

# Como aprender as construções passivas sem regras

Rosângela Gabriel\*

O objetivo principal da pesquisa desenvolvida é contribuir para a compreensão da natureza da linguagem e do cérebro através do estudo de um único fenômeno lingüístico. Assim como muitos outros processos cognitivos, o fenômeno de aquisição e uso de construções sintáticas envolve a consideração simultânea de um grande número de informações de origens diversas. A fim de investigar *como as crianças aprendem as construções passivas* e, mais especificamente, por que as crianças aprendem as construções passivas, ou ainda, que fatores contribuem para a aquisição das construções passivas, foram empregadas duas técnicas de investigação: análise de dados empíricos e simulação dos dados empíricos em redes neuronais conexionistas.

## 1 Dados empíricos de falantes monolíngües de português e inglês

Os dados empíricos foram colhidos com crianças e adultos falantes monolíngües de inglês britânico e português brasileiro. Os sujeitos do estudo Produção em Inglês foram 46 crianças, divididas em 4 grupos (3-4 anos de idade, 5-6 anos, 7-8 e 9-10 anos), e 10 adultos. Já os sujeitos do estudo Produção em Português foram 80 crianças e 20 adultos. O *design* do experimento envolveu 12 cenas nas quais 12 vezes um agente foi topicalizado e 12 vezes um não-agente foi topicalizado. As cenas foram mostradas aos sujeitos através de um desenho animado em um monitor de vídeo. Após cada cena, o investigador pediu ao participante "Me fala sobre o (*agente*)" e "Me fala sobre o (*não-agente*)". As respostas dos sujeitos foram então classificadas como (1) ativas, (2) passivas, (3) não-passivas e (4) não-evento, como ilustrado pelos exemplos:

\* PUCRS - Pontifícia Universidade Católica do Rio Grande do Sul.

- (1) O gato assustou o rato.                      (3) O rato ganhou um susto do gato  
(2) O rato foi assustado pelo gato.        (4) O rato tá com medo.

Os resultados demonstraram que as respostas dos sujeitos são fortemente influenciadas pelo tópico dado pelo interlocutor. A manutenção do tópico discursivo no início da frase parece ser o critério usado pelos sujeitos na escolha dentre as estruturas sintáticas disponíveis. Essa tendência dos sujeitos pela manutenção do tópico discursivo já havia sido documentada por estudos anteriores, entre eles Clark e Begun (1968), Bates e Devescovi (1989), Marchman et al. (1991), Ferreira (1994) e Bates et al. (1995). Os dados indicam ainda que, apesar de as crianças serem sensíveis ao tópico discursivo, elas nem sempre têm disponível ou facilmente acessível<sup>1</sup> uma construção que permita a manutenção do tópico discursivo no início da frase. Logo, gostaríamos de sugerir que as crianças aprendem as construções passivas a fim de terem disponível uma construção que permite que um não-agente ocupe a posição de tópico frasal.

Mas como as crianças descobrem que devem manter o tópico discursivo? Talvez as crianças nasçam equipadas de algum tipo de regra ou princípio que indique que o tópico discursivo deve ser mantido. Ou talvez as crianças identifiquem um determinado padrão no *input* linguístico a que estão expostas e passem a apresentar um comportamento condizente com esse padrão. A fim de testar essas hipóteses, usamos uma técnica de investigação indireta, baseada na simulação de aprendizagem em redes neuronais conexionistas. Será possível desenvolver um modelo conexionista que aprenda a preservar o tópico discursivo, sem lançar mão de nenhuma regra explícita que o instrua a assim proceder?

## 2 Simulando um *input* equilibrado

A tarefa desta rede neuronal conexionista é processar uma cena e descrevê-la em uma estrutura ativa ou passiva, de acordo com o personagem topicalizado. O objetivo é desenvolver um modelo que seja capaz de aprender a preservar o tópico discursivo e usar uma estrutura frasal que permita sua manutenção na primeira posição da sentença. Mais especificamente, estamos interessados na troca entre estruturas ativas e passivas dependendo dos papéis semânticos (agente ou não-agente) anteriormente topicalizados no discurso. Obviamente, a troca entre duas estruturas frasais pressu-

põe sua disponibilidade para o falante, isto é, elas precisam ter sido previamente adquiridas e precisam ser facilmente acessáveis. Vários problemas tornam esse processo difícil para o jovem falante. Em primeiro lugar, nas duas línguas investigadas, isto é, Inglês e Português, a estrutura ativa é mais freqüente do que a passiva no *input* linguístico, causando a consolidação de uma ordem semântica *default*:<sup>2</sup> agente  $\Rightarrow$  ação  $\Rightarrow$  paciente. O uso da estrutura passiva requer a habilidade de inverter essa seqüência de palavras e de reorganizar os papéis semânticos: paciente  $\Rightarrow$  ação  $\Rightarrow$  agente. Portanto, a fim de entender e produzir frases passivas, a criança tem que abrir mão de uma primeira (e falsa) generalização baseada na amostra de frases ativas, que poderia ser assim verbalizada "o primeiro elemento em uma frase é aquele que faz algo".

Em segundo lugar, a passiva em ambas as línguas, Português e Inglês, é uma estrutura sintática complexa envolvendo o uso do particípio passado e eventualmente a preposição "por" (em Inglês *by*). Lidar com esses elementos não é uma tarefa corriqueira, uma vez que o particípio em ambas as línguas apresenta formas regulares e irregulares, cuja aquisição representa um desafio para o jovem aprendiz, como demonstrado em vários estudos (entre outros, Rumelhart e McClelland, 1986; Plunkett e Marchman, 1991; Pinker, 1999). A ambigüidade da preposição "por" é também um desafio, porque o jovem falante é exposto a frases em que "por" introduz um agente, como em "es flores foram plantadas pela vovó", assim como frases em que "por" introduz um adjunto adverbial, como em "As flores foram plantadas pela manhã" (Liversedge et al., 1998). Finalmente, o jovem falante tem que ser sensível à exigência discursiva de manutenção do tópico. Crianças de apenas 3 anos de idade parecem ter desenvolvido essa sensibilidade, como observado em nossos estudos de produção linguística (ver também Bates e Devescovi, 1989; Marchman et al., 1991). De fato, mesmo crianças muito jovens tentam manter o tópico discursivo na primeira posição sentencial através do uso de uma estrutura frasal apropriada.

Nosso modelo conexionista de manutenção do tópico não dá conta de toda a complexidade envolvida na aquisição de frases ativas e passivas. Vários aspectos da aquisição e processamento dessas construções não serão abordados aqui. Em nossa simulação, nenhuma tentativa é feita para representar mudanças de tempos verbais e concordância entre sujeito-verbo. Todos os verbos foram

<sup>1</sup> Apesar de a palavra "acessável" não constar nos dicionários de Língua Portuguesa, optamos pelo seu uso, uma vez que a palavra acessível carrega o sentido de "algo de fácil acesso", ao passo que o sentido que buscamos é "algo que pode ser acessado".

<sup>2</sup> Optamos pela manutenção da palavra em Inglês *default* por considerarmos que palavras como "canônica" ou "não-marcada" alterariam ligeiramente o significado de nossa afirmação.

representados na forma em Inglês do passado simples/particípio (por exemplo, *licked*<sup>3</sup>), sendo que verbos que não apresentam a mesma forma no passado simples e no particípio não foram usados. Uma vez que o Inglês é a língua usada nessa simulação, nenhuma alteração na concordância sujeito-verbo é necessária, já que os verbos estão conjugados no pretérito simples.<sup>4</sup> A fim de manter a rede tão simples quanto possível, o verbo auxiliar (ou seja *be* ou *get* – “ser” em Português) não foi representado na simulação, nem o foram os artigos (isto é *the* – o(s)/a(s) em Português). A simulação imita os dados dos falantes adultos do estudo Produção em Inglês. Quando solicitados a falar sobre o agente em uma cena, 100% das respostas dos falantes adultos de Inglês Britânico estavam na estrutura ativa. Por outro lado, quando solicitados a falar sobre o não-agente, os mesmos sujeitos forneceram aproximadamente 100% de respostas na estrutura passiva. Reproduzindo esse padrão, o número de frases ativas e passivas nessa simulação é equilibrado, e a estrutura frasal altamente associada com o personagem topicalizado. A tarefa da rede é alternar a estrutura frasal de acordo com qual personagem da cena é topicalizado.

## 2.1 Representação do input e output

Um léxico de 5 nomes, 5 verbos e a preposição *by* foi usado no modelo.<sup>5</sup> Assim como com sujeitos humanos, o conhecimento do significado das palavras não é questionado e se assume que a rede (e os seres humanos) conhece o nome dos personagens e os verbos que descrevem as ações. O *input* consiste em uma cena constituída de três elementos (dois personagens e uma ação) e um tópico (o agente ou o não-agente). Uma representação localista simples é usada aqui, com apenas uma única unidade de *input* ativada em uma dada palavra. Cada palavra é representada por um vetor de 15-bit em que apenas um *bit* está ligado (1) e os 14 restantes estão desligados (0). As palavras estão concatenadas formando um vetor de 60-bit, representando um par cena/tópico no *input*. A rede é treinada em 80 cenas, cada uma apresentada com um agente e com um não-agente topicalizado, resultando em 160 pares cena/tópico.

<sup>3</sup> Em Inglês, a maioria dos verbos no pretérito simples (Simple Past Tense) e no particípio (Past Participle) são homófonos, por exemplo *kiss* ⇒ *kissed* ⇒ *kissed*, mas não *break* ⇒ *broke* ⇒ *broken*.

<sup>4</sup> Em Inglês, os verbos conjugados no Simple Past Tense não sofrem alteração decorrente da mudança de pessoa do discurso, como é o caso em Português.

<sup>5</sup> Um conjunto suplementar de 5 nomes foi dado a rede para uma possível tarefa de generalização.

O *output* desejado consiste de frases ativas e passivas simples, cada uma com três palavras lexicais (dois nomes e um verbo) e a palavra gramatical *by* no caso das frases passivas. Cada palavra lexical é representada por um vetor de 15-bit, enquanto a preposição é representada por um vetor de 2-bit. Portanto a corrente de *output* representando a frase toda é formada por um vetor de 47-bit. A fim de manter o comprimento da corrente de *output* uniforme nas frases ativas e passivas, uma palavra ‘vazia’ representada por duas unidades inativas é usada nas frases ativas. A rede é treinada em uma amostra de 80 frases ativas e 80 passivas. A tarefa em cada etapa é descrever uma dada cena em uma forma ativa ou passiva, dependendo de qual papel temático (agente ou não-agente) é topicalizado.

## 2.2 Arquitetura da rede

Uma rede *feedforward* de três camadas completamente conectadas, com 60 unidades *input*, 30 unidades *hidden* e 47 unidades *output*, foi construída para desempenhar a tarefa. Não existem conexões entre os núdulos dentro de uma mesma camada nem conexões de *feedback* para os núdulos de camadas anteriores. É usado um algoritmo de aprendizagem conhecido como *backpropagation of error*,<sup>6</sup> cujo objetivo é ajustar o peso das conexões durante o treinamento da rede a fim de aprender um único conjunto de pesos, de modo que qualquer padrão de *input* vá produzir o correto padrão de *output*.

## 2.3 Treinando a rede

O treinamento é composto de etapas em que a rede recebe pares de cena/tópico e uma descrição do evento ora na forma ativa, ora na forma passiva. A rede é treinada para produzir uma frase a partir da cena/tópico no *input*. Existem 160 padrões de *input* (pares cena/tópico) e 160 padrões de *output* (frases). Portanto são necessários 160 ciclos para que a rede seja exposta a todo o conjunto de treinamento, o que constitui uma “época”. A rede foi treinada por 6.000 impulsos ou *sweeps*, ou seja, aproximadamente 37 épocas. A rede foi treinada randomicamente, sem substituição, a fim de garantir que todos os padrões do *input* fossem apresentados à rede em uma dada época. A taxa de aprendizagem foi fixada em 0,1 e a taxa *momentum* em 0,9.

<sup>6</sup> O algoritmo de aprendizagem conhecido como *backpropagation of error* é usado em Rumelhart, Hinton e Williams (1986), Plunkett e Elman (1997) e inúmeras redes conexionistas.

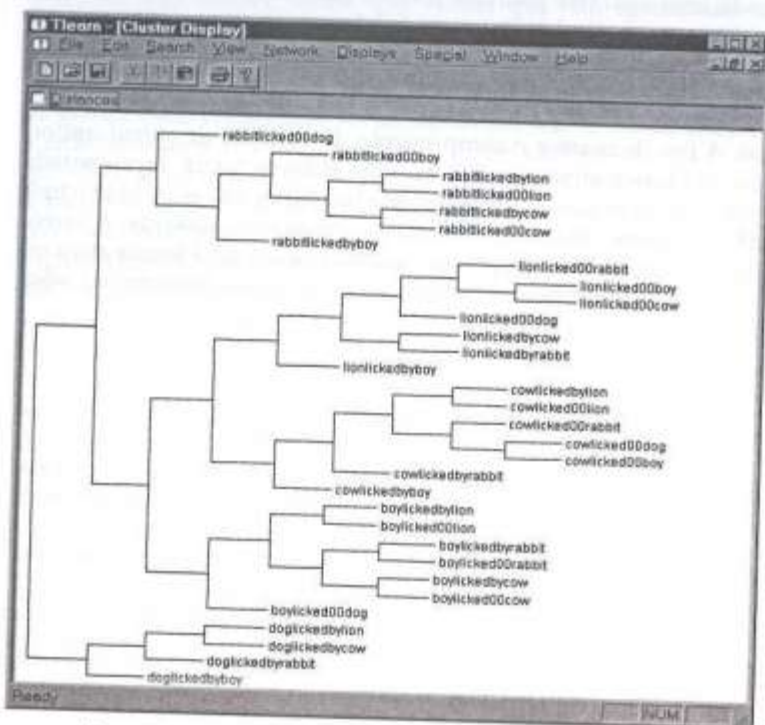


Figura 1. Análise por grupos de 32 vetores usando a ativação das unidades *hidden* após 6.000 impulsos de treinamento

Testamos a rede com o conjunto de treinamento usando o peso das conexões após 6.000 impulsos. Todas as 160 frases produzidas pela rede se igualaram à sentença alvo, isto é, a rede produziu uma frase ativa toda vez que um agente foi topicalizado e uma frase passiva quando um não-agente foi topicalizado. Como a rede aprende a manter o tópico discursivo dado no *input*? A fim de analisar a solução da rede, observamos como a estrutura dos padrões de *input* semelhantes é modificada como resultado da passagem da camada de *input* para as representações internas. Tomando um subconjunto de 32 vetores<sup>7</sup> da ativação das unidades *hidden*, uma

<sup>7</sup> Em uma situação ideal, gostaríamos de visualizar os 160 vetores usados no conjunto de treinamento; contudo, restrições no tamanho do papel e da fonte para impressão exigem uma versão mais modesta do diagrama por grupos exibido aqui. Portanto, o

*cluster analysis* ou análise por grupos hierárquicos foi desenvolvida após a testagem, esboçando um diagrama em forma de árvore das estruturas semelhantes. Como pode ser visto na Figura 1, a rede agrupou as frases de acordo com qual palavra se encontra na posição tópico, independentemente de seu papel como agente ou não-agente na cena e da estrutura frasal usada.

### 3 Conclusão

É possível desenvolver um modelo conexionista capaz de aprender a preservar o tópico discursivo sem lançar mão de regras explícitas que o instruem a assim proceder. Contudo, é prematuro afirmar que a solução encontrada pelo modelo computacional seja idêntica ou mesmo semelhante à estratégia usada por seres humanos na aquisição da linguagem. O que é possível dizer com relativa segurança é que o *input* linguístico carrega informações de natureza explícita e implícita. Se um modelo computacional é capaz de apreender essas informações, por que não o seria também o cérebro humano? Uma das coisas que a nossa rede conexionista *feed-forward* mostra é que redes simples são capazes de descobrir representações internas úteis e interessantes para tarefas linguísticas. Obviamente, modelos computacionais mais sofisticados são necessários para dar conta das múltiplas informações envolvidas na aquisição e processamento da linguagem e nas demais tarefas cognitivas. No entanto, acreditamos que, ao congregarmos dados comportamentais, neuronais e computacionais, as redes conexionistas constituam uma ferramenta útil na busca de respostas e de novas questões.

### Referências bibliográficas

- BATES, E.; DEVESCOVI, A. Crosslinguistic studies of sentence production. In: MACWHINNEY, B.; BATES, E. *The crosslinguistic study of sentence processing*. Cambridge: Cambridge University Press, 1989, p. 225-253.
- BATES, E.; HARRIS, C.; MARCHMAN, V.; WULFECT, B.; KRITCHEVSKY, M. Production of complex syntax in normal ageing and Alzheimer's disease. *Language and Cognitive Processes*, v. 10, n. 5, p. 487-539, 1995.
- CLARK, H. H.; BEGUN, J. S. The use of syntax in understanding sentences. *The British Journal of Psychology*, v. 59, n. 3, p. 219-229, 1968.

leitor precisa acreditar que o mesmo padrão apresentado na amostra de 32 vetores é observado no diagrama contendo os 160 vetores.

- FERREIRA, F. Choice of passive voice is affected by verb type and animacy. *Journal of Memory and Language*, 33, p. 715-736, 1994.
- LIVERSEDGE, S. P.; PICKERING, M. J.; BRANIGAN, H. P.; VANGOMPEL, R. P. G. Processing arguments and adjuncts in isolation and context: the case of by-phrase ambiguities in passives. *Journal of Experimental Psychology: Learning, Memory and Cognition*, v. 24, n. 2, p. 461-475, 1998.
- MARCHMAN, V. A.; BATES, E.; BURKARDT, A.; GOOD, A. B. Functional constraints in the acquisition of the passive: toward a model of the competence to perform. *First Language*, 11, p. 65-92, 1991.
- PINKER, S. *Words and rules: the ingredients of language*. London: Weidenfeld & Nicolson, 1999.
- PLUNKETT, K.; ELMAN, J. L. *Exercises in Rethinking Innateness: a handbook for connectionist simulations*. London: The MIT Press, 1997.
- PLUNKETT, K.; MARCHMAN, V. U-shaped learning and frequency effects in a multi-layered perceptron: implications for child language acquisition. *Cognition* 38, 43-102, 1991.
- RUMELHART, D. E.; MCCLELLAND, J. L. On learning the past tenses of English verbs. In: MCCLELLAND, J. L.; RUMELHART, D. E.; THE PDP RESEARCH GROUP. *Parallel Distributed Processing: explorations in the microstructure of cognition*. London: The MIT Press, 2, p. 216-271, 1986.
- RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning internal representations by error propagation. In: MCCLELLAND, J. L.; RUMELHART, D. E.; THE PDP RESEARCH GROUP. *Parallel Distributed Processing: explorations in the microstructure of cognition*. London: The MIT Press, 1, p. 318-362, 1986.