GRAMÁTICA CATEGORIAL ATRAVÉS DE ESTRUTURA DE CARACTERÍSTICAS

Luiz Arthur Pagani *

Introdução

Apesar de não ser um modelo de análise lingüística tão conhecido como a Gramática Gerativa, ou mesmo como alguns paradigmas alternativos adotados em áreas específicas (como a *Head-Driven Phrase Structure Grammar* e a *Lexical Functional Grammar*, na Lingüística Computacional; ou a Gramática Funcional, na Lingüística Antropológica), a Gramática Categorial (GC) tem atraído alguns lingüistas ao oferecer uma relação transparente entre as operações sintáticas e suas respectivas interpretações semânticas. Além disso, a GC ainda dispõe de boas ferramentas para a indagação sobre a ontologia categorial necessária à investigação lingüística; ou seja, ao contrário da Gramática de Estrutura Sintagmática (GES), na GC as categorias não precisam ser arbitrariamente estipuladas, já que elas podem ser recursivamente definidas a partir de algumas poucas categorias básicas.

Assim, um dos objetivos do presente artigo é o de fazer uma breve apresentação da GC, principalmente em relação à mencionada transparência entre análise sintática e interpretação semântica. No entanto, além de

* Universidade Federal do Paraná.

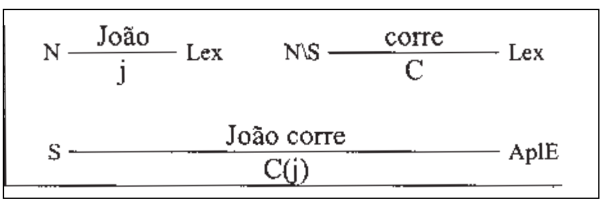
apresentarmos as derivações no chamado estilo de Prawitz, que tem sido amplamente adotado pela GC, introduziremos também uma outra maneira de se representar as derivações: por meio da chamada Estrutura de Características (EC). Além de representar as mesmas informações derivacionais do estilo de Prawitz, a EC ainda permite unificar num só ambiente representacional todas as etapas do processo da análise gramatical: todo o léxico, todas as regras categorias e, como dissemos, a própria análise gramatical. Dessa forma, a única operação empregada para construir uma análise gramatical é a unificação.

Gramática Categorial e Diagrama de Prawitz

Modelo AB

Numa das primeiras versões do modelo AB (assim chamado em homenagem a seus precursores: Ajdukiewicz e Bar-Hillel), a GC dispunha apenas de duas categorias básicas e uma única regra de aplicação sem direcionalidade (Ajdukiewicz, 1935). Um exemplo desse tipo de análise gramatical pode ser visto no diagrama 1. (Na verdade, no diagrama 1 já estamos empregando a notação direcional; o sistema categorial de Ajdukiewicz não era direcional porque seu objeto era a notação polonesa para o cálculo de predicados, onde a expressão funcional sempre aparece mais à esquerda seguida por seus argumentos, de forma que a questão da direcionalidade não se aplica).

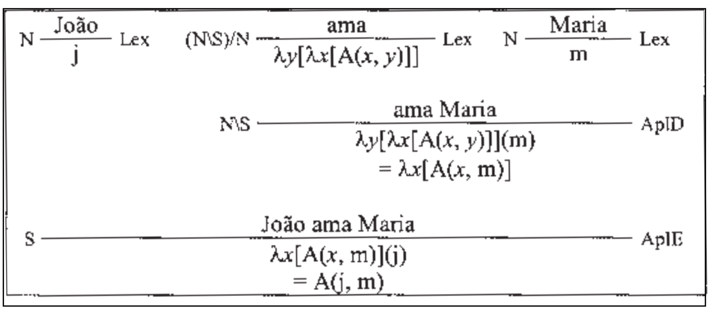
DIAGRAMA 1 - DERIVAÇÃO DE "JOÃO CORRE" COM REGRA DE APLICAÇÃO PARA A ESQUERDA



Como se pode observar pelo índice ao lado esquerdo das linhas horizontais sob as expressões, a categoria de "corre" é complexa e representa uma função que toma o nome "João" ("N") como argumento, resultando numa sentença ("S"). Pelo lado direito das mesmas linhas horizontais, sabemos que "João" e "corre" são itens lexicais ("Lex"), e que a sentença "João corre" é construída a partir dessas duas expressões por meio da regra de aplicação ("AplE"). Finalmente, adotando-se uma semântica denotacional, a representação da denotação de cada expressão aparece logo abaixo das respectivas linhas horizontais; assim, como de costume, "João" denota um determinado indivíduo (representado aqui por "j") e "corre" denota uma função característica de indivíduos a valores de verdade (ou seja, uma função que aplicada a indivíduos resulta em um valor de verdade, representada aqui por "C"), a denotação de "João corre" é um valor de verdade, também resultado da aplicação da denotação de "corre" à denotação de "João" (representada aqui por uma fórmula do cálculo de predicados: "C(j)").

Mais tarde, Bar-Hillel (1953) constatou a necessidade de distinguir duas direções de aplicação, já que nas línguas naturais alguns predicados buscam seus argumentos de ambos os lados. Ainda tomando um verbo como exemplo, se quisermos manter a análise tradicional de que ele toma primeiro um objeto para depois se relacionar com seu sujeito, sua categoria precisa ser "(N\S)/N". Assim, a análise da sentença "João ama Maria" ficaria como no diagrama 2.

DIAGRAMA 2 - DERIVAÇÃO DE "JOÃO AMA MARIA" COM REGRAS DE APLICAÇÃO PARA AMBOS OS LADOS



Nesse outro diagrama, o verbo "ama" é do tipo mencionado acima, e denota uma relação entre dois indivíduos. Para mantermos a versão tradicional do cálculo de predicados, introduzimos na representação denotacional o operador lambda, que adapta a ordem de combinação dos argumentos às posições argumentais do predicado (em muitas versões da CG, a opção é a de alterar a ordem das posições argumentais, de forma que a mesma fórmula seria reescrita como "(A m) j)"; essa mudança não afeta os predicados unários, mas se quisermos manter o mesmo padrão, a denotação de "corre" também pode ser representada por "l*x*[C(*x*)]" (conseqüentemente, a denotação de "João corre" seria "l*x*[C(*x*)](j)", que depois de reduzida daria o mesmo "C(j)"). Por uma questão de exaustividade, as operações de redução do operador lambda também estão representadas no diagrama.

Essas regras de aplicação para a direita (*forward*) e para a esquerda (*backward*) foram esquematicamente representadas por Wood (1993, p. 36) pelas seguintes fórmulas:

1.	X/Y:*f*	Y:*a*	→	X:*f(a)*
2.	Y:*a*	Y\X:*f*	→	X:*f(a)*

Na aplicação à direita (AplD, fórmula 1), uma expressão de categoria complexa "X/Y", cuja denotação é a função "*f*", seguida de uma expressão de categoria "Y", cuja denotação é o indivíduo "*a*", resulta numa expressão de categoria "X", cuja denotação é "*f(a)*" (a aplicação da função *f* ao indivíduo *a*). Com a aplicação à esquerda (AplE, fórmula 2), a função e o argumento trocam de lugares entre si (o que resulta na mudança do conectivo categorial), mas o resultado continua o mesmo: a categoria e a denotação da expressão resultante ainda é aplicação da categoria e da denotação da expressão funcional na categoria e na denotação da expressão argumental.

Com a comprovação de que esse modelo gramatical apresentava equivalência fraca com a gramática de estrutura sintagmática (Bar-Hillel; Gaifman; Shamir, 1960), a GC perdeu um pouco de interesse, até que a difusão do modelo de Lambek (1958), que acrescentava ao modelo AB mais uma regra binária (composição), além de outras duas regras unárias (permutação e promoção) renovou-se a curiosidade pelas novas possibilidades oferecidas (principalmente as computacionais).

Modelo clássico

Nesse novo modelo proposto por Lambek (também conhecido como CG clássica), ainda mais do que no outro, a noção estruturalista de constituência é substituída pela noção de conexidade. Segundo essa nova noção, “dentro de uma seqüência legal, qualquer subseqüência é ‘conexa’, ou seja, é um constituinte, numa determinada derivação, se, em algum ponto desta derivação, ela estiver rotulada por uma única categoria” (Wood, 1993, p. 23). Uma das principais vantagens da noção de conexidade em relação à GES é que se torna desnecessário declarar os símbolos iniciais da gramática.

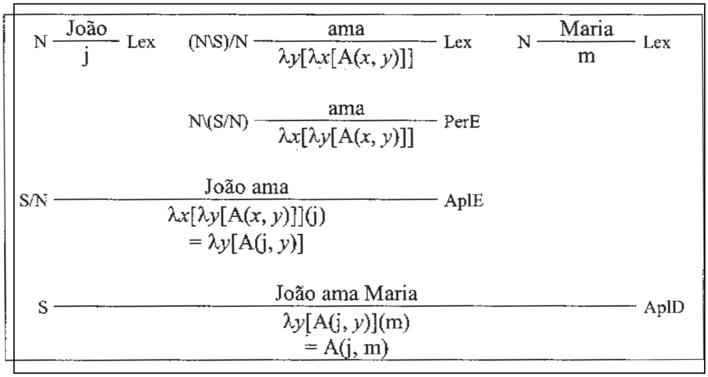
Usando a mesma representação esquemática, Wood (1993, p. 37) apresenta a regra da permutação (*swapping*) através das seguintes fórmulas:

3. $X\backslash(Y/Z): \lambda v_x[\lambda v_z[f(v_x, v_z)]] \rightarrow (X\backslash Y)/Z: \lambda v_z[\lambda v_x[f(v_x, v_z]]$
4. $(X\backslash Y)/Z: \lambda v_z[\lambda v_x[f(v_x, v_z)]] \rightarrow X\backslash(Y/Z): \lambda v_x[\lambda v_z[f(v_x, v_z]]$

Normalmente, essas regras são mais conhecidas como regras de associatividade; no entanto, segundo Wood (1993, p. 37; que, por sua vez, atribui essa ressalva a Oehrle, sem mencionar referências), “a associatividade é caracteristicamente definida como uma propriedade de um único operador, mas esta é uma relação entre dois operadores” (os conectivos à direita e à esquerda). Essas regras de permutação (PerD e PerE) apenas invertem a ordem na qual um funtor se combina com seus argumentos. Assim, na fórmula 3 (a de permutação à direita), um funtor que tomava primeiro um argumento à sua esquerda e depois outro à sua direita é transformado num funtor que toma seu primeiro argumento pela direita e depois o segundo pela esquerda; já na fórmula 4, acontece o inverso: um funtor que se combinava primeiro pela esquerda e depois pela direita passa a se combinar primeiro pela direita e só depois pela esquerda. Como antes, os operadores lambda são responsáveis apenas pela readequação do ordenamento dos argumentos às posições argumentais do funtor.

Por intermédio da regra de permutação é possível chegarmos a uma derivação incremental (na qual cada expressão básica é imediatamente integrada à análise, sem precisar ser combinada com nenhuma outra expressão posterior) para a sentença “João ama Maria”, que já recebeu uma análise não incremental no diagrama 2.

DIAGRAMA 3 - DERIVAÇÃO INCREMENTAL DE “JOÃO AMA MARIA” COM REGRAS DE PERMUTAÇÃO E DE APLICAÇÃO



No diagrama 3, a expressão “João ama” é uma expressão conexa, mesmo que não seja um constituinte numa GES. E mais, ela é uma expressão com denotação: ela denota o conjunto dos indivíduos que são amados por João. Quanto à derivação, ao contrário do que acontece no diagrama 2, no qual “ama” só pode ser integrado à análise depois de ter sido combinado a “Maria”, depois de permutar a ordem de combinação de “ama” (“PerE”), esse verbo pode se juntar primeiro a “João” e depois então a “Maria”, de forma que seu resultado final tanto em relação à categoria quanto à denotação é o mesmo obtido antes.

Com a introdução das regras de composição (ComD e ComE), também será possível oferecer uma derivação incremental para a análise de “um menino ama Maria”, como se pode ver no diagrama 4. Esquematicamente, essas regras são representadas pelas seguintes fórmulas (Wood, 1993, p. 39):

5. $X/Y{:}\,f \qquad Y/Z{:}\,g \qquad \rightarrow \qquad X/Z{:}\ \lambda v_z[f(g(v_z))]]$
6. $Z\backslash Y{:}\,g \qquad Y\backslash X{:}\,f \qquad \rightarrow \qquad Z\backslash X{:}\ \lambda v_z[f(g(v_z))]]$

Segundo essas fórmulas, duas expressões podem ser combinadas quando suas categorias forem ambas funtoras e uma delas buscar um argumento do mesmo tipo que o resultado da outra; essa combinação resulta na composição funcional das denotações de cada uma das expressões.

No diagrama 4, a expressão “um menino” é formada através de uma aplicação à direita, de forma a resultar num quantificador generalizado, como tem sido feito depois de Montague; para isso, as categorias e as denotações de “um” e de “menino” foram postuladas para permitir esse resultado (as denotações são as mesmas que se encontram, por exemplo, em Dowty, Wall e Peters (1981, p. 105-110) – com uma pequena adaptação apenas na de “menino”; em relação às categorias, optou-se por atribuir a “menino” a categoria “Nc”, mas também poderíamos tê-la atribuído à categoria “N” (o que ainda permitiria uma composição, entre “menino” e a expressão subseqüente, que a escolha categorial feita aqui não permite) – com isso, a categoria de “um” pode ser deduzida como sendo “(S/(N\S))/Nc”, já que a categoria do quantificador generalizado é “(S\(N/S)”).

DIAGRAMA 4 - DERIVAÇÃO INCREMENTAL DE “UM MENINO AMA MARIA” COM AS REGRAS DE COMPOSIÇÃO E DE APLICAÇÃO

$$
{}^{(S/(N\backslash S))/Nc}\frac{\text{um}}{\lambda P[\lambda Q[\exists x[P(x)\ \&\ Q(x)]]]}\,{}^{\text{Lex}}
\qquad \text{Nc}\ \frac{\text{menino}}{\lambda y[\mathrm{M}(y)]}\ \text{Lex}
\qquad {}^{(N\backslash S)/N}\frac{\text{ama}}{\lambda w[\lambda z[\mathrm{A}(z, w)]]}\,{}^{\text{Lex}}
\qquad \text{N}\ \frac{\text{Maria}}{\text{m}}\ \text{Lex}
$$

$$
\text{S/(N\\S)}\ \frac{\text{um menino}}{\begin{array}{c}\lambda P[\lambda Q[\exists x[P(x)\ \&\ Q(x)]]](\lambda y[\mathrm{M}(y)]) \\ = \lambda Q[\exists x[\lambda y[\mathrm{M}(y)](x)\ \&\ Q(x)]] \\ = \lambda Q[\exists x[\mathrm{M}(x)\ \&\ Q(x)]]\end{array}}\ \text{AplD}
$$

$$
\text{S/N}\ \frac{\text{um menino ama}}{\begin{array}{c}\lambda w[\lambda Q[\exists x[\mathrm{M}(x)\ \&\ Q(x)]](\lambda z[\mathrm{A}(z, w)])] \\ = \lambda w[\exists x[\mathrm{M}(x)\ \&\ \lambda z[\mathrm{A}(z, w)](x)]] \\ = \lambda w[\exists x[\mathrm{M}(x)\ \&\ \mathrm{A}(x, w)]]\end{array}}\ \text{ComD}
$$

$$
\text{S}\ \frac{\text{um menino ama Maria}}{\begin{array}{c}\lambda w[\exists x[\mathrm{M}(x)\ \&\ \mathrm{A}(x, w)]](\mathrm{m}) \\ = \exists x[\mathrm{M}(x)\ \&\ \mathrm{A}(x, \mathrm{m})]\end{array}}\ \text{AplD}
$$

No entanto, convém observar que mesmo sem as regras de composição seria possível derivar a análise dessa mesma sentença, ainda que não incrementalmente, recorrendo apenas às regras de aplicação, como se vê no diagrama 5. Neste diagrama, duas aplicações à direita juntam, respectivamente “um” com “menino” e “ama” com “Maria”; depois, uma terceira aplicação à direita concatena “um menino” e “ama Maria”. O resultado, porém, tanto categorial quanto denotacional, é o mesmo que o obtido na derivação do diagrama 4. Ainda que alguns vejam nessa multiplicidade de opções uma desvantagem, outros apreciam essa característica da GC por permitir uma aproximação mais simples entre interpretação semântica e estrutura prosódica, por exemplo.

DIAGRAMA 5 - DERIVAÇÃO NÃO INCREMENTAL DE “UM MENINO AMA MARIA” APENAS COM REGRAS DE APLICAÇÃO

$$\text{(S/(N\backslash S))/Nc}\ \frac{\text{um}}{\lambda P[\lambda Q[\exists x[P(x)\ \&\ Q(x)]]]}\ \text{Lex} \qquad \text{Nc}\ \frac{\text{menino}}{\lambda y[\mathrm{M}(y)]}\ \text{Lex} \qquad \text{(N\backslash S)/N}\ \frac{\text{ama}}{\lambda w[\lambda z[\mathrm{A}(z, w)]]}\ \text{Lex} \qquad \text{N}\ \frac{\text{Maria}}{\mathrm{m}}\ \text{Lex}$$

$$\text{S/(N\backslash S)}\ \frac{\text{um menino}}{\begin{array}{l}\lambda P[\lambda Q[\exists x[P(x)\ \&\ Q(x)]]](\lambda y[\mathrm{M}(y)]) \\ = \lambda Q[\exists x[\lambda y[\mathrm{M}(y)](x)\ \&\ Q(x)]] \\ = \lambda Q[\exists x[\mathrm{M}(x)\ \&\ Q(x)]]\end{array}}\ \text{AplD} \qquad \text{N\backslash S}\ \frac{\text{ama Maria}}{\begin{array}{l}\lambda w[\lambda z[\mathrm{A}(z, w)]](\mathrm{m}) \\ = \lambda z[\mathrm{A}(z, \mathrm{m})]\end{array}}\ \text{AplD}$$

$$\text{S}\ \frac{\text{um menino ama Maria}}{\begin{array}{l}\lambda Q[\exists x[\mathrm{M}(x)\ \&\ Q(x)]](\lambda z[\mathrm{A}(z, \mathrm{m})]) \\ = \exists x[\mathrm{M}(x)\ \&\ \lambda z[\mathrm{A}(z, \mathrm{m})](x)] \\ = \exists x[\mathrm{M}(x)\ \&\ \mathrm{A}(x, \mathrm{m})]\end{array}}\ \text{AplD}$$

Finalmente, o último par de regras a ser apresentado aqui será o de promoção (ProD e ProE). Na representação esquemática de Wood (1993, p. 42), as regras de promoção são expressas pelas seguintes fórmulas:

7. X: a → Y/(X\Y): $\lambda v[v(a)]$
8. X: a → (Y/X)\Y: $\lambda v[v(a)]$

Por meio destas regras, é possível chegarmos a uma segunda derivação incremental para “João ama Maria”, sem que seja preciso recorrer à permutação dos argumentos de “ama”, como no diagrama 3. No diagrama 6, abaixo, a promoção à direita de “João” resulta na categoria “S/(N\S)” e na transformação de sua denotação (“j”) num quantificador generalizado (“l*P*[*P*(j)]”). Como o denominador da categoria resultante dessa promoção e o numerador da categoria de “ama” são iguais, a regra de composição pode ser empregada; assim, forma-se a expressão “João ama”, cuja categoria é “S/N” e cuja denotação (depois de duas reduções) corresponde ao conjunto dos indivíduos que João ama (ou, alternativamente, à propriedade de ser amado por João). O último passo da derivação é exatamente igual ao do diagrama 3.

DIAGRAMA 6 - DERIVAÇÃO INCREMENTAL ALTERNATIVA PARA “JOÃO AMA MARIA” COM AS REGRAS DE PROMOÇÃO E DE COMPOSIÇÃO

$$\mathrm{N}\ \frac{\text{João}}{\mathrm{j}}\ \mathrm{Lex} \qquad \mathrm{(N\backslash S)/N}\ \frac{\text{ama}}{\lambda y[\lambda x[\mathrm{A}(x, y)]]}\ \mathrm{Lex} \qquad \mathrm{N}\ \frac{\text{Maria}}{\mathrm{m}}\ \mathrm{Lex}$$

$$\mathrm{S/(N\backslash S)}\ \frac{\text{João}}{\lambda P[P(\mathrm{j})]}\ \mathrm{ProD}$$

$$\mathrm{S/N}\ \frac{\text{João ama}}{\lambda y[\lambda P[P(\mathrm{j})](\lambda x[\mathrm{A}(x, y)])]}\ \mathrm{ComD}$$
$$= \lambda y[\lambda x[\mathrm{A}(x, y)](\mathrm{j})]$$
$$= \lambda y[\mathrm{A}(\mathrm{j}, y)]$$

$$\mathrm{S}\ \frac{\text{João ama Maria}}{\lambda y[\mathrm{A}(\mathrm{j}, y)](\mathrm{m})}\ \mathrm{AplD}$$
$$= \mathrm{A}(\mathrm{j}, \mathrm{m})$$

Na apresentação de Wood (1993, p. 46), ainda se apresenta uma terceira regra unária – a divisão – que ela mesma observa ter sido classificada por Lambek “entre uma quantidade de outras regras “prováveis” num sistema baseado em aplicação, permutação e composição”; ou seja, uma regra derivável das outras. Já que ela não é uma regra básica da GC, resolvemos não apresentá-la aqui.

Estrutura de características

Segundo Covington (1994, p. 123), “uma estrutura de características é um conjunto de características (atributos) e valores com no máximo um valor para cada característica”. Como uma estrutura de características (EC) geralmente é representada por uma matriz de atributos e valores (existem outras representações, mas elas serão comentadas aqui), a matriz em 9 é um exemplo de

EC bem formada, porque para a característica "a" há um valor "b" e à característica "c" corresponde o valor "d"; por outro lado, a matriz em 10 não é uma EC bem formada porque há dois valores ("b" e "c") para a característica "a". Apesar da representação de Covington também ser matricial, ela é um pouco diferente da convenção que estamos empregando aqui, que é fortemente inspirada por Copestake (2002); as principais diferenças são: 1) Covington separa característica e valores por meio de dois pontos (":") e espaçamento, enquanto Copestake usa apenas o espaçamento; 2) Covington representa as variáveis através de letras maiúsculas (como no Prolog), já Copestake antepõe a elas o símbolo "#"; e 3) na notação de Covington não há nenhum recurso para abreviar as listas, o que é feito por Copestake pelos parênteses angulados ("⟨" e "⟩").

9. $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ 10. $\begin{bmatrix} a & b \\ a & c \end{bmatrix}$

As características são sempre representadas por símbolos atômicos, mas os valores podem ser representados por um símbolo atômico (como nos exemplos 9 e 10) ou por outra EC. Essa distribuição recursiva permite que algumas características sejam agrupadas numa matriz como valor de uma determinada característica subordinante, como em 11, em que as características "i" e "k" (junto com seus respectivos valores) constituem a EC que é valor da característica "h"; este par de característica e valor, por sua vez, compõe a EC correspondente ao valor da característica "c", junto com as características "d" e "f" (novamente junto com seus respectivos valores); finalmente, a característica "c" (e seu valor) forma junto com a característica "a" (e seu valor) a matriz de 11. Enquanto conjunto, a ordem em que os pares de características e valores aparecem dentro de uma matriz não afeta a sua identidade; assim, as matrizes de 11 e 12 são idênticas, apesar do ordenamento diferente.

$$11.\ \begin{bmatrix} a & b \\ c & \begin{bmatrix} d & e \\ f & g \\ h & \begin{bmatrix} i & j \\ k & l \end{bmatrix} \end{bmatrix} \end{bmatrix} \qquad 12.\ \begin{bmatrix} c & \begin{bmatrix} f & g \\ h & \begin{bmatrix} i & j \\ k & l \end{bmatrix} \\ d & e \end{bmatrix} \\ a & b \end{bmatrix}$$

A única operação sobre EC que será empregada na construção das análises gramaticais é a de unificação, representada pelo símbolo "∏". A unificação é operação que toma duas EC, formando uma terceira EC; o resultado da unificação é basicamente a união de todos os pares de características e valores das EC de entrada. Ainda segundo Covington (1994, p. 125):

- Para unificar duas EC, unifique os valores de todas as características.
- Se uma característica ocorre apenas em uma EC e não na outra, simplesmente a inclua na EC resultante.
- Se uma característica ocorre am ambas as EC, unifique seus valores:
  - Para unificar valores representados por símbolos atômicos, é preciso que eles sejam iguais; caso contrário, a unificação falha.
  - Para unificar uma variável com qualquer coisa, simplesmente a faça igual à coisa.
  - Para unificar valores que sejam EC, aplique todo esse processo recursivamente.

Os exemplos que o próprio Covington nos oferece são as unificações em 13 e 14. Na primeira unificação, como o par de característica e valor "a b" só aparece na primeira EC, ele também aparece no resultado; como a característica "c" tem como valor uma variável na primeira EC, no resultado essa característica assume o valor dessa mesma característica na segunda EC; e, finalmente, como o par de característica e valor "h i" só aparece na segunda EC, ele é repetido no resultado. Na segunda unificação, a característica "a" tem como valor uma variável na segunda EC ("#y", que ainda está ligada a uma outra posição dentro da mesma EC) e "p" na primeira, o par de característica e valor "a p" é incluído no resultado; como o par de característica e valor "b p" só aparece na segunda EC, ele é copiado no resultado; e, para terminar, como a característica "c" apresenta uma variável ("#x") como valor na primeira EC e uma outra EC como valor na segunda, essa outra EC será o valor de “c” no resultado (observando a unificação do valor da variável “#y”, que faz com que o valor da característica “d” seja “p”).

$$13.\ \begin{bmatrix} a & b \\ c & \#x \end{bmatrix} \prod \begin{bmatrix} c & \begin{bmatrix} d & e \\ f & g \end{bmatrix} \\ h & i \end{bmatrix} = \begin{bmatrix} a & b \\ c & \begin{bmatrix} d & e \\ f & g \end{bmatrix} \\ h & i \end{bmatrix}$$

$$14.\ \begin{bmatrix} a & p \\ c & \#x \end{bmatrix} \prod \begin{bmatrix} a & \#y \\ b & q \\ c & \begin{bmatrix} d & \#y \\ e & f \end{bmatrix} \end{bmatrix} = \begin{bmatrix} a & p \\ b & q \\ c & \begin{bmatrix} d & p \\ e & f \end{bmatrix} \end{bmatrix}$$

Um exemplo de unificação inconsistente, representada pelo símbolo “⊥”, ocorreria se tentássemos unificar os resultados da unificação em 13 e 14, como se pode ver em 15. Como a característica “a” recebe valores atômicos diferentes em cada uma das EC (“b”, na primeira, e “p”, na segunda), isso é suficiente para a unificação falhar; mas ainda há uma segunda causa para a inconsistência desta unificação: os valores da característica “d” na matriz que é valor de “c” (“e”, na primeira, e “p”, na segunda).

GC em EC

A forma mais evidente de se empregar uma EC para representar uma análise gramatical feita por meio da CG é pela estipulação de cinco características, para as seguintes informações representadas num diagrama de Prawitz: 1) expressão; 2) sua categoria; 3) sua denotação; 4) regra; e 5) constituintes. (Na verdade, como todas as matrizes sempre terão pelo menos quatro dessas características, já estamos no domínio das EC padronizadas; o termo em inglês é *typed*, que normalmente é traduzido como "tipado", mas me parece que "padronizado" reflete melhor o conceito, sem recorrer a um neologismo desnecessário. Sobre as EC padronizadas, consultar Carpenter, 1992 ou Copestake, 2002.

Aplicação

Assim, a mesma análise para "João corre" apresentada ao estilo de Prawitz no diagrama 1 pode ser representada por uma EC como a da matriz 1. Nesta matriz, podemos ver que "João corre" é uma expressão ("exp"), cuja categoria ("cat") é "S" e cuja denotação ("den") é "C(j)" (ao contrário da outra, nessa representação não se oferece a história derivacional das reduções do operador λ), devido à regra ("reg") de aplicação à esquerda ("AplE"); em sua constituição ("con"), essa aplicação envolve a expressão "João", que é um item lexical ("reg Lex"), cuja categoria é "N" e cuja denotação é "j", e a expressão "corre", que também é um item lexical, cuja denotação é "$\lambda x[C(x)]$" e cuja categoria funtora resulta ("res") num "S" tomando um "N" à sua esquerda ("esq"). (Uma maneira alternativa de se representar as categorias funtoras seria por meio da estipulação de três características: a do resultado, a da direção e a do argumento; preferimos, no entanto, a que pareceu mais econômica.) Enquanto unidades independentes, esses itens lexicais estão representados na matriz 2, de "João", e na matriz 3, de "corre".

MATRIZ 1 - EC COM ANÁLISE DE "JOÃO CORRE"

$$\begin{bmatrix} \text{exp} & \text{João corre} \\ \text{reg} & \text{AplE} \\ \text{con} & \left\langle \begin{bmatrix} \text{exp} & \text{João} \\ \text{reg} & \text{Lex} \\ \text{cat} & \text{N} \\ \text{den} & \text{j} \end{bmatrix}, \begin{bmatrix} \text{exp} & \text{corre} \\ \text{reg} & \text{Lex} \\ \text{cat} & \begin{bmatrix} \text{res} & \text{S} \\ \text{esq} & \text{N} \end{bmatrix} \\ \text{den} & \lambda x[C(x)] \end{bmatrix} \right\rangle \\ \text{cat} & \text{S} \\ \text{den} & C(j) \end{bmatrix}$$

MATRIZ 2 - EC COM REPRESENTAÇÃO LEXICAL DE "JOÃO"

$$\begin{bmatrix} \text{exp} & \text{João} \\ \text{reg} & \text{Lex} \\ \text{cat} & \text{N} \\ \text{den} & \text{j} \end{bmatrix}$$

MATRIZ 3 - EC COM REPRESENTAÇÃO LEXICAL DE "CORRE"

$$\begin{bmatrix} \text{exp} & \text{corre} \\ \text{reg} & \text{Lex} \\ \text{cat} & \begin{bmatrix} \text{res} & \text{S} \\ \text{esq} & \text{N} \end{bmatrix} \\ \text{den} & \lambda x[C(x)] \end{bmatrix}$$

Dessa forma, a regra de aplicação à esquerda também pode ser representada por uma EC, como na matriz 4. Nesta matriz, a regra de aplicação à esquerda ("reg AplE") é definida pela concatenação (representada pelo símbolo "•") de duas expressões ("exp #exp1•#exp2"), de modo que a categoria da primeira é a mesma exigida como argumento esquerdo da segunda ("con.rest.cat.esq"; é assim que se designa um caminho numa matriz: "esq" é uma característica de "cat", que por sua vez é uma característica de "rest", que é característica de "con"; "rest" faz parte da estrutura em lista abreviada pelos parênteses angulados); a categoria da expressão concatenada tem o mesmo valor que o da característica "res" na segunda expressão ("con.rest.cat.res"), e sua denotação será a aplicação da denotação da segunda ("#f") à da primeira ("#a"). Os efeitos da regra são obtidos todos por meio da ligação das variáveis, e o resultado da análise gramatical na matriz 1 decorre da unificação das EC dos itens lexicais da matriz 2 e da matriz 3 com as duas posições na lista que é valor da característica "con" na matriz 4.

MATRIZ 4 - EC PARA A REGRA DE APLICAÇÃO À ESQUERDA

$$\begin{bmatrix} \text{exp} & \#\text{exp1}\bullet\#\text{exp2} \\ \text{reg} & \text{AplE} \\ \text{con} & \left\langle \begin{bmatrix} \text{exp} & \#\text{exp1} \\ \text{cat} & \#\text{y} \\ \text{den} & \#\text{a} \end{bmatrix}, \begin{bmatrix} \text{exp} & \#\text{exp2} \\ \text{cat} & \begin{bmatrix} \text{res} & \#\text{x} \\ \text{esq} & \#\text{y} \end{bmatrix} \\ \text{den} & \#\text{f} \end{bmatrix} \right\rangle \\ \text{cat} & \#\text{x} \\ \text{den} & \#\text{f}(\#\text{a}) \end{bmatrix}$$

A definição da aplicação à direita é bem parecida com a EC na matriz 4, bastando apenas inverter a ordem das categorias e das denotações na lista que é valor da característica "con", como se pode ver na matriz 5.

MATRIZ 5 - EC PARA A REGRA DE APLICAÇÃO À DIREITA

$$\begin{bmatrix} \text{exp} & \#\text{exp1}\bullet\#\text{exp2} \\ \text{reg} & \text{ApID} \\ \text{con} & \left\langle \begin{bmatrix} \text{exp} & \#\text{exp1} \\ \text{cat} & \begin{bmatrix} \text{res} & \#X \\ \text{dir} & \#Y \end{bmatrix} \\ \text{den} & \#f \end{bmatrix}, \begin{bmatrix} \text{exp} & \#\text{exp2} \\ \text{cat} & \#Y \\ \text{den} & \#a \end{bmatrix} \right\rangle \\ \text{cat} & \#X \\ \text{den} & \#f(\#a) \end{bmatrix}$$

Considerando as representações para os itens lexicais “ama” e “Maria” como sendo os da matriz 6 e da matriz 7, a análise gramatical de “ama Maria”, que decorre da unificação dessas EC com a lista da característica “con” na EC da aplicação à direita, pode ser representada pela matriz 8.

MATRIZ 6 - EC PARA REPRESENTAÇÃO LEXICAL DE “AMA”

$$\begin{bmatrix} \text{exp} & \text{ama} \\ \text{reg} & \text{Lex} \\ \text{cat} & \begin{bmatrix} \text{res} & \begin{bmatrix} \text{res} & S \\ \text{esq} & N \end{bmatrix} \\ \text{dir} & N \end{bmatrix} \\ \text{den} & \lambda y[\lambda x[A(x,y)]] \end{bmatrix}$$

MATRIZ 7 - EC PARA REPRESENTAÇÃO LEXICAL DE “AMA”

$$\begin{bmatrix} \text{exp} & \text{Maria} \\ \text{reg} & \text{Lex} \\ \text{cat} & N \\ \text{den} & m \end{bmatrix}$$

MATRIZ 8 - EC COM ANÁLISE GRAMATICAL DE “AMA MARIA”

$$
\begin{bmatrix}
\text{exp} & \text{ama Maria} \\
\text{reg} & \text{AplD} \\
\text{con} & \left\langle \begin{bmatrix}
\text{exp} & \text{ama} \\
\text{reg} & \text{Lex} \\
\text{cat} & \begin{bmatrix} \text{res} & \begin{bmatrix} \text{res} & S \\ \text{esq} & N \end{bmatrix} \\ \text{dir} & N \end{bmatrix} \\
\text{den} & \lambda y[\lambda x[A(x,y)]]
\end{bmatrix}, \begin{bmatrix}
\text{exp} & \text{Maria} \\
\text{reg} & \text{Lex} \\
\text{cat} & N \\
\text{den} & m
\end{bmatrix} \right\rangle \\
\text{cat} & \begin{bmatrix} \text{res} & S \\ \text{esq} & N \end{bmatrix} \\
\text{den} & \lambda x[A(x,m)]
\end{bmatrix}
$$

E a mesma análise gramatical de “João ama Maria”, representada ao estilo de Prawitz no diagrama 2, aparecerá em EC como na matriz 9.

MATRIZ 9 - EC COM DERIVAÇÃO DE “JOÃO AMA MARIA” (APENAS COM AS REGRAS DE APLICAÇÃO)

$$
\begin{bmatrix}
\text{exp} & \#\text{exp} \\
\text{reg} & \text{PerD} \\
\text{con} & \left\langle \begin{bmatrix}
\text{exp} & \#\text{exp} \\
\text{cat} & \begin{bmatrix} \text{res} & \begin{bmatrix} \text{res} & \#Y \\ \text{dir} & \#Z \end{bmatrix} \\ \text{esq} & \#X \end{bmatrix} \\
\text{den} & \lambda\# v_x[\lambda\lambda v_z[\# f(\# v_x, \# v_z)]]
\end{bmatrix} \right\rangle \\
\text{cat} & \begin{bmatrix} \text{res} & \begin{bmatrix} \text{res} & \#Y \\ \text{esq} & \#X \end{bmatrix} \\ \text{dir} & \#Z \end{bmatrix} \\
\text{den} & \lambda\# v_z[\lambda\lambda v_x[\# f(\# v_x, \# v_z)]]
\end{bmatrix}
$$

Permutação

Em EC, as regras de permutação devem ser expressas como na matriz 10 (à esquerda) e na matriz 11 (à direita).

MATRIZ 10 - EC PARA REGRA DE PERMUTAÇÃO À ESQUERDA

$$\begin{bmatrix} \text{exp} & \#\text{exp} \\ \text{reg} & \text{PerE} \\ \text{con} & \left\langle \begin{bmatrix} \text{exp} & \#\text{exp} \\ \text{cat} & \begin{bmatrix} \text{res} & \begin{bmatrix} \text{res} & \#Y \\ \text{esq} & \#X \end{bmatrix} \\ \text{dir} & \#Z \end{bmatrix} \\ \text{den} & \lambda\# v_z[\lambda\lambda\, v_x[\# f(\# v_x, \# v_z)]] \end{bmatrix} \right\rangle \\ \text{cat} & \begin{bmatrix} \text{res} & \begin{bmatrix} \text{res} & \#Y \\ \text{dir} & \#Z \end{bmatrix} \\ \text{esq} & \#X \end{bmatrix} \\ \text{den} & \lambda\# v_x[\lambda\lambda\, v_z[\# f(\# v_x, \# v_z)]] \end{bmatrix}$$

MATRIZ 11 - EC PARA REGRA DE PERMUTAÇÃO À DIREITA

$$\begin{bmatrix} \text{exp} & \#\text{exp} \\ \text{reg} & \text{PerD} \\ \text{con} & \left\langle \begin{bmatrix} \text{exp} & \#\text{exp} \\ \text{cat} & \begin{bmatrix} \text{res} & \begin{bmatrix} \text{res} & \#Y \\ \text{dir} & \#Z \end{bmatrix} \\ \text{esq} & \#X \end{bmatrix} \\ \text{den} & \lambda\# v_x[\lambda\lambda\, v_z[\# f(\# v_x, \# v_z)]] \end{bmatrix} \right\rangle \\ \text{cat} & \begin{bmatrix} \text{res} & \begin{bmatrix} \text{res} & \#Y \\ \text{esq} & \#X \end{bmatrix} \\ \text{dir} & \#Z \end{bmatrix} \\ \text{den} & \lambda\# v_z[\lambda\lambda\, v_x[\# f(\# v_x, \# v_z)]] \end{bmatrix}$$

Para ilustrar o emprego da EC para a regra de permutação, vamos apresentar apenas a parte relevante da derivação incremental de “João ama Maria”, apresentada no diagrama 3. Na matriz 12, é possível observar a aplicação de “ama” a “João”, depois da primeira ter sido permutada à esquerda; nesse ponto, “João ama” poderia ser aplicado a “Maria”, mas como a EC para essa derivação ficaria muito grande e como a aplicação já foi explicada, sua construção fica como sugestão de exercício.

MATRIZ 12 - EC PARA DERIVAÇÃO DE “JOÃO AMA” COM PERMUTAÇÃO

$$
\left[\begin{array}{ll}
\text{exp} & \text{João ama} \\
\text{reg} & \text{AplE} \\
\text{con} & \left\langle \left[\begin{array}{ll} \text{exp} & \text{João} \\ \text{reg} & \text{Lex} \\ \text{cat} & \text{N} \\ \text{den} & \text{j} \end{array}\right], \left[\begin{array}{ll}
\text{exp} & \text{ama} \\
\text{reg} & \text{PerE} \\
\text{con} & \left\langle \left[\begin{array}{ll} \text{exp} & \text{ama} \\ \text{reg} & \text{Lex} \\ \text{cat} & \left[\begin{array}{ll} \text{res} & \left[\begin{array}{ll} \text{res} & \text{S} \\ \text{esq} & \text{N} \end{array}\right] \\ \text{dir} & \text{N} \end{array}\right] \\ \text{den} & \lambda y[\lambda x[\mathrm{A}(x,y)]] \end{array}\right] \right\rangle \\
\text{cat} & \left[\begin{array}{ll} \text{res} & \left[\begin{array}{ll} \text{res} & \text{S} \\ \text{dir} & \text{N} \end{array}\right] \\ \text{esq} & \text{N} \end{array}\right] \\
\text{den} & \lambda x[\lambda y[\mathrm{A}(x,y)]]
\end{array}\right] \right\rangle \\
\text{cat} & \left[\begin{array}{ll} \text{res} & \text{S} \\ \text{dir} & \text{N} \end{array}\right] \\
\text{den} & \lambda y[\mathrm{A}(\mathrm{j},y)]
\end{array}\right]
$$

Promoção

Para representar as regras de promoção, precisamos da matriz 13 (à direita) e da matriz 14 (à esquerda).

MATRIZ 13 - EC PARA REGRA DE PROMOÇÃO À DIREITA

$$\begin{bmatrix} \text{exp} & \#\text{exp} \\ \text{reg} & \text{ProD} \\ \text{con} & \left\langle \begin{bmatrix} \text{exp} & \#\text{exp} \\ \text{cat} & \#X \\ \text{den} & \text{a} \end{bmatrix} \right\rangle \\ \text{cat} & \begin{bmatrix} \text{res} & \#Y \\ \text{dir} & \begin{bmatrix} \text{res} & \#Y \\ \text{esq} & \#X \end{bmatrix} \end{bmatrix} \\ \text{den} & \lambda v[v(a)] \end{bmatrix}$$

MATRIZ 14 - EC PARA REGRA DE PROMOÇÃO À ESQUERDA

$$\begin{bmatrix} \text{exp} & \#\text{exp} \\ \text{reg} & \text{ProE} \\ \text{con} & \left\langle \begin{bmatrix} \text{exp} & \#\text{exp} \\ \text{cat} & \#X \\ \text{den} & \text{a} \end{bmatrix} \right\rangle \\ \text{cat} & \begin{bmatrix} \text{res} & \#Y \\ \text{esq} & \begin{bmatrix} \text{res} & \#Y \\ \text{dir} & \#X \end{bmatrix} \end{bmatrix} \\ \text{den} & \lambda v[v(a)] \end{bmatrix}$$

Como já foi dito, com as regras de promoção, temos uma alternativa para a derivação incremental, além das regras de permutação. No entanto, essa alternativa incremental ainda exige as regras de composição, por isso o exemplo de derivação incremental será postergado até depois da apresentação destas regras. Por enquanto, fiquemos com uma segunda derivação de "João corre", como se pode ver na matriz 15.

MATRIZ 15 - EC PARA DERIVAÇÃO DE "JOÃO CORRE" COM REGRA DE PROMOÇÃO

$$\begin{bmatrix} \text{exp} & \text{João corre} \\ \text{reg} & \text{ApID} \\ \text{con} & \left\langle \begin{bmatrix} \text{exp} & \text{João} \\ \text{reg} & \text{ProD} \\ \text{con} & \left\langle \begin{bmatrix} \text{exp} & \text{João} \\ \text{reg} & \text{Lex} \\ \text{cat} & N \\ \text{den} & \text{j} \end{bmatrix} \right\rangle \\ \text{cat} & \begin{bmatrix} \text{res} & S \\ \text{dir} & \begin{bmatrix} \text{res} & S \\ \text{esq} & N \end{bmatrix} \end{bmatrix} \\ \text{den} & \lambda P[P(\text{j})] \end{bmatrix}, \begin{bmatrix} \text{exp} & \text{corre} \\ \text{reg} & \text{Lex} \\ \text{cat} & \begin{bmatrix} \text{res} & S \\ \text{esq} & N \end{bmatrix} \\ \text{den} & \lambda x[\text{C}(x)] \end{bmatrix} \right\rangle \\ \text{cat} & N \\ \text{den} & \text{C}(\text{j}) \end{bmatrix}$$

Composição

Finalmente, as regras de composição podem ser representadas pela matriz 16 (à direita) e pela matriz 17 (à esquerda).

MATRIZ 16 - EC PARA REGRA DE COMPOSIÇÃO À DIREITA

$$\begin{bmatrix} \text{exp} & \#\text{exp1}\bullet\#\text{exp2} \\ \text{reg} & \text{ComD} \\ \text{con} & \left\langle \begin{bmatrix} \text{exp} & \#\text{exp1} \\ \text{cat} & \begin{bmatrix} \text{res} & \#X \\ \text{dir} & \#Y \end{bmatrix} \\ \text{den} & \#f \end{bmatrix}, \begin{bmatrix} \text{exp} & \#\text{exp2} \\ \text{cat} & \begin{bmatrix} \text{res} & \#Y \\ \text{dir} & \#Z \end{bmatrix} \\ \text{den} & \#g \end{bmatrix} \right\rangle \\ \text{cat} & \begin{bmatrix} \text{res} & \#X \\ \text{dir} & \#Z \end{bmatrix} \\ \text{den} & \lambda v_z[f(g(v_z))] \end{bmatrix}$$

MATRIZ 17 - EC PARA REGRA DE COMPOSIÇÃO À ESQUERDA

$$\begin{bmatrix} \text{exp} & \#\text{exp1}\bullet\#\text{exp2} \\ \text{reg} & \text{ComE} \\ \text{con} & \left\langle \begin{bmatrix} \text{exp} & \#\text{exp1} \\ \text{cat} & \begin{bmatrix} \text{res} & \#Y \\ \text{esq} & \#Z \end{bmatrix} \\ \text{den} & \#g \end{bmatrix}, \begin{bmatrix} \text{exp} & \#\text{exp2} \\ \text{cat} & \begin{bmatrix} \text{res} & \#X \\ \text{esq} & \#Y \end{bmatrix} \\ \text{den} & \#f \end{bmatrix} \right\rangle \\ \text{cat} & \begin{bmatrix} \text{res} & \#X \\ \text{esq} & \#Z \end{bmatrix} \\ \text{den} & \lambda v_z[f(g(v_z))] \end{bmatrix}$$

De volta ao exemplo incremental, prometido na seção anterior, a derivação de “João ama” pelas regras de promoção e composição pode ser vista na matriz 18 (mais uma vez, a derivação de toda a sentença “João ama Maria” fica como sugestão de exercício).

MATRIZ 18 - EC PARA DERIVAÇÃO INCREMENTAL DE “JOÃO AMA” COM REGRAS DE PROMOÇÃO E DE COMPOSIÇÃO

$$
\begin{bmatrix}
\text{exp} & \text{João ama} \\
\text{reg} & \text{ComD} \\
\text{con} & \left\langle
\begin{bmatrix}
\text{exp} & \text{João} \\
\text{reg} & \text{ProD} \\
\text{con} & \left\langle \begin{bmatrix} \text{exp} & \text{João} \\ \text{reg} & \text{Lex} \\ \text{cat} & \text{N} \\ \text{den} & \text{j} \end{bmatrix} \right\rangle \\
\text{cat} & \begin{bmatrix} \text{res} & \text{S} \\ \text{dir} & \begin{bmatrix} \text{res} & \text{S} \\ \text{esq} & \text{N} \end{bmatrix} \end{bmatrix} \\
\text{den} & \lambda P[P(\text{j})]
\end{bmatrix},
\begin{bmatrix}
\text{exp} & \text{ama} \\
\text{reg} & \text{Lex} \\
\text{cat} & \begin{bmatrix} \text{res} & \text{S} \\ \text{esq} & \text{N} \end{bmatrix} \\
\text{den} & \lambda y[\lambda x[\text{A}(x, y)]]
\end{bmatrix}
\right\rangle \\
\text{cat} & \text{N} \\
\text{den} & \lambda y[A(\text{j}, y)]
\end{bmatrix}
$$

Conclusões

A princípio, de um ponto de vista estritamente gramatical, não parece haver nenhuma diferença entre a representação de uma derivação por meio de um diagrama ao estilo de Prawitz ou de uma matriz de EC. No entanto, do ponto de vista da construção da análise gramatical, como já foi dito, as EC oferecem um ambiente representacional no qual todas as informações lingüísticas podem ser agrupadas (sem que se perca a modularidade de seus diferentes tipos). Assim, ao permitir a construção de representações lingüísticas completas apenas por meio da operação de unificação, as EC reintegram os diversos níveis de análise lingüística, sem descaracterizá-los. Aqui, essa reintegração foi representada apenas pelos níveis sintáticos e semânticos, mas as EC poderiam facilmente ser expandidas para incluírem informações sobre os níveis fonéticos, morfológicos e mesmo pragmáticos; e cada um desses níveis ainda poderiam ser expandidos em quantos subníveis fossem necessários (por exemplo, o nível fonético poderia incluir uma representação prosódica, além da segmental; no nível semântico, além do nível extensional apresentado, poderíamos incluir uma representação intensional).

Justamente por oferecer esse ambiente integrado, essa representação em EC pode se constituir num bom modelo psicolingüístico, pelo menos no que se tem chamado de psicolingüística computacional. Além de desfrutar da incrementalidade, que vem junto com a GC, nessa representação faz mais sentido em se falar em processamento lingüístico, já que todos os níveis de análise lingüística estão sujeitos ao mesmo procedimento: a unificação. E mesmo uma questão computacional controversa que é a integração de parâmetros contínuos e discretos passa a ser irrelevante, porque as EC não distinguem o tipo dos valores das características: a princípio, eles podem tanto ser discretos (como todos os que foram usados no presente texto) como contínuo (de modo que os valores das características contínuas fossem representados, por exemplo, por números reais). Infelizmente, essa questão não poderá ser aprofundada aqui, já que precisaríamos redefinir a unificação; no texto, ela foi definida apenas para EC de características discretas, o que de forma alguma é uma necessidade epistemológica das EC.

Contudo, observando mais detidamente, mesmo do ponto de vista exclusivamente lingüístico, a representação em EC pode oferecer soluções para questões que a GC sozinha tem dificuldade em resolver. A questão da concordância entre o sujeito e o predicado, por exemplo, pode ser facilmente resolvida incluindo nas representações dos itens lexicais informações sobre a concordância que, caso não unifiquem, impedem algumas derivações inadequadas.

Dessa maneira, bastaria incluir nas representações dos itens lexicais nominais a característica “conc”, que por sua vez deve ser constituída por outras duas características: “gen” (para a concordância de gênero), que pode receber os valores “masc”, para o masculino, e “fem”, para o feminino; e “num” (para a concordância de número), que pode receber os valores “sing”, para o singular, e “plur”, para o plural.

Assim, com “o” representado pela matriz 19 e “menino” representado pela matriz 20, a derivação de “o menino” seria representada pela matriz 21 (em todas as matrizes abaixo, algumas características foram abreviadas, o que é representado pelas reticências, de forma a reduzir as representações apenas à questão discutida; caso contrário o tamanho das matrizes dificultaria a visualização dos exemplos).

MATRIZ 19 - REPRESENTAÇÃO LEXICAL DE "O" COM CARACTERÍSTICAS PARA CONCORDÂNCIA NOMINAL

$$\begin{bmatrix} \text{exp} & \text{o} \\ \text{reg} & \text{Lex} \\ \text{conc} & \begin{bmatrix} \text{gen} & \text{masc} \\ \text{num} & \text{sing} \end{bmatrix} \\ \text{cat} & \text{K} \\ \text{den} & \text{K} \end{bmatrix}$$

MATRIZ 20 - REPRESENTAÇÃO LEXICAL DE "MENINO" COM CARACTERÍSTICAS PARA CONCORDÂNCIA NOMINAL

$$\begin{bmatrix} \text{exp} & \text{menino} \\ \text{reg} & \text{Lex} \\ \text{conc} & \begin{bmatrix} \text{gen} & \text{masc} \\ \text{num} & \text{sing} \end{bmatrix} \\ \text{cat} & \text{K} \\ \text{den} & \text{K} \end{bmatrix}$$

MATRIZ 21 - DERIVAÇÃO DE "O MENINO" COM CONCORDÂNCIA NOMINAL

$$\begin{bmatrix} \text{exp} & \text{o menino} \\ \text{reg} & \text{AplD} \\ \text{con} & \left\langle \begin{bmatrix} \text{exp} & \text{o} \\ \text{reg} & \text{Lex} \\ \text{conc} & \begin{bmatrix} \text{gen} & \text{masc} \\ \text{num} & \text{sing} \end{bmatrix} \\ \text{cat} & \text{K} \\ \text{den} & \text{K} \end{bmatrix}, \begin{bmatrix} \text{exp} & \text{menino} \\ \text{reg} & \text{Lex} \\ \text{conc} & \begin{bmatrix} \text{gen} & \text{masc} \\ \text{num} & \text{sing} \end{bmatrix} \\ \text{cat} & \text{K} \\ \text{den} & \text{K} \end{bmatrix} \right\rangle \\ \text{cat} & \text{K} \\ \text{den} & \text{K} \end{bmatrix}$$

Já para os itens lexicais "as" e "meninas", representados respectivamente pela matriz 22 e pela matriz 23, resultariam numa derivação para "as meninas" como na matriz 24.

MATRIZ 22 - REPRESENTAÇÃO LEXICAL PARA "AS" COM CARACTERÍSTICAS PARA CONCORDÂNCIA NOMINAL

$$\begin{bmatrix} \text{exp} & \text{as} \\ \text{reg} & \text{Lex} \\ \text{conc} & \begin{bmatrix}\text{gen} & \text{fem}\\ \text{num} & \text{plur}\end{bmatrix} \\ \text{cat} & K \\ \text{den} & K \end{bmatrix}$$

MATRIZ 23 - REPRESENTAÇÃO LEXICAL DE "MENINAS" COM CARACTERÍSTICAS PARA CONCORDÂNCIA NOMINAL

$$\begin{bmatrix} \text{exp} & \text{meninas} \\ \text{reg} & \text{Lex} \\ \text{conc} & \begin{bmatrix}\text{gen} & \text{fem}\\ \text{num} & \text{plur}\end{bmatrix} \\ \text{cat} & K \\ \text{den} & K \end{bmatrix}$$

MATRIZ 24 - DERIVAÇÃO DE "AS MENINAS" COM CONCORDÂNCIA NOMINAL

$$\begin{bmatrix} \text{exp} & \text{as meninas} \\ \text{reg} & \text{AplD} \\ \text{con} & \left\langle \begin{bmatrix} \text{exp} & \text{as} \\ \text{reg} & \text{Lex} \\ \text{conc} & \begin{bmatrix}\text{gen} & \text{fem}\\ \text{num} & \text{plur}\end{bmatrix} \\ \text{cat} & K \\ \text{den} & K \end{bmatrix}, \begin{bmatrix} \text{exp} & \text{meninas} \\ \text{reg} & \text{Lex} \\ \text{conc} & \begin{bmatrix}\text{gen} & \text{fem}\\ \text{num} & \text{plur}\end{bmatrix} \\ \text{cat} & K \\ \text{den} & K \end{bmatrix} \right\rangle \\ \text{cat} & K \\ \text{den} & K \end{bmatrix}$$

Para se chegar a esses resultados, a única alteração que precisaria ser feita na regra de aplicação à direita é a inclusão da unificação da característica de concordância das expressões que constituem a aplicação; depois disso, eles decorrem exclusivamente da unificação (ou não) dessa característica especificada nos itens lexicais. Observe que seria impossível derivar uma EC para "o meninas" ou para "as menino", já que a escolha do artigo vai impor seu gênero e número ao substantivo. É importante ressaltar que este é apenas um exemplo didático, e não deve ser avaliado como uma solução séria da concordância; qualquer solução minimamente relevante ainda precisaria incluir a concordância verbo-nominal e avaliar como essa distinção pode ser incluída sem afetar a generalidade das regras, o que não é nada trivial.

Para encerrar, é preciso dizer que nada do que foi apresentado aqui foi empregado para analisar dados efetivos do português, além dos poucos exemplos apresentados. Para conhecer os limites e as qualidades efetivas desse paradigma, é preciso empregá-lo massivamente para produzir análises de vários tipos de fenômenos. As ferramentas elementares para essa tarefa foram apresentadas aqui, e a partir de agora é preciso formar um grupo que as coloque em uso, para sabermos efetivamente onde elas produzem bons resultados e onde elas ainda precisarão ser ajustadas.

RESUMO

No presente texto, apresenta-se uma maneira de empregar a Estrutura de Características para representar tanto a Gramática Categorial quanto as análises gramaticais construídas por meio delas. Para atingir esse objetivo, depois de introduzirmos dois modelos complementares da Gramática Categorial (o Modelo AB e o Modelo Clássico), algumas noções elementares de Estrutura de Característica serão explicadas de forma a descrever a representação da Gramática Categorial e de análises gramaticais através de matrizes de Estrutura de Característica; o texto termina com algumas conclusões a respeito desse tipo de representação.

Palavras-chave: gramática categorial, estrutura de característica, análise gramatical.

ABSTRACT

In this text, it will be presented a way to employ Feature Structure in order to represent both Categorial Grammar and parsing. To reach this goal, after the introduction of two complementary Categorial Grammar models (AB Model and Classic Categorial Grammar), some basic notions about Feature Structure will be explained with the aim of describing how to parse expressions using Categorial Grammar and Feature Structure. The text ends with some conclusions about this kind of representation.

Key-words: categorial grammar, feature structure, parsing.

REFERÊNCIAS

AJDUKIEWICZ, K. Die syntaktische Konnexität. *Studia Philosophica*, v. 1, p. 1-27, 1935. (traduzida como: Syntactic connection. In: McCALL, S. (Ed.). *Polish Logic*. Oxford: Oxford University Press, 1967. p. 207-231).

BAR-HILLEL, Y. A quasi-arithmetical notation for syntactic description. *Language*, v. 29, p. 47-58, 1953. (reimpresso em: BAR-HILLEL, Y. *Language and Information*. Reading: Addison-Wesley, 1964. p. 61-74).

_____; GAIFMAN, C.; SHAMIR, E. On categorical and phrase structure grammars. *The Bulletin of the Research Council of Israel*, v. 9F, p. 1-16, 1960. (reimpresso em: BAR-HILLEL, Y. *Language and Information*. Reading: Addison-Wesley, 1964. p. 99-115).

CARPENTER, B. *The Logic of Typed Feature Structures*. Cambridge: Cambridge University Press, 1992.

COPESTAKE, A. *Implementing Typed Feature Structure Grammars*. Stanford: CSLI, 2002.

COVINGTON, M. *Natural Language Processing for Prolog Programmers*. Englewood Cliffs: Prentice Hall, 1994.

DOWTY, D.; WALL, R.; PETERS, S. *Introduction to Montague Semantics*. Dordrecht: Kluwer, 1981.

LAMBEK, J. The mathematics of sentence structure. *American Mathematical Monthly*, v. 65, p. 154-170, 1958. (reimpresso em: BUSZKOWSKI, W.; MARCISZEWSKI, W.; VAN BENTHEM, J. (Eds.). *Categorial Grammar*. Amsterdam: John Benjamins, 1988).

WOOD, M. *Categorial Grammars*. London: Routledge, 1993.