

Impressões lingüísticas sobre duas axiomatizações para a Gramática Categorial

Luiz Arthur Pagani¹

¹ Departamento de Lingüística, Letras Vernáculas e Clássicas – UFPR
Rua General Carneiro, n. 460, 11o. andar – 80.060–150 Curitiba – PR

arthur@ufpr.br

Abstract. *In the present essay, two different but logically equivalent axiomatizations of Categorical Grammar will be compared from an exclusively linguistic perspective: 1) the version of the so called reduction rules, and 2) the version of Lambek calculus. In order to do that, after a first section of introduction, each axiomatization will be separately presented in the second section. In the third section both versions will be commented in relation to the type of linguistic knowledge representation they allow. At the conclusion in the fourth section the reduction rules version is argued to be the best linguistic option. Although the observations are linguistically motivated, their conclusions affect the nature of the linguistic knowledge (especially the lexical one) to be represented in any parser for Categorical Grammar.*

Resumo. *No presente texto, compara-se de uma perspectiva exclusivamente lingüística duas axiomatizações diferentes, mas logicamente equivalentes, da Gramática Categorial: 1) a versão das chamadas regras de redução e 2) a do cálculo de Lambek. Para isso, depois de uma primeira seção de introdução, cada uma dessas axiomatizações é apresentada separadamente numa segunda seção. Na terceira seção, as duas versões são comentadas em relação ao tipo de representação do conhecimento lingüístico em cada uma delas. Por fim, na quarta seção, apresenta-se como conclusão a preferência pela axiomatização das regras de redução. E ainda que as observações sejam lingüisticamente motivadas, suas conclusões afetam a natureza do conhecimento lingüístico (especialmente o lexical) que precisa ser representado em analisadores para Gramáticas Categoriais.*

1. Introdução

No presente texto,¹ discutiremos de um ponto de vista exclusivamente lingüístico duas axiomatizações ligeiramente diferentes da Gramática Categorial, mas que são logicamente equivalentes. Numa dessas versões, a gramática é definida por um conjunto de seis pares de regras de redução (aplicação, permutação,² composição, promoção,³ divisão

¹Agradeço a meus companheiros do Laboratório de Lingüística, Lógica e Computação, da Universidade Federal do Paraná, por oferecer um ambiente propício ao estudo e à discussão de questões relacionadas à Gramática Categorial; agradeço em especial a Rodrigo Tadeu Gonçalves por algumas sugestões diretas sobre o presente texto.

²Normalmente, o termo usado aqui é “associatividade”, do inglês *associativity*, mas seguindo uma observação de Oehrle, citada em [Wood, 1993, p. 37], prefiro usar o termo “permutação”, traduzindo o inglês *swapping*, que remete à troca da ordem em que os argumentos se combinam com seu funtor.

³A escolha desse termo ainda é mais complicada do que a do anterior, porque em inglês mesmo o conceito é mencionado através de mais de um termo: *raising* [Wood, 1993, p. 42], *lifting* [Moortgat, 1988,

do funtor principal e divisão do funtor subordinado) que associa certas operações semânticas às respectivas operações de combinação categorial; na segunda versão, essas mesmas regras são teoremas deriváveis a partir de um único axioma e de dois pares de regras de inferência, ou dedução — um par de regras para cada conectivo categorial (/ e \).⁴

Nesta apresentação, no entanto, não nos deteremos no aspecto lógico ou dedutivo das axiomatizações. O principal objetivo aqui não é discutir a decidibilidade ou a precisão axiomática de cada um dos dois sistemas, mas sim chamar a atenção para certos aspectos ontológicos e epistemológicos da representação do conhecimento lingüístico. Para isso, ao invés de postularmos alguma espécie de primazia algébrica ou computacional de um sistema em relação ao outro (que talvez nem exista, se eles forem mesmo logicamente equivalentes), vamos decidir essa primazia a partir de determinadas características que cada um desses sistemas apresenta em relação ao tipo de explicação que ele pode sugerir para o trabalho de um lingüista.

2. Gramáticas Categoriais

Segundo [Moortgat, 1988, pp. 1–2], uma Gramática Categorial se distingue de outras teorias lingüísticas muito semelhantes (tais como a Gramática de Estrutura Sintagmática Generalizada (*Generalized Phrase Structure Grammar*, GPSG) ou a Gramática de Estrutura Sintagmática Conduzida pelo Núcleo (*Head-Driven Phrase Structure Grammar*, HPSG)) por apresentar as seguintes quatro características:

- *Lexicalismo*. As teorias gramaticais que se concentram na estrutura superficial compartilham uma tendência em deslocar para o léxico a carga explicativa que, em outras teorias, seria atribuída ao componente sintático. Ao desenvolver uma noção mais ampla de estrutura categorial, por exemplo, a GPSG torna desnecessário um componente transformacional como o da Gramática Gerativa Clássica. A Gramática Categorial avança um passo a mais em direção ao lexicalismo, tornando desnecessário o próprio componente sintagmático. A informação sintática é completamente projetada a partir das estruturas categoriais atribuídas aos itens lexicais. Na sua forma mais pura, a Gramática Categorial identifica o léxico como o único local para as estipulações específicas às línguas. A sintaxe é uma álgebra livre: uma combinatória universal conduzida pelas estruturas categoriais complexas.
- *Estrutura de função e argumento*. A contribuição categorial mais específica para a teoria das categorias é a de que as expressões incompletas são modeladas, sintática e semanticamente, como funtores. As dependências elementares entre as expressões, que determinam fenômenos como a regência, o controle e a concordância, são todas definidas através da hierarquia entre funções e argumentos, e não por sua configuração estrutural.
- *Flexibilidade dos constituintes*. A Gramática Categorial Clássica, assim como a Gramática de Estrutura Sintagmática, atribui uma única estrutura de constituintes

p. 11] e *shifting* [Carpenter, 1997, p. 100]. Preferimos então o termo “promoção”, mas em [Neto, 1999] o termo usado é “elevação”; só achamos inadequado o uso de “alçamento”, porque esse termo tem sido usado na Gramática Gerativa para designar um outro tipo de fenômeno (o movimento de um constituinte para uma posição mais alta na árvore de estrutura sintagmática da expressão).

⁴Na verdade, no cálculo de Lambek há um terceiro conectivo ($\bullet$), o que exige mais um par de regras de inferência; no entanto, a ausência desse terceiro conectivo e de suas respectivas regras de dedução não afetam as questões discutidas no presente texto. Além disso, ambas as axiomatizações são compostas efetivamente por esquemas de axiomas e de regras de inferência, que ainda precisam ser preenchidos por categorias para se tornarem axiomas e regras; no entanto, tomaremos aqui a liberdade terminológica de chamá-los apenas de axiomas e de regras de inferência.

tes a uma expressão não-ambígua. As teorias categoriais generalizadas substituem essa noção de constituintes por outra mais flexível, oferecendo um inventário mais amplo de operações combinatórias que configuram um cálculo da mudança de tipo. Uma expressão não-ambígua é associada a um conjunto de derivações equivalentes. A coordenação booleana generalizada funciona como uma técnica experimental que revela os constituintes alternativos ocultos.

- *Composicionalidade.* A relação entre a álgebra sintática e a álgebra semântica é um homomorfismo, ou seja, uma relação que preserva a estrutura, na qual cada operação sintática corresponde a uma operação semântica. A Gramática Categorical Clássica incorpora uma forma de composicionalismo bastante forte, baseada na correspondência entre a regra de redução sintática central e a aplicação funcional na semântica. Os sistemas categoriais generalizados ampliam esta forma forte de composicionalidade para o cálculo da mudança de tipo, executando assim o programa da interpretação conduzida pelo tipo.

Uma consequência dessas características é que muitos dos fenômenos lingüísticos que exigiram da Gramática Gerativa, por exemplo, a postulação de categoria vazia, deslocamento e eliminação de estrutura, poderão ser monotonicamente resolvidos na Gramática Categorical sem recorrer a operações destrutivas como estas.

2.1. Regras de Redução

Numa de suas versões mais difundidas entre os poucos lingüistas que adotam a Gramática Categorical, normalmente ela é definida através de um conjunto de seis pares de regras de redução⁵ que associam uma operação de combinação categorial e uma operação de construção da representação semântica.

Nessa versão, seguindo [Moortgat, 1988, p. 11],⁶ os seis pares de regras de redução são definidos da seguinte maneira:

Regras de redução

R1 Aplicação

$$\begin{aligned} X/Y : f, Y : a &\Rightarrow X : f(a) \\ Y : a, Y \setminus X : f &\Rightarrow X : f(a) \end{aligned}$$

R2 Composição

$$\begin{aligned} X/Y : f, Y/Z : g &\Rightarrow X/Z : \lambda v[f(g(v))] \\ Z \setminus Y : g, Y \setminus X : f &\Rightarrow Z \setminus X : \lambda v[f(g(v))] \end{aligned}$$

R3 Permutação

$$\begin{aligned} (Z \setminus X)/Y : f &\Rightarrow Z \setminus (X/Y) : \lambda v_1[\lambda v_2[f(v_2)(v_1)]] \\ Z \setminus (X/Y) : f &\Rightarrow (Z \setminus X)/Y : \lambda v_1[\lambda v_2[f(v_2)(v_1)]] \end{aligned}$$

R4 Promoção

⁵Na verdade, esta é a versão mais ampla dessa vertente. A primeira versão da Gramática Categorical, proposta por [Ajdukiewicz, 1935] tinha apenas uma das regras de aplicação, porque ela não era direcional; numa das primeiras aplicações mais lingüisticamente motivadas [Bar-Hillel, 1953], conhecida como modelo AB, apenas o par de regras de aplicação era usado; numa outra versão, conhecida como Gramática Categorical Livre [Cohen, 1967], além das regras de aplicação, aparecem também as regras de permutação, de composição e de promoção. No entanto, o modelo mais empregado modernamente é mesmo esse de seis pares de regras, que também aparece num formato um pouco alterado na formulação de [Steedman, 1988], chamada de Gramática Categorical Combinatória.

⁶As expressões do cálculo lambda que representam a interpretação semântica receberam uma notação um pouco mais explícita aqui. No texto original, o escopo do operador lambda era marcado com um ponto, como em $\lambda P.P(x)$, o que é bastante usual; no entanto, como em fórmulas muito longas pode ficar difícil perceber o escopo do operador, preferi uma notação na qual o escopo recebe um marcador de início e de fim, como em $\lambda P[P(x)]$ (onde os colchetes marcam inequivocamente o começo e o final do escopo do operador).

$$X : a \Rightarrow Y/(X \setminus Y) : \lambda v[v(a)]$$

$$X : a \Rightarrow (Y/X) \setminus Y : \lambda v[v(a)]$$

R5 Divisão (funtor principal)

$$X/Y : f \Rightarrow (X/Z)/(Y/Z) : \lambda v_1[\lambda v_2[f(v_1(v_2))]]$$

$$Y \setminus X : f \Rightarrow (Z \setminus Y) \setminus (Z \setminus X) : \lambda v_1[\lambda v_2[f(v_1(v_2))]]$$

R6 Divisão (funtor subordinado)

$$X/Y : f \Rightarrow (Z/X) \setminus (Z/Y) : \lambda v_1[\lambda v_2[v_1(f(v_2))]]$$

$$Y \setminus X : f \Rightarrow (Y \setminus Z)/(X \setminus Z) : \lambda v_1[\lambda v_2[v_1(f(v_2))]]$$

De acordo com estas regras, e considerando que as expressões “Pedro”, “ama” e “Maria” correspondem respectivamente aos pares de categoria sintática e representação semântica ‘ $N : p$ ’, ‘ $(N \setminus S)/N : A$ ’ e ‘ $N : m$ ’, podemos representar a estrutura da sentença “Pedro ama Maria” através de um diagrama como o da Figura 1.⁷

$$\frac{\frac{\text{Pedro}}{N : p} \quad \frac{\text{ama}}{(N \setminus S)/N : A} \quad \frac{\text{Maria}}{N : m}}{N \setminus S : A(m)} \quad R1$$

$$\frac{\quad}{S : A(m)(p)} \quad R1$$

Figura 1: Derivação de “Pedro ama Maria” apenas com R1

Esse diagrama da Figura 1 representa a demonstração de que a sentença “Pedro ama Maria” é uma sentença que denota a relação de amar que se estabelece de Pedro para Maria (‘ $S : A(m)(p)$ ’), a partir das regras de redução e de três premissas: 1) “Pedro” é um nome que denota o indivíduo Pedro (‘ $N : p$ ’), 2) “ama” é um predicado de dois lugares que denota a relação de amar (‘ $(N \setminus S)/N : A$ ’) e 3) “Maria” é um nome que denota o indivíduo Maria (‘ $N : m$ ’).

Nesse sentido, esse diagrama não é muito diferente de um diagrama em árvore, como os que são associados a uma Gramática de Estrutura Sintagmática. A grande diferença, no entanto, é que, a partir de uma Gramática de Estrutura Sintagmática para “Pedro ama Maria”, conseguiríamos construir uma única árvore para essa sentença; já com uma Gramática Categórica, como a apresentada acima, poderíamos chegar a uma outra estrutura para a mesma sentença, chegando à mesma representação semântica final, através de uma sequência de regras diferente da anterior, como podemos ver no diagrama da Figura 2.⁸

Poderíamos chegar ainda a um terceiro diagrama para a mesma sentença “Pedro ama Maria”, como na Figura 3, também com a mesma representação semântica. Na verdade, na Gramática Categórica, podemos encontrar um número infinito de derivações equivalentes para a mesma expressão, o que normalmente é chamado de ambigüidade espúria, que já foi considerado uma das principais falhas da Gramática Categórica, mas que pode ser facilmente controlada através de uma exigência de normalização das derivações (como em [Carpenter, 1997, pp. 160–164]).

⁷Nestes diagramas, conhecidos como dedução ao estilo de Prawitz, as barras horizontais relacionam uma conclusão e suas premissas, de forma que a conclusão aparece debaixo da barra e as premissas sobre ela; ao lado direito da barra registra-se a regra empregada na inferência. No entanto, a inserção dos itens lexicais não segue bem esse padrão: as expressões linguísticas aparecem acima das suas respectivas barras na primeira linha do diagrama, e debaixo delas são registradas suas respectivas categorias e representações semânticas. Finalmente, algumas derivações exigem ainda a introdução de suposições, que são apresentadas entre colchetes numerados com um índice, que marca o escopo entre a sua introdução e a sua eliminação.

⁸A redução- β é uma das principais operações do cálculo- λ , e pode ser caracterizada pela seguinte fórmula: $\lambda v[F](a) \Rightarrow F[v \mapsto a]$, que pode ser lida como ‘um termo- λ ($\lambda v[F]$), aplicado a outro termo (a) é equivalente ao termo no escopo do operador com as ocorrências livres da variável v substituídas pelo termo a ($F[v \mapsto a]$)’ [Carpenter, 1997, p. 50].

$$\begin{array}{c}
\text{Pedro} \qquad \qquad \qquad \text{ama} \qquad \qquad \qquad \text{Maria} \\
\frac{N : p}{Lx} \quad \frac{(N \setminus S)/N : A}{Lx} \quad \frac{N : m}{Lx} \\
\hline
\frac{N \setminus (S/N) : \lambda x[\lambda y[A(y)(x)]]}{R3} \\
\hline
\frac{S/N : \lambda x[\lambda y[A(y)(x)]](p)}{R1} \\
\qquad \qquad \qquad =_{red.\beta} \lambda y[A(y)(p)] \\
\hline
\frac{S : \lambda y[A(y)(p)](m)}{R1} \\
\qquad \qquad \qquad =_{red.\beta} A(m)(p)
\end{array}$$

Figura 2: Derivação de “Pedro ama Maria” com R1 e R3

$$\begin{array}{c}
\text{Pedro} \qquad \qquad \qquad \text{ama} \qquad \qquad \qquad \text{Maria} \\
\frac{N : p}{Lx} \quad \frac{(N \setminus S)/N : A}{Lx} \quad \frac{N : m}{Lx} \\
\hline
\frac{S/(N \setminus S) : \lambda P[P(p)]}{R4} \\
\hline
\frac{S/N : \lambda x[\lambda P[P(p)](A(x))]}{R2} \\
\qquad \qquad \qquad =_{red.\beta} \lambda x[A(x)(p)] \\
\hline
\frac{S : \lambda x[A(x)(p)](m)}{R1} \\
\qquad \qquad \qquad =_{red.\beta} A(m)(p)
\end{array}$$

Figura 3: Derivação de “Pedro ama Maria” com R1, R2 e R4

No entanto, do ponto de vista lingüístico, pode-se perceber uma pequena diferença entre o diagrama da Figura 1, por um lado, e os diagramas das Figuras 2 e 3, por outro: se considerarmos que as operações realizadas pelas regras de redução equivalem também a concatenações das seqüências fonológicas, na Figura 1 teríamos uma estrutura prosódica correspondente a “(Pedro (ama Maria))” (onde os parênteses encerram os constituintes prosódicos concatenados), enquanto que nas Figuras 2 e 3 teríamos a seguinte estrutura prosódica: ‘((Pedro ama) Maria)’.

Apesar de semanticamente equivalentes, essas duas estruturas prosódicas apresentam características sintáticas e discursivas distintas. Por exemplo, apenas a segunda estrutura aceitaria uma continuação como “E não Márcia”, no sentido de que não é a Márcia, e sim a Maria, a pessoa que o Pedro ama; a primeira estrutura prosódica, ao contrário, é compatível com uma continuação como “E não Paulo”, no sentido de que a pessoa que ama a Maria não é o Paulo, e sim o Pedro. Isso justifica, do ponto de vista lingüístico, a distinção das infinitas derivações aparentemente equivalentes de “Pedro ama Maria” em dois grupos: 1) o das que concatenam primeiro “Pedro” e “ama”, e 2) o das que concatenam primeiro “ama” e “Maria”.

2.2. Cálculo de Lambek

Ainda segundo [Moortgat, 1988, p. 2],

o cálculo de Lambek substitui o conjunto de regras de redução categorial que foram propostas na literatura (Aplicação, Composição, Promoção, etc.) por uma noção geral de derivabilidade, a partir da qual as leis de redução são consideradas teoremas. A derivabilidade é definida na forma de axiomatização de seqüentes, o que reduz as derivações categoriais a deduções lógicas com base nos procedimentos de prova desenvolvidos originalmente por Gentzen, em seu trabalho sobre o cálculo proposicional intuicionístico.

Devido à facilidade notacional, ao invés dos diagramas de derivação de seqüentes, vamos adotar aqui a mesma notação de derivação da Dedução Natural usada nos diagramas anteriores para a sentença “Pedro ama Maria”. Assim, ao invés de seis pares de regras, no cálculo de Lambek, precisamos apenas de um par de regras para cada conectivo: um de introdução e outro de eliminação do conectivo, como nos esquemas da Figura 4, abaixo, adaptados de [Carpenter, 1997, pp. 153 e 156].

$$\begin{array}{c}
 \begin{array}{c}
 \frac{X/Y : f \quad Y : a}{X : f(a)} \quad /E \\
 \text{(a) Eliminação de } /
 \end{array}
 \quad
 \begin{array}{c}
 \frac{Y : a \quad Y \backslash X : f}{X : f(a)} \quad \backslash E \\
 \text{(b) Eliminação de } \backslash
 \end{array}
 \quad
 \begin{array}{c}
 \frac{\vdots \quad [Y : v]^n}{X : F} \quad /I^n \\
 \text{(c) Introdução de } /
 \end{array}
 \quad
 \begin{array}{c}
 \frac{[Y : v]^n \quad \vdots}{X : F} \quad \backslash I^n \\
 \text{(d) Introdução de } \backslash
 \end{array}
 \end{array}$$

Figura 4: Esquemas para eliminação e introdução dos conectivos

Assim, nesta outra axiomatização, “as leis de redução que foram introduzidas antes como primitivos, passam a ser *teoremas*; ou seja, inferências válidas da lógica dos conectivos categoriais” [Moortgat, 1988, p. 27].

Em relação ao diagrama da Figura 1, com a derivação da sentença “Pedro ama Maria” apenas com a regra R1, a derivação empregando apenas os esquemas de eliminação dos conectivos, na Figura 5, não apresenta nenhuma diferença.

$$\begin{array}{c}
 \text{Pedro} \quad \text{ama} \quad \text{Maria} \\
 \frac{}{N : p} \quad Lx \quad \frac{}{(N \backslash S)/N : A} \quad Lx \quad \frac{}{N : m} \quad Lx \\
 \hline
 \frac{}{A(m)} \quad /E \\
 \hline
 \frac{}{A(m)(p)} \quad \backslash E
 \end{array}$$

Figura 5: Derivação de “Pedro ama Maria” apenas com eliminação

No entanto, o diagrama equivalente à derivação com R1 e R3 apresenta uma diferença essencial, que é a suposição de uma variável da categoria ‘N’, que aparece entre colchetes na derivação da Figura 6.

$$\begin{array}{c}
 \text{Pedro} \quad \text{ama} \quad \text{Maria} \\
 \frac{}{N : p} \quad Lx \quad \frac{}{(N \backslash S)/N : A} \quad Lx \quad [N : x]^1 \quad \frac{}{N : m} \quad Lx \\
 \hline
 \frac{}{N \backslash S : A(x)} \quad /E \\
 \hline
 \frac{}{S : A(x)(p)} \quad \backslash E \\
 \hline
 \frac{}{S/N : \lambda x[A(x)(p)]} \quad /I^1 \\
 \hline
 \frac{}{S : \lambda x[A(x)(p)](m)} \quad /E \\
 \quad =_{red.\beta} A(m)(p)
 \end{array}$$

Figura 6: Derivação de “Pedro ama Maria” equivalente à com R1 e R3

Para se chegar a um diagrama equivalente ao da derivação com R1, R2 e R4, na Figura 3, é necessário não apenas a suposição de uma variável que consome um dos argumentos de “ama”, mas é preciso supor também uma variável que se aplique a “Pedro”

para que, depois de ‘descarregada’ a suposição, tenhamos transformado “Pedro” numa função que toma um ‘ $N \setminus S$ ’ para resultar num ‘ S ’, como podemos ver na Figura 7.

$$\begin{array}{c}
\begin{array}{ccc}
\text{Pedro} & \text{ama} & \text{Maria} \\
\hline
N : p & [N \setminus S : P]^1 & (N \setminus S)/N : A & [N : x]^2 & N : m \\
\hline
S : P(p) & \setminus E & N \setminus S : A(x) & /E \\
\hline
S/(N \setminus S) : \lambda P[P(p)] & /I^1 & & & \\
\hline
S : \lambda P[P(p)](A(x)) & & & & /E \\
\quad =_{red.\beta} A(x)(p) & & & & \\
\hline
S/N : \lambda x[A(x)(p)] & & & & /I^2 \\
\hline
S : \lambda x[A(x)(p)](m) & & & & /E \\
\quad =_{red.\beta} A(m)(p) & & & &
\end{array}
\end{array}$$

Figura 7: Derivação de “Pedro ama Maria” equivalente à com R1, R2 e R4

3. Comparações

Para um lingüista acostumado com as árvores de uma Gramática de Estrutura Sintagmática, não é difícil ver, num diagrama como o da Figura 1, os itens lexicais claramente representados na parte superior da derivação. Logo abaixo desses itens lexicais, também são claras as representações das operações sintáticas e semânticas reunindo esses itens lexicais em expressões cada vez mais complexas. Como já foi dito, mesmo em diagramas como os da Figura 2 e 3, com derivações incomuns para uma Gramática de Estrutura Sintagmática, podemos relacionar as regras de permutação (R3) e de promoção (R4) a uma operação que afeta a estrutura prosódica da expressão, trocando de lugar fronteiras fonológicas e reagrupando seus constituintes entoacionais.

Já em relação aos diagramas das Figuras 6 e 7, a que tipo de coisas poderíamos relacionar as suposições entre colchetes?

Numa primeira tentativa de responder esta pergunta, poderíamos querer ver nessas suposições as chamadas categorias vazias da Gramática Gerativa, já que elas parecem introduzir apenas uma entidade da representação semântica (uma variável que será posteriormente abstraída pelo operador- λ) que não corresponde a nenhuma expressão lingüística. No entanto, na Gramática Gerativa, as categorias vazias são itens lexicais. Nas derivações da Gramática Categorial, os itens lexicais são claramente representados, no topo do diagrama, como uma expressão lingüística sobre uma barra horizontal, debaixo da qual constam sua categoria e sua representação semântica. É apenas na introdução dos itens lexicais que aparecem expressões lingüísticas sobre estas barras; no restante da derivação sobre as barras ficam apenas as premissas das conclusões localizadas sob a barra. Assim, era de se esperar que, se as suposições fossem categorias vazias, elas fossem introduzidas nos diagramas através de uma barra com o rótulo lexical com a parte acima da barra vazia; portanto, fica impossível relacionar estas suposições às categorias vazias.

Além disso, ainda que essa constatação não seja feita em nenhum lugar, os itens lexicais parecem corresponder a axiomas. Ou seja, eles precisam ser definidos em algum lugar, e eles não podem ser derivados a partir de nenhum outro teorema da Gramática Categorial. Mas, nas regras de introdução dos conectivos, as suposições não correspondem a nenhum axioma ou teorema; pelo contrário, nos diagramas, a suposição nunca aparece

debaixo de uma barra horizontal, como observamos. Dessa maneira, não é apenas às categorias vazias que as suposições não podem corresponder: elas também não podem ser comparadas a nenhum item lexical.

Contudo, mesmo que tivéssemos descoberto uma natureza mais lingüísticamente motivada para o papel das suposições nas regras de introdução de conectivos, ainda precisaríamos encontrar a motivação lingüística que justificasse a maior quantidade de regras empregadas nos diagramas do cálculo de Lambek, em relação aos diagramas equivalentes da versão das regras de redução.

Se observarmos os diagramas das Figuras 2 e 3, podemos constatar que ambos são formados apenas através de três aplicações de regras: uma aplicação de R3 e duas de R1, no diagrama da Figura 2, e uma aplicação de R4, uma de R2 e uma de R1, no diagrama da Figura 3. E em ambos os casos, é fácil relacionar as aplicações das regras unárias (R3 e R4) a operações fonológicas que afetam a organização dos constituintes prosódicos: em ambas as derivações a aplicação das regras unárias faz com que a estrutura prosódica se torne ‘((Pedro ama) Maria)’; ao contrário da derivação na qual essas regras não atuam, na Figura 1, cuja estrutura prosódica é ‘(Pedro (ama Maria))’.

Já nas derivações equivalentes do cálculo de Lambek, nas Figuras 6 e 7, as mesmas derivações são realizadas respectivamente através de quatro e seis aplicações das regras. Aqui, fica impossível equiparar cada uma dessas aplicações das regras a qualquer operação fonológica, já que há uma mesma diferença prosódica a ser relacionada à aplicação de uma eliminação e de uma introdução, no diagrama da Figura 6, e à aplicação de duas eliminações e duas introduções, no diagrama da Figura 7.

A diferença fica ainda mais ressaltada quando a derivação é apresentada com o cálculo de seqüentes, de Gentzen, como se pode ver na Figura 8.⁹

$$\frac{\frac{N : m \Rightarrow N : m}{\text{Id}} \quad \frac{\frac{N : p \Rightarrow N : p}{\text{Id}} \quad \frac{S : A(m)(p) \Rightarrow S : A(m)(p)}{\text{Id}}}{N : p, N \backslash S : A(m) \Rightarrow S : A(m)(p)} \backslash \text{Esq}}{N : p, (N \backslash S) / N : A, N : m \Rightarrow S : A(m)(p)} / \text{Esq}$$

Figura 8: Derivação de “Pedro ama Maria” com seqüentes

Na derivação da Figura 8, fica difícil para um lingüista reconhecer o que poderia corresponder a um item lexical, e principalmente identificar a entrada lexical relativa ao verbo “ama”: como a introdução de “Maria” e de “Pedro” é feita por duas instâncias do axioma da identidade (Id), não seria de esperar que “ama” também fosse introduzido por outra instância do mesmo axioma? No cálculo de seqüentes, não. Apenas as expressões atômicas aparecem nas instâncias do axioma de identidade: os nomes “Maria” e “Pedro”, e a sentença “Pedro ama Maria”. Como a categoria de “ama” é funcional, ele só aparece na conclusão da última inferência. Mas observe que, nessa representação, as expressões lingüísticas propriamente ditas sequer aparecem nos diagramas: vemos apenas suas categorias e suas representações semânticas, mas em lugar nenhum podemos perceber as expressões “Pedro”, “ama” ou “Maria”.

E se já é difícil identificar os próprios itens lexicais, é ainda mais difícil relacionar com o diagrama de seqüentes as operações de concatenação prosódica apontadas anteriormente. A maneira mais simples de apresentar as derivações equivalentes às com a permutação (R3) e com promoção e composição (R4 e R2), seria demonstrando-as separadamente, e depois substituindo na derivação os itens lexicais pelas equivalências

⁹Infelizmente, por falta de espaço, não será possível apresentar aqui a formalização do cálculo de Lambek com seqüentes, que pode ser encontrada em [Moortgat, 1988], [Morrill, 1994] e [Carpenter, 1997].

demonstradas, através da regra de corte (*cut rule*). Mas a que tipo de operação lingüística poderiam corresponder essa demonstração paralela e a própria regra de corte? Com efeito, não é possível relacioná-las lingüisticamente a nada.¹⁰

4. Conclusão

Através da observação de como os itens lexicais e uma operação lingüística (a concatenação de constituintes prosódicos) poderiam ser identificados nos diagramas de derivação da Gramática Categórica, o que se conclui é que não apenas “as derivações no cálculo associativo de Lambek são representadas mais economicamente por derivações da dedução natural ao estilo de Prawitz” [Morrill, 1994, p. 80], mas que sua versão com os seis pares de regras de redução oferecem um ambiente mais propício para a reflexão mais lingüisticamente motivada.

Assim, mesmo que “a perspectiva dos seqüentes seja uma base particularmente lúcida para a discussão de questões essenciais como as de derivabilidade e de decidibilidade” [Moortgat, 1988, p. 27], a equivalência entre os dois sistemas garante que as descobertas feitas para a versão com os seqüentes possa ser imediatamente transferida para a versão das regras de redução, que é mais adequada ao trabalho do lingüista.

$$\begin{array}{c}
 \frac{}{\text{Pedro} - N : p} Lx \quad \frac{}{\text{ama} - (N \setminus S) / N : A} Lx \quad \frac{}{\text{Maria} - N : m} Lx \\
 \hline
 \frac{}{(ama \text{ Maria}) - N \setminus S : A(m)} R1 \\
 \hline
 \frac{}{(\text{Pedro} (ama \text{ Maria})) - S : A(m)(p)} R1
 \end{array}$$

Figura 9: Explicitando a concatenação prosódica apenas com R1

$$\begin{array}{c}
 \frac{}{\text{Pedro} - N : p} Lx \quad \frac{}{\text{ama} - (N \setminus S) / N : A} Lx \quad \frac{}{\text{Maria} - N : m} Lx \\
 \hline
 \frac{}{\text{ama} - N \setminus (S / N) : \lambda x[\lambda y[A(y)(x)]]} R3 \\
 \hline
 \frac{}{(\text{Pedro} \text{ ama}) - S / N : \lambda x[\lambda y[A(y)(x)](p)]} R1 \\
 \quad \quad \quad =_{red.\beta} \lambda y[A(y)(p)] \\
 \hline
 \frac{}{((\text{Pedro} \text{ ama}) \text{ Maria}) - S : \lambda y[A(y)(p)](m)} R1 \\
 \quad \quad \quad =_{red.\beta} A(m)(p)
 \end{array}$$

Figura 10: Explicitando a concatenação prosódica com R1 e R3

$$\begin{array}{c}
 \frac{}{\text{Pedro} - N : p} Lx \quad \frac{}{\text{ama} - (N \setminus S) / N : A} Lx \quad \frac{}{\text{Maria} - N : m} Lx \\
 \hline
 \frac{}{\text{Pedro} - S / (N \setminus S) : \lambda P[P(p)]} R4 \\
 \hline
 \frac{}{(\text{Pedro} \text{ ama}) - S / N : \lambda x[\lambda P[P(p)](A(x))]} R2 \\
 \quad \quad \quad =_{red.\beta} \lambda x[A(x)(p)] \\
 \hline
 \frac{}{((\text{Pedro} \text{ ama}) \text{ Maria}) - S : \lambda x[A(x)(p)](m)} R1 \\
 \quad \quad \quad =_{red.\beta} A(m)(p)
 \end{array}$$

Figura 11: Explicitando a concatenação prosódica com R1, R2 e R4

¹⁰Novamente por limitação de espaço, também não apresentaremos os diagramas de seqüentes equivalentes às derivações das Figuras 2 e 3. Um diagrama com exemplo do uso da regra de corte para introdução da promoção do sujeito pode ser encontrado em [Carpenter, 1997, p. 147].

Nesse sentido, para encerrar, vamos apresentar uma adaptação para o estilo de Prawitz da representação que [Morrill, 1994, pp. 110–129] desenvolve usando o estilo de Ficht. As principais diferenças desta representação para as apresentadas antes são duas: 1) os itens lexicais ocupam o lugar das premissas nas demonstrações, e 2) as expressões lingüísticas aparecem explicitamente concatenadas. Os diagramas correspondentes às derivações da Figura 1, 2 e 3 podem ser vistos, respectivamente, nas Figuras 9, 10 e 11.

Nestes três últimos diagramas, podemos ver claramente os dois principais pontos ressaltados durante as discussões apresentadas aqui:

- Os itens lexicais correspondem a axiomas, introduzidos sob uma barra sem nada sobre ela; do ponto de vista lingüístico, a consequência é que os itens lexicais são independentes: eles não dependem diretamente de nenhuma operação da Gramática Categorial — pelo contrário, são os itens lexicais que afetam a análise lingüística representada nas derivações, assim como o lexicalismo preconiza.
- As operações de concatenação das expressões lingüísticas estão explicitamente expressas, assim como as operações de combinação categorial e de unificação das representações semânticas; assim, em cada passo da derivação, sabemos de cada expressão lingüística construída sua categoria e sua interpretação semântica, relacionando claramente expressões lingüísticas e suas respectivas interpretações semânticas, exatamente como exige a composicionalidade.

Dessa maneira, ainda que aparentemente a discussão acima tenha sido fundamentada pela ontologia das entidades lingüísticas (mais especificamente, pela natureza axiomática dos itens lexicais) e pela epistemologia das operações lingüísticas (não apenas as de combinação categorial e de interpretação semântica, mas também as de aglutinação de constituintes prosódicos), as conclusões a que acabamos de chegar afetam diretamente a representação desse tipo de conhecimento na elaboração de analisadores gramaticais que sirvam como modelo para o comportamento lingüístico humano: a implementação de analisadores para Gramáticas Categoriais que se pretendam psicologicamente realísticos deve representar os itens lexicais como axiomas de uma álgebra livre.

Referências

- Ajdukiewicz, K. (1935). Die syntaktische konnexität. *Studia Philosophica*, 1:1–27.
- Bar-Hillel, Y. (1953). A quasi-arithmetical notation for syntactic description. *Language*, 29:47–58.
- Carpenter, B. (1997). *Type-Logical Semantics*. The MIT Press, Cambridge, Massachusetts.
- Cohen, J. M. (1967). The equivalence of two concepts of categorial grammar. *Information and Control*, 10:475–484.
- Moortgat, M. (1988). *Categorial Investigations — Logical and Linguistic Aspects of the Lambek Calculus*. Foris, Dordrecht.
- Morrill, G. V. (1994). *Type Logical Grammar — Categorial Logic of Signs*. Kluwer, Dordrecht.
- Neto, J. B. (1999). Introdução à gramática categorial. UFPR, Curitiba.
- Steedman, M. (1988). Combinators and grammars. In Oehrle, R., Bach, E., and Wheeler, D., editors, *Categorial Grammars and Natural Language Structures*, pages 417–442. Reidel, Dordrecht.
- Wood, M. M. (1993). *Categorial Grammars*. Routledge, London.