



# A AQUISIÇÃO DA FRICATIVA /ʒ/ E O ALGORITMO GRADUAL DE APRENDIZAGEM: EVIDÊNCIAS EMPÍRICAS VIA VALORES DE RESTRIÇÕES

Felipe Flores Kupske<sup>1</sup>  
Cristiano Egger Veçossi<sup>2</sup>  
Giovana Ferreira-Gonçalves<sup>3</sup>

## 1. Introdução

Os avanços da neurociência e da área computacional têm trazido contribuições para diversos campos do conhecimento. Na área da aquisição da linguagem, é crescente o uso de recursos digitais em favor da compreensão de dados empíricos. A relevância de tais programas é comprovada à medida que, quando alimentados com dados, eles passam a se comportar de modo semelhante ao aprendiz real.

Nesse sentido, o presente trabalho apresenta os achados obtidos a partir da rodagem de dados empíricos em um programa computacional. Trata-se do OT Soft 2.3, desenvolvido por Tesar, Hayes e Zuraw (2003), pensado para facilitar as análises via Teoria da Otimidade (TO) com o uso de algoritmos de aprendizagem, no caso desse estudo, o *Gradual Learning Algorithm* (GLA) – algoritmo de aquisição gradual.

O foco de estudo é o processo de aquisição da fricativa /ʒ/, acompanhado por meio de seis coletas realizadas com o sujeito M. Isolados os casos de ocorrência de /ʒ/, submetemos tais dados ao programa computacional OTSoft. Após a rodagem, observamos que o programa, junto ao algoritmo selecionado,

---

<sup>1</sup> Mestrando da Universidade Federal de Santa Maria (UFSM). kupske@gmail.com

<sup>2</sup> Mestre em Letras/UFSM. Professor Substituto do Centro de Educação da UFSM. cristiano\_letras@yahoo.com.br

<sup>3</sup> Professora Adjunto da Universidade Federal de Pelotas (UFPeL). gfgb@terra.com.br



forneceu um *ranking* de restrições condizente com os estágios da gramática do aprendiz, sendo capaz de traduzir a variação na produção do sujeito M. para um conjunto de restrições com pesos diferenciados.

Assim, nosso artigo encontra-se organizado em seis seções básicas. Além da introdução, em 2, apresentamos informações de cunho teórico acerca da aquisição da linguagem sob o enfoque da Teoria da Otimidade (TO). Em 3, concluímos o referencial teórico, tratando da aquisição da fricativa /ʒ/ no português brasileiro. Em 4, apresentamos informações sobre a metodologia de nosso estudo. Em 5, tratamos da descrição e análise dos dados. Por fim, a seção 6 traz as conclusões a que chegamos com nosso estudo.

## 2. Aquisição da linguagem e a Teoria da Otimidade

De acordo com Del Ré (2006, p. 18), o pressuposto fundamental do paradigma empirista - base para o desenvolvimento da Teoria da Otimidade em um viés conexionista -, com relação à aquisição da linguagem, é a inexistência da mente. Desse modo, todo conhecimento humano provém da experiência, da associação entre estímulos e respostas. No interior deste paradigma, encontram-se as teorias behaviorista e conexionista.

Em termos aquisicionais, para o behaviorismo, a criança desenvolve seu conhecimento linguístico por meio dos procedimentos de estímulo e resposta, imitação e reforço. Também inserido no paradigma empirista, o conexionismo apresenta-se mais abrangente e com maior poder explicativo, visto que tenta elucidar como se dá o aprendizado a partir da consideração dos efeitos da experiência empírica sobre as redes neuronais do cérebro. Conforme Bonilha (2007, p. 189): *[...] o conexionista procura 'ver', através da rede neuronal, exatamente o que está acontecendo nesse processo que transforma um input em um output específico.*

De acordo com Poersch (2007, p. 245), o conexionismo, entendido como um modelo cognitivo, apresenta como pressuposto fundamental a ideia de que a

Verba  
Volant

VERBA VOLANT

Volume 1 – Número 1 – julho - dezembro 2010 – ISSN 2178-4736

informação não se encontra em um módulo específico do cérebro, mas, ao contrário, é armazenada de maneira distribuída nos neurônios. Além disso, o processamento da informação é concebido como sendo em paralelo, ao contrário do modo serial como este é compreendido no modelo simbólico.

É a partir da ideia de que o cérebro humano funciona como um computador, no qual os neurônios, tal como os processadores, encontram-se interligados, que os estudiosos dessa área realizam suas pesquisas. Nesse sentido, são realizados experimentos computacionais, nos quais as redes são alimentadas com dados, visando-se, assim, a simular o funcionamento do cérebro de um aprendiz diante do contato com produções verbais empíricas de sua língua. É por meio da equiparação dos resultados encontrados em pesquisas realizadas com crianças em fase de aquisição e aqueles encontrados em estudos computacionais que se comprova a validade dessas simulações.

No processo de aprendizagem, seja qual for o tipo de habilidade a ser desenvolvida, exerce importante papel o algoritmo de aprendizagem, o qual, segundo Poersch (2007, p. 256), *ajusta a força das conexões [neurônios] na rede até ser alcançado o desempenho desejado.*

## 2.1. Algoritmos de aprendizagem

Conforme Poersch (2005, p. 175), modelos conexionistas podem ser treinados para executar uma grande variedade de tarefas como, por exemplo, prever a re-aparição de um objeto atrás de uma tela, transformar um verbo em sua forma passada, prever a palavra seguinte em uma sentença etc. Em cada caso, o algoritmo de aprendizagem ajusta a força das conexões na rede até que uma *performance* adulta/madura seja atingida. Assim, a rede pode ser analisada para checarmos como esta executa uma dada tarefa, fornecendo, assim, fontes de hipóteses de como os adultos a executariam. Também é possível examinar o perfil desenvolvimental da rede ao monitorá-la em intervalos de tempo regulares.



Se as atitudes de uma rede no processo de aprendizagem podem nos fornecer informações acerca da criança no processo de aquisição, talvez o monitoramento possa nos dizer muito sobre os diferentes estágios percorridos pela criança no processo.

Redes aprendem ao trocar a força das conexões em resposta a uma atividade neuronal. Um dos trabalhos clássicos acerca da aquisição da linguagem é o de Rumelhart e McClelland (1986). Os autores utilizam uma rede neuronal de camada única, possibilitando que a rede adquira as formas regular e irregular dos verbos do inglês. O dispositivo de aprendizagem é dirigido pelos dados, sem recorrer a um sistema simbólico de regras. Ao apresentar formas de verbos regulares e irregulares à rede, essa aprende fazendo ajustes aos valores de suas conexões, ao comparar o que produziu com a forma pretendida. Para tal ajuste, um algoritmo de aprendizagem é aplicado, conduzindo à produção de uma forma desejada gradualmente.

Basicamente são dois os algoritmos utilizados com fins de análise linguística: o CDA (*Constraints Demotion Algorithm – TESAR & SMOLESNKY, 1996, 1998, 2000*) e o GLA (*Gradual Learning Algorithm – BOERSMA & HAYES, 2001*). Bonilha (2007) aponta a incapacidade do CDA em lidar com os dados empíricos da aquisição fonológica e a capacidade promissora do GLA em lidar com os mesmos dados. Neste sentido, adotamos e aplicamos o GLA em seis coletas do banco de dados LIDES, buscando dados que corroborem os achados da autora.

O GLA é um aprendiz *online* conduzido pelo erro, sendo alimentado com um par de *input-output* correto a cada processamento. O algoritmo, então, determina o *output* ótimo para um dado *input*, conforme o estado em que a gramática se encontra. Se o *output* gerado é diferente das referências/dados de entrada, a aprendizagem é provocada. O GLA atualiza os valores das restrições, subtraindo um valor X dos valores das restrições violadas pelo candidato ótimo - forma correta/alvo - e adicionando o mesmo valor X a todas as restrições que são violadas pelo candidato subótimo – forma com erro produzida pelo aprendiz.



Verba  
Volant

VERBA VOLANT

Volume 1 – Número 1 – julho - dezembro 2010 – ISSN 2178-4736

De maneira mais detalhada, segundo Boersma e Hayes (2001), o GLA tenta localizar um valor de *ranking* empiricamente apropriado para cada restrição analisada. Podemos dividir esse processo em seis partes: estado inicial; passo 1: referência; passo 2: geração; passo 3: comparação; passo 4: ajuste; e estado final.

No estado inicial, as restrições possuem valores de *ranking* estipulados pelo linguista. Dessa forma, todas as restrições podem começar com os mesmos valores, ou podem incorporar valores distintos (fidelidade baixa ou alta, por exemplo). Então, no passo 1, o algoritmo é alimentado com uma referência de aquisição como, por exemplo, produções de adultos que um aprendiz escuta em seu ambiente e assume como corretas. Já que o algoritmo é dirigido pelo erro, o passo 2 envolve checar o que a gramática atual gera para a forma de *input* assumida. No passo 3, caso a forma gerada pela gramática seja igual à forma alimentada ou forma esperada, nada mais é feito, mas, caso haja uma má combinação, um ajuste da gramática ocorrerá. Especificamente, o algoritmo compara as violações dos dados de referência alimentados com o que é atualmente gerado pela gramática. O algoritmo toma essa má combinação como um sinal para alterar a gramática para que esta gere o candidato esperado. A alteração será na forma de uma mudança nos valores das restrições, produzindo um novo *ranking*. O passo 4 é encarregado de fazer os últimos ajustes em todas as restrições que envolvam violações não canceladas pelo processo. Com ainda mais exposição aos dados de aprendizagem, chegamos ao estado final, no qual o algoritmo se aplica ciclicamente por meio dos passos de 1 a 4.

## 2.2. A Teoria da Otimidade

Considerando que o conexionismo é uma teoria cognitiva, e não uma teoria linguística, torna-se necessário recorrer a uma teoria que considere os avanços da neurociência, tomando-os como pressupostos fundamentais, mas que seja

aplicável à descrição e análise linguística. Nesse sentido, adquire especial relevância a Teoria da Otimidade (TO).

Em termos formais, há, por meio da TO, a possibilidade de se aplicar certos pressupostos do paradigma conexionista à análise linguística. Mesmo que, em sua versão *standard* (PRINCE e SMOLENSKY, 1993; MCCARTHY e PRINCE, 1993), a TO ainda se encontre embasada em alguns pressupostos gerativistas, é bastante perceptível que esta teoria, desde seu nascimento, já se encontra influenciada por ideias de cunho conexionista.

A TO considera que a gramática de uma língua é formada por um ranking de restrições, todas elas violáveis pelo aprendiz. Nesse sentido, aprender determinada língua significa adquirir o ranqueamento específico relativo a essa língua. Não sendo derivacional, a teoria (em sua versão inicial, híbrida) considera a existência de uma representação subjacente, da qual um dispositivo chamado GEN gera em paralelo um conjunto de *outputs* possíveis, sendo que todos eles são avaliados por um outro dispositivo chamado EVAL, a partir de um conjunto de restrições<sup>4</sup> (CON), as quais se encontram organizadas (ranqueadas) de maneira diferente em cada língua. Cabe assinalar que os dispositivos mencionados (GEN, EVAL e CON), na TO *standard*, são ainda considerados parte da Gramática Universal, aspecto que sofre alteração na Teoria da Otimidade Conexionista (COT), proposta por Bonilha (2004).

Considerando, portanto, a existência de um conjunto de candidatos a *output*, de restrições ranqueadas de determinada maneira e de um *input*, a forma ótima é resultado da interação entre candidatos a *output* e *input*, via restrições

---

<sup>4</sup> Com relação às restrições, costuma-se considerar que são de dois tipos: marcação e fidelidade. As restrições de marcação, ranqueadas mais acima na hierarquia, garantem que formas mais simples ou mais frequentes na língua (menos marcadas, portanto) sejam produzidas. As restrições de fidelidade implicam a relação entre *input* e *output*, de modo que, quando ranqueadas mais acima na hierarquia, visam a garantir que as formas produzidas sejam as mais próximas possíveis da língua-alvo. Desse modo, acredita-se que, no início do processo de aquisição, marcação domine fidelidade, haja vista que a criança, embora omita ou insira segmentos (violando, portanto, as restrições de fidelidade), tende a produzir o que é menos marcado na língua.



Verba  
Volant

VERBA VOLANT

Volume 1 – Número 1 – julho - dezembro 2010 – ISSN 2178-4736

hierarquizadas. O candidato ótimo será aquele que violar as restrições ranqueadas mais abaixo na hierarquia.

Conforme Bonilha (Idem, p. 51), na COT, considera-se que a fonologia não está na chamada forma subjacente, conforme proposta de teorias gerativas, ou seja, as unidades fonológicas emergem da hierarquia de restrições. Desse modo, na COT, a forma subjacente não é vista como um nível de representação fonológica separado, mas, ao contrário, provém do *input* (*output* do adulto), via ranqueamento de restrições. Conforme Bonilha (Ibid., p. 54), adquirir o fonema implica, simplesmente, ranquear as restrições de marcação abaixo das restrições de fidelidade. Assim, o *input* acionado será “o mais fiel possível à forma ouvida pela criança”.

Outra mudança, em relação ao modelo standard de 1993, proposta pela autora, é a substituição das funções GEN e EVAL pelo Otimizador, responsável pela emergência da forma ótima, na qual está contido o *input*.

Coerente com uma perspectiva conexionista, para a COT, restrições são universais apenas enquanto potencialidade, desta forma, o processo de aquisição da linguagem inclui não apenas o reordenamento das restrições, visando à hierarquia alvo, mas a emergência das restrições que constituem a gramática.

A versão conexionista proposta por Bonilha (2004) pode ser considerada em um viés estocástico da Teoria da Otimidade, pois o GLA, proveniente deste modelo, é, hoje, o algoritmo mais adequado a ser aplicado à análise dos dados.

Na TO estocástica, as restrições recebem valores numéricos, entretanto, cada vez que a gramática é utilizada para avaliar um conjunto de candidatos, os valores são convertidos em um *ranking* correspondente. A diferença numérica entre as restrições é relevante após tal conversão: se C1 tem um valor maior que C2, então o *ranking* correspondente é  $C1 \gg C2$ , independente do tamanho da diferença  $C1 - C2$ . A distância numérica, na verdade, tem importância no processo de conversão em si. Antes da conversão de números em um *ranking*, cada um desses é “perturbado”, adicionando-se um número positivo ou negativo diferente retirado de uma distribuição normal. Em avaliações sucessivas,





restrições que possuem valores numéricos suficientemente próximos uma da outra irão variar em seus *rankings*. Um aspecto atrativo desta TO é, justamente, o fato de vir acompanhada pelo GLA (COETZEE e PATER, 2008).

### 3. A aquisição da fricativa /ʒ/

Conforme a literatura da área, o segmento /ʒ/ pertence à classe das fricativas. Em termos de ordem de aquisição, tal classe de sons segue as plosivas e nasais, havendo, no interior do grupo das fricativas, diferenças entre /f/ e /v/, considerados de aquisição mais precoce e /s/, /z/, /ʃ/ e /ʒ/, de aquisição mais tardia (OLIVEIRA, 2004, p. 83).

A pesquisa de Oliveira (2002) tratou dos segmentos fricativos /f/, /v/, /ʃ/ e /ʒ/ a partir de um *corpus* composto por 1527 palavras, extraídas das produções de 103 crianças, na faixa etária de 1:0 e 3:8, e analisadas por meio do programa estatístico VARBRUL.

Com relação ao fonema /ʒ/, conforme a autora, esse começa a ser produzido aos 1:6, de modo que se encontra adquirido apenas aos 2:6. No caso do português brasileiro, tal segmento só compõe *onset* silábico (inicial ou medial). Como estratégias de reparo aplicadas pela criança na tentativa de produzir tal segmento, estão a omissão do fonema ou da sílaba portadora deste e a substituição por outro segmento.

Nos dados da autora, a maioria dos casos de omissão de /ʒ/ ocorreram em sílaba pré-pretônica e, em apenas uma palavra, em sílaba tônica. Como substituições, são observados principalmente processos de anteriorização – produção de [z] –, desonorização - produção de [j] –, substituição do valor do traço [contínuo] – troca de /ʒ/ por [d] – e do traço [coronal] – produção de [v].

Como fatores favoráveis à produção de /ʒ/, o programa VARBRUL apontou os seguintes contextos: (i) seguinte, /u/; (ii) precedente, coda com /n/; (iii) posição na palavra, *onset* absoluto; (iv) número de sílabas, trissílabas e (v) tonicidade, preservação da sílaba pretônica.





#### 4. Metodologia

Neste artigo, analisamos dados de aquisição do sujeito M., sexo masculino, com idade entre 1:9 e 2:2, cujo desenvolvimento fonológico foi acompanhado em seis coletas longitudinais. O *corpus* é oriundo do banco de dados LIDES<sup>5</sup>(Linguagem Infantil em Desenvolvimento).

A análise aqui realizada diz respeito à fricativa /ʒ/, observada em todas as palavras nas quais houve possibilidade de produção (posições de onset inicial e onset medial). As coletas, mensais, foram realizadas de outubro de 2006 a abril de 2007, com exceção do mês de fevereiro de 2007, no qual não houve coleta.

A partir da observação dos dados, isolamos as ocorrências nas quais havia a possibilidade de produção de /ʒ/. Feito isso, submetemos tais dados ao programa computacional OT Soft 2.3, desenvolvido por Hayes, Tesar e Zuraw (2003). Além dos dados, fornecemos ao programa restrições que acreditávamos ser necessárias para dar conta da análise e suas respectivas violações, conforme (1):

(1)

\*σ[w]: proibindo glide dorsal em início de sílaba.

\*[j]: proibindo fricativa [coronal, - anterior, -sonora]

\*[ʒ]: proibindo fricativa [coronal, - anterior, + sonora]

\*[v]: proibindo fricativa [labial, + sonora]

MAX I/O: segmentos do *input* devem ser preservados no *output*.

IDENT [sonoro], IDENT [consonantal], IDENT [anterior]: identidade entre as formas de *input* e *output* no que se refere aos traços [sonoro], [consonantal] e [anterior], respectivamente.

---

<sup>5</sup> O banco é coordenado pelas professoras Márcia Cristina Zimmer (UCPel) e Giovana Ferreira-Gonçalves (UFPeI).



Com o programa computacional, realizamos dois procedimentos. O primeiro visava a descobrir o modo como as restrições dadas estariam ranqueadas em cada uma das coletas, bem como o grau de influência de cada restrição sobre os estágios de gramática do aprendiz. Para a realização deste, o *software* foi programado de modo que as restrições de fidelidade receberam o valor zero e as restrições de marcação valor 100. A plasticidade inicial e final receberam os valores 3 e 001, respectivamente, com um valor de ruído 2. Foram simuladas um milhão de tentativas de aprendizagem e as formas foram analisadas 1000 vezes por item lexical.

Com o segundo procedimento, objetivávamos testar o aprendiz. Para isso, fizemos mais 5 aplicações do GLA. Diferentemente das aplicações até aqui descritas, não rodamos os dados com a produção do sujeito M. marcadas como ótimas, mas sim as produções corretas esperadas no Português Brasileiro. O programa foi alimentado com o mesmo conjunto de restrições, mesmos pesos, ruído e números de testagens das aplicações para os dados de M.. Rodamos as restrições para cada item lexical utilizado no estudo e fizemos uma rodagem que contemplou todos os itens simultaneamente.

A descrição e a análise dos dados são apresentadas na seção seguinte.

## **5. Descrição e análise dos dados**

### **5.1 A aquisição de /ʒ/ pelo sujeito M.**

Os quadros relativos a cada uma das coletas estão evidenciados em (2). Na primeira coluna, apresentamos a forma-alvo. Na segunda, a transcrição fonética da forma que foi produzida pelo sujeito. Em cada uma das coletas, nos casos em que houve mais de uma ocorrência da mesma forma, há a indicação do número de vezes que a ocorrência foi produzida.

(2)

a) Coleta realizada em 03/10/06 – idade: 1:9

1. Jogar	[ʒo'ga] (1x)
2. João	[vu'ãw] (5x), [wu'ãw] (1x), [u'ãw] (1x)

b) Coleta realizada em 09/11/06 – idade: 1:10

1. João	[wu'ãw] (2x)
---------	--------------

c) Coleta realizada em 06/12/06 – idade: 1:11

1. Jogar	[ʒa] (2x)
2. João	[wu'ãw] (17x), [ʒu] (2x)
3. Juju	[wu'vu] (8x)

d) Coleta realizada em 31/01/07 – idade: 2:0

1. Jogar	[ʒo'ga] (1x)
----------	--------------

e) Coleta realizada em 16/03/07 – idade: 2:1

1. João	[wu'ãw] (2x)
2. Juju	[wu'vu] (3x)
3. Colégio	[ku'leʒiw] (2x) [ʒu du'εʒiw] (1x) [ʒu'leʒyu] (1x)

f) Coleta realizada em 19/04/07 – idade: 2:2

1. João	[wu'ãw] (1x)
2. Jogar	[ʒo'ga] (2x)
3. Juju	[vu'vu] (4x)
4. Colégio	[ku'εʒyu] (1x)

Nos quadros apresentados, observamos as palavras que foram encontradas nos dados do sujeito M., bem como as estratégias de reparo empregadas pelo sujeito na tentativa de produzir o segmento /ʒ/. Os itens lexicais produzidos foram, em ordem decrescente quanto ao número de ocorrências: *João* (31 ocorrências), *Juju* (15 ocorrências), *jogar* (6) e *colégio* (5). Assim, considerando todas as

tentativas de produção deste segmento, tanto em onset inicial quanto em onset medial, o *corpus* de nosso estudo é composto por 57 palavras. Tendo em vista que, no caso de *Juju*, há duas vezes o fonema /ʒ/, nosso corpus apresenta 74 contextos de possível produção de tal segmento.

Com relação às estratégias de reparo aplicadas por M. na tentativa de produzir o fonema /ʒ/, elas são apresentadas no gráfico 1.

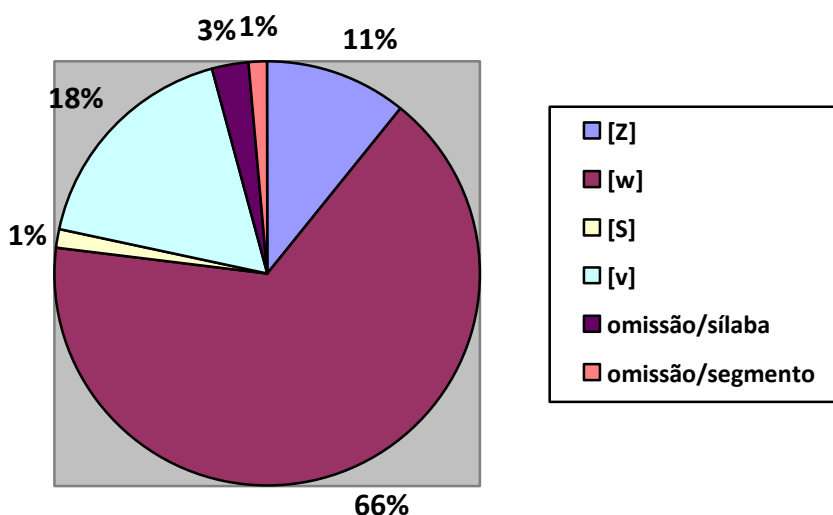


Gráfico 1 – Estratégias de reparo aplicadas por M. na produção de /ʒ/

Por meio do gráfico, observamos que a criança produz corretamente o segmento /ʒ/ em apenas 11% das ocorrências. A substituição pelo glide posterior é a estratégia de reparo mais utilizada pela criança na tentativa de produzir a fricativa [-anterior]. Em segundo lugar, está a substituição por [v] e, em menor número, a substituição por [j] (dessonorização), e o apagamento do segmento ou da sílaba portadora dele.

No quadro 1, apresentamos os valores numéricos, com relação a cada um dos itens lexicais encontrados nos dados, referentes às estratégias utilizadas pelo sujeito M. na tentativa de produzir o segmento /ʒ/.

Item lexical	Estratégia aplicada	Número de ocorrências	Percentual
<i>João</i>	Produção correta	0	0%
	Substituição por glide [w]	23	79,3%
	Substituição por [v]	5	17,2%
	Omissão do segmento	1	3,5%
<i>Juju</i> <sup>6</sup>	Produção correta	0	0%
	Substituição por glide [w]	26	76,5%
	Substituição por [v]	8	23,5%
<i>jogar</i>	Produção correta	3	50%
	Omissão da sílaba que contém o segmento	2	33%
	Substituição por [j]	1	17%
<i>colégio</i>	Produção correta	5	100%

Quadro 1 – Itens lexicais e estratégias de reparo aplicadas

O quadro permite perceber que, à exceção das produções referentes aos itens *jogar* e *colégio*, nos quais há, respectivamente, cinquenta e cem por cento de produções corretas, para os demais itens, há predomínio da substituição por glide, seguido pela substituição por [v], ou seja, substituição por segmentos labiais. Tal fato parece ser explicado pelo contexto vocálico seguinte, pois, em [ʒu'ãw] e [ʒu'ʒu], temos a presença da vogal dorsal/labial [u], o que pode estar

<sup>6</sup> Considerando que nesta palavra há dois contextos possíveis para a produção de /ʒ/, os valores numéricos apresentados dizem respeito às estratégias aplicadas ao segmento em ambos os contextos.



Verba  
Volant

VERBA VOLANT

Volume 1 – Número 1 – julho - dezembro 2010 – ISSN 2178-4736

servindo como gatilho para a presença do glide labial e da fricativa labial nas estratégias de reparo.

A substituição do traço [coronal], com o surgimento de [v], também é constatada nos dados de Oliveira (2002), no entanto, de acordo com os resultados da autora, a vogal [u] serviria como contexto favorável à produção de /ʒ/, o que não foi confirmado nos dados de M.. Na verdade, a criança pode estar apresentando uma tendência a sequências que apresentem o mesmo ponto de articulação em suas produções.

Já no que concerne às omissões, sejam essas relacionadas ao segmento ou à sílaba, seguem o padrão verificado em Oliveira (2002), ou seja, ocorrem quando o segmento alvo ocupa a sílaba pretônica.

Cabe assinalar que a produção correta de /ʒ/ em cem por cento das ocorrências da palavra *colégio* sugere que, nesse caso, a aquisição tenha sido dada por item lexical. Podemos supor que esta palavra seja produzida frequentemente no contexto no qual a criança vive, de modo que este tenha sido um dos fatores responsáveis pela aquisição tão precoce pelo sujeito M.. Outro fator que pode estar militando é o fato de o contexto vocálico seguinte apresentar o segmento coronal [i], o que pode ser facilitador na realização da fricativa, igualmente coronal, /ʒ/.

## 5.2. Comparação entre a aquisição de /ʒ/ pelo sujeito M. e pela máquina

No primeiro procedimento realizado, os dados produzidos pela criança foram submetidos ao programa *OT Soft*, objetivando descobrir os ranqueamentos – estágios da gramática do aprendiz – de restrições referentes a cada uma das coletas.

Com relação à primeira coleta, apresentamos, no quadro 2, na coluna à esquerda, os dados produzidos pelo sujeito M. e, à direita, a hierarquia de restrições fornecida pelo programa a partir da análise de tais dados.



Dados do sujeito M.	Hierarquia dada pelo programa
<u>J</u> oão = 5 ocorrências com [v]	100,000 *[j]
1 ocorrência com [w]	75,638 MAX I/O
1 omissão do segmento	75,478 *σ [w],
<u>j</u> ogar = 1 ocorrência com [ʒ]	74,965 *[v]
	73,920 *[ʒ]
	0,000 IDENT [+ sonoro]
	-24,522 IDENT [+ consonantal]
	-25,035 IDENT [- anterior]

Quadro 2 – Dados da 1ª. coleta de M. e hierarquia de restrições

Conforme o ordenamento das restrições apresentado à direita, podemos ver que a restrição de fidelidade IDENT [+sonoro] possui valor zero e que a restrição de marcação \*[j] está ranqueada acima de todas as outras, apresentando peso 100,000. Dessa forma, o segmento /j/ jamais seria produzido nessa gramática. Podemos notar, também, que MAX, \*σ [w], \*[v] e \*[ʒ] possuem valores entre 73,920 e 75,638, isto é, menos de 1,718 pontos de diferença, o que caracteriza variação linguística com a produção de [w], [v], [ʒ] e do apagamento do segmento /ʒ/.<sup>7</sup> Dado o seu baixo valor, a restrição IDENT [- anterior] sugere uma preferência para a produção de [v], pois esse é [+anterior]. Como se vê na coluna à direita, é isso que ocorre nos dados de M., já que ele produz, na maior parte das vezes, o segmento [v], havendo, também, a produção do glide dorsal e de [ʒ].

Para a coleta 2, os dados do aprendiz e o ranqueamento apresentado pelo GLA podem ser visualizados no quadro 2:

<sup>7</sup> Tal como outros trabalhos da área e, conforme o manual de instruções do programa computacional utilizado, estamos considerando a diferença de até 10,000 pontos entre o peso de uma restrição e o de outra como indicativo de variação linguística.





DADOS DO SUJEITO M.	HIERARQUIA DADA PELO PROGRAMA	
<u>João</u> = 2 ocorrências com [w]	100,000	*[ç]
	100,000	*[j]
	100,000	*[v]
	59,080	MAX I/O
	40,920	*σ [w],
	0,000	IDENT [+ sonoro]
	0,000	IDENT [- anterior]
	-59,080	IDENT [+ consonantal]

Quadro 2 – Dados da 2ª. coleta e hierarquia de restrições

Dada esta gramática da segunda coleta de M., podemos claramente ver que os outputs [ç], [j] e [v] não seriam produzidos, já que as restrições de marcação proibindo tais produções se encontram com valor 100,000 e, conseqüentemente, no topo do ranqueamento. O fato de MAX I/O dominar \*σ [w] indica a preferência pela produção do glide em detrimento do apagamento. Embora ambas as restrições possuam valores relativamente próximos, esses apresentam mais de 10,000 pontos de diferença, o que não caracteriza variação. O fato de IDENT [+consonantal] estar ranqueada mais abaixo e possuir o menor dos valores também indica a preferência pela produção do glide, já que as outras duas restrições IDENT possuem valores zero.

Dando continuidade a nossa descrição, para a terceira coleta, são observados os seguintes dados e restrições.

DADOS DO SUJEITO M.	HIERARQUIA DADA PELO PROGRAMA	
<u>João</u> = 17 ocorrências com [w]	100,000	*[ç]
<u>Juju</u> = 10 ocorrências com [w]	100,000	*[j]
<u>jogar</u> = 2 omissões do segmento	100,000	*[v]
	65,080	IDENT [- anterior]
	53,080	MAX I/O
	46,920	*s [w],
	0,000	IDENT [+ sonoro]
	-53,080	IDENT [+ consonantal]

Quadro 3 – Dados da 3ª. coleta e hierarquia de restrições



Como na gramática da segunda coleta, podemos ver que \*[ɜ], \*[j] e \*[v] possuem valores de 100,000 e estão no topo da hierarquia, impossibilitando a produção dos respectivos segmentos. IDENT [- anterior], MAX I/O e \*σ [w] estão ranqueados com valores intermediários, entretanto, MAX I/O e \*σ [w] possuem 7,000 pontos de diferença entre si, estando ranqueados abaixo de IDENT [- anterior], o que configura a variação entre o apagamento e a produção do glide em detrimento de [v]. Como nenhum candidato violou a restrição de fidelidade IDENT [+sonoro], esta permanece com valor zero, e o número negativo de IDENT [+ consonantal] representa a preferência pela produção do glide. Desse modo, vemos que o ordenamento das restrições reflete o estágio atual da gramática do aprendiz: ele produz, na grande maioria das vezes, o glide [w], realizando, também, o apagamento como estratégia de reparo.

Com relação à quarta coleta, o quadro 4 apresenta os dados do aprendiz e o ranqueamento gerado pelo programa a partir dos dados empíricos.

DADOS DO SUJEITO M.	HIERARQUIA DADA PELO PROGRAMA
jogar = 1 ocorrência com [j]	100,000    *[ɜ]
	100,000    *[v]
	100,000    *σ [w],
	57,000    MAX I/O
	43,000    *[j]
	0,000    IDENT [- anterior]
	0,000    IDENT [+ consonantal]
	-57,000    IDENT [+ sonoro]

Quadro 4 – Dados da 4ª. coleta e hierarquia de restrições

Observamos que a hierarquia gerada pelo programa referente à quarta coleta mostra-se condizente com os dados empíricos de M.. Nesta coleta, na tentativa de produzir o segmento-alvo, o aprendiz acaba produzindo a correspondente desvozeada [j]. Desse modo, a hierarquia dada pelo programa



reflete tal situação, visto que as restrições de marcação referentes aos segmentos [ʒ], [v] e [w] estão ranqueadas mais acima na hierarquia do aprendiz, com peso máximo de 100,000, o que impede que os fonemas em questão sejam produzidos. Já a restrição de fidelidade IDENT [+ sonoro] encontra-se na última posição da hierarquia, o que permite que o aprendiz, por um processo de dessonorização, produza [ʃ]. A restrição de marcação que proíbe a produção da fricativa anterior surda tem um valor de peso baixo – 43,000 –, encontrando-se ranqueada numa posição mais inferior na hierarquia, o que corrobora a produção realizada pela criança.

O quadro 5 apresenta os dados do sujeito M. e o ranqueamento gerado pelo programa computacional referentes à quinta coleta.

DADOS DO SUJEITO M.	HIERARQUIA DADA PELO PROGRAMA
<i>João</i> = 2 ocorrências com [w]	100,000 * [ʃ]
<i>Juju</i> = 6 ocorrências com [w]	100,000 * [v]
<i>colégio</i> = 4 ocorrências com [ʒ]	81,000 MAX I/O
	67,000 * [ʒ]
	52,000 *σ [w],
	33,000 IDENT [- anterior]
	0,000 IDENT [+ sonoro]
	-48,000 IDENT [+ consonantal]

Quadro 5 – Dados da 5ª. coleta e hierarquia de restrições

A hierarquia dada pelo programa evidencia a presença das restrições \* [ʃ], \* [v] e MAX I/O ranqueadas mais acima, com valores de peso altos. Isso evidencia o porquê de não haver, nos dados relativos a esta coleta, a produção da fricativa anterior desvozeada e do segmento fricativo [v], bem como o apagamento do segmento. O fato de a criança aplicar como principal estratégia de reparo a produção do glide posterior deve-se à posição de IDENT [+ consonantal], a qual se encontra na última posição da hierarquia, com valor de peso baixo. A restrição

de fidelidade que proíbe a produção do segmento-alvo encontra-se na quarta posição da escala, sendo seguida pela que prediz a proibição do glide em posição de onset silábico. Isso justifica o porquê de haver maior produção do glide, seguido da produção de [ʒ]: \*σ [w] encontra-se abaixo de \*[ʒ] na hierarquia. Deve-se salientar, no entanto, que seria esperado – pelo o que aponta a literatura da área acerca da proximidade dos valores das restrições para que a variação possa ocorrer – que os valores das restrições \*[ʒ] e \*σ [w] apresentassem uma diferença de, no máximo, 10,000, não os 15,000 constatados. Tal fato pode acenar para a possibilidade de se repensar os valores máximos entre duas restrições postulados para as formas variáveis.

O quadro 6 mostra os dados relativos à sexta coleta.

DADOS DO SUJEITO M.	HIERARQUIA DADA PELO PROGRAMA
<i>João</i> = 1 ocorrência com [w]	100,000 *[j]
<i>Juju</i> = 8 ocorrências com [v]	84,000 MAX I/O
<i>jogar</i> = 2 ocorrências com [ʒ]	72,362 *σ [w],
<i>colégio</i> = 1 ocorrência com [ʒ]	72,196 *[v]
	71,442 *[ʒ]
	0,000 IDENT [+ sonoro]
	-27,638 IDENT [+ consonantal]
	-27,804 IDENT [- anterior]

Quadro 6 – Dados da 6<sup>a</sup>. coleta e hierarquia de restrições

Com relação aos dados da sexta coleta, o ordenamento gerado pelo programa parece refletir a variabilidade de estratégias aplicadas pelo sujeito. A criança produz, na maior parte das vezes, o segmento [v], o que justifica que a restrição IDENT [- anterior] esteja na última posição da hierarquia. A restrição \*[j] é a que se encontra ranqueada mais acima, com valor de peso máximo, o que garante a não-produção da fricativa anterior desvozeada nesta coleta. Estando MAX I/O ranqueada logo abaixo, há a garantia de que não há apagamento dentre

os dados. Na sequência, estão as restrições de marcação que proíbem a produção do glide em onset silábico, de [v] e de [ʒ]. É interessante observarmos, nesse caso, os pesos dessas restrições –  $*\sigma$  [w] = 72,362;  $*[v]$  = 72,196;  $*[ʒ]$  = 71,442 –, os quais, em virtude de apresentarem menos de 10,000 pontos de diferença, apontam para a presença de formas variáveis. Isso justifica a coexistência de [w], [v] e [ʒ] nos dados do aprendiz nesta coleta.

Feitas as análises referentes às seis coletas, por fim, realizamos o segundo procedimento, visando a testar o algoritmo. Para isso, fizemos mais cinco aplicações do GLA. Diferentemente das aplicações até aqui descritas, não rodamos os dados com a produção do sujeito M. marcadas como ótimas, mas sim com as produções corretas esperadas no português brasileiro. O programa foi alimentado com o mesmo conjunto de restrições, mesmos pesos, ruído e números de testagens das aplicações para os dados de M.. Rodamos as restrições para cada item lexical utilizado no estudo e fizemos uma rotação que contemplou todos os itens simultaneamente. Os ordenamentos obtidos são apresentados no quadro 7.

Todos os itens lexicais	Colégio	João
100,000 $*[j]$	100,000 $*[j]$	00,000 $*[j]$
100,000 $*[v]$	100,000 $*[v]$	100,000 $*[v]$
100,000 $*\sigma$ [w],	100,000 $*\sigma$ [w],	100,000 $*\sigma$ [w],
58,000 MAX I/O	60,000 MAX I/O	58,442 MAX I/O
42,000 $*[ʒ]$	40,000 $*[ʒ]$	41,558 $*[ʒ]$
6,000 IDENT [- anterior]	0,000 IDENT [+ sonoro]	0,000 IDENT [+ sonoro]
0,000 IDENT [+ sonoro]	0,000 IDENT [- anterior]	0,000 IDENT [- anterior]
0,000 IDENT [+ consonantal]	0,000 IDENT [+ consonantal]	0,000 IDENT [+ consonantal]
Jogar	Juju	
100,000 $*[j]$	00,000 $*[j]$	
100,000 $*[v]$	100,000 $*[v]$	
100,000 $*\sigma$ [w],	100,000 $*\sigma$ [w],	
59,080 MAX I/O	57,000 MAX I/O	
40,920 $*[ʒ]$	57,000 IDENT [- anterior]	
0,000 IDENT [+ sonoro]	43,000 $*[ʒ]$	
0,000 IDENT [- anterior]	0,000 IDENT [+ sonoro]	
0,000 IDENT [+ consonantal]	0,000 IDENT [+ consonantal]	

Quadro 7 – Segundo procedimento: hierarquias de restrições



Podemos perceber que o ordenamento atingido para todos os itens lexicais, com uma pequena mudança de valores, é o mesmo, exceto em *Juju*. Enquanto para os outros itens lexicais \*[3] domina IDENT [- anterior], para *Juju*, IDENT [- anterior] domina \*[3]. Importante referir que, nos casos de substituição para esse item, o sujeito M. produzia um [v], violando, justamente, IDENT [- anterior]. Essa divergência pode, também, servir para explicar a validade do conjunto de restrições propostos para todos os itens lexicais no que concerne o segmento [3], pois as violações de [v] em *Juju*, junto aos outros itens lexicais, fazem com que a restrição IDENT [-anterior] receba valor 6,000 e fique, desta forma, em posição de dominância em relação a IDENT [+ sonoro] e IDENT [+consonantal], o que não acontece na análise dos itens *João*, *Colégio* e *Jogar* isoladamente.

## 6. Conclusões

Com relação ao primeiro procedimento realizado via algoritmo, vemos que os valores de restrições e os ordenamentos propostos pelo GLA para cada coleta traduzem os dados do sujeito M., tanto no que concerne às produções efetivas como à variação na produção do segmento aqui abordado, corroborando, assim, os achados de Bonilha (2007).

Além disso, por meio do segundo procedimento – testagem do aprendiz virtual –, podemos comprovar, também, que a gama de restrições adotada neste estudo dá conta dos dados de M., pois todas são passíveis de violação em alguma das produções realizadas pela criança. As restrições são, portanto, ativas e relevantes para uma análise da aquisição de /3/ via TO.



### Referências bibliográficas

BONILHA, G. *Aquisição fonológica do português brasileiro: uma abordagem conexionista da teoria da Otimidade*. Tese (Doutorado em Letras) – Faculdade de Letras, PUCRS, Porto Alegre, 2004.

\_\_\_\_\_. Contribuições do paradigma conexionista à Teoria da Otimidade. In: POERSCH, José Marcelino. ROSSA, Adriana Angelim (Orgs.). *Processamento da Linguagem e Conexionismo*. Santa Cruz do Sul: EDUNISC, 2007.

BOERSMA, P; HAYES, B. *Empirical Tests of the Gradual Learning Algorithm*. *Linguistic Inquiry* 32, 45-86, 2001.

COETZEE, A; PATER, J. The place of variation in phonological theory. In: Riggle and Yu (eds), *The handbook of phonological theory*, Goldsmith, 2008.

DEL RÉ, Alessandra. A pesquisa em aquisição da linguagem: teoria e prática. In: \_\_\_\_\_. *Aquisição da linguagem: uma abordagem psicolinguística*. São Paulo: Contexto, 2006.

HAYES, Bruce; TESAR, Bruce; ZURAW, Kie. *OTSoft 2.1*, software package. Disponível em: <<http://www.linguistics.ucla.edu/people/hayes/otsoft/>>. 2003.

McCARTHY, John; PRINCE, Alan S. *Prosodic Morphology I: Constraint Interaction and Satisfaction*. New Brunswick: Rutgers University Center for Cognitive Science, 1993.

OLIVEIRA, Carolina Cardoso. *Aquisição das fricativas /f/, /v/, /S/ e /Z/ do Português Brasileiro*. 2002. Dissertação (Mestrado em Letras) – Faculdade de Letras, PUCRS, Porto Alegre, 2002.

\_\_\_\_\_. Sobre a aquisição das fricativas. In: LAMPRECHT, Regina Ritter (Org.). *Aquisição fonológica do português: perfil de desenvolvimento e subsídios para terapia*. Porto Alegre: Artmed, 2004.

POERSCH, M. *A new paradigm for learning language: Connectionist artificial intelligence*. *Linguagem & Ensino*, vol. 8, No. 1, 2005 (161-183)

\_\_\_\_\_. Simulações conexionistas: a inteligência artificial moderna. In: POERSCH, José Marcelino. ROSSA, Adriana Angelim (Orgs.). *Processamento da Linguagem e Conexionismo*. Santa Cruz do Sul: EDUNISC, 2007.

PRINCE, Alan, SMOLENSKY, Paul. *Optimality Theory - Constraint Interaction in Generative Grammar*. RuCCs Technical report 2, 1993.





RUMELHART, D. E., MCCLELLAND, J. L. On learning the past tense of English verbs. In: MCCLELLAND, J.D, RUMELHART, D. E. *Parallel distributed processing: explorations in the microstructure of cognition: psychological and biological models*. v. 2. Cambridge: MIT, 1986.

TESAR, B; SMOLENSKY, P. *Learnability in Optimality Theory (long version)*. ROA - 156, 1996. [<http://rucss.rutgers.edu/roa.html>]

\_\_\_\_\_. Learnability in Optimality Theory. **Linguistic Inquiry** 29, p. 229-68, 1998.

\_\_\_\_\_. *The learnability of Optimality Theory*. Cambridge, MA: MIT Press, 2000.

## RESUMO

O presente trabalho apresenta os achados obtidos a partir da rodagem de dados empíricos em um programa computacional. Trata-se do OT Soft 2.3, desenvolvido por Hayes, Tesar e Zuraw (2003), pensado para facilitar as análises via Teoria da Otimidade com o uso de algoritmos da aprendizagem. No caso desse estudo, verificamos a eficácia do programa, ao ser alimentado com os dados do sujeito M., correspondentes a um período de seis coletas, na faixa etária de 1:9 a 2:2. Para isso, o *Gradual Learning Algorithm* (GLA) foi escolhido. A análise dos dados evidenciou que o programa se mostrou relevante, haja vista que foi capaz de fornecer as hierarquias de restrições correspondentes aos estágios de aquisição de M., bem como revelou a hierarquia do adulto com relação ao segmento /ʒ/ no português brasileiro.

**PALAVRAS-CHAVE:** Teoria da Otimidade, Algoritmo de Aquisição Gradual, aquisição segmental

## ABSTRACT

The present work presents the findings obtained from the application of empirical data in a computational program, the OT Soft 2.3. The program was developed by Tesar, Hayes e Zuraw (2003) to facilitate linguistic analysis via Optimality Theory with the use of learning algorithms. In regards this study, we verified the program efficacy as it was fed with Subject's M. data which was accompanied through six data collections from age 1:9 to 2:2. In order to do so, the Gradual Learning Algorithm (GLA) was chosen. The data analysis evidenced that the program showed to be relevant, as it was capable to provide constraints hierarchies correspondent to the acquisition stages of M., revealed as well the adult hierarchy regarding the segment /ʒ/ in Brazilian Portuguese.

**KEY-WORDS:** Optimality Theory, Gradual Learning Algorithm, segmental acquisition