saídas esperadas		
<i>Output activations using t-groupD-328.wts and groupD.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
m n x (target: d n x)		I		ws	I	
m o x (target: e o x)		yu		wr	yu	
m p x (target: d p x)		he		ws	he	
e q x (target: e q x)	wr	we		wr	we	
m o x (target: e o x)		yu		wr	yu	
m r x (target: e r x)		th		wr	th	
m n x (target: h n t)		I		dd	I	st
m o x (target: h o t)		yu		dd	yu	st
m p x (target: h p t)		he		dd	he	st
e q x (target: h q t)	wr	we		dd	we	st
e o x (target: h o t)	wr	yu		dd	yu	st
m r x (target: h r t)		th		dd	th	st
<i>Output activations using t-groupD-364.wts and groupD.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
m n x (target: d n x)		I		ws	I	
m o t (target: e o x)		yu	st	wr	yu	
m p t (target: d p x)		he	st	ws	he	
e q x (target: e q x)	wr	we		wr	we	
e o x (target: e o x)	wr	yu		wr	yu	
e r x (target: e r x)	wr	th		wr	th	
m n t (target: h n t)		I	st	dd	I	st
m o t (target: h o t)		yu	st	dd	yu	st

g p t (target: h p t)	ds	he	st	dd	he	st
e q x (target: h q t)	wr	we		dd	we	st
e o t (target: h o t)	wr	yu	st	dd	yu	st
m r t (target: h r t)		th	st	dd	th	st

Output activations using t-groupD-400.wts and groupD.data (Novel Data *) using Output Translation mapp.out (Euclidean Distance)*

m n x (target: d n x)		I		ws	I	
m o x (target: e o x)		yu		wr	yu	
m p x (target: d p x)		he		ws	he	
e q x (target: e q x)	wr	we		wr	we	
m o x (target: e o x)		yu		wr	yu	
e r x (target: e r x)	wr	th		wr	th	
m n t (target: h n t)		I	st	dd	I	st
m o t (target: h o t)		yu	st	dd	yu	st
m p t (target: h p t)		he	st	dd	he	st
e q x (target: h q t)	wr	we		dd	we	st
m o t (target: h o t)		yu	st	dd	yu	st
e r x (target: h r t)	wr	th		dd	th	st

Output activations using t-groupD-436.wts and groupD.data (Novel Data *) using Output Translation mapp.out (Euclidean Distance)*

m n x (target: d n x)		I		ws	I	
m o x (target: e o x)		yu		wr	yu	
m p t (target: d p x)		he	st	ws	he	
e q x (target: e q x)	wr	we		wr	we	
m o x (target: e o x)		yu		wr	yu	
m r x (target: e r x)		th		wr	th	
h n t (target: h n t)	dd	I	st	dd	I	st
h o t (target: h o t)	dd	yu	st	dd	yu	st
h p t (target: h p t)	dd	he	st	dd	he	st
m q t (target: h q t)		we	st	dd	we	st
m o t (target: h o t)		yu	st	dd	yu	st
h r t (target: h r t)	dd	th	st	dd	th	st

Output activations using t-groupD-472.wts and groupD.data (Novel Data *) using Output Translation mapp.out (Euclidean Distance)*

m n x (target: d n x)		I		ws	I	
m o x (target: e o x)		yu		wr	yu	
m p x (target: d p x)		he		ws	he	
e q x (target: e q x)	wr	we		wr	we	
e o x (target: e o x)	wr	yu		wr	yu	
e r x (target: e r x)	wr	th		wr	th	
h n t (target: h n t)	dd	I	st	dd	I	st
m o t (target: h o t)		yu	st	dd	yu	st
h p t (target: h p t)	dd	he	st	dd	he	st
h q t (target: h q t)	dd	we	st	dd	we	st
m o t (target: h o t)		yu	st	dd	yu	st
h r t (target: h r t)	dd	th	st	dd	th	st

Output activations using t-groupD-508.wts and groupD.data (Novel Data *) using Output Translation mapp.out (Euclidean Distance)*

m n t (target: d n x)		I	st	ws	I	
m o x (target: e o x)		yu		wr	yu	
m p t (target: d p x)		he	st	ws	he	
m q x (target: e q x)		we		wr	we	
m o x (target: e o x)		yu		wr	yu	
m r x (target: e r x)		th		wr	th	
h n t (target: h n t)	dd	I	st	dd	I	st
h o t (target: h o t)	dd	yu	st	dd	yu	st

h p t (target: h p t)	dd	he	st	dd	he	st
h q t (target: h q t)	dd	we	st	dd	we	st
h o t (target: h o t)	dd	yu	st	dd	yu	st
h r t (target: h r t)	dd	th	st	dd	th	st
<i>Output activations using t-groupD-544.wts and groupD.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
m n x (target: d n x)		I		ws	I	
m o x (target: e o x)		yu		wr	yu	
m p x (target: d p x)		he		ws	he	
e q x (target: e q x)	wr	we		wr	we	
e o x (target: e o x)	wr	yu		wr	yu	
e r x (target: e r x)	wr	th		wr	th	
h n t (target: h n t)	dd	I	st	dd	I	st
h o t (target: h o t)	dd	yu	st	dd	yu	st
h p t (target: h p t)	dd	he	st	dd	he	st
h q t (target: h q t)	dd	we	st	dd	we	st
h o t (target: h o t)	dd	yu	st	dd	yu	st
h r t (target: h r t)	dd	th	st	dd	th	st
<i>Output activations using t-groupD-580.wts and groupD.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
m n x (target: d n x)		I		ws	I	
m o x (target: e o x)		yu		wr	yu	
m p x (target: d p x)		he		ws	he	
m q x (target: e q x)		we		wr	we	
m o x (target: e o x)		yu		wr	yu	
e r x (target: e r x)	wr	th		wr	th	
h n t (target: h n t)	dd	I	st	dd	I	st
h o t (target: h o t)	dd	yu	st	dd	yu	st
h p t (target: h p t)	dd	he	st	dd	he	st
h q t (target: h q t)	dd	we	st	dd	we	st
h o t (target: h o t)	dd	yu	st	dd	yu	st
h r t (target: h r t)	dd	th	st	dd	th	st
<i>Output activations using t-groupD-616.wts and groupD.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
m n x (target: d n x)		I		ws	I	
e o x (target: e o x)	wr	yu		wr	yu	
m p x (target: d p x)		he		ws	he	
e q x (target: e q x)	wr	we		wr	we	
e o x (target: e o x)	wr	yu		wr	yu	
e r x (target: e r x)	wr	th		wr	th	
h n t (target: h n t)	dd	I	st	dd	I	st
h o t (target: h o t)	dd	yu	st	dd	yu	st
h p t (target: h p t)	dd	he	st	dd	he	st
h q t (target: h q t)	dd	we	st	dd	we	st
h o t (target: h o t)	dd	yu	st	dd	yu	st
h r t (target: h r t)	dd	th	st	dd	th	st
<i>Output activations using t-groupD-651.wts and groupD.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
d n x (target: d n x)	ws	I		ws	I	
e o x (target: e o x)	wr	yu		wr	yu	
d p x (target: d p x)	ws	he		ws	he	
e q x (target: e q x)	wr	we		wr	we	
e o x (target: e o x)	wr	yu		wr	yu	
e r x (target: e r x)	wr	th		wr	th	
h n t (target: h n t)	dd	I	st	dd	I	st
h o t (target: h o t)	dd	yu	st	dd	yu	st

h p t (target: h p t)	dd	he	st	dd	he	st
h q t (target: h q t)	dd	we	st	dd	we	st
h o t (target: h o t)	dd	yu	st	dd	yu	st
h r t (target: h r t)	dd	th	st	dd	th	st

- **Group E (Past Continuous)**

Códigos	Saídas			Saídas esperadas		
<i>Output activations using t-groupE-694.wts and groupE.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
d n x (target: d n w)	ws	I		ws	I	ng
e o x (target: e o w)	wr	yu		wr	yu	ng
c p x (target: d p w)	is	he		ws	he	ng
e q x (target: e q w)	wr	we		wr	we	ng
e o x (target: e o w)	wr	yu		wr	yu	ng
e r x (target: e r w)	wr	th		wr	th	ng
<i>Output activations using t-groupE-736.wts and groupE.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
d n w (target: d n w)	ws	I	ng	ws	I	ng
e o x (target: e o w)	wr	yu		wr	yu	ng
m p w (target: d p w)		he	ng	ws	he	ng
e q x (target: e q w)	wr	we		wr	we	ng
e o x (target: e o w)	wr	yu		wr	yu	ng
e r x (target: e r w)	wr	th		wr	th	ng
<i>Output activations using t-groupE-778.wts and groupE.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
d n w (target: d n w)	ws	I	ng	ws	I	ng
e o x (target: e o w)	wr	yu		wr	yu	ng
d p w (target: d p w)	ws	he	ng	ws	he	ng
e q w (target: e q w)	wr	we	ng	wr	we	ng
e o x (target: e o w)	wr	yu		wr	yu	ng
e r w (target: e r w)	wr	th	ng	wr	th	ng
<i>Output activations using t-groupE-820.wts and groupE.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
d n x (target: d n w)	ws	I		ws	I	ng
e o x (target: e o w)	wr	yu		wr	yu	ng
d p w (target: d p w)	ws	he	ng	ws	he	ng
e q x (target: e q w)	wr	we		wr	we	ng
e o x (target: e o w)	wr	yu		wr	yu	ng
e r x (target: e r w)	wr	th		wr	th	ng
<i>Output activations using t-groupE-862.wts and groupE.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
d n w (target: d n w)	ws	I	ng	ws	I	ng
e o w (target: e o w)	wr	yu	ng	wr	yu	ng
d p w (target: d p w)	ws	he	ng	ws	he	ng
e q w (target: e q w)	wr	we	ng	wr	we	ng
e o w (target: e o w)	wr	yu	ng	wr	yu	ng
e r w (target: e r w)	wr	th	ng	wr	th	ng

- **Group F (Present Perfect)**

Códigos	Saídas			Saídas esperadas		
<i>Output activations using t-groupF-911.wts and groupF.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
anx (target: inv)	am	I		hv	I	ir/ed
iou (target: iov)	hv	yu	ed	hv	yu	ir/ed
mpx (target: jpv)		he		hs	he	ir/ed
iqx (target: iqv)	hv	we		hv	we	ir/ed
iou (target: iov)	hv	yu	ed	hv	yu	ir/ed
mr x (target: ir v)		th		hv	th	ir/ed
<i>Output activations using t-groupF-959.wts and groupF.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
anx (target: inv)	am	I		hv	I	ir/ed
iou (target: iov)	hv	yu	ed	hv	yu	ir/ed
mpx (target: jpv)		he		hs	he	ir/ed
m q x (target: i q v)		we		hv	we	ir/ed
iou (target: iov)	hv	yu	ed	hv	yu	ir/ed
mr x (target: ir v)		th		hv	th	ir/ed
<i>Output activations using t-groupF-1007.wts and groupF.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
anx (target: inv)	am	I		hv	I	ir/ed
mo x (target: i o v)		yu		hv	yu	ir/ed
mpx (target: jpv)		he		hs	he	ir/ed
i q v (target: i q v)	hv	we	ir	hv	we	ir/ed
mo x (target: i o v)		yu		hv	yu	ir/ed
mr x (target: ir v)		th		hv	th	ir/ed
<i>Output activations using t-groupF-1055.wts and groupF.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
anu (target: inv)	am	I	ed	hv	I	ir/ed
iou (target: iov)	hv	yu	ed	hv	yu	ir/ed
mpv (target: jpv)		he	ir	hs	he	ir/ed
i q v (target: i q v)	hv	we	ir	hv	we	ir/ed
iou (target: iov)	hv	yu	ed	hv	yu	ir/ed
mr u (target: ir v)		th	ed	hv	th	ir/ed
<i>Output activations using t-groupF-1103.wts and groupF.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
anu (target: inv)	am	I	ed	hv	I	ir/ed
iou (target: iov)	hv	yu	ed	hv	yu	ir/ed
jp v (target: j p v)	hs	he	ir	hs	he	ir/ed
i q v (target: i q v)	hv	we	ir	hv	we	ir/ed
iou (target: iov)	hv	yu	ed	hv	yu	ir/ed
ir v (target: ir v)	hv	th	ir	hv	th	ir/ed
<i>Output activations using t-groupF-1151.wts and groupF.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
anu (target: inv)	am	I	ed	hv	I	ir/ed
iou (target: iov)	hv	yu	ed	hv	yu	ir/ed
mp u (target: j p v)		he	ed	hs	he	ir/ed
i q u (target: i q v)	hv	we	ed	hv	we	ir/ed
iou (target: iov)	hv	yu	ed	hv	yu	ir/ed
mr v (target: ir v)		th	ir	hv	th	ir/ed
<i>Output activations using t-groupF-1199.wts and groupF.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
anu (target: inv)	am	I	ed	hv	I	ir/ed
iou (target: iov)	hv	yu	ed	hv	yu	ir/ed
mp u (target: j p v)		he	ed	hs	he	ir/ed
i q u (target: i q v)	hv	we	ed	hv	we	ir/ed

io u (target: io v)	hv	yu	ed	hv	yu	ir/ed
ir v (target: ir v)	hv	th	ir	hv	th	ir/ed
<i>Output activations using t-groupF-1247.wts and groupF.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
in v (target: in v)	hv	I	ir	hv	I	ir/ed
io u (target: io v)	hv	yu	ed	hv	yu	ir/ed
mp v (target: jp v)		he	ir	hs	he	ir/ed
iq v (target: iq v)	hv	we	ir	hv	we	ir/ed
io u (target: io v)	hv	yu	ed	hv	yu	ir/ed
ir v (target: ir v)	hv	th	ir	hv	th	ir/ed
<i>Output activations using t-groupF-1295.wts and groupF.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
in u (target: in v)	hv	I	ed	hv	I	ir/ed
io u (target: io v)	hv	yu	ed	hv	yu	ir/ed
mp u (target: jp v)		he	ed	hs	he	ir/ed
iq v (target: iq v)	hv	we	ir	hv	we	ir/ed
io u (target: io v)	hv	yu	ed	hv	yu	ir/ed
ir v (target: ir v)	hv	th	ir	hv	th	ir/ed
<i>Output activations using t-groupF-1343.wts / t-groupF-1391.wts and groupF.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
in u (target: in v)	hv	I	ed	hv	I	ir/ed
io u (target: io v)	hv	yu	ed	hv	yu	ir/ed
ip u (target: jp v)	hv	he	ed	hs	he	ir/ed
iq v (target: iq v)	hv	we	ir	hv	we	ir/ed
io u (target: io v)	hv	yu	ed	hv	yu	ir/ed
ir v (target: ir v)	hv	th	ir	hv	th	ir/ed
<i>Output activations using t-groupF-1439.wts and groupF.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
in u (target: in v)	hv	I	ed	hv	I	ir/ed
io u (target: io v)	hv	yu	ed	hv	yu	ir/ed
jp u (target: jp v)	hs	he	ed	hs	he	ir/ed
iq v (target: iq v)	hv	we	ir	hv	we	ir/ed
io u (target: io v)	hv	yu	ed	hv	yu	ir/ed
ir v (target: ir v)	hv	th	ir	hv	th	ir/ed

- **Group G (Past Perfect)**

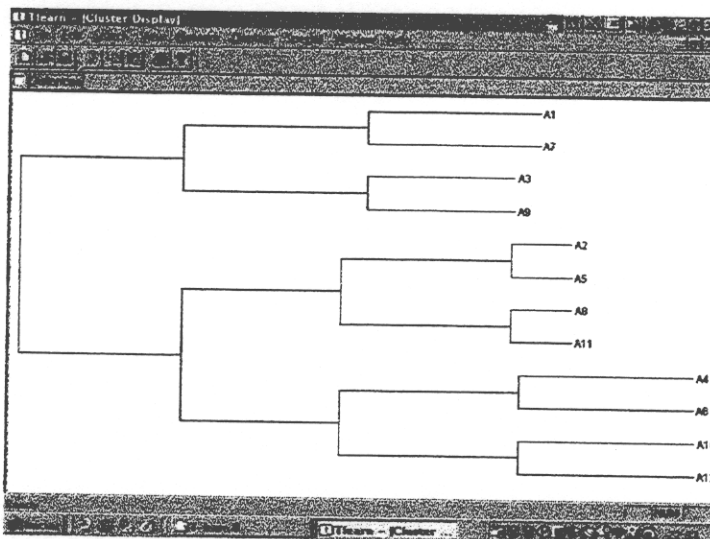
Códigos	Saídas			Saídas esperadas		
<i>Output activations using t-groupG-1494.wts and groupG.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
kn u (target: kn v)	hd	I	ed	hd	I	ir/ed
mo u (target: ko v)		yu	ed	hd	yu	ir/ed
kp u (target: kp v)	hd	he	ed	hd	he	ir/ed
mq v (target: kq v)		we	ir	hd	we	ir/ed
eo u (target: ko v)	wr	yu	ed	hd	yu	ir/ed
mr v (target: kr v)		th	ir	hd	th	ir/ed
<i>Output activations using t-groupG-1548.wts and groupG.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
kn v (target: kn v)	hd	I	ir	hd	I	ir/ed
mo u (target: ko v)		yu	ed	hd	yu	ir/ed
kp v (target: kp v)	hd	he	ir	hd	he	ir/ed
mq v (target: kq v)		we	ir	hd	we	ir/ed
mo u (target: ko v)		yu	ed	hd	yu	ir/ed
mr v (target: kr v)		th	ir	hd	th	ir/ed

<i>Output activations using t-groupG-1602.wts and groupG.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
k n v (target: k n v)	hd	I	ir	hd	I	ir/ed
m o u (target: k o v)		yu	ed	hd	yu	ir/ed
k p v (target: k p v)	hd	he	ir	hd	he	ir/ed
k q v (target: k q v)	hd	we	ir	hd	we	ir/ed
m o u (target: k o v)		yu	ed	hd	yu	ir/ed
k r v (target: k r v)	hd	th	ir	hd	th	ir/ed
<i>Output activations using t-groupG-1656.wts and groupG.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
k n v (target: k n v)	hd	I	ir	hd	I	ir/ed
m o u (target: k o v)		yu	ed	hd	yu	ir/ed
m p u (target: k p v)		he	ed	hd	he	ir/ed
m q v (target: k q v)		we	ir	hd	we	ir/ed
m o u (target: k o v)		yu	ed	hd	yu	ir/ed
m r v (target: k r v)		th	ir	hd	th	ir/ed
<i>Output activations using t-groupG-1710.wts and groupG.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
k n v (target: k n v)	hd	I	ir	hd	I	ir/ed
m o u (target: k o v)		yu	ed	hd	yu	ir/ed
k p u (target: k p v)	hd	he	ed	hd	he	ir/ed
k q v (target: k q v)	hd	we	ir	hd	we	ir/ed
m o u (target: k o v)		yu	ed	hd	yu	ir/ed
k r v (target: k r v)	hd	th	ir	hd	th	ir/ed
<i>Output activations using t-groupG-1764.wts and groupG.data (* Novel Data *) using Output Translation mapp.out (Euclidean Distance)</i>						
k n v (target: k n v)	hd	I	ir	hd	I	ir/ed
k o v (target: k o v)	hd	yu	ir	hd	yu	ir/ed
k p v (target: k p v)	hd	he	ir	hd	he	ir/ed
k q v (target: k q v)	hd	we	ir	hd	we	ir/ed
k o u (target: k o v)	hd	yu	ed	hd	yu	ir/ed
k r v (target: k r v)	hd	th	ir	hd	th	ir/ed

ANEXO G - A ANÁLISE DE AGRUPAMENTO

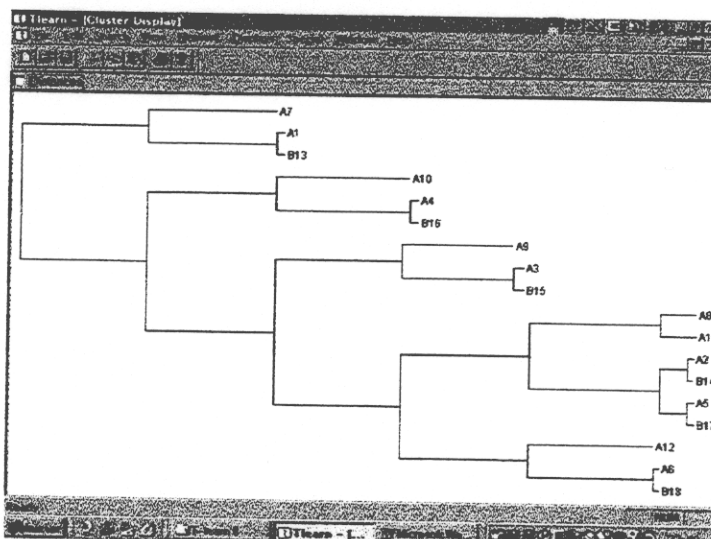
• Grupo A

A2 = ARE + YOU(2PS)
 A5 = ARE + YOU(2PP)
 A8 = DO + YOU(2PS) + TALK
 A11 = DO + YOU(2PP) + TALK



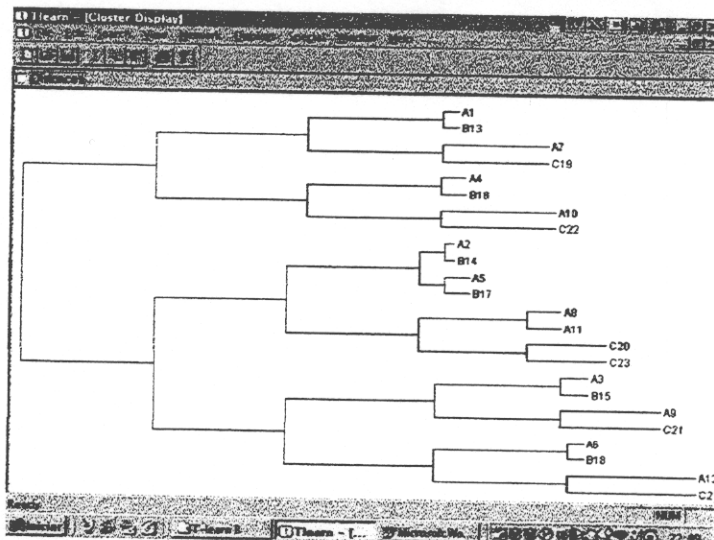
• Grupo B

A8 = DO + YOU(2PS) + TALK
 A11 = DO + YOU(2PP) + TALK
 A2 = ARE + YOU(2PS)
 B14 = ARE + YOU(2PS) + TALKING
 A5 = ARE + YOU(2PP)
 B17 = ARE + YOU(2PP) + TALKING



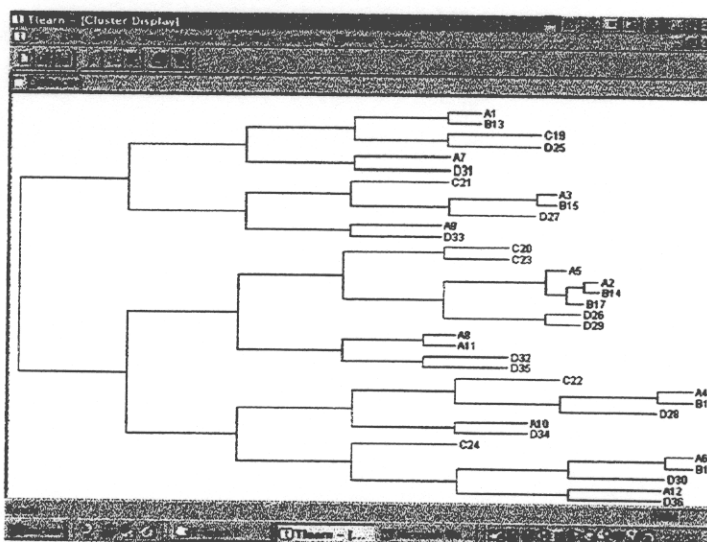
• Grupo C

A2 = ARE + YOU(2PS)
 B14 = ARE + YOU(2PS) + TALKING
 A5 = ARE + YOU(2PP)
 B17 = ARE + YOU(2PP) + TALKING
 A8 = DO + YOU(2PS) + TALK
 A11 = DO + YOU(2PP) + TALK
 C20 = CAN + YOU(2PS) + TALK
 C23 = CAN + YOU(2PP) + TALK



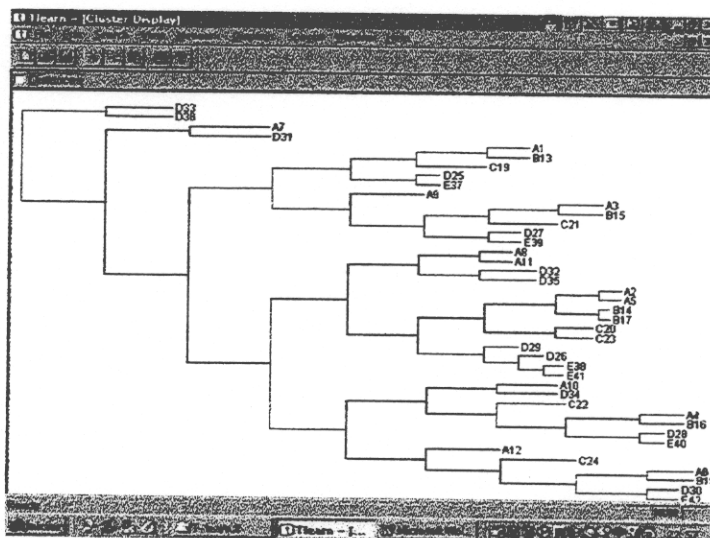
- **Grupo D**

C20 = CAN + YOU(2PS) + TALK
 C23 = CAN + YOU(2PP) + TALK
 A5 = ARE + YOU(2PP)
 A2 = ARE + YOU(2PS)
 B14 = ARE + YOU(2PS) + TALKING
 B17 = ARE + YOU(2PP) + TALKING
 D26 = WERE + YOU(2PS)
 D29 = WERE + YOU(2PP)
 A8 = DO + YOU(2PS) + TALK
 A11 = DO + YOU(2PP) + TALK
 D32 = DID + YOU(2PS) + TALK
 D35 = DID + YOU(2PP) + TALK



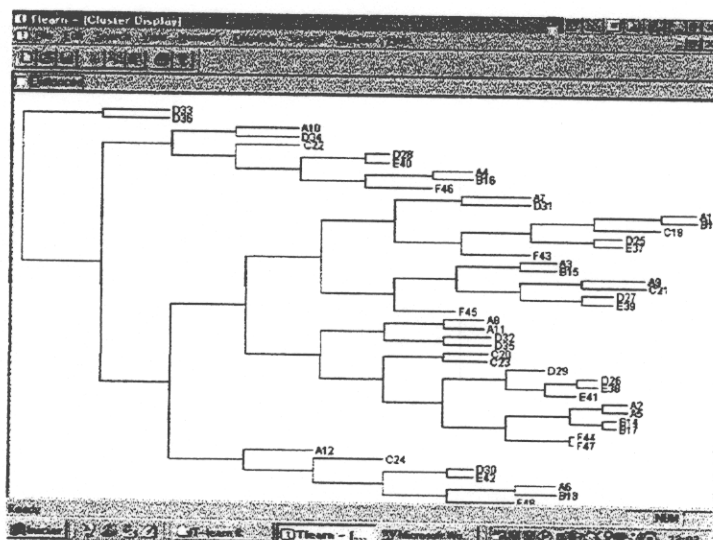
- **Grupo E**

A8 = DO + YOU(2PS) + TALK
 A11 = DO + YOU(2PP) + TALK
 D32 = DID + YOU(2PS) + TALK
 D35 = DID + YOU(2PP) + TALK
 A2 = ARE + YOU(2PS)
 A5 = ARE + YOU(2PP)
 B14 = ARE + YOU(2PS) + TALKING
 B17 = ARE + YOU(2PP) + TALKING
 C20 = CAN + YOU(2PS) + TALK
 C23 = CAN + YOU(2PP) + TALK
 D29 = WERE + YOU(2PP)
 D26 = WERE + YOU(2PS)
 E38 = WERE + YOU(2PS) + TALKING
 E41 = WERE + YOU(2PP) + TALKING



- **Grupo F**

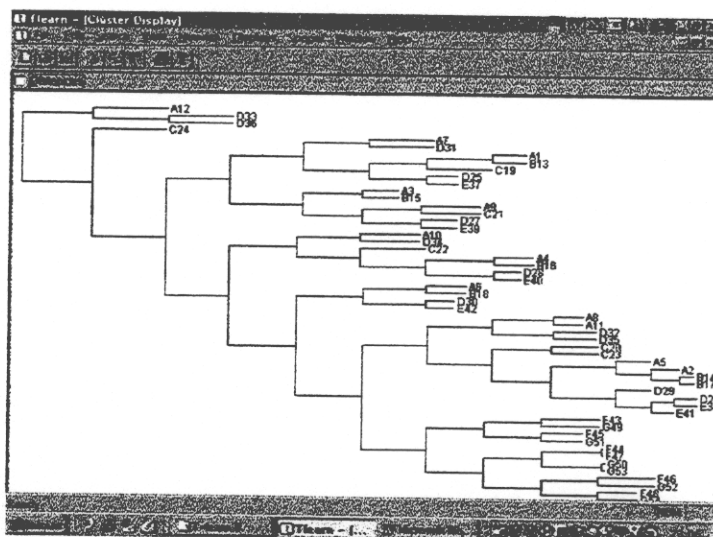
A8 = DO + YOU(2PS) + TALK
 A11 = DO + YOU(2PP) + TALK
 D32 = DID + YOU(2PS) + TALK
 D35 = DID + YOU(2PP) + TALK
 C20 = CAN + YOU(2PS) + TALK
 C23 = CAN + YOU(2PP) + TALK
 D29 = WERE + YOU(2PP)
 D26 = WERE + YOU(2PS)
 E38 = WERE + YOU(2PS) + TALKING
 E41 = WERE + YOU(2PP) + TALKING
 A2 = ARE + YOU(2PS)
 A5 = ARE + YOU(2PP)



B14 = ARE + YOU(2PS) + TALKING
 B17 = ARE + YOU(2PP) + TALKING
 F44 = HAVE + YOU(2PS) + TALKED*
 F47 = HAVE + YOU(2PP) + TALKED

- Grupo G

A8 = DO + YOU(2PS) + TALK
 A11 = DO + YOU(2PP) + TALK
 D32 = DID + YOU(2PS) + TALK
 D35 = DID + YOU(2PP) + TALK
 C20 = CAN + YOU(2PS) + TALK
 C23 = CAN + YOU(2PP) + TALK
 A5 = ARE + YOU(2PP)
 A2 = ARE + YOU(2PS)
 B14 = ARE + YOU(2PS) +
 TALKING
 B17 = ARE + YOU(2PP) +
 TALKING
 D29 = WERE + YOU(2PP)
 D26 = WERE + YOU(2PS)
 E38 = WERE + YOU(2PS) +
 TALKING
 E41 = WERE + YOU(2PP) +
 TALKING



(*) Este padrão também contempla os verbos irregulares.