

Resenha de *Lexikalisch-funktionale Grammatik. Eine Einführung am Beispiel des Französischen mit computerlinguistischer Implementierung*

Tito Lívio Cruz Romão¹

Francisco G. dos Santos Nogueira²

Programa de Pós-Graduação em Letras, Universidade Federal do Ceará, Fortaleza, CE, Brasil

Resenha da obra:

SCHWARZE, Christoph; ALENCAR, Leonel F. de. **Lexikalisch-funktionale Grammatik. Eine Einführung am Beispiel des Französischen mit computerlinguistischer Implementierung**. 1. ed. Tübingen: Stauffenburg, 2016. v. 1. 278p.

No ano de 2016, os pesquisadores Dr. Christoph Schwarze (professor emérito da Universidade de Constança/Alemanha) e Dr. Leonel F. de Alencar (professor associado da Universidade Federal do Ceará) publicaram pela editora Stauffenburg (Tübingen, Alemanha) o livro *Lexikalisch-funktionale Grammatik. Eine Einführung am Beispiel des Französischen mit computerlinguistischer Implementierung* [tradução livre: Gramática léxico-funcional. Uma introdução baseada na língua francesa com implementação linguístico-computacional]. De início, cumpre destacar que, no contexto dos diversos modelos gramaticais em que atualmente se insere o mais elevado nível de desenvolvimento das pesquisas em Linguística, a LFG, como vertente da gramática gerativa, desponta como um modelo que permite implementação computacional, possibilitando simultaneamente uma rigorosa revisão das análises gramaticais. Trata-se de uma ferramenta, como muito bem ilustram os autores do livro em epígrafe, a que têm recorrido, por exemplo, estudantes, teóricos e pesquisadores das áreas de Informática, Processamento de Linguagem Natural (PLN), Linguística Teórica e Linguística Computacional.

É igualmente digno de menção que, se os estudos visando a uma implementação de gramáticas no âmbito do PLN para línguas como o alemão e o inglês já se encontra bastante avançado, as pesquisas específicas voltadas para o português ainda estão caminhando a

¹ Doutor (UFSC), Professor: UFC. Orcid: <https://orcid.org/0000-0002-3195-3600>

E-mail: cruzromao@terra.com.br

² Mestre (UFC), doutorando: UFC. Orcid: <https://orcid.org/0000-0002-7896-9140>

E-mail: herrnogueira@gmail.com

passos comparativamente mais lentos. Entendemos que essa obra de cunho teórico, técnico e acadêmico, elaborada e publicada por Schwarze & Alencar (2016), justamente por trabalhar com a implementação de uma língua românica, o francês, poderá inspirar o surgimento de diversos trabalhos voltados para a implementação de gramáticas automáticas do português. Como ilustração, na Universidade Federal do Ceará (UFC) esse modelo de investigação teórico-prática já vem sendo aplicado por Alencar na orientação de mestrandos e doutorandos no contexto do Programa de Pós-Graduação em Linguística (PPGL) e também na orientação de bolsistas do Programa Institucional de Bolsas de Iniciação Científica.

Corroborando as palavras dos próprios autores, a publicação desse livro tem como um de seus objetivos primordiais o de conectar a descrição e a análise sintática a um sistema automático que permita testar a modelagem da LFG quanto à sua consistência e à sua adequação empírica. Para alcançar esse objetivo, Schwarze & Alencar lançam mão do *Xerox Linguistic Environment (XLE)*, um *software* desenvolvido pela empresa Xerox na década de 1990, que, como atestam os autores, já se mostrou eficaz na análise e na geração de sentenças com base na LFG, representando o estado da arte atual no desenvolvimento e *parsing* de gramáticas no formalismo da LFG. Para a consecução de sua pesquisa e publicação do livro, os autores fundamentaram suas hipóteses e alcançaram seus objetivos apoiados em autores como Berman & Frank (1996), Butt *et al.* (1999), Bresnan (2001), Falk (2001), Beesley & Karttunen (2003), dentre outros.

O livro ora analisado obedece a uma divisão em oito capítulos que na verdade são chamados de lições. No prefácio, os dois autores já esclarecem que sua opção por essa segmentação do livro em lições foi uma escolha de cunho meramente didático. Considerando-se a própria natureza de uma implementação gramatical para PLN, o trabalho de Schwarze & Alencar precisou restringir-se a um fragmento gramatical, dando ênfase, por exemplo, às seguintes propriedades tipicamente românicas da língua francesa: o pronome sujeito clítico, a sintaxe do adjetivo, a voz passiva e o *passé composé*. Seguindo a mesma orientação didática do recorte gramatical, também foi preciso limitar o léxico abordado a valências e significados canônicos, para não extrapolar os limites técnicos ensejados pelo fragmento gramatical proposto. Ressalte-se, ainda, que no que tange à morfologia flexional, recorreu-se, em larga escala, a um léxico de formas plenas (*full-form lexicon*). Ademais, nos fragmentos de XLE também foram levados em consideração, com algumas simplificações, conhecimentos oriundos do projeto *Parallel Grammar (ParGram)*, que tem como foco, a partir de diferentes famílias linguísticas, o desenvolvimento de gramáticas LFG comparáveis desenvolvidas em XLE (SCHWARZE & ALENCAR, 2016, p. V).

Retomando conceitos da linguística chomskyana e da informática teórica, os autores recordam que nesses dois campos de estudo a gramática é caracterizada como um aparato (*device*) destinado à produção de sentenças e/ou sequência de elementos (*strings*) e que, além disso, a tarefa de reconhecimento e/ou a análise dessas componentes são confiadas a um autômato ou máquina automática. Não obstante, Schwarze & Alencar (2016) ressaltam que, segundo sua concepção, uma gramática não é nem um *parser* (analisador sintático

automático) nem um gerador automático, ou seja: uma gramática não gera nenhum algoritmo procedural visando ao reconhecimento e à análise e/ou à produção de sentenças; em vez disso, constitui uma estrutura declarativa de dados da qual diferentes programas, tais como *parsers* e geradores, podem fazer uso. Ressaltam ainda que um *parser*, em sua função de programa procedural que executa uma série finita de instruções segundo um determinado algoritmo, assume a tarefa de verificar se é possível produzir um determinado *input* conforme as regras e as entradas lexicais da gramática; ademais, se esse *input* revelar-se gramatical, o *parser* também tem a tarefa de determinar que estrutura, conforme essa gramática, deverá ser atribuída a tal *input*. Se uma determinada gramática apresentar o formato exigido, automaticamente o XLE a transforma em um *parser* que poderá ser aplicado à análise de sentenças. Os autores ainda lembram que, se a gramática de algum modo contrariar as rigorosas especificações de formato, o processamento será interrompido, e uma mensagem de erro será enviada. Pode-se então concluir que o XLE não logrará elaborar um *parser* para uma gramática que apresente esse tipo de defeito. A partir de uma gramática codificada de forma correta, o XLE também será capaz de produzir um gerador automático.

No tocante às análises do francês e à apresentação do modelo da LFG, o livro segue um trabalho preliminar realizado por Schwarze (1996), que, para esse novo propósito, passou por uma profunda revisão. Por outro lado, a implementação dos fragmentos gramaticais e das componentes de análise morfológica foi submetida a uma completa reelaboração a cargo de Alencar. Por entendermos que esse trabalho é de fundamental importância como estímulo para que outros pesquisadores derivem trabalhos semelhantes atinentes a outras línguas neolatinas – neste caso, especialmente ao português do Brasil –, recomendamos uma leitura dessa obra que certamente vale como referência para futuras investigações no campo do desenvolvimento de fragmentos de gramáticas segundo a LFG.

Conforme a metodologia adotada no livro, entre as lições 1 e 6, os autores desenvolvem uma gramática LFG visando a cobrir diversas áreas centrais da língua francesa; para alcançar seu objetivo, procedem à criação de fragmentos gramaticais, também chamados de minigramáticas. Mediante esse procedimento, os aprendizes que fizerem uso desse método poderão expandir autonomamente as minigramáticas iniciadas. Outro detalhe relevante a destacar na abordagem adotada no livro consiste em que é apresentada, sempre no final de cada capítulo-lição, uma lista de bibliografia especializada relativa às temáticas ali abordadas de maneira particular. A seguir, apresentaremos, de forma resumida, um panorama dos conteúdos de cada uma das lições, acompanhados, sempre que necessário, de algum comentário.

Na Lição 1, o modelo gramatical da LFG é apresentado: a) como um instrumento heurístico que estimula a observação exata de fatos sintáticos e morfológicos; b) como um método usado para apresentar e avaliar enunciados sobre fatos linguísticos e sua contextualização; c) como um modelo explícito, pois permite que gramáticas LFG sejam implementadas e examinadas no computador com a ajuda do sistema XLE, que gera um

parser a partir de uma gramática LFG. Conforme as explicações, o emprego tanto dessa ferramenta quanto de convenções notacionais compatíveis permite que muitas gramáticas descritas no *ParGram* possam ser cotejadas de modo pormenorizado. Para o tratamento reservado à flexão e à composição de palavras, recorre-se a um *software* adicional, o *Xerox Finite State Tools* (XFST). A primeira lição também aborda os seguintes tópicos: a) as origens teóricas da LFG: a gramática de estrutura sintagmática (PSG), a teoria das valências e a teoria das funções gramaticais; b) a implementação de uma minigramática no XLE; e c) bibliografia especializada em análise de constituintes, gramática da dependência e valência, PSG, *parsing*, XLE, *ParGram* e morfologia de estados finitos. Ao abordar a PSG, retoma-se a concepção – já presente na gramática tradicional – de que a gramática de estrutura sintagmática é uma concretização da ideia de que uma sentença não tão-somente é uma estrutura linear, como também possui uma estrutura hierárquica. Utiliza-se como exemplo esta sentença francesa: *Le chevalier cherche la fée* ('o cavaleiro procura a fada'). Em seguida, apresenta-se a estrutura sintagmática dessa sentença como diagrama arbóreo, frisando-se que os constituintes da estrutura sintagmática formam nós. Numa fase posterior, aborda-se, de forma resumida, a teoria das valências, relacionando-a à LFG e ressaltando-se as correspondências entre os termos originalmente franceses e seus correspondentes na LFG, tais como *actants* = funções regidas e *circumstants* = adjuntos. Ainda no campo da valência, destaca-se que a LFG é uma gramática de unificação: cada traço é um atributo com seu valor. A unificação exige que os valores de um atributo dado que se encontram num sintagma não se contradigam. Desse modo, à guisa de exemplo, a congruência dentro do sintagma nominal (NP) do francês consiste em que os valores de gênero e número de artigos, adjetivos e substantivos precisam ser compatíveis. Sem recorrer a regras, o princípio da unificação possibilita que sejam reconhecidas e refutadas violações contra as leis fundamentais de concordância, como neste exemplo: **les belle fées* (tradução literal: **as bela fadas*). Quanto à implementação da PSG em XLE, são apresentadas as informações necessárias para a configuração de uma gramática nesse sistema, a saber: título da seção, categoria-raiz da gramática, entradas lexicais, regras de estrutura sintagmática e codificação de caracteres (padrão internacional utilizado: utf-8).

Na Lição 2, Schwarze & Alencar abordam os seguintes tópicos: a) concordância gramatical; b) estrutura de constituintes (estrutura C) e estrutura funcional (estrutura F); c) integralidade e coerência; d) codificação da informação funcional: léxico e anotações; e) implementação no XLE; f) indicação de bibliografia especializada. De maneira direta, essa lição ressalta a importância do tratamento a ser dado à concordância entre sujeito e predicado, tendo-se em consideração que o sujeito da sentença francesa concorda em número e pessoa com o verbo finito e em gênero e número com o adjetivo predicativo. Todavia, nessa lição somente é abordada a concordância entre o sujeito e o verbo. Ao discorrer sobre a estrutura C e a estrutura F, os autores explicitam que a primeira nada mais é que a estrutura que se pode produzir com uma PSG, e que a segunda estrutura é projetada a partir da primeira. Para fazer jus às exigências da unificação, as estruturas F precisam respeitar duas condições: integralidade e coerência. Quanto à codificação das informações

contidas na estrutura funcional, esta dispõe de duas fontes: a primeira são as indicações dos traços gramaticais no léxico e a segunda, as anotações nas regras da estrutura C. Tais anotações prestam informações sobre a estrutura funcional da sentença, abrangendo as funções gramaticais dos constituintes (sujeito, objeto etc.). Sobre a implementação no XLE, Schwarze & Alencar evidenciam nessa lição que para testar uma gramática em XLE se aconselha o desenvolvimento de um pacote de testes formado de duas componentes parciais: um conjunto de sentenças que a gramática deverá produzir, ou seja, um pacote de testes positivos, e um conjunto de sentenças não-gramaticais que a gramática não deverá produzir, isto é, um pacote de testes negativos. No final da lição, os autores recomendam uma seleção de bibliografia especializada sobre pronomes clíticos e a noção de núcleo na LFG.

A Lição 3 é iniciada com a apresentação de sete exemplos de sentenças francesas, a partir das quais são desenvolvidas as seguintes etapas práticas: a) esquemas estruturais; b) regras de constituintes para um fragmento expandido do francês; c) equações e setas; d) implementação no XLE; e) indicação de bibliografia especializada sobre valência preposicional, pronomes expletivos e verbos impessoais, análise de adjetivos monopredicacionais em função atributiva e predicativa, contração de preposições e determinantes em francês, sintaxe formal e semântica de sintagmas preposicionais etc. Ao abordarem os esquemas estruturais, a intenção dos autores é que a gramática ampliada produza as frases em questão, atribuindo-lhes estruturas C e F. Para tanto, faz-se mister a introdução das respectivas regras e entradas lexicais, o que é feito de forma gradativa. Em seguida, elabora-se uma ampliação das regras anteriormente formuladas na Lição 2, em que agora surge a categoria AP (sintagma adjetival). Em seguida, os autores detêm-se sobre a necessidade de explicarem o uso de símbolos notacionais comuns em XLE, tais como “^” e “!” , que na notação-padrão do formalismo da LFG são escritos como seta ascendente (\uparrow) e descendente (\downarrow), respectivamente, assumindo sua função matemática de apontar para traços específicos. No final da lição, as indicações bibliográficas concernem às seguintes temáticas: preposições com casos, distinção entre preposições lexicais e não-lexicais, pronomes expletivos e verbos impessoais, análise de adjetivos como predicados monovalentes em função atributiva e predicativa, dentre outras.

Iniciando com exemplos práticos da língua francesa, a Lição 4 propõe-se a expandir o fragmento de gramática abordando os seguintes temas: a) sintagmas adjetivais; b) sentenças complementares com *que* e *si*; c) sintagmas preposicionais como adjuntos do sintagma verbal (VP); d) novas entradas lexicais; e) novas regras; f) implementação no XLE; e) indicação de bibliografia especializada. Nessa lição, primeiramente é abordada a questão do uso predicativo ou adnominal de adjetivos na língua francesa. Essa distinção é importante para o desenvolvimento da estrutura C correspondente a cada situação envolvendo o uso de adjetivos dessa ou daquela forma. Em seguida, são abordadas as chamadas sentenças com CP, os sintagmas complementadores, ou seja, aquelas que necessitam de uma conjunção que as conecte à oração principal. Para a exemplificação no desenvolvimento do fragmento de

gramática francesa, os autores restringiram-se, por questão de simplificação da modelagem, às conjunções integrantes *que* e *si*. Para alimentar a ampliação da gramática, também são abordados sintagmas preposicionais com preposições com valor semântico empregados como complementos predicativos (XCOMP) de verbos copulativos. No tocante às novas entradas lexicais, nessa lição é apresentada uma lista de verbos transitivos facultativos, verbos com complemento, além de uma série de substantivos, adjetivos, advérbios de grau etc. Na sequência, é explanada a implementação no XLE dos fenômenos tratados na lição e, por fim, apresentada uma lista de títulos bibliográficos recomendados sobre os seguintes temas: preposições em francês, modelagem na LFG da distribuição sintática de adjetivos franceses visando à implementação no XLE, distinção entre homonímia e polissemia, utilização de padrões na codificação de generalizações linguísticas, dicionários eletrônicos da língua francesa, dentre outros.

Propondo-se a expandir ainda mais o fragmento de gramática do francês almejado em sua obra, Schwarze & Alencar iniciam a Lição 5 voltando sua atenção para os chamados *verbos de controle*, ou seja, aqueles verbos que regem uma construção com um infinitivo. Em seguida é apresentada uma série de sentenças em francês em que surgem verbos de controle, com o intuito de processá-las a partir da gramática automática desenvolvida pelos autores. Na sequência, são abordadas as seguintes etapas: a) a estrutura C das construções com verbos de controle; b) a estrutura F e a representação das relações de controle; c) entradas lexicais (lista de alguns verbos de controle, tais como: *devoir* ‘dever’, *pouvoir* ‘poder’, *savoir* ‘saber’, *vouloir* ‘querer’, *écouter* ‘escutar’, *entendre* ‘ouvir’, *voir* ‘ver’, *cesser* ‘cessar’, *commencer* ‘começar’, *continuer* ‘continuar’, *croire* ‘crer’ etc.); d) regras destinadas ao registro das novas estruturas C e às expansões do sintagma verbal (VP); e) implementação no XLE; f) indicação de bibliografia especializada sobre diferentes temas, tais como: controle na LFG, tokenização de forma genérica, tokenizar com XFST, abrigar a tokenização na arquitetura XLE etc.

Em sua tarefa de expandir mais e mais os fragmentos da gramática proposta para o francês, os autores abordam na Lição 6 a forma passiva e o *passé composé*, a partir de nove sentenças francesas apresentadas no início do capítulo-lição. Na parte destinada à voz passiva, é tratada a homonímia do verbo *être*, mais precisamente: *être* como verbo cópula *versus être* como verbo auxiliar na voz passiva. A discussão também envolve o complemento-agente regido pela preposição *par*. Em seguida, são analisadas as estruturas C e F de algumas sentenças francesas na voz passiva. Ademais, são apresentadas as novas entradas lexicais, abrangendo alguns participios passivos, ressaltando-se que a gramática se restringe às formas que surgem nas sentenças apresentadas no início da Lição 6. No próximo passo, são abordados os seguintes tópicos: o *passé composé*, o participio passado, a escolha do verbo auxiliar, e, por fim, a concordância entre o verbo *être* e o sujeito, e entre o objeto e o verbo *avoir*. Como ocorre em cada lição, é apresentada a implementação no XLE e, por fim, sugerida uma seleção bibliográfica especializada, englobando estes temas: adjetivo *versus* voz passiva, modularização gramatical e demais questões envolvendo o desenvolvimento de

gramáticas a partir de um *software*, implementação da regra da voz passiva em XLE, dentre outros.

Abre-se a Lição 7 com algumas considerações importantes: a) até então o fragmento gramatical abordara apenas uma pequena fração do sistema de tempos verbais franceses; b) o léxico restringira-se, no tocante aos tempos simples, somente ao presente do indicativo e do subjuntivo, bem como à codificação, em muitos casos, apenas da terceira pessoa do singular; e, por último, c) na codificação do léxico, os autores em geral se limitaram àqueles verbos necessários à realização do *parsing* das sentenças-exemplo provenientes dos conjuntos de testes positivos e negativos. Segundo Schwarze & Alencar, uma expansão do léxico de verbos, no nível em que se encontra a gramática, representaria um procedimento muito moroso, pois, para praticamente cada uma das formas do paradigma, isso demandaria a criação de uma nova entrada. Na Lição 7, ainda anunciam uma expansão do fragmento incluindo as diferentes formas simples do indicativo e do subjuntivo de um subgrupo de verbos com a terminação do infinitivo em *-er*, tais como: *arriver* ‘chegar’, *acheter* ‘comprar’, *compléter* ‘completar’, *changer* ‘mudar’, *avancer* ‘avançar’, dentre outros. Observe-se que a lista inclui verbos inteiramente regulares (p. ex.: *arriver*), mas também há alguns que apresentam alternância fonológica no radical ao longo da conjugação (*compléter*, *je complète* ‘eu completo’, *nous complétons* ‘nós completamos’) e outros com alternâncias de ordem estritamente ortográfica, mas não fonológica (*nous changeons* ‘nós mudamos’, *vous changez* ‘vós mudais’). Para realizar o *parsing* de sentenças que contenham esses verbos, os autores adiantam que não recorrerão a entradas de formas plenas, e sim à implementação de um mecanismo, no âmbito da morfologia de estados finitos, que permita a completa análise e geração automática de todas as formas dos tempos simples selecionados e das formas do infinitivo e do particípio dos respectivos verbos. O mecanismo utilizado é um *transdutor lexical*. Na etapa seguinte, são tratados os seguintes temas: a) considerações sobre a morfologia; b) modelagem de um fragmento da conjugação francesa na morfologia de estados finitos; c) indicação de bibliografia especializada sobre: alternâncias de radicais verbais no francês, línguas formais e teoria dos autômatos, morfologia de estados finitos etc.

Após a implementação de um transdutor lexical na lição precedente, aplicado a um excerto da conjugação verbal francesa, a Lição 8, correspondente ao último capítulo do livro, tem como objetivo mostrar como esse transdutor pode ser acoplado à gramática XLE do francês objeto da pesquisa de Schwarze & Alencar, visando a executar o *parsing* de alguns tipos de sentenças na voz ativa e passiva e com verbos em um dos tempos simples do indicativo e do subjuntivo ou no *passé composé*. De forma separada, são abordados os seguintes tópicos: integração de um transdutor lexical ao fragmento XLE; b) implementação do *passé composé* e reformulação da regra da voz passiva; c) testagem do fragmento de gramática; d) indicação de bibliografia especializada sobre os seguintes temas: integração de um transdutor lexical em uma gramática XLE, implementação da regra da voz passiva, implementação da seleção de verbos auxiliares e concordância do particípio no *passé composé* formado em dependência do auxiliar, dentre outros.

Reiteramos que o estudo apresentado por Schwarze & Alencar (2016) tem o mérito de apresentar, de maneira gradual, um modelo de implementação, no âmbito da LFG e mediante a ferramenta XLE, de fragmentos de uma gramática francesa. Com a apresentação de cada nova etapa procedural em uma nova lição, a metodologia utilizada mostra os passos que vão desde a morfologia flexional, entendida como morfologia de estados finitos, até sua integração em fragmentos sintáticos. Por se tratar de uma abordagem voltada para uma língua românica, identificamos, nessa obra, um grande potencial de conteúdos que podem ser explorados por estudiosos da LFG interessados em aplicar esse método e as respectivas ferramentas à língua portuguesa, em especial ao português do Brasil.

Destaque-se, por fim, que os fragmentos de gramática apresentados nas diferentes lições do livro estão livremente disponíveis para *download* no seguinte endereço: <https://github.com/lfg-french-grammar>. O livro conta também com um blogue, mantido por Alencar, com informações adicionais, inclusive uma lista de correções: <http://lfg-book.blogspot.com/>.

Referências

- BEESLEY, K. R.; KARTTUNEN, L. *Finite State Morphology*. Stanford: CSLI Publications, 2003.
- BRESNAN, J. *Lexical-Functional Syntax*. Oxford: Blackwell, 2001.
- BUTT, M. *et al. A Grammar Writer's Cookbook*. Stanford, California: CSLI Publications, 1999.
- FALK, Y. *Lexical-Functional grammar*. Stanford, California: CSLI Publications, 2001.
- FRANK, A. Eine LFG-Grammatik des Französischen. In: BERMAN, J.; FRANK, A. (Org.). *Deutsche und französische Syntax im Formalismus der LFG*. Tübingen: Niemeyer, 1996. p. 97-244.
- SCHWARZE, C. *Lexikalisch-funktionale Grammatik: Eine Einführung in 10 Lektionen mit französischen Beispielen*. Konstanz: Fachbereich Sprachwissenschaft der Universität Konstanz, 1996.

Recebido em: 30/09/2019.

Aceito em: 30/11/2019.