

UMA INTERLÍNGUA CONEXIONISTA*

A CONNECTIONIST INTERLANGUAGE

FERNANDO G. FERREIRA JUNIOR¹

RESUMO: Neste artigo é feita uma reflexão sobre o fenômeno dos erros lingüísticos, sua possível natureza e sua importância para os processos de aprendizagem de línguas (ASL). Após discutir o conceito de ‘interlíngua’ (Selinker 1972;1996) aponto semelhanças e diferenças conceituais entre a idéia de uma interlíngua e a filosofia conexionista para a aquisição e processamento da linguagem, buscando compatibilizar o construto ‘interlíngua’ com redes conexionistas. Concluo especulando que o processamento da linguagem parece ser basicamente regido por um sistema probabilístico.

Palavras-chave: conexionismo; aquisição de segunda língua; processos cognitivos.

ABSTRACT: In this paper a reflection is made about the pervasive phenomenon of linguistic errors, their possible nature and the role they seem to play in the processes of second language acquisition (SLA) in particular. An analogy is drawn between the concept of an ‘interlanguage’, as originally proposed by Selinker (1972;1996), and connectionist models of language acquisition and processing. Language processing, it is suggested here, seems to be governed mainly by a probabilistic system.

Keywords: connectionism; second language acquisition; cognitive processes.

INTRODUÇÃO

Perhaps in this difficult business of trying to understand IL competence in all its core and peripheral complexity, we are on the verge of getting away from our previous narrowness.

Larry Selinker (1996:110)

* Este artigo, recentemente reformulado, foi apresentado como comunicação no 52º Seminário do GEL (Grupo de Estudos Lingüísticos do Estado de São Paulo) realizado no IEL/Unicamp, Campinas, em julho de 2004.

¹ Professor do CEFET-Ouro Preto e doutorando em Estudos Lingüísticos pela UFMG. Contato com o autor: jr_fernando@hotmail.com

Neste artigo farei, inicialmente, uma breve reflexão sobre o fenômeno do ‘erro’, mais especificamente os ‘erros’ lingüísticos, sua possível natureza e sua importância para os processos de aprendizagem de línguas (ASL)². Em seguida discuto o conhecido conceito de interlíngua, desenvolvido por Larry Selinker em 1972, conceito este que busca explicar o processo de ASL como sendo algo sistemático e gradual, revelado pela natureza também sistemática dos erros lingüísticos em L2 (refletido por sua instância maior que é a *fossilização* lingüística). Possíveis semelhanças e diferenças conceituais entre a noção de interlíngua e a filosofia conexionista para a aquisição e processamento da linguagem são apontadas, numa tentativa de compatibilizar o construto interlíngua (i.e., este sistema específico) com redes conexionistas. Retomo, ao final do texto, a discussão acerca do papel dos “erros” lingüísticos da perspectiva conexionista, na qual o erro passa a ser visto como um produto resultante, e desta forma um bom indicativo, de processos contínuos de aprendizagem/desenvolvimento. Finalmente, concluo especulando que o processamento da linguagem talvez seja inteiramente regido por um sistema probabilístico.

1. O ‘ERRO’ NA AQUISIÇÃO DE LÍNGUAS

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

Todas as definições dadas acima interessam à discussão ora feita, pois, obviamente, definem a palavra ‘erro’. No entanto, as definições fornecidas em 5 e 6 tornam-se mais relevantes diante da compreensão de ‘erro’ da perspectiva conexionista discutida neste trabalho. Estas definições expressam, de certa forma, uma noção de erro na qual quem ‘erra’ (uma rede, um sistema ou um aprendiz) o faz numa tentativa lógica, portanto correta, de satisfazer restrições múltiplas com base no estabelecimento de padrões probabilísticos/estatísticos (estocásticos) de ocorrência/frequência de uma dada experiência. Esta discussão será retomada mais adiante.

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

² Ao longo deste trabalho as expressões ‘aquisição de segunda língua’ (ASL) e ‘aquisição de língua estrangeira’ (ALE), e respectivamente as palavras L2 e LE, são tratadas como sinônimas.

De fato, por um longo período a ‘análise de erros’ e a ‘análise contrastiva’ influenciaram teorizações sobre o fenômeno dos erros em ASL. A possibilidade de prever, no processo de aquisição, os casos em que ocorreriam fenômenos como transferência negativa ou ‘interferência’ de estruturas da L1, provenientes do contraste entre a L1 e a L2 estudada, era algo que tornava a hipótese contrastiva atraente para a aprendizagem (e ensino) bem-sucedidos de uma L2, como já destacavam Weinreich (1953) e Pit Corder (1969; 1973). No entanto, apontam Widdowson (1996) e Ellis (1997), a explicação do fenômeno do erro como sendo uma simples interferência de estruturas da L1, ou um mau hábito que deveria ser evitado, não era satisfatória do ponto de vista teórico. Além disto, observava-se que os erros em ASL eram sistemáticos e recorrentes, e muitos deles comuns a todos os aprendizes da L2, independentemente de sua L1.

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

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

Rumelhart & McClelland desenvolveram um modelo conexionista (um programa de computador) para o processamento de verbos da língua inglesa que reproduz o fenômeno da curva em U observado na aprendizagem do passado dos verbos em inglês. O processo de aprendizagem revela, basicamente, três estágios: 1) inicialmente o uso, pela criança, de um número limitado de verbos de alta frequência, geralmente verbos irregulares, e o uso correto do passado de cada verbo (i.e., as crianças parecem usar os verbos separadamente, tratando-os item por item; 2) em um segundo estágio, à medida que a criança entra em contato com mais e mais insumo, ela começa a generalizar, tratando todos os verbos como regulares. Neste estágio há um processo de supergeneralização de regras/padrões abstratos subjacentes às formas verbais, ou seja, busca de aprendizagem sistemática. Nesta fase ocorrem muitos erros, resultantes de uma busca constante por sistematização através da generalização de padrões diferentes. 3) o último estágio revela a coexistência de formas verbais regulares e irregulares, conseqüência de um processo de categorização a partir da sistematização de padrões diferentes. O quadro 1 abaixo, adaptado de Ellis (1997:23), ilustra bem o fenômeno relatado no estudo aqui descrito:

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

Quadro 1

Estágios na aquisição do tempo passado do verbo 'eat'

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

O estudo acima relatado também revela que, apesar do fenômeno ser observado em estágios aparentemente estanques, o pleno desenvolvimento ou a plena aquisição das estruturas da linguagem (revelado pela completitude da curva em U) é altamente irregular e não-linear. Esta inevitável irregularidade e não-linearidade é justamente o que produz comportamentos erroneamente taxados de erros – quando analisados de uma perspectiva puramente comportamental ou de 'desempenho' lingüístico, como podemos observar nos exemplos do quadro 1 acima. Da perspectiva do sistema lingüístico do aprendiz de uma língua os erros tornam-se acertos, pois, como destacado no início deste trabalho, revelam uma busca lógica por sistematicidade ou regularização.

Finalmente, é importante observar a diferença crucial entre a aprendizagem de itens em separado e a aprendizagem sistemática, ou a aprendizagem das regras/padrões abstratos subjacentes ao uso dos itens lingüísticos (sejam eles sufixos de passado, palavras, estruturas, etc.). E isto é evidenciado por estudos conduzidos com bebês no qual são usadas gramáticas artificiais, como o trabalho de Gómez & Gerken (1999; 2000)³.

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

³ Por falta de espaço não serão aqui detalhados os procedimentos metodológicos de estudos psicolingüísticos experimentais envolvendo bebês, como os adotados no experimento ora relatado.

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

Observe que há uma semelhança *superficial* entre o padrão ABA com a ordem de categorias *substantivo-verbo-substantivo*. Ou seja, as relações envolvendo categorias ocorrem em um nível superior de abstração (as quais não podem basear-se simplesmente na percepção (física) de identidade)⁴.

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

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

2. O CONCEITO DE ‘INTERLÍNGUA’

Um conceito que, de certa maneira, representa um grande avanço nos estudos em ASL no tocante à busca pela compreensão da “natureza do erro” ou do processo de “fossilização” é o conceito de interlíngua (IL), proposto por Larry Selinker em 1972. Em um artigo seminal publicado naquele ano, Selinker aponta (como revelado pelos estudos de

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

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

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

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

É importante destacar que a estrutura latente descrita neste artigo não possui nenhum programa genético, se comparada com a proposta de Lenneberg: não há um equivalente direto com nenhum conceito do tipo ‘gramática universal’, não há nenhuma garantia de que esta estrutura latente será ativada; não há nenhuma garantia de que tal estrutura venha a se tornar uma estrutura de fato de qualquer língua natural (i.e., não há nenhuma garantia que a aprendizagem será bem sucedida), e é bem possível que exista uma superposição entre tal estrutura latente para a aquisição da linguagem com outras estruturas do nosso intelecto.⁵ (1972:212).

No entanto, para explicar os cerca de 5% dos aprendizes que desenvolvem uma competência lingüística e atingem um desempenho idênticos ao de um falante nativo da L2 estudada, Selinker lança mão de uma suposta ativação da Gramática Universal (UG)

⁵ Minha tradução de “It is important to state that with the latent structure described in this paper as compared to Lenneberg’s, there is no genetic timetable; there is no direct counterpart to any grammatical concept such as ‘universal grammar’; there is no guarantee that this latent structure will be activated at all; there is no guarantee that the latent structure will be ‘realized’ into the actual structure of any natural language (i.e., there is no guarantee that attempted learning will prove successful), and there is every possibility that an overlapping exists between this latent language acquisition structure and other intellectual structures”.

chomskiana (que conteria todo o conteúdo proposicional) através de um mecanismo de aquisição de línguas, um LAD (*language acquisition device*), presente no conceito análogo de *estrutura lingüística latente* proposto por Lenneberg e adotado por Selinker em seu artigo (1972: 211-12).⁶

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

Veremos, a seguir, que a idéia de uma *interlíngua* (principalmente como sugerida nas reformulações recentes de Selinker) é altamente compatível com modelos conexionistas para o processamento de sistemas estruturados. A filosofia norteadora dos modelos conexionistas parece dar conta, ou pelo menos vislumbra a possibilidade de fornecer boas explicações, do processo de aquisição da linguagem, como discutirei no item 4 a seguir. O conceito de interlíngua, acima de tudo, redireciona a teorização em ASL para aspectos importantes do processamento cognitivo dos aprendizes.

3. A INTERLÍNGUA VISTA COMO UMA REDE CONEXIONISTA ALTAMENTE ESTRUTURADA

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

⁶ Veja White (2003) para uma discussão acerca do conceito interlíngua de uma perspectiva inatista (i.e., chomskiana) no âmbito dos estudos em ASL.

Numa tentativa de interpretar o texto de Selinker (por sinal bastante hermético, como bem destacado por Ellis⁷) à luz da filosofia conexionista, farei uma comparação do que o autor chama de *estrutura psicológica latente* com a arquitetura e o funcionamento de modelos conexionistas. Desta forma, tento sugerir que modelos conexionistas explicam, de maneira bastante convincente, e partindo de uma base neurofisiológica, o processo de ASL como algo parecido com o construto selinkiano de uma “interlíngua”, caracterizado pelo que chamei acima de ‘unidades psicológicas emergentes’. Como sugerido posteriormente pelo próprio Selinker (1996), também acredito haver fortes semelhanças entre ambas as teorizações. Com este intuito, será dada, a seguir, uma breve explicação sobre a natureza e o funcionamento de uma rede conexionista. A discussão baseia-se, principalmente, em Elman (1990), Elman et al. (1998) e Ferreira Junior (2005).

Como descrevem Elman et al. (1998) e Elman (1990), uma rede conexionista⁸ é formada, basicamente, por unidades interligadas umas às outras através de conexões, um enfoque, portanto, fortemente apoiado na neurofisiologia humana. O processamento dessa rede é totalmente distribuído e em paralelo (daí o termo sinônimo de rede PDP, ou seja, processamento distribuído em paralelo). Modelos conexionistas, com uma arquitetura interna como a descrita acima, são construídos na forma de programas de computador. Uma vez em funcionamento, o sistema estabelece processos de regulação entre as unidades que compõem a rede (na forma de ajuste das forças/pesos de conexão) provocando, assim, a estruturação do sistema. O ‘conhecimento’ da rede está justamente nas forças/pesos definidos localmente entre as unidades (i.e, através da constante ativação ou inibição destas). Assim, surge um desempenho baseado em padrões de ativação entre unidades da rede, padrões estes que estabelecem configurações específicas a cada ativação.

Há, essencialmente, três tipos de unidades: as unidades de insumo, as unidades ocultas e as unidades de saída. É na camada oculta que as representações são criadas e o conhecimento é gerado. A camada oculta também faz com que a rede *aprenda* através da *retropropagação* dos erros. A rede aprende baseando-se no *cálculo* entre a diferença de ativação (o seu peso) na unidade de saída com o peso ideal daquela unidade. A rede estabelece associações através da regulação entre o seu peso/forma atual e o peso/forma ideal (*target*), este último informado, inicialmente, pela experiência com o meio (ou, nesta nossa discussão, a L2).

Há uma questão importante nesta discussão. Como destacam Elman (1990) e Elman et al (1998), as associações de padrões idênticos nos primeiros modelos conexionistas (tipo o Perceptron) só podiam ocorrer em um campo espacial linear, ou bidimensional. Isto fazia com que a rede ignorasse o fato de, às vezes, formas diferentes poderem revelar padrões idênticos (e vice-versa, como discutido anteriormente na nota 4 deste artigo). Modelos do

⁷ Ellis (1997:127) destaca, na bibliografia comentada ao final de seu livro, que o artigo de Selinker “is not easy to read, but it gave SLA the term ‘interlanguage’ and it contains a rich seam of theoretical ideas that is still being mined today”.

⁸ O conceito sinônimo de “redes neurais”, implicando em uma possível plausibilidade biológica devido à metáfora do cérebro em funcionamento, através de neurônios e sinapses, também é usado nas teorizações de cunho conexionista sobre a estrutura e o funcionamento do sistema cognitivo humano.

tipo Perceptron tinham apenas duas camadas (níveis) de unidades, ou seja, unidades de entrada e unidades de saída, e isto tornava o comportamento da rede algo parecido com o estímulo/resposta do behaviourismo. Em um modelo Perceptron as associações são puramente mecânicas. É necessário, portanto, um nível de processamento que gere representações abstratas.

Para não cair em um mero ‘associacionismo computacional’, Elman et al (1998) propõem uma “rede recorrente simples”, ou SRN (*simple recurrent network*), que possui, além de uma camada de unidades intermediárias (ocultas), responsável pelo aspecto “abstração” ou “representações abstratas” da rede, uma camada ou unidades de *contexto*, adicionando, assim, o aspecto tempo ou uma memória à rede (devido à recursividade). Assim, nas redes recorrentes as unidades são informadas não apenas pelo insumo externo a elas, mas também por ativações prévias da própria unidade que ficam “registradas” nas unidades de contexto, aumentando o poder de autoregulação. E isto é uma importante contribuição conexionista para a discussão sobre autonomia. A rede produz, assim, insumo que, por sua vez, regulará a própria rede. Também, inerente a uma rede recorrente é a sua capacidade de “inferir” (este aspecto é consoante com a aprendizagem em vários domínios e reflete a capacidade de “fazer previsões”, certamente consequência direta do estabelecimento de protótipos ou configurações padrão). A título de ilustração vemos, na Figura 1 a seguir, um exemplo de uma rede recorrente simples com camada de contexto:

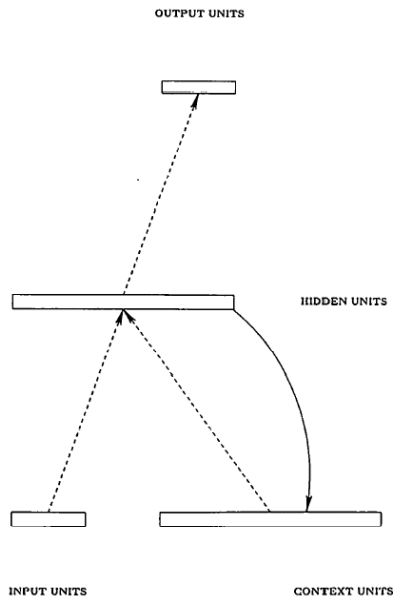


Figura 1

Uma rede recorrente simples (SRN) com unidade de contexto. Os retângulos representam as camadas de unidades.

In: Jeffrey Elman. Finding Structure in Time. *Cognitive Science*, 14, 1990:184.

Processos de associação de padrões numa rede conexcionista tipo SRN tornam-se mais dinâmicos, pois o que é “similar” não mais será definido em um ambiente bidimensional (baseado puramente na forma) mas multidimensional (sendo consideradas também relações espaciais). Desta maneira, aumenta-se a capacidade de representação e, conseqüentemente, de associação. Podemos dizer que tais redes também exibem um comportamento inteligente, pois possuem níveis diferentes de abstração/representação do conhecimento gerado.

Os pesos entre as unidades (e a conseqüente estabilização da rede ou a aprendizagem propriamente dita) em um modelo conexcionista são autoreguláveis. Assim a rede vai, aos poucos, autoregulando-se e refinando positivamente o limiar de ativação (em outras palavras, *refinar* significa corrigir a diferença (o erro) entre produção atual e produção ideal). Visto desta maneira, o erro (e a sua constante minimização) torna-se *fundamental* para a rede atingir a produção ideal; o fenômeno da curva em U nos processos de aprendizagem reflete justamente isso, como podemos observar no trabalho sobre a aprendizagem do passado dos verbos irregulares de Rumelhart & McClelland, cujos ‘erros’ típicos vimos na tabela 1 mais acima. A rede atinge a generalização de padrões a partir da exposição gradual ao padrão ideal desejado. A cada produção da rede é feita uma associação com o padrão ideal. O cálculo da diferença entre os pesos refaz o caminho de forma inversa. Esta *retropropagação* faz com que o *cálculo* retorne para as unidades ocultas, de forma que haja uma regulação e um reajuste visando às futuras ativações (este é um processo de constante minimização do erro por parte da rede). Desta forma, a rede aprende a aprender, através da diminuição da sensibilidade ao erro (i.e., torna-se um sistema mais tolerante e menos maleável). Como destacam Bod, Hay & Jannedy (2003:08), a aquisição da linguagem envolve “uma atualização contínua de distribuições probabilísticas”. Esta também parece ser a essência do processo de especialização de domínios (ou modularização).

O funcionamento da rede é sempre influenciado pelos mais variados tipos de restrições encontrados; no nosso caso, a L1 do aprendiz, o insumo da L2 ao qual está exposta, bem como as próprias restrições criadas pela rede ao longo de seu desenvolvimento, ou seja, a sua língua interna ou ‘interlíngua’. É a partir destas restrições que a rede emerge ou, em outros termos, o sistema lingüístico é construído. É a partir de processos de restrições de restrições que se dá a regulação do sistema, levando em última instância à autoregulação (autonomia) e à consolidação da aprendizagem da língua.

Para concluir esta breve explicação de uma rede conexcionista, é importante destacar que em uma rede do tipo PDP o importante são as representações locais, i.e., as conexões entre as unidades. É a partir da força inicialmente estabelecida localmente que se dá um processo de equilíbrio da rede. E é a partir do equilíbrio da rede que surgem as representações distribuídas. Ou seja, representações locais geram representações emergentes (ou ativas) e distribuídas, determinadas sempre, por sua vez, por aquelas mesmas configurações localmente e constantemente (re)definidas através de pesos (forças) entre as conexões.

Que comparações podemos, então, estabelecer entre a proposta de Selinker e os modelos conexionistas? Em ambos os casos há, inicialmente, um “alvo” externo, que é o insumo em L2 (ou ainda a experiência como um todo)⁹. Com o passar do tempo, e devido a constante complexificação da rede, a informação da L2 passa a determinar ajustes locais cada vez mais finos entre unidades específicas. Este processo de “calibração”, que busca minimizar o ‘erro’, é feito com base em processos probabilístico-estatísticos. A externalização deste processo de calibração se dá na forma de ‘erros’ no desempenho do sistema (ou na produção lingüística do aprendiz), como vimos na tabela 1 acima.

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

Portanto, nenhuma teorização inatista sobre um determinado período crítico de aprendizagem parece ser necessária. Tais “períodos” são uma consequência do próprio desenvolvimento ou estruturação da rede; ou seja, uma vez que certa configuração, ou protótipos (fonéticos, sintáticos, etc..), se estabelece, o sistema praticamente se fecha (como destacado no parágrafo anterior). Em termos neurofisiológicos tal fenômeno equivaleria à diminuição observada na neuroplasticidade humana (cf.: Elman et. al., 1998:389).¹⁰

4. CONSIDERAÇÕES FINAIS: A LINGUAGEM COMO UM SISTEMA PROBABILÍSTICO

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

⁹ Embora passe aqui a idéia de língua como um ‘alvo’, destaco, em tempo, ser este não um alvo estático, mas um em constante movimento e altamente dinâmico!

¹⁰ Entretanto, no tocante a aprendizagem de línguas, uma “desfossilização” é algo possível por meio de instrução com foco na forma e um “estímulo exagerado”. Como observa Ellis (2004:63), um modelo conexionista desenvolvido por McClelland et al (2002, apud Ellis) mostrou-se eficaz no clássico fenômeno da discriminação dos fonemas /l/ e /r/ na aprendizagem do inglês por parte de aprendizes japoneses.

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

Como apontam inúmeros estudos que utilizam modelos conexionistas, a linguagem parece mesmo possuir propriedades probabilísticas. Os modelos conexionistas exibem um desempenho baseado em gradação (contínuo) no qual o limiar de força/peso de ativação estabelece, gradualmente, categorizações; ou seja, os modelos “são capazes de fazer distinções sutis e mesmo categorizar ao longo de dimensões que parecem ser, por natureza, contínuas” (Elman et. al.,1998:53).

Alguns estudos que podemos destacar, e que apontam para processos de base estatístico-probabilística, são o estudo sobre a percepção visual de palavras, de McClelland e Rumelhart (1981), que mostra o papel das probabilidades transicionais na percepção de palavras (*word-boundary perception*); o estudo de categorização de Elman (1990) que aponta gradientes na categorização de verbos e substantivos; o estudo de Gómez & Gerken (1999;2000) que aponta o papel das probabilidades transicionais na identificação de padrões em seqüências gramaticais, revelando, assim, que é possível “aprender” padrões sintáticos abstratos; bem como o já citado estudo sobre os verbos irregulares, desenvolvido por Rumelhart & McClelland (1986), que mostra a existência de regularidades e subregularidades dentro da categoria tradicionalmente chamada de “verbos irregulares”.

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

REFERÊNCIAS BIBLIOGRÁFICAS

- BOD, R., HAY, J. & JANNEDY, S. (2003). Introduction. In: Rens Bod, Jennifer Hay & Stefanie Jannedy. *Probabilistic Linguistics*. Cambridge, MA: Bradford/MIT Press. pp. 01-10.
- CORDER, P. (1967). The significance of learners' errors. *International Review of Applied Linguistics*, 5. pp. 161-169.
- _____. (1973). Linguistics and the language teaching syllabus. In: J.P.B Allen and P. Corder (eds). *The Edinburg Course in Applied Linguistics*. London: Oxford University Press. pp. 275-284.
- (2001). *Dicionário Houaiss da Língua Portuguesa*. Rio de Janeiro: Editora Objetiva.

- ELMAN, J. L. (1990). Finding structure in time. *Cognitive Science*, 14. pp. 179-211.
- ELMAN, J. et al. (1998). *Rethinking Innateness: a connectionist perspective on development*. Massachusetts: MIT Press.
- ELLIS, N. C. (2004). *The processes of second language acquisition*. In: Van Patten et. al. (Eds.). *Form-Meaning Connections in Second Language Acquisition*. New Jersey: LEA. pp. 49-76.
- ELLIS, R. (1997). *Second Language Acquisition*. Oxford: Oxford University Press.
- FERREIRA JUNIOR, F. G. (2005). Dos grupos de discussão às redes neurais: Reflexões sobre o desenvolvimento de um léxico mental. *Revista Brasileira de Linguística Aplicada*, Vol.5, n.2, pp. 231-252.
- GÓMEZ, R. & GERKEN, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, 70. pp. 109-135.
- _____. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Science*, Vol.4, n.5, pp. 178-186.
- MCCLELLAND, J. L. & RUMELHART, D. (1981). E. An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, 88. pp. 287-330.
- RUMELHART, D. & MCCLELLAND, J. (1986). On learning the past tenses of English verbs. In: D.E. Rumelhart & J. L. McClelland (orgs) In: *Parallel Distributed Processing (Explorations in the Microstructure of Cognition)*, Vol.1. MIT Press. pp. 216-271.
- SEIDENBERG, M. (1997). Language Acquisition and Use: Learning and Applying Probabilistic Constraints. *Science*, 275. pp. 1599-1603.
- SELINKER, L. (1972). Interlanguage. *IRAL*, 10. pp. 209-31.
- _____. (1969). On the notion of 'IL competence' in early SLA research: an aid to understanding some baffling current issues. In: Gillian Brown et. al. (orgs). *Performance & Competence in Second Language Acquisition*. Cambridge: Cambridge University Press. 1996. pp. 92-113.
- WHITE, L. (2003). *On the Nature of Interlanguage Representation: Universal Grammar in the Second Language*. In: Doughty, C & Long., M. (orgs). *The Handbook of Second Language Acquisition*. Oxford, UK: Blackwell. pp. 19-42
- WEINREICH, U. (1953). *Languages in contact*. The Hague: Mouton.
- WIDDOWSON, H. G. (1996). *Linguistics*. Oxford: Oxford University Press.

Recebido: 06/09/2005

Aceito: 29/08/2006