TEORIA DA OTIMIDADE, GRAMÁTICA HARMÔNICA E RESTRIÇÕES CONJUNTAS

Ubiratã Kickhöfel ALVES*

- RESUMO: A Teoria da Otimidade, *Standard* (PRINCE; SMOLENSKY, 1993) ou Estocástica (BOERSMA; HAYES, 2001), opera sob a noção de dominância estrita. Diferencia-se, nesse aspecto, do modelo teórico da Gramática Harmônica (LEGENDRE; MIYATA; SMOLENSKY, 1990; SMOLENSKY; LEGENDRE, 2006), na qual a avaliação do candidato ótimo considera o caráter cumulativo de todas as violações incorridas por cada um dos candidatos a *output*. Ao considerarmos tal caráter cumulativo da Gramática Harmônica, questionamos a efetiva necessidade de formação de restrições conjuntas à luz de tal modelo teórico. Frente a tal questionamento, foram realizadas simulações computacionais, através do *software* Praat (BOERSMA; WEENINK, 2009), à luz dos algoritmos de aprendizagem vinculados aos dois modelos. As respostas fornecidas pelos algoritmos evidenciam que o modelo da Gramática Harmônica consegue convergir em sistemas que só se mostrariam passíveis de aprendizagem, via OT Estocástica, através de restrições conjuntas. Os resultados apontados incitam a discussão a respeito do papel do mecanismo de Conjunção Local sob a Gramática Harmônica, bem como evidenciam a necessidade de uma reflexão acerca das implicações do uso de um ou outro modelo de análise linguística.

- PALAVRAS-CHAVE: Teoria da Otimidade. Gramática Harmônica. Conjunção Local. Restrições Conjuntas. Algoritmos de Aprendizagem.

Introdução

A noção de dominância estrita é um dos principais aspectos caracterizadores da Teoria da Otimidade, tanto em sua versão *Standard* (PRINCE; SMOLENSKY, 1993) quanto sob sua versão Estocástica (BOERSMA; HAYES, 2001). Sob tal aspecto, candidatos que apresentam marcas de violação na restrição mais altamente ranqueada são automaticamente eliminados na disputa pelo *status* de ótimo, independentemente do número de marcas de violação incorridas por este mesmo candidato nas restrições mais baixas na hierarquia. Em outras palavras, considerando-se a existência de um *ranking* de restrições do tipo $C_1 >> C_2 >> C_3$, candidatos que violam C_1 serão automaticamente excluídos, pois a violação da restrição mais alta caracteriza-se como condição suficiente para

* UFRGS – Universidade Federal do Rio Grande do Sul. Instituto de Letras – Departamento de Línguas Modernas. Porto Alegre – RS – Brasil. 91570-000 – ukalves@gmail.com.

a eliminação. Ao seguirmos essa linha de análise, podemos ter, como *output* ótimo, um candidato que incorra um número muito maior de violações do que as incorridas pelos candidatos eliminados; em tal caso, tais violações, ainda que em número muito maior, ocorrem sobre as restrições de *status* mais baixo na hierarquia. Em suma, com a eliminação de um candidato através da violação da restrição mais altamente ranqueada, restrições mais baixas mostram-se, portanto, irrelevantes.

Uma consequência direta da noção de ranqueamento estrito está na impossibilidade de, através de restrições, expressar a gramática de sistemas que admitem uma ou outra estrutura marcada, mas não duas violações de marcação em um mesmo domínio. Para exemplificarmos o aqui dito, consideremos três restrições, C_1, C_2 e C_3, sendo as duas primeiras de marcação e a última de fidelidade[1]. Através do *ranking* $C_3 >> C_2 >> C_1$, vemos que a língua permite *outputs* fiéis que violem C_2 ou C_1, conforme expresso em (1).

(1)

	C_3	C_2	C_1
☞ a		*	
b	*!		
a_1	*!		
☞ b_1			*

Considerando-se o ranking em questão, candidatos que violam ambas as restrições também se sagram ótimos, conforme vemos em (2).

(2)

	C_3	C_2	C_1
a_2	*!		
☞ b_2		*	*

Como explicitar sistemas que, conforme já afirmado, permitem uma ou outra estrutura marcada, mas não a violação das duas restrições de marcação em um

[1] Ao seguirmos Fukazawa e Miglio (1998), consideramos que o processo de conjunção deve ser limitado a restrições de uma mesma família, o que explica definirmos tanto C_1 quanto C_2 como restrições de marcação. A presença de C_3 no exemplo em questão caracteriza o conflito clássico da OT entre marcação e fidelidade.

mesmo domínio?[2] Promover o *status* hierárquico de ambas as restrições não pode ser a solução, uma vez que o ranqueamento em questão permitiria apenas candidatos não marcados, e, conforme já afirmado, o sistema desejado deve levar, também, a *outputs* fiéis quando apenas uma das restrições de marcação esteja sendo violada. A impossibilidade de expressar o sistema em questão advém da própria noção de dominância estrita: o candidato 'a_2' não pode se sagrar ótimo em função de a marca de violação de C_3 já eliminá-lo, independentemente das marcas de violação de 'b_2' sobre C_1 ou C_2. O fato de o perdedor estar violando uma ou as duas restrições de marcação mostra-se irrelevante, visto que fidelidade, mais altamente ranqueada para garantir a emergência dos candidatos 'a' e 'b_1', está sendo desrespeitada.

Para dar conta de sistemas que aceitem, sob uma mesma gramática, os *outputs* 'a', 'b_1' e 'a_2', a solução encontrada, na Teoria da Otimidade, encontra-se no mecanismo de Conjunção Local. Proposto por Smolensky (1995), o esquema de conjunção de restrições pode ser da seguinte forma definido:

(3)

> Conjunção Local de Restrições (SMOLENSKY, 1995): a Conjunção Local de duas restrições C_1 e C_2 em um domínio D, $C_1 \&_D C_2$ é violada sempre que ambas as restrições C_1 e C_2 são violadas dentro de um dado domínio.

À luz da Teoria da Otimidade, restrições conjuntas associam duas restrições já baixas no *ranking*. Para verificarmos o papel exercido por uma restrição conjunta, o *tableau* em (4), de caráter explicativo, tem o objetivo de demonstrar a ação da Conjunção Local em uma hierarquia de restrições:

(4)

	[C_1 & C_2]	C_3	C_2	C_1
☞ a_2		*		
b_2	*!		*	*

Vemos, a partir dos pares de candidatos acima, a ação da conjunção das restrições C_1 e C_2 formando uma restrição conjunta que, estando altamente ranqueada, exclui o candidato 'b_2'. Considerando-se que C_1 e C_2 sejam restrições de marcação, concluímos que a língua em questão permite uma estrutura

[2] Para exemplos de situações hierárquicas semelhantes, encontráveis nas línguas do mundo, aconselhamos a leitura de Itô e Mester (1998). Os autores, dentre vários outros exemplos, mencionam a neutralização de obstruintes vozeadas em coda do alemão, língua que permite a violação, ou da restrição *NoCoda* ou de **Voiced Obstruent*, mas não a violação de ambas as restrições ao mesmo tempo.

marcada ou outra, mas não as duas ocorrendo ao mesmo tempo no mesmo domínio. A restrição conjunta [C_1 & C_2] impede justamente esses casos. Em outras palavras, ainda que C_1 e C_2 já estejam baixas no *ranking*, elas podem, ainda, ter seus efeitos sentidos contra a restrição C_3. Cabe dizer que, ainda que restrições como [C_1 & C_2] consigam expressar a ação conjunta de duas restrições que já se encontram baixas no sistema, não parece haver, entre os estudiosos, um consenso acerca dos aspectos basilares da Teoria da Conjunção Local. Questões como a universalidade do Operador de Conjunção '&' e a formação das restrições na aquisição de linguagem, além de tentativas de limitação de tal operador, com vistas a impedir uma ação exagerada e um acréscimo desnecessário do número de restrições, encontram-se em pleno debate na literatura (ITO; MESTER, 1998; BAKOVIC, 1999, 2000; FUKAZAWA, 1999, 2001; FUKAZAWA; MIGLIO, 1998; MORETON; SMOLENSKY, 2002; LUBOWICZ, 2002, 2006, 2007; BONILHA, 2003, 2005; ALVES, 2008).

Uma vez que a Conjunção Local é o resultado da noção de dominância estrita da OT, questionamo-nos se um modelo de análise que não use desse critério pode vir a não carecer da formação de restrições adicionais, para expressar o efeito cumulativo de duas restrições mais baixas. Um modelo que opera sob restrições, mas desconsidera a noção de dominância estrita, é a Gramática Harmônica (LEGENDRE; MIYATA; SMOLENSKY, 1990; SMOLENSKY; LEGENDRE, 2006). Em tal modelo, que tem sido empregado em uma série de trabalhos recentes (PATER, 2005, 2008, 2009; PATER; JESNEY; TESSIER, 2007; TESSIER, 2007; BOERSMA; PATER, 2008; COETZEE; PATER, 2009; GOLDRICK; DALAND, 2009; ALVES, 2009) como uma abordagem alternativa à OT, a organização das restrições dá-se por pesos numéricos que apresentam caráter cumulativo na escolha do candidato ótimo, caracterizando "efeitos de gangue" entre as restrições. Dessa forma, na escolha do candidato ótimo, todas as restrições exercem papel importante: é possível, dessa forma, que o candidato ótimo seja um que viole restrições com maior valor numérico (mais altas no sistema) em função de os candidatos perdedores apresentarem um número alto de violações nas restrições com *status* mais baixo na gramática da língua em questão. Em outras palavras, a violação da restrição mais alta não é eliminatória nesse modelo teórico. Considerada tal característica, um exercício teórico importante diz respeito à verificação da possibilidade de ação conjunta de restrições: é possível que um candidato que viole a restrição mais alta sagre-se ótimo em função de os candidatos perdedores estarem violando, individualmente, um conjunto de restrições mais baixas, as quais, somadas, exercerão um efeito cumulativo que poderá vir a apresentar um peso superior ao possibilitado pela violação da restrição mais alta[3].

[3] Conforme será visto em breve, para a determinação do peso da Harmonia de um candidato, que determina os efeitos de gangue, é necessário considerar todas as restrições que violam tal candidato, uma vez que todas elas terão seus pesos somados.

Ao indagarmos acerca da possibilidade de o caráter cumulativo da HG exercer papel semelhante ao das restrições conjuntas que operam sob os modelos de dominância estrita, o presente trabalho tem por objetivo verificar a capacidade de o modelo da Gramática Harmônica operar sem a necessidade de restrições conjuntas. Tal verificação se fará possível por meio da simulação do processo de aquisição de linguagem, ao utilizarmos, na simulação computacional, os algoritmos associados a cada um dos modelos teóricos: o OT-GLA (BOERSMA; HAYES, 2001), algoritmo associado à Teoria da Otimidade Estocástica[4], e o HG-GLA (BOERSMA; PATER, 2008; BOERSMA; WEENINK, 2009), vinculado à Gramática Harmônica. Por meio da simulação do processo de aquisição nesses dois algoritmos, hipotetizamos que, sem a restrição conjunta, o algoritmo vinculado à OT Estocástica não conseguirá convergir, ou seja, nunca atingirá o estágio final da gramática quando o sistema a ser atingido, ainda que permita a emergência de candidatos fiéis com violações sobre uma ou outra restrição de marcação, rejeite, por sua vez, *outputs* fiéis que apresentem violações sobre essas duas restrições de marcação referentes a um mesmo domínio. Por sua vez, hipotetizamos que a Gramática Harmônica conseguirá convergir no estágio final da gramática em questão, utilizando-se apenas três restrições, C_1, C_2, e C_3, de modo a dispensar a Conjunção Local dentro desse modelo teórico. Os resultados das simulações em HG propiciarão, portanto, não somente uma discussão acerca de como se dá a avaliação de candidatos nesse construto teórico, mas, sobretudo, uma reflexão acerca da existência do Operador de Conjunção Local '&' em tal modelo de análise linguística.

O artigo é organizado da seguinte forma: apresentamos, no que segue, os principais aspectos teóricos que diferenciam a OT Estocástica da Gramática Harmônica, além de serem feitos esclarecimentos acerca dos algoritmos associados a cada um dos modelos. Estabelecidas tais diferenças, apresentamos, então, o resultado das simulações computacionais do processo de aquisição da gramática-alvo, através do *software Praat – Version 5.1.13* (BOERSMA; WEENINK, 2009), sob ambos os algoritmos. O trabalho é encerrado com uma discussão teórica acerca das diferenças entre a Teoria da Otimidade e a Gramática Harmônica, voltada não somente para a avaliação dos candidatos e para os conjuntos de restrições de ambos os modelos, mas também para as implicações de realizarmos análises linguísticas sob um ou outro modelo teórico.

[4] Neste trabalho, utilizamos a Teoria da Otimidade em sua versão Estocástica (BOERSMA; HAYES, 2001), uma vez que o *Constraint Demotion Algorithm* (TESAR; SMOLENSKY, 1993, 1996, 1998, 2000), algoritmo de aprendizagem ligado à OT Standard, não é capaz de dar conta de dados variáveis. Para maiores detalhes acerca da superioridade do *Gradual Learning Algorithm*, algoritmo desenvolvido à luz da OT Estocástica, aconselhamos a leitura de Bonilha (2005).

Referencial teórico

Antes de iniciarmos a descrição e discussão dos índices numéricos fornecidos pelos algoritmos de aprendizagem associados aos modelos da Teoria da Otimidade e da Gramática Harmônica, julgamos importante explicitar as diferenças entre esses dois modelos teóricos, sobretudo no que diz respeito à escolha do candidato ótimo. Dessa forma, nas seções que seguem, descreveremos os principais aspectos da Teoria da Otimidade Estocástica e da Gramática Harmônica, respectivamente. Através da descrição individual de cada modelo, acreditamos deixar ainda mais clara, também, a questão de investigação que norteia o presente estudo: a verificação de o modelo da Gramática Harmônica, ao contrário da Teoria da Otimidade, dispensar o mecanismo de Conjunção Local.

A OT Estocástica

Na OT Estocástica (BOERSMA; HAYES, 2001), as restrições recebem valores numéricos para atuarem ao longo de uma escala. Cada vez em que há avaliação de candidatos, tais valores são convertidos em um ranqueamento correspondente. O ranqueamento em questão, resultante da conversão dos valores numéricos em hierarquia, segue as mesmas premissas de avaliação do modelo *Standard* da OT, a partir do qual o candidato ótimo é aquele que obedece às restrições mais altamente ranqueadas, independentemente do número de violações incorridas por tal candidato às restrições de *status* mais baixo na gramática. Vejamos tal noção de dominância estrita no *tableau* que segue:

(5)

	C_2	C_1
[*Output1*]	*!	
☞ [*Output2*]		***

De acordo com o *tableau* em (5), o candidato ótimo é [*Output2*], pois ele não viola a restrição mais altamente ranqueada na hierarquia, ainda que tenha desrespeitado a restrição mais baixa três vezes. Nas análises à luz da OT, a eliminação dos candidatos se dá a partir da restrição mais alta: assim, [*Output 1*] já é eliminado pela restrição C_2, o que fica claro pela marca de violação fatal "!", que simboliza a exclusão do candidato. Tal aspecto, conforme veremos em breve, constitui a principal característica que diferencia a Teoria da Otimidade do modelo da Gramática Harmônica. Não interessa para a OT Estocástica o número de violações incorridas por uma violação mais baixa, se restrições mais altas estiverem sendo violadas; se uma restrição mais alta estiver sendo violada por dado um candidato, tal candidato será, indiscutivelmente, excluído.

Conforme explicam Coetzee e Pater (2009), um aspecto importante da OT Estocástica diz respeito ao fato de ela ser acompanhada de uma teoria de aprendizagem, sendo vinculada a um algoritmo chamado de Algoritmo de Aprendizagem Gradual (GLA, do inglês *Gradual Learning Algorithm*). De acordo com os princípios de funcionamento do algoritmo em questão, o aprendiz recebe um mapeamento *input-output* de cada vez, e o estado corrente da gramática determina o *output* ótimo. Quando o *output* gerado difere dos dados da evidência positiva, o aprendizado acontece. O GLA atualiza o valor das restrições, de modo a subtrair um valor *x*[5] dos valores das restrições que são mais violadas na forma correta do que no "erro" do aprendiz, além de adicionar um valor *x* a todas as restrições que são violadas no candidato com erro. A hierarquia de restrições, conforme vimos, é estabelecida em função dos valores a serem assumidos pelas restrições de tal escala numérica.

Ao nos referirmos ao GLA, julgamos fundamental ressaltar o seu caráter estocástico, de acordo com o qual o ranqueamento é afetado por um dado valor de ruído (*noise*) estatístico[6] a cada momento de avaliação de candidatos. O valor numérico das restrições a ser promovido ou demovido pelo algoritmo corresponde ao ponto central de uma faixa ou gama de valores probabilísticos que podem vir a ser assumidos pela restrição em questão, em um dado momento de produção. Em função do ruído, a cada momento de fala, as restrições podem assumir um índice numérico distinto, caracterizado por Boersma e Hayes (2001) como "ponto de seleção". Em avaliações (momentos de produção linguística) sucessivas, restrições que apresentam tais valores centrais próximos um do outro poderão variar em termos de ranqueamento. Assim, é possível que, em um dado momento de produção, uma restrição C_1 assuma um ponto de seleção mais alto do que C_2, enquanto que, em outros momentos, C_2 assuma um valor de ponto de seleção mais alto do que C_1[7], ainda que, por exemplo, o valor central de C_1 seja superior ao de C_2.

---

[5] O valor *x*, que corresponde à taxa de incremento/decremento do algoritmo, é definido, na simulação computacional, através do valor de *plasticidade*, que pode ser definido pelo pesquisador no *software Praat*.

[6] O valor de ruído *default* do algoritmo computacional sugerido em Boersma e Hayes (2001) é 2.0.

[7] Boersma e Hayes (2001) chamam a atenção para o fato de que, ainda que o ponto de seleção possa compreender qualquer índice numérico dentro da faixa de valores, ele é mais provável de assumir pesos numéricos mais próximos do ponto central de tal gama de valores (ou, conforme chamam Boersma e Hayes (2001), do *ranking value* – "valor de ranqueamento" ou "valor central"). Por exemplo, considerando-se uma restrição C_1, que apresenta valor central 67 e uma faixa que vai de 62 a 72, é mais provável que o ponto de seleção venha a assumir um índice numérico tal como 66, 67 ou 68 em vez de 62 ou 72, ainda que esses últimos sejam também probabilisticamente possíveis.

Conseguimos, desse modo, expressar a ocorrência de *outputs* variáveis em uma língua: a variação ocorre porque, em alguns momentos de conversão dos valores numéricos em *rankings*, as restrições apresentam uma relação hierárquica X >> Y, enquanto que, em outros momentos, a relação Y >> X pode ocorrer. Isso somente acontece quando ambas as restrições apresentam valores centrais bastante próximos, de modo que o valor de ruído propicie que, em alguns momentos de avaliação, X apresente valor mais alto do que Y e, em outros momentos, Y consiga expressar valores mais altos do que X. Considerar que uma restrição domina categoricamente outra, de modo que a segunda nunca possa vir a apresentar um valor de ponto de seleção (e, consequentemente, um *status* hierárquico) mais alto do que a primeira, significa que, na escala contínua em questão, as duas restrições apresentam valores centrais bastante afastados, para que não haja um *overlap* em suas gamas de possíveis valores de ponto de seleção e, desse modo, não seja possível uma inversão hierárquica no momento da avaliação.

O GLA revela-se um algoritmo poderoso se comparado ao seu antecessor, o Algoritmo de Demoção de Restrições (vinculado à OT *Standard*), por ser capaz de modelar o processo de variação linguística. Apesar dessa grande vantagem, uma série de trabalhos (PATER, 2005, 2008; PATER; JESNEY; TESSIER, 2007; TESSIER, 2007; BOERSMA; PATER, 2008) demonstra a incapacidade de tal algoritmo em convergir, ou seja, em chegar a resultados numéricos que sejam convertidos em uma gramática[8]. Em outras palavras, "[...] existem línguas que podem ser representados pela OT (e, portanto, pela OT Estocástica), mas não conseguem ser aprendidas pelo GLA." (BOERSMA; PATER, 2008, p.2). Dadas as limitações do algoritmo vinculado ao modelo em questão, Boersma e Pater (2008) apontam como solução um algoritmo baseado em um modelo que desconsidera a noção de ranqueamento com dominância estrita de restrições, pelo fato de tomar como premissa a ação cumulativa de todas as restrições envolvidas na avaliação dos candidatos. Tal algoritmo, capaz de convergir em padrões que não conseguiam ser expressos pela OT Estocástica, mostra-se vinculado aos preceitos da Teoria de Gramática Harmônica (LEGENDRE; MIYATA; SMOLENSKY, 1990; SMOLENSKY; LEGENDRE, 2006).

A Gramática Harmônica

Ainda que o primeiro trabalho à luz da Gramática Harmônica tenha sido publicado em 1990 (LEGENDRE; MIYATA; SMOLENSKY, 1990), antes mesmo

[8] Para uma discussão mais detalhada das limitações de convergência do algoritmo vinculado à OT Estocástica, aconselhamos a leitura da obra de Pater (2008), em que o autor descreve o chamado "Problema dos *Rankings* WLW", considerado o maior caso de não convergência do algoritmo em questão.

do texto fundador da Teoria da Otimidade (PRINCE; SMOLENSKY, 1993), foi em 2006, com o lançamento da obra *The harmonic mind* (SMOLENSKY; LEGENDRE, 2006), que o modelo começa a ser retomado pelos estudiosos, sobretudo como uma alternativa a algumas das limitações de convergência apresentadas pela OT Estocástica, conforme já discutido.

Assim como na OT Estocástica, o modelo da Gramática Harmônica também opera com restrições que apresentam caráter numérico. Entretanto tal modelo diferencia-se da OT Estocástica sobretudo no que diz respeito à avaliação do candidato ótimo. Conforme explicam Jesney e Tessier (2007), a HG, de modo semelhante à OT, apresenta três componentes: GEN, módulo da gramática capaz de tomar um *input* e criar infinitas possibilidades de forma de saída para tal representação; CON, o conjunto universal de restrições violáveis; e EVAL, módulo da gramática que realiza a avaliação dos candidatos, capaz de apontar o candidato ótimo a partir dos preceitos de seleção do modelo teórico. A diferença entre a Teoria da Otimidade e a Gramática Harmônica reside neste último componente, uma vez que os preceitos de avaliação do candidato ótimo são distintos sob cada um dos dois modelos.

A diferença entre a OT Estocástica e a HG, em termos de avaliação, encontra-se na noção da dominância estrita que é a base da OT, mas não da HG. Conforme já dissemos, sob a concepção de dominância estrita, as restrições com *status* mais baixo fazem-se relevantes apenas nos casos em que as restrições mais altas não se mostram capazes de decidir o *output* ótimo em função de empates entre candidatos frente a tais restrições. Isso fica claro em (6), em que apresentamos um *tableau* sob o modelo da Teoria da Otimidade:

(6)

	C_3	C_2	C_1
[*Output1*]	*!		
☞ [*Output2*]		*	*
[*Output3*]		**!	

No *tableau* em (6), a restrição C_3 não se mostra capaz de decidir o candidato ótimo em função de um empate entre [*Output2*] e [*Output3*] sob tal restrição. A decisão, então, recai para a restrição C_2, que define [*Output2*] como o resultado da gramática. Uma vez que a decisão foi tomada pela restrição C_2, a restrição C_1 não exerce papel algum na escolha do candidato ótimo.

Na Gramática Harmônica, independentemente de seus pesos, todas as restrições exercem papel na escolha da forma de saída. Tais restrições, no momento de avaliação do candidato ótimo, não são convertidas em ranqueamentos estritos; permanecem, dessa forma, com seus valores numéricos, que desempenharão efetivo papel no cálculo que leva à escolha do candidato ótimo. Sob a HG, o candidato selecionado pela gramática é aquele que apresenta o maior valor numérico de Harmonia (H). Tal valor é obtido ao multiplicarmos cada marca de violação do candidato pelo valor da restrição violada e, após isso, somarmos todos os resultados dessas multiplicações referentes ao candidato em questão. O *tableau* em (7), elaborado com base no exemplo fornecido por Pater (2009), demonstra como se dá a escolha do candidato ótimo, à luz da HG.

(7)

	2	1	
	C_2	C_1	H
[*Output1*]	*		-2
[*Output2*]	*	*	-3
☞ [*Output3*]		*	-1
[*Output4*]		**	-2

Com base em Legendre, Sorace e Smolensky (2006), consideramos, no modelo em questão, que cada violação das restrições corresponde a um índice negativo no cálculo da harmonia do candidato. Visto o *tableau* acima, vemos que o candidato ótimo é [*Output3*], pelo fato de apresentar o índice mais alto de Harmonia, ou seja, o mais próximo de 0.

Ainda que o *tableau* em (7) consiga demonstrar a lógica de avaliação dos candidatos à luz da HG, poderíamos pensar, em princípio, que a adoção desse princípio de avaliação resultaria, em todos os casos, nos mesmos *outputs* ótimos obtidos através de uma avaliação via Teoria da Otimidade. De fato, ao considerarmos o *tableau* em (7) à luz da OT Estocástica, sob uma perspectiva de dominância estrita, o candidato ótimo também seria o mesmo. Isso ocorre porque, à luz da OT Estocástica, os pesos em questão seriam convertidos no *ranking* $C_2 >> C_1$. Dado esse *ranking*, a restrição mais alta excluiria os dois primeiros candidatos, sendo que a decisão cairia para a restrição C_1, que avaliaria como ótimo o candidato com menor número de violações. Em outras palavras, na Teoria da Otimidade, em função de um empate de candidatos na restrição mais alta, a restrição mais baixa exerce, também, papel na gramática.

Entretanto, o diferenciado funcionamento de EVAL à luz da Teoria da Otimidade Estocástica e da Gramática Harmônica pode resultar em diferentes *outputs* ótimos, em função do modelo adotado. Para evidenciarmos a diferença entre os dois modelos, apresentamos o *tableau* em (8), elaborado com base em Boersma e Pater (2008, p.27):

(8)

	1.5	1.0
	C_2	C_1
HG: ☞ [*Output1*]	*	
OT: ☞ [*Output2*]		**

Conforme vemos no *tableau* em (8), enquanto que uma avaliação à luz da noção de dominância estrita da OT leva à produção de [*Output2*], a submissão dos candidatos aos princípios de avaliação da Gramática Harmônica leva à emergência de [*Output1*]. A diferença reside no fato de o candidato [*Output2*], tomado como menos harmônico pela HG e ótimo pela OT, violar duas vezes a restrição C_1, que apresenta peso mais baixo no sistema. À luz da dominância estrita, a restrição C_1 não exerce efeito algum na eliminação dos candidatos para a escolha do ótimo, pois a decisão do *output* efetivamente produzido já foi feita pela restrição C_2. Entretanto, para a avaliação da HG, todas as restrições exercem efeito no cálculo do valor de Harmonia. A dupla violação de C_1 levou o candidato em questão a exibir um valor de Harmonia de -2, diferentemente do candidato [*Output1*] que apresenta uma Harmonia de -1.5, ou seja, mais próxima de 0. Fica claro, assim, o fato de que a adoção de um ou outro modelo de análise exerce implicações diretas na escolha do candidato ótimo.

Ainda no que diz respeito à avaliação dos candidatos à luz da Gramática Harmônica, precisamos mencionar o efeito cumulativo possível de ser exibido pelas restrições, na escolha do *output* ótimo. Para tal verificação, apresentamos o *tableau* em (9), que expressa tal efeito de cumulatividade[9].

---

[9] Para maiores detalhes acerca das vantagens da noção de cumulatividade, em sua capacidade de expressar efeitos de *output* que não seriam atingidos através da Teoria da Otimidade, aconselhamos a leitura de Pater (2009) e de Coetzee e Pater (2009).

(9)

	1.5	1	1	
	C_3	C_2	C_1	H
HG: ☞[*Output1*]	*			-1.5
OT: ☞ [*Output2*]		*	*	-2.0

Vemos, no *tableau* em (9), novamente um candidato ótimo que não seria selecionado sob a avaliação à luz dos preceitos da Teoria da Otimidade. Ainda que viole a restrição com maior peso, o candidato [*Output1*] é selecionado como ótimo, uma vez que [*Output2*] viola duas restrições que, ao terem seus pesos somados, levam a um valor de harmonia ainda mais baixo do que o apresentado por [*Output1*]. Encontramos em (9), portanto, a ação da cumulatividade de todas as restrições que exibem o fenômeno referido como "Efeito de Gangue" (PATER; JESNEY; TESSIER, 2006; JESNEY; TESSIER, 2007; BOERSMA; PATER, 2008; PATER, 2009; COETZEE; PATER, 2009).

A Gramática Harmônica também se encontra associada a um Algoritmo de Aprendizagem. A simulação computacional de tal Algoritmo encontra-se, também, disponibilizada através do *software Praat*[10]. O algoritmo em questão, assim como o GLA, apresenta um valor de ruído, que, somado aos pesos das restrições, pode dar conta da variação nas formas de *output*. Em função de tal característica, Boersma e Pater (2008) denominam tal algoritmo de HG-GLA, pelo fato de, assim como na versão do algoritmo OT-GLA de Boersma e Hayes (2001), expressar a gradualidade do processo de aquisição e as possíveis variações nas formas de *output*. No que diz respeito ao funcionamento, os dois algoritmos diferenciam-se nos procedimentos de incremento/decremento do valor das restrições: ao passo que, no OT-GLA, o acréscimo/decréscimo dos pesos se caracteriza pela adição de um valor *x* (correspondente à plasticidade, na simulação computacional) ao valor da restrição em questão, o procedimento de modificação dos índices numéricos das restrições no algoritmo da Gramática Harmônica considera, ainda, o número de violações que uma dada restrição sofre, de modo que uma restrição com um número maior de violações venha a sofrer acréscimos ou decréscimos de maior intensidade a cada rodada do algoritmo.

A variação nos padrões de *output* à luz do HG-GLA ocorre, dessa forma, quando dois candidatos apresentam valores de Harmonia bastante próximos. Nesses casos, é possível que, em função das alterações causadas pelo valor de

[10] De acordo com Coetzee e Pater (2009), o algoritmo em questão foi vinculado ao programa a partir da versão 5.0.01 no ano de 2007.

ruído, em alguns momentos de avaliação, o candidato [*Output1*] apresente uma harmonia superior ao de [*Output2*], enquanto que, em outros casos, o candidato [*Output2*] apresente uma harmonia maior que a de [*Output1*]. Veja-se o exemplo, a seguir, que ilustra o acima afirmado:

(10)[11]

Valores das restrições (centrais)	Ponto de seleção 1ª avaliação	Ponto de seleção 2ª avaliação
C_3: 103	102	100
C_2: 76	76	79
C_1: 24	24	26

1ª avaliação

	102	76	24	
	C_3	C_2	C_1	H
[*Output1*]	*			-102
☞ [*Output2*]		*	*	-100

2ª avaliação

	100	79	26	
	C3	C2	C1	H
☞ [*Output1*]	*			-100
[*Output2*]		*	*	-105

Vemos que o valor central da restrição C_3 (103) já se encontra bastante superior ao apresentado por C_2 (76). À luz da OT Estocástica, tal diferença numérica acentuada já seria suficiente para que pudéssemos determinar o ranqueamento categórico C_3>>C_2, a partir do qual [*Output2*] seria invariavelmente escolhido. Entretanto, para a Gramática Harmônica, o que interessa não é a distância de duas restrições específicas, mas sim a

11 Os valores das restrições em (10) foram por nós estipulados para fins de uma melhor exemplificação do argumento.

diferença no valor de Harmonia entre os candidatos. Ainda que o valor de C_3 esteja bastante afastado do de C_2, a soma dos valores de C_2 e C_1 encontra-se bastante próxima do valor de C_3. Dessa forma, não é surpreendente o fato de que, em função do ruído aplicado em cada um dos momentos de avaliação dos candidatos em algumas avaliações [*Output1*] apresente uma harmonia mais baixa do que [*Output2*], conforme vemos no *tableau* que representa a primeira avaliação, enquanto que o contrário ocorra em outros momentos de produção linguística. A possibilidade de variação, a partir do ruído estocástico aplicado, concretiza-se em função de os valores de harmonia dos candidatos, formados a partir dos "*efeitos de gangue*" das restrições, estarem bastante próximos. Dessa forma, para não haver variação nos padrões de *output*, é preciso que o valor de peso de C_3 se encontre suficientemente afastado do valor da **soma** de C_2 e C_1, uma vez que essas duas últimas restrições agem em conjunto na seleção do candidato ótimo.

Ao considerarmos a noção de cumulatividade, característica do modelo de Gramática Harmônica, questionamo-nos, pois, se restrições conjuntas precisam ser efetivamente propostas nesse modelo. Uma vez que, para a Gramática Harmônica, todas as restrições acabam exercendo papel na escolha do candidato ótimo, ao contribuírem com seus pesos no cálculo dos valores de Harmonia, julgamos necessário checar a possibilidade de obtermos um sistema de gramática que exiba o mesmo efeito da restrição conjunta sem necessariamente com ela contar. Uma vez que a formação de uma restrição conjunta deve sempre operar sob a premissa de não redundância (FUKAZAWA, 1999, 2001; FUKAZAWA; MIGLIO, 1998; BONILHA, 2003, 2005), caso venhamos a atingir os efeitos da ação conjunta de restrições através da cumulatividade, será possível dispensar o operador de conjunção de restrições '&', no modelo teórico em questão. A verificação de tal possibilidade bem como as suas implicações teóricas constituem questões a serem discutidas ao longo do presente trabalho.

As simulações computacionais

Com vistas a verificar a necessidade ou não de restrições conjuntas, também, sob o modelo da Gramática Harmônica, realizamos simulações computacionais à luz dos algoritmos de aprendizagem vinculados à OT Estocástica e à Gramática Harmônica, através do *software Praat* (BOERSMA; WEENINK, 2009). Sob ambos os algoritmos, alimentamos o sistema, primeiramente, com três restrições, C_1, C_2 (marcação) e C_3 (fidelidade). O sistema-alvo a ser atingido por ambos os modelos era o mesmo: uma língua que permitisse a violação ou de C_1 ou de C_2, mas não das duas restrições ao mesmo tempo. Para expressarmos a impossibilidade de atingirmos tal sistema via OT, retratamos, novamente, os *tableaux* que expressam o esquema de Conjunção Local previamente apresentados em (1) e (2):

(11)

	C_3	C_1	C_2
☞ a		*	
b	*!		
a_1	*!		
☞ b_1			*
☹ a_2	*!		
☞ b_2		*	*

Conforme aponta a literatura da OT, somente C_1, C_2 e C_3 não seriam capazes de levar a um sistema que leve à emergência do candidato 'a_2', de modo que uma restrição conjunta venha a ser necessária. Por outro lado, conforme também já afirmado, hipotetizamos que somente essas três restrições já são capazes de permitir que o algoritmo vinculado à HG consiga convergir no sistema que apresenta o comportamento desejado. São retratados, no que segue, os resultados das simulações desenvolvidas sob ambos os algoritmos.

Simulações com a Teoria da Otimidade

Conforme já dito, para a simulação via GLA, informamos o algoritmo da existência de três restrições, C_1, C_2 e C_3. Para determinarmos o valor inicial das restrições, determinamos que C_1 e C_2 teriam a função de restrições de marcação, ao passo que C_3 funcionaria como uma restrição de fidelidade Dessa forma, ao considerarmos que, no estágio inicial da aquisição da linguagem, Marcação >> Fidelidade (DEMUTH, 1995; LEVELT, 1995; PATER; PARADIS, 1996; SMOLENSKY, 1996a, 1996b; GNANADESIKAN, 2004; LEVELT; VAN DE VIJVER, 2004; DAVIDSON, JUSCZYK; SMOLENSKY, 2004), determinamos que C_1 e C_2 tivessem peso igual a 100, ao passo que C_3 apresentasse o valor inicial equivalente a 0.

Como sistema a ser adquirido, determinamos que os candidatos ótimos deveriam ser fiéis ao *input* (ou seja, não violar C_3), podendo violar C_1 ou C_2, mas nunca ambos. Se o candidato violasse ambas as restrições, o *output* infiel, que desrespeita C_3, deveria sagrar-se alvo.

Após alimentar o algoritmo com 100000 dados[12], os valores das restrições são os seguintes:

(12)

Valor de ranqueamento	
C_1	6408.527
C_2	6409.665
C_3	6408.514

Observamos, pelo alto valor de todas as restrições, que o algoritmo não foi capaz de convergir em um estágio final. Conforme explicam Pater, Jesney e Tessier (2007), a não convergência de um algoritmo revela-se através dos altíssimos valores numéricos assumidos pelas restrições. Conforme podemos ver nos resultados em (12), os valores das restrições C_1, C_2 e C_3 encontram-se na faixa de 6000, valores extremamente altos, sobretudo ao considerarmos que o estágio inicial de aquisição correspondia aos pesos de 100 e 0. Tais valores altos são explicados pelo fato de que, na busca pela gramática-alvo, o algoritmo tende a promover cada vez mais os valores das restrições sem nunca atingir o estágio final. De fato, se o algoritmo fosse rodado novamente, os valores numéricos de tais restrições seriam, ainda, mais altos. Pater, Jesney e Tessier (2007) ainda apontam como outro aspecto caracterizador da não convergência a grande proximidade entre os valores de todas as restrições, o que também pode ser visto em (12). Em outras palavras, vemos, nos resultados da simulação aqui feita, que, uma vez que o algoritmo associado à OT Estocástica não consegue chegar ao sistema-alvo sem restrições conjuntas, os valores das restrições componentes da gramática são cada vez mais aumentados, sem que tais restrições atinjam efetivos pesos que levem à gramática em questão.

Em suma, a não convergência do algoritmo demonstra a incapacidade, por parte da Teoria da Otimidade, de expressar o poder de restrições que, ainda que baixas no *ranking*, exercem ação cumulativa. Com apenas 3 restrições, C_1, C_2 e C_3, apenas as seguintes possibilidades de *output* são possíveis: *outputs* sempre fiéis, independentemente das violações sobre marcação; ou candidatos ótimos infiéis, que tendam a respeitar uma ou outra restrição de marcação; ou até ambas. Não é possível, com essas três restrições, expressar um sistema em que fidelidade apenas se faça valer caso uma de duas restrições de marcação (independentemente de qual delas) esteja sendo violada, sendo tal fidelidade, portanto, desrespeitada pelo candidato escolhido caso ambas as restrições de marcação apresentem violações.

[12] Em todas as simulações, no que diz respeito à definição do número de repetições e dos valores de plasticidade, plasticidade inicial e replicações por plasticidade, operamos com os valores *default* fornecidos pelo *software Praat*.

Estando clara a necessidade de evocação de uma restrição conjunta, realizamos uma outra simulação do processo de aquisição à Luz da OT Estocástica, em que acrescentamos uma outra restrição (C_4), que exerceria o papel de uma conjunção entre C_1 e C_2. Tal restrição também recebeu o valor inicial de 100, assim como as outras restrições que a compõem. Esperávamos garantir, dessa forma, a convergência para o mesmo sistema que havia sido previamente alimentado ao algoritmo: a língua admite *outputs* ótimos que violam ou C_1 ou C_2, mas não ambos.

Ao considerarmos que, segundo a Teoria da Conjunção Local, restrições conjuntas devem ser formadas apenas quando as suas restrições componentes já apresentarem um *status* mais baixo na gramática, um exercício interessante seria determinar, na definição do estágio inicial do algoritmo, que C_2 e C_1 já apresentam um valor inicial menor do que 100. Isso significaria dizer que a restrição conjunta C4 seria formada somente apenas os aprendizes receberem uma quantidade de *input* que já permitisse a demoção dos valores das restrições de marcação. Ressaltemos, entretanto, que tal implementação referente à determinação do estágio inicial no algoritmo não exerce efeito algum na definição do estágio final da gramática (pois, conforme testado, os valores finais sob ambas as possibilidades mantêm-se os mesmos), conforme apresentado em (13). Em função de tal constatação, bem como por fins de delimitação, definimos, no presente trabalho, valores iguais (100) para a caracterização dos pesos iniciais de C_1, C_2 e C_4[13].

Após informarmos o algoritmo acerca do mesmo sistema-alvo, os resultados numéricos fornecidos pelo OT-GLA bem como os tableaux resultantes desses valores são a seguir expressos:

(13)

C_4	100.00	101.091
C_3	75.500	75.823
C_2	63.440	67.213
C_1	61.060	61.459

13 No que diz respeito à Teoria de Gramática Harmônica, considerando-se o estágio inicial de aquisição em que as restrições de marcação apresentam pesos bastante superiores aos de fidelidade, podemos dizer que as estruturas marcadas somente começam a emergir após a demoção de marcação e a promoção de fidelidade. Ou seja, ainda que, em termos de mecanismos formais, o efeito de gangue esteja disponível desde o primeiro estágio de aquisição de linguagem, ele somente começa a se fazer sentir à medida que as crianças vão descobrindo os pesos corretos das restrições, através da promoção de fidelidade, motivada pela evidência positiva com exemplares de estruturas marcadas, movimentação essa prevista pelo algoritmo.

	101.091	75.823	67.213	61.459
	C_4	C_3	C_2	C_1
☞ a_2		*		
b_2	*!		*	*

Verificamos, na simulação em questão, a capacidade de o algoritmo convergir, isto é, encontrar um estado numérico final para as restrições que compõem a gramática. Conforme podemos ver em (13), C_3 apresenta o valor de 75.500, que se mostra mais de 10 pontos superior aos valores de C_1 (63.440) e C_2 (61.060). À luz do Algoritmo de Aprendizagem Gradual de Boersma e Hayes (2001), uma distância superior a 10 pontos nos valores de restrições indica que não haverá, nos diversos momentos de avaliação (produção linguística), a variação hierárquica entre as restrições que apresentam tal distância. Dessa forma, C_3 domina categoricamente C_1 e C_2, e permite que um *output* que viole uma das duas restrições venha a se sagrar como ótimo. C_1 e C_2, por apresentarem valores numéricos bastante próximos (63.440 e 61.060, respectivamente), podem variar em *status* hierárquico em diferentes momentos de produção linguística; tal fato, entretanto, não exerce efeitos no resultado final da gramática, uma vez que ambas já apresentam um valor bastante inferior ao de C_3.

A proibição à violação de duas restrições é expressa por C_4, que exerce o papel de restrição conjunta no sistema em questão. É importante notar que tal restrição não sofreu nenhuma promoção ou demoção por parte do algoritmo, uma vez que seu valor inicial já era altamente ranqueado. A restrição conjunta, com um valor bastante superior e afastado do das outras restrições, não permite que, em situação alguma, o candidato que viole as duas restrições de marcação, ao mesmo tempo, venha a emergir. De fato, no caso em questão, a restrição conjunta deve dominar tanto C_3 quanto as restrições individuais que a compõem, pré-requisito que se mostra plenamente obedecido após a aplicação do algoritmo[14].

Em suma, comprovamos, através da simulação à luz da Teoria da Otimidade Estocástica, uma questão já bastante discutida na literatura: restrições conjuntas são necessárias no modelo da Teoria da Otimidade, pelo fato de, na avaliação dos candidatos, as restrições não apresentarem ação cumulativa, por operarem sob a noção de dominância estrita. No que segue, contrastaremos tais resultados com os índices numéricos obtidos do algoritmo associado à Gramática Harmônica.

[14] É importante notar que tal condição foi obedecida em função da evidência positiva fornecida ao algoritmo. De fato, a aplicação computacional do algoritmo em si não diferencia C_4 das outras restrições; isso quer dizer que o algoritmo trata restrições conjuntas da mesma maneira que trata restrições simples. A determinação de C_4 como restrição conjunta é, na verdade, uma determinação extrínseca por parte do analista, uma vez que o mecanismo de conjunção local, conforme já dissemos, não se encontra incorporado ao algoritmo.

Simulações com a Gramática Harmônica

Uma vez que o modelo da Gramática Harmônica opera a partir da noção de restrições que agem cumulativamente para a determinação do índice de Harmonia de cada candidato, hipotetizamos que o algoritmo se mostrará capaz de convergir em uma gramática que, ao contar apenas com as restrições C_1, C_2 e C_3, permita *outputs* ótimos que violem ou C_1 ou C_2, mas não as duas juntas. Para a simulação computacional[15], foram fornecidas exatamente as mesmas informações que haviam sido providas ao OT-GLA, tanto no que diz respeito à descrição estrutural das restrições e ao estado inicial da gramática, quanto ao sistema-alvo a ser atingido. Os resultados da simulação, à luz do algoritmo associado à Gramática Harmônica, são apresentados em (14).

(14)

valor de ranqueamento	ponto de seleção
C_3 75.628	76.807
C_2 61.587	62.804
C_1 62.784	60.809

	C_3	C_2	C_1	Harmonia
☞ a		*		-62.804
B	*!			-76.807
a_1	*!			-76.806
☞ b_1			*	-60.809
☞ a_2	*!			-76.807
b_2		*	*	-123.612

Verificamos que o algoritmo conseguiu, mesmo com apenas três restrições, convergir no estágio final desejado em que os candidatos 'a', 'b_1' e 'a_2' sagram-se ótimos, a partir de uma única gramática. Isso se deve à ação cumulativa

[15] Dentre as opções de perfil de Gramática Harmônica fornecidos pelo Praat, foi selecionada a opção "Linear OT". Reconhecemos, entretanto, que, para a simulação em questão, na qual as restrições não atingem valores numéricos negativos, as opções "*Harmonic Grammar*" e "*Positive HG*" seriam igualmente eficientes. Para uma vantagem da opção "*Linear OT*" sobre as outras opções de modalidades de Gramática Harmônica oferecidas pelo *software Praat*, aconselhamos a leitura de Boersma e Pater (2008) e Coetzee e Pater (2009).

das restrições C_2 e C_1: na disputa entre 'a_2' e 'b_2', é o primeiro candidato que se sagra ótimo, uma vez que o seu valor de Harmonia ('a_2': - 76.807) é bem superior (mais próximo de zero) ao do candidato 'b_2', cuja harmonia (-123.612) é o resultado dos valores de ponto de seleção de C_2 e C_1, ou seja, da ação cumulativa de duas restrições. O valor de 'a_2' (-76.807), por sua vez, corresponde ao valor do ponto de seleção de C_3, única restrição violada por este candidato. Ressaltemos, ainda, que os valores de Harmonia de 'a_2' e 'b_2' se mostram bastante afastados, o que caracteriza, à luz da HG, o caráter categórico, invariável, da forma de saída.

A partir da observação de (14), cabe ressaltar, ainda, que o resultado obtido pela HG não se faria possível à luz da Teoria da Otimidade. Uma vez que a OT Estocástica converte os valores numéricos em *rankings* de restrições, que operam sob a noção de dominância estrita, 'a_2' nunca poderia vir a emergir como ótimo, no sistema em questão, pelo fato de violar a restrição mais altamente ranqueada. Sob a perspectiva da OT, C_1 e C_2 já não exercem papel algum, uma vez que C_3 já sofreu uma violação fatal. Por sua vez, para a HG, todas as restrições exercem efeito no cálculo da Harmonia. A partir da ação cumulativa de restrições (o que pôde ser visto em (14), através do "efeito de gangue" entre C_1 e C_2), é possível, ao contrário do que ocorre na OT, garantir a emergência de a_2.

Ainda que, na simulação que acabamos de retratar, já tenha sido sugerido o caráter redundante de uma possível restrição conjunta, considerando-se as restrições utilizadas na Gramática Harmônica, realizamos, também sob o algoritmo vinculado à HG, uma simulação que acrescentasse uma restrição conjunta ao sistema, assim como havíamos feito ao operarmos com a OT Estocástica. Novamente, os valores do sistema inicial bem como os da gramática-alvo foram os mesmos realizados na simulação à luz do modelo anterior. Os resultados dos valores das restrições são apresentados em (15).

(15)

	valor de ranqueamento	ponto de seleção
C_4	100	101.087
C_3	74.654	72.697
C_1	63.446	62.964
C_2	61.900	61.864

	C_4	C_3	C_1	C_2	Harmonia
☞ a				*	-61.864
B		*			-72.697
a_1		*			-72.697
☞ b_1			*		-62.964
☞ a_2		*			-72.697
b_2	*		*	*	-225.915

Também com a inclusão da restrição conjunta, o algoritmo consegue convergir para a gramática desejada. Entretanto, ao compararmos os resultados de ambas as simulações realizadas à luz da HG, fica bastante claro o caráter redundante da restrição conjunta representada por C_4, uma vez que a sua retirada do sistema não implica diferenças nos padrões de *output* que se sagram como ótimos. De fato, ao observarmos a gramática adquirida em (15), vemos que uma possível retirada de C_4 do sistema (o que levaria à gramática expressa na simulação anterior) não afetaria nenhum padrão de *output*. A restrição conjunta, dessa forma, exibe caráter redundante na Gramática Harmônica, em função do caráter cumulativo de todas as restrições, no estabelecimento dos valores de Harmonia de cada candidato a *output*.

Em suma, ao passo que a Conjunção de Restrições pode vir a se mostrar necessária na Teoria da Otimidade, o modelo da Gramática Harmônica, que não opera sob a noção de dominância estrita, dispensa a ação do mecanismo de conjunção local. Confirmada a hipótese de que a análise à luz da Gramática Harmônica dispensaria a formação da restrição conjunta, é preciso refletir a respeito das diferenças entre OT Estocástica e Gramática Harmônica na análise dos sistemas linguísticos. Tal reflexão teórica será desenvolvida no que segue.

Discussão e considerações finais

Neste trabalho, apresentamos os resultados de simulações computacionais do processo de aquisição de linguagem à luz de dois distintos modelos teóricos: a Teoria da Otimidade Estocástica e a Gramática Harmônica. Verificamos a capacidade de convergência, por parte de ambos os algoritmos, em sistemas em que um candidato fiel com violações em ambas as restrições de marcação (C_1 e C_2) não possa emergir como ótimo, ainda que a produção de candidatos fiéis que apresentem violações em apenas uma das restrições seja permitida. Para isso, realizamos simulações em cada um dos algoritmos sob duas condições: (a) simulações com apenas três restrições, C_1, C_2 (marcação) e C_3 (fidelidade); (b) simulações com quatro restrições, contando, portanto, com a restrição C_4, que corresponde à conjunção de C_1 e C_2.

Os resultados das simulações computacionais com os dois algoritmos demonstraram que, na condição experimental expressa em (a), apenas a Gramática Harmônica conseguiu chegar ao sistema-alvo. Conforme hipotetizávamos, a noção de dominância estrita, que rege a OT Estocástica, impede que obtenhamos um sistema que permita candidatos fiéis que incorram a violação ou de C_1, ou de C_2, mas não de ambas. Por outro lado, tal efeito é atingido através da ação conjunta das restrições na avaliação dos candidatos à luz da HG. Uma vez que, para a Gramática Harmônica, todas as restrições acabam exercendo papel na escolha do candidato ótimo, ao contribuírem com seus pesos no cálculo dos valores de Harmonia, conseguimos obter um sistema de gramática que exibe o mesmo efeito da restrição conjunta sem necessariamente com ela contar. Na OT Estocástica, por sua vez, tal condição só foi atingida sob a condição experimental expressa em (b): de fato, a gramática desejada só conseguiu ser atingida ao ser acrescentada a restrição C_4, que corresponderia à conjunção de C_1 e C_2.

Os resultados das simulações representam o início de uma discussão a respeito da necessidade de restrições conjuntas, bem como a ação do operador '&' sob ambos os modelos. Através das simulações realizadas à luz da HG-GLA, verificamos a capacidade de convergência na gramática esperada mesmo sem a existência de restrições conjuntas; de fato, à luz do HG-GLA, o acréscimo da restrição C_4 apresentou caráter completamente redundante, uma vez que o sistema já havia sido capaz de convergir apenas com as restrições C_1, C_2 e C_3.

Frente a tais constatações, devemos discutir a necessidade de considerarmos, como constituinte do quadro CON na Gramática Harmônica, o operador de conjunção local '&'. Uma vez que a formação de uma restrição conjunta deve sempre operar sob a premissa de não redundância (FUKAZAWA, 1999, 2001; FUKAZAWA; MIGLIO, 1998; BONILHA, 2003, 2005; ALVES, 2008), ou seja, não pode haver uma restrição conjunta que atinja exatamente os mesmos resultados que seriam alcançados por restrições simples, os efeitos de cumulatividade da HG dispensam o fenômeno de conjunção de restrições. Assim sendo, chegamos à conclusão de que a HG se diferencia da OT não somente no que diz respeito à ação de EVAL; de fato, os resultados das simulações apontam, também, para diferenças consideráveis no módulo CON de cada um dos modelos, uma vez que o operador '&' e todas as restrições passíveis de serem por ele formadas tornam-se desnecessárias na HG.

Acreditamos que as conclusões aqui apresentadas exercem consequências de caráter formal bastante importantes. Primeiramente, dispensar a ação de restrições conjuntas implica um menor número de restrições para a análise de modo a caracterizar a análise à luz da HG como mais econômica e formalmente simples. Mais do que isso, dispensar a ação do operador '&' pode vir a representar a resolução de um dos temas de discussão mais polêmicos da OT: a Conjunção Local. Conforme já expresso, até o presente momento dos avanços das análises em OT, parece não haver um consenso, na literatura, acerca dos critérios de formalização de restrições conjuntas, bem como das limitações que devem ser impostas sobre o mecanismo formador de tais restrições. A análise via HG, dessa forma, distancia-se de tal discussão, uma vez que não precisa recorrer a tal mecanismo.

Mais estudos são necessários a partir de tais constatações, para que possamos discutir a natureza de CON sob a Gramática Harmônica. Nesse sentido, acreditamos que as diferenças entre os dois modelos possam ser mais bem entendidas também a partir do estudo das teorias de base cognitiva que servem para o estabelecimento de cada uma das propostas teóricas. Ao considerar os pesos de todas as violações em cada uma das restrições, a Gramática Harmônica retoma os pressupostos das simulações conexionistas voltadas à aquisição linguística. Assim, no que concerne à avaliação de candidatos, a HG abandona um dos aspectos de caráter mais fortemente simbólico da Teoria da Otimidade: a dominância estrita. Por sua vez, os efeitos de gangue e ações cumulativas da HG retomam a noção de um processamento distribuído, em que os diversos pesos atuam em conjunto na escolha da forma de *output* efetivamente produzida. De fato, uma vez que tal modelo opera sob a noção de pesos, os sistemas harmônicos da HG podem ser facilmente simulados em redes conexionistas, de modo a constituírem "[...] um mediador entre as descrições dos níveis mais altos de cognição, que caracterizam a Teoria da Otimidade plenamente simbólica, e os níveis mais baixos de cognição, definidos puramente em termos de redes conexionistas." (LEGENDRE; SORACE; SMOLENSKY, 2006, p.339).

Conforme verificado no presente trabalho, diferentes maneiras de avaliar o candidato ótimo têm implicações sobre o próprio quadro de restrições possíveis de ser adotadas em cada um dos modelos. A HG, nesse sentido, dispensa restrições conjuntas, constatação de grande relevância ao procedermos a um levantamento das diferenças entre essas duas propostas teóricas. Dadas as diferenças não somente nos procedimentos de análise, mas nas teorias de base e nos componentes que caracterizam as restrições, verificamos, dessa forma, uma instigante agenda de pesquisas futuras que, ao se voltar para as implicações do emprego de uma ou outra Teoria, mostra-se de suma importância para os pesquisadores voltados ao entendimento dos modelos de análise linguística. Em suma, as diferentes formas de avaliação do candidato ótimo à luz da OT e da HG representam mais do que mecanismos procedimentais de análise. Elas refletem, também, as concepções de "cognição" e de "gramática" que servem como alicerces teóricos para o desenvolvimento de tais propostas.

ALVES, U. K. Optimality theory, harmonic grammar, and conjoined constraints. *Alfa*, São Paulo, v.54, n.1, p.237-263, 2010.

- *ABSTRACT: Optimality Theory (OT), in both Standard (PRINCE; SMOLENSKY, 1993) and Stochastic (BOERSMA; HAYES, 2001) versions, operates under Strict Domination. It differs from Harmonic Grammar (LEGENDRE, MIYATA; SMOLENSKY, 1990; SMOLENSKY; LEGENDRE, 2006), in which the choice of the optimal output is based on the cumulative effect obtained from the ranking values of all violated constraints. As we consider the role of this "gang effect" in Harmonic Grammar (HG), we inquire whether conjoined constraints are really necessary under this framework. To answer this question, we ran computer simulations with the OT and HG learning algorithms in Praat (BOERSMA; WEENKINK, 2009). The results provided by the two algorithms show that the HG model is able to converge to systems that would only be achievable, under OT, via conjoined constraints. These results, in turn, motivate a discussion on the role of Local Conjunction in HG, besides shedding light on the debate on the implications of using either OT or HG.*
- *KEYWORDS: Optimality Theory. Harmonic Grammar. Local Conjunction. Conjoined Constraints. Learning algorithms.*

REFERÊNCIAS

ALVES, U. K. A epêntese vocálica na aquisição das plosivas finais do inglês (L2): tratamento pela OT Estocástica e pela Gramática Harmônica. In: SIMPÓSIO SOBRE VOGAIS, 2., 2009, Belo Horizonte. *Anais*... Belo Horizonte: UFMG: SIS-VOGAIS, 2009. Disponível em: <http://relin.letras.ufmg.br/probravo/pdf_sisvogais/parlato.pdf/> Acesso em: 19 ago. 2009.

_______. *A aquisição das sequências finais de obstruintes do inglês (L2) por falantes do Sul do Brasil*: análise via Teoria da Otimidade. 2008. 335f. Tese (Doutorado em Letras) – Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2008.

BAKOVIĆ, E. *Harmony, dominance and control*. 2000. 327f. Ph.D. Dissertation (Doctor of Linguistics) – Rutgers University, New Jersey, 2000.

_______. Assimilation to the unmarked. In: ALEXANDER, J.; HAN, N.-R.; FOX, M. M. (Ed.). *Proceedings of the 23rd annual penn linguistics colloquium.* Philadelphia: University of Pennsylvania, 1999. p.1-16.

BOERSMA, P.; HAYES, B. Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry,* Cambridge, v.32, n. 1, p. 45-86, winter 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. Disponível em: <http://www.fon.hum.uva.nl/praat/> Acesso em: 19 ago. 2009.

BONILHA, G. F. G. *Aquisição fonológica do português brasileiro*: uma abordagem conexionista da Teoria da Otimidade. 2005. 389f. Tese (Doutorado em Letras) – Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 2005.

______. Conjoined constraints and phonological acquisition. *Journal of Portuguese Linguistics*, Lisboa, v.2, n.2, p.7-30, 2003.

COETZEE, A. W.; PATER, J. The place of variation in phonological theory. In: GOLDSMITH, J.; RIGGLE, J.; YU, A. (Org.). *The handbook of Phonological Theory*. 2. ed. Malden: Blackwell, 2009. p.1-38.

DAVIDSON, L.; JUSCZYK, P.; SMOLENSKY, P. The initial and final states: theoretical implications and experimental explorations of Richness of the Base. In: KAGER, R.; PATER, J.; ZONNEVELD, W. (Ed.). *Fixing priorities*: constraints in phonological acquisition. Cambridge: Cambridge University Press, 2004. p.321-368.

DEMUTH, K. Markedness and the development of prosodic structure. In: THE NORTH-EAST LINGUISTICS SOCIETY, 1995, Amherst. *Proceedings*... Amherst: GLSA Publications, 1995. v.25, p.13-25.

FUKAZAWA, H. Local conjunction and extending Sympathy Theory: OCP effects in Yucatec Maya. In: LOMBARDI, L. *Segmental Phonology in Optimality Theory*: constraints and representations. Cambridge: Cambridge University Press, 2001. p.231-260.

______. *Theoretical implications of OCP effects on features in Optimality Theory*. 1999. 296 f. Ph.D. Dissertation (Doctor of Linguistics) – University of Maryland, College Park, Maryland, 1999.

FUKAZAWA, H.; MIGLIO, V. Restricting conjunction to constraint families. In: SAMIIAN, V. (Ed.). *Proceedings of WECOL'96*. Fresno: California State University, 1998. p.102-117.

GNANADESIKAN, A. Markedness and faithfulness constraints in child phonology. In: KAGER, R.; PATER, J.; ZONNEVELD, W. *Constraints in Phonological Acquisition*. Cambridge: Cambridge University Press, 2004. p.73-108.

GOLDRICK, M.; DALAND, R. Linking speech errors and phonological grammars: insights from Harmonic Grammar networks. *Phonology*, Cambridge, v.26, p.147-185, 2009.

ITÔ, J.; MESTER, A. Markedness and Word Structure: OCP Effects in Japanese. *Rutgers Optimality Archive*, New Brunswick, n.255, 1998. Disponível em: <www.roa.rutgers.edu>. Acesso em: 20 jul. 2007.

JESNEY, K.; TESSIER, A. Re-evaluating learning biases in Harmonic Grammar. In: BECKER, M. (Org.). *University of Massachusetts Ocasional 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. (Org.). REGIONAL MEETING OF THE CHICAGO LINGUISTIC SOCIETY, 26, 1990, Chicago. *Proceedings*... Chicago: Chicago Linguistic Society, 1990. p.1-16.

LEGENDRE, G; SORACE, A.; SMOLENSKY, P. The Optimality-Theory – Harmonic Grammar Connection. In: SMOLENSKY, P.; LEGENDRE, G. *The Harmonic Mind*. Boston: MIT, 2006. p.339-402.

LEVELT, C. C. *Unfaithful kids*: place of articulation patterns in early vocabularies. Maryland: University of Maryland, 1995.

LEVELT; C. C.; VAN de VIJVER, R. Syllable types in cross-linguistic and developmental grammars. In: KAGER, R.; PATER, J.; ZONNEVELD, W. *Constraints in Phonological Acquisition*. Cambridge: Cambridge University Press, 2004. p.204-218.

LUBOWICZ, A. Restricting local conjunction. In: OLD WORLD CONFERENCE IN PHONOLOGY 2, 2007, Norway. Disponível em: <http://wwwrcf.usc.edu/~lubowicz/docs/ocp-2-handoutwebpage.pdf. 2005>. Acesso em: 22 jul. 2007.

______. Locality of Conjunction. In: ALDERETE, J.; HAN, C.; KOCHETOV, A. (Org.). WEST COAST CONFERENCE ON FORMAL LINGUISTICS 24, 2006, Somerville. *Proceedings...* Somerville: Cascadilla Press, 2006. p.254-262.

______. Derived environment effects in Optimality Theory. *Lingua*, Amsterdam, v.112, n.4, p.243-280, 2002.

MORETON, E.; SMOLENSKY, P. Typological consequences of local constraint conjunction. In: WEST COAST CONFERENCE ON FORMAL LINGUISTICS, 21., 2002, Cambridge. *Proceedings...* Cambridge: Cascadilla Press, 2002.

PATER, J. Weighted constraints in Generative Linguistics. *Cognitive Science*, Norwood, v.33, p.999-1035, 2009.

______. Gradual learning and convergence. *Linguistic Inquiry*, Cambridge, v.39, n.2, p. 334-345, 2008.

______. Non-convergence in the GLA and variation in the CDA. *Rutgers Optimality Archive*, New Brunswick, n.780, 2005. Disponível em: <http://roa.rutgers.edu>. Acesso em: 22 jul. 2007.

PATER, J.; JESNEY, K.; TESSIER, A. Phonological acquisition as Weighted Constraint Interaction. In: BELIOKOVA, A.; MERONI, L.; UMEDA, M. (Ed.). CONFERENCE ON GENERATIVE APPROUCHS TO LANGUAGE ACQUISITION– NORTH AMERICA, 2., 2007. *Proceedings*... Sommerville: Cascadilla Proceedings Project, 2007.

PATER, J; PARADIS, J. Truncation without templates in child phonology. In: STRINGFELLOW, A.; CAHANA-AMITAY, D.; HUGHES, E.; ZUKOWSKI, A. (Org.). ANNUAL BOSTON UNIVERSAL CONFERENCE ON LANGUAGE DEVELOPMENT, 20., Sommerville, 1996. *Proceeding*... Somerville: Cascadilla Press, 1996.

PRINCE, A; SMOLENSKY, P. *Optimality Theory:* constraint interaction in generative grammar. Baltimore: The Johns Hopkins University, 1993.

SMOLENSKY, P. On the internal structure of the constraint component Con of UG. *Rutgers Optimality Archive,* New Brunswick, n.118, 1996a. Disponível em: <http://roa.rutgers.edu>. Acesso em: 29 jul. 2007.

______. The Initial State and 'Richness of the Base' in Optimality Theory. *Rutgers Optimality Archive*, New Brunswick, n.118, 1996b. Disponível em: <http://roa.rutgers.edu>. Acesso em: 22 jul. 2007.

SMOLENSKY, P.; LEGENDRE, G. *The harmonic mind*: from neural computation to Optimality-Theoretic grammar. Cambridge: MIT, 2006.

TESAR, B.; SMOLENSKY, P.. *Learnability in Optimality Theory*. Cambridge: MIT, 2000.

______. Learnability in Optimality Theory. *Linguistic Inquiry*, Cambridge, v.29, p.229-268, 1998.

______. *Learnability in Optimality Theory*: long version. *Rutgers Optimality Archive,* New Brunswick, n.156, 1996. Disponível em: <http://roa.rutgers.edu>. Acesso em: 22 jul. 2007.

______. The learnability of Optimality Theory. In: ARANOVICH, R. et al. *Proceedings of the Thirteenth West Coast Conference on Formal Linguistics.* Stanford: CSLI, 1993. p.122-137,

TESSIER, A. *Biases and Stages in Phonological Acquisition*. 2007. 305f. Ph.D. Dissertation (Doctor of Linguistics) – University of Massachusetts, Amherst, 2007.

Recebido em setembro de 2009.
Aprovado em dezembro de 2009.