

ESCOPO E “VARIÁVEIS LIGADAS TÍPICAS” DO PORTUGUÊS BRASILEIRO

Sergio Menuzzi*

Introdução

Meu objetivo com o presente trabalho é fazer uma discussão preliminar da seguinte observação factual: embora ambas as formas *seu* e *dele* possam se referir a um antecedente como *Quase todo rapaz*, como em (1) abaixo, em contextos como (2), somente *dele* pode – *seu* em (2a), sendo obrigatória ou preferencialmente interpretado como *a mãe de quase todo rapaz*, e não como *quase todo rapaz*:¹

- (1) a. [_{SN} *Quase todo rapaz*] se preocupa com *sua* namorada.
- b. [_{SN} *Quase todo rapaz*] se preocupa com a namorada *dele*.
- (2) a. ?? [_{SN} A mãe de [*quase todo rapaz*]] se preocupa com *sua* namorada.
- b. [_{SN} A mãe de [*quase todo rapaz*]] se preocupa com a namorada *dele*.

* Pontifícia Universidade Católica – Rio Grande do Sul. menuzzi@puers.br

¹ Daqui por diante, os SNs sublinhados em uma frase estão em relação de dependência anafórica, isto é, de interpretação referencial.

Com respeito ao problema com a frase em (2a), é importante deixar claro, de saída, que nada tem a ver com o fato de *seu* poder ser ambíguo – enquanto que *dele* desfaria a ambigüidade, como frequentemente observa a gramática tradicional. Se a possível ambigüidade de *seu* fosse o problema, uma frase como (2c) abaixo também deveria ser julgada “ruim” ou “marginal”, já que *seu* pode se referir tanto a *quase todo rapaz* quanto a *uma namorada*; entretanto, (2c) é uma frase perfeitamente boa em qualquer uma das duas interpretações:

- (2) c. *Quase todo rapaz quer uma namorada que enfrente sua mãe.*

O contraste entre (1a) e (2a) – e entre as possibilidades de interpretação de *seu* e *ele* – é, na verdade, um caso particular de uma oposição mais geral. Recentemente (em Menuzzi, 2003), eu apresentei uma generalização que parece caracterizar o uso de formas pronominais para expressar o que se costuma chamar de “interpretação de variável ligada” (noção que explicarei na segunda seção):

- (3) *Economia Morfossintática das Variáveis Ligadas.*
Se uma língua dispõe de formas pronominais alternativas para uma determinada posição sintática, então a forma favorecida (ou exigida) para interpretação de variável ligada naquela posição é a forma “mais econômica morfossintaticamente”.

Ilustra esta generalização o fato de que, embora o pronome *ele* possa ser usado, em certos contextos, como uma alternativa ao pronome possessivo *seu* (cf. o livro *dele/seu livro*) e ao reflexivo *si* (cf. *João sempre carrega uma arma consigo/com ele*), as formas *seu* e *si* são preferenciais quando a interpretação desejada é a de “variável ligada”. E estas formas são “mais econômicas morfossintaticamente” do que *ele*: enquanto *ele* contém informação sobre a pessoa gramatical, o gênero e o número de seu antecedente, *seu* e *si* só contém informação acerca da pessoa (ver Menuzzi, 2003, seção 4, para discussão; aqui, ver quinta seção).

Dada a generalização em (3), pode-se então chamar de “variáveis ligadas típicas” aquelas formas pronominais que, em contextos de alternância com outros pronomes, favorecem a interpretação de variável ligada – em português (brasileiro, ao menos), *seu* e *si* são exemplos de “variáveis ligadas típicas”, por oposição a *ele*. A questão de saber por que a interpretação de variável ligada prefere formas pronominais econômicas é um mistério cuja solução desconheço. No presente trabalho, eu gostaria de mostrar que o contraste ilustrado em (1) e (2) acima adiciona à coleção mais um mistério relacionado com as “variáveis ligadas típicas”.

Como todo e qualquer semanticista praticante sabe, as noções de “escopo” e “variável ligada” estão intimamente ligadas, aliás por definição em lógica standard (ver Dowty et al., 1981; Chierchia, McConnell-Ginet, 1990, por exemplo; aqui, ver a quarta seção para a noção de *escopo*). Procurarei mostrar, entretanto, que, em português (brasileiro) ao menos, não há correspondência completa entre a sintaxe do escopo e a possibilidade de uso de uma “variável ligada típica”. Esta conclusão levanta algumas questões conceituais muito interessantes acerca das relações entre sintaxe e semântica, como tentarei indicar ao longo deste artigo.

A discussão que segue se organiza assim: a segunda, terceira e quarta seções apresentam de modo breve e informal as noções que julgo importantes para compreender a relevância do contraste entre (1) e (2) – as noções de “interpretação de variável ligada”, “c-comando” e “escopo”, respectivamente;² na quinta seção, mostro que o contraste ilustrado em (1)-(2) se generaliza para outros contextos e outras “variáveis ligadas típicas” do português (brasileiro), incluindo os chamados “sujeitos nulos” ou “ocultos”; na seção final, indico brevemente por que o contraste observado é um mistério.

A Interpretação de Pronomes: Co-referência *versus* Interpretação como Variável Ligada

Pronomes e outros SNs que possuem um “antecedente” podem, em princípio, ter dois tipos de interpretação. Podem ser *co-referenciais* ao

2 Estas três seções tentam apresentar, de modo acessível ao estudante que tenha tido um curso introdutório em sintaxe gerativa e tenha um conhecimento básico de lógica, resultados bem conhecidos na literatura sobre quantificação e interpretação de pronomes; são, portanto, dispensáveis para quem é familiarizado com o assunto.

antecedente, denotando o mesmo indivíduo que ele. Nesta interpretação, a frase (4a) abaixo teria uma “forma lógica” que pode ser abreviada como em (4b), em que o pronome *ele* é “traduzido” pela “constante individual” *João*:³

- (4) a. *João* [_{SV} disse que Maria gosta *dele*].
b. *João* (_{SV} λx (x disse que Maria gosta de João)).

Formas pronominais podem ainda ser interpretadas como *variáveis ligadas*, em cujo caso representam uma “posição aberta” em um predicado, que pode ser preenchida por qualquer indivíduo que esteja incluído na denotação do antecedente. Esta distinção pareceria não fazer sentido para a frase (4a), já que o antecedente, *João*, denota apenas um indivíduo. Ainda assim, (4a) possui uma tal interpretação, que pode ser abreviada por (4c) abaixo, em que a “posição aberta” correspondente a *ele* é “traduzida” pela variável “x”; esta é “ligada” pelo operador lâmbda “ λx ”, que é satisfeito, por sua vez, por *João*:

- (4) c. *João* (_{SV} λx (x disse que Maria gosta de x)).

Desde Sag (1976) e Williams (1977), é bem sabido que as duas interpretações acima referidas são facilmente identificadas em contextos de elipse do SV, como o demonstra a ambigüidade de uma frase como (5) abaixo. Qualquer falante de

3 A “forma lógica” de uma frase é uma representação informal que expressa numa linguagem lógica – o chamado cálculo de predicados de primeira ordem com quantificação – certos aspectos da interpretação semântica da frase, como os elementos referenciais, as relações de predicado-argumento e o escopo de operadores lógicos (incluindo a ligação de variáveis). Em sintaxe gerativa, “forma lógica” tem também um outro significado técnico – é um nível de representação gramatical, como a estrutura-de-superfície e a estrutura profunda. Ver, aqui, a quarta seção, May (1985) e Chierchia e McConnell-Ginet (1990) para discussão detalhada; no capítulo 7 de Chierchia e McConnell-Ginet, também se encontra uma boa apresentação dos “operadores lâmbda”, isto é, de operadores como o representado por “ λx ” em (4b). Para fins de leitura, pode-se traduzir “ λx ” pela expressão “... é um x tal que ...”.

português percebe que o SV elíptico de (5) (isto é, o SV2) pode ser interpretado como em (6a), em cujo caso temos um exemplo da chamada “interpretação com identidade estrita” do pronome *ele* (em inglês, *strict identity reading*): o “*ele* elíptico” tem exatamente a mesma identidade que o *ele* expresso, isto é, ambos denotam o mesmo indivíduo, João. Esta interpretação de (5) corresponde à interpretação co-referencial do pronome, cf. a “forma lógica” em (6b):

- (5) *João* [_{SV1} disse que Maria gosta *dele*], e Paulo também [_{SV2} ____].
- (6) a. *João* [_{SV1} disse que Maria gosta *dele*, *João*],
e Paulo também [_{SV2} disse que a Maria gosta *dele*, *João*].
b. João (_{SV1} λx (x disse que Maria gosta de João)),
e Paulo também (_{SV2} λx (x disse que Maria gosta de João)).

Por outro lado, a frase em (5) também pode ser interpretada como (7a) abaixo. Neste caso, temos o que a literatura chama de “interpretação com identidade ‘relaxada’” do pronome *ele* (em inglês, *sloppy identity reading*): nesta interpretação de (5), o “*ele* elíptico” é, num certo sentido, interpretado do mesmo modo que o “*ele* expresso” – ambos são “idênticos” aos sujeitos de suas respectivas orações; mas não há “identidade estrita”, isto é, o “*ele* elíptico” *não* denota o mesmo indivíduo que o “*ele* expresso”. É esta interpretação de (5) que é capturada pela “interpretação de variável ligada” do pronome *ele*, representada na “forma lógica” em (7b):⁴

- (7) a. *João* [_{SV1} disse que Maria gosta *dele*, *João*],
e *Paulo* também [_{SV2} disse que Maria gosta *dele*, *Paulo*].
b. João (_{SV1} λx (x disse que Maria gosta de x)),
e Paulo também (_{SV2} λx (x disse que Maria gosta de x)).

4 Müller e Negrão (1996) sustentam que o pronome *ele* não pode ser interpretado como uma variável ligada em português brasileiro. Justamente devido a fatos como a ambigüidade de (5), eu creio que esta conclusão é incorreta, pelo menos para alguns dialetos – incluindo o meu próprio. Veja também Müller (1997; 2003) e Menuzzi (1999, seção 4.2; 2003) para discussão mais detalhada.

Em resumo, pode-se dar conta da ambigüidade de uma frase como (5) – ela pode ser interpretada como (6a) ou (7a) – se se presume que: a) no caso de elipse do SV, deve-se interpretar um SV elíptico “copiando” a “forma lógica” de um SV que o preceda (como em (6b) ou (7b), que o SV2 é copiado de SV1); b) pronomes podem ser interpretados co-referencialmente ou como variáveis ligadas. Para nós, aqui, a observação fundamental é: pode-se dizer que um pronome é interpretado como uma “variável ligada” se, em um contexto de elipse de SV, ele puder ser interpretado por “identidade ‘relaxada’”, como na interpretação de (5) descrita em (7). Ou seja, toda a vez que tivermos um caso de *sloppy identity* em um contexto de elipse de SV, é porque temos um pronome sendo interpretado como uma variável ligada.

Uma observação adicional importante acerca da interpretação dos pronomes é a de que embora possam ser ambíguos entre uma interpretação co-referencial e uma “interpretação de variável ligada”, esta possibilidade só existe com certos antecedentes. Especificamente, é preciso que o SN antecedente denote um indivíduo – possa ser *referencial*, como *João* e *Paulo* nos exemplos acima. Entretanto, certos SNs, como *quase todo rapaz* em (8) abaixo, não são referenciais; antes, sua interpretação depende de uma “operação quantificacional” sobre um conjunto de indivíduos, razão pela qual são chamados de *quantificacionais* (ver Chierchia e McConnell-Ginet, especialmente capítulos 3 e 9):

- (8) *Quase todo rapaz* [_{SV} diz que Maria gosta dele].

Com antecedentes quantificacionais, que não denotam indivíduos – não são referenciais – é claro que um pronome só pode ser interpretado como uma variável ligada. De fato, não há uma interpretação para a frase em (9) abaixo que corresponda a (10a), isto é, a uma “leitura com identidade estrita” do pronome (ou seja, em que o pronome expresso e o elíptico denotariam o mesmo indivíduo); antes, a única interpretação possível para o pronome em (9) é aquela em que há “identidade ‘relaxada’”, como em (10b) – em que o pronome é uma “variável ligada”, cf. a “forma lógica” em (10c) (‘#’ em (10a) significa que esta frase não descreve um significado possível de (9)):

- (9) *Quase todo rapaz* [_{SV} diz que Maria gosta dele]; mas o Paulo não [_{SV} ____].

- (10) a. # *Quase todo rapaz* diz que Maria gosta *dele mesmo*, do *próprio rapaz*, mas o Paulo não diz que a Maria gosta *daquele mesmo rapaz*.
 b. *Quase todo rapaz* diz que Maria gosta *dele mesmo*, o *próprio rapaz*, mas o Paulo não diz que a Maria gosta *dele mesmo*, o *próprio Paulo*.
 c. Quase todo rapaz (_{SV} λx (x diz que Maria gosta de x)), mas o Paulo não (_{SV} λx (x disse que Maria gosta de x)).

Assim, se queremos limitar nossas observações aos casos de “interpretação de variável ligada” de um pronome, basta verificarmos os casos em que o antecedente é quantificacional – nestes casos, por definição, não haverá uma “interpretação co-referencial” disponível.

Variáveis Ligadas e C-Comando

Uma outra descoberta fundamental feita pela lingüística contemporânea acerca da “interpretação de variável ligada” dos pronomes é a de que ela depende de uma relação estrutural específica; a interpretação co-referencial dos pronomes, por outro lado, não depende da mesma relação (ver Reinhart, 1983, capítulo 5, para discussão e precedentes).⁵ Pode-se ilustrar esta observação por meio do contraste entre (11) e (12) abaixo:

- (11) a. *Quase todo rapaz do bairro* diz que Maria gosta *dele*.
 b. * [_{SN} *Aquela menina que admira quase todo rapaz do bairro*] diz que Maria gosta *dele*.

5 Uma “relação estrutural” é uma relação definida com base na estrutura de constituintes de frase, isto é, nos agrupamentos de unidades sintáticas e sua organização hierárquica. Em sintaxe gerativa, usa-se freqüentemente uma *árvore* para representar a estrutura de constituintes da frase: uma representação gráfica em que cada agrupamento é representado por um *nó rotulado* com o tipo sintático do constituinte (SN, SV, SP etc.) e ao qual as unidades de que é composto – também nós rotulados – são ligadas por meio de *ramos* ou *galhos*. Veja a discussão de (16a,b) abaixo para ilustração; para apresentação detalhada, veja Radford (1981; 1988).

- (12) a. *O João diz que Maria gosta dele.*
b. [_{SN} Aquela menina que admira *o João*] diz que Maria gosta *dele*.

(11a) mostra, como (8) acima, que um SN quantificacional pode ser um antecedente de um pronome; (11b) mostra, por outro lado, que esta possibilidade parece não existir se o SN quantificacional estiver contido dentro de algum outro constituinte (um outro SN, em (11b)). Esta restrição não se aplica a antecedentes referenciais, como se vê em (12b): embora *o João* ocorra exatamente no mesmo contexto que *quase todo rapaz do bairro* em (11b), ainda assim ele serve perfeitamente de antecedente para *ele*. A impossibilidade de *quase todo rapaz do bairro* ser o antecedente de *ele* em (11b) se deve justamente ao fato de este antecedente quantificacional requer que o pronome seja interpretado como uma variável ligada; mas esta interpretação só é possível se o SN quantificacional e o pronome estiverem numa relação estrutural específica, o que se dá em (11a), mas não em (11b).

Que a restrição vista em (11b) diz respeito à “interpretação de variável ligada” do pronome, e não a relações envolvendo um tipo específico de antecedente – isto é, SNs quantificacionais –, é revelado pelo contraste interpretativo entre os exemplos em (13):

- (13) a. *O João* [_{SV} diz que Maria gosta *dele*]; *o Paulo* também [_{SV} ___].
b. [_{SN} Aquela menina que admira *o João*] [_{SV} diz que Maria gosta *dele*]; *o Paulo* também [_{SV} ___].

Como vimos antes, a elipse do SV com um antecedente referencial como *João* resulta numa ambigüidade: o pronome pode ser interpretado tanto co-referencialmente quanto como uma variável ligada. De fato, (13a) (do mesmo modo que (5) acima) é ambígua: o “pronome elíptico” pode ser interpretado tanto como João (ou seja, Paulo também diz que Maria gosta de *João*: interpretação co-referencial do pronome) quanto como Paulo (ou seja, Paulo diz que Maria gosta *dele mesmo*, *Paulo*: interpretação do pronome como variável ligada). Entretanto, a frase (13b) não possui a mesma ambigüidade: o “pronome elíptico” só pode ser interpretado *como João, e não como Paulo*. Isto é, (13b) só pode ser interpretada como (14a) abaixo, e não como (14b):

- (14) a. A menina que admira *o João* diz que Maria gosta *dele, João*; Paulo também diz que a Maria gosta *dele, João*.
 b. # A menina que admira *o João* diz que Maria gosta *dele, João*; Paulo diz que a Maria gosta *dele, Paulo*.

A ausência de uma interpretação de (13b) que corresponda à (14b) mostra que o pronome não pode ser interpretado como uma “variável ligada” em (13b), embora seu antecedente seja um SN referencial. Ou seja, exatamente no mesmo contexto estrutural em que um SN quantificacional não pode ligar um pronome (isto é, (11b)), este pronome também não pode ser interpretado como uma variável ligada se o antecedente for um SN referencial.

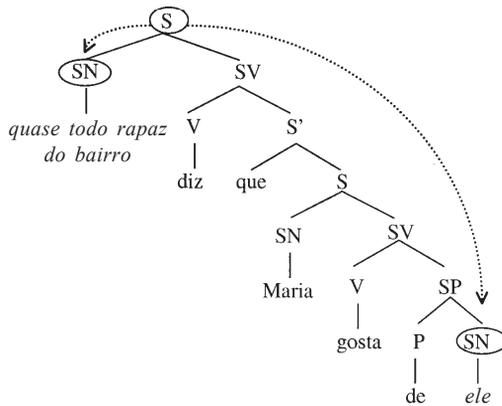
Assim, o problema com (11b) acima não resulta do fato de o antecedente ser um SN quantificacional, mas da impossibilidade de se poder estabelecer a relação estrutural necessária à “interpretação de variável ligada” do pronome, restrição que formulo em (15a) abaixo. Desde Reinhart (1976) (mas veja especialmente Reinhart, 1983), acredita-se que a relação estrutural relevante seja a de *c-comando*, cuja definição (adaptada) aparece em (15b):

- (15) a. *Restrição Estrutural à Interpretação de Variável Ligada:*
 Um pronome Y só pode ser interpretado como uma variável ligada a um antecedente X se X *c-comanda* Y.
 b. Um nó X *c-comanda* outro nó Y se e somente se o primeiro constituinte máximo (S, SN, SV, SP etc.) que domina X também domina Y.⁶

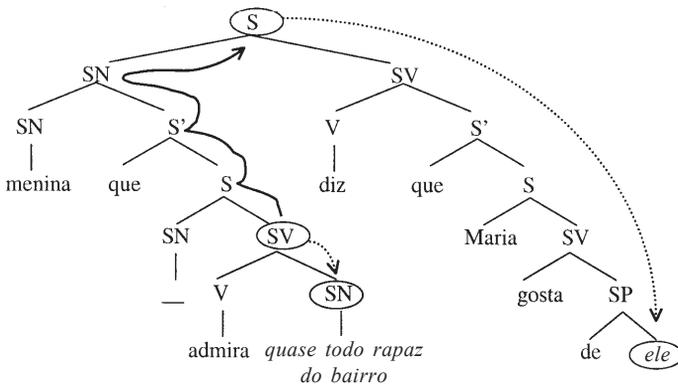
6 Definição de acordo com Reinhart (1983, p. 18 et seq.). Quanto a “dominar”, diz-se que um nó X domina outro nó Y em uma árvore se existe um caminho exclusivamente descendente que vai de X a Y. Por exemplo, o nó S circundado em (16b) abaixo domina todos os outros nós porque sempre há um caminho que desce dele até os demais nós em (16a), o que vale para o pronome *ele* (cf. o caminho representado pela seta pontilhada). Já o nó SV circundado domina o SN *quase todo rapaz do bairro* (cf. seta pontilhada), mas não domina o pronome *ele*: o caminho que leva do SV circundado até o pronome precisa inicialmente subir até o nó S circundado (cf. seta contínua) para só então descer de S até o pronome – o caminho não seria completamente descendente.

Comparemos agora as estruturas das frases (11a) e (11b), para vermos como em (11a) há a relação de c-comando entre *qualquer rapaz do bairro* e o pronome – permitindo a interpretação de variável ligada do último –, enquanto que em (11b) não há esta relação: (16a) abaixo é a estrutura de (11a), e (16b), a estrutura de (11b):

(16) a.



(16) b.



Em (16a), o S circundado é o primeiro constituinte máximo que domina o SN *quase todo rapaz do bairro* (o “X” da definição em (15b)), como a seta pontilhada indica; este S também domina o pronome *ele* (o “Y” da definição em (15b)), como a outra seta pontilhada indica. Assim, o primeiro constituinte máximo que domina *quase todo rapaz do bairro* também domina *ele*; logo, de acordo com a definição em (15b), *quase todo rapaz do bairro* c-comanda *ele*, que pode ser, então, interpretado como uma variável ligada (por *quase todo rapaz do bairro*, isto é, este SN pode ser interpretado como o antecedente de *ele* em (16a)).

Em (16b), por outro lado, o primeiro constituinte máximo que domina o SN *quase todo rapaz do bairro* é o SV circundado, *que não domina o pronome ele* (como a seta contínua mostra, seria necessário subir na árvore até o S circundado para só então descer até pronome). Se o primeiro constituinte máximo que domina o SN *quase todo rapaz do bairro* não domina o pronome *ele*, então, de acordo com a definição em (15b), *quase todo rapaz do bairro* não c-comanda *ele*. Logo, este pronome *não* pode ser interpretado como uma variável ligada (por *quase todo rapaz do bairro*, isto é, este SN *não* pode ser interpretado como o antecedente de *ele* em (16a)).

Em resumo, a interpretação de variável ligada de um pronome está sujeita a uma restrição estrutural, a de que o antecedente c-comande o pronome (cf. (15a)). Esta observação é importante porque em lógica/semântica, “ser uma variável ligada” é ser uma posição aberta em um predicado no “escopo” de um operador, que fornece à variável um “valor semântico” (ver Chierchia e McConnell-Ginet, capítulo 3). Tipicamente, “operadores” são SNs quantificacionais (embora outros operadores também possam ligar uma variável; por exemplo, em (4) acima, o único operador disponível é o “operador lambda”), e o “escopo” de um operador é justamente o “domínio estrutural” sobre o qual atua. As perguntas que surgem são: Qual a relação entre a noção lógica de *escopo* de um operador e a noção sintática de *c-comando*? É possível reduzir a restrição estrutural sobre a interpretação de variável ligada de um pronome (15a) ao escopo estrutural dos operadores? É claro que, do ponto de vista teórico, a resposta desejável seria *sim*: afinal, se “ser uma variável ligada” é “estar no escopo de um operador”, as restrições sobre ambas as relações deveriam ser as mesmas – o que mostraria que há correspondência completa entre sintaxe e semântica neste aspecto de nosso conhecimento lingüístico. Procurarei mostrar mais adiante (na quinta seção) que esta correspondência se quebra, entretanto, para uma classe de elementos pronominais. Na próxima seção, discuto rapidamente sob que condições estruturais um operador – ou melhor, uma classe de “operadores típicos”, os SNs quantificacionais – pode tomar escopo.

Escopo e C-Comando

A análise semântica mais comum dos “operadores quantificacionais” – que segue a teoria da quantificação da lógica standard – toma a oração como o “domínio de operação” semântica de um SN quantificacional, ou seu *escopo* (ver Chierchia, McConnell-Ginet, 1990, capítulo 3). Por exemplo, a frase em (10a) teria como “forma lógica” uma representação como (10b); nesta representação, as expressões lógicas correspondentes aos SNs *quase todo rapaz do bairro* e *uma menina* são os operadores quantificacionais “Para quase todo x, x um rapaz” e “Existe um y, y uma menina, tal que”; estes operadores se combinam com uma “fórmula aberta”, a oração “x gosta de y”, e resultam numa “proposição completa” – isto é, uma frase cujas condições de verdade são completamente definidas e podem ser verificadas:

- (17) a. [_S *Quase todo rapaz do bairro* gosta de *uma menina*].
b. (Para quase todo x, x um rapaz do bairro) (existe um y, y uma menina, tal que) [_S x gosta de y].

Diz-se, assim, que em (17b) o escopo dos operadores “Para quase todo x, x um rapaz” (correspondente a *quase todo rapaz* em (17a)) e “Existe um y, y uma menina, tal que” (correspondente a *uma menina*) é a fórmula “x gosta de y” – correspondente, portanto, à oração cujo núcleo é *gostar* em (17a). Estando dentro do escopo destes operadores, as variáveis “x” e “y” são a eles “ligadas” – sua interpretação é dada pelo operador com o qual estão identificadas.

Observe agora que o escopo, ou “domínio”, de um operador *não* corresponde necessariamente à oração em que ocorre “*na superfície*”: isto é o que acontece em (17a), mas não necessariamente em (18a) abaixo, que é ambígua entre as interpretações representadas por (17b) e (17c):

- (17) a. [_S *Muitas meninas* querem [_S namorar *quase todo rapaz do bairro*]].
b. (Existe muito x, x uma menina, tal que) [_S x quer [_S (para quase todo y, y um rapaz do bairro) [_S x namorar y]]].

c. (Para quase todo y, y um rapaz do bairro) (existe muito x, x uma menina, tal que) [_s x quer [_s x namorar y]].

Na interpretação representada em (17b), a frase (17a) significa, aproximadamente, que cada uma das muitas meninas quer namorar quase todos os meninos do bairro. Na interpretação representada em (17c), a frase (17a) significa, aproximadamente, que quase todos os meninos do bairro possuem um fã clube – cada um deles é admirado por muitas meninas.

Agora, a observação importante é a seguinte: na “superfície” da frase (17a), *muitas meninas* está dentro da oração cujo núcleo é *querer*, enquanto que *quase todo rapaz do bairro* está dentro da oração cujo núcleo é *namorar*; se o escopo de um quantificador correspondesse necessariamente a oração em que ocorre na superfície, o escopo de *muitas meninas* deveria ser a oração cujo núcleo é *querer*, enquanto que o escopo de *quase todo rapaz do bairro* deveria ser a oração cujo núcleo é *namorar*. De fato, isto corresponde a *uma* das interpretações possíveis de (17a), a que corresponde a (17b). Mas há pelo menos uma *outra* interpretação para a frase em que o escopo de *quase todo rapaz do bairro* não corresponde à oração em que ocorre na superfície: a interpretação representada por (17c) (em que *quase todo rapaz do bairro* toma como escopo a oração principal, cujo núcleo é *querer*).

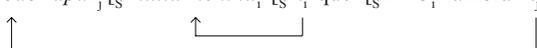
Existem diferentes teorias sobre o modo como os quantificadores tomam escopo nas línguas naturais (ver Reinhart, 1997 para uma síntese e discussão). Em sintaxe gerativa, o tratamento mais comum é postular-se que SNs quantificacionais sofrem “*movimento abstrato*” (*Quantifier Raising*), dando origem a uma *representação gramatical diferente da que corresponde à estrutura-de-superfície* da frase. Nesta outra representação, o escopo de operadores é representado “transparentemente” – o que se vê é o que é –, resultando numa árvore muito semelhante ao que seria a “forma lógica” da frase; por isso (entre outras razões), chama-se à representação gramatical resultante do “movimento abstrato” de quantificadores justamente de *forma lógica* da frase (ver May, 1985 e referências lá citadas). De acordo com esta

teoria, teríamos as seguintes formas lógicas correspondentes às interpretações em (17b) e (17c) da frase (17a):⁷

- (18) a. forma lógica correspondente à (17b):

$$[_s \textit{Muita menina}_i [_s t_i \textit{ quer } [_s \textit{ quase todo rapaz}_j [_s \textit{ PRO}_i \textit{ namorar } t_j]]]]]$$

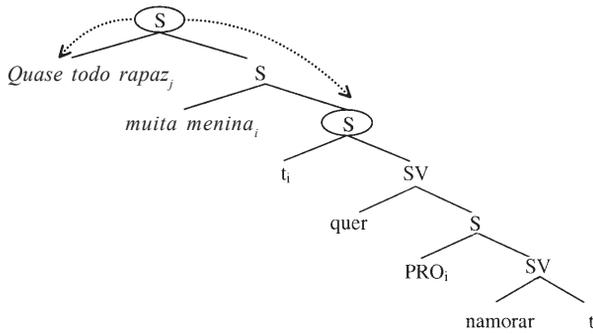
- b. forma lógica correspondente à (17c):

$$[_s \textit{ Quase todo rapaz}_j [_s \textit{ muita menina}_i [_s t_i \textit{ quer } [_s \textit{ PRO}_i \textit{ namorar } t_j]]]]]$$


O importante para nossa discussão aqui é que, do ponto de vista desta análise da sintaxe do escopo, é possível definir “escopo” em termos de *c-comando*, isto é, da mesma relação estrutural necessária à interpretação de variável ligada de um pronome. Isto é possível porque, em forma lógica, depois do movimento abstrato acima referido, um operador sempre *c-comanda* seu “domínio de operação”, como se vê, por exemplo, na árvore em (19a) abaixo, correspondente à (18b):

7 Nas representações em (18) abaixo, usa-se “t” para representar o elemento vazio, a “variável”, que ocupa o lugar ocupado em estrutura-de-superfície pelo quantificador deslocado (“t” é abreviatura do inglês *trace*, “vestígio”). Esta “variável” é ligada pelo quantificador, o que é representado pela co-indexação de ambos (os índices são as letras subscritas). Observe-se também que a regra de *Quantifier Raising* é tradicionalmente concebida como um “movimento de adjunção” a S, ou seja, o quantificador é deslocado para o começo da S e forma com ela um novo constituinte de natureza oracional, também rotulado com “S” (cf. May, 1985). Finalmente, PRO é um outro elemento vazio, de natureza pronominal; é o sujeito de orações infinitivas (cf. Chomsky, 1981). Na frase em questão, PRO também é interpretado como uma variável ligada.

- (19) a. *forma lógica* (18b) (correspondente à (17c)):

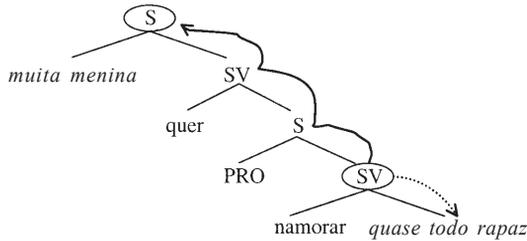


Em (19a), o primeiro constituinte máximo que domina *quase todo rapaz do bairro* é um S que também domina o S correspondente à fórmula “[x quer [x namorar y]]”, isto é, à oração que *quase todo rapaz do bairro* toma como escopo em fórmula lógica; portanto, pode-se dizer que, em forma lógica, *quase todo rapaz do bairro* c-comanda a S que toma como escopo. O que vale para este caso vale em geral, e por isso, em sintaxe gerativa, define-se “escopo” em termos de c-comando, como em (19b):

- (19) b. Y é o *domínio de escopo* de X se e somente se X c-comanda Y em forma lógica.

É importantíssimo observar que, de acordo com (19b), a definição de escopo baseada na noção de c-comando tem de ser limitada à forma lógica da frase – isto é, à representação resultante do movimento abstrato dos quantificadores. Isto se deve ao fato de que freqüentemente os quantificadores *não* c-comandam a oração que tomam como escopo *na estrutura-de-superfície* da frase. É o que acontece com a estrutura-de-superfície correspondente a (19a), cuja árvore é (19c) abaixo:

(19) c. *estrutura-de-superfície* de (17a):



Como vimos antes, *quase todo rapaz do bairro* pode tomar como escopo a oração principal de (17a) (cf. a interpretação representada em (17c)-(19a)). Entretanto, na estrutura-de-superfície da frase, representada em (19c) acima, este operador não c-comanda a oração principal – a S circundada: o primeiro constituinte máximo que domina o operador é o SV circundado, que *não* domina a S principal. Assim, não é possível definir a noção de “escopo de um quantificador” em termos de c-comando se olharmos apenas a estrutura-de-superfície da frase; antes, *escopo deve ser definido como em (19b), em termos de c-comando em forma lógica – depois do movimento abstrato dos quantificadores.*

Qual a importância da observação de que a noção de escopo pode ser definida em termos de c-comando? Bom, a importância desta observação deveria estar mais ou menos clara dado o que discutimos até aqui: a “interpretação de variável ligada” está sujeita a uma restrição estrutural que exige que o antecedente do pronome o c-comande; mas “ser interpretado como uma variável ligada” é, por definição – em lógica standard e em semântica das condições de verdade – estar no escopo de um operador; ora, se a noção de escopo se define por meio da relação estrutural de c-comando, explica-se de modo óbvio por que a “interpretação de variável ligada” também depende da relação de c-comando. De quebra, esta conclusão revelaria que, pelo menos no que diz respeito à relação entre quantificação e interpretação de pronomes, haveria isomorfia entre sintaxe e semântica: as condições estruturais que atuam na interpretação de pronomes seriam redutíveis às operações necessárias para a semântica da quantificação. Como veremos na próxima seção, entretanto, alguns fatos do português (brasileiro) – contrastes como o que há entre (1) e (2) acima – revelam que esta conclusão pode ser precipitada.

Antes de retornarmos a esta discussão, é preciso fazer uma observação final, muito importante, sobre o escopo de quantificadores, ou melhor, sobre a operação sintática com a qual os quantificadores adquirem escopo em forma lógica – o “movimento abstrato” ou *Quantifier Raising*. Sendo uma operação sintática de movimento, *Quantifier Raising* deveria estar sujeito a certas limitações, pois sabidamente operações que deslocam constituintes não podem operar em qualquer contexto: há as chamadas “ilhas” para movimento, isto é, contextos em que uma operação de movimento não pode se aplicar (cf. Ross, 1967; ver Haegeman, 1994, capítulo 7, para uma apresentação atualizada). De fato, em muitas versões da teoria do escopo baseada em *Quantifier Raising*, tenta-se reduzir as restrições que se observa no escopo de quantificadores às restrições impostas à operação de movimento (ver Reinhart, 1997, para uma síntese e discussão crítica).

Nesta perspectiva, o fato de que o SN quantificacional *quase todo rapaz do bairro* não pode ser o antecedente do pronome em (11b) acima – repetida como (20a) abaixo – não seria mais o resultado da restrição de c-comando sobre a interpretação de variável ligada (cf. (15a) acima); antes, o que é preciso explicar agora é por que *Quantifier Raising* não pode se aplicar a (20a), derivando a forma lógica em (20b):

- (20) a. * [_S [_{SN} A menina que admira *quase todo rapaz do bairro*] diz que Maria gosta dele].
 b. * *Quase todo rapaz do bairro*_i [_S [_{SN} a menina que admira t_i] diz que Maria gosta dele].

Note que, se a forma lógica (20b) fosse possível, *quase todo rapaz do bairro* c-comandaria *ele* em forma lógica – e *ele* deveria poder ser interpretado como ligado a *quase todo rapaz do bairro*. Presumivelmente, o que exclui (20b) é alguma restrição sobre *Quantifier Raising* – neste caso, não problemática, já que (20b) envolve movimento em uma “ilha” muito conhecida, a ilha do NP complexo – que se aplica a todos os tipos de movimento, inclusive, aparentemente, a *Quantifier Raising* (ver Reinhart, 1997; sobre o caráter particular da “Ilha do NP Complexo”, ver Menuzzi, 2000; Eick, 2002).

O reverso do raciocínio acima descrito também é válido: naqueles contextos em que um SN quantificacional *pode* adquirir um determinado escopo, somos obrigados a concluir, nesta teoria, que *Quantifier Raising* pode se aplicar.

Considere, por exemplo, uma frase como (21a) abaixo: nela, o quantificador *quase todo rapaz* é interpretado como tendo escopo sobre toda a oração; inclusive, o pronome *ele* pode ser perfeitamente interpretado como uma variável ligada a *quase todo rapaz*, sendo a frase como um todo interpretada como em (21b); ou seja, na análise que descrevemos antes, é preciso postular uma forma lógica como (21c):

- (21) a. [_S [_{SN} A mãe de *quase todo rapaz*] confiaria cegamente *nele*]
 b. (Para quase todo x, x um rapaz) [_S [_{SN} a mãe de x] confiaria cegamente em x]
 c. *Quase todo rapaz*_i [_S [_{SN} a mãe de t_i] confiaria cegamente *nele*].
- 

Assim, com base nas possibilidades de escopo de *quase todo rapaz* em (21a) (cf. (21b)), concluímos que *Quantifier Raising* pode se aplicar a um operador que está dentro do SN sujeito. Este tipo de observação é importante porque exclui a possibilidade de que outros problemas com a interpretação de variável ligada neste contexto tenham a ver com restrições à operação de *Quantifier Raising*. Por exemplo, como veremos a seguir, o pronome reflexivo *si* é uma forma tipicamente interpretada como variável ligada; não pode, entretanto, substituir o pronome *ele* em (21a), como vemos em (22):

- (22) * [_S [_{SN} A mãe de *quase todo rapaz*] confiaria cegamente em *si*].

Obviamente, *quase todo rapaz* não pode ser um antecedente de *si* em (22); o que a discussão de (21) indica é que o problema com (22) não pode ser atribuído a alguma restrição a *Quantifier Raising*. Voltarei a este ponto, e também às propriedades de *si*, na próxima seção.

Em resumo, vimos nesta seção que a noção de escopo de um quantificador pode ser definida em termos da noção de c-comando se tiver como base a forma lógica de uma frase. Apontei brevemente também para a importância desta conclusão no que diz respeito a tema deste artigo: de um ponto de vista mais específico, ela oferece uma explicação natural para a restrição estrutural que encontramos sobre a interpretação de pronomes como variáveis ligadas (15a);

de um ponto de vista mais geral, esta explicação permitiria sustentar que, ao menos no que diz respeito à relação entre quantificação e interpretação de pronomes, haveria isomorfia completa entre sintaxe e semântica – conclusão que, por sua vez, seria de suma importância para o estudo das relações entre sintaxe e semântica (ponto que comentarei ligeiramente na seção final). Assim, diante da perspectiva do que vim apresentando, a pergunta fundamental é: será que, no que concerne a línguas naturais, as restrições que atuam na determinação do *escopo de operadores* de fato coincidem as restrições sobre *variáveis ligadas*? Acredito que o comportamento do que chamei de “variáveis ligadas típicas” indica que a resposta para esta pergunta pode ser negativa.

“Variáveis Ligadas Típicas”: C-Comando em Estrutura-de-Superfície *versus* Escopo

Como mencionei antes, em meu artigo de 2003 argumento que uma certa classe de elementos pronominais está associada de modo peculiar à interpretação de variável ligada. Os elementos desta classe alternam com outras formas pronominais em determinados contextos sintáticos, sendo mais “econômicos” morfossintaticamente que suas alternativas. Em português brasileiro, há pelo menos três destas formas: o reflexivo *si*, o possessivo *seu* e aquilo que a gramática tradicional chama de *sujeito oculto* e que é conhecido na literatura gerativista pelo termo *sujeito nulo*. Vejamos o comportamento de cada uma destas formas no que diz respeito à interpretação de variável ligada.

O reflexivo *si*

A forma *si* alterna com o pronome *ele* em contextos como (23) abaixo – ambas podem, em princípio, ser usadas em português (brasileiro); e *si* é mais “econômico” morfossintaticamente que *ele*: *ele* informa a pessoa gramatical, o gênero e o número de seu antecedente; *si* informa apenas a pessoa gramatical:

- (23) *O João só confia {em si/nele}.*

Quando uma tal alternância é possível, a forma pronominal “mais econômica” tem necessariamente, ou favorece fortemente, uma interpretação de variável ligada, enquanto que a outra forma tende a favorecer ou, ao menos, permitir a interpretação co-referencial (cf. a generalização descrita em (3) acima). No caso de *si*, por exemplo, a interpretação de variável ligada é obrigatória, como vemos pela ausência de ambigüidade numa frase como (24a) abaixo, que só pode ser interpretada como (24b) (que corresponde a interpretação de *si* como uma variável ligada) e não como (24c) (em que *si* teria uma interpretação co-referencial: ver a segunda seção):

- (24) a. *O Paulo só confia em si, e o João também.*
b. *O Paulo só confia nele mesmo, Paulo, e o João só confia nele mesmo, João.*
c. # *O Paulo só confia nele mesmo, Paulo, e o João também só confia em Paulo.*

Já a forma alternativa a *si*, o pronome *ele*, no mesmo contexto, pode ser interpretado (ao menos para mim) tanto co-referencialmente quanto como uma variável ligada – isto é, a frase (25a) abaixo (idêntica a (24a), exceto por ter *ele* no lugar de *si*) pode ser interpretada tanto como (25b) (= (24b)) quanto como (25c) (= (24c)):

- (24) a. *O Paulo só confia nele (nestas situações), e o João também.*
b. *O Paulo só confia nele mesmo, Paulo, e o João só confia nele mesmo, João.*
c. *O Paulo só confia nele mesmo, Paulo, e o João também só confia em Paulo.*

Ou seja, a forma *si* está relacionada de modo particular à “interpretação de variável ligada”: *si* requer esta interpretação, enquanto que a forma

alternativa *ele* pode ou não tê-la. O reflexivo *si* é, portanto, o tipo de forma a que chamei de “variável ligada típica” – é uma forma econômica morfossintaticamente que, quando usada, exige ser interpretada como uma variável ligada (cf. a generalização em (3)). O que há de interessante acerca destas formas – tendo em vista a discussão do presente artigo – é que elas permitem testar a hipótese de que, no que diz respeito a relação entre quantificação e interpretação de pronomes como variáveis, haveria isomorfia completa das condições estruturais.

De acordo com esta hipótese, esperamos que, se a sintaxe de uma língua possui expressões de alguma forma especializadas para a interpretação como variável ligada, então tais expressões deveriam ser usadas exatamente sob as mesmas condições em que há escopo de um operador – afinal, como vimos antes, “ser uma variável ligada” é, por definição, “estar no escopo de um operador”. Pois as “variáveis ligadas típicas” são justamente o tipo de expressão que procuramos: são especializadas para a interpretação de variável ligada. Como veremos a seguir, entretanto, *si* e as demais variáveis ligadas típicas do português (brasileiro) parecem ir contra a hipótese da isomorfia.

Considere o caso da forma *si*. Vimos na seção anterior que o fato de que *quase todo rapaz* pode ser o antecedente do pronome *ele* em (25a) abaixo indica duas coisas: que *ele* é uma variável ligada (única interpretação possível com um antecedente quantificacional, cf. segunda seção) e que, além disso, está no escopo de *quase todo rapaz* – há c-comando em forma lógica (veja discussão de (21a,b,c) acima). De fato, isto é confirmado por contextos de elipse do SV, como em (25b) abaixo, em que se pode ter, perfeitamente, uma “interpretação com identidade ‘relaxada’” para o pronome – isto é, (25b) pode perfeitamente ser interpretada como (25c) (que corresponde à interpretação em que o pronome em (25b) é uma variável ligada: ver segunda seção):

- (25) a. [_{SN} A mãe de *quase todo rapaz*] confiaria cegamente *nele*.
 b. [_{SN} A mãe de *quase todo rapaz*] [_{SV} confiaria cegamente *nele*], mas a mãe do *João* não [_{SV} ____].
 c. A mãe de *quase todo rapaz* confiaria cegamente no *seu filho*, mas a mãe do *João* não confiaria *nele, João*.

Assim, (25a) é um contexto em que a forma interpretada como variável ligada – *ele*, neste caso – está no escopo de um operador – *quase todo rapaz*. Mas, como já havíamos ilustrado em (22) e repetimos em (26a) abaixo, neste mesmo contexto o escopo de *quase todo rapaz* não é suficiente para permitir o uso de uma “variável ligada típica”, *si*: esta forma exige, antes, c-comando em estrutura-de-superfície, como em (26b):

- (26) a. * [_{SN} A mãe de *quase todo rapaz*] confiaria cegamente em *si*.
b. [_{SN} *Quase todo rapaz*] confiaria cegamente em *si*.

Neste ponto, um leitor pouco familiarizado com a literatura contemporânea sobre pronomes reflexivos, mas conhecedor das gramáticas tradicionais, poderia pensar: “É claro: o antecedente de um pronome reflexivo deve, necessariamente, ser o sujeito da frase! É por isso que *quase todo rapaz* pode ser o antecedente de *si* em (26b), mas não em (26a)!” Mas esta explicação para o contraste em (26) teria vida curta: considere agora o contraste entre (27a) e (27b) abaixo:

- (27) a. Aquele psicanalista [_{SV} coloca *todo paciente* diante de *si*].
b. *Aquele psicanalista [_{SV} ridiculariza *todo paciente*] diante de *si*.

Parece-me claro que *todo paciente* pode ser o antecedente de *si* em (27a) (a frase pode significar “aquele psicanalista coloca todo paciente diante do próprio paciente – isto é, fã-lo enfrentar-se a si próprio”), mas não em (27b) (a frase *não* pode significar “aquele psicanalista ridiculariza todo paciente em frente ao próprio paciente”),⁸ pois este contraste elimina a explicação tradicional para o uso de *si*: o antecedente em (27a) não é o sujeito da frase, mas o *objeto* – e a frase é perfeitamente boa.

8 Em ambas, o antecedente de *si* pode ser *aquele psicanalista*, o que é compatível com o que estamos discutindo aqui: a posição de sujeito c-comanda *si* tanto em (27a) quanto em (27b).

Mas se o antecedente de *si* não precisa ser o sujeito da frase, por que não pode ser *quase todo rapaz* em (26a)? Pela mesma razão que *todo paciente* não pode ser o antecedente de *si* em (27b): embora *si* esteja no escopo de *todo paciente* em (27b), ele não é c-comandado por este SN em estrutura-de-superfície. Para demonstrar que *si* está no escopo de *todo paciente* em (27b), podemos utilizar os mesmos testes que utilizamos em (25)-(26). Troquemos a forma *si* pela forma alternativa *ele*: o fato do pronome *ele* poder ter como antecedente *todo paciente* em (28a) abaixo indica que é uma variável ligada no escopo de *todo paciente* ((28a) pode significar que o psicanalista ridiculariza todos os seus pacientes na frente dos próprios pacientes). E, como se vê em (28b), esta conclusão é confirmada pelo teste da elipse: (28b) pode ter a interpretação correspondente a (28c) – isto é, à interpretação em que o pronome em (28b) seria uma variável ligada.⁹

- (28) a. Aquele psicanalista [_{sv} ridiculariza *todo paciente*] diante *dele* (, sem qualquer constrangimento).
 b. O teu psicanalista ridiculariza *quase todo paciente* diante *dele* (, sem qualquer constrangimento); o da Maria, só *alguns*.
 c. O teu psicanalista ridiculariza *quase todo paciente* diante do *próprio paciente*; o psicanalista da Maria só ridiculariza *alguns pacientes* diante dos *próprios*.

Assim, a impossibilidade de que *todo paciente* seja o antecedente de *si* em (27b) não pode ser atribuída a ausência de escopo – de c-comando por *todo paciente* em forma lógica. Como eu disse antes, pode-se recorrer à explicação alternativa que sugeri também para o contraste em (26): *si*, uma “variável ligada típica”, exige não apenas estar no escopo de seu operador – isto é, ser c-comandado por ele –, mas isto deve acontecer em estrutura-de-superfície. Com efeito, é bem conhecido na literatura sintática o contraste entre o uso do sintagma locativo *diante de si* em (27a) e (27b) (ver Reinhart, 1983, capítulo 3, para síntese): há evidência de que o locativo em (27a), um complemento do verbo, está estruturalmente “mais próximo” do objeto verbal do que em (27b), onde ele é um

9 No caso de (28b), não teríamos uma elipse de SV, mas uma outra construção, chamada em inglês de *gapping*. Para uma discussão de vários casos de “apagamento” de constituintes, ver Sag (1976).

adjunto adverbial. Este fato levou Reinhart (1983, p. 60 et seq.) a propor que há c-comando de *si* por *todo paciente* em (27a), mas não em (27b).¹⁰

Neste ponto, um leitor informado acerca da literatura contemporânea sobre reflexivos poderia objetar: “É claro que *si* exige que seu antecedente o c-comande em estrutura-de-superfície! O reflexivo *si* é o que se chama, em teoria da ligação, de um ‘anafórico’: um tipo de SN que está sujeito ao ‘Princípio A’ desta teoria, isto é, que precisa ser ‘ligado’ – o que inclui ser c-comandado pelo antecedente – dentro de um domínio local” (cf. Chomsky, 1981; 1986; ver Menuzzi, 1999, capítulo 1, para síntese e discussão crítica). O que enfraquece esta explicação alternativa para a inaceitabilidade de (26a) e (27b), entretanto, é que ela *não* se aplica a outras “variáveis ligadas típicas”, que não se comportam como “anafóricos” – não, ao menos, no sentido da teoria da ligação padrão. É o que passo a demonstrar.

O possessivo *seu*

Considere, por exemplo, o pronome possessivo *seu*: como é bem sabido na literatura sobre o português brasileiro (ver Menuzzi, 1996; Müller, 1997, para referências e discussão), o possessivo alterna com a forma *dele* – que inclui o pronome *ele* – e parece favorecer a interpretação de variável ligada. Isto pode ser visto em sua distribuição em *corpus* de fala, que revela o padrão resumido no seguinte paradigma:

10 Um dos argumentos para distinguir estruturalmente as duas ocorrências do sintagma locativo em (28) foi apontado por Chomsky na obra que inaugurou a revolução gerativista, *Syntactic Structures* (Chomsky, 1957): complementos locativos, por estarem mais “intimamente ligados” ao verbo, resistem ao deslocamento de dentro do SV, como vemos pela estranheza de (ib) abaixo:

- (i) a. Aquele psicanalista colocava os pacientes diante de um espelho.
- b. ??Diante de um espelho, aquele psicanalista colocava os pacientes.

Adjuntos locativos, por outro lado, por não estarem “ligados intimamente” ao verbo, seriam modificadores da oração, o que explicaria sua mobilidade sintática e a ausência de qualquer problema para a sua anteposição, como se vê em (iib) abaixo:

- (ii) a. Aquele psicanalista ridicularizava os pacientes diante de sua secretária.
- b. Diante de sua secretária, aquele psicanalista ridicularizava os pacientes.

- (29) a. *Todo mundo* quer ver {*seu* filho/??o filho *dele*} numa boa escola.
 b. *Todo pai* quer ver {*seu* filho/o filho *dele*} numa boa escola.
 c. *João* quer ver {??*seu* filho/o filho *dele*} numa boa escola.

Com certos antecedentes quantificacionais (que Oliveira e Silva, 1991, chamam de “totalmente gerais”), exemplificados por *todo mundo* em (29a), só o possessivo *seu* é usado; com outros (que Oliveira e Silva chamam de “indefinidos não-referenciais”), exemplificados por *todo pai* em (29b), tanto *seu* quanto *dele* podem ser usados; e, finalmente, com SNs tipicamente referenciais (que Oliveira e Silva chamam de “totalmente definidos”), exemplificados por nomes próprios como *João* em (29c), praticamente apenas a forma *dele* é usada. Este padrão distribucional mostra claramente que *seu* é favorecido – ou, talvez, usado exclusivamente – na interpretação de variável ligada: praticamente só ocorre com antecedentes quantificacionais.¹¹ Pois além de favorecer a interpretação de variável ligada, *seu* é “mais econômico” morfossintaticamente que *dele*: enquanto *seu* é como *si*, fornecendo informação apenas sobre a pessoa gramatical de seu antecedente, *dele* fornece informação sobre a pessoa gramatical, o número e o gênero. Ou seja, *seu* para fazer parte da generalização descrita em (3) acima: é uma “variável ligada típica”.

Assim, o pronome possessivo *seu* é, em parte, como o reflexivo *si*: ambos são “variáveis ligadas típicas”. Entretanto, *seu não* é como *si* no que diz respeito ao “Princípio A” da teoria da ligação: *seu não* exige um antecedente “dentro de um domínio local”, como se vê pelo contraste entre (30b) e (31b) abaixo:

- (30) a. [_s *Todo aluno desta escola* pensa primeiro em *si*].
 b. * [_s *Todo aluno desta escola* quer [_s que alguém pense primeiro em *si*]].
- (31) a. [_s *Todo aluno desta escola* faz sempre *seus* trabalhos].
 b. [_s *Todo aluno desta escola* quer [_s que alguém faça sempre *seus* trabalhos]].

¹¹ É claro que isto não explica outros aspectos do paradigma em (29), como por exemplo o fato de que *dele* não é usado com antecedentes como *todo mundo*. Para uma discussão deste aspecto específico, ver Müller (1997) e Menuzzi (1999, seção 4.2).

(30) mostra que o “domínio local” em que *si* deve ter seu antecedente é o da primeira oração que o contém; (31) mostra que *seu* não exige que seu antecedente esteja dentro deste domínio. Portanto, embora ambos *si* e *seu* sejam “variáveis ligadas típicas”, somente *si* é um “anafórico” no sentido da teoria da ligação padrão.

O ponto importante, agora, é: se *seu* mostra as mesmas restrições estruturais que *si* quanto à relação com o antecedente, devemos concluir que estas restrições se devem ao fato de ambas as formas serem “variáveis ligadas típicas” – já que *seu não* é um “anafórico” da teoria da ligação. De fato, a restrição que vimos com *si* – tem de ser c-comandado em estrutura-de-superfície por seu antecedente – também se aplica a *seu*. Isto pode ser ilustrado pelo paradigma com *seu* análogo ao paradigma para *si* em (25)-(26) acima:

- (32) a. * [_{SN} A mãe de [*quase todo rapaz*]] se preocupa com *sua* namorada.
b. [_{SN} A mãe de [*quase todo rapaz*]] se preocupa com a namorada *dele*.
c. A mãe de *quase todo rapaz* se preocupa com a namorada *dele*, mas a mãe do *João* não.
d. A mãe de *quase todo rapaz* se preocupa com a namorada de *seu filho*, mas a mãe do *João* não se preocupa com a namorada *dele, João*.

(32a) mostra que, se um SN quantificacional está contido no sujeito da oração, então ele não pode ser o antecedente de *seu* dentro de um complemento verbal – contexto em que o SN quantificacional não c-comanda *seu*.¹² Neste mesmo contexto, o mesmo antecedente poderia ser o antecedente do pronome *ele*, como indica (32b). Isto demonstra que é possível, neste contexto, ter uma forma pronominal com interpretação de variável ligada, que deve, portanto, estar no escopo do SN quantificacional (isto é, é c-comandada por ele em forma lógica). (32c) confirma esta análise: a frase de fato tem uma interpretação com “identidade ‘relaxada’” do pronome, ou seja, pode significar o mesmo que (32d). Conclui-se que a única razão pela qual (32a) é inaceitável (na interpretação em que *quase*

¹² Para mim, a única interpretação possível de (32a) é aquela em que a mãe de quase todo rapaz se preocupa com a namorada *dela, a mãe*.

todo rapaz seria o antecedente de *seu*) é por *seu* ser uma “variável ligada típica”, e este tipo de elemento pronominal exige c-comando do antecedente em estrutura-de-superfície.

Não é de surpreender, portanto, que o paradigma em (27)-(28) também pode ser reproduzido com *seu*:

- (33) a. A professora [coloca *todo aluno* na sua posição].
 b. ?? A professora [elogia *todo aluno*] no *seu* boletim.
 c. A professora [elogia *todo aluno*] no boletim *dele*.
 d. A professora [elogia *todo aluno*] no boletim *dele*, mas a diretora só elogia *os melhores*.
 e. A professora [elogia *todo aluno*] no boletim *dele*, mas a diretora só elogia *os melhores alunos* nos boletins *deles*.

(33a) mostra que um SN objeto quantificacional pode ser antecedente de *seu* dentro de um locativo complemento; (33b) mostra que esta possibilidade é (fortemente) desfavorecida quando *seu* está dentro de um locativo adjunto.¹³ (33c) indica que o problema não é com o escopo de *todo aluno* em (33b), já que se a forma pronominal é *ele*, a interpretação que se tenta obter com (33b) é possível. E (33d) confirma que o pronome *ele* está sendo interpretado como uma variável ligada em (33c), pois a frase tem a interpretação com “identidade ‘relaxada’” – isto é, (33d) pode ser interpretada como (33e). Assim, a conclusão de novo deve ser que *seu*, uma “variável ligada típica”, exige que seu antecedente a c-comande em estrutura-de-superfície.

Em resumo, *seu* parece ter essencialmente as mesmas propriedades que *si*: a) alterna com outra forma pronominal, *ele*, em certos contextos; b) nestes contextos, favorece ou exige a interpretação de variável ligada; c) exige que seu antecedente o c-comande em estrutura-de-superfície – escopo, ou c-comando em forma lógica, não sendo suficiente. Entretanto, como procurei mostrar, esta última propriedade de *seu* não pode ser atribuída ao fato de que esta forma é um “anafórico”; assim, o comportamento do possessivo indica que a restrição que exige, para certas formas pronominais, que sejam c-comandadas em estrutura-

¹³ Para mim, ao menos, o boletim só pode ser da professora em (33b), o que é um contra-senso pragmático, mas, ainda assim, a única interpretação possível.

de-superfície, deve ser atribuída ao fato de tais formas se qualificarem como “variáveis ligadas típicas”. E, é claro, se queremos obter generalidade, devemos concluir que isto vale para *si* também. E, ao que parece, para sujeitos nulos do português brasileiro, como veremos a seguir.

Sujeitos Nulos

Os chamados “sujeitos nulos” ou *pro*’s (“prozinhos”) da sintaxe gerativa – “sujeitos ocultos”, para a gramática tradicional – também alternam com formas pronominais plenas; no caso da terceira pessoa, com a forma *ele*, como se vê em (34) abaixo (o sujeito oculto é representado, no que segue, pela lacuna “__”):

- (34) *João* disse que { __ / *ele* } chegou cedo.

E, como aponto em meu artigo de 2003 (cf., também, referências lá citadas), há razões para acreditar que os sujeitos nulos – ao menos os de terceira pessoa – também são “variáveis ligadas típicas”. Em primeiro lugar, das duas formas alternantes em (34), é a que é a “mais econômica” morfossintaticamente: fornece informação sobre a pessoa e o número do antecedente (de acordo com a flexão de concordância verbal: compare com “Os rapazes disseram que __ chegaram cedo”), mas não sobre o gênero. Além disso, tende a ser interpretada como uma variável ligada em contextos como (34): (35a) tende a ser interpretada (fortemente, para mim) como (35b) – com “identidade ‘relaxada’” – e a interpretação co-referencial correspondente a (35c) é difícil de ser obtida. O pronome *ele*, por outro lado, permite facilmente a interpretação co-referencial em (34), o que é ilustrado por (36):

- (35) a. *João* disse que __ chegou cedo, e *Paulo* também.
b. *João* disse que *ele*, *João*, chegou cedo, e o *Paulo* disse que *ele*, *Paulo*, chegou cedo.

- c. # *João* disse que *ele*, *João*, chegou cedo, e *Paulo* também disse que *João* chegou cedo.
- (36) a. *João* disse que *ele* chegou cedo, e *Paulo* também.
 b. *João* disse que *ele*, *João*, chegou cedo, e o *Paulo* disse que *ele*, *Paulo*, chegou cedo.
 c. *João* disse que *ele*, *João*, chegou cedo, e *Paulo* também disse que *João* chegou cedo.

Também é importante observar que, embora tenha propriedades semelhantes aos chamados “anafóricos” da teoria da ligação, os sujeitos nulos de terceira pessoa do português (brasileiro) *não* são “anafóricos” no mesmo sentido que reflexivos; em particular, não precisam ser ligados em um domínio, como se vê em (37):¹⁴

- (37) [_s *João* disse que [_s ninguém aceitou que [_s __ tivesse escolhido Maria como secretária]]].

Em suma, o que agrupa os sujeitos nulos com o reflexivo *si* e o possessivo *seu* é que são todos elementos pronominais que, competindo com o pronome *ele*, favorecem ou exigem a interpretação de variável ligada – são todos “variáveis ligadas típicas”. Com efeito, encontramos com os sujeitos nulos as mesmas restrições estruturais que antes havíamos visto operarem sobre *si* e *seu*: sujeitos nulos exigem ou, ao menos, pedem fortemente que seu antecedente os c-comande em estrutura-de-superfície. É o que mostra o paradigma em (38) (correspondente a (25)-(26) para *si* e (32) para *seu*) – deixo ao leitor os detalhes da discussão:

- (38) a. [_{SN} *Quase todo rapaz*] diz que __ é um bom menino.
 b. ??[_{SN} *A mãe de quase todo rapaz*] diz que __ é um bom menino.

¹⁴ Embora relações locais possam facilitar sua interpretação: ver Figueiredo Silva (1996, capítulo 4) para discussão em que se sugere que *pro* em português brasileiro possui propriedades de um elemento “anafórico”.

- c. [_{SN} A mãe de [quase todo rapaz]] diz que *ele* é um bom menino.
- d. [A mãe de *quase todo rapaz*] diz que *ele* é um bom menino, mas a mãe do *João* não.
- e. [A mãe de *quase todo rapaz*] diz que *seu filho* é um bom menino, mas a mãe do *João* não diz que *João* é um bom menino.

Assim, o comportamento dos sujeitos nulos parece confirmar a conclusão a que havíamos chegado antes com a discussão sobre o reflexivo *si* e o possessivo *seu*: “variáveis ligadas típicas” exigem, ou ao menos preferem fortemente, que seu antecedente as c-comande em estrutura-de-superfície. O fato de estarem no escopo do antecedente – isto é, de serem por eles c-comandadas em forma lógica – parece não ser suficiente para permitir seu uso. Isto, como procurei mostrar, surpreende, já que a restrição não se aplica às formas pronominais que competem com as “variáveis ligadas típicas” – por exemplo, *ele* em português (brasileiro): para estas formas, estar no escopo do antecedente é suficiente para serem interpretadas como variáveis ligadas.

Conclusão

Neste artigo, procurei mostrar que certas formas pronominais, a que chamei de “variáveis ligadas típicas”, são misteriosas porque embora sejam especializadas para expressar a interpretação de variável ligada, não podem ser usadas exatamente nos mesmos contextos em que a interpretação de variável ligada é permitida. Em particular, vimos que não basta para as “variáveis ligadas típicas” estarem no escopo de seu antecedente – isto é, serem por ele c-comandadas em forma lógica; antes, exigem c-comando em estrutura-de-superfície. Esta situação é misteriosa em pelo menos dois sentidos.

Do ponto de vista empírico, é curioso que a restrição estrutural mencionada coincida com uma propriedade morfossintática: Por que as formas que são especializadas para a interpretação de variável ligada, simultaneamente, são formas “econômicas” morfossintaticamente e exigem c-comando em estrutura-de-superfície? Qual a relação entre estas duas propriedades? E qual a relação delas com a interpretação de variável ligada? Tanto quanto sei, não há

resposta para estas perguntas (mas ver Reuland, 2001, para discussão de questões semelhantes e para sugestão de análise).

Do ponto de vista conceitual – ou da articulação geral da teoria lingüística –, o mistério me parece ainda mais interessante. Se as línguas naturais desenvolvem expressões especializadas para uma certa função – no caso tratado aqui, a interpretação de variável ligada –, seria de se esperar que as propriedades desta função determinassem o comportamento da expressão especializada. Aliás, esta intuição básica é a mesma que está por trás do que se chama de “iconicidade” nos estudos funcionalistas da linguagem (ver, por exemplo, Givón, 1990). Assim, um elemento especializado na interpretação de variável ligada deveria ter seu comportamento determinado por aquilo que define esta interpretação. Como vimos, o que define a possibilidade de um elemento ser interpretado como uma variável ligada é o fato dele estar no escopo de um operador. Entretanto, surpreendentemente, estar no escopo de um operador não é suficiente para “variáveis ligadas típicas”.

O resultado disto é uma ausência de correspondência direta entre categorias da semântica e da sintaxe, que se pode resumir no seguinte quadro:

<i>Tipos de Expressão</i> (Sintaxe)		<i>Tipos de Interpretação</i> (Semântica)
Variáveis Ligadas Típicas (<i>si, seu, pro</i>)	← →	Variável Ligada
Pronomes (<i>ele</i>)	← →	Co-referência

O caso das “variáveis ligadas típicas” é, então, mais um exemplo de um fato mais geral, o de que não há correspondência de um-a-um entre categorias da sintaxe e categorias da semântica, fato que coloca vários problemas interessantes. Um deles é da unificação: se não é possível reduzir a sintaxe à semântica (nem vice-versa, evidentemente), qual a natureza da relação que há entre estes dois componentes lingüísticos? É ela arbitrária, motivada, ou, ainda, parcialmente motivada? E, é claro, esta última questão levanta outra, sobre as relações causais que há entre sintaxe e semântica: se as línguas desenvolvem instrumentos de expressão especializados para certas funções de significado – presumindo que seja este o caso das “variáveis ligadas típicas” –, como e por

que estes instrumentos adquirem “vida própria”, isto é, adquirem propriedades que não parecem ser redutíveis ao que os motivou originalmente? Estas são questões extremamente interessantes, que merecem discussão profunda e detalhada. Aqui, deixo para o leitor a indicação dos trabalhos de Frederick Newmeyer, que desenvolve de modo interessante estes tópicos a partir da perspectiva gerativista (ver, por exemplo, Newmeyer, 1992; 2000).

RESUMO

O artigo procura mostrar que certas formas pronominais, chamadas de “variáveis ligadas típicas”, levantam questões importantes acerca das relações entre sintaxe e semântica. Tais formas pronominais – o reflexivo *si*, o possessivo *seu* e os *sujeitos nulos* do português brasileiro – são especializadas para expressar a interpretação de variável ligada (cf. Menuzzi, 2003); no entanto, não podem ser usadas exatamente nos mesmos contextos em que esta é permitida. Em particular, não basta para as “variáveis ligadas típicas” estarem no “escopo” de seu antecedente – isto é, serem por ele c-comandadas em forma lógica. Antes, as “variáveis ligadas típicas” exigem c-comando em estrutura-de-superfície. Esta falta de correspondência entre categorias da sintaxe e da semântica indica uma relação complexa entre ambos os componentes da gramática.

Palavras-chave: escopo, pronomes, interpretação de variável ligada, interface sintaxe-semântica, português brasileiro.

ABSTRACT

This paper tries to show that certain pronominal forms, called “typical bound variables”, raise important issues concerning the relation between syntax and semantics. The relevant forms – the reflexive *si*, the possessive *seu* and *null subjects* in Brazilian Portuguese – are specialized in expressing the bound variable interpretation (cf. Menuzzi, 2003); however, they cannot be used in exactly the same contexts in which that interpretation is allowed. In particular, it is not enough for “typical bound variables” to be in the “scope” of their antecedents – that is, to be c-commanded by them at logical form. This mismatch between syntactic and semantic categories points to a complex relationship between syntax and semantics.

Key-words: scope, pronouns, bound variable interpretation, syntax-semantics interface, Brazilian Portuguese.

REFERÊNCIAS

CHIERCHIA, G.; McCONNELL-GINET, S. *Meaning and Grammar: An Introduction to Semantics*. Cambridge (EUA): MIT Press, 1990.

CHOMSKY, N. *Syntactic Structures*. Haia: Mouton, 1957.

_____. *Lectures on Government and Binding*. Dordrecht (Holanda): Foris, 1981.

_____. *Knowledge of Language: Its Nature, Origin and Use*. Nova Iorque: Praeger, 1986.

DOWTY, D.; WALL, R. E.; PETERS, S. *Introduction to Montague Semantics*. Dordrecht (Holanda): Reidel, 1981.

EICK, P. F. *Efeitos da condição de subyacência em português brasileiro*. Porto Alegre, 2002. Tese (Doutorado) - Pontifícia Universidade Católica -RS.

FIGUEIREDO SILVA, M. C. *A posição sujeito em português brasileiro em frases finitas e infinitas*. Campinas: Editora da Unicamp, 1996.

GIVÓN, T. *Syntax: A Functional-Typological Introduction*. Amsterdã: John Benjamins, 1990. v. 2.

HAEGEMAN, L. *Introduction to Government and Binding Theory*. 2. ed. Cambridge (Inglaterra): Blackwell, 1994.

MAY, R. *Logical Form: Its Structure and Derivation*. Cambridge (EUA): MIT Press, 1985.

MENUZZI, S. 3rd Person Possessives in Brazilian Portuguese: On the Syntax-Discourse Relation. In: BOTLEY, S. et al. (Eds.). *Approaches to Discourse Anaphora: Proceedings of the Discourse Anaphora and Resolution Colloquium (DAARC96)*. Lancaster (Inglaterra): UCREL Technical Papers 8, Lancaster University, 1996. p. 191-210.

_____. *Binding Theory and Pronominal Anaphora in Brazilian Portuguese*. Haia (Holanda): Holland Academic Graphics, 1999. (Série “LOT Dissertations”, v. 30)

_____. Sobre a Evidência para a Maturação de Universais Lingüísticos. *Letras de Hoje*, Porto Alegre, v. 36, n. 3, p. 141-166, 2001.

_____. Sobre as Opções Anafóricas para Antecedentes Genéricos e para Variáveis Ligadas. *Letras de Hoje*, Porto Alegre, v. 38, n. 1, p. 125-144, 2003.

MÜLLER, A. *A Gramática das Formas Possessivas no Português do Brasil*. Campinas, 1997. Tese (Doutorado) - Universidade Estadual de Campinas.

_____. Anaphora and Genericity in Brazilian Portuguese. *Letras de Hoje*, Porto Alegre, v. 38, n. 1, p. 109-124, 2003.

_____: NEGRÃO, E. As Mudanças no Português Brasileiro: Substituição ou Especialização? *DELTA*, n. 12, p. 125-152, 1996.

NEWMeyer, F. Iconicity and Generative Grammar. *Language*, n. 68, p. 756-796, 1992.

_____. *Language Form and Language Function*. Cambridge (EUA): MIT Press, 2000.

RADFORD, A. *Transformational Syntax: A Student's Guide to Chomsky's Extended Standard Theory*. Cambridge (Inglaterra): Cambridge University Press, 1981.

_____. *Transformational Syntax: A First Course*. Cambridge (Inglaterra): Cambridge University Press, 1988.

REINHART, T. *The Syntactic Domain of Anaphora*. Cambridge, 1976. Tese (Doutorado) - MIT.

_____. *Anaphora and Semantic Interpretation*. Londres: Croom Helm, 1983.

_____. Quantifier Scope: How Labor is Divided between QR and Choice Functions. *Linguistics and Philosophy*, n. 20, p. 335-397, 1997.

_____. Strategies of Anaphora Resolution. In: BENNIS, H.; EVERAERT, M.; REULAND, E. (Eds.). *Interface strategies*. Amsterdã: North Holland, 2000. p. 295-325.

REULAND, E. Primitives of Binding. *Linguistic Inquiry*, n. 32, p. 439-492, 2001.

ROSS, J. R. *Constraints on Variables in Syntax*. Cambridge, 1967. Tese (Doutorado) - MIT.

SAG, I. *Deletion and Logical Form*. Cambridge, 1976. Tese (Doutorado) - MIT.

WILLIAMS, E. Discourse and Logical Form. *Linguistic Inquiry*, n. 15, p. 131-153, 1977.