

EXTRAÇÃO DO SUJEITO: A ANÁLISE DE STEEDMAN E ANÁLISES DA TEORIA GERATIVA

Thais Chaves *

Em seu livro *Surface structure and interpretation*, Mark Steedman propõe um modelo de gramática categorial, a gramática categorial combinatória, e revê alguns fenômenos conhecidos, apresentados pela teoria gerativa, propondo um outro tipo de análise, dentro de seu modelo. Assim, o autor pretende mostrar que a sua teoria pode dar conta dos mesmos fenômenos de que a teoria gerativa trata, fazendo uso de menos primitivos e sem fazer apelo à noção de *movimento*. Neste artigo, comparo a análise proposta por Steedman (1996) para os fenômenos de extração de sujeitos e a análise fornecida pela teoria gerativa. A análise de Steedman se encontra no item “Assimetrias entre Sujeito e Objeto”, do Capítulo 3 de seu livro. Aqui, vou me limitar a alguns casos de extração do sujeito, sem pretender dar conta de tudo o que Steedman apresenta nesse item.

Os dados em questão são do seguinte tipo: frases como (1) são agramaticais, enquanto a frase em (2) é perfeita.

- (1) *Who_i do you think [t_i that t_i saw Bill] ?
- (2) Who_i do you think [t_i that Bill saw t_i] ?

* Mestranda em Linguística na FFLCH-USP.

Steedman apresenta, para comparação, a análise que Chomsky propõe em seu livro *Barriers*, de 1986. No sistema desenvolvido nesse livro, fenômenos desse tipo são explicados através de movimentos. Os itens lexicais são projetados do léxico dentro de certas relações e, através de movimentos, se reorganizam dentro da estrutura sintática permitida por aquela língua. Para se formarem questões em inglês, por exemplo, os pronomes interrogativos devem se mover de sua posição de origem para a posição de especificador de CP, no início da sentença.¹ Assim, em (1) o pronome interrogativo *who* é gerado na posição de sujeito do verbo *see* e posteriormente se move até a posição de especificador de CP; em (2), o pronome é gerado na posição de objeto do verbo e se move para essa mesma posição. Nos dois casos, o movimento é necessário para formar uma sentença interrogativa.

O movimento é um recurso teórico largamente utilizado na teoria gerativa. Desde o trabalho de Ross (1967), foram observadas certas restrições gerais em relação aos tipos de movimento. Ross fez um trabalho descritivo, em que ele levanta as proibições sobre o movimento encontradas em diversas línguas. Um movimento “proibido” é aquele que gera uma sentença agramatical e que, portanto, não deveria ser previsto pelo sistema. Os locais de onde o movimento – ou extração – é impossível, ou pelo menos difícil, são conhecidos como *ilhas*. Assim, temos vários contextos sintaticamente diferentes que proíbem a extração, como sintagmas nominais complexos, sujeitos sentenciais, estruturas coordenadas e outros.

A teoria gerativa tentou dar conta desse conjunto de fenômenos de maneira unificada. O primeiro mecanismo a ser proposto tentava dar conta da impossibilidade de extração desses constituintes através de uma proibição de movimentos *longos*. Assumindo-se que as sentenças se compõem em estruturas hierarquicamente organizadas, e que os constituintes estabelecem relações locais entre si, pode-se estabelecer uma noção de *distância* entre as posições em que os constituintes são gerados e as posições para as quais eles devem se mover. Essa distância vai depender, então, do tipo de configuração sintática em que o constituinte foi gerado: quanto mais *encaixado* ele estiver em um outro constituinte, por exemplo, mais difícil será o movimento para uma posição externa a esse constituinte.

O problema para a teoria gerativa, então, era como definir quais distâncias eram excessivas. Criou-se um princípio, o princípio de subjacência, que tentava estipular quantos nós-barreira um constituinte podia *ultrapassar*, ou

1 CP é a projeção máxima da categoria funcional C(omplementizador).

atravessar, ao se movimentar; os nós-barreira eram os constituintes NP e IP (para línguas como o inglês), ou NP e CP (para línguas como o italiano e o português).

É a partir da postulação do princípio de subjacência que as assimetrias entre sujeito e objeto começam a ficar mais claras. Em um conjunto de frases como (3),

(3)

?*Quem_i o Pedro soube d(e) [_{SN} a notícia [_{CP} de que t_i consertou o seu carro]]?

??O que_i o Pedro soube d(e) [_{SN} a notícia [_{CP} de que o mecânico consertou t_i]]?

*Como_i o Pedro soube d(e) [_{SN} a notícia [_{CP} de que o mecânico consertou o seu carro t_i]]?

a partir do princípio de subjacência, esperaríamos que todas as sentenças fossem igualmente ruins, já que os componentes movidos estariam atravessando o mesmo número de nós-barreira; entretanto, as sentenças apresentam diferentes graus de aceitabilidade. Embora todas as sentenças sejam marginais, a sentença da qual se extraiu o objeto é mais aceitável do que a sentença da qual se extraiu o sujeito, enquanto a sentença da qual se extraiu o adjunto é totalmente impossível com a interpretação desejada.

Da mesma forma, em casos como (4),

(4)

a. Who_i [_{IP} did John ask [_{CP} whether [_{IP} Mary will help t_i]]]?

b. Who_i [_{IP} did John ask [_{CP} whether [_{IP} t_i will help Mary]]]?

o movimento do constituinte atravessa o mesmo número de nós-barreira, mas a extração do sujeito gera sentenças menos aceitáveis do que aquelas em que há uma extração do objeto.

Para dar conta dessas assimetrias, criou-se o princípio das categorias vazias (the empty category principle - ECP).

Um constituinte, ao se mover, deixa um vestígio cuja interpretação é dependente, ou seja, o vestígio de um constituinte movido funciona como uma variável presa. Assim, em (5):

(5)

a. a man [who_i [_{IP} I think [_{CP} that [_{IP} Keats likes t_i]]]

b. *a man [who_i [_{IP} I think [_{CP} that [_{IP} t_i likes Keats]]]

o índice *i* indica que os vestígios de movimento são interpretados como tendo a mesma referência que os sintagmas movidos.

O princípio das categorias vazias define o conceito de regência apropriada e estipula que categorias vazias precisam ser apropriadamente regidas.

Criam-se então condições para a existência das categorias vazias resultantes do movimento de constituintes. Dessa forma, o movimento de dentro de certos contextos criaria categorias vazias “ilegais” e seria, portanto, bloqueado por esse princípio.

A regência apropriada ocorre em duas situações: (i) regência lexical, que é a regência por um núcleo; e (ii) regência por antecedente, em que o vestígio é c-comandado por um antecedente, e não está “muito longe” dele.

Como a regência lexical é a relação que se estabelece entre um núcleo e seu complemento, essa definição disjuntiva do ECP faz com que o vestígio do objeto de um verbo seja sempre apropriadamente regido, enquanto os vestígios de sujeitos e adjuntos só podem satisfazer o ECP se forem regidos por antecedência. Uma possível definição de regência por antecedência é:

Regência por antecedência:

α (categoria movida) rege β por antecedência se:

(a) α c-comanda β

(b) no máximo uma categoria bloqueadora domina β mas não α
(categorias bloqueadoras: IP e DP)

(c) não há nenhuma posição-C preenchida que c-comande minimamente β e não c-comande α .

Assim, partindo dessas definições, em (2), o vestígio t_i é lexicalmente regido pelo verbo *like*, o que o desobriga a ser regido por antecedência. Já em (1), o vestígio t_i não é lexicalmente regido, já que não é o complemento do verbo, e não é regido por antecedência, já que a presença de *that* impede esse tipo de regência. Assim, esse vestígio viola o ECP e a construção é agramatical.

Steedman (1996) discute o fenômeno das assimetrias entre sujeitos e objetos, entendidas como restrições sobre dependências à distância. Steedman afirma que os fenômenos que são tratados como conseqüências do princípio das categorias vazias (empty category principle - ECP) têm uma explicação direta dentro da gramática categorial combinatória e são o resultado da direcionalidade dos argumentos do verbo em uma língua SVO. Nas gramáticas categoriais, os predicados são entendidos como funtores que exigem argumentos. Um verbo transitivo, por exemplo, pode ser entendido como um funtor que toma um sintagma nominal como argumento e gera um outro funtor, que, por sua vez, toma um outro sintagma nominal como argumento para gerar uma sentença. Esses funtores podem ser especificados quanto à direção em que tomam seus argumentos; no caso do verbo transitivo, por exemplo, o verbo poderia ser especificado para tomar um sintagma nominal objeto à sua

direita, e o funtor resultante poderia ser caracterizado por tomar um sintagma nominal sujeito à sua esquerda.² Assim, em uma língua como o inglês, a entrada lexical de um verbo transitivo traz informações sobre a direção em que os argumentos desse verbo devem aparecer, e as regras combinatórias que se aplicarão sobre esse funtor são sensíveis a essa direcionalidade. Dentro desse contexto, são as características do próprio verbo e do sistema combinatório que impedem que se extraia um sujeito, enquanto a extração de um objeto é permitida.

Steedman apresenta o seguinte exemplo:

- (6) a. a man who(m) [I think that]_{S/S} [Keats likes]_{S/NP}
 b. *a man who(m) [I think that]_{S/S} [likes Keats]_{S/NP}

O autor explica que, para que fosse possível extrair da posição de sujeito, com o que a sentença (17b) seria gramatical, seria necessário que a gramática do inglês dispusesse de uma regra composicional diferente das quatro regras combinatórias que ele apresentou no capítulo anterior. Essa regra seria a regra composicional *forward crossing*, $>B_x$,³ para compor as categorias *S/S* e *S/NP*.

Steedman argumenta que essa regra seria permitida pela teoria, mas que a sua utilização geraria uma outra língua, que não o inglês, a qual não teria a propriedade de configuracionalidade, que é crucial para o inglês. Sem essa propriedade, não haveria qualquer regra de ordenação de palavras, o que permitiria sentenças *embaralhadas*. Assim, se não é possível utilizar uma regra combinatória que permita a formação das categorias derivadas *S/S* e *S/NP* no caso acima, a impossibilidade de se extrair o sujeito é uma parte necessária do sistema dessa língua, e do sistema de línguas SVO configuracionais em geral.

A análise de Steedman, como ele mesmo aponta, é muito mais direta para dar conta da proibição de extração de sujeitos do que a análise proposta pela teoria gerativa. Com a gramática categorial combinatória, damos conta

2 As relações de direcionalidade podem ser expressas por diversos tipos de notação. Steedman (1996) apresenta à esquerda o funtor principal, e define a direcionalidade através de barras: “/” define que o funtor toma um argumento à direita, enquanto “\” indica que o funtor toma um argumento à esquerda. Um verbo como *eat*, usado como exemplo por Steedman (1996, p. 9), seria representado pela seguinte notação:

(2) *eats* := (S/NP)/NP,

onde (S\NP) é o funtor que obtemos quando o verbo toma um argumento NP à direita. A notação do funtor indica que ele vai gerar uma sentença (S) se tomar um argumento NP à esquerda.

3 A regra *forward crossing* é a regra de composição funcional (106b), apresentada no capítulo 2 de Steedman (1996, p. 43) e sua forma é: $X/Y \ YZ \Rightarrow_B \ XZ$.

do contraste entre frases como (1) e (2), ou (3) e (4), sem fazer apelo a categorias vazias ou a movimento de categorias; as relações se estabelecem entre constituintes adjacentes, e as propriedades desses constituintes devem ser satisfeitas para que a sentença resultante seja reconhecida como uma sentença pertencente à língua em questão.

Entretanto, se comparamos (1) e (7), vemos que a frase sem o complementizador *that* permite a extração, já que (7) é gramatical.

(1) *Who do you think [t_i that t_i saw Bill] ?

(7) Who do you think [t_i t_i saw Bill] ?

Para a teoria gerativa, a frase (7) é gramatical porque a extração do sujeito inteiro não é proibida em princípio; ela gerará frases agramaticais se, na sentença formada, o vestígio desse sujeito não for apropriadamente regido. A *condição do sujeito*, uma das ilhas de Ross, mostra que a extração de constituintes de dentro de sujeitos sentenciais gera sentenças agramaticais, mas não proíbe a extração do sujeito como um todo. Observa-se, entretanto, que a presença do complementizador *that* parece bloquear sempre a regência apropriada – é o que chamamos de *that-trace effect*.

Para explicar o mesmo fato, Steedman cria um mecanismo específico: o autor considera que a possibilidade de extração de sujeitos em inglês (quando o complementizador não está presente) é uma exceção à regra geral que impede a extração (*fixed subject constraint*). O autor lembra que não se pode introduzir uma regra de composição *crossing forward*, mesmo que a restringíssemos muito, pois seria impossível evitar que ela fosse excessivamente produtiva. Assim, para Steedman, a única maneira de dar conta desses fatos seria introduzindo-os no léxico, de forma que um verbo como *think* tivesse, além das categorias óbvias, uma outra categoria específica para construções de extração de sujeito.

Assim, os verbos que permitem a extração do sujeito de seu complemento sentencial são marcados no léxico com uma categoria *-LEX*; da mesma forma, seus argumentos devem ser marcados com o mesmo traço, o que impede, segundo Steedman, que estes argumentos sejam saturados por um complemento com materiais lexicais, de qualquer tipo. Esse traço funciona da mesma maneira que os traços de concordância que Steedman propõe em seu sistema. Uma consequência dessa marcação é que a interpretação semântica das categorias formadas, que nas gramáticas categoriais sempre segue imediatamente paralela à sintaxe, vai determinar que o objeto do verbo marcado com o traço *-LEX* deve ser interpretado como o sujeito de seu complemento (sentencial). Nesse sistema, os argumentos não-sujeitos que podem ser extraídos permanecem subespecificados para o atributo *LEX*.

Para Steedman, seu sistema apresenta vantagens por concentrar todos os problemas no léxico; dessa forma, ele dá conta dos dados sem precisar utilizar categorias vazias em sua teoria, e sem o aparato teórico da gerativa, que reformula a noção de regência e cria o conceito de regência apropriada. Dentro do quadro teórico das gramáticas categoriais, o tamanho do léxico ou a complexidade de sua composição não representam problemas, já que essas não são teorias com uma preocupação imediata com a realidade mental do sistema da gramática. Aos olhos de um gerativista, entretanto, a solução em que se postula um traço lexical para dar conta de um fenômeno sintático parece, no mínimo, *ad hoc*. Se analisarmos, novamente, as frases em (8),

(8)

- a. *Who do you think that saw Bill?
- b. Who do you think saw Bill?
- c. Who do you think Bill saw?
- d. *Which rock star_i was [_{CP} that the police would arrest t_i] expected?⁴

veremos que, para dar conta desse conjunto de dados, a teoria gerativa (P&P), na versão com a qual Steedman compara seu modelo, vai se utilizar do ECP e, portanto, do conceito de regência apropriada, com uma estipulação que afirma que o complementizador bloqueia a regência apropriada. O modelo de Steedman, embora dê conta de forma imediata de uma parte dos dados, cria novas entradas lexicais para os verbos e complementos envolvidos em frases com extração de sujeito, e proíbe a extração de dentro de sujeitos sentenciais, como em (8d), por meio de uma estipulação. Além disso, da forma como Steedman analisa as sentenças com o complementizador *that*, não fica claro qual o papel que ele representa na sentença, e por que as frases em que ele não ocorre são tão diferentes das frases em que ele ocorre. O autor não comenta em nenhum momento as semelhanças entre os dois tipos de frases.

Se comparamos apenas as análises propostas para as sentenças tomadas como exemplo por Steedman, os dois modelos parecem equivalentes, já que propõem uma solução geral e uma exceção. Entretanto, na análise proposta pela teoria gerativista, todos os fenômenos de extração são explicados de forma paralela, já que são gerados pelo mesmo tipo de mecanismo sintático. Ao mesmo tempo, os conceitos envolvidos no ECP, como c-comando ou regência por antecedência, são independentemente motivados e necessários também para outros módulos da teoria, sendo responsáveis por vários mecanismos de interpretação, na teoria de ligação, e tendo efeitos sobre quase todos os fenômenos sintáticos.

4 Exemplo (12a) do capítulo 4 de Roberts, I. *Comparative Syntax*. p. 187.

O próprio fenômeno que está na base das diferenças entre as frases (1) e (7), o *that-trace effect*, era um fenômeno especialmente problemático no modelo proposto em *Barriers*, que por sua vez também era um modelo problemático, por sua complexidade; para explicar o *that-trace effect*, Chomsky criou o conceito de *minimalidade*. Esse conceito foi posteriormente reformulado por Rizzi (1990), que criou o novo conceito de *minimalidade relativizada*, reformulando o modelo como um todo e tornando o conceito de *localidade* um conceito central para a teoria gerativa. Se o modelo de *Barriers* acabou por ser superado, ele abriu caminho para essa reformulação, que levou a uma teoria ainda mais geral e que dá conta de muitos fenômenos diversos fazendo uso de um número cada vez menor de primitivos. Assim, parece-nos que, de uma forma global, a explicação fornecida pela teoria gerativa para a extração de sujeitos, mesmo no modelo utilizado para comparação, é mais homogênea e faz menos apelo a mecanismos específicos do que o modelo proposto por Steedman.

RESUMO

Neste trabalho, comparo a análise de Steedman (1996) com a análise de Chomsky (1986) para o fenômeno da extração de sujeitos. Steedman (1996) propõe uma análise no quadro da gramática categorial combinatória, explicando o fenômeno através da direcionalidade de funtores e propondo que verbos que tomam como complemento orações das quais o sujeito é “extraído” são especialmente marcados no léxico para fazer parte dessa configuração. Chomsky (1986) propõe um modelo dentro do qual as impossibilidades de extração dependem da configuração sintática em que os constituintes estão inseridos, e apresenta uma exceção, o *that-trace effect*, que é explicado como um efeito de minimalidade.

Palavras-chave: gramática categorial, sintaxe, extração do sujeito.

ABSTRACT

In this paper, I compare the analysis proposed by Steedman (1996) for subject extraction to that proposed by Chomsky (1986). The analysis by Steedman (1996) is inserted in the framework of combinatory categorial grammar; it rules out sentences where subject extraction has taken place as impossible sentences, given the directionality

of the functors in a language like English. In the cases where subject extraction is possible, the verb is specially marked in the lexicon to allow this possibility. Chomsky (1986) proposes a model in which extraction is determined by the syntactic configuration in which the constituents are generated; one exception to the general case, the *that-trace effect*, is explained in terms of minimality.

Key-words: categorial grammar, syntax, subject extraction.

REFERÊNCIAS BIBLIOGRÁFICAS

- CHOMSKY, Noam. *Barriers*. Cambridge, Massachusetts: The MIT Press, 1986.
- STEEDMAN, Mark. *Surface structure and interpretation*. Cambridge, Massachusetts: The MIT Press, 1996.
- RIZZI, Luigi. *Relativized minimality*. Cambridge, Massachusetts: The MIT Press, 1990.
- ROBERTS, Ian. *Comparative syntax*. Londres: Arnold, 1996.
- ROSS, John R. *Constraints on variables in syntax*. Cambridge, Massachusetts, 1967. Tese (Doutorado) - MIT.