

A INTERPRETAÇÃO DOS PROSSINTAGMAS

JOSÉ BORGES NETO
LUIZ ARTHUR PAGANI
UFPR

RESUMO: No presente texto, apresentamos a solução que Ilari oferece, no início dos anos 80, para problemas de interação entre uma anáfora nominal e outra que Ilari chama de prossintagma verbal, em sentenças como *Pedro surpreendeu sua mulher e Antônio fez o mesmo*, para a qual há três interpretações em relação a que mulher está sendo surpreendida por quem: 1) a de Pedro por Pedro e por Antônio (leitura de identidade estrita); 2) a de Pedro por Pedro, e a de Antônio por Antônio (leitura de identidade fajuta); ou 3) a mulher de um terceiro homem por Pedro e por Antônio (leitura dêitica). Na seqüência, avaliaremos a proposta de Ilari; e, como acreditamos que haja atualmente uma solução melhor dentro da Gramática Categorial (modelo escolhido por Ilari), indicaremos as insuficiências da solução inicialmente proposta e sugerimos uma nova alternativa, baseada no conectivo categorial definido por Jäger.

ABSTRACT: In the present essay we present Ilari's solution to solve problems of interaction between nominal and verbal anaphors (called *verbal prophrase* by Ilari, dating back to the beginning of the eighties), on sentences such as *John surprised his wife and Bill does the same*, which has three interpretations in relation to whose wife is surprised by whom: i) John's wife by John and Bill (strict identity lecture), ii) John's wife by John, and Bill's wife by Bill (sloppy identity lecture), or iii) a third man's wife by John and Bill (deitic lecture). In the sequence we evaluate Ilari's proposal, and since we believe that there are nowadays better solutions in Categorial Grammar (the model chose by Ilari) we present the insufficiencies of the solution initially proposed, and we will suggest a new alternative based on Jäger's categorial connective.

1. INTRODUÇÃO

Num texto notável, publicado em 1981, Rodolfo Ilari [13] discutiu (entre outras coisas) um fenômeno semântico conhecido como “ambigüidade de identidade estrita ou fajuta” (em inglês, *strict/sloppy identity*). Além de mostrar que a proposta de tratamento das relações anafóricas presente em Jackendoff [14] é absolutamente inadequada para o tratamento destes casos de ambigüidade, Ilari constrói um tratamento para os dados do português no quadro de uma Gramática Categorial (GC), usando a teoria proposta por David Lewis [20].

A análise proposta por Ilari, no entanto, embora basicamente correta do ponto de vista empírico, sofre com a escolha teórica feita: a versão da GC proposta por Lewis apresenta problemas de arquitetura insanáveis, que se transmitem aos mecanismos explicativos propostos por Ilari.

Nossa proposta, no presente texto, é fazer a revisão da análise de Ilari no quadro de uma versão mais moderna – e mais consistente – da GC.

2. A AMBIGÜIDADE DE IDENTIDADE ESTRITA E FAJUTA

Identificada por Peter Geach [8], a questão da identidade estrita ou fajuta tem a ver com a interpretação de sentenças como as seguintes:¹

- (1) Pedro bateu em sua mulher e José fez o mesmo.
- (2) Pedro bateu em sua mulher e José também.
- (3) Pedro declarou que ele não tinha escrito a carta comprometedora e José fez o mesmo.
- (4) Somente Pedro fala de si mesmo.
- (5) Pedro bateu em sua mulher e José não.

Todas as sentenças admitem, ao menos, duas interpretações. As sentenças (1) e (2) são equivalentes e podem significar tanto que Pedro e José bateram, cada um, em sua própria mulher, quanto significar que Pedro e José bateram na mulher de Pedro. A sentença (3) pode significar que Pedro e José declararam não terem escrito a carta comprometedora e também pode significar que Pedro declarou que ele (ele próprio ou uma terceira pessoa) não tinha escrito a carta comprometedora e José confirmou a declaração. É importante destacar que, neste caso, se Pedro declara que uma terceira pessoa – digamos, João – escreveu a carta, o que José faz é declarar exatamente a mesma coisa. Em (4), podemos reconhecer uma leitura em que se afirma que ninguém, a não ser Pedro, fala de si mesmo e uma outra leitura em que se afirma que ninguém, a não ser Pedro, fala de Pedro. A sentença (5), de certo modo, tem um comportamento semelhante às sentenças (1) e (2), exceto que a ação de José é exatamente contrária à de Pedro, ou seja, se Pedro bate em sua mulher, José ou não bate na mulher de Pedro ou não bate em sua própria mulher.

Embora todas as cinco sentenças sejam exemplos de ambigüidade de identidade estrita ou fajuta, elas apresentam estruturas sintáticas distintas. Vamos começar com (1) e (2).

Em (1) temos um caso de prossintagma verbal, na terminologia de Ilari. A expressão *fazer o mesmo* é um anafórico que tem um SV como antecedente. Exatamente como um pronome retoma um SN no co-texto (ou no contexto), um prossintagma verbal retoma um SV.² Em (2), por outro lado, não temos um prossintagma verbal. A expressão *também* não é um anafórico que tem o SV como antecedente, uma vez que pode co-ocorrer com ele (ou com partes dele). Os casos de (6) confirmam isso:

- (6) a. Pedro bateu em sua mulher e José também.
b. Pedro bateu em sua mulher e José também bateu.
c. Pedro bateu em sua mulher e José também bateu em sua mulher.

¹ A maior parte dos exemplos é do texto de Ilari (a exceção é o exemplo 5).

² Ilari enumera uma série de outros prossintagmas verbais, além de *fazer o mesmo*, como *acontecer o mesmo*, *dizer o mesmo*, *dar-se o mesmo*, *valer o mesmo*, etc. Cada um deles caracteriza uma relação temática diferente entre o SN sujeito e o SV da sentença.

Com os verdadeiros prossintagmas verbais, essa co-ocorrência é impossível.

- (7) a. Pedro bateu em sua mulher e José fez o mesmo.
- b. Pedro bateu em sua mulher e *José fez o mesmo bateu.
- c. Pedro bateu em sua mulher e *José fez o mesmo bateu em sua mulher.

Segundo Ilari, então, nos casos como (2) (com o *também*) podemos falar em elipse do SV e nos casos como (1) (com prossintagmas verbais) podemos falar em substituição do SV.

O comportamento de (5) é idêntico ao de (2), isto é, trata-se de um caso de elipse de SV. Os exemplos de (8) confirmam isso:

- (8) a. Pedro bateu em sua mulher e José não.
- b. Pedro bateu em sua mulher e José não bateu.
- c. Pedro bateu em sua mulher e José não bateu em sua mulher.

Em (3) temos novamente um caso de prossintagma verbal, mas em (4) não temos nem elipse de SV nem prossintagma verbal. A ambigüidade reconhecida em (4) deve resultar da composição de uma sentença com reflexivo (*Pedro fala de si mesmo*) e uma “partícula de exclusão” (*somente*).

Em todos os casos, no entanto, a ambigüidade de identidade estrita ou fajuta reconhecida nos cinco exemplos depende crucialmente da presença, na sentença, de um elemento anafórico que admita uma dupla interpretação. Em (1), (2) e (5), esse elemento é o possessivo *sua*; em (3), esse elemento é o pronome *ele*; e em (4) é o reflexivo *si mesmo*. A ambigüidade, portanto, não tem origem nem no prossintagma verbal (nos casos (1) e (3)), nem no SV elíptico (nos casos (2) e (5)); nem na partícula de exclusão (no caso (4)).

Ilari fixa sua atenção nos prossintagmas verbais e apresenta uma análise para eles que procuraremos reproduzir abaixo.

3. AGCUSADAPORILARI

Antes de apresentar a análise dos dados feita por Ilari, cremos que cabe uma apresentação do quadro teórico utilizado. Se a análise fosse efetuada num quadro teórico mais conhecido, poderíamos nos dispensar desse trabalho, mas a teoria assumida por Ilari é suficientemente desconhecida para que esta “introdução” se faça necessária.

A GC surgiu nos anos 1930, pelas mãos do lógico polonês Kazimierz Ajdukiewicz [1] e ficou praticamente desconhecida dos lingüistas até os anos 1950, quando foi proposta como modelo de análise lingüística pelo lógico israelense Yehoshua Bar-Hillel [2]. O impacto da proposta de Bar-Hillel foi praticamente nulo, embora seu texto tenha sido publicado na *Language*, a revista “oficial” da LSA (*Linguistic Society of America*).

Em 1958, Joachim Lambek publicou um texto [19]³ em que tentava mostrar como era possível tratar as expressões das línguas naturais num modelo formal de caráter rigorosa-

³ É interessante observar que, em nota, Lambek agradece aos comentários feitos por Chomsky; no entanto, apesar do conhecimento sobre a Gramática Categorial, esta parece ter influenciado pouco os rumos da Gramática Gerativa.

mente matemático. Sua proposta era de uma GC direcional, já apontada por Bar-Hillel – que permitia superar as limitações relacionadas com a questão da ordem das palavras nas sentenças. A questão da ordem das palavras era ignorada no algoritmo de Ajdukiewicz e era apresentada como um problema real para a adoção de uma GC nos modelos descritivos das línguas naturais por Bar-Hillel. Lambek, no entanto, era um matemático e escrevia para matemáticos: sua teoria também ficou praticamente desconhecida dos lingüistas.

No início dos anos 1970, como conseqüência do debate entre os gerativistas chomskyanos e os membros do movimento que se chamou Semântica Gerativa, a GC voltou à cena dos estudos lingüísticos – agora numa versão “transformacional”: a gramática de David Lewis. A proposta de Lewis, embora tenha sido originalmente publicada numa revista de filosofia, foi recolhida numa antologia [6] que, juntamente com textos de filósofos e psicólogos, incluía textos importantes de lingüistas bem conhecidos. Deste modo, a proposta de David Lewis – mesmo que não se possa dizer que teve grande impacto nos estudos lingüísticos – ficou sendo a versão de GC mais conhecida dos lingüistas.⁴

O tratamento dado por Ilari aos fenômenos ligados à ambigüidade de identidade estrita e fajuta foi feito no quadro da GC “transformacional”, de David Lewis. Vejamos em que consiste esta gramática.

3.1. A Teoria de Lewis

Uma GC, como a de Lewis [20], à qual Ilari recorre, é um mecanismo que associa as expressões de uma língua natural a um conjunto de categorias sintático-semânticas que se dividem em categorias primitivas (S , N e C) e categorias derivadas (p. ex. S/N).

Como no algoritmo de Ajdukiewicz, o conjunto de categorias primitivas é o menor possível,⁵ enquanto o número de categorias derivadas é aberto. As categorias derivadas são obtidas por uma regra recursiva que diz que se α e β são categorias (primitivas ou derivadas), $\frac{\alpha}{\beta}$ também é categoria.

A concatenação das expressões, para obtermos expressões “maiores” (mais complexas), é regida por uma única regra, R_1 , que tem a seguinte formulação:⁶

(9) R_1 :

Toda seqüência formada por uma expressão da categoria $A/B_1; : : : ; B_n$ seguida de expressões das categorias $B_1; : : : ; B_n$ é uma expressão bem formada da categoria A .

⁴ O texto de Lewis abre o livro organizado por Bárbara Hall-Partee sobre a Gramática de Montague [22]. Há uma apresentação sucinta da teoria de Lewis em [4].

⁵ Para Ajdukiewicz, só há duas categorias primitivas: N e S . Lewis acrescenta uma terceira categoria C para distinguir nomes próprios e comuns, que têm distribuições sintáticas diferentes em inglês.

⁶ Esta regra R_1 é conhecida como *regra de cancelamento de frações*, uma vez que funciona da mesma forma que a regra de cancelamento de numeradores e denominadores idênticos durante a multiplicação de frações da aritmética ordinária. Foi Bar-Hillel [2] quem lhe atribuiu esse nome. Observe-se que, por razões de praticidade tipográfica, ao invés de escrevermos o índice de uma categoria derivada como uma fração, estamos escrevendo esses índices com o auxílio de uma barra diagonal (ou seja, em vez de, $\frac{\alpha}{\beta}$ escrevemos α/β).

A gramática de Lewis conta com três categorias primitivas: a categoria das *sentenças* (*S*), que contém as expressões que denotam valores de verdade; a categoria dos *nomes individuais* – ou “próprios” – (*N*), que contém as expressões que denotam indivíduos; e a categoria dos *nomes comuns* (*C*), que abriga as expressões que denotam classes de indivíduos.

As expressões são sempre ordenadas com as *funções* (expressões funtoras) antecedendo os seus *argumentos* (expressões argumentais). As expressões pertencentes às categorias primitivas são sempre argumentos, enquanto as expressões pertencentes às categorias derivadas podem ser funtores ou argumentos.

Supondo que o nome próprio Pedro pertença à categoria *N* e que o verbo *correr* pertença à categoria *S/N*, a seqüência de palavras *correr Pedro*, pela aplicação da regra R_1 , pertencerá à categoria *S*.

Esse mecanismo permite gerar um conjunto de expressões que, posteriormente, precisam sofrer a ação de um componente transformacional para resultar em expressões da língua natural. Para melhor compreensão do funcionamento da teoria, vamos reproduzir aqui um fragmento da gramática do português presente em [13, p. 119].

(10) Léxico:

{ Pedro, Maria, ... }	∈ <i>N</i>
{ homem, menino, ... }	∈ <i>C</i>
{ dorme, foge, ... }	∈ <i>S/N</i>
{ o, todo, algum, este, ... }	∈ <i>N/C</i>
{ elogia, ... }	∈ (<i>S/N</i>)/ <i>N</i>

Com este léxico (categorizado) e com a regra R_1 podemos obter o seguinte conjunto de expressões bem formadas (EBFs), dentre outras:

(11) Expressões bem formadas:

a. Pedro	∈ <i>N</i>
b. menino	∈ <i>C</i>
c. dorme Pedro	∈ <i>S</i>
d. o menino	∈ <i>N</i>
e. dorme o menino	∈ <i>S</i>
f. elogia o menino	∈ <i>S/N</i>
g. elogia o menino Pedro	∈ <i>S</i>

Vejamos o processo de derivação de um exemplo.

(12) Derivação de (11c):

Passo 1:	$\left\{ \begin{array}{l} \text{dorme} \in S/N \\ \text{Pedro} \in N \end{array} \right\}$	(pelo léxico)
Passo 2:	$\left\{ \begin{array}{l} \text{Pedro dorme} = N \bullet S/N \\ \text{dorme Pedro} = S/N \bullet N \end{array} \right\}$	(concatenações possíveis)

Passo 3:	$N \bullet S/N = \text{não EBF}$	(impossível aplicar R_1)
Passo 4:	$S/N \bullet N = \text{EBF}$	(R_1 , resultando numa expressão da categoria S)

A este conjunto de expressões geradas pelo mecanismo categorial (que têm um papel semelhante ao papel das *estruturas profundas* da teoria-padrão da Gramática Gerativa: servem de *input* para a interpretação semântica) aplica-se um conjunto de regras transformacionais (mais ou menos equivalente às “*scrambling rules*” da teoria-padrão) que reordenam os elementos da expressão de forma a obtermos as sentenças da língua.

No fundo, a GC de Lewis é apenas uma proposta alternativa para o componente categorial (ou BASE) da teoria-padrão. Como na Gramática Gerativa, o componente categorial, juntamente com o léxico, gera as estruturas profundas, que serão convertidas em estruturas superficiais pelas regras transformacionais. Na gramática de Lewis – descendente direta do algoritmo de Ajdukiewicz – as estruturas profundas representam a estrutura semântica da sentença, escrita em notação polonesa.⁷

Em síntese, então, a gramática de Lewis consiste basicamente numa GC (essencialmente idêntica ao algoritmo de Ajdukiewicz) que “gera” um conjunto de “estruturas profundas” que podem ser facilmente interpretadas (traduzidas, projetadas) em expressões do cálculo de predicados que, por sua vez, servem de representações para os “significados”.

Essas “estruturas profundas” sofrem a ação de um componente transformacional para que possamos obter as “estruturas superficiais” correspondentes, que serão interpretadas foneticamente.

Em outras palavras, o modelo de gramática proposto por Lewis é essencialmente o modelo chomskiano conhecido como “teoria-padrão”, com um mecanismo categorial no lugar das regras de reescritura.

A análise que Ilari propõe para os casos de ambigüidade de identidade estrita e fajuta dependem crucialmente dessa “forma” da gramática de Lewis. Passemos à análise.

3.2. A análise de Ilari

Tomemos dois exemplos.

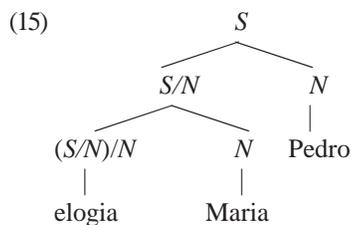
(13) Pedro elogia Maria

(14) Pedro se elogia

Já vimos acima que *elogiar* é da categoria $(S/N)/N^8$ e que *Pedro* e *Maria* são N . A análise de (13), então, é relativamente simples e pode ser representada pela árvore seguinte:

⁷ Em linhas gerais, a notação polonesa – usada para notar as expressões dos cálculos lógicos sem necessitar de parênteses ou de outros sinais de pontuação para evitar as ambigüidades – sempre escreve o funtor antes dos argumentos.

⁸ A notação categorial de Ilari difere ligeiramente da de Lewis, com o primeiro usando parêntesis angulados onde o segundo sua parêntesis comuns; há também um excesso de parentetização na notação de Ilari. Por isso, preferimos aqui manter a notação de Lewis.



O primeiro problema de Ilari, então, é categorizar o reflexivo *se*, de (14).

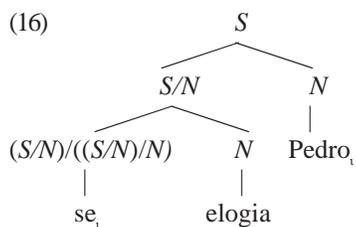
Uma primeira alternativa – óbvia – é incluir o *se* na categoria *N*. Esta alternativa, no entanto, não nos explicaria o fato de que o *se*, ao contrário dos outros membros da categoria *N*, não pode ocupar a posição de sujeito da sentença. Em outras palavras, embora possa muitas vezes ocorrer em posições típicas de *N*, nem sempre isso acontece.⁹

Nas palavras de Ilari [13, ps. 124-125]:

As duas alternativas mais óbvias para o aparecimento de *se* apenas na posição de objeto consistem (1) em incluir *se* numa categoria operador (ao invés de incluí-lo entre os possíveis argumentos de um verbo transitivo) e (2) em atribuir seu aparecimento a um componente transformacional que operaria sobre estruturas categoriais com forma especificada.

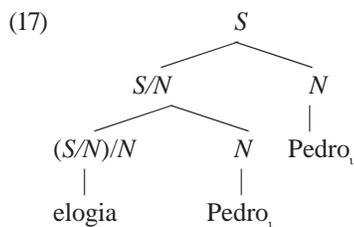
Ilari vai assumir as duas alternativas, simultaneamente, e propor que existem em português dois *se* distintos. Sua hipótese assume uma “ambigüidade” inerente do reflexivo relacionada a sua dupla origem sintática. Reproduzimos aqui a hipótese:

- (A) As duas interpretações de *se* correspondem a duas origens sintáticas distintas;
- (B) Numa primeira origem possível, *se* aparece desde a estrutura profunda, onde é um operador que forma verbos intransitivos a partir de verbos transitivos; sua categoria exprime-se como $(S/N)/((S/N)/N)$. Confirma a estrutura profunda de *Pedro se elogia* em (16), abaixo.



- (C) Numa segunda origem possível, *se* não aparece representado na estrutura profunda; é um vestígio, deixado por uma transformação que cancela o segundo de dois *N* idênticos, em estruturas como (17).

⁹ Ilari não explicita o fato de que o *se*, mesmo sendo tipicamente um objeto direto, só aparece antes do verbo transitivo nos seus exemplos. Como a questão do ordenamento dos clíticos ultrapassa os limites propostos para o presente texto, ela não será abordada tampouco aqui, apesar de ser uma questão de extrema pertinência.

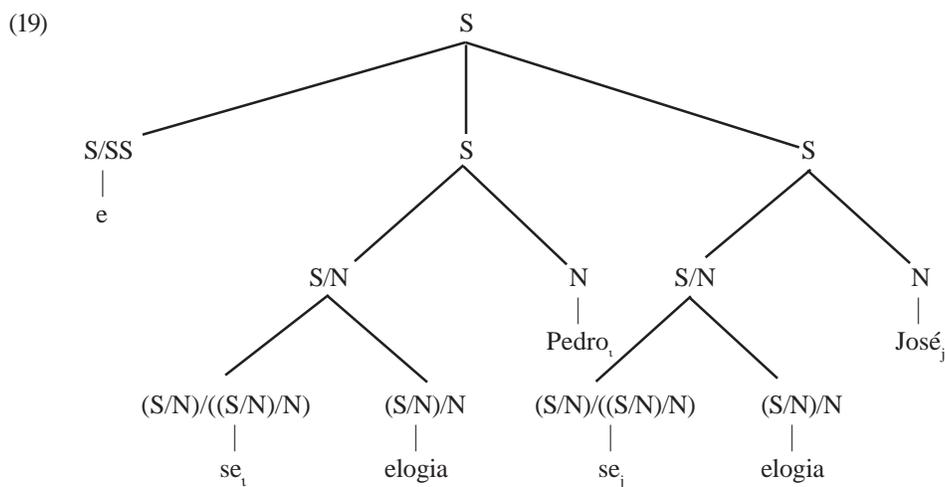


O passo seguinte é analisar os casos (1)–(5) a partir desta análise do reflexivo.¹⁰ Vejamos o exemplo (18).

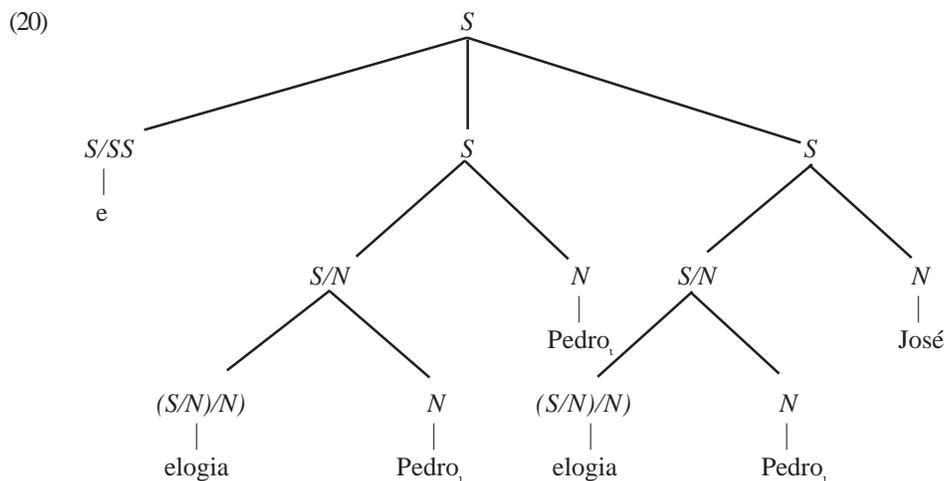
(18) Pedro se elogia e José faz o mesmo

Para Ilari, *fazer o mesmo* é uma expressão anafórica que retoma um *S/N* presente na sentença e sua origem é uma transformação que substitui “o segundo de dois *S/N*s que sejam *variantes alfabéticas*” (p. 138).

Assim, as duas leituras de (18) podem ser obtidas a partir das duas seguintes estruturas profundas:



¹⁰ Ilari mostra que é possível um tratamento para os possessivos exatamente nos mesmos termos. Não vamos reproduzir sua argumentação aqui.



Tanto em (19) quanto em (20), o segundo *S/N* “idêntico” (*se elogia* em (19) e *elogia Pedro* em (20)) é substituído transformacionalmente por *faz o mesmo*. Como a interpretação semântica se faz sobre as estruturas profundas, as duas estruturas superficiais resultantes, apesar de idênticas, terão distintos significados.¹¹

Este é, em linhas gerais e sem fazer total justiça à análise apresentada por ele, o tratamento proposto por Ilari para a ambigüidade de identidade estrita e fajuta.¹² A análise dos outros casos de ambigüidade (*Somente Pedro se elogia, Pedro se elogia e Maria também*, etc.) segue o mesmo raciocínio.

3.3. Comentários sobre a análise de Ilari

A análise dos casos de ambigüidade de identidade estrita e fajuta presente em [13] é obviamente datada e não poderia ser adotada hoje como uma explicação plausível para esses fenômenos. E são várias as razões para que afirmemos isso. Vejamos algumas delas.

3.3.1. O modelo da “Semântica Gerativa”

O modelo de gramática adotado por David Lewis (e por Ilari) entrou em desuso e caiu no esquecimento desde o fim dos anos 1970.

A Semântica Gerativa (SG) surgiu no final dos anos 1970 com toda a força e, na época, parecia ser uma verdadeira alternativa ao modelo, puramente sintático, proposto por Chomsky. Embora a SG surgisse como um mero rótulo que agrupava uma série de análises de fenômenos lingüísticos pontuais, produzidas por diversos lingüistas que, em princípio,

¹¹ Um conjunto de transformações de “*scrambling*” se encarrega de “colocar em ordem” as palavras de forma a obtermos as sentenças do português.

¹² A análise de Ilari envolve o cálculo- λ e consegue dar conta de estruturas encaixadas e de sintagmas nominais quantificados. Estamos ignorando aqui estas questões.

acreditavam estar trabalhando no interior da Gramática Gerativa (GG), as modificações que essas análises propunham no modelo original da GG eram tantas e de tal monta que pareceu, aos lingüistas da época, que se estava diante de uma nova forma de encarar as línguas naturais.

Num claro fenômeno de “luta pelo poder”, os lingüistas que propunham essas modificações no modelo chomskiano¹³ incentivaram esse entendimento e passaram a se considerar um “grupo revolucionário” que estava propondo um novo paradigma (no sentido kuhiano do termo) em substituição ao paradigma chomskiano.

No embate que se seguiu entre os dois “paradigmas” em confronto – embate duro, que perpassou os anos 1970 e que chegou a ser chamado de “guerras lingüísticas”¹⁴ – dois resultados principais se registraram: (i) as propostas da SG foram rechaçadas e praticamente todos os lingüistas que as apoiavam ou retornaram ao modelo da GG ou passaram a se dedicar a outros assuntos; (ii) as alterações que Chomsky propôs para o modelo da GG – como forma de enfrentar as críticas feitas pela SG – deram uma nova direção à teoria, que a encaminhou no sentido da desvalorização das regras (principalmente das regras transformacionais) e na valorização dos princípios gerais. Em suma, do confronto entre a SG e a GG resultou, por um lado, a teoria dos Princípios e Parâmetros¹⁵ e, por outro lado, o abandono completo das teses principais da SG.

Embora haja muita intuição de alta qualidade nos trabalhos dos semanticistas gerativos – muitos deles sendo recuperados hoje pela GG – os mecanismos analíticos utilizados por eles não são (e não podem ser) mais utilizados pelos lingüistas. A SG atribuía uma força extrema às regras transformacionais, que podiam fazer de tudo e em qualquer lugar e momento da derivação das sentenças. Os trabalhos seminiais de Peters e Ritchie [23, 24, 25] descreditaram completamente essa forma de organizar uma gramática.

O simples fato, então, de ser conduzida no interior de um quadro teórico “antigo” já nos levaria a sugerir que a análise de Ilari mereceria, no mínimo, uma revisão num quadro teórico mais consistente. Mas não é só.

3.3.2. As regras transformacionais supostas

A análise de Ilari repousa crucialmente sobre uma regra transformacional de *supressão de SN idêntico* e de regras, também transformacionais, de “*scrambling*”, que reordenam os elementos terminais da estrutura superficial antes que as regras fonológicas façam sua interpretação.

Embora possíveis no quadro da SG – e mesmo no quadro da GG dos anos 1970 – essas regras transformacionais não são mais usadas hoje.

As sucessivas restrições ao poder das regras transformacionais conjugadas à integração, cada vez mais forte, entre regras e princípios acabaram por eliminar completamente do repertório de regras possíveis numa Gramática Transformacional uma transformação como a de supressão de SN idêntico. Nas versões mais modernas da GG (e só podemos

¹³ James McCawley, Paul Postal, Charles Fillmore, John Robert Ross e George Lakoff, entre outros.

¹⁴ Ver, por exemplo, [10]

¹⁵ Esse processo, obviamente, levou algum tempo e comportou alguns “estágios” intermediários.

falar da GG, uma vez que a SG não mais existe), não há mais regras transformacionais que apaguem elementos. Na Gramática de Princípios e Parâmetros a única ação admitida para regras transformacionais é a de *mover* elementos.¹⁶

Em outras palavras, a hipótese de tratamento de Ilari para os casos de ambigüidade de identidade estrita e fajuta é claramente inviável, se considerarmos os quadros teóricos “transformacionalistas” da atualidade.

Resta-nos ver, no entanto, se não há uma “salvação” para a hipótese de Ilari no quadro teórico das GC.

3.3.3. As GC

O tratamento das relações anafóricas sempre foi um grande desafio para as GC. Dentre outras coisas, porque, de modo geral, essas gramáticas não admitem transformações.

As GC são gramáticas radicalmente lexicalistas e sua organização é basicamente a de “item-e-processo”, ao contrário das Gramáticas Transformacionais (GG e SG) que são modelos sintáticos e de organização por “item-e-arranjo”.¹⁷ Vejamos cada uma dessas características separadamente.

Uma gramática lexicalista supõe que as propriedades – sintáticas e semânticas – estão contidas nos itens lexicais e que, fora dos itens lexicais, não há nada. Cada item lexical contém, além de suas propriedades sintático-semânticas, as formas de concatenação com outros itens admitidas. Ou seja, não há uma “sintaxe autônoma” (nem uma “semântica autônoma”), independente do léxico.

Como consequência dessa postura lexicalista, o modelo de organização da gramática é “item-e-processo”: a sintaxe e a semântica são vistas como processos de concatenação de itens em que um deles (o funtor) “opera” sobre o outro (o argumento) para a obtenção de uma nova construção, que poderá ser um novo funtor ou um novo argumento.

Ao contrário, as teorias transformacionais (GG e SG) admitem uma sintaxe (ou uma semântica, no caso de SG) autônoma, completamente independente do léxico; e seu modelo de organização é “item-e-arranjo”: no caso de GG, a sintaxe constrói uma estrutura (uma “árvore”, um “arranjo”) e os itens lexicais são, num momento lógico posterior, inseridos nos nós terminais da estrutura. Estrutura e itens lexicais são completamente separados (são construídos separadamente e possuem natureza e comportamento completamente distintos).

É justamente esta diferença no modelo de organização (“item-e-processo” *versus* “item-e-arranjo”) que permite a existência de *categorias vazias* nas teorias transformacionais e impede sua existência nas GC (que propriedades sintático-semânticas um “item” vazio poderia ter?). Da mesma forma, é esta diferença de organização que permite que as teorias transformacionais admitam os movimentos, impedidos nas GC (um “item” pode mover-se de um lugar para outro na estrutura, mas se não há estrutura “fora” dos itens, não há para onde se moverem os itens).

¹⁶ Na verdade, a única transformação admitida é a regra “Mova α ”, que permite mover qualquer coisa para qualquer lugar, desde que o movimento seja permitido pelos princípios gerais.

¹⁷ “Item-e-arranjo” e “item-e-processo” são termos cunhados por Hockett [11] para designar diferentes modos de modelar a morfologia das línguas naturais. A utilização destes termos para caracterizar diferentes teorias sintáticas se deve a Schmerling [26].

De novo, a inadequação do quadro teórico utilizado por Ilari é patente. A gramática de Lewis não é, de fato, uma GC na extensão do termo. É uma Gramática Transformacional que substitui um mecanismo de reescritura por um mecanismo de natureza categorial, sem assumir compromissos maiores com esse mecanismo categorial.

Creemos, com isso, que comprovamos a afirmação inicial de que a hipótese de Ilari é datada e que os fenômenos abordados por ela estariam ainda à espera de uma explicação.

4. ANÁLISES ALTERNATIVAS

Vamos fazer um “vôo”, breve e panorâmico, sobre algumas hipóteses alternativas, que pretendem dar conta dos fenômenos relacionados à identidade estrita e fajuta. Não temos nenhuma pretensão de exaustividade – nem vertical, pelo aprofundamento de uma ou outra hipótese; nem horizontal, pela pretensão de abordar todas, ou mesmo uma boa parte, das hipóteses existentes – nem de chegar a alguma solução para a questão da identidade. Interessa-nos, apenas, apontar algumas linhas de raciocínio seguidas pelos lingüistas interessados na questão da identidade estrita e fajuta para fundamentar a reflexão que se seguirá.

4.1. O cálculo de Lambek

Na versão da GC conhecida como cálculo de Lambek, ao estilo da Dedução Natural, principalmente naquela associada a expressões do cálculo- λ como representação semântica, sugerida por van Benthem [3, cap. 7], as estruturas gramaticais das expressões analisadas são representadas através de árvores de demonstração construídas através de algumas poucas regras – um par de regra para cada um dos três conectivos categoriais: “/”, “\” e “•”; como não usaremos aqui o conectivo “•”, apresentamos abaixo os pares de regras apenas para os outros dois conectivos.

$\frac{Y:A \quad Y \setminus X:F}{X:(FA)} \quad E \setminus$	$\frac{X/Y:F \quad Y:A}{X:(FA)} \quad E /$
$\frac{\frac{\overline{Y:v} \quad \vdots}{\vdots \quad \vdots}}{X:T} \quad \Gamma^i$	$\frac{\frac{\vdots \quad \overline{Y:v}}{\vdots \quad \vdots}}{X:T} \quad \Gamma^i$
$\frac{}{X \setminus Y: \lambda v.T} \quad \Gamma^i$	$\frac{}{X \setminus Y: \lambda v.T} \quad \Gamma^i$

Nas regras de eliminação dos conectivos, o que se diz sintaticamente é que uma expressão da categoria funcional do tipo X/Y (ou $Y \setminus X$) se concatena com uma expressão da

categoria Y depois, no caso do conectivo “/” (antes, no caso do conectivo “\”), de forma que a expressão complexa resultante é da categoria X; na parte semântica, o significado da expressão concatenada é a aplicação da função F, relativa à expressão de categoria funcional, ao argumento A da outra expressão, resultando em (FA).

As regras de introdução dos conectivos é um pouco mais complexa, e envolvem a introdução de hipóteses que precisam ser eliminadas para que a derivação seja bem formada. Assim, se pudermos postular, num momento arbitrário de uma derivação, uma categoria Y, associada semanticamente a uma variável ν que ainda não tenha sido usada na derivação (marcada pelo índice i), e na continuação da derivação chegamos depois a uma categoria X, cuja denotação é representada pelo termo T, então podemos “descarregar” a hipótese i , introduzindo um dos dois conectivos categoriais (Y\X, caso a hipótese tenha sido formulada na periferia esquerda da derivação; X/Y, caso a hipótese tenha ocorrido na periferia direita); o resultado semântico desta operação é sempre uma abstração- λ da variável ν , introduzida pela hipótese, sobre o termo T ($\lambda\nu:T$).

Como estamos recorrendo a termos- λ , sempre é útil dispormos ainda da redução- β , que simplifica termos complexos, sem afetar a sua identidade. Essa redução garante que termos com o formato ($\lambda\alpha: \gamma\beta$) é equivalente a $\lambda^{\gamma\alpha\rightarrow\beta}$ (este último termo equivale ao termo γ com todas as instâncias livres da variável α substituídas pelo termo β ; normalmente, ainda seria preciso garantir que não estaríamos fazendo ligações indevidas de variáveis nesta substituição, mas como tomaremos o cuidado de sempre usar variáveis novas na derivação, esse cuidado passa a ser desnecessário).

Sem entrar em muitos detalhes técnicos, pode-se dizer que o cálculo de Lambek constitui o “núcleo” de uma GC. Foi Lambek quem conseguiu dar uma formulação rigorosamente matemática para as GC propostas por Ajdukiewicz [1], Bar-Hillel[2], Geach [9] e Flynn [7], de forma que as regras presentes nessas gramáticas deixassem de ser axiomas e passassem a ser deduzidas, como teoremas, dos comportamentos dos conetivos.

Assim, um exemplo é o da dedutibilidade da chamada “promoção de tipo” (*type raising*) no cálculo de Lambek, sem que ela precise ser estipulada, como tem sido em algumas teorias semânticas. Segundo a promoção de tipo, um N como *Pedro*, que denota o indivíduo p , pode ser transformado num $S/(N\backslash S)$, denotando $\lambda P_0:(P_0 p)$. Dessa maneira, em versões mais clássicas da GC (em que a promoção de tipo é chamada de R3), a sentença *Pedro corre* seria deduzida da seguinte maneira:

$$\begin{array}{c}
 \begin{array}{cc}
 \text{Pedro} & \text{corre} \\
 \hline
 N : p & M\backslash S : C
 \end{array} \\
 \begin{array}{c}
 \text{Lex} \qquad \text{Lex} \\
 \hline
 \text{R3} \\
 S/(N\backslash S) : \lambda P_0.(P_0 p)
 \end{array} \\
 \hline
 S : (\lambda P_0.(P_0 p) C) \\
 =_{\text{red}\beta} (C p) \\
 \text{R1}
 \end{array}$$

No entanto, a regra R3 não precisa ser estipula no cálculo de Lambek, já que sua validade pode ser demonstrada a partir daquelas regras de introdução e eliminação apresentadas antes, como podemos ver abaixo.

$$\frac{\frac{\frac{X : A \quad \frac{\frac{}{X \setminus Y : P_0} 1}{\quad} E \setminus}{\quad} I^{-1}}{Y : (P_0 A)} E \setminus}{\quad} I^{-1}}{Y / (X \setminus Y) : \lambda P_0 . (P_0 A)}$$

Portanto, a derivação da mesma sentença *Pedro corre* pode ser feita sem a estipulação da promoção de tipo, usando diretamente o esquema acima, como se pode ver na seguinte árvore:

$$\frac{\frac{\frac{\frac{\text{Pedro}}{N : p} \text{Lex} \quad \frac{\frac{\frac{}{N \setminus S : P_0} 1}{\quad} E \setminus}{\quad} I^{-1}}{S : (P_0 p)} E \setminus}{\quad} I^{-1}}{S / (N \setminus S) : \lambda P_0 . (P_0 p)} E /}{\quad} E /}{S : (\lambda P_0 . (P_0 p) C) =_{\text{red}; \beta} (C p)}$$

Esta derivação serve apenas como ilustração do tipo de operação permitida pelo cálculo de Lambek, já que ela não é uma derivação normal, o que indica que ela é mais complexa do que seria necessário; mas mesmo esse tipo de constatação sobre a complexidade desnecessária das derivações (o que permite reduzir as chamadas “ambigüidades espúrias”, que sempre foi um grande problema para as versões clássicas da GC) só foi possível a partir do sistema de Lambek.

4.2 Primeira tentativa (intuitiva, mas frustrada)

A partir da disponibilização do cálculo- λ , a primeira tentativa para representar o significado do pronome *se* poderia ser através da fórmula $\lambda P : \lambda x . ((P x) x)$, que funcionaria basicamente como um operador que toma um predicado de dois argumentos independentes e o transforma num predicado que unifica seus dois argumentos de uma única vez através de duas instâncias da mesma variável, seguindo uma intuição de que o pronome *se* recupera na posição de objeto o mesmo referente da posição de sujeito. É preciso destacar que atribuindo a *se* a categoria $(N \setminus S) / ((N \setminus S) / N)$, não é mais necessário postular nenhuma transformação, ao contrário da explicação que Ilari oferecia. Assim a derivação de *Pedro_i se_i elogia* ficaria como:

$$\begin{array}{c}
 \text{Pedro} \qquad \qquad \qquad \text{se} \qquad \qquad \qquad \text{elogia} \\
 \hline
 \text{N} : p \quad \text{Lex} \quad \frac{\text{N} \setminus \text{S} / ((\text{N} \setminus \text{S}) / \text{N}) : \lambda P_0. \lambda x_0. ((P_0 x_0) x_0)}{\text{N} \setminus \text{S} : (\lambda P_0. \lambda x_0. ((P_0 x_0) x_0) E)} \quad \text{Lex} \quad \frac{\text{N} \setminus \text{S} / \text{N} : E}{\text{N} \setminus \text{S} : (\lambda P_0. \lambda x_0. ((P_0 x_0) x_0) E)} \\
 \hline
 \text{N} \setminus \text{S} : (\lambda P_0. \lambda x_0. ((P_0 x_0) x_0) E) \\
 =_{red, \beta} \lambda x_0. ((E x_0) x_0) \\
 \hline
 \text{S} : (\lambda x_0. ((E x_0) x_0) p) \\
 =_{red, \beta} ((E p) p)
 \end{array}$$

Para lidar então com a sentença *Pedro se elogia* e *José faz o mesmo*, poderíamos atribuir a *faz o mesmo* uma categoria que o fizesse uma função que toma um sujeito antes (*José*, no caso), que depois toma a conjunção *e*, para então tomar um predicado, também antes (*se elogia*, no caso) e finalmente tomar um segundo sujeito (*Pedro*), para resultar numa sentença. Do ponto de vista semântico, o significado de *faz o mesmo* precisa ser uma função que toma o significado do seu sujeito imediato, para resultar numa função que toma o significado da conjunção, para resultar numa função que toma o significado do predicado anterior, para resultar numa função que toma o sujeito do predicado anterior, e posiciona tudo de forma a obtermos a distribuição adequada entre os argumentos e as funções finais. Assim, supondo que *e* seja um $(S \setminus S) / S : \lambda P_1. \lambda P_0. P_0 \wedge P_1$, precisamos que *faz o mesmo* corresponda a $\mathcal{M}(((S \setminus S) / S) \setminus ((\mathcal{M} \setminus S) \setminus (\mathcal{M} \setminus S))) : \lambda x_3. \lambda x_2. \lambda x_1. \lambda x_0. ((x_2 (x_1 x_3)) (x_1 x_0))$. A complexidade notacional torna difícil a leitura do termo, mas na Figura 1, abaixo, podemos ver como ele se comporta na derivação da sentença *Pedro se elogia* e *José faz o mesmo*, na interpretação em que *Pedro_i* elogia *Pedro_i* e *José_j* elogia *José_j*.¹⁸

Ainda que esta solução dê conta de uma das leituras (especificamente, a leitura da identidade fajuta), com os recursos do cálculo de Lambek, não parece possível conseguir qualquer atribuição categorial às expressões básicas que fosse capaz de derivar a outra leitura (a da identidade estrita, em que *Pedro se elogia* e *José faz o mesmo* seria interpretada como se o *Pedro_i* e o *José* elogiassem o *Pedro_i*).

4.3 A análise de Jäger

Em três artigos [15, 17, 16] e um livro [18], Gerhard Jäger propõe uma solução para a questão da anáfora numa GC.¹⁹ Basicamente, sua proposta consiste em acrescentar ao conjunto de conetivos presentes no cálculo de Lambek um novo conetivo (“”) – que ele atribui a Pauline Jacobson – cuja função específica é a de retomar anaforicamente Expressões presentes no “discurso prévio”. Intuitivamente, podemos dizer que este conetivo apresenta as seguintes propriedades:

¹⁸ Nas derivações, as partes identificadas por *D* indicam uma abreviação daquela parte da derivação (ou porque aquela subderivação já foi feita em outro lugar, ou porque ela é irrelevante naquele momento); portanto, *D* cumpre uma função semelhante à dos triângulos nas árvores de estrutura sintagmática.

¹⁹ Na verdade, Jäger utiliza uma versão da GC conhecida como *Type-Logic Grammar*. Para maiores detalhes, ver [21], [5] e o próprio [18].

- a) Uma expressão da categoria $\alpha|\beta$ é uma expressão que precisa de um antecedente da categoria β para se comportar como uma expressão da categoria α .
- b) A denotação de uma expressão da categoria $\alpha|\beta$ é uma função que tem o domínio nas denotações de β e o contra-domínio nas denotações de α .

Do ponto de vista formal, este novo conectivo exige o acréscimo de um novo par de regras, ao cálculo de Lambek, para regulamentar a sua introdução e a sua eliminação nas derivações. Estas novas regras são:²⁰

$$\boxed{\begin{array}{c} Y|X:F \\ [X:A] \dots \frac{\quad}{Y:(FA)} \quad E| \end{array}}$$

²⁰ Por motivos estilísticos relativos ao formato das derivações, já que explicitamos a redução- β , estamos escolhendo escrever sempre a categoria antes da representação semântica; Jäger, ao contrário, escreve primeiro a representação semântica e depois a categoria. Esta, no entanto, é apenas uma divergência notacional que não afeta as observações feitas.

Pedro	se elogia	e	José	faz o mesmo	
———— Lex	———— D	———— Lex	———— Lex	———— D	
N : p	N\S : $\lambda x_4.((E x_4) x_4)$	(S\S)/S : $\lambda P_1. \lambda P_0. P_0 \wedge P_1$	N : j	$N(((S\S)/S)((N\S)(N\S))) :$ $\lambda x_3. \lambda x_2. \lambda x_1. \lambda x_0. ((x_2 (x_1 x_3)) (x_1 x_0))$	D
				$((S \S)/S)((N \S)(N \S)) :$ $(\lambda x_3. \lambda x_2. \lambda x_1. \lambda x_0. ((x_2 (x_1 x_3)) (x_1 x_0)) j)$ $=_{red:\beta} \lambda x_2. \lambda x_1. \lambda x_0. ((x_2 (x_1 j)) (x_1 x_0))$	E\
				$(N \S)(N \S) :$ $(\lambda x_2. \lambda x_1. \lambda x_0. ((x_2 (x_1 j)) (x_1 x_0)) \lambda P_1. \lambda P_0. P_0 \wedge P_1)$ $=_{red:\beta} \lambda x_1. \lambda x_0. ((\lambda P_1. \lambda P_0. P_0 \wedge P_1 (x_1 j)) (x_1 x_0))$ $=_{red:\beta} \lambda x_1. \lambda x_0. (\lambda P_0. P_0 \wedge (x_1 j) (x_1 x_0))$ $=_{red:\beta} \lambda x_1. \lambda x_0. (x_1 x_0) \wedge (x_1 j)$	E\
		$N \S : (\lambda x_1. \lambda x_0. (x_1 x_0) \wedge (x_1 j) \lambda x_4. ((E x_4) x_4))$ $=_{red:\beta} \lambda x_0. (\lambda x_4. ((E x_4) x_4) x_0) \wedge (\lambda x_4. ((E x_4) x_4) j)$ $=_{red:\beta} \lambda x_0. ((E x_0) x_0) \wedge (\lambda x_4. ((E x_4) x_4) j)$ $=_{red:\beta} \lambda x_0. ((E x_0) x_0) \wedge ((E j) j)$			E\
		$S : (\lambda x_0. ((E x_0) x_0) \wedge ((E j) j) p)$ $=_{red:\beta} ((E p) p) \wedge ((E j) j)$			E\

Figura 1: Derivação com reflexivo como “unificador” de argumentos

$Y_1 X:F_1$	$Y_2 X:F_2$	$Y_n X:F_n$
$\overline{\quad}^{i_1}$	$\overline{\quad}^{i_2}$	$\overline{\quad}^{i_n}$
$Y_1:(F_1 v)$	$Y_2:(F_2 v)$	$Y_n:(F_n v)$
⋮	⋮	⋮
⋮	⋮	⋮
$Z:T$		
$Z X.\lambda vT$		
$I^{(i_1, i_2, \dots, i_n)}$		

Na solução de Jäger, os pronomes anafóricos são expressões da categoria $N|N$ e denotam a função identidade sobre indivíduos; ou seja, aos pronomes anafóricos se associa o termo $\lambda x_n.x_n$. Os prossintagmas verbais (como *fazer o mesmo*) são expressões da categoria $(N\backslash S)|(N\backslash S)$ e sua denotação é associada ao termo $\lambda P_n.P_n$.²¹ Como em inglês os pronomes reflexivos ocorrem na posição canônica do objeto direto, eles também podem receber a mesma categoria dos pronomes que ocorrem na posição de sujeito; nos nossos exemplos em português, por outro lado, como os pronomes reflexivos são clíticos que antecedem o verbo,²² eles não só não podem ser atribuídos à mesma categoria como ainda precisam receber uma interpretação semântica diferente, que faça jus à sua distribuição.

Então, um pronome como *se* terá que pertencer à categoria $((N\backslash S)/((N\backslash S)/N))|N$ e sua representação semântica precisará ser $\lambda x_n.\lambda P_m.(P_m x_n)$.²³

De volta à análise, uma sentença como (18), que reproduzimos abaixo como (21), poderia receber as análises que passaremos a apresentar a seguir.

(21) Pedro se elogia e José faz o mesmo Primeiro, como expressão da categoria $((N\backslash S)/((N\backslash S)/N))|N$ (e, portanto, como pronome anafórico), o *se* procuraria um antecedente da categoria N para que pudesse assumir as funções de um $(N\backslash S)/((N\backslash S)/N)$. O antecedente óbvio é o N *Pedro*, o que nos deixaria com uma leitura de (21) equivalente a (22) abaixo:

(22) Pedro_i se_i elogia e José faz o mesmo.

²¹ Estamos usando, a partir de agora, variáveis específicas para identificar seus tipos (antes elas estavam sendo usadas sem que seus tipos estivessem explicitamente representados); assim, x_n indica uma variável de tipo e (variável de indivíduo) e P_n indica uma variável de tipo $\langle e, t \rangle$ (variável de conjunto de indivíduos) – em que n deve ser um número inteiro ainda não usado para identificar variáveis na derivação, de forma a se evitar ligação indevida de variáveis.

²² Evidentemente, a posição dos clíticos é uma questão muito mais complicada, e explicar seu ordenamento na sentença é um problema que ultrapassa os limites dos objetivos do presente texto.

²³ Diferentemente da escolha de Jäger, em que o reflexivo *se* comportava exclusivamente como um argumento, agora ele é um funtor, que primeiro toma como argumento o antecedente anafórico, e depois um predicado binário, para se tornar um predicado unário reflexivizado. Do ponto de vista semântico, o reflexivo do português passa por um tipo de promoção de tipo.

Em seguida, o prossintagma *fazer o mesmo*, como uma expressão $(N \setminus S) \setminus (N \setminus S)$, procuraria um antecedente da categoria $N \setminus S$ para assumir as funções de um $N \setminus S$. E o antecedente óbvio é o $N \setminus S$ *se_i elogia*. O resultado final é o que temos em (23):

(23) Pedro_i [*se_i elogia*]_j e José [*faz o mesmo*]_j.

Se fizéssemos acompanhar a análise sintática da interpretação semântica correspondente, o resultado seria:

(24) $((E p) p) \wedge ((E p) j)$

A representação estrutural em (23) e a fórmula em (24) apresentam a leitura de identidade estrita. Esta leitura pode ser obtida de acordo com a Figura 2.

Para obtermos a leitura de identidade fajuta, o procedimento é um pouco diferente. Primeiro, o prossintagma *fazer o mesmo* procura um antecedente de categoria $N \setminus S$ e encontra a expressão *se elogia*. O resultado desta operação está em (25):

(25) Pedro [*se elogia*]_i e José [*faz o mesmo*]_j.

Como *se elogia* e *faz o mesmo* são correferentes nesta análise estrutural, eles são funcionalmente equivalentes, e sua estrutura pode ser alternativamente concebida como em (26).

(26) Pedro [*se elogia*]_i e José [*se elogia*]_j.

Note-se que, na interpretação semântica de (26), chegamos a duas ocorrências do $N \setminus S$ *se elogia*. Semanticamente, portanto, (26) pode ser parcialmente representado como (27):

(27) $((E x) p) \wedge ((E y) j)$

No passo seguinte, cada ocorrência do pronome anafórico *se* procura um antecedente de categoria N e o resultado final é o que está em (28), com a semântica de (29).²⁴

(28) Pedro_j [*se_j elogia*]_i e José_k [*se_k elogia*]_i

(29) $((E p) p) \wedge ((E j) j)$

A derivação da leitura de identidade fajuta é apresentada, então, na Figura 3.²⁵

Na verdade, uma vantagem que a solução de Jäger oferece, e que Ilari nem chegou a mencionar em seu texto, é a de identificar que sentenças como *Pedro bate em sua mulher*

²⁴ Toda a explicação acima, claro, é informal e não corresponde aos passos derivacionais apresentados na Figura 3.

²⁵ Desta vez, para economizar espaço, a representação da conjunção foi substituída pelo sinal “. . .”; mas ela é exatamente a mesma já usada aqui em duas ocasiões anteriores. Ainda para economizar espaço, a aplicação funcional e a redução- β em E³ também foram omitidas.

e *José faz o mesmo* apresentam uma terceira interpretação, além das leituras da identidade estrita e fajuta: o pronome *sua* pode ser interpretado deiticamente, caso em que a sua retomada anafórica pelo prossintagma verbal *faz o mesmo* recuperaria esse mesmo antecedente dêitico.²⁶ Esses passos intuitivos podem ser acompanhados nas três representações abaixo:

²⁶ Conforme apontado por um dos pareceristas anônimos, o tratamento oferecido aqui para o reflexivo *se* permite a derivação de uma leitura dêitica para *Pedro se elogia*; além disso, o reflexivo apresenta algumas limitações para recuperar seu antecedente (como em *o pai do João se elogia e João disse que Pedro se elogia*, em que João não pode ser antecedente de nenhuma das ocorrências de *se* nestas sentenças), que nossa análise não captura. No entanto, como o objetivo do presente texto não é dar conta da distribuição precisa do reflexivo, mas sim o da interação de duas expressões anafóricas (*se* ou *sua*, numa primeira sentença, e *faz o mesmo*, numa segunda sentença), não nos preocuparemos com estas questões aqui. Para evitar qualquer complicação, estamos trocando o exemplo para o pronome possessivo, que efetivamente apresenta a leitura dêitica. Uma possibilidade de solução destas questões talvez envolvesse o uso de conectivos categoriais modais, como em [21, ps. 157-160]; outra alternativa seria manter a sobregeração, restringindo-a através de fatores pragmáticos, como em [12]. Mas isso tudo estaria muito além das preocupações do presente texto.

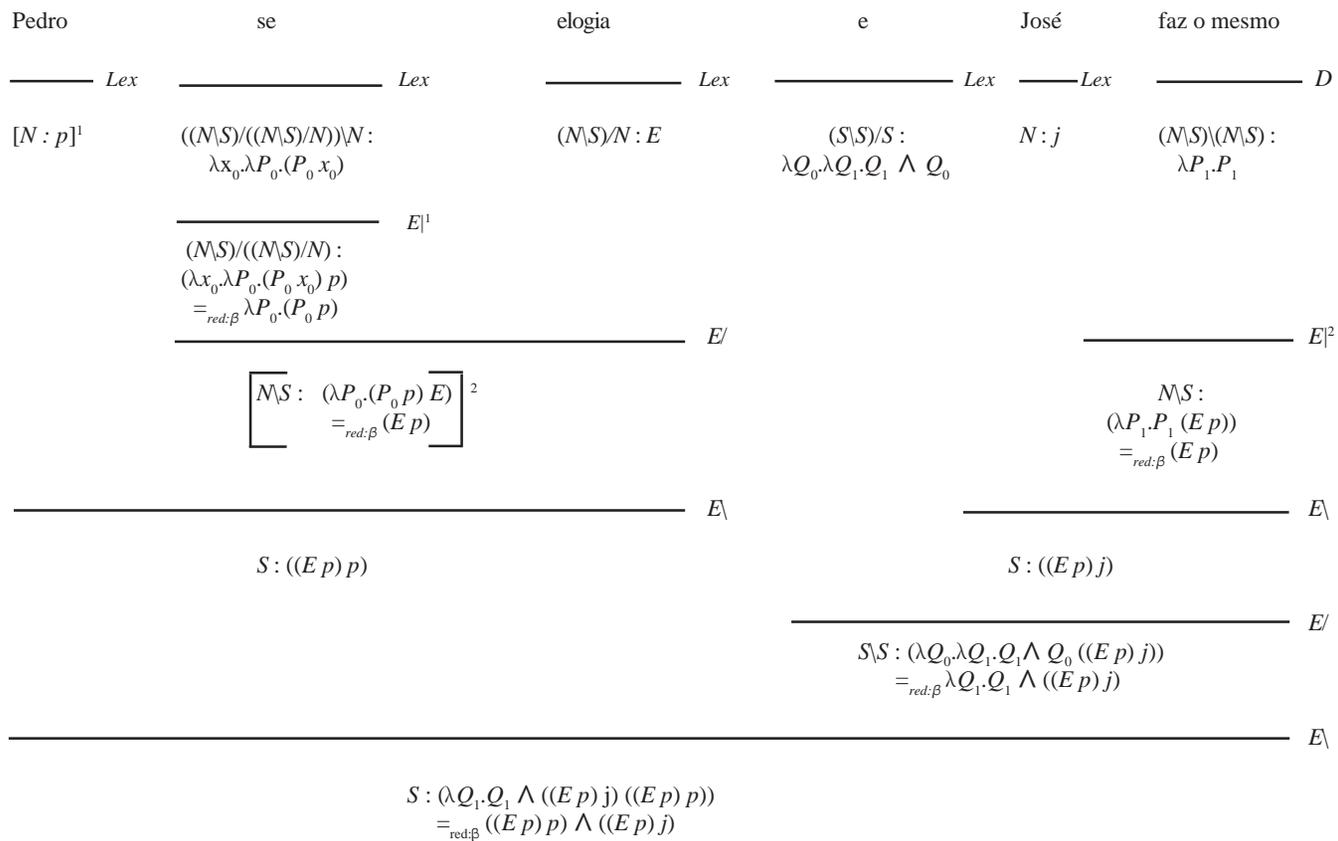


Figura 2: Derivação da leitura de identidade estrita para *Pedro se elogia e José faz o mesmo*

Pedro		se		elogia		e		José		faz o mesmo	
N : p	Lex	[N : x ₁] ²	1	(N\S)/((N\S)/N) N : λ _{x₀} .λP ₀ .(P ₀ x ₀)	Lex	(N\S)/N : E	Lex	N : j	Lex	(N\S) (N\S) : λP ₁ .P ₁	D
				(N\S)/((N\S)/N) : (λ _{x₀} .λP ₀ .(P ₀ x ₀) x ₁) = _{red:β} λP ₀ .(P ₀ x ₁)						N\S : λx ₁ .(E x ₁) x ₁)	E\ ³
				N\S : (λP ₀ .(P ₀ x ₁) E) = _{red:β} (E x ₁)						S : (λx ₁ .(E x ₁) x ₁) j) = _{red:β} ((E j) j)	E\
				S : ((E x ₁) x ₁)						S\S : (λQ ₀ .λQ ₁ .Q ₁ ∧ Q ₀ ((E j) j)) = _{red:β} λQ ₁ .Q ₁ ∧ ((E j) j)	E/
				[N\S : λx ₁ .(E x ₁) x ₁] ³							E\
				S : (λx ₁ .(E x ₁) x ₁) p) = _{red:β} ((E p) p)							E\
				S : (λQ ₁ .Q ₁ ∧ ((E j) j) ((E p) p)) = _{red:β} ((E p) p) ∧ ((E j) j)							E\

Figura 3: Derivação da leitura de identidade fajuta para *Pedro se elogia e José faz o mesmo*

(30) Pedro_i [bate em sua_j mulher]_i e José_k [faz o mesmo]_i

(31) Pedro_i bate em sua_j mulhe e José_k bate em sua_j mulher

(32) $((B(Mx_0))p) \wedge ((B(Mx_0))j)$

Para facilitar a derivação (evitando ter de lidar com a preposição), vamos trocar o exemplo (aproveitando para escolher um sentença mais “politicamente correta”): *Pedro surpreende sua mulher e José faz o mesmo*. As leituras de identidade estrita e fajuta para o pronome possessivo sua seriam completamente semelhantes às da sentença *Pedro se elogia e José faz o mesmo*; já sua interpretação dêitica é obtida através do uso da regra I], como podemos constatar na Figura 4.²⁷

Com a representação semântica obtida acima, podemos ainda facilmente explicar uma sentença como Antônio, *Pedro surpreende sua mulher e José faz o mesmo*. Eliminando o conectivo anafórico, o resultado da combinação da “sentença anafórica” com *Antônio é S*: $((S(Ma))p) \wedge ((S(Ma))j)$; ou seja, ‘Pedro surpreende a mulher de Antônio e José surpreende a mulher de Antônio’.

5. CONCLUSÕES E PROSPECÇÕES

A principal observação sobre as análises apresentadas acima está relacionada ao pioneirismo de Ilari. Apesar de Jäger [18, p. 76] afirmar que “a primeira tentativa de integrar a resolução da anáfora à Gramática Categorial se deve a Anna Szabolcsi” [27], num texto de 1989, a data da publicação do texto de Ilari é 1981! No entanto, como o próprio Ilari [13, p. 56] diz, suas principais idéias sobre o assunto já tinham sido “expostas, em forma de comunicação, numa mesa-redonda sobre ‘Os componentes sintático e semântico na descrição das línguas naturais’, organizada no âmbito de [sic] XXX Reunião Anual da Sociedade Brasileira para o Progresso da Ciência” (Ilari não diz, mas essa reunião da SBPC aconteceu em 1978); como ele não lista entre seus objetivos, naquela ocasião, a apresentação de uma solução categorial, apenas menciona o fenômeno empírico da ambigüidade estrita e fajuta, a preferência por uma solução não pragmática para essa ambigüidade e a incapacidade dos recursos formais utilizados por Jackendoff [14], vemos uma possibilidade de que a solução categorial possa ter sido incluída entre a apresentação em 1978 e sua publicação em 1981. Claro, Ilari escreveu em português, enquanto Szabolcsi escreveu em inglês.

Em segundo lugar, gostaríamos de comparar a nossa solução com a de Ilari, pois ambas exploram criativamente os recursos de boas teorias à nossa disposição nos respectivos momentos históricos. Em 1981, Ilari recorre à proposta de Lewis para encontrar uma

²⁷ Estamos tomando a liberdade de sobrecarregar o termo *M*, usando-o não só predicativamente, mas também funcionalmente; assim, na derivação acima, (Mx_0) não é uma proposição aberta ‘ x_0 é mulher’, mas sim a expressão relacional ‘a mulher de x_0 ’. Esta liberdade não causa prejuízo à análise adotada aqui, porque ela pode ser entendida como uma espécie de abreviação de um termo como ‘ $x_1:(Mx_1) \wedge ((DEx_1)x_0)$ ’, em que *M* corresponde a um predicado ‘ser mulher’ e *DE* é um predicado que representa a parte relacional.

solução para os reflexivos que não estava naquele autor; nós, hoje, recorreremos à proposta de Jäger para resolver o mesmo problema, agora sem necessidade de transformações, mas também precisamos encontrar novas atribuições categoriais para o reflexivo clítico no português, que ocupa uma posição diferente do reflexivo no inglês (assim, não pudemos aceitar a identidade como representação semântica do reflexivo, e precisamos descobrir

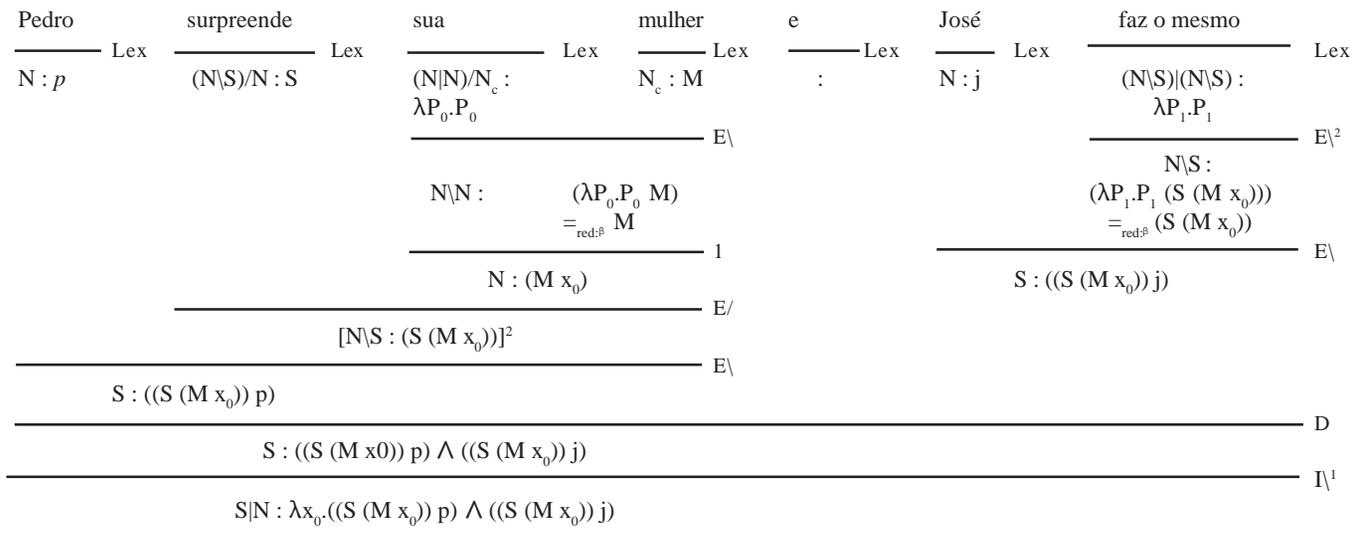


Figura 4: Derivação da leitura dêitica para Pedro surpreende sua mulher e José faz o mesmo 21

uma função que correspondesse à sua interpretação; além disso, não pudemos aceitar também que ele era apenas um elemento nominal, já que, enquanto anáfora ($\dots|N$), ele se combina com um $(M|S)/N$ para resultar num $M|S$. Por infelicidade de Ilari, ele escolheu uma versão que recorria a transformações, e como as transformações sofreram um “ataque”, mesmo dentro da própria GG, e acabaram sendo completamente repudiadas na GC, sua solução acabou desprestigiada em ambos os programas. Aguardaremos outros desenvolvimentos da Lingüística para sabermos a que acidentes históricos a nossa solução, apresentada aqui, será submetida.

Ainda sobre comparações, destacamos duas diferenças entre a solução de Jäger e a nossa, e mais três entre a nossa e a de Ilari.

Para Jäger, que estava apresentando sua proposta exemplificada com o inglês, bastava que o pronome reflexivo *himself* fosse um $N|N : \lambda x.x$, porque ele busca um N como antecedente e se comporta como um N no resto da derivação, correspondendo à função de identidade. No português, essa análise não se sustenta porque o reflexivo é clítico e antecede o verbo (pelo menos nos poucos exemplos que herdamos de Ilari); portanto, depois de buscar o seu antecedente ($\dots|N$), optamos por fazer com que o clítico do português fosse uma função que toma um verbo transitivo direto ($(M|S)/N$) e nos retorna uma outra função $(M|S)$ que toma um N e resulta num S . Na parte semântica, não foi possível aproveitar a identidade, porque nosso clítico precisa se tornar um quantificador generalizado, depois de encontrar seu antecedente anafórico, para poder se combinar com um predicador.

Jäger também não diz explicitamente em seus textos a que categoria o pronome possessivo deve pertencer, nem qual é sua interpretação; podemos até deduzir de algumas derivações, como [18, p. 188], que ele deve ser um $(N|N)/N_c : \lambda P_n.P_n$, porque *his* precisa se combinar com o nome comum *paper* para resultar no elemento que receberá o antecedente anafórico como argumento. Nas derivações em que *his paper* aparece, sua subderivação não é apresentada (na verdade, no livro, ela é equivocadamente atribuída ao léxico), o que nos obrigou a desmembrá-la.

Na comparação com a solução de Ilari, este postula a necessidade de dois *se* diferentes: um introduzido lexicalmente, e outro definido transformacionalmente; na nossa solução, precisamos sempre de apenas um único *se* ($((M|S)/((M|S)/N))|N : \lambda x_n.\lambda P_m.(P_m x_n)$), que tem sua resolução anafórica resolvida no processo da construção da derivação. Portanto, nossa solução é ontologicamente mais simples do que a de Ilari.

No texto de Ilari, o prossintagma verbal *fazer o mesmo* era introduzido transformacionalmente pelo apagamento de sintagmas equivalentes (*Pedro se elogia e José se elogia vira Pedro se elogia e José faz o mesmo*; ou *Pedro elogia Pedro e José elogia Pedro vira Pedro se elogia e José faz o mesmo*). Portanto, em sua análise, Ilari deixa de empregar o recurso categorial para tentar resolver o problema; como esse recurso foi empregado na nossa solução $((M|S)|(NnS) : \lambda P_n.P_n)$, de forma que precisamos novamente de uma única expressão e nenhuma transformação.

Em seu texto, Ilari não parecia preocupado com a terceira interpretação de *Pedro bate em sua mulher e José faz o mesmo*, em que o pronome possessivo pode ser interpretado deiticamente; portanto, ele não apresentava uma derivação para esta interpretação. Em nossa análise, esta terceira interpretação decorre automaticamente das escolhas categoriais

e dos procedimentos derivacionais, de forma que a interpretação dêitica é construída empregando a regra de introdução de |.

Depois de nos “vangloriarmos” do que fizemos de bom, falta agora reconhecer que há ainda muita coisa por fazer. Passaremos a listar algumas das tarefas que nos pareceram mais evidentes.²⁸

Nos exemplos usados por Ilari, e que adotamos aqui, os reflexivos só ocupam a posição pré-verbal; mas não seria difícil nos depararmos com exemplos como *Pedro feriu-se e José fez o mesmo*, no qual o reflexivo aparece depois do verbo, o que exigiria uma outra atribuição categorial (o que tornaria a introduzir dois reflexivos diferentes: um que aparece antes e outro depois do verbo) ou uma solução categorial radicalmente diferente que lidasse com as variabilidades distribucionais (recorrendo a modalizações de permutação). Ainda nesse mesmo sentido, seria preciso considerar sentenças com verbos intrinsecamente reflexivos, como *Pedro suicidou-se e José fez o mesmo*, que só admitem a interpretação da identidade fajuta (já que José não pode suicidar Pedro).

Outro fenômeno do mesmo tipo, para o qual se espera que as análises propostas aqui ajudem a explicar, é o das construções com os chamados pronomes relativos. É possível que sintagmas nominais como *a bola que o menino furou*, *a bola que bateu no menino e a bola que o menino furou ela* também possam ser integrados nesta mesma perspectiva, mas esse trabalho ainda precisa ser feito.

Ainda que Ilari apontasse a distinção entre prossintagmas verbais, como *fazer o mesmo*, e indicadores de elipses de SV, como *também e não*, nem ele nem nós oferecemos qualquer tentativa de explicação para a elipse. Numa observação superficial, pode-se prever que a elipse poderia ser explicada pelos mesmos recursos anafóricos usados aqui, de forma que o elemento indicador da elipse, quando estivesse acompanhado do material não elidido, se combinasse com este antes de virar uma expressão anafórica; quando efetivamente ocorrer a elisão, a derivação talvez precisasse introduzir uma hipótese correspondente ao material elidido, que depois precisaria ser “descarregada” para que a derivação pudesse ser concluída. Mas esta reflexão é completamente subjetiva (para não usar o termo “fantasiosa”, empregado por Ilari [13, p. 72]), e precisaria ser explicitamente testada através da construção das respectivas derivações, junto com as atribuições categoriais adequadas.

Na solução apresentada aqui, sempre introduzimos *fazer o mesmo* sem mostrar como seus elementos básicos colaboram nas derivações em que eles aparecem. Numa abordagem composicional, o que se espera é que digamos explicitamente quais deveriam ser as atribuições categoriais de *fazer*, *o* e *mesmo*, de forma que o seu resultado como $(N \setminus S) \setminus (N \setminus S) : \lambda P_m . P_n$ decorresse diretamente daquelas atribuições aos elementos básicos. Estivemos evitando propositalmente essa questão porque o próprio Ilari já nos alertava, em seu texto

²⁸ O mesmo parecerista anônimo alega ter um julgamento, compartilhado por alguns falantes que ele diz ter consultado, segundo o qual sentenças como *Pedro se elogia e Antônio faz o mesmo* apresentariam somente a interpretação fajuta. Para uma semântica formal, esta parece ser uma questão de preferência de análise, e não propriamente uma questão de interpretação: ainda que os falantes não percebam imediatamente a leitura estrita, ela parece poder ser elicitada por um contexto adequado. No entanto, claro, esta é uma questão empírica que também mereceria mais estudos, mas que pode envolver aspectos pragmáticos, psicolingüísticos ou sociolingüísticos, e por isso não a consideraremos na lista a seguir.

[13, ps. 89 e 90], que a escolha do verbo do prossintagma verbal “depende de propriedades lexicais do verbo que é objeto de ‘retomada’” e que “cada SV está para o prossintagma que o substitui como uma expressão de sentido mais específico está para uma expressão de sentido mais geral”, pois ainda é preciso descobrir se esta relação de hiperonímia é suficiente, ou se ainda outras relações lexicais podem interferir nesta escolha.

Finalmente, nossa proposta sobre o vocativo só funciona, como no exemplo apresentado, quando ele antecede a sentença; se ele fosse invertido, como em *Pedro suprême sua mulher, e José faz o mesmo, Antônio*, não há como a relação catafórica ser adequadamente estabelecida. Com esta observação, estamos reconhecendo que nossa solução importa de Jäger uma ferramenta estritamente anafórica, que não pode ser integrada de forma alguma com a catáfora. Assim, se existem elementos lingüísticos catafóricos que são essencialmente aqueles mesmos que também são usados nas anáforas, então era de se esperar que anáforas e catáforas pudessem ser integradas. Há uma esperança ainda de que esse mesmo problema posicional entre anáforas e catáforas possa ser resolvido pelos mesmos recursos para lidar com a variabilidade distribucional dos reflexivos.

Finalmente, como observado por um dos pareceristas anônimos, nossa análise efetivamente sobregera a interpretação dêitica para o reflexivo *se*; seria ainda preciso procurar alguma solução para impedir ou filtrar essa interpretação indevida, como admitimos em nota.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] AJDUKIEWICZ, Kazimierz. (1935). Die syntaktische konnexität. *Studia Philosophica*, 1:1-27.
- [2] BAR-HILLEL, Yehoshua. (1953). A quasi-arithmetical notation for syntactic description. *Language*, 29:47-58.
- [3] VAN BENTHEM, Johan. (1986). *Essays in Logical Semantics*. Reidel, Dordrecht.
- [4] BORGES NETO, José. (1979). Da gramática categorial. *Construtura*, 17:5-19.
- [5] CARPENTER, Bob. (1997). *Type-Logical Semantics*. The MIT Press, Cambridge.
- [6] DAVIDSON, Donald and HARMAN, Gilbert, editors. (1972). *Semantics of Natural Language*. Reidel, Dordrecht.
- [7] FLYNN, M. (1983). A categorial theory of structure building. In G. Gazdar, E. Klein, and G. Pullum, editors, *Order, Concord and Constituency*, pages 139-174. Foris, Dordrecht.
- [8] GEACH, Peter. (1962). *Reference and Generality*. Cornell University Press, Cornell.
- [9] _____. A program to syntax. In Davidson and Harman [6], pages 483-497.
- [10] HARRIS, R. A. (1993). *The Linguistic Wars*. Oxford University Press, Oxford.
- [11] HOCKETT, Charles F. (1954). Two models of grammatical description. *Word*, 10:210-231.
- [12] HUANG, Yan. (2004). Anaphora and the pragmatics-syntax interface. In Laurence R. Horn and Gregory Ward, editors, *The Handbook of Pragmatics*, pages 288-314. Blackwell, Malden, MA.

- [13] ILARI, Rodolfo. (1981). Os reflexivos e a interpretação dos prossintagmas. *Cadernos de Estudos Lingüísticos*, 2:56-156.
- [14] JACKENDO, Ray. (1972). *Semantic Interpretation in Generative Grammar*. The MIT Press, Cambridge.
- [15] JÄGER, Gerhard. (1997). Anaphora and ellipsis in type-logical grammar. In P. Dekker, M. Stokhof, and Y. Venema, editors, *Proceedings of the 11th Amsterdam Colloquium*, pages 175-180, Amsterdam. University of Amsterdam. Este texto pode ser encontrado no site <http://www.ling.uni-potsdam.de/~jaeger/paperac97.pdf>, acessado no dia 26/01/2004.
- [16] _____. (2001). Anaphora and quantification in categorial grammar. In M. Moortgat, editor, *Logical Aspects of Computational Linguistics*, pages 70-90. Springer, Berlin. Este texto pode ser encontrado no site <http://www.ling.uni-potsdam.de/~jaeger/lacl.pdf>, acessado no dia 26/01/2004.
- [17] _____. (2001). Indefinites and sluicing: A type logical approach. In R. Rooy and M. Stokhof, editors, *Proceedings of the 13th Amsterdam Colloquium*, pages 114-119, Amsterdam. University of Amsterdam. Este texto pode ser encontrado no site <http://www.ling.uni-potsdam.de/~jaeger/paperac01.pdf>, acessado no dia 26/01/2004.
- [18] _____. (2005). *Anaphora and Type Logical Grammar*. Springer, Dordrecht.
- [19] LAMBEK, Joachim. (1958). The mathematics of sentence structure. *American Mathematical Monthly*, 65:154-169.
- [20] LEWIS, David. General semantics. In Davidson and Harman [6], pages 169-218.
- [21] MORRILL, Glyn. (1994). *Type Logical Grammar*. Kluwer, Dordrecht.
- [22] PARTEE, Barbara Hall. editor. (1976). *Montague Grammar*. Academic Press, New York.
- [23] PETERS, S. and RITCHIE, R. (1969). A note on the universal base hypothesis. *Journal of Linguistics*, 5:150-152.
- [24] _____. (1971). On restricting the base component of transformational grammars. *Information and Control*, 18:483-501.
- [25] _____. (1973). On the generative power of transformational grammars. *Information Sciences*, 6:49-83.
- [26] SCHMERLING, Susan F. (1983). Two theories of syntactic categories. *Linguistics and Philosophy*, 6:393-421.
- [27] SZABOLCSI, Anna. (1989). Bound variables in syntax (are there any?). In Renate Bartsch, Johan van Benthem, and Peter van Emde Boas, editors, *Semantics and Contextual Expressions*, pages 295-318. Foris, Dordrecht.