

**A PRODUÇÃO DA VOGAL EPENTÉTICA NO  
PORTUGUÊS BRASILEIRO POR COLOMBIANOS:  
UMA ANÁLISE VIA GRAMÁTICA HARMÔNICA**  
THE PRODUCTION OF THE EPENTHETIC VOWEL IN BRAZILIAN  
PORTUGUESE BY COLOMBIAN LEARNERS:  
AN ANALYSIS VIA HARMONIC GRAMMAR

Roberta Quintanilha Azevedo<sup>1</sup>

Carmen Lúcia Matzenauer<sup>2</sup>

Ubiratã Kickhöfel Alves<sup>3</sup>

**Resumo:** *Este artigo tem por objetivo analisar a produção de segmentos plosivos em codas mediais de palavras, por colombianos, aprendizes de Português Brasileiro (PB), que são susceptíveis à produção de segmentos epentéticos. Para a constituição do corpus, foi proposto um instrumento composto de palavras cognatas do Português e Espanhol para a leitura em frases veículo, gravadas no Programa AUDACITY 1.3.5 (ANSI). Foram escolhidos, para participar como sujeitos da pesquisa, quatro estudantes colombianos que permaneceram no Brasil por um período médio de seis meses. Foi confirmada a hipótese de que, mesmo em palavras cognatas nas duas línguas analisadas, há diferente tratamento às plosivas em coda, dependendo do sistema linguístico que está sendo utilizado pelo falante de Espanhol aprendiz de PB. Tal fato foi captado, à luz do algoritmo de aprendizagem vinculado ao modelo formal da Gramática Harmônica - HG (LEGENDRE, MIYATA & SMOLENSKY, 1990; SMOLENSKY & LEGENDRE, 2006).*

**Palavras-chave:** *Epêntese Vocálica, Gramática Harmônica, Aquisição de L2.*

**Abstract:** *This study aims to analyze the production, by Colombian Learners of Brazilian Portuguese (BP), of stop segments in word-mid codas, liable for the production of epenthetic segments. In order to build the corpus, an instrument composed of Portuguese and Spanish cognate words was proposed, so that vehicle phrases could be recorded in AUDACITY 1.3.5 (ANSI) software. Four Colombian students, who remained in Brazil for a period of six months, were chosen to take part in the investigation. We confirmed the hypothesis that, even in cognate words, there is a different treatment to plosives in coda*

---

1 Doutoranda do Programa de Pós-Graduação em Letras da UCPEL.

2 Professora do Programa de Pós-Graduação em Letras da UCPEL, pesquisadora do CNPq.

3 Professor do Instituto de Letras da UFRGS, pesquisador do CNPq.

*position in both languages, depending on the language system which is being used by the speaker. This finding was formalized, according to Harmonic Grammar – HG (LEGENDRE, MIYATA & SMOLENSKY, 1990; SMOLENSKY & LEGENDRE, 2006).*

**Keywords:** *Vowel Epenthesis, Harmonic Grammar, L2 Acquisition .*

## **Introdução**

A ocorrência de epêntese nas várias línguas do mundo é fenômeno conhecido, mas ainda não esgotado nas discussões da área da fonologia. Para este trabalho, trataremos a epêntese como um processo fonológico caracterizado pela inserção de um som vocálico no *output* das palavras, som esse não presente no *input*. Tal fenômeno vai servir para tornar menos marcada a estrutura silábica de uma dada língua.

Em Collischonn (1997), é proposto que, na silabação do Português Brasileiro (PB), uma consoante não apta em posição silábica de ataque ou coda – consoante perdida – não seria ligada a nenhum nó silábico, o que desencadearia a criação de uma outra sílaba, desprovida temporariamente de núcleo vocálico. Esse seria o contexto perfeito para a inserção de um núcleo vocálico, como podemos perceber em palavras como *et[i]nia* e *ac[i]ne*.

Procuramos neste trabalho, portanto, analisar a produção da vogal epentética [i] em um contexto de plosivas em codas mediais de palavras, por parte de falantes nativos de Espanhol (colombianos), aprendizes de Português Brasileiro, à luz do modelo formal da Gramática Harmônica (HG). Essas plosivas em coda formam uma sequência com a consoante *onset* da sílaba seguinte, a qual se constitui no que se denomina “contato silábico” (sequência heterossilábica). Segundo Murray & Vennemann (1983), em qualquer contato silábico na sequência  $C_1 \ \$ \ C_2$  há preferência de que  $C_1$  exceda  $C_2$  em sonoridade. Essa distância de sonoridade entre  $C_1$  e  $C_2$  é estudada por Gouskova (2004), conforme discussão mostrada na seção 2.1 deste artigo. Uma vez que estudos recentes, por nós desenvolvidos (cf. Azevedo, Alves e Matzenauer, 2012), já demonstraram a capacidade da OT Estocástica em formalizar os efeitos do contato silábico na aquisição da L2, interessa-nos verificar, neste estudo, se o modelo da Gramática Harmônica se mostra também capaz de propiciar esta formalização, bem como discutir os mecanismos formais a partir dos quais o modelo operará, frente a tal tarefa.

Fizeram parte da pesquisa quatro estudantes colombianos que estiveram na cidade de Pelotas (RS/Brasil), realizando intercâmbio por um período médio de seis meses. Eles foram submetidos a uma tarefa de leitura de sequências recorrentes que continham o contexto fonológico analisado, que facilitava o aparecimento do fenômeno da epêntese no Português Brasileiro. Após a verificação acústica, os dados gravados foram dispostos para análise através do modelo da Gramática Harmônica, que consideramos ser um instrumento teórico relevante para formalizar o fenômeno levantado.

A fim de atender ao nosso objetivo geral de, à luz da HG, analisar a produção da vogal epentética em um contexto de plosivas em codas mediais de palavras, por parte de aprendizes de Português Brasileiro, que são falantes nativos de Espanhol (colombianos), apresentamos as Questões Norteadoras a seguir:

- 1 – Qual(is) fator(es) linguístico(s) contribui(em) para a ocorrência da epêntese diante das sequências de segmentos /pt, pn, kn, kt, tn/, em codas mediais de palavras no Português Brasileiro, por falantes nativos do Espanhol Colombiano?
- 2 – Que organização de restrições dá conta das formas de *output* produzidas pelos aprendizes, à luz da Gramática Harmônica?
- 3 – O modelo da Gramática Harmônica consegue formalizar com adequação o processo de aquisição de L2?

Nas próximas seções, apresentaremos os pressupostos teóricos que fornecem subsídios e suporte ao desenvolvimento do trabalho, e realizaremos a análise desse fenômeno com base no modelo teórico da Gramática Harmônica.

## ***1 Fundamentação Teórica***

### **1.1 Pressupostos da Gramática Harmônica e seu Algoritmo de Aprendizagem Gradual (HG-GLA)**

O modelo da Gramática Harmônica, assim como a Teoria da Otimidade Estocástica, trabalha com as restrições com um caráter numérico. Além disso, como ocorre desde a Teoria da Otimidade considerada clássica - conforme explicam Jesney & Tessier (2007) -, vai apresentar os módulos: GEN (*Generator*), CON (*Constraint*) e EVAL (*Evaluator*).

Entretanto, tal modelo diferencia-se, fundamentalmente, na avaliação do melhor *output* feita por EVAL, já que vai considerar, na escolha do candidato ótimo, o caráter cumulativo das violações de cada candidato,

considerando os pesos numéricos atribuídos às restrições universais. Será escolhido como ótimo o candidato cujo valor de Harmonia for mais alto.

Basicamente, o princípio da HG é determinar para cada *input* um *output*, através da resolução de um conflito entre restrições, com a avaliação cumulativa dos valores atribuídos a essas restrições, conforme podemos ver no *tableau* a seguir.

Output/Restrições	(40) Restrição A	(30) Restrição B	(15) Restrição C	H
☞ [Output 1]	*		*	-55
[Output 2]		***		-90

Tableau 1 - Exemplo de tableau no modelo da Gramática Harmônica

O *tableau* (1) mostra que o módulo da gramática, EVAL, é dado por um valor de Harmonia (H), que é obtido ao multiplicarmos cada violação do candidato (\*), que apresenta valor negativo, pelo valor de Ponto de Seleção (que corresponde ao valor da Restrição (R) que vem a ser assumido a cada momento de produção linguística). Após tal operação, devem ser somados todos os resultados dessas multiplicações, referentes ao candidato.

A noção de cumulatividade das restrições violadas por um dado candidato, para a avaliação do candidato ótimo, no modelo da HG, vai distanciar tal teoria do modelo que opera sob uma noção de dominância estrita na avaliação do candidato ótimo, conforme ocorre na Teoria da Otimidade Estocástica<sup>4</sup>.

O algoritmo de aprendizagem nesse modelo, denominado de HG-GLA, é disponibilizado no *software* PRAAT (BOERSMA & WEENINK, 2009), e serve para expressar o processo de aquisição e as possíveis variações do *output*. Tal algoritmo vai apresentar um valor de ruído, que é somado ao peso das restrições.

Cabe ressaltarmos, ainda, que tal algoritmo é do tipo *error-driven* (guiado pelo erro), o que permite que as restrições se movimentem até o sujeito chegar ao seu alvo. Dessa forma, o algoritmo vai alterar o valor numérico das restrições quando o *output* da sua gramática se mostrar diferente do padrão encontrado na linguagem ambiente (erro). Esse índice de alteração se dá em função de um valor de plasticidade<sup>5</sup> que é adotado pelo algoritmo. Assim, o algoritmo vai diminuir (demover) o valor numérico

4 Na OT Estocástica, ao contrário do que ocorre na HG, há a dominância estrita entre as restrições que constituem a hierarquia que responde pela escolha de cada *output*.

5 Plasticidade é um valor numérico através do qual o algoritmo vai ajustar o *ranking* das restrições. Esse valor é fornecido durante a simulação computacional, através da função *Learn* do PRAAT.

das restrições que são violadas pelas formas encontradas de produção e vai aumentar (promover) o valor das restrições violadas pelo *output* “errado”, na hierarquia de Restrições vigente.

Logo, cada restrição apresentará, na verdade, dois valores numéricos: o valor que corresponde ao ponto central da faixa de valores - valor central ou valor de ranqueamento/*ranking value* (VC) – e o ponto de seleção (PS), que é um valor assumido dentro da faixa de valores, correspondente ao valor exibido no momento de avaliação dos candidatos.

É graças a esses valores centrais que o algoritmo vai representar se os *outputs* são variáveis ou categóricos. Valores centrais bem afastados entre restrições, ou seja, valores com distância superior a 10 pontos (ex.: Restrição A → VC = 50 / PS = Qualquer valor de 45 a 55, Restrição B → VC = 89 / PS = Qualquer valor de 84 a 94) representam o resultado categórico, pois diferentes momentos de produção linguística não causarão cruzamento na faixa de valores destas restrições. Em contrapartida, um resultado da gramática fornecida pelo PRAAT, com valores centrais das restrições com distância inferior a 10 pontos (ex.: Restrição A → VC = 50 / PS = Qualquer valor de 45 a 55, Restrição B → VC = 58 / PS = Qualquer valor de 53 a 63) estaria sujeito a cruzamento da faixa de valores a serem assumidos por essas restrições, o que poderia alterar o seu ordenamento e, conseqüentemente, o candidato ótimo, caracterizando a variação nos *outputs*, conforme veremos nas simulações realizadas na Seção 2, da Análise dos Dados.

Na simulação, optamos, ainda, pela especificação do algoritmo denominada *Linear OT*. Já que os valores de Harmonia podem ter pesos negativos, caso não utilizássemos o recurso *Linear OT*, poderíamos incorrer no problema denominado de “*harmonic bounding*”<sup>6</sup>. De fato, a soma para que se chegue ao valor de Harmonia, após multiplicarmos as violações incorridas com os pesos das restrições, pode oferecer um sério engano, conforme podemos verificar no *tableau 2*.

	R1 (1.0)	R2 (-2.0)	H
☞ Output1	-1	-2	+3
Output2		-1	+2

Tableau 2 - O problema harmonic bounding exemplificado em tableau

O *Tableau 2* mostra que uma restrição com peso negativo acabou por erroneamente dar ao *output 1* o status de candidato ótimo, quando, na verdade, de-

6 Lembramos que *harmonic bounding* é um problema da *OT Standard*, não sendo exclusivo da versão Estocástica desse modelo teórico. Tal fenômeno é aqui referido pelo fato de ser, na utilização do algoritmo vinculado ao *software Praat*, o que motiva a escolha da opção “*Linear OT*”, em lugar da opção “*HG*”.

veria ser o *output 2* pois, em função de a restrição ter peso negativo, o resultado da multiplicação com o índice de violações (asterisco) será positivo. Contudo, graças à seleção da função *Linear OT* no PRAAT, pode-se impedir tal problema, visto que o recurso em questão impede que pesos negativos exerçam papel no cálculo do valor da Harmonia. Conforme será visto na análise do presente trabalho, tal recurso será de grande pertinência em nossa análise.

No que diz respeito à aplicação do modelo da HG em dados de aquisição de L2, pequeno, ainda, é o número de estudos desenvolvidos em nosso país. Em Alves (2010), temos acesso a uma discussão acerca do processo de aquisição da linguagem à luz dos modelos da HG e da OT Estocástica e a necessidade do uso de restrições conjuntas. Através das simulações realizadas em ambos os modelos, foi verificada a capacidade de convergência na HG de forma mais econômica que na OT Estocástica, visto seu caráter cumulativo vir a diminuir, e, no caso do artigo em questão, até dispensar, o uso de restrições conjuntas na análise, concluindo que os modelos podem estar diferenciados não apenas na forma de avaliar o candidato ótimo (módulo EVAL), mas no conjunto de restrições, módulo CON.

Considerando a pesquisa de Azevedo, Alves e Matzenauer (2012) sobre a epêntese no Português como língua estrangeira por falantes nativos do Espanhol no modelo da OT Estocástica, para os mesmos dados ora utilizados, o presente trabalho poderá vir a prover maiores insumos para a discussão teórica proposta em Alves (2010), no que se refere às possíveis diferenças no processo de aquisição entre os modelos da OT Estocástica e HG. Além disso, julgamos pertinente a tarefa de propor uma análise, à luz do Modelo da Gramática Harmônica, de dados já previamente analisados, de maneira bem-sucedida, a partir da Teoria da Otimidade Estocástica. Assim, além das respostas às questões norteadoras propostas, é possível uma consideração mais ampla na conclusão, sobre a eficiência do modelo de análise cujo funcionamento estamos investigando neste trabalho.

## ***2 A Análise dos Dados***

### **2.1 Dados da Análise**

As tabelas a seguir representam o tratamento estatístico dos dados de produção de palavras cognatas no Português Brasileiro e no Espanhol Colombiano, ou seja, palavras pertencentes ao léxico das duas línguas, constituídas pelas sequências de plosivas surdas em codas mediais [p.n, k.n, t.n, p.t, k.t], dispostas em frases simples no Espanhol Colombiano e no

PB. As palavras utilizadas na análise para a coleta dos dados foram: *raptto*, *corrupto*, *estricnina*, *tecnicismo*, *Vietnam*, *etnia*, *cacto*, *convicto*, *apneia* e *hipnose* (para detalhes da coleta de dados e escolha das palavras cognatas, ver Azevedo, 2011).

Consoante em coda	Epêntese	Apagamento
Consoante em coda [k.n, p.n, t.n] 100%	Epêntese [ki.n, pi.n, ti.n] 0	Apagamento total [ø.n] 0
Consoante em coda [k.t, p.t] 100%	Epêntese [ki.t, pi.t] 0	Apagamento total [ø.t] 0

Tabela 1 – Produção das sequências [p.n, k.n, t.n, p.t, k.t] no Espanhol Colombiano [Fig. 1]

Consoante em coda	Epêntese	Apagamento
Consoante em coda [k.n, p.n, t.n] 86,52%	Epêntese [ki.n, pi.n, ti.n] 13,48%	Apagamento total [ø.n] 0
Consoante em coda [k.t, p.t] 98,96%	Epêntese [ki.t, pi.t] 1,04%	Apagamento total [ø.t] 0

Tabela 2 - Produção das sequências [p.n, k.n, t.n, p.t, k.t] no Português Brasileiro [Fig. 2]

Os dados apresentados nas Tabelas 1 e 2 demonstram que palavras como *etnia*, composta de consoante plosiva em coda, seguida de consoante nasal ou plosiva no *onset* da sílaba seguinte, nas produções em frases no Espanhol - Tabela 1 - tiveram 100% de produção da plosiva em coda, sem apagamento ou epêntese.

Já no Português Brasileiro - Tabela 2 - vemos que houve processo de epêntese, e que o total de ocorrência de vogal epentética encontrada é maior em um contato de plosiva em coda seguida de nasal (13,48%), em comparação com um *plateau* de sonoridade entre plosivas (1,04%) como, por exemplo, ocorre na palavra *corrupto*. Assim, constatamos que o contato silábico estaria influenciando na ocorrência de epêntese pelos aprendizes de Português Brasileiro. Logo, a discussão das restrições que segue deverá refletir essa ideia.

## 2.2 Formalização das restrições

As restrições podem pertencer a duas famílias: Fidelidade e Marcação. As restrições de fidelidade devem buscar a preservação do *input*. No caso de plosiva em posição de coda medial de palavra, argumentamos que as

restrições de fidelidade DEP - oposição à ocorrência de epêntese - e MAX - oposição à ocorrência de apagamento - (McCARTHY E PRINCE, 1995) conseguem dar conta dos padrões encontrados.

A decisão para a escolha das restrições de marcação partiu da preocupação em caracterizarmos o nível prosódico com que estamos trabalhando, especificamente em um contexto de plosivas em posição de coda medial. O Mecanismo de Alinhamento Harmônico (HA), de Prince & Smolensky (1993), foi escolhido então, pois esse mecanismo propiciará um conjunto de restrições que pode mostrar quais segmentos se apresentam mais aptos a ocorrer em coda silábica a partir de uma escala linguística.

Retomando a constatação de que a ocorrência da vogal epentética no PB, por parte dos sujeitos, falantes nativos de Espanhol Colombiano, é consequência de contato silábico, levantamos, também, restrições que vão servir para dar conta da influência da sonoridade causada pelo *onset* da próxima sílaba na inserção de epêntese nos nossos dados. Assim, baseamo-nos no Mecanismo de Alinhamento Relacional de Gouskova (2004), que vai combinar a escala de harmonia referente à sonoridade de coda com a de *onset*, determinando quais distâncias são mais marcadas entre os segmentos heterossilábicos que propomos para este estudo.

O Mecanismo de Alinhamento Relacional de Gouskova (2004) transforma em índices numéricos as diferenças de sonoridade entre coda e *onset* da sílaba seguinte, uma vez que, em termos de contato silábico, é melhor que ocorra uma redução do valor do segmento da coda para o *onset* da próxima sílaba. Basicamente, é um esquema de Distância de Sonoridade entre duas consoantes (\*DISTANCE), onde \*Dist (0) >> \*Dist (-1) >> \*Dist (-2), já que, nesse caso, quanto maior a queda de sonoridade, mais harmônica é a distância. Em outras palavras, é melhor termos um encontro [k.t] do que um [k.n], ou um encontro [p.t] a um [p.n].

A figura a seguir representa o Mecanismo de Alinhamento Relacional de Gouskova (2004):

Mecanismo de Alinhamento Relacional – com base em Gouskova (2004)  
W = glídes N = nasais L = líquidas T = obstruintes

w.t	w.n l.t	w.l l.n n.t	w.w l.l n.n t.t	l.w n.l t.n	n.w t.l	t.w
-3	-2	-1	0	+1	+2	+3

Fig. 3

Cada coluna representa um estrato e as sequências que ocupam o mesmo estrato terão valor igual de distância de sonoridade (Dist). Os valores com sinal negativo (-3, -2 e -1) indicam que está havendo queda de sonoridade entre a coda e o *onset* da sílaba seguinte, representando as sequências

mais harmônicas como, por exemplo, em w.t (glide em coda e obstruinte no *onset* da sílaba seguinte – [saw. t<sub>o</sub>]). O sinal positivo, ao contrário, representa um aumento de sonoridade (+1, +2 e +3), como em [ak.ne], representando os valores mais marcados. Quanto ao estrato com valor zero, que separa os valores positivos dos negativos, este indica um platô de sonoridade, ou seja, apresenta a possibilidade de termos na coda um segmento com a mesma sonoridade do *onset* da sílaba seguinte, como é o caso de [kak.to]. Dessa forma, os valores dos estratos mais à esquerda são mais harmônicos ou menos marcados com relação aos valores mais à direita.

Gouskova (2004) não sugere relações de estringência, mas um *ranking* fixo entre tais restrições. Em nossa análise, pensamos, para esta e as demais restrições utilizadas, numa organização que trabalhe esta relação em estringência, para que possamos demonstrar as relações de marcação existentes entre os membros de uma escala, formalizando a ordem de aquisição de cada um desses membros.

Conforme podemos ver no esquema de Gouskova, a escala utilizada (glides, nasais, líquidas e obstruintes) não separa as consoantes plosivas das fricativas. Assim, para que pudéssemos perceber as diferenças entre as sequências heterossilábicas [p.t, k.t] e [p.n, k.n, t.n] com relação às sequências com fricativa em coda, que também são obstruintes, mas não são responsáveis por epêntese no PB, acrescentamos as restrições conjuntas e restrições de proibição a plosivas em coda, necessárias para se oporem apenas às sequências com plosivas em coda ( $\text{Dist } \{1, 0\} \& \text{*}\{\text{stop}\}_{\text{coda}}$  e  $\text{*}\text{Dist } \{1\} \& \text{*}\{\text{stop}\}_{\text{coda}}$ ), que apresentam comportamento diferente das demais possíveis obstruintes em coda.

Em suma, a partir dos processos de Alinhamento Harmônico e Relacional, obtém-se um grande número de restrições, para expressar as escalas primitivas que as originam. Uma vez que a HG tem, por premissa, considerar, na avaliação do candidato ótimo, todas as restrições incluídas na análise, considera-se, nesse sentido, que a formação de restrições estringentes advindas de escalas primitivas pode representar um desafio adicional para o modelo de HG. Assim, julgamos necessário ver como o modelo dará conta de todas essas restrições. As restrições de marcação e de fidelidade formalizadas para a análise são apresentadas a seguir:

#### RESTRIÇÕES DE FIDELIDADE

MAX e DEP

#### RESTRIÇÕES DE MARCAÇÃO

Mecanismo de Alinhamento Relacional em estringência

$\text{*}\text{Dist } \{+3\}$ ,  $\text{*}\text{Dist } \{+3, +2\}$ ,  $\text{*}\text{Dist } \{+3, +2, +1\}$ ,  $\text{*}\text{Dist } \{+3, +2, +1, 0\}$ ,  $\text{*}\text{Dist } \{+3, +2, +1, 0, -1\}$ ,  $\text{*}\text{Dist } \{+3, +2, +1, 0, -1, -2\}$ ,  $\text{*}\text{Dist } \{+3, +2, +1, 0, -1, -2, -3\}$

### Mecanismo de Alinhamento Harmônico em estringência

\*{stop}<sub>coda</sub>, \*{stop, fric}<sub>coda</sub>, \*{stop, fric, nas}<sub>coda</sub>, \*{stop, fric, nas, liq}<sub>coda</sub>,  
\*{stop, fric, nas, liq, glide}<sub>coda</sub>

### Restrições Conjuntas

\*Dist {1, 0} & \*{stop}<sub>coda</sub> e \*Dist {1} & \*{stop}<sub>coda</sub>.

## 2.3 A simulação computacional dos dados de aquisição

Para formalizarmos o estágio desenvolvimental dos aprendizes colombianos em direção ao alvo (aquisição do PB), com as produções da vogal epentética ocorrendo variavelmente (representado na Simulação 3), foi necessário, primeiramente, verificar se o algoritmo seria capaz de convergir na L1 dos aprendizes, que caracteriza um caso de produção categórica, de 100% de plosiva em coda - Espanhol Colombiano - (representado na Simulação 1), bem como na língua estrangeira, que apresenta 100% de produção da epêntese vocálica - Português Brasileiro - (representado na Simulação 2). Isso porque, ao considerarmos que o primeiro estágio de aquisição da L2 é o sistema da L1, as simulações com o algoritmo de aprendizagem precisam expressar, primeiramente, a aquisição da língua materna, conforme veremos a seguir.

### 2.3.1 Simulação 1 - Aquisição do Espanhol Colombiano (L1)

a) A aquisição do Espanhol Colombiano (L1) via Gramática Harmônica

Estamos trabalhando com 7 sequências de segmentos consonantais (“t.t” – obstruente + obstruente, “t.n” – obstruente + nasal, “n.t” – nasal + obstruente, “l.t” – líquida + obstruente, “w.t” – glide + obstruente, [s.t], [s.n]) que representam as possibilidades de contato silábico entre os segmentos que compõem a escala de sonoridade de Clements (1990) – glides (w), nasais (n), líquidas (l), obstruintes (t), com exceção da vogal.

Cada sequência que utilizamos representará uma distância de sonoridade dada pelo Mecanismo de Alinhamento Relacional de Gouskova (2004), com o cuidado de termos apenas sequências compostas de um segmento consonantal em coda e uma obstruente ou nasal em onset da sílaba seguinte. Somadas a estas sequências da escala “Dist”, ainda temos [s.t] e [s.n], que, dentre as sequências de obstruintes utilizadas (“t.t” – obstruente + obstruente, “t.n” – obstruente + nasal), têm comportamento diferente dos demais segmentos, pois, em posição de coda medial no PB, não são passíveis de epêntese.

Dando início à primeira simulação, alimentamos o sistema com dados que tratam da aquisição do Espanhol Colombiano, ou seja, que expressem, para todas as sequências de segmentos utilizadas, 100% da produção dos segmentos

em coda, sem epêntese ou apagamento, tal como deve ocorrer no Espanhol.

Como o Espanhol Colombiano é a língua materna dos nossos aprendizes de Português Brasileiro, o sistema é programado a responder a um estágio inicial de um bebê adquirindo as consoantes em coda no Espanhol Colombiano, conforme verificamos no conjunto de *tableaux* 1.

	<i>ranking value</i>	<i>disharmony</i>	<i>plasticity</i>
*Dist {+3, +2}	100.000	102.090	1.000000
*Dist {+3}	100.000	95.380	1.000000
Dep	88.626	92.441	1.000000
Max	89.792	88.833	1.000000
*Dist {1} & {stop}coda	39.061	40.793	1.000000
*{stop}coda	11.140	13.855	1.000000
*Dist {1,0} & {stop}coda	11.140	12.086	1.000000
*Dist {+3, ...+1}	9.423	9.048	1.000000
*{stop, fric}coda	-31.453	-29.687	1.000000
*Dist {+3, ... 0}	-31.453	-31.262	1.000000
*Dist {+3, ...-1}	-53.405	-51.027	1.000000
*{stop, ... nas}coda	-53.405	-55.718	1.000000
*{stop, ... liq}coda	-66.469	-63.629	1.000000
*Dist {+3, ...-2}	-66.469	-68.799	1.000000
*Dist {+3, ...-3}	-78.418	-76.188	1.000000
*{stop, ... glide}coda	-78.418	-77.640	1.000000

	*Dist{+3, +2}	*Dist{+3}	MAX	DEP	*Dist{1}&{stop}coda	*Dist{1,0}&{stop}coda	*Dist{+3, ...+1}	*{stop}coda	*{stop, fric}coda	*Dist{+3, ...0}	*{stop, ...nas}coda	*Dist{+3, ...-1}	*{stop, ...liq}coda	*Dist{+3, ...-2}	*{stop, ...glide}coda	*Dist{+3, ...-3}	
p.t/k.t					*	*	*	*	*	*	*	*	*	*	*	*	-25941
pit/kit				*													-92441
t			*														-88833
p.n/k.n/t.n					*	*	*	*	*	*	*	*	*	*	*	*	-75782
pi.n/ki.n/ti.n				*													-92441
n			*														-88833
n.t										*	*	*	*	*	*	*	0
nit				*													-92441
t			*														-88833
lt													*	*	*	*	0
lit				*													-92441
t			*														-88833
w.t															*	*	0
wit				*													-92441
t			*														-88833
s.t								*	*	*	*	*	*	*	*	*	0
sit				*													-92441
t			*														-88833
s.n						*	*	*	*	*	*	*	*	*	*	*	-9048
sin				*													-92441
n			*														-88833

Conjunto de Tableaux 1 - Simulação 1 HG / Aquisição do Espanhol Colombiano (0% de epêntese)

Os candidatos ótimos são aqueles que apresentam a articulação do segmento consonantal em coda, como em ra[p].to ou a[k].ne. Já que estamos

ainda trabalhando com *outputs* categóricos, além de apresentarem valores superiores aos dos demais candidatos, os vencedores mantiveram a distância necessária de 10 pontos entre os valores de Harmonia dos demais candidatos a *output*<sup>7</sup>. Por mais que, a cada avaliação, tenhamos valores diferentes de ponto de seleção para as restrições, tais diferenças não chegarão a causar variação entre os candidatos perdedores e os candidatos ótimos, dada a distância nos valores de Harmonia entre estes candidatos perdedores e os vencedores.

No que diz respeito ao ordenamento das restrições, operamos com as restrições de fidelidade DEP e MAX, e as restrições de marcação \*Dist{+3} e \*Dist{+3, +2} - que se referem a sequências inexistentes no Português Brasileiro e no Espanhol Colombiano - com os valores mais altos de ranqueamento.

É interessante observar que o algoritmo manteve as restrições de marcação referentes às sequências mais marcadas (\*Dist{+3,...+1}, \*Dist{1} & \*{stop}<sub>coda</sub>, \*Dist{1,0} & \*{stop}<sub>coda</sub> e \*{stop}<sub>coda</sub>) com valor positivo, em detrimento das demais restrições de marcação, que ficaram com valores negativos, sendo excluídas do cálculo do valor de H, de acordo com a função *Linear* OT do PRAAT, cujo funcionamento já foi aqui explicado. O algoritmo deu conta desta demanda ao excluir do cálculo as restrições de caráter mais geral pertencentes ao conjunto de restrições estridentes, mais baixas na hierarquia, facilitando a DEP e MAX a tarefa de virem a ter de assumir valores tão altos que, no cálculo da H, os mantivessem vencedores diante dos candidatos produzidos com a consoante em coda. Essa situação poderia estar denunciando também que as restrições que receberam valor negativo estão redundantes na análise, já que a formação destas restrições, ainda que, em princípio, pudesse ser considerada desnecessária, ocorreu a partir de mecanismos formais de formação de restrições com base em escalas primitivas, como o Alinhamento Harmônico e o Alinhamento Relacional. Maiores detalhes acerca de tal artifício, por parte do Algoritmo, serão discutidos ao final deste trabalho.

Dessa forma, em se tratando da aquisição da língua materna dos aprendizes, em um sistema estridente como o aqui apresentado, o algoritmo conseguiu convergir na HG para uma simulação que sugere *outputs* categóricos. DEP e MAX apresentaram valores de ponto de seleção que permitiram que

7 Conforme explicado na Seção 1.1, no modelo da Gramática Harmônica, o que vai definir se o candidato é categórico ou variável é a distância do valor de Harmonia (H) do candidato ótimo, com relação aos demais. Valores de Harmonia (H) com diferença igual ou superior a 10 pontos representam um candidato ótimo categórico; valores de Harmonia com diferença inferior a 10 pontos, representam os candidatos variáveis.

os candidatos que os violam atingissem valores de Harmonia superiores aos daqueles que violam as restrições de marcação (acima de 10 pontos), não permitindo, portanto, que houvesse variação nos candidatos ótimos.

### 2.3.2 Simulação 2 - Aquisição do Português Brasileiro (L2)

#### b) A aquisição do Português Brasileiro (L2) via Gramática Harmônica

A aquisição plena do PB vai ter, no estágio final de aquisição, a produção de 100% de epêntese após as consoantes plosivas em posição de coda. Como estamos trabalhando com a escala de Clements (1990), estamos tratando as plosivas em conjunto com as fricativas e africadas, sob a categoria das obstruintes (t). No PB, esta classificação, em princípio, não parece ser desejável, porque dentre as fricativas temos o segmento [s], perfeitamente possível em posição de coda (ex.: “cos.ta”), em comparação aos demais segmentos que ensejam o acréscimo da epêntese como correção a uma situação indesejada, como em “a.f[i].ta” ou “ca.[ki].to”.

Ao alimentarmos o sistema com os valores das restrições obtidos na simulação anterior (estágio final do Espanhol Colombiano e inicial para o Português Brasileiro), o algoritmo foi informado de que, no sistema-alvo, *inputs* com plosivas seguidas de outra obstruinte ou nasal devem ser epentetizados, e que *inputs* com a fricativa [s], bem como as sequências compostas de “n.t” (nasal em coda e obstruinte em *onset* da sílaba seguinte), “l.t” (líquida em coda e obstruinte em *onset*), “w.t” (glide em coda e obstruinte em *onset* da sílaba seguinte) devem manter-se sem epêntese.

O conjunto de *Tableaux 2* possibilita a verificação da resposta do algoritmo da HG para a aquisição plena do Português Brasileiro (L2).

	<i>ranking value</i>	<i>disharmony</i>	<i>plasticity</i>
*Dist {+3}	100.000	101.427	1.000000
*Dist {+3, +2}	100.000	98.707	1.000000
Max	89.792	87.900	1.000000
Dep	60.725	59.613	1.000000
*Dist {l} & {stop}coda	41.838	40.943	1.000000
*{stop}coda	39.041	40.482	1.000000
*Dist {l,0} & {stop}coda	39.041	38.616	1.000000
*Dist {+3, ...+1}	12.200	8.197	1.000000
*{stop, fric}coda	-3.552	-2.626	1.000000
*Dist {+3, ... 0}	-3.552	-4.345	1.000000
*Dist {+3, ...-1}	-25.504	-24.888	1.000000
*{stop, ... nas}coda	-25.504	-26.827	1.000000
*Dist {+3, ...-2}	-38.568	-37.087	1.000000
*{stop, ... liq}coda	-38.568	-39.476	1.000000
*{stop, ... glide}coda	-50.517	-51.347	1.000000
*Dist {+3, ...-3}	-50.517	-53.283	1.000000

	*Dist{+3}	*Dist{+3, +2}	MAX	DEP	*Dist{1}&[stop] coda	*[stop] coda	*Dist{1,0}&[stop] coda	*Dist{+3,...+1}	*[stop,fric] coda	*Dist{+3,...0}	*Dist{+3,...-1}	*[stop,...nas] coda	*Dist{+3,...-2}	*[stop,...lq] coda	*[stop,...glide] coda	*Dist{+3,...-3}	
p.t/k.t					*	*			*	*	*	*	*	*	*	*	-7909
☞ pit/ki.t			*														-5961
t				*													-8790
p.n/k.n/t.n					*	*	*	*	*	*	*	*	*	*	*	*	-12823
☞ pin/ki.n/ti.n			*														-5961
n				*													-8790
☞ n.t											*	*	*	*	*	*	
ni.t			*														-5961
t				*													-8790
☞ lt													*	*	*	*	
li.t			*														-5961
t				*													-8790
☞ w.t															*	*	
wi.t			*														-5961
t				*													-8790
☞ s.t								*	*	*	*	*	*	*	*	*	
si.t			*														-5961
t				*													-8790
☞ s.n							*	*	*	*	*	*	*	*	*	*	-819
si.n			*														-5961
n				*													-8790

Conjunto de Tableaux 2 - Simulação 2 HG / Aquisição do Português Brasileiro (100% de epêntese)

Ao lembrarmos que, neste modelo, o que importa são os valores de H, não é surpresa termos DEP com um valor de ranqueamento que o faz aparecer, numericamente, acima das demais restrições de marcação, relacionadas às sequências formadas de segmentos consonantais em coda. Se o que importa é DEP assumir valor superior ao da soma das restrições que violam os candidatos com plosiva em coda, ao mesmo tempo em que exhibe um valor inferior à Harmonia dos demais candidatos, não podemos nos ater unicamente aos valores do *ranking*.

Ao olharmos para os valores de Harmonia dos candidatos ótimos no conjunto de *Tableaux 2*, vemos que esta meta foi alcançada e que, ainda, foi possível garantir a distância necessária entre o valor de Harmonia (H) dos candidatos ótimos e dos candidatos perdedores, distância essa que faz com que não tenhamos variação nos candidatos ótimos.

No que se refere à restrição de fidelidade MAX e às restrições de marcação \*Dist {+3} e \*Dist {+3, +2}, reafirmamos que, não coincidentemente, em todas as simulações - aquisição de L1, aquisição de L2 e Interlíngua -, tais restrições sempre serão promovidas com relação às demais.

Em outras palavras, se observarmos todos os conjuntos de *tableaux* nesta Análise, veremos que o valor de ranqueamento para as restrições

\*Dist{+3} e \*Dist{+3, +2} será sempre o máximo, igual a 100.000. Isto quer dizer que, sob nenhuma hipótese, tais restrições serão demovidas, frente às sequências presentes na evidência positiva desta análise “t.t”, ”t.n”, ”n.t”, ”l.t”, ”w.t”, [s.t] e [s.n]. Quanto à restrição de fidelidade MAX, essa vai assumir valores mais altos, em torno de 88.000, pois não basta que ela esteja alta na hierarquia, mas, sim, que o valor de Harmonia do candidato que exhibe apagamento seja mais negativo do que o valor do candidato ótimo.

Para que possamos ver candidatos ótimos com epêntese apenas seguidos de obstruintes em coda – ra.[**pi**].to a.[**ki**].ne - (com exceção das sequências compostas de fricativa em coda [s.n] e [s.t], como nas palavras “asno” e “costa”), só precisamos que DEP assuma valor maior, de modo que o candidato com epêntese apresente um valor de H superior ao das sequências com plosiva em coda.

Estar acima quer dizer que DEP deve levar o candidato com epêntese a ter valor menos negativo, ou mais próximo a zero, com relação aos valores de Harmonia calculados. Por outro lado, ao mesmo tempo, o candidato que viola unicamente DEP deve ser mais negativo do que os candidatos que exibem as demais sequências [s.t], [s.n], n.t, l.t, w.t, não epentetizadas no PB. Como a HG trabalha com a cumulatividade dos pesos das restrições, seria difícil DEP alcançar valor mais negativo do que as restrições referentes à sequência [s.n], por exemplo, que requerem a soma de \*Dist{+3,...+1}, \*{stop, fric}<sub>coda</sub>, \*Dist{+3,...0}, \*Dist{+3,...-1}, \*{stop,...nas}<sub>coda</sub>, \*Dist{+3,...-2}, \*{stop,... liq}<sub>coda</sub>, \*{stop,... glide}<sub>coda</sub>, \*{+3,... -3}<sub>coda</sub>, não fosse a estratégia utilizada pelo algoritmo de fazer com que estas restrições atinjam valor negativo, sendo anuladas do cálculo de H.

Utilizando na HG a opção “*Linear OT*”, no programa PRAAT, assumimos que restrições com pesos negativos devem ser excluídas do cálculo de H. Para esta situação de aquisição do PB, em que DEP teve de assumir duas posições diante das restrições de marcação postas em estringência - na qual a violação de uma restrição mais específica implica violação às restrições mais gerais - esta opção foi importante. Foi possível, dessa forma, que as consoantes em coda fossem adquiridas no PB, com a particularidade de a aquisição das plosivas, em detrimento das demais consoantes em coda, exigirem a inserção da epêntese para a aquisição plena do Português Brasileiro.

Assim, em se tratando da aquisição da língua estrangeira dos aprendizes colombianos (Português Brasileiro), o algoritmo conseguiu convergir na HG para uma simulação que sugere *outputs* categóricos.

### 2.3.3 Simulação 3 – Interlíngua (*outputs* variáveis)

c) A aquisição dos *outputs* variáveis (Interlíngua) via Gramática Harmônica

Nesta simulação, vamos analisar o tratamento do algoritmo de aprendizagem frente à gramática em desenvolvimento dos aprendizes, com *ou-*

*tputs* variáveis. O algoritmo de aprendizagem da HG foi informado de que o estágio inicial da interlândia é o resultado final da L1. O conjunto de *Tableaux 3* demonstra a gramática em desenvolvimento dos estudantes, a partir do algoritmo de aprendizagem vinculado à HG.

	<i>ranking value</i>	<i>dizharmory</i>	<i>plasticity</i>
*Dist {+3}	100.000	98.003	1.000000
*Dist {+3, +2}	100.000	94.449	1.000000
Max	91.695	92.463	1.000000
Dep	67.922	71.107	1.000000
*Dist {1,0} & {stop}coda	29.941	31.129	1.000000
*{stop}coda	29.941	29.750	1.000000
*Dist {1} & {stop}coda	3.560	7.220	1.000000
*Dist {+3, ... 0}	-12.652	-11.276	1.000000
*{stop, fric}coda	-12.652	-17.022	1.000000
*Dist {+3, ...+1}	-26.078	-24.064	1.000000
*Dist {+3, ...-1}	-34.604	-32.458	1.000000
*{stop, ... nas}coda	-34.604	-36.958	1.000000
*{stop, ... liq}coda	-47.668	-47.225	1.000000
*Dist {+3, ...-2}	-47.668	-51.571	1.000000
*{stop, ... glide}coda	-59.617	-60.210	1.000000
*Dist {+3, ...-3}	-59.617	-61.978	1.000000

	*Dist{+3}	*Dist{+3, +2}	MAX	DEP	*Dist{1,0}&{stop}coda	*{stop}coda	*Dist{1}&{stop}coda	*Dist{+3, ...0}	*{stop, fric}coda	*Dist{+3, ...+1}	*Dist{+3, ...-1}	*{stop, ...nas}coda	*{stop, ...liq}coda	*Dist{+3, ...-2}	*{stop, ...glide}coda	*Dist{+3, ...-3}
♂ p.t/kt					*	*										
pi.t/ki.t				*												
t		*														
♂ p.n/k.n/t.n					*	*	*	*	*	*	*	*	*	*	*	*
pi.n/ki.n/ti.n				*												
n		*														
♂ nt										*	*	*	*	*	*	*
ni.t				*												
t		*														
♂ lt													*	*	*	*
li.t				*												
t		*														
♂ wt														*	*	*
wi.t				*												
t		*														
♂ st								*	*	*	*	*	*	*	*	*
si.t				*												
t		*														
♂ sn								*	*	*	*	*	*	*	*	*
si.n				*												
n		*														

Conjunto de *Tableaux 3* - Simulação 3 HG / Interlândia (outputs variáveis)

Olhando para os valores de Harmonia de cada candidato no conjunto de *Tableaux 3*, temos que os candidatos ótimos são “t.n” e [t]nia, “t.t” ca[k]to, “l.t” a[l]to, “w.t”, [s.t] a[s]tro, [s.n] a[s]no, “n.t” a[n]ta. Porém, se observarmos além da indicação do candidato ótimo, ao considerarmos o seu valor de Harmonia com relação aos candidatos perdedores, percebemos que “t.n” [p.n, k.n, t.n] e “t.t” [p.t, k.t] não são categóricos, diferentemente do que ocorre com as demais sequências. Os candidatos ótimos com a plosiva em

coda (“t.t” = -60.879 e “t.n” = -68.099) expressam um valor de H tão próximo aos candidatos com epêntese (t[i].t e t[i].n = -71.107) que indicam que, em algum momento de produção, as formas de saída oriundas do conjunto de *Tableaux* 3 para estas sequências com plosiva em coda podem mudar.

Conforme já vimos, estes valores, que determinam que tais candidatos ótimos devem variar, não são aleatórios, uma vez que foram obtidos a partir da análise acústica que realizamos com base nas produções dos aprendizes de PB nestas sequências, demonstrados na Tabela 2. Logo, as restrições responsáveis pelo valor de Harmonia de “t.n” e “t.t” devem organizar-se numericamente de forma a responder a esta situação. O conjunto de *Tableaux* 4, a seguir, tem o propósito de demonstrar esta variação entre os candidatos, com a produção da plosiva em coda e com epêntese.

	ranking value	disharmony	plasticity
+Dist {+3}	100.000	101.473	1.000000
+Dist {+3, +2}	100.000	99.538	1.000000
Max	91.693	91.336	1.000000
Dep	67.922	66.935	1.000000
+Dist {1,0} & {stop}coda	29.941	32.449	1.000000
+{stop}coda	29.941	31.761	1.000000
+Dist {1} & {stop}coda	3.560	3.068	1.000000
+Dist {+3, ... 0}	-12.652	-10.625	1.000000
+{stop, fric}coda	-12.652	-10.717	1.000000
+Dist {+3, ... +1}	-26.078	-27.142	1.000000
+Dist {+3, ... -1}	-34.604	-33.527	1.000000
+{stop, ... nas}coda	-34.604	-35.988	1.000000
+Dist {+3, ... -2}	-47.668	-47.249	1.000000
+{stop, ... liq}coda	-47.668	-50.064	1.000000
+Dist {+3, ... -3}	-59.617	-56.615	1.000000
+{stop, ... glide}coda	-59.617	-57.577	1.000000

	*Dist{+3}	*Dist{+3, +2}	MAX	DEP	*{stop}coda	*Dist{1,0}&{stop}coda	*Dist{1}&{stop}coda	*{stop,fric}coda	*Dist{+3,...0}	*Dist{+3,...+1}	*Dist{+3,...-1}	*{stop,...nas}coda	*Dist{+3,...-2}	*{stop,...liq}coda	*{stop,...glide}coda	*Dist{+3,...-3}	
p.t.k.t					*	*	*	*	*	*	*	*	*	*	*	*	-64210
pi.t/ki.t				*													-66935
t		*															-91336
p.n/k.n/t.n					*	*	*	*	*	*	*	*	*	*	*	*	-67278
pi.n/ki.n/ti.n				*													-66935
n		*															-91336
nt									*	*	*	*	*	*	*	*	0
ni.t				*													-66935
t		*															-91336
lt													*	*	*	*	0
li.t				*													-66935
t		*															-91336
w.t															*	*	0
wi.t				*													-66935
t		*															-91336
s.t							*	*	*	*	*	*	*	*	*	*	0
si.t				*													-66935
t		*															-91336
s.n							*	*	*	*	*	*	*	*	*	*	0
si.n				*													-66935
n		*															01226

Conjunto de Tableaux 4 - Simulação 3 HG / Interlíngua – Segunda Rodada

Se compararmos os valores do ponto de seleção (*disharmony*) do conjunto de *Tableaux* 4 com os valores obtidos na 1ª rodada desta simulação na HG, que trata da interlíngua dos estudantes colombianos (conjunto de *Tableaux* 3), veremos que as mudanças numéricas dos pesos das restrições fizeram com que o candidato com epêntese (“t[i].n”) emergisse, atingindo valor de H (-66.935) superior ou mais próximo de zero, com relação ao candidato com a plosiva em coda “t.n” (-67.278). É dessa forma que o algoritmo de aprendizagem mostrou que os candidatos ótimos relacionados à plosiva em coda não são categóricos na HG, em função da proximidade dos valores de H entre os candidatos.

Com relação aos *outputs* ótimos [s.t], [s.n], “n.t”, “w.t”, “l.t”, ao levarmos em conta que as restrições  $\ast\{\text{stop, fric}\}_{\text{coda}}$ ,  $\ast\{\text{stop, ...nas}\}_{\text{coda}}$ ,  $\ast\{\text{stop, ...liq}\}_{\text{coda}}$ ,  $\ast\{\text{stop, ...glide}\}_{\text{coda}}$  e  $\ast\text{Dist}\{+3, \dots +1\}$ ,  $\ast\text{Dist}\{+3, \dots 0\}$ ,  $\ast\text{Dist}\{+3, \dots -1\}$ ,  $\ast\text{Dist}\{+3, \dots -2\}$ ,  $\ast\text{Dist}\{+3, \dots -3\}$ , que estão relacionadas a estas sequências heterossilábicas com consoante diferente de plosiva em coda, apresentam valor negativo, concluímos que DEP poderia expressar qualquer valor positivo como peso, já que o valor negativo da restrição está indicando um valor de H igual a zero para as sequências formadas de consoante diferente de plosiva em coda. Essa situação pode também ser uma estratégia do modelo para dizer que tais restrições de marcação estão tendo uma função redundante, no momento em que se referem a sequências com o mesmo comportamento diante da sua exposição à coda medial de palavras.

Também é importante o fato de lidarmos com um *ranking* que vai organizar as restrições em termos de relações de marcação, preservando as relações de estringência. Nesta simulação, que trabalha com os dados de produção dos estudantes colombianos no PB (interlíngua), sempre vamos ter as restrições referentes à sequência “t.n” mais próximas a DEP, o que indica que esta sequência, além de mais marcada, está assumindo maior ocorrência de epêntese (13,48%), com relação à “t.t” (1,04%) e, consequentemente, com relação às demais sequências, que mantêm suas restrições de marcação sempre abaixo na hierarquia para que se mantenham categóricas. Por sua vez, as restrições conjuntas  $\ast\text{Dist}\{1\} \& \ast\{\text{stop}\}_{\text{coda}}$  e  $\ast\text{Dist}\{1,0\} \& \ast\{\text{stop}\}_{\text{coda}}$ , que caracterizam e diferenciam as sequências heterossilábicas de plosiva em coda das demais sequências, foram fundamentais neste processo, indicando que, mesmo em um modelo que trabalha com a cumulatividade dos dados, restrições conjuntas, que expressam a ação conjunta de duas restrições, são importantes.

Assim, em se tratando da aquisição da língua estrangeira dos aprendizes (Português Brasileiro), o algoritmo conseguiu convergir na HG para

uma simulação que sugere *outputs* categóricos e variáveis. Dessa forma, visualizamos a capacidade do modelo em lidar com essas duas situações em paralelo.

### 3 Considerações Finais

O trabalho propôs analisar a produção da vogal epentética à luz do modelo da Gramática Harmônica, num contexto de plosivas em codas mediais de palavras, por parte de aprendizes de Português Brasileiro, que são falantes nativos de Espanhol (colombianos). Para tanto, foram formuladas três Questões Norteadoras:

A primeira Questão Norteadora, que já pressupunha a ocorrência da epêntese nos *outputs* dos aprendizes no Português Brasileiro após consoantes plosivas em codas mediais de palavras em sequências não-tautossilábicas, questionava qual(is) fator(es) linguístico(s) contribui(ram) para a ocorrência dessa epêntese.

Com os dados tratados estatisticamente, conforme Tabelas 1 e 2, foi possível constatar que o contato silábico estaria influenciando na ocorrência da epêntese na produção dos sujeitos. Destaca-se, nesse sentido, o contato com uma consoante nasal no *onset* da sílaba seguinte, como na palavra *apneia*. A conclusão a essa questão foi fundamental para que pudessemos estabelecer as restrições que dessem conta do fenômeno que estamos analisando e, assim, responder à segunda Questão Norteadora, que questionava que organização de restrições responde pelas formas de *output* produzidas pelos aprendizes.

Primeiramente, verificamos que o peso das restrições para a formalização do fenômeno é fundamental, pois através dele calculamos o valor de Harmonia, definidor do candidato ótimo. Na L1 dos aprendizes, as restrições de fidelidade DEP e MAX assumem um peso que as mantém afastadas das restrições de marcação, de tal forma que, mesmo se fizéssemos avaliações sucessivas, não veríamos variação entre as restrições. A demerção das restrições de marcação demonstrou que o Espanhol Colombiano sugere que as restrições de fidelidade assumam as posições mais altamente ranqueadas de forma a caracterizar a organização das restrições no Espanhol.

As restrições de marcação \*Dist {+3} e \*Dist {+3, +2}, não coincidentemente, em todas as simulações - aquisição de L1, aquisição de L2 e Interlíngua - mantêm-se altas com relação às demais, pois nunca serão violadas pelas sequências de segmentos propostos. Como estamos traba-

lhando somente com sequências existentes no Português Brasileiro (PB) e no Espanhol Colombiano (EC), e tais restrições contemplam contatos não existentes nem no PB, nem no EC (como, por exemplo, uma consoante plosiva em coda seguida de um glide ou uma consoante nasal em coda, também seguida de um glide), essas receberão sempre os valores mais altos de ranqueamento, de modo que seus valores não se modifiquem de simulação para simulação, pelo fato de tais restrições não sofrerem promoção ou demção em função da evidência positiva.

Na aquisição plena da língua estrangeira - Português Brasileiro -, percebemos alterações importantes para que se pudesse assumir outro candidato ótimo e, também, categórico. Na aquisição plena do PB, assumimos que o candidato ótimo deveria ser aquele com 100% de epêntese e sem variação; logo, DEP teve de ser demovido, passando de um valor de ranqueamento de 88626 na aquisição completa do Espanhol Colombiano (Simulação 1) para 60725, para que a epêntese pudesse emergir como a estratégia de reparo capaz de adaptar, ao padrão de PB, os *inputs* que resultariam em plosivas em codas mediais das palavras dadas. Logo, o algoritmo de aprendizagem teve de atualizar os pesos das restrições - apesar de os *rankings* permanecerem com as restrições de fidelidade dominando marcação - para que, à luz da HG, os valores de Harmonia do candidato ótimo com epêntese fossem superiores aos dos candidatos perdedores.

A simulação 3, que pretendia demonstrar os *outputs* variáveis das produções dos aprendizes no PB, teve de novamente restabelecer os pesos das restrições. Merece destaque, no que diz respeito a esta simulação, a diferença fundamental de que tal simulação resultou em valores de Harmonia próximos para o candidato dado como ótimo e o candidato com epêntese, proximidade essa que indicava a probabilidade de variação entre tais candidatos, permitindo a produção variável do segmento epentético. O algoritmo consegue, assim, retratar os estágios desenvolvimentais por que passaram os aprendizes, ao mostrar, inclusive, o papel fundamental da restrição de fidelidade DEP, no processo de aquisição das plosivas em posição de coda medial no Português Brasileiro e Espanhol Colombiano.

Abordamos, na sequência, a terceira Questão Norteadora, que, consequentemente, questionava se o modelo da Gramática Harmônica consegue formalizar com adequação o processo de aquisição de L2. Confirmamos que a hierarquia das restrições utilizadas nas simulações, dada pelo peso das restrições, se mostrou capaz de defender não somente as gramáticas categóricas da língua materna (com 0% de epêntese) e da língua estrangeira (com 100% de epêntese), mas, também, a gramática de *outputs* variáveis, da interlíngua.

Podemos acrescentar, ainda, que o modelo da HG, diante das restrições de marcação propostas:  $\ast\{\text{stop, fric}\}\text{coda}$ ,  $\ast\{\text{stop, ...nas}\}\text{coda}$ ,  $\ast\{\text{stop, ...liq}\}\text{coda}$ ,  $\ast\{\text{stop, ...glide}\}\text{coda}$  e  $\ast\text{Dist}\{+3, \dots, 0\}$ ,  $\ast\text{Dist}\{+3, \dots, -1\}$ ,  $\ast\text{Dist}\{+3, \dots, -2\}$ ,  $\ast\text{Dist}\{+3, \dots, -3\}$ , demonstrou para todas as simulações que, para a emergência do candidato a *output* ótimo, talvez não fossem necessárias estas restrições.

Cabe salientarmos que a formalização de tais restrições é advinda do processo de Alinhamento Relacional, uma vez que as restrições em questão remetem a entes de primitivos escalares. Através dos processos de Alinhamento Harmônico e Relacional, produzimos, portanto, restrições que expressam a natureza das escalas primitivas, ainda que muitas das restrições constituintes deste sistema de oposições (em ranking fixo ou estrinência, como foi aqui o caso) não venham a ter efetivo poder decisivo na escolha do candidato ótimo.

Em uma OT Estocástica, conforme visto na análise de Azevedo, Alves e Matzenauer (2012), a própria dominância estrita, característica do sistema de avaliação do modelo em questão, se mostra suficiente para excluir qualquer ação destas possíveis restrições, que, apesar de presentes por imposição dos processos de Alinhamento Harmônico e Relacional, acabam por não exercer papel no ranking. Já em um modelo como a Gramática Harmônica, em que todas as restrições podem vir a ter um papel na definição de Harmonia, cabe-nos discutir que estratégias podem ser lançadas pelo algoritmo para que algumas restrições, que não deveriam apresentar papel decisivo na gramática e se mostram presentes apenas por uma questão de formalização, tenham seus efeitos excluídos da definição de Harmonia. Considerando-se o recurso de *Linear OT* do algoritmo vinculado à HG, já descrito neste trabalho, a possibilidade de exclusão da ação destas restrições no cálculo da Harmonia pode ser prevista, uma vez que tais restrições desnecessárias podem receber valores negativos. Julgamos que, ao fornecer valores negativos às restrições em todas as simulações desta última análise, o modelo pode estar acusando o caráter desnecessário de tais restrições para explicar o processo que estávamos analisando. Tais restrições, outrossim, têm sua formalização fundamentada nos mecanismos de formação de restrições a partir de escalas primitivas, conforme já explicamos. De qualquer modo, verificamos, dessa forma, que a Gramática Harmônica, ainda que considere o peso de todas as restrições e não opere sob a noção de dominância estrita, também apresenta, assim como a Teoria da Otimidade Estocástica, artifícios formais que permitem destacar a existência de restrições que, para uma dada análise em específico, se fazem desnecessárias.

Concluímos, assim, que a análise de aquisição da língua estrangeira aqui realizada, à luz do modelo da Gramática Harmônica, permitiu uma explicação formal para os padrões de produção encontrados em nossos dados, dados esses, também discutidos em Azevedo, Alves e Matzenauer (2012), sob o modelo da Teoria da Otimidade Estocástica. Tal modelo teórico revela-se, portanto, adequado para a formalização do processo de aquisição fonológica de segunda língua. Demonstrada a capacidade, por parte deste modelo, de dar conta dos dados desenvolvimentais, mostraram-se necessárias, em estudos futuros, discussões teóricas que, ao irem além da verificação sobre a capacidade de formalização de dados tanto sob o modelo da TO Estocástica quanto da HG, se voltem, sobretudo, para as concepções de linguagem que embasam cada um desses modelos. Acreditamos que este trabalho, ao demonstrar que a formalização sob a HG se faz tão viável quanto à realizada via Teoria da Otimidade Estocástica, se tenha mostrado uma contribuição ao representar o ponto de partida para que os pesquisadores se sintam motivados a discutir as questões de base envolvidas em cada uma dessas distintas maneiras de analisar os dados linguísticos.

## BIBLIOGRAFIA

- ALVES, U. *Teoria da Otimidade, Gramática Harmônica e Restrições Conjuntas*. *Alfa:Revista de Linguística (UNESP. São José do Rio Preto. Impresso)*. V.54, p.237-263, 2010.
- AZEVEDO, R. ALVES, U. MATZENAUER, C. A Epêntese na Produção de Plosivas em Codas Mediais do Português Brasileiro por Colombianos: uma abordagem com base em restrições. (jan./jun. 2012). *Letras & Letras, Uberlândia*, v.28, n.1, p.327-345, 2012.
- AZEVEDO, R. A Epêntese no Português Brasileiro (L2), em segmentos plosivos em codas mediais, por falantes nativos do Espanhol colombiano (L1): uma análise via Teoria da Otimidade Estocástica e Gramática Harmônica. Dissertação de Mestrado em Linguística Aplicada inédita. Pelotas: UCPEL, 2011.
- BOERSMA, P. HAYES, B. Empirical Tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32, 45-86, 2001.
- BOERSMA, P.; PATER, J. Convergence Properties of a Gradual Learning Algorithm for Harmonic Grammar. Amsterdam: University of Amsterdam, UMass Amherst, 2008. (Manuscript).

- BOERSMA, P. WEENINK, D. PRAAT: doing phonetics by computer: version 5.1.07.2009 [computer program]. Disponível em <http://www.praat.org/>.
- CLEMENTS, G. N. The Role of the Sonority Cycle in Core Syllabification. In: KINGSTON, J. & BECKMAN, M. (Orgs.). *Papers in Laboratory Phonology I*. Cambridge: CUP, p. 283-333, 1990.
- COLLISCHONN, G. *Análise Prosódica da Sílabas em Português*. Tese de doutorado inédita. Porto Alegre: PUCRS, 1997.
- GOUSKOVA, M. Relational Hierarchies in Optimality Theory: The case of Syllable Contact. *Phonology* 21:2, p. 201-250, 2004.
- JESNEY, K.; TESSIER, A. Re-evaluating Learning Biases in Harmonic Grammar. In: BECKER, M. (Org.). *University of Massachusetts Occasional Papers in Linguistics*, 36: *Papers in Theoretical and Computational Phonology*. Amherst: GLSA, 2007. p.1-42.
- LEGENDRE, G.; MIYATA, Y.; SMOLENSKY, P. Can connectionism contribute to syntax? Harmonic Grammar, with an application. In: ZIOLKOWSKI, M.; NOSKE, M. DEATON, K. (eds.). *Proceedings of the 26th Regional Meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society, 1990.
- MCCARTHY, A.; PRINCE, A. Faithfulness and Reduplicative Identity. In: BECKMAN, J. et al (eds.). *University of Massachusetts Occasional Papers in Linguistics* 18. Amherst, MA: GLSA Publications, 1995. [Disponível: ROA 60].
- MURRAY, R.; VENNEMANN, T. Sound Change and Syllable Structure [problems] in Germanic Phonology. *Language* 59, 1983.
- PRINCE, A; SMOLENSKY, P. Optimality Theory: Constraint Interaction in Generative Grammar. Technical Report, Rutgers University and University of Colorado at Boulder, 1993. Revised version published by Blackwell, 2004. [Disponível: ROA-573].
- SMOLENSKY, P.; LEGENDRE, G. *The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar*. Cambridge: MIT, 2006.

*Recebido em: 23/02/2013; Aceito em: 07/06/2013*