

RELATÓRIO TÉCNICO
PROJETO
“GRAMMAR PLAY”: REVISÕES E AMPLIAÇÕES

EQUIPE:

Sérgio Menuzzi (pesquisador responsável, UFRGS)

Gabriel Othero (bolsista de doutorado, PUCRS)

Gabriela Conteratto (bolsista de doutorado, PUCRS)

Maurício Piccini (assessor de programação)

Financiado pelo CNPq, Modalidade APQ, Processo 402810/2004-0

Porto Alegre, Outubro de 2006

1. Introdução

De acordo com a proposta submetida ao CNPq (Edital no. 32/2004, Processo 402810/2004-0; proposta em anexo), o projeto “Grammar Play: Revisões e Ampliações” tinha como objetivo estender o domínio de aplicação do programa “Grammar Play”, um aplicativo de parsing automático do Português Brasileiro que começou a ser desenvolvido em 2004 por Gabriel de Ávila Othero, sob orientação de Sergio Menuzzi. Especificamente, o projeto se propunha a:

a) resolver alguns dos vários problemas técnicos então identificados no “Grammar Play”, entre os quais: problemas da interface gráfica destinada ao uso didático do “Grammar Play”; problemas de implementação do *parser* gramatical propriamente dito, como a conhecida incompatibilidade das estruturas de recursão à esquerda (p.ex., $N' \rightarrow N' XP$, como em [[[*livro interessante* N'] *de Chomsky* N'] *sobre sintaxe* N']) com os métodos de *parsing top-down*;

b) ampliar o domínio da operação do *parser* sintático do “Grammar Play”, implementando regras sintáticas para estruturas oracionais mais complexas do que

orações simples, incluindo estruturas com verbos auxiliares bem como para a complementação e a coordenação oracionais;

c) integrar às representações lexicais do *parser* informação sobre a estrutura de argumentos dos itens lexicais e suas restrições seletivas;

d) estender o domínio de operação do aplicativo para a interpretação semântica das frases, incorporando a ele um componente de identificação e representação das relações de predicado/argumento (em particular, das funções semânticas, ou “papéis temáticos”, de agente, paciente, etc.).

Em função das mudanças de vínculo empregatício do responsável pelo projeto (Sergio Menuzzi deixou a PUCRS em agosto de 2005, trabalhou sob contrato temporário com a UNICAMP entre agosto de 2005 e fevereiro de 2006 e, finalmente, assumiu o posto de professor adjunto na UFRGS em março de 2006), a implementação da proposta aprovada pelo CNPq só iniciou em outubro de 2005, sendo que suas atividades foram dadas como encerradas em outubro de 2006. Neste período, trabalharam no projeto Sergio de Moura Menuzzi (pesquisador principal), Gabriel Othero, Gabriela Conteratto (bolsistas do PPGL/PUCRS) e Maurício Piccini (assessor de programação). Com relação aos objetivos do projeto, pode-se dizer, em síntese, que todos os objetivos da proposta foram ao menos parcialmente atingidos.

2. Resultados Técnicos

2.1 Melhorias na interface gráfica

Numerosos ajustes foram feitos para melhorar o Grammar Play, tornando-o mais “portável” e mais *user-friendly*. Os principais são os seguintes:

- As fontes do input e do output foram alteradas para melhor visualização.
- O programa de interface tinha originalmente 17.7 Mb; a versão atual tem apenas 2.3 Mb, o que tornou mais fácil e rápida a instalação do Grammar Play.
- Todos os rótulos de categorias agora aparecem em vermelho na estrutura de colchetes rotulados, melhorando sua visualização.
- Os rótulos **v_bar**, **sn**, **n_bar**, etc., utilizados pelo Prolog, também eram utilizados na interface gráfica. Na versão atual, eles foram substituídos por **V'**, **SN**, **N'**, etc., rótulos utilizados por sintaticistas.

- O menu de **Opções** passou a ter quatro comandos: **Ver estrutura arbórea**, **Ver Gramática**, **Ver Léxico** e **Créditos**. Assim, o usuário passa a ter acesso, por meio da interface gráfica, a todos os principais componentes do Grammar Play. No topo da janela do programa, o usuário agora pode ver o nome e a versão do Grammar Play.
- O alinhamento de nós da estrutura arbórea das sentenças agora permite visualizar as árvores tal como são comumente desenhadas por sintaticistas; em particular, não é mais utilizado o alinhamento dos “nós nucleares” entre si (como aparece em Haegeman 1994), mas, como na maior parte das publicações em sintaxe, o nó-mãe é centralizado e equidistante dos nós-filhos.
- É possível visualizar as árvores de sentenças de qualquer extensão: foram implementados botões de rolagem que permitem visualizar as partes da árvore que não couberam na janela automática (como acontece com páginas muito grandes nos browsers da internet).

2.2 *O problema da recursão-à-esquerda*

Uma das principais melhorias obtidas para o parser gramatical propriamente dito ainda não foi implementada no Grammar Play, mas o programa-protótipo está pronto e precisa apenas ser adaptado e integrado ao Grammar Play: trata-se de uma solução adequada para o problema das regras de recursão à esquerda.

É preciso lembrar que, na versão original do “Grammar Play”, o conhecido problema da recursão-à-esquerda, comum aos *parsers top-down*, especialmente aos implementados em Prolog (cf. Pereira & Shieber 1987; Gazdar & Mellish 1989, 1996, entre outros), foram evitados por meio de uma solução *ad hoc*: o uso de “símbolos posições”, isto é, rotulação não-categorial em regras de reescrita X-barras (ver Othero 2004 para maiores detalhes). Embora o problema não pareça incomodar aos lingüistas computacionais (cf. Gazdar & Mellish 1989, p.157 e ss., entre outros), ele atrapalha o Grammar Play, cujo objetivo é tentar implementar, da maneira mais direta possível, teorias lingüísticas estabelecidas (ver Othero 2004 e Othero & Menuzzi 2006).

Há, é claro, alguns lingüistas computacionais que se sentem incomodados com a transformação de gramáticas recursivas em não-recursivas; mas estes incomodados, em geral, sugerem que a alternativa é utilizar o processamento *bottom-up* (Convington

1994, caps. 3 e 6; Bouma & Villada 2002/2003; Blackburn et al. 2006), alternativa incompatível com as escolhas do Grammar Play: em função do seu caráter didático, o Grammar Play faz uso do recurso das *Definite Clause Grammars* – o que permite converter regras de reescrita automaticamente em programa de parsing em Prolog; essas regras, é claro, atuam de modo *top-down, left-to-right*.

Assim, diante das escolhas feitas para o Grammar Play, o que seria preciso era uma solução que, embora adotando o recurso das DCGs, ainda assim permitisse evitar o problema da recursão-à-esquerda sem os “símbolos postiços” (solução que precisa de várias regras de reescrita *ad hoc*, artificiais – não apenas de difícil compreensão para estudantes, mas também sem fundamento lingüístico). A solução desenvolvida para o Grammar Play é, na verdade, a conciliação de dois modos alternativos de conceber a teoria X-barra que são encontrados na literatura: (a) como um conjunto limitado de diferentes tipos de categorias sintagmáticas projetadas a partir do mesmo núcleo (como na concepção *standard*, em que um núcleo X possui uma projeção X^0 , projeções intermediárias X' ou X^1 e projeções máximas X'' ou X^2 ou XP, cf. Chomsky 1970, Stowell 1981, Haegeman 1994); (b) como um conjunto infinito de categorias do mesmo tipo (em que, na verdade, as distinções entre X^0 , X' e XP são “contextuais”, e não parte dos princípios básicos de projeção sintagmática, como em Radford 1981, Stuurman 1985 e Chomsky 1985).

Concretamente, o programa-protótipo desenvolvido criou, para as categorias intermediárias (isto é, para X' , a “categoria recursiva” na teoria X-barra *standard*), um “subíndice de projeção” identificado por um número natural – isto é, cada subíndice n , embora resultado de uma regra recursiva, ainda assim pode ser sempre distinguido do subíndice $n-1$ que o precede e ao qual a regra se aplicou. Assim, no protótipo, embora as regras de reescrita identifiquem ainda diferentes tipos de categorias (cf. a versão (a) da teoria X-barra acima mencionada), elas também tratam as categorias intermediárias como um conjunto potencialmente infinito de categorias do mesmo tipo (cf. a versão (b) da teoria X-barra), sendo assim capazes de distinguir “várias ocorrências” de uma categoria intermediária. Por exemplo, enquanto [*livro de Chomsky*] é analisado como a categoria $N'(0)$, [*livro de Chomsky sobre sintaxe*] é analisado como $N'(1)$ (composto de [$N'(0) + PP$]), e [*livro de Chomsky sobre sintaxe com uma capa vermelha*] como $N'(2)$ (composto de [$N'(1) + PP$]), e assim sucessivamente.

A solução brevemente descrita acima nos parece satisfatória em vários sentidos: (a) é implementável com DCGs; (b) permite tratar algumas categorias como recursivas,

exatamente como requer a teoria X-barras e a sintaxe das línguas naturais, em geral; (c) é compatível com outros aspectos da teoria X-barras (por exemplo, permite codificar a diferença entre categorias recursivas e não-recursivas); (d) tem fundamento lingüístico (é uma implementação de certas concepções do componente recursivo da estrutura sintagmática); e, finalmente, (e) não incorre no artificialismo e na falta de economia das soluções que recorrem aos “símbolos posições”.

Para comparação, considere-se brevemente a gramática em (1), implementada na versão original do Grammar Play para a recursão à esquerda de adjetivos (isto é, para estruturas como [*aparelhos eletrônicos multifuncionais contemporâneos*]), em face da solução em (2), adotada no protótipo desenvolvido (as regras em (1) e (2) foram adaptadas para comparação, com informação desnecessária eliminada):

(1) N' no Grammar Play original (“falsa recursão” com símbolos posições)

- (a) $n_bar([N]) \rightarrow n(N)$.
- (b) $n_bar([N_bar, SAdj]) \rightarrow n_bar(N_bar), sadj(SAdj)$.
- (c) $n_bar([X, SAdj]) \rightarrow x(X), sadj(SAdj)$.
- (d) $x([N]) \rightarrow n(N)$.
- (e) $x([Y, SAdj]) \rightarrow y(Y), sadj(SAdj)$.
- (f) $y([N]) \rightarrow n(N)$.
- (g) $y([Z, SAdj]) \rightarrow z(Z), sadj(SAdj)$.
- (h) $z([N]) \rightarrow n(N)$.

(2) N' em programa-protótipo (recursão com subíndice numérico)

- (a) $n_bar([N], 0) \rightarrow n(N)$.
- (b) $n_bar([N_bar, SAdj], I) \rightarrow n_bar(N_bar, J), sadj(SAdj), \quad \{I = J + 1\}$.

Note-se que, salvo pelos índices numéricos (0 em (2a), e I e J em (2b)), a solução em (2) expressa a gramática da recursão dos adjetivos basicamente como as regras que se encontram em qualquer manual de sintaxe X-barras: (a) $N' \rightarrow N$; e (b) $N' \rightarrow N' SAdj$. Por outro lado, a gramática em (1) faz uso de 8 regras para obter, no máximo, SNs com três adjetivos modificadores – não há recursão de fato, apenas um “simulacro”.

2.3 Ampliações na cobertura sintática do parser

Em face da concentração no problema da recursão-à-esquerda e nas melhorias da interface gráfica, foi pequeno o progresso em outras áreas da implementação da gramática do Grammar Play. Em particular, praticamente não houve extensão no poder de análise sintática do Grammar Play (para além da criação de instrumento de implementação da recursão): nenhuma das pretendidas extensões relativas a estruturas oracionais mais complexas foi desenvolvida.

Por outro lado e como compensação, deve-se mencionar o início da implementação de um recurso previsto para uma outra fase do projeto: o de um componente responsável pela flexão de verbos. Conforme esclarecido no projeto, a versão original do Grammar Play só era capaz de analisar frases simples, isto é, com uma só forma verbal. Além disso, esta forma verbal devia ser necessariamente uma forma de 3a. pessoa do singular do presente do indicativo, que era a forma de entrada lexical dos verbos no Grammar Play. Essa limitação será em breve superada, já que está sendo desenvolvido um componente gramatical que analisa as formas verbais encontradas no input e determina, antes de executar a análise sintática do input, se são formas flexionais de algum dos verbos registrados no léxico do Grammar Play. Assim, espera-se que nas próximas versões do Grammar Play já seja possível submeter frases com qualquer forma verbal.

Finalmente, mencionamos ainda, com relação à implementação das regras sintáticas, algumas outras pequenas melhorias resultantes de uma revisão da gramática encontrada na versão original do Grammar Play:

- Algumas regras foram revistas. Havia regras que não diziam respeito a frases simples (como é a proposta original do Grammar Play), como esta, para frases de situação (e.g. *Fogo! Socorro!*): **S** --> **SN**. Essa e outras duas regras equivocadas foram eliminadas.
- Foi eliminado o sistema de “gramáticas alternativas”, que possibilitava, de certa forma, “corrigir análises erradas” e apresentar mais de uma análise para sentenças ambíguas (ver Othero 2004 para maiores detalhes). Eliminando este sistema, o Grammar Play ficou mais leve e menos dependente do usuário. Entretanto, o mesmo recurso era utilizado para buscar uma análise gramatical alternativa para frases-input; sem o recurso, o Grammar Play não pode mais oferecer análises alternativas para frases ambíguas, por exemplo – um problema que deverá ser corrigido nas próximas versões do programa.

2.4 Melhorias no léxico e incorporação de componente semântico

Também as ampliações no tratamento do léxico e o desenvolvimento de um componente semântico ficaram aquém do que foi planejado. Especificamente, com relação à incorporação, nas entradas lexicais, de informação sobre a estrutura de argumentos e sobre outras restrições seletivas, o que se concretizou foi uma representação *standard* para a estrutura de argumentos dos verbos, apenas. Por “representação *standard*” entende-se aqui aquela em que se informam os papéis semânticos exigidos pelos verbos, bem como o “modo de associação” com os argumentos sintáticos do verbo (isto é, com seu sujeito, objeto, etc.).

Para os rótulos dos papéis semânticos, adotou-se simplesmente a lista de papéis encontrada na maioria dos manuais de sintaxe (Radford 1988, Raposo 1992, Haegeman 1994, etc.). Por exemplo, um verbo inacusativo como *cair* recebeu em sua entrada lexical uma representação para o papel semântico de seu único argumento – o papel de Tema; já o verbo inergativo *caminhar* recebeu uma representação para o papel semântico de Agente; e o verbo transitivo *matar*, para os papéis de Agente e Tema.

Para a associação entre papéis temáticos de um verbo e seus argumentos manifestos, adotou-se um procedimento inspirado na teoria “projetiva” proposta por Higginbotham (1985) e Speas (1986). Especificamente, os papéis semânticos são codificados como uma “hierarquia”, por meio da ordem linear esquerda-direita, na entrada lexical; essa hierarquia corresponde, grosseiramente, à hierarquia das funções gramaticais sujeito<objeto direto<objeto indireto<outros. Por exemplo, na entrada de *cair*, representa-se a hierarquia (Tema); como o primeiro papel temático corresponde à primeira função, isto é, ao sujeito, o sistema identificará o sujeito de uma oração com o verbo *cair* como sendo um Tema. Com *matar*, codifica-se a hierarquia (Agente, Tema) correspondente a Sujeito-Objeto; e com o verbo *receber*, a hierarquia (Meta, Tema, Fonte) correspondente a Sujeito-Objeto-Objeto Indireto, etc.

A associação propriamente dita entre essa informação e os argumentos sintáticos do verbo é obtida a partir da análise sintática executada pelas regras de reescrita. Com base na estrutura sintagmática da frase, é possível identificar: (a) qual NP é o sujeito ou o objeto da frase e (b) qual é o verbo predicador e, portanto, qual sua estrutura de argumentos. As duas informações podem, então, ser combinadas e, assim, obtém-se o resultado da atribuição temática de uma frase. Em (3) abaixo apresenta-se um pequeno

trecho do programa-protótipo onde este sistema foi inicialmente testado (trecho correspondente à atribuição de papel temático de sujeito com verbos “intransitivos”):

(3) Atribuição temática em programa-protótipo

% Identificação das Funções Gramáticas Básicas:

subject(NP,S) :- s(s(NP,_),S,[]).

predicator(V,S) :- s(s(NP,VP),S,[]), np(NP,S,X), vp(VP,X,[]), vp(vp(V),X,[]).

% Regras de atribuição temática:

role_subject(NP,Role,S) :- subject(NP,S), predicator(V,S), lex(_,_ ,Role,_ ,_).

% Léxico:

lex(walks, v, agent,_ ,_).

lex(arrives, v, theme,_ ,_).

Nesse sistema, as “regras de atribuição temática” permitem interrogar, para frases como “John walks” ou “John arrives”, qual é o seu sujeito e qual é o papel temático dele (em Prolog: “?- role_subject(NP, Role, [John, walks].” ou “?- role_subject(NP, Role, [John, arrives].”), e a resposta será “John” e “Agente” ou “Tema”, respectivamente (em Prolog: “NP = John” e “Role = agent”, ou “NP = John” e “Role = theme”).

Evidentemente, dispondo dessa informação, é possível implementar em Prolog um novo sistema em que o output será uma “representação da estrutura semântica básica” da frase – a de suas relações de predicado-argumento. Por exemplo, uma vez que o sistema acima é capaz de indicar, para a frase “John walks”, seu predicator, o sujeito e seu papel temático, pode-se facilmente imaginar meios de juntar estas informações numa representação do tipo “existe(E) & evento(E, walking) & agente_de(E, john)” – isto é, numa “forma lógica” bem próxima à representação de eventos encontrada em Carlson (1990). Um dos objetivos da proposta submetida ao CNPq era a de implementar esta representação no Grammar Play, mas este objetivo não foi atingido, tendo-se chegado à etapa preliminar acima descrita.

No entanto, outras melhorias relativas ao léxico do Grammar Play também foram realizadas:

- A maioria dos itens lexicais da gramática foi revista. Havia alguns que tinham sido classificados erroneamente (substantivos masculinos que estavam

classificados como substantivos femininos, adjetivos que estavam classificados como substantivos, etc.). Foi feita uma primeira revisão para eliminar os erros.

- O arquivo original com a gramática do Grammar Play foi dividido em dois: um contendo os itens lexicais e outro, as regras sintagmáticas e de inserção lexical. Dessa maneira, ficou mais fácil de consultar a gramática e trabalhar com o dicionário de itens lexicais.
- Foram eliminadas a distinção entre os advérbios (de tempo, modo, intensidade...) e as várias regras sintagmáticas pertinentes, em face da inadequação do tratamento então oferecido pelo Grammar Play. Parece mais adequado, no momento, utilizar um conjunto único de regras sintagmáticas para advérbios, sendo que as restrições distribucionais devem interagir com a informação lexical – nos moldes das regras de atribuição temática. Entretanto, esta parte da gramática sequer foi experimentada em protótipos.

3. Prospectos Imediatos

Há, evidentemente, várias tarefas e frentes de pesquisa a serem enfrentadas no desenvolvimento do Grammar Play, mas pode-se enumerar as seguintes como as prioritárias para o futuro imediato:

- completar a implementação do sistema de recursão numérica, do sistema de atribuição de papéis temáticos e do sistema de flexão verbal, com a adaptação e incorporação definitiva dos protótipos desenvolvidos no período de outubro de 2005 a outubro de 2006;
- revisar o *parser* sintático, com a inclusão de regras sintagmáticas mais apropriadas para advérbios (dando conta das diferentes distribuições destes) e sintagmas preposicionais (eliminando o sistema *ad hoc* adotado de Pagani 2004), bem como ampliar o *parser* sintático, com a incorporação das várias estruturas oracionais inicialmente previstas para 2005-2006.

4. Produção Intelectual

São resultados diretamente relacionados ao desenvolvimento do projeto de pesquisa financiado pelo CNPq os seguintes trabalhos:

4.1 Livros

OTHERO, G. A. . "Teoria X-barras: descrição do português e aplicação computacional". São Paulo: Contexto, 2006. v. 1. 160 p.

4.2 Artigos, textos em anais de congressos, etc.

MENUZZI, S. M. Prefácio do livro "Teoria X-barras: descrição do português e aplicação computacional". São Paulo: Contexto, 2006.

OTHERO, G. A. . Sintaxe X-barras: uma aplicação computacional. In: 7o. Círculo de Estudos Lingüísticos do Sul - CELSUL, 2006, Pelotas. 7o Círculo de Estudos Lingüísticos do Sul - Resumos, 2006.

OTHERO, G. A. ; MENUZZI, S. M. . Sintaxe X-barras - uma aplicação computacional. Working Papers em Lingüística, PPL/UFSC, Florianópolis, 2006. (no prelo)

4.3 Apresentações em seminários, congressos, etc.

OTHERO, G. A. "Sintaxe computacional: uma aplicação em Prolog para um parser em língua portuguesa". I Seminário do NES - Núcleo de Estudos em Semântica Lexical, Universidade Federal de Santa Catarina – UFSC, Florianópolis, 2006. (Participações em eventos/Seminário).

OTHERO, G. A. "Sintaxe X-barras: uma aplicação computacional". I Colóquio de Lingüística dos Alunos de Pós-Graduação da PUCRS, Porto Alegre, 2006.

OTHERO, G. A., CONTERATTO, G. B. H. "O léxico na interface sintaxe\semântica: perspectivas e limitações computacionais". VI Semana de Letras, FALE/PUCRS, Porto Alegre, 2006. (Participações em eventos/Outra).

OTHERO, G. A. "Sintaxe X-barras: uma aplicação computacional". 7o. Círculo de Estudos Lingüísticos do Sul – CELSUL, UCPEL, Pelotas RS. 2006. (Participações em eventos/Outra).

5. Resumo do Relatório Financeiro

O orçamento original do projeto, constante na proposta submetida ao CNPq, era de R\$ 17.000,00 (inferior ao limite de R\$ 20.000,00 do Edital CNPq no. 32/2004). O valor aprovado e concedido pelo CNPq foi de R\$ 5.000,00, os quais foram despendidos como segue:

DESPESA	VALOR (R\$)
- compra de equipamento em 18/11/2005: 01 computador laptop DELL Latitude 110L, com garantia de 1 ano	3.422,71
- compra de equipamento em 30/12/2005: 01 Pen drive Kingston 512MB	240,00
- pagamento de serviços, período de março a outubro de 2006: assessoria técnica em programação Delphi (para a interface gráfica)	1.200,00
TOTAL:	4.862,71
A DEVOLVER AO CNPq:	137,29

6. Bibliografia

- BLACKBURN, P. et al. *Learn PROLOG Now!* Londres: College Publications, 2006.
- BOUMA, G. & VILLADA, B. Shift Reduce Parsing. Hand-out, 2002/2003. Disponível em http://www.let.rug.nl/~gosse/nlp1/shift_reduce1.pdf
- CHOMSKY, N. Remarks on Nominalization. In R. A. JACOBS & P. S. ROSENBAUM, orgs. *Readings in English Transformational Grammar*. Waltham, Mass: Ginn & Co, 1970.
- CHOMSKY, N. Bare Phrase Structure. In G. WEBELHUTH, org. *Government and Binding and the Minimalist Program*. Oxford & Cambridge USA: Blackwell, 1995.
- COVINGTON M. A. *Natural Language Processing for Prolog Programmers*. New Jersey: Prentice Hall, 1994.
- GAZDAR, G. & MELLISH, C. *Natural Language Processing in Prolog: an Introduction to Computational Linguistics*. Addison-Wesley: Workingham, 1989.
- HAEGEMAN, L. *Introduction to Government and Binding Theory*. 2 ed. Oxford: Blackwell, 1994.
- HIGGINBOTHAM, J. On Semantics. *Linguistic Inquiry* 16, 547-93, 1985.
- OTHERO, G. de Á. *Grammar Play: um Parser Sintático em Prolog para a Língua Portuguesa*. Porto Alegre: PUCRS. Dissertação de Mestrado, 2004.
- OTHERO, G. de Á. *Teoria X-Barra: Descrição do Português e Aplicação Computacional*. São Paulo: Contexto, 2006.
- PARSONS, T. *Events in the Semantics of English*. Cambridge: The MIT Press, 1990.
- PEREIRA, F. C. N. & SHIEBER, S. M. *Prolog and Natural Language Analysis*. CSLI Lecture Notes 10. Stanford: Chicago University Press, 1987.
- RADFORD, A. *Transformational Syntax*. Cambridge (Inglaterra): Cambridge University Press, 1981.
- RADFORD, A. *Transformational Grammar*. Cambridge (Inglaterra): Cambridge University Press, 1988.
- RAPOSO, E. *Teoria da Gramática: a Faculdade da Linguagem*. Lisboa: Ed. Caminho, 1992.
- SPEAS, M. Adjunctions and projections in syntax. Tese de doutorado. Cambridge, Mass.: MIT, 1996.

- STOWELL, T. *Origins of Phrase Structure*. Tese de doutorado. Cambridge, Mass.: MIT, 1981.
- STUURMAN, F. *Phrase structure theory in generative grammar*. Dordrecht: Foris, 1985.