

A PASSIVA POSSESSIVA EM PORTUGUÊS: UMA ABORDAGEM LEXICALISTA COM UMA IMPLEMENTAÇÃO COMPUTACIONAL

LA PASIVA POSESIVA EN PORTUGUÉS: UN ENFOQUE LEXICALISTA CON UNA
IMPLEMENTACIÓN COMPUTACIONAL

THE POSSESSIVE PASSIVE IN PORTUGUESE: A LEXICALIST APPROACH WITH A
COMPUTATIONAL IMPLEMENTATION

Leonel Figueiredo de Alencar
Universidade Federal do Ceará

RESUMO: A passiva possessiva constitui um dos fenômenos gramaticais mais discutidos das línguas do leste asiático. Em português, embora comum, foi discutida, segundo parece, unicamente por Lunguinho (2011, 2013, 2016), numa abordagem transformacional, utilizando exemplos construídos. No presente artigo, contrapomo-nos a essa abordagem tanto no plano teórico quanto no metodológico. Por um lado, propomos uma análise puramente lexicalista no quadro dos modelos não transformacionais LFG e LDG. Por outro, recorreremos a evidências extraídas de textos autênticos. Essa dupla estratégia permite explicar dados que não se encaixam na proposta de Lunguinho. Argumentamos que as propriedades distintivas dessa passiva resultam de uma regra lexical que estende a estrutura argumental de *ter* pela incorporação de uma variável de predicado à sua forma semântica. Essa variável é vinculada a um argumento predicativo, tornando *ter* um verbo de controle do objeto do tipo *equi*.

PALAVRAS-CHAVE: Passiva possessiva. Passiva não canônica. Gramática Léxico-Funcional (LFG). Decomposição de predicados. Linguística computacional.

RESUMEN: La pasiva posesiva es uno de los fenómenos gramaticales más discutidos en los idiomas de Asia oriental. En portugués, aunque es común, fue discutido, aparentemente, solo por Lunguinho (2011, 2013, 2016), en un enfoque transformacional utilizando ejemplos contruidos. En este artículo, nos oponemos a este enfoque tanto teórica como metodológicamente. Por un lado, proponemos un análisis puramente lexicalista basado en los modelos no transformacionales LFG y LDG. Por otro lado, recurrimos a evidencias extraídas de textos auténticos. Esta doble estrategia nos permite explicar datos que no se ajustan a la propuesta de Lunguinho. Argumentamos que las propiedades distintivas de esta variedad pasiva resultan de una regla léxica que extiende la estructura de argumento del verbo *ter* 'tener' al incorporar una variable predicativa a su forma semántica. Esta variable está vinculada a un argumento predicativo, convirtiendo *ter* en un verbo *equi* de control de objeto.

PALABRAS CLAVE: Pasiva posesiva. Pasiva no canónica. Gramática Léxico-Funcional (LFG). Descomposición de predicados. Linguística computacional.

ABSTRACT: The possessive passive is one of the most discussed grammatical phenomena of East Asian languages. In Portuguese, although common, it has apparently been discussed only by Lunguinho (2011, 2013, 2016) in a transformational approach with constructed examples. In this paper, we oppose this approach both theoretically and methodologically. On the one hand, we propose a purely lexicalist analysis based on the non-transformational LFG and LDG models. On the other, we resort to evidence extracted from authentic texts. This twofold strategy allows us to explain data that do not fit Lunguinho's proposal. We argue that the distinctive properties of this passive variety result from a lexical rule that extends the argument structure of the verb *ter* 'have', by incorporating a predicate variable into its semantic form. This variable is linked to a predicative argument, turning *ter* into an object control *equi* verb.

KEYWORDS: Possessive passive. Non-canonical passive. Lexical-Functional Grammar (LFG). Predicate decomposition. Computational linguistics

1 INTRODUÇÃO

Em *Aurora*, Kim Stanley Robinson conduz o leitor pelas peripécias de uma viagem intergaláctica (ROBINSON, 2015). Um dos aspectos mais interessantes dessa história é que, na maior parte, é narrada pela inteligência artificial (IA) da espaçonave, que assiste aos seus habitantes na resolução dos mais variados problemas, com quem interage em linguagem natural. Para tanto, o sistema tem à disposição todo o conhecimento humano armazenado digitalmente em textos, cujos zettabytes de informações é capaz de extrair, organizar e aplicar em tarefas cognitivas tão complexas como a elaboração de uma narrativa literária.

É incerto se algum dia um sistema de IA poderá, pela autoaprendizagem, alcançar a maestria de um narrador como o próprio Robinson. No entanto, sistemas de resolução de perguntas (Q&A) há algum tempo auxiliam técnicos de manutenção de aviões, interagem com consumidores de empresas de transportes, fornecem informações imobiliárias, turísticas e meteorológicas etc. (KOTSEV, 2010). Também constitui realidade uma das estratégias em que alguns desses sistemas se baseiam: a extração automática de informações (IE) a partir de textos em linguagem natural, que integra dezenas de softwares de aplicação geral ou especializados para determinados domínios e tornou-se indispensável em vários setores de atividade, contribuindo para a análise de grandes conjuntos de dados (*big data*). Talvez o caso de maior sucesso na utilização dessas duas tecnologias seja o Watson, da IBM, que em 2011 venceu dois ex-campeões do *Jeopardy*, programa televisivo norte-americano de perguntas e respostas (WATSON, 2017).

A Q&A e a IE integram a subárea do processamento de linguagem natural (PLN) designada compreensão textual. A IE consiste em converter, em estruturas de dados, informações estruturadas de forma apenas implícita nos textos (MEHLER; LOBIN, 2004, p. 1). Um dos componentes da arquitetura de sistemas de compreensão textual, como o Watson, é um analisador sintático automático

(*parser*), cujas análises constituem entrada para a construção das representações semânticas das sentenças (MCCORD; MURDOCK; BOGURAEV, 2012).

Consideremos os fragmentos a seguir, extraídos de textos reais:

- (1) [...] os torcedores tinham quebrado os vidros do quiosque. (Google)
- (2) Vidraças da sede do banco RBS foram quebradas por manifestantes. (Google)
- (3) Bancos da região central tiveram vidraças quebradas pelos manifestantes. (Google)

Intuitivamente, (1)-(3) expressam, sob diferentes roupagens, um mesmo tipo de situação, modelado na IE por meio de um *template*, estrutura de dados constituída de *slots* referentes aos diferentes papéis semânticos de uma relação. Esses *slots* são preenchidos automaticamente por entidades extraídas de textos em linguagem natural, permitindo que informações codificadas de forma não estruturada alimentem uma ontologia (KIRSTEIN-JOST, 2010), constituam respostas para usuários de sistemas de Q&A etc. Por exemplo, podemos esquematizar as relações semânticas expressas por (1)-(3) por meio do molde DEPREDAÇÃO, abrangendo eventos do tipo *quebrar*, *destruir*, *deprestar* etc. e incluindo os papéis PERP(*retador*), ALVO e POSS(*uido*)R (CHAMBERS; JURAFSKY, 2011). A aplicação desse molde a (1)-(3) revela a estrutura subjacente comum às diferentes configurações superficiais:

- (4) PERP: torcedores, ALVO: vidros, POSSR: quiosque
- (5) PERP: manifestantes, ALVO: vidraças, POSSR: banco RBS
- (6) PERP: manifestantes, ALVO: vidraças, POSSR: bancos da região central

Esse exemplo evidencia a necessidade de as arquiteturas dos sistemas de IE e Q&A levarem em conta os diferentes tipos de construções sintáticas. O sistema de IE proposto por Chambers e Jurafsky (2011), por exemplo, prescinde da elaboração manual de *templates*, os quais extrai de textos e preenche automaticamente. Para tanto, recorre, entre outros recursos, a um *parser*, utilizado para rotular as funções gramaticais das ocorrências dos verbos que compõem os *templates* e converter todas as sentenças passivas nas correspondentes ativas. Desse modo, adotando essa abordagem, a extração de (5) a partir de (2) implicaria transformar essa sentença na correspondente ativa, nos moldes, portanto, de (1). Uma vez que (3) expressa o mesmo tipo de relação que (2), também precisaria passar por transformação análoga.

A sentença (3) exemplifica uma construção do português que, apesar de comum tanto na variedade brasileira quanto na europeia, salvo a abordagem de Lunguinho (2011; 2013; 2016) no quadro do Programa Minimalista (PM), aparentemente, ainda não foi investigada. Pelo contrário, rotulada passiva possessiva, sua contraparte em línguas do leste asiático, onde é pervasiva, tem sido amplamente discutida (HUANG, 1999, p. 51). Em coreano e japonês, que formam a passiva sinteticamente, a mesma forma verbal derivada é utilizada tanto na passiva direta (correspondente a (2)) quanto na possessiva (correspondente a (3)) (OSHIMA, 2004). Em ambas as passivas, o participante realizado como sujeito na ativa é demovido para uma função oblíqua. Na possessiva, porém, o objeto direto do verbo na ativa é preservado, sendo promovido a sujeito um participante que não integra a estrutura argumental de partida. O termo *passiva possessiva* decorre da restrição de que deve subsistir algum tipo de relação de posse entre os referentes do sujeito e do objeto direto nessa construção.

Neste trabalho, analisamos a passiva possessiva do português, tanto em sua dimensão gramatical quanto semântica, a partir de dados reais, extraídos de textos recentes dos mais diversos gêneros e registros da variedade brasileira. O modelo que fundamenta a primeira dimensão da análise é a Gramática Léxico-Funcional (LFG), vertente não transformacional da teoria gerativa que confere um lugar central às funções gramaticais e ao léxico. O nível de representação fundamental do modelo é a estrutura funcional, que constitui entrada para a representação do significado das sentenças. Tratamos dos aspectos semânticos da passiva possessiva sob a perspectiva da Gramática de Decomposição Lexical (LDG), teoria lexicalista da vinculação argumental e das alternâncias verbais, baseada na decomposição de predicados (WUNDERLICH, 2000).

ma das vantagens da LFG em relação a outras abordagens gerativas, como o PM, é a completa formalização, facilitando a simulação computacional de fenômenos gramaticais específicos ou a compilação de *parsers*. Por conta disso e do *design* propício ao processamento semântico, a LFG tem sido utilizada para a compreensão textual, em sistemas de IE (NOVICHKOVA; EGOROV; DARASELIA, 2003), Q&A (BOBROW et al., 2007) etc. Um atrativo da LFG para estudantes e pesquisadores de linguística é o *Xerox Linguistic Environment* (XLE), um ambiente amigável de desenvolvimento de gramáticas, gratuitamente disponível para fins não comerciais. Parte substancial da arquitetura do modelo pode ser implementada nesse sistema em gramáticas de menor ou maior extensão, sem necessidade de conhecimentos aprofundados de informática. O XLE permite compilar *parsers* a partir dessas gramáticas e aplicá-los a sentenças individuais ou *corpora* inteiros, oferecendo diversos recursos para avaliar os resultados. Desse modo, uma determinada proposta de análise de um dado recorte gramatical pode ser automaticamente testada em um vasto conjunto de dados. Analogamente, abordagens alternativas de um mesmo conjunto de fenômenos podem ser automaticamente comparadas quanto à complexidade. Essa metodologia de avaliação não é praticável no caso de propostas elaboradas em teorias não completamente formalizadas e, portanto, não diretamente implementáveis, como o PM¹.

A análise gramatical proposta neste trabalho foi implementada no XLE no âmbito da BrGram (ALENCAR, 2013), uma gramática de média cobertura sintática, capaz de analisar exemplos não triviais como (1)-(3), preenchendo lacuna da gramática de Santos (2014), que não contempla a passiva possessiva². Desse modo, a BrGram pode contribuir de forma direta para sistemas de compreensão textual.

Na próxima seção, expomos o quadro teórico que fundamenta a análise gramatical e a implementação computacional da passiva possessiva do português da Seção 4. Essa proposta oferece uma solução para diversas questões suscitadas por estudos prévios sobre o fenômeno em coreano, japonês e português, resumidos na Seção 3. A última seção apresenta as conclusões e oferece sugestões para trabalhos futuros.

2 A GRAMÁTICA LÉXICO-FUNCIONAL

A LFG filia-se à gramática gerativa. No entanto, opõe-se aos modelos propostos por Chomsky, como o PM, dos quais diverge em relação às transformações sintáticas. Na LFG, apenas admitem-se transformações no léxico, via regras lexicais.

¹ Segundo Falk (2001, p. 65), o PM é uma teoria *semiformal*.

² A BrGram está disponível neste endereço: <<https://github.com/LFG-PTBR/BrGram>>.

"Os torcedores tinham quebrado os vidros do quiosque."

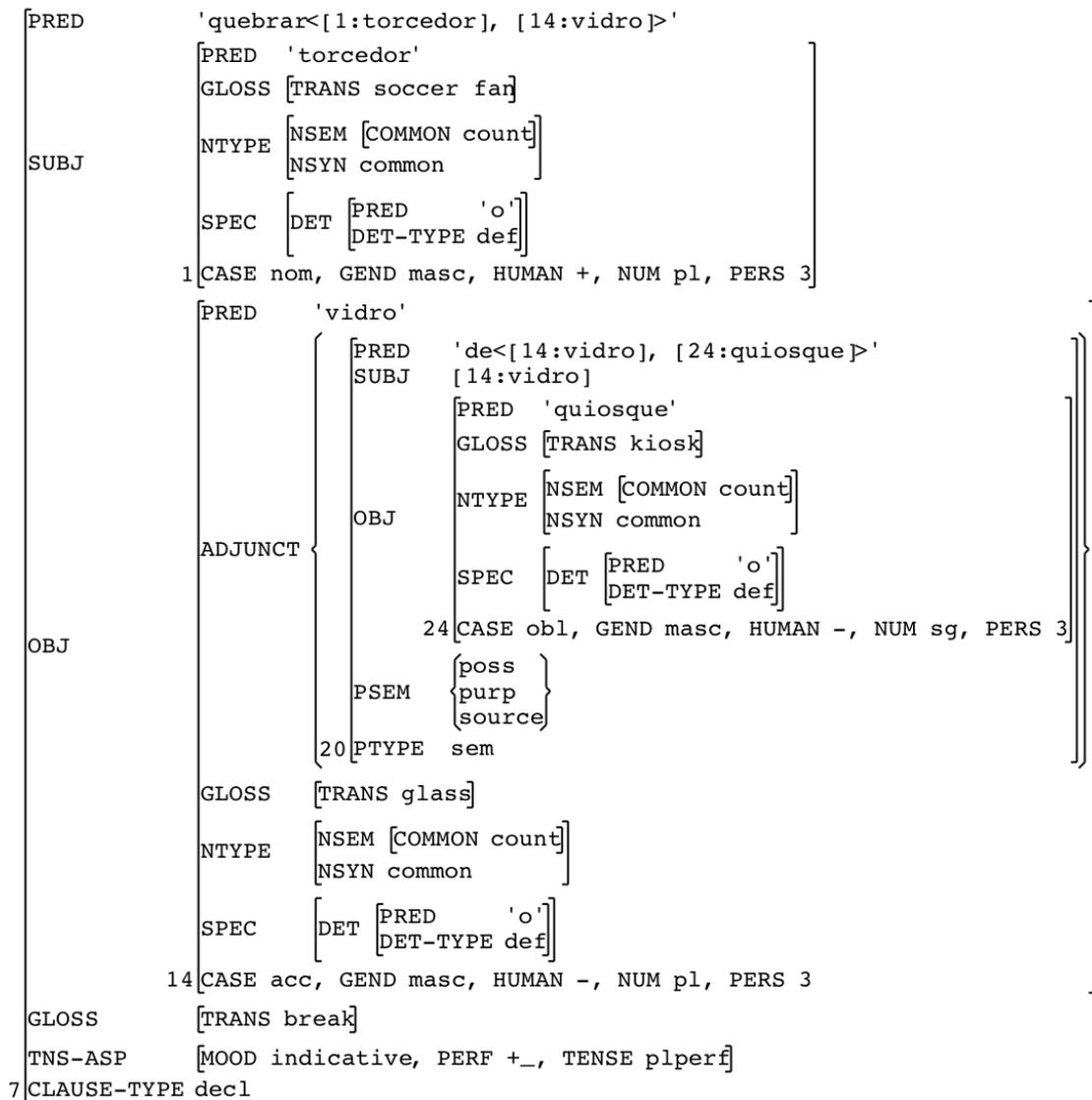


Figura 1: Estrutura F do exemplo (1) gerada pela BrGram

Fonte: Elaborada pelo autor.

Outra diferença em relação às abordagens transformacionais resulta do Princípio da Integridade Lexical (doravante PIL), conforme o qual as palavras são os átomos da sintaxe. Desse modo, processos do componente sintático não podem manipular elementos que não constituem palavras, por exemplo núcleos como *v**, adotado por Lunguinho (2011; 2013; 2016) em sua análise da passiva possessiva no quadro do PM. Por conta do PIL, a LFG desloca, para o componente lexical, a codificação de uma ampla gama de fenômenos analisados por meio de processos sintáticos nos modelos transformacionais. O léxico assume, dessa forma, uma enorme importância nessa teoria.

A LFG postula três projeções fundamentais para representar as propriedades gramaticais de uma sentença, das quais apenas as duas primeiras estão implementadas no XLE e, por conseguinte, na BrGram: (i) estrutura sintagmática ou de constituintes (estrutura C), (ii) estrutura funcional (estrutura F) e (iii) estrutura argumental (estrutura A). A primeira é representada por meio de diagramas arbóreos. A segunda codifica, entre outras propriedades, as funções gramaticais. A terceira constitui uma lista de traços que codificam os aspectos sintaticamente relevantes dos papéis semânticos de um predicado. Outras projeções estão previstas no modelo, como a estrutura fonológica, a estrutura semântica e a estrutura informacional (FALK, 2001).

A projeção mais importante, no presente contexto, é a estrutura F, que constitui *input* para a construção da representação do significado sentencial. A principal informação representada nessa estrutura consiste nas funções gramaticais, como SUBJ (sujeito), OBJ (objeto direto) etc. Nesse nível são também codificadas propriedades como tempo, modo, aspecto e voz.

A estrutura F constitui uma matriz de atributos e valores, em que um determinado atributo pode ter um valor atômico ou uma outra estrutura F como valor, como na Figura 1, que segue a notação de King (2004), a quem remetemos para a expansão das abreviaturas utilizadas nas figuras do presente artigo. Designamos por meio de f_1, f_2, \dots, f_n as estruturas F de índices 1, 2, ..., n. Na Figura 1, f_7 tem os atributos PRED (predicado), SUBJ e OBJ. O valor do primeiro é a forma semântica (7), no qual *quebrar* constitui um predicado de dois argumentos, saturados por f_1 e f_{14} , i.e., pelos valores dos atributos SUBJ e OBJ, cujas formas semânticas são os lemas *torcedor* e *vidro*. Abstraindo das informações modo-temporais, (7) corresponde, na lógica de predicados, a *quebrar(t,v)*, em que *t* e *v* se referem às entidades denotadas pelo SUBJ e OBJ.

(7) 'quebrar<[1:torcedor],[14:vidro]>'

O outro exemplo de função gramatical da Figura 1 é ADJUNCT, que engloba não só adjuntos adnominais, como nesse caso, mas também adjuntos adverbiais. Diferentemente de SUBJ e OBJ, ADJUNCT não é subcategorizado. O valor de ADJUNCT é um conjunto de estruturas F, representado por meio de chaves. No exemplo em tela, esse conjunto tem um único membro, cuja forma semântica é dada pela preposição *de*, constituindo a fórmula (8). Essa preposição expressa posse (*poss*), finalidade (*purp*) ou fonte (*source*) (KING, 2004)³. Num sistema de compreensão textual, a desambiguação entre essas diferentes acepções fica a cargo de um módulo posterior na cadeia de processamento.

(8) 'de< [14:vidro], [24: quiosque]>'

Fórmulas como (7) e (8) podem ser usadas em um sistema de IE para preenchimento de *templates* como DEPREDACÃO por meio da simples vinculação dos argumentos dessas fórmulas aos diferentes *slots*: o PERP e o ALVO são, respectivamente, o primeiro e o segundo argumento de (7), enquanto o POSSR é o segundo argumento de (8).

A Figura 1 exemplifica o tratamento dado na BrGram aos auxiliares utilizados nas perífrases verbais do tipo de (1), analisados como instâncias da categoria Flexão (doravante I), desprovidas de significação lexical e, portanto, de atributo PRED (FALK, 2001). Desse modo, não constituem domínio predicacional nem oracional próprio, contribuindo para a estrutura F da sentença apenas com traços como pessoa, número, tempo, modo, aspecto etc. No exemplo em questão, há um único domínio oracional, constituído pelo verbo *quebrar*. O auxiliar *ter*, implementado como núcleo I, contribui apenas com os traços *MOOD=indicative* e *TENSE=plperf*, especificando o modo e o tempo da sentença como indicativo e mais-que-perfeito.

Em (9), temos exemplo de entrada lexical na variante notacional do formalismo da LFG utilizada no XLE, no caso para a forma *quebrado* do exemplo (1). Nessa entrada, a primeira linha expressa a categoria lexical (V=verbo) seguida do predicado *quebrar*, cujos argumentos devem ser realizados pelo SUBJ e OBJ. Essa fórmula específica, dessa maneira, também a valência sintática ou moldura de subcategorização do verbo. Na segunda linha da entrada, cada uma das duas siglas precedidas de "@" constitui uma invocação de molde, um recurso do XLE análogo às sub-rotinas ou funções de linguagens de programação como C ou Python. Os moldes @PSTPT e @ACT caracterizam essa forma verbal como participio passado ativo.

³ Essa lista está longe de esgotar o amplo leque de relações coberto pela preposição *de*. Novos valores serão incluídos em versões futuras da BrGram.

(9) quebrado V * (^ PRED)= 'quebrar<(^ SUBJ)(^ OBJ)>'
@PSTPT @ACT.

Na BrGram, diferentes classes valenciais são codificadas por meio de moldes, simplificando o processo de elaboração de entradas lexicais. Para os verbos divalentes do tipo de (9), que governam um OBJ e são apassiváveis, define-se o molde TRNS (transitivo) na forma de (10). Esse molde toma como parâmetro o lema verbal, representado pela variável P. Invocando esse molde com o parâmetro *quebrar*, pode-se condensar a codificação da forma semântica de (9) como (11).

(10) TRNS(P) = (^ PRED)= 'P<(^ SUBJ)(^ OBJ)>'.
(11) quebrado V * @(TRNS quebrar)

Outra função gramatical subcategorizada relevante na nossa análise da passiva possessiva é o complemento predicativo XCOMP, que se caracteriza pela posição vazia do seu sujeito na estrutura C, preenchida na estrutura F por uma função gramatical nuclear do verbo matriz, estabelecendo uma relação de *controle* entre essas duas funções. Em (12)-(14), a função controladora é o OBJ do verbo matriz. O núcleo de um XCOMP pode ser qualquer categoria lexical (N, V, A etc.).

(12) O prefeito [...] convenceu os manifestantes a terminarem o protesto. (Google)

(13) A juíza [...] julgou improcedente a reclamação trabalhista. (Google)

(14) [...] o sapateiro viu o bandido fugir em uma bicicleta [...]. (Google)

"O prefeito convenceu os manifestantes a terminarem o protesto."

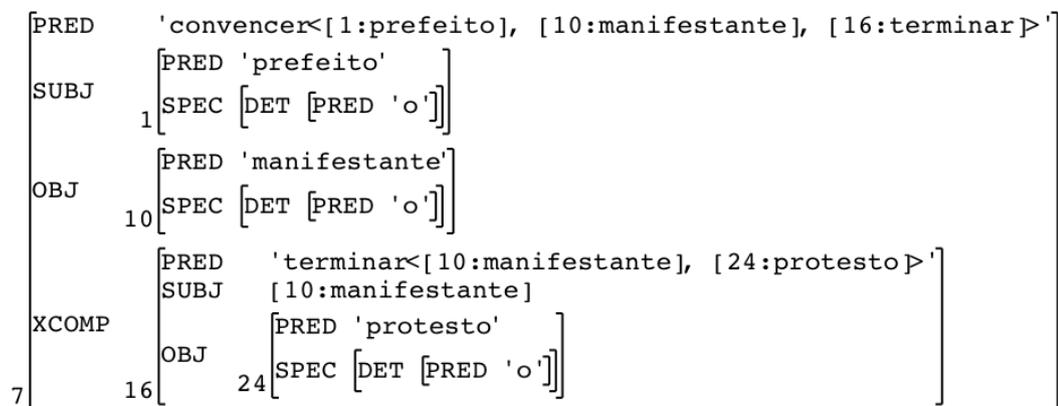


Figura 2: Estrutura F simplificada gerada pela BrGram para o exemplo (12)

Fonte: Elaborada pelo autor

A Figura 2 exemplifica o funcionamento do controle. Nesse exemplo, os argumentos do predicado *convencer* são preenchidos por f1, f10 e f16, valores, respectivamente, dos atributos SUBJ, OBJ e XCOMP. A estrutura f16 tem como representação semântica o

predicado diádico *terminar*, preenchido por f10 e f24, valores dos atributos SUBJ e OBJ. Desse modo, esse exemplo tem uma estrutura bioracional, constituída por uma oração matriz nucleada pelo verbo *convencer* e uma oração encaixada, argumento desse verbo, nucleada pelo verbo *terminar*. Pelo contrário, a estrutura da Figura 1 é mono-oracional. Como a estrutura F constitui *input* para projeção da estrutura S, a estrutura da Figura 2 projeta uma estrutura S bipredicacional, enquanto a estrutura F da Figura 1 projeta uma estrutura S monopredicacional.

A LFG modela a relação entre as funções controladora e controlada por meio de uma equação nas entradas lexicais, como na segunda linha de (15). Essa equação assegura, concomitantemente, a concordância entre essas duas funções gramaticais, bloqueando exemplos como (16).

(15) convenceu V * (^ PRED)='convencer<(^ SUBJ)(^ OBJ)(^ XCOMP)>'
 (^ XCOMP SUBJ)=(^ OBJ)

(16) *O prefeito convenceu o manifestante a terminarem o protesto.

Na BrGram, também a codificação de verbos de controle é simplificada por meio de moldes que permitem a modelação das propriedades comuns às diferentes classes de verbos desse tipo. Por exemplo, na elaboração de entradas para verbos de controle do objeto, a utilização do molde (17) permite condensar as informações das duas linhas de código de (15) em uma única linha, como exemplificado em (18).

(17) O-CTRL(P) = (^ PRED)='P<(^ SUBJ)(^ OBJ)(^ XCOMP)>'
 (^ XCOMP SUBJ)=(^ OBJ).

(18) convenceu V * @(O-CTRL convencer)

Os verbos de controle classificam-se em duas classes, conforme a relação da função controladora com o verbo matriz: verbos do tipo *equi* e verbos de alçamento. O verbo de (15) enquadra-se na primeira classe, na qual a controladora realiza argumento semântico (i.e. papel teta) do verbo matriz, o que não ocorre nos verbos da segunda classe, caso dos verbos aspectuais (*começar, cessar* etc.) e dos verbos causativos do tipo de *fazer*. Por exemplo, em (19), não é o verbo matriz que impõe restrições de seleção ao seu OBJ, mas o encaixado. Funções gramaticais subcategorizadas que não realizam papel teta são representadas na LFG fora dos parênteses angulares, como em (20).

(19) A iniciativa fez os ladrões fugirem. (Google)

(20) fez V * (^ PRED)='fazer<(^ SUBJ)(^ XCOMP)>(^ OBJ)'
 (^ XCOMP SUBJ)=(^ OBJ)

A Figura 3 exemplifica a análise de *ser* da passiva comum como verbo de alçamento, defendida por Morais (1988), Patejuk e Przepiórkowski (2014) etc.⁴. O predicado mais externo é *ser*, que governa duas funções: um XCOMP-PRED e um SUBJ, constituindo apenas o primeiro um argumento semântico. Na f23, temos praticamente a mesma fórmula (7) da contraparte ativa dessa sentença, a única diferença é que, na estrutura passiva, o primeiro lugar argumental é realizado pela estrutura F do OBL-AG, o agente da passiva, enquanto o segundo é preenchido pela estrutura F do SUBJ. As duas fórmulas expressam a mesma relação semântica, resultando no mesmo *template* preenchido de (4).

"Os vidros do quiosque tinham sido quebrados pelos torcedores."

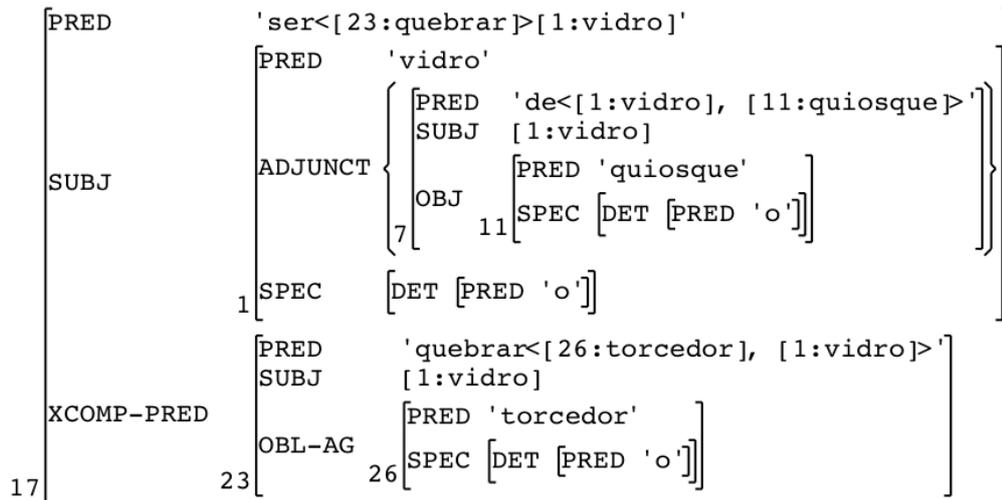


Figura 3: Estrutura F simplificada gerada pela BrGram para a contraparte passiva do exemplo da Figura 1

Fonte: Elaborada pelo autor

Por conta do PIL, a LFG adota uma análise puramente lexical das diáteses verbais, como passivização, alçamento de possuidor etc., explicando-as com base nas estruturas argumentais dos verbos. Esse nível de representação é manipulado por regras lexicais, processos que incidem sobre as representações fonológicas, sintáticas e/ou semânticas de entradas do léxico, gerando novas entradas. Dispensam, portanto, a codificação individual de entradas lexicais, conferindo economia à elaboração do léxico. No XLE, a modelação de regras lexicais responsáveis por diáteses, como a passiva, recorre ao modelo da LFG do início da década de 1980, segundo o qual a passiva envolve a manipulação de funções gramaticais, conforme as transformações de (21), passíveis de implementação no XLE por meio de um molde (CROUCH et al., 2011). Na primeira transformação, o OBJ vira SUBJ. A segunda envolve duas alternativas, conectadas pela disjunção lógica “|”: ou o SUBJ é transformado em OBL-AG ou sofre apagamento (indicado por NULL).

$$(21) (^{\wedge} \text{OBJ}) \rightarrow (^{\wedge} \text{SUBJ}) \\ \{ (^{\wedge} \text{SUBJ}) \rightarrow (^{\wedge} \text{OBL-AG}) \mid (^{\wedge} \text{SUBJ}) \rightarrow \text{NULL} \}$$

3 ANÁLISES PRÉVIAS

Diante da proximidade tipológica entre coreano e japonês (BUSSMANN, 2002, p. 383), não surpreende que essas línguas compartilhem o mesmo padrão sintático superficial na passiva possessiva. No entanto, conforme Oshima (2004), são muitas as diferenças entre as duas línguas nesse quesito. Em primeiro lugar, o leque de relações entre os referentes do sujeito e do objeto é bem mais amplo em japonês, que licencia “qualquer relação pragmaticamente significativa” (p. 16). Em segundo, sujeitos inanimados

⁴ Isso é matéria de controvérsia na LFG. Bresnan (2001, p. 116), por exemplo, classifica esse elemento em inglês como I.

são admitidos na construção japonesa, restrita em coreano a sujeitos animados. Em terceiro, a passiva possessiva japonesa (tal como a passiva direta) não implica que o sujeito seja adversamente afetado, característica essa da passiva indireta, tipo inexistente em coreano. Pelo contrário, as duas passivas do coreano implicam que o sujeito, quando animado, é adversamente afetado. Em quarto, apenas o sufixo apassivador do japonês é produtivo, sua contraparte em coreano restringe-se a formas lexicalizadas. A quinta diferença diz respeito à ambiguidade dos sufixos apassivadores. Em coreano, esse sufixo funciona também como causativizador. Em japonês, em vez disso, dependendo do tipo de relação entre o sujeito e o objeto, determinadas sentenças passivas podem ser interpretadas tanto como possessivas quanto como indiretas. A última diferença refere-se à quantidade de estratos da construção em cada língua. No japonês, a passiva possessiva, assim como os outros dois tipos, constitui uma estrutura semântica biestratal, expressando uma relação estativa de falta de controle de alguém sobre a ação de outrem, envolvendo três participantes: um ator, um *undergoer* e um “efeito”, do qual o segundo participante funciona como ator (OSHIMA, 2003). Essa complexidade da passiva japonesa no nível semântico, porém, não se projeta no plano sintático. Um radical verbal apassivado funciona, sintaticamente, como um único verbo, com uma moldura valencial que especifica um sujeito, um complemento dativo e um acusativo. Desse modo, verbos apassivados do japonês apresentam estrutura sintática e semântica análoga à dos verbos causativizados (OSHIMA, 2003). Já em coreano, ambas as variantes da passiva são sintática e semanticamente monoestratais. O argumento adicional da passiva possessiva resulta, para Oshima (2004), de alçamento do possuidor, que destaca como um processo comum na língua.

Em sua tese sobre verbos auxiliares, Lunguinho (2011) compara os dois tipos de passiva do português, a “canônica” com auxiliar *ser* e a “não canônica” ou “adversativa” com auxiliar *ter*. Ambas são formadas com o particípio passivo, restrito a verbos com um argumento externo e um interno direto, com o qual concorda, distinguindo-se do particípio perfeito, que não sofre essa restrição argumental nem exibe concordância. Outras propriedades em comum entre as duas passivas são a demissão do argumento externo e a promoção do argumento interno direto. A partir de exemplos gramaticais e agramaticais construídos, ele levanta as diferenças entre os dois tipos, para os quais propõe uma explicação no quadro do PM. Como essa proposta foi reformulada em estudos posteriores (LUNGUINHO, 2013, 2016), de que trataremos em seguida, nos limitamos aqui aos aspectos aos quais nos contrapomos na nossa proposta.

O primeiro aspecto consiste na posição do argumento interno relativamente ao particípio. Para Lunguinho (2011), este deve subseguir aquele, como em (3), pelo que (22) seria agramatical.

(22) *A porta teve consertada a maçaneta (pelo maceneiro). (p. 43)

O segundo aspecto refere-se ao licenciamento de uma categoria vazia (*ec*) na posição de possuidor, como em (23), que constituiria vestígio deixado pelo alçamento do possuidor, passível de ocorrer também de dentro de um argumento interno indireto ou externo, como em (24) e (25), respectivamente.

Segundo Lunguinho (2011, p. 74), “apenas as passivas canônicas permitem que possuidor e possuído se mantenham unidos como um constituinte”. Ele admite não ter explicação para a obrigatoriedade do movimento do possuidor para especificador da categoria T(empo), que impediria a derivação de sentenças como (26).

(23) ele_i teve conversas *ec_i* interceptadas pela PF (Google)

(24) Byron_i teve cinco anos acrescentados à pena *ec_i*.⁵

⁵ Baseado na seguinte ocorrência extraída do Google, análoga ao exemplo (63a) de Lunguinho (2011, p. 72): “Byron Moreno está preso e deve ter cinco anos acrescentados à pena.”

(25) [O Pedro]_i teve as contas pagas *ec*_i pelo pai *ec*_i. (p. 72)

(26) *Nesse momento tem o suspeito a casa revistada pela polícia. (p. 76, n. 27)

Lunguinho afirma utilizar *possuidor* como um termo genérico para cobrir também casos em que a categoria vazia não corresponde, necessariamente, ao possuidor de uma relação de posse alienável (v. (72)) ou inalienável (v. (44)), constituindo argumento de nome relacional (v. (84)) ou de relação todo-parte (v. (3)) ou tema de deverbais (v. (65)).

O terceiro aspecto concerne à realização do possuidor como um pronome lexical, não havendo diferença de marcação relativamente à realização como categoria vazia:

(27) Ele teve uma conversa dele grampeada e divulgada. (Google)

(28) Ela teve uma conversa sua [...] impressa pelo diretor da escola. (Google)

O quarto aspecto refere-se às propriedades do auxiliar *ter*, que atribuiria papel temático ao seu sujeito e licenciaria o acusativo do argumento interno do participio. Admitindo que essa caracterização contraria a sua própria definição de auxiliar como categoria incapaz de selecionar argumento e atribuir papel temático, Lunguinho (2011, p. 80) a justifica pela incorporação do auxiliar *ser* em *ter*. Ele não especifica, porém, o papel temático atribuído por *ter*, limitando-se a caracterizá-lo como “compatível com a interpretação de afetação” (p. 77) geralmente associada, segundo ele, ao argumento na posição de sujeito.

Para Lunguinho (2013, p. 2), (29) é ambígua entre as leituras (30a) e (30b), às quais correspondem as perguntas (31a) e (31b), respectivamente:

(29) O João tem dois artigos citados pelo Chomsky.

(30) a. O João possui dois artigos citados pelo Chomsky.
b. Dois artigos do João são citados pelo Chomsky.

(31) a. O que o João tem?
b. Por quem o João tem o trabalho citado?

Lunguinho (2013) denomina a construção de (29), na leitura (30a), de possessiva e, na leitura (30b), como anteriormente, de passiva não canônica. Como evidenciam (31a) e (31b), a cada leitura subjaz uma estrutura sintática diferente. Na primeira interpretação, o verbo *ter* (doravante *ter*₁) expressa uma relação de posse entre seus argumentos externo e interno, enquanto o participio passado constitui um modificador desse último.

Na segunda interpretação de (29), o referente de [o João] não funcionaria como possuidor de [dois artigos], mas como autor. Nesse caso, *ter* funcionaria como auxiliar (doravante *ter*₂), analogamente a *ser* na passiva canônica. Contudo, a passiva não canônica

discreparia da canônica por não admitir o preenchimento da posição de sujeito por expletivo (v. (32)) e exibir um argumento adicional.

(32) São/*Têm citados pelo Chomsky dois artigos do João. (p. 2)

Para explicar os aspectos comuns e as diferenças entre as duas passivas, Lunguinho propõe uma análise no quadro do PM, que toma como ponto de partida a derivação da passiva canônica representada em (33), base da leitura (30b).

(33) [_{vSER} [_{VoicE} [_{PartP} [_{DP} [_{DP} O João] dois artigos] [_{Part} citados [_{VP...}]]] [_{VoicE} por [_{v*P} [_{DP} o Chomsky]] (p.2)

As diferenças entre as duas passivas decorreriam de um núcleo v^* capaz de introduzir um argumento, licenciando o alçamento do possuidor à posição de sujeito, e valorar o Caso do argumento interno do participio. Esse núcleo projetaria estrutura dominando a projeção do auxiliar *ser* representada em (33), da qual derivaria ter_2 pela concatenação de v^* e *ser*.

Lunguinho (2016) foca a interpretação (30b), propondo, para os dois tipos de passiva, a mesma análise de Lunguinho (2013). Três pontos desse trabalho mais recente merecem ser destacados. O primeiro refere-se aos tipos de relação que o argumento extra da passiva não canônica entretém com o argumento interno do participio. No caso de tema de deverbais, Lunguinho (2016) denomina a relação de *complementação*. A relação expressa por nome relacional não é mais mencionada. O segundo ponto concerne à afetação do argumento na posição de sujeito de ter_2 . Lunguinho (2016) é mais categórico do que Lunguinho (2011), tratando essa propriedade não como acidental, mas essencial. Finalmente, o terceiro ponto refere-se a exemplos como (34), considerados agramaticais devido à ausência de relação semântica entre o argumento na posição de sujeito de *ter* e o argumento interno do participio:

(34) *O Paulo teve as contas da Maria pagas pela Ana. (LUNGUINHO, 2016, p. 13)

4 ANÁLISE LÉXICO-FUNCIONAL E DECOMPOSICIONAL

Nesta seção, propomos respostas com base na LFG e na LDG, para questões sobre a passiva possessiva em português suscitadas pelas abordagens da seção anterior. Os aspectos lexicais e sintáticos da proposta são implementados no âmbito da BrGram.

A primeira questão diz respeito à ambiguidade de (29) entre (30a) e (30b), que, conforme Lunguinho (2013), correlaciona-se com a oposição entre dois tipos de papel do referente do sujeito em cada estrutura. Na primeira, seria possuidor, enquanto, na segunda, autor. Na verdade, essa diferenciação de papel é ortogonal à ambiguidade estrutural. Decorre não da estrutura sintática da sentença, mas da ambiguidade das expressões que denotam relação de posse *lato sensu*, quando o segundo argumento (doravante POSSM) é referido por expressão nucleada por substantivo como *livro, artigo, escultura, vídeo, filme* etc. Por exemplo, em todos os casos de (35), o primeiro argumento da relação (doravante POSSR) pode constituir tanto “autor” quanto “possuidor” do POSSM. Essa ambiguidade é menos plausível nos mesmos contextos com substantivos do tipo *pedra, cachorro, terreno* etc.

- (35) a. O João tem dois artigos.
 b. os dois artigos do João
 c. os dois artigos dele
 d. os seus dois artigos

Segundo Pustejovsky (1995), a representação semântica de substantivos que designam artefatos do tipo de *livro*, *artigo* etc. contém a especificação de um atributo AGENTE cujo valor é o tipo de evento responsável pela produção da entidade. No caso de (35), trata-se de um evento do tipo *escrever*. Em estruturas desse tipo, o POSSR pode, então, ser identificado com o primeiro argumento desse predicador. Pelo contrário, em *sua pedra* ou *o cachorro dela*, a interpretação do POSSR como autor não emerge porque esses substantivos, pelo menos na sua especificação por *default*, não têm valor para o atributo AGENTE.

Desse modo, (30a) e (30b) são ambíguas entre uma interpretação em que o referente de [*o João*] é “autor” e outra em que é “possuidor”. Consequentemente, (29) não possui apenas duas leituras, como propõe Lunguinho (2013), mas quatro, uma vez que tanto a estrutura com *ter*₁ quanto a com *ter*₂ expressam uma relação entre os referentes de [*o João*] e [*dois artigos*] tal que aquele tanto pode ser “autor” quanto “possuidor” deste.

Qual a diferença semântica entre as duas leituras de (29), dado que não se refere ao tipo de relação entre os referentes do sujeito e do objeto? A primeira diferença concerne à classe aspectual de *ter*₁ e *ter*₂. Na leitura (30a), a classe aspectual de (29) é estado, como evidencia (36a). Nesse exemplo, *possuir* não admite a forma progressiva marcada pela perífrase de gerúndio, o que, segundo Cançado e Amaral (2016, p. 151), constitui critério diagnóstico dos verbos de estado. Pelo contrário, na leitura (30b), a classe aspectual de (29) não é estado, como evidencia a gramaticalidade de (36b). Em consonância com isso, a única leitura de (36c) é (36b). Outro teste para identificar verbos de estado, conforme Cançado e Amaral (2016, p. 152), é a inadequação do seu uso como resposta a perguntas do tipo de (37a), como vemos em (37b). A resposta (37c) só é admissível na leitura (37d), corroborando a classificação aspectual de *ter*₁ como estativo e a não inclusão nessa classe de *ter*₂.

- (36) a. *O João está possuindo dois artigos citados pelo Chomsky.
 b. Dois artigos do João estão sendo citados pelo Chomsky.
 c. O João está tendo dois artigos citados pelo Chomsky.

- (37) a. O que aconteceu?
 b. #O João possuiu dois artigos citados pelo Chomsky.
 c. O João teve dois artigos citados pelo Chomsky.
 d. Dois artigos do João foram citados pelo Chomsky.

Propomos que *ter*₂ não é marcado para aspecto. A classe aspectual da passiva possessiva, analogamente ao que ocorre na passiva comum, é determinada pelo verbo do qual deriva o particípio, como evidenciam (38)-(41). O verbo *usar* de (38) enquadra-se nas atividades, por ser dinâmico e atélico, admitindo a forma progressiva, enquanto *pintar*, *decorar* e *limpar* de (39)-(41) constituem *accomplishments*, pois são dinâmicos e télicos, diferenciando-se de *arrancar* de (42), que constitui *achievement* por não possuir intervalos internos (CANÇADO; AMARAL, 2016, p. 153-167).

(38) O prefeito diz que soube que estava tendo o nome usado por golpistas em meados de março de 2015. (Google)

(39) O Home Office [...] teve as paredes pintadas de tinta látex preta fosca. (Google)

(40) A pequena casa de taipa [...] teve suas paredes decoradas com as ostras. (Google)

(41) O site teve as “pichações” limpadas pela tecnologia da empresa por volta das 9 horas. (Google)

(42) A vítima conta que carregava a carteira embaixo do braço, quando a teve arrancada pelo infrator. (Google)

A segunda questão refere-se à categoria sintática de *ter*₂. Linguinho (2011) analisa-o no quadro do PM como um auxiliar atípico, que discrepa dos demais membros dessa categoria por selecionar um argumento, no caso o externo, atribuir-lhe papel teta e valorar Caso, especificamente o acusativo do argumento interno do participio, explicitado no clítico de (42). Esse comportamento corresponde, na LFG, ao de um verbo de alçamento do objeto, que seleciona semanticamente o seu SUBJ, mas licencia apenas sintaticamente o seu OBJ. Por outro lado, o verbo *ser* da passiva comum tem sido analisado no âmbito da LFG como um verbo de alçamento do sujeito. Nesse cenário, configuram-se duas alternativas para a categorização de *ter*₂ no quadro da LFG: (i) auxiliar ou (ii) verbo de alçamento. Conforme (i), a construção seria monoestratal, como em coreano, ao passo que (ii) implicaria uma estrutura biestratal, tal como a estrutura semântica da construção equivalente em japonês. Não optamos por nenhuma dessas duas categorizações, classificando *ter*₂, em vez disso, como verbo de controle do objeto do tipo *equi*, tal como *convencer* em (15) e (18), constituindo uma estrutura F biestratal, como exemplificado na Figura 4, estruturalmente análoga à Figura 2.

Na BrGram, o molde (43) codifica, por meio da disjunção de duas invocações de molde, as propriedades de *ter*₁ e *ter*₂. Os dois moldes disjuntivamente invocados estão definidos em (10) e (17), respectivamente. A relação de controle identifica o OBJ de *ter*₂ com o SUBJ do XCOMP, assegurando a concordância do participio. Na Figura 5, temos a estrutura F da interpretação de (44) com *ter*₁, implausível no contexto em tela.

(43) VERB-TER={"ter1" @(TRNS ter) | "ter2" @(O-CTRL ter)}.

"Gaúcho teve o cabelo elogiado pelo Valdir."

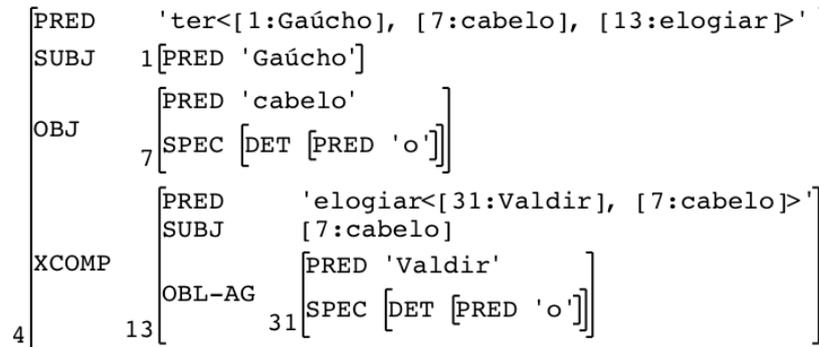


Figura 4: Estrutura F simplificada gerada pela BrGram para a interpretação mais plausível de (44)

Fonte: Elaborada pelo autor

"Gaúcho teve o cabelo elogiado pelo Valdir."

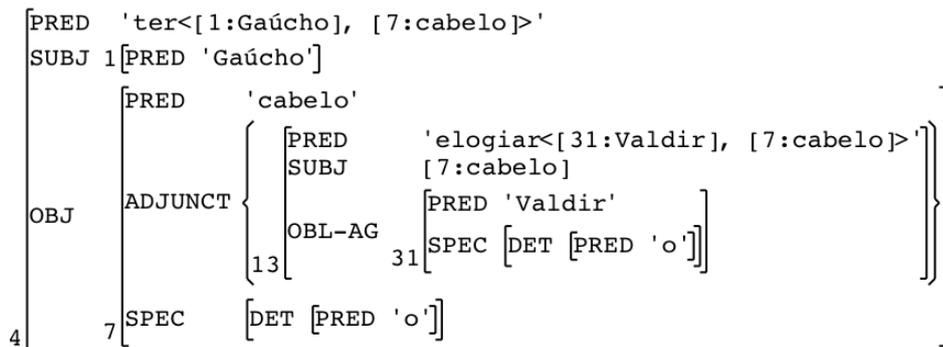


Figura 5: Estrutura F simplificada gerada pela BrGram para a interpretação menos plausível de (44)

Fonte: Elaborada pelo autor

Conforme (43), ter_2 é um verbo pleno, do qual não só o SUBJ, mas também o OBJ, realiza argumento semântico, governando também um XCOMP, realizado pelo particípio passivo, cujo SUBJ é controlado pelo OBJ. Como a função controladora é argumento semântico do verbo matriz, trata-se de verbo do tipo *equi*. Na Figura 4, temos uma estrutura bipredicacional (e, portanto, bioracional), cujo primeiro nível de predicação é constituído pelo predicado *ter* de três lugares, preenchidos pelas estruturas F do SUBJ, do OBJ e do XCOMP, respectivamente. O segundo nível predicacional é constituído pelo valor do atributo PRED do XCOMP, que é o predicado diádico *elogiar*, cujos argumentos são saturados, respectivamente, pelas estruturas F do OBL-AG e do SUBJ do XCOMP, a mesma do OBJ de *ter*.

A terceira questão refere-se à representação semântica de ter_2 . Para tanto, propomos, com base LDG, uma representação da forma semântica (SF) da sentença da Figura 4, parafraseada em (45). Na LDG, a SF de um verbo se restringe aos aspectos do significado relevantes para a vinculação argumental. As relações temporais entre subeventos são especificadas apenas no nível da estrutura conceptual (CS). Desse modo, em vez de operadores de relações temporais (PUSTEJOVSKY, 1995, p. 69-70), a LDG utiliza, na SF, a conjunção lógica &.

(44) Gaúcho [...] teve o cabelo elogiado pelo Valdir. (Google)

(45) O Valdir elogiou o cabelo de Gaúcho.

Uma vez que essas duas sentenças são logicamente equivalentes, compartilham a mesma SF. Seguindo Wunderlich (2000), entre outros, traduzimos *elogiar* como $\lambda y \lambda x [ELOGIAR(x,y)]$ e utilizamos POSS para representar a preposição *de*. Abstraindo da flexão modo-temporal, propomos (46) como SF de (44) e (45).

(46) POSS(gaúcho,o.cabelo)&ELOGIAR(valdir,o.cabelo)

(47) i.	λQ	λy	λx	$[POSS(x,y) \& Q(y)]$
ii.	XCOMP OBJ	SUBJ		

A partir de (46), resulta, por abstração lambda, a SF de ter_2 em (47i) com a vinculação argumental de (47ii). O primeiro argumento de (47i) é uma variável do tipo $\langle e,t \rangle$, i.e., uma função que recebe uma entidade e retorna um valor de verdade, os outros dois são entidades. Dado que a variável Q predica sobre o argumento interno de POSS, realizado como OBJ, é essa função que constitui a função controladora do SUBJ do XCOMP.

Mostramos em (48)-(52) como o significado de (45) deriva de (47) por aplicação funcional e redução lambda a partir da composição de *ter* com seus argumentos (WITT, 1998, p. 47). Assumimos $\lambda v \lambda u [ELOGIAR(v,u)]$ como SF do particípio passivo, idêntica à do verbo na forma ativa, exceto pela inversão dos abstratores lambda (STERNEFELD, 2006, p. 673). Desse modo, representando como $[[x]]$ o significado de uma expressão x , $[[\textit{elogiado}]]$ se compõe primeiro com o OBL-AG para formar $[[\textit{elogiado pelo Valdir}]]$ em (48), com o que $[[\textit{ter}]]$ se compõe em (49), resultando em $[[\textit{ter elogiado pelo Valdir}]]$ em (50), que, por sua vez, é aplicado a $[[\textit{o cabelo}]]$ em (51). O resultado é $[[\textit{ter o cabelo elogiado pelo Valdir}]]$ em (52), que, aplicado a $[[\textit{gaúcho}]]$, resulta em $[[\textit{Gaúcho ter o cabelo elogiado pelo Valdir}]]$ em (53). Num sistema de IE, por meio do mesmo procedimento utilizado para derivar (53), o significado de (3) pode ser computado, resultando na fórmula em (54), a partir da qual o molde DEPREDAÇÃO pode ser facilmente preenchido, resultando em (6). Um procedimento análogo aplica-se a (1) e (2).

(48) $\lambda v \lambda u [\text{ELOGIAR}(v,u)](\text{valdir})$

(49) $\lambda Q \lambda y \lambda x [\text{POSS}(x,y) \& Q(y)](\lambda u [\text{ELOGIAR}(\text{valdir},u)])$

(50) $\lambda y \lambda x [\text{POSS}(x,y) \& \lambda u [\text{ELOGIAR}(\text{valdir},u)](y)]$

(51) $\lambda x \lambda y [\text{POSS}(x,y) \& \text{ELOGIAR}(\text{valdir},y)](\text{o.cabelo})$

(52) $\lambda x [\text{POSS}(x,\text{o.cabelo}) \& \text{ELOGIAR}(\text{valdir}, \text{o.cabelo})](\text{gaúcho})$

(53) $\text{POSS}(\text{gaúcho},\text{o.cabelo}) \& \text{ELOGIAR}(\text{valdir}, \text{o.cabelo})$

(54) $\text{POSS}(\text{bancos...},\text{vidraças}) \& \text{QUEBRAR}(\text{manifestantes}, \text{vidraças})$

Vejamos agora que argumentos favorecem, sob a perspectiva da LFG, a nossa proposta, comparativamente à abordagem de Lunguinho. Em primeiro lugar, esta viola o PIL, pois envolve a manipulação, na sintaxe, de categorias, como o núcleo v^* , que não constituem palavras. Em segundo lugar, ao analisar a passiva possessiva como uma sucessão de movimentos, envolvendo a extração do possuidor de dentro de argumento do participípio e seu alçamento à posição de SUBJ de *ter*, é incompatível com a natureza não transformacional da sintaxe na LFG. Em terceiro lugar, um auxiliar, na LFG, não tem representação semântica sob a forma de um atributo PRED. Desse modo, não pode ter grade temática nem moldura de subcategorização. Caso *ter*₂ seja analisado como núcleo I, não pode atribuir papel temático algum nem governar função gramatical. Os papéis POSSR e POSSM têm de ser selecionados, nesse caso, pelo participípio, que deve governar as funções gramaticais que os realizam, ou seja, o SUBJ e o OBJ.

Uma inconveniência dessa análise é que não é contemplada por (21), que gera as molduras de subcategorização dos participípios da passiva canônica. Para analisar a passiva possessiva no quadro dessa proposta, essa regra precisa ser reformulada, de modo a incluir uma alternativa em que o OBJ é preservado e um novo argumento (o POSSR), a ser realizado como SUBJ, é introduzido na moldura. Com essa reformulação, a complexidade do léxico aumenta proporcionalmente ao número de participípios passivos, porque cada elemento desse grupo passa a ter duas representações, uma para a passiva canônica e outra para a possessiva. Nossa proposta é, portanto, menos complexa computacionalmente, uma vez que utiliza, na passiva possessiva, os mesmos participípios da canônica. A única complexidade que introduz no léxico se refere à inclusão de *ter*₂ em (43).

A análise de *ter*₂ como núcleo I apresenta um problema também de ordem sintática. Em (55) e (56), apresentamos as estruturas de I' e V' de (1) e (3) geradas pela BrGram. Em (57), reformulamos (56) em termos da análise de *ter*₂ como núcleo I. Observe que, em (57), contrariamente a (55), o DP complemento de V é realizado à esquerda desse núcleo, contrariando a ordem canônica de núcleos e complementos no português, uma língua de núcleo inicial. A implementação dessa análise é desvantajosa por necessitar da introdução, no aparato de regras sintagmáticas, de uma regra que gere VP com núcleo final. Nossa proposta em (56) não padece desse tipo de problema, pois obedece rigorosamente ao parâmetro do núcleo inicial: tanto o núcleo V quanto o Part(icipípio) precedem os seus respectivos complementos.

(55) os torcedores [_I tinham] [_{VP} [_V quebrado] [_{DP} os vidros do quiosque]]]

(56) bancos... [_V tiveram] [_{DP} vidraças] [_{PartP} [_{Part'} [_{Part} quebradas] [_{PP} pelos...]]]]]

(57) bancos... [_I tiveram] [_{VP} [_V [_{DP} vidraças] [_V quebradas] [_{PP} pelos...]]]]]

Há também um argumento de natureza semântica a favor da nossa análise, em detrimento da análise de *ter*₂ como I ou como verbo de alçamento. Consideremos os exemplos de (58) e suas respectivas paráfrases em (59).

(58) a. O livro teve uma página rasgada pelo vândalo.
b. Uma página do livro foi rasgada pelo vândalo.

(59) a.#O lago teve uma página rasgada pelo vândalo.
b.#Uma página do lago foi rasgada pelo vândalo.

A sentença (58a) é perfeitamente normal, correspondendo à paráfrase (58b). O exemplo (59a) difere de (58a) apenas no tipo de entidade referida pelo SUBJ. No entanto, é semanticamente anômalo, assim como sua paráfrase em (59b). A que se deve esse contraste? Nossa proposta oferece uma explicação natural para a normalidade de (58), por um lado, e a anomalia de (59), por outro. Conforme (43) e (47), *ter*₂ seleciona semanticamente tanto o SUBJ quanto o OBJ, que realizam os dois argumentos do predicado POSS. A anomalia do segundo grupo de exemplos decorre, naturalmente, da dificuldade de se conceber uma relação de posse entre um lago e uma página, enquanto POSSR e POSSM, respectivamente, dessa relação, ao contrário do que ocorre com o primeiro grupo.

Nas propostas que analisam *ter*₂ como I ou verbo de alçamento do objeto, o POSSM não é argumento semântico desse núcleo, que, conseqüentemente, não expressa a relação de posse inerente à construção. Na abordagem de Lunguinho, essa relação é abrigada em argumento do participípio, do qual o possuidor é extraído, operação incompatível com a LFG. No quadro dessa teoria, para manter as análises de *ter*₂ como núcleo funcional I ou verbo de alçamento, é necessário codificar, de alguma forma, a relação POSS no participípio passivo. Como essa categoria não expressa essa relação na passiva canônica, faz-se necessário estipular uma extensão semântica para incluí-la. Como esse processo deve aplicar-se a todo participípio passivo, é evidente que essa solução incorre no mesmo problema de complexidade anteriormente apontado. Outra questão que as duas abordagens alternativas enfrentam se refere ao licenciamento do argumento adicional no papel de POSSR. Se *ter*₂ for considerado I, é o participípio passivo que precisa licenciar esse argumento. Na análise de *ter*₂ como verbo de alçamento, o POSSR pode ser licenciado por esse verbo. Nesse caso, porém, tem-se uma assimetria que parece injustificável no tratamento dos dois participantes da relação POSS, uma vez que o POSSM é licenciado pelo participípio. Uma maneira de contornar essa dificuldade seria codificar POSS no participípio, como na abordagem que trata *ter*₂ como I, e estipular, na esteira da abordagem de Lunguinho, um novo papel semântico para o argumento na posição de sujeito de *ter*₂, correferente do POSSR. Veremos mais adiante, contudo, que esse argumento não exerce papel de afetado.

Nossa proposta, pelo contrário, dispensa todas essas estipulações. O participípio da possessiva passiva é o mesmo da comum. A relação de posse que diferencia o primeiro do segundo tipo de passiva decorre de *ter*₂, que, como *ter*₁, codifica POSS, responsável pelos papéis semânticos POSSR e POSSM do SUBJ e do OBJ. Desse modo, nossa proposta correlaciona de forma natural a diferença semântica entre as duas passivas (i.e. presença do predicado POSS) à diferença sintática entre elas (i.e. presença de *ter*). Sob essa perspectiva, a passiva possessiva constitui mera diátese de *ter*, ao nosso ver resultado de um processo geral de aumento de valência que Wunderlich (2000) modela por meio da regra lexical de extensão semântica ARG, responsável por construções resultativas

como *Ela cortou os cabelos curtos* e, segundo nos parece, subjacente a diáteses análogas de vários outros verbos do tipo *equi*, como *convencer, julgar, ver* etc. (ver (12)-(14)).

A quarta questão foca a ordenação de constituintes. Para Lunguinho (2011), exemplos como (22b) são agramaticais, contrariamente às evidências de (60)-(62).

(60) [...] o brasileiro “Aquarius” [...] teve elogiada a atuação de Sonia Braga pelo crítico. (Google)

(61) [...] o Prefeito de Canela teve vetado pela Câmara de Vereadores o seu pedido de autorização para viajar à Itália (Google)

(62) [...] o Bitfinex teve roubado por hackers 65 milhões de dólares. (Google)

As ocorrências (13), (63) e (64) evidenciam que, em verbos de controle do objeto, o XCOMP pode preceder o OBJ. Como na nossa análise essas duas funções gramaticais são governadas por *ter*₂ e o português geralmente não estabelece uma ordem rígida entre complementos verbais, são esperados exemplos como (60)-(62) com a ordem XCOMP OBJ, ao lado de exemplos com a ordem inversa OBJ XCOMP como (3), (38) e (39). Na análise de Lunguinho (2011), porém, exemplos do primeiro tipo não são licenciados.

(63) A categoria quer resolvida na mesa de negociação a cláusula de compensação dos dias parados. (Google)

(64) [...] a pernambucana viu chegarem às lojas só um LP, “A Rainha da Ciranda” (1977), e um CD. (Google)

Contrariamente à análise de Lunguinho (2013) de (32), (65) evidencia que a passiva possessiva licencia SUBJ final. Como evidencia (66), a agramaticalidade daquele exemplo decorre da não realização do OBJ.

(65) Além de João Vaccari, tiveram as condenações confirmadas pelo tribunal o marqueteiro João Santana [...], a mulher dele, Mônica Moura, e o lobista Zwi Skornicki. (Google)

(66) Têm trechos citados pelo Chomsky dois artigos do João.

Outra divergência da nossa proposta em relação à de Lunguinho (2011) refere-se a (26), considerado por ele agramatical. Embora sem comprovação de textos reais, julgamos gramatical esse exemplo, haja vista ocorrências como (67), que evidenciam que a inversão do sujeito é licenciada em verbos do tipo *equi*.

(67) Por despacho proferido no dia 19 de Dezembro de 2001, julgou o juiz a arguição da executada improcedente. (Google)

A quinta questão refere-se aos tipos de relação de posse. Vimos que a passiva possessiva coreana se restringe à posse inalienável, enquanto a japonesa admite ampla gama de relações. Para o português, Lunguinho (2011) propõe um leque de cinco tipos. Nossa análise em (47) prediz espectro de relações coincidente com o de outros elementos com o predicado POSS na sua composição, como *ter*₁ e os possessivos, predição corroborada pelo fato de tanto o possessivo sintético quanto o analítico poderem explicitar essa rela-

ção, como em (27), (28), (40) e (61). Desse modo, no caso de deverbal, o POSSM não se limita ao papel tema, como propõe Lunguinho, mas pode exercer também, analogamente aos possessivos, outros papéis, como agente (v. (23), (27) e (61)), recipiente (v. (68)) etc.

(68) Os setores de imprensa dos Ministérios tiveram ampliadas as suas atribuições. (Google)

A sexta questão concerne à realização do POSSM, que, no coreano e no japonês, pelos exemplos de Oshima (2003; 2004), restringe-se ao OBJ, restrição que incorporamos na nossa análise em (47). No entanto, (24) de Lunguinho (2011), abonado por (69)-(71), evidencia que o POSSM pode se realizar por um argumento interno indireto do participio.

(69) No comprimento, o novo Palio teve adicionados mais 28 mm. (Google)

(70) a comissão teve acrescentadas em seu nome [...] as palavras “Direitos Humanos”. (Google)

(71) O piloto [...] teve adicionados 10s ao seu tempo final de prova. (Google)

As ocorrências (72)-(74) evidenciam que o POSSM pode se realizar também como ADJ(UNCT) do OBJ. Com efeito, em (74), por exemplo, não foi a largura do carro que foi ampliada, mas a da sua grade dianteira.

(72) Carmem teve o carro dela e [o _{ADJ} do filho]] arrastados pela correnteza. (Google)

(73) [...] o Luverdense [...] teve quebrados [os vidros _{ADJ} do ônibus que o levou ao Estádio]] [...]. (Google)

(74) [...] o novo [Audi] TT teve [a largura _{ADJ} da grade]] ampliada [...]. (Google)

Para contemplar a realização do POSSM por outras funções que não OBJ, propomos a noção de *posse mediata* (MPOSS), que definimos em (75) na linguagem de programação Prolog, analogamente à definição de descendente de Blackburn, Bos e Striegnitz (2006, p. 52)⁶. Parafraçando (75) informalmente, traduzindo POSS como *ter*, x tem mediatamente y (i) se x tem y ou (ii) se z tem y e x tem mediatamente z. Reformulamos, portanto, a SF de *ter*₂ em (47) como (76), que cobre todos os casos de (72)-(74), por exemplo, em (73), de *POSS(ônibus, vidros)* e *POSS(luverdense, ônibus)*, decorre *MPOSS(luverdense, vidros)*. A aplicação de (76) a (34) em

(77) evidencia que essa sentença, contrariando Lunguinho (2016), é, na verdade, gramatical, por aceitável numa situação em que subsistisse POSS(p,m), como em (78).

(75) $mposs(X,Y):- poss(X,Y).$
 $mposs(X,Y):- poss(Z,Y), mposs(X,Z).$

(76) $\lambda Q\lambda y\lambda x[MPOSS(x,y)\&Q(y)]$

⁶ Em Prolog, “:-” corresponde ao condicional se.

(77) MPOSS(p,c)&PAGAR(a,c)&POSS(m,c)

(78) As contas da Maria do Paulo foram pagas pela Ana.

A representação (76) contempla também casos como (24) e (69)-(71). Aplicando-a na composição do significado de (24), obtemos (79), em que abreviamos os nomes dos argumentos e decomponemos ACRESCENTAR. Como vimos, representações da forma $c1 \& c2$ não explicitam a precedência (<) entre as eventualidades denotadas por cada proposição. Em (46), temos $c1 < c2$. Em (79), pelo contrário, temos $c2 < c1$, porque, preenchida a condição POSS(byron,a.pena), MPOSS(byron,5.anos) só se verifica com POSS(a.pena,5.anos).

(79) $\exists x[\text{MPOSS}(\text{byron}, 5.\text{anos}) \& \text{CAUSE}(x, \text{BECOME}(\text{POSS}(\text{a.pena}, 5.\text{anos})))]$

Nossa proposta contempla também ocorrências como (80)-(82), que não são passíveis de derivação por meio de alçamento de possuidor. Com efeito, carecem de configuração $[_{DP} [_{DP} \text{possuidor}] \text{possuído}]$ do qual o possuidor possa ser extraído nos moldes de (33), subjacente a toda passiva possessiva, segundo Lunguinho (2013). O SUBJ nessa variante construcional não é um possuidor correferente a uma *ec*, como em (69), que possa ser substituída por um possessivo, como em (70) ou (71). A representação de (82) em (83) evidencia a analogia com (24). Em ambos os casos, subsiste $c2 < c1$. A única diferença é que MPOSS(d3048,p) denota posse imediata, decorrendo de POSS(d3048,p), conforme a primeira regra de (75).

(80) [...] o Hino [...] teve adicionados os versos de Ovídio Saraiva [...]. (Google)

(81) LISP [...] teve adicionados recursos de linguagens imperativas [...]. (Google)

(82) O Decreto n. 3.048 [...] teve um parágrafo acrescentado pelo Decreto n. 5.545 [...]. (Google)

(83) MPOSS(d3048,p) & CAUSE(d5545,BECOME(POSS(d3048,p)))

A sétima questão diz respeito à alegada afetação do referente do SUBJ de *ter*₂. Enquanto a construção coreana equivalente implica que o sujeito é adversamente afetado, em japonês isso só ocorre na passiva indireta, que exclui relação de posse entre os referentes do SUBJ e do OBJ. A representação semântica que propomos em (47) prediz, contrariamente a Lunguinho (2016), que não há necessariamente esse tipo de implicação, o que se comprova por (84) e (85). Se há afetação ou não, depende do participio, em interação com fatores pragmático-discursivos. Essa é uma característica em comum com a passiva canônica, que não expressa, necessariamente, afetação do paciente ou tema (CANÇADO; AMARAL, 2016).

(84) Cada dente teve a largura mensurada por três vezes. (Google)

(85) A ex-primeira-dama Marisa Leticia teve o nome excluído da ação após a sua morte. (Google)

A última questão refere-se à realização do POSSR como possessivo, alternando com a sua não realização; comparem-se (23), (39) e (69) com (27), (28), (40) e (71), o que parece contrariar a análise de Lunguinho (2011, 2013). De fato, se o alçamento do possuidor é

obrigatório na passiva não canônica, como explicar que, como possessivo, concomitantemente forme constituinte com o POSSM, configuração restrita à passiva canônica? Na nossa proposta, que prescinde do movimento de constituintes, esse problema inexistente. O verbo ter_2 codifica relação MPOSS entre o SUBJ e o OBJ, e o possessivo codifica relação POSS entre esses dois argumentos ou entre o SUBJ e uma outra função gramatical.

5 CONCLUSÃO

A passiva possessiva constitui um dos fenômenos mais discutidos das línguas do leste asiático. Em coreano e japonês, por exemplo, que formam a passiva morfologicamente, essa diátese distingue-se pelo argumento adicional como SUBJ, a preservação do OBJ e a relação de posse entre eles. Este trabalho propõe uma abordagem dessa construção em português, que, abstraindo da formação perifrástica da passiva, compartilha essas características.

Para investigar as propriedades gramaticais e semânticas da passiva possessiva do português, formulamos oito questões a partir das características das correspondentes coreana e japonesa, conforme Oshima (2003; 2004), e da proposta de Lunguinho (2011; 2013; 2016), aparentemente a única sobre o fenômeno em português. Sob a perspectiva de duas teorias formais lexicalistas, a LFG e a LDG, propomos soluções ancoradas em ocorrências em textos autênticos. Enquanto a primeira teoria fundamenta a descrição dos aspectos gramaticais, a segunda embasa a descrição da interface sintaxe-semântica por meio da decomposição de predicados.

Argumentamos, sob a ótica da LFG, que existem pelo menos duas variantes de *ter* em português: ter_1 biargumental e ter_2 da passiva possessiva. Este é analisado por Lunguinho como um auxiliar atípico, resultado da concatenação de v^* e *ser*, núcleo aquele responsável pelo argumento externo de ter_2 , que seria afetado, e o acusativo do argumento interno do participípio. Ao contrário, defendemos que ter_2 é um verbo de controle do objeto do tipo *equi*, cujo XCOMP é realizado pelo participípio passivo, analogamente à análise de *ser* da passiva comum como verbo de alçamento do sujeito. Nossa proposta é consistente com ocorrências em que o sujeito de ter_2 não é afetado, impossíveis para Lunguinho, que também exclui, entre outras variantes posicionais abonadas em textos autênticos, exemplos em que o XCOMP precede o OBJ. Na nossa proposta, esses dois complementos de ter_2 são passíveis de permutação, tal como noutros verbos do tipo *equi*.

Com base na LDG, propomos que ter_2 resulta da aplicação, sobre ter_1 , da operação ARG, que adjuge, ao predicado de entrada, um predicado adicional, subjazendo a diversas diáteses verbais. Admitindo $\lambda y \lambda x [\text{POSS}(x,y)]$ como representação semântica de ter_1 , propomos, inicialmente, $\lambda Q \lambda y \lambda x [\text{POSS}(x,y) \& Q(y)]$ como representação de ter_2 , da qual derivam a moldura de subcategorização e a equação de controle do objeto que caracterizam esse lexema. Dessa representação, decorre também a atribuição dos papéis POSSR e POSSM ao SUBJ e ao OBJ, respectivamente, pelo que o SUBJ não é necessariamente afetado. Por conta de POSS, o leque de relações semânticas entre SUBJ e OBJ é o mesmo de ter_1 e de outros elementos possessivos. Dessa forma, as propriedades gramaticais e semânticas distintivas da passiva possessiva em relação à comum não decorrem de elementos que não constituem palavras, o que violaria o PIL, mas da integração do participípio passivo numa estrutura bipredicacional com *ter* em vez de *ser*.

Na nossa proposta, o participípio da passiva possessiva é o mesmo da comum, sendo, portanto, computacionalmente mais econômica que as alternativas que codificam as propriedades distintivas dessa construção no participípio, as quais exigem a duplicação de participípios passivos para cada verbo apassivável.

Para contemplar também casos em que o POSSM é realizado por função outra que não o OBJ de ter_2 , substituímos POSS na representação desse verbo por MPOSS, que expressa posse mediata. Essa solução cobre igualmente casos em que não há uma configuração de possuidor e possuído do qual o primeiro possa ser extraído e que, portanto, não são deriváveis pela abordagem de Lunguinho. Oferece também uma explicação para a realização concomitante do POSSR como possessivo, que, ao nosso ver, constitui um problema para a abordagem de Lunguinho.

Dado o caráter completamente formalizado da LFG, a análise gramatical foi implementada no âmbito da BrGram (ALENCAR, 2013). Isso permitirá, por um lado, testar a análise automaticamente em um grande volume de dados, como pretendemos num

trabalho próximo. Por outro, representa uma potencial contribuição para aplicações de PLN direcionadas à compreensão textual, como IE e Q&A.

Várias questões tiveram de ficar de fora deste artigo, que focou os aspectos gramaticais e semânticos da passiva possessiva. Por exemplo, as funções discursivas dessa diátese, que, ao promover a SUBJ o possuidor de argumento de verbo do qual não é participante, constitui estratégia para conferir-lhe status de tópico. Outras questões interessantes relacionam-se aos aspectos conceituais, históricos, dialetais, tipológicos e tradutológicos.

REFERÊNCIAS

ALENCAR, L. F. de. BrGram: uma gramática computacional de um fragmento do português brasileiro no formalismo da LFG. In: BRAZILIAN SYMPOSIUM IN INFORMATION AND HUMAN LANGUAGE TECHNOLOGY – STIL, 9., 2013. Fortaleza. *Proceedings...* Fortaleza: Sociedade Brasileira de Computação, 2013. p. 183-188.

BLACKBURN, P.; BOS, J.; STRIEGNITZ, K. *Learn Prolog Now!* Londres: College Publications, 2006.

BOBROW, D. G. et al. PARC's bridge and question answering system. In: GRAMMAR ENGINEERING ACROSS FRAMEWORKS, 2007. *Proceedings...* Stanford: CSLI, 2007. p.46-66.

BRESNAN, J. *Lexical-functional syntax*. Malden: Blackwell, 2001.

BUSSMANN, H. (Org.). *Lexikon der Sprachwissenschaft*. 3.ed. Stuttgart: Kröner, 2002.

CANÇADO, M.; AMARAL, L. *Introdução à semântica lexical*. Petrópolis: Vozes, 2016.

CHAMBERS, N.; JURAFSKY, D. Template-based information extraction without the templates. ASSOCIATION FOR COMPUTATIONAL LINGUISTICS, 49., 2011. *Proceedings...* Stroudsburg: Association for Computational Linguistics, 2011. p. 976-986.

CROUCH, D. et al. *XLE documentation*. Palo Alto: Palo Alto Research Center, 2011. Disponível em: http://www2.parc.com/isl/groups/nltt/xle/doc/xle_toc.html. Acesso em: 22 fev. 2016.

FALK, Y. N. *Lexical-functional grammar: an introduction to parallel constraint-based syntax*. Stanford: CSLI, 2001.

HUANG, C.-T. J. Chinese passives in comparative perspective. *The Tsing Hua Journal of Chinese Studies*, Hsinchu, v. 29, n. 4, p. 423-509, 1999.

KIRSTEIN-JOST, S. *Auflösung von Anaphern im Rahmen der Informationsextraktion für Ontologie-Management im Bereich Life Sciences*. Hamburg: Dr. Kovač, 2010.

KOTSEV, E. *Interaktive Dialogsysteme: Sprachwissenschaftliche Grundlagen und Untersuchungen zu Nutzungsmöglichkeiten*. Saarbrücken: VDM Verlag, 2010.

LUNGUINHO, M. V. S. *Verbos auxiliares e a sintaxe dos domínios não-finitos*. 2011. 225 f. Tese (Doutorado em Linguística) – Universidade de São Paulo, São Paulo, 2011.

_____. *Construções participiais com o verbo 'ter'*. [S.l.]: ANPOLL, 2013. Disponível em: <http://anpoll.org.br/gt/teoria-da-gramatica-gttg/wp-content/uploads/sites/14/2013/06/Marcus-Lunguinho.pdf>. Acesso em: 25 set. 2017.

_____. *Applicatives and 'ter' passives*. Workshop de Linguística Formal, Brasília, 2016. *Abstract booklet...* Brasília: Universidade de Brasília, 2016. p. 12-14.

MCCORD, M. C.; MURDOCK, J. W.; BOGURAEV, B. K. Deep parsing in Watson. *IBM Journal of Research and Development*, Armonk, v. 56, n. 3/4, p. 1-15, 2012.

MEHLER, A.; LOBIN, H. Aspekte der texttechnologischen Modellierung. In: _____ (Org.). *Automatische Textanalyse: Systeme und Methoden zur Annotation und Analyse natürlichsprachlicher Texte*. Wiesbaden: VS Verlag für Sozialwissenschaften, 2004. p. 1-21.

MORAIS, M.A.C.R.T. *A passivização no português: uma abordagem léxico-funcional*. 1988. 181 f. Dissertação (Mestrado em Letras) –Universidade Federal do Paraná, Curitiba, 1988.

NOVICHKOVA, S; EGOROV, S.; DARASELIA, N. MedScan: a natural language processing engine for MEDLINE abstracts. *Bioinformatics*, Oxford, v. 19, n. 13, p. 1699-1706, 2003.

OSHIMA, D. Y. Out of Control: A Unified Analysis of Japanese Passive Constructions. INTERNATIONAL CONFERENCE ON HPSG, 9., 2003. *Proceedings...* Stanford: CSLI, 2014. p. 245-265.

_____. *Adversity and Korean/Japanese passives: constructional analogy*. [Nagoya]: Nagoya University, 2004. Disponível em: <http://www.gsid.nagoya-u.ac.jp/oshima/docs/kjpassives.pdf>. Acesso em: 24 maio 2017.

PATEJUK, A.; PRZEPIÓRKOWSKI, A. In favour of the raising analysis of passivisation. INTERNATIONAL LEXICAL-FUNCTIONAL GRAMMAR CONFERENCE, 19., 2014. *Proceedings...* Stanford: CSLI, 2014. p. 461-481.

PUSTEJOVSKY, J. *The Generative Lexicon*. Cambridge: MIT Press, 1995.

ROBINSON, K. S. *Aurora*. Londres: Orbit, 2015.

SANTOS, A.F. *Uma gramática LFG-XLE para a análise sintática profunda do português*. 2014. 178 f. Tese (Doutorado em Linguística) - Universidade Federal do Ceará, Fortaleza, 2014.

STERNEFELD, W. *Syntax: eine morphologisch motivierte generative Beschreibung des Deutschen*. Tübingen: Stauffenburg, 2006. 2 v.

WATSON (computer). In: WIKIPEDIA. [S.l.]: [s.n.], 2017. Disponível em: [https://en.wikipedia.org/wiki/Watson_\(computer\)](https://en.wikipedia.org/wiki/Watson_(computer)). Acesso em: 16 dez. 2017

WITT, J. Kompositionalität und Regularität im System der Partikelverben mit *ein-*. In: Olsen, S. (Org.). *Semantische und konzeptuelle Aspekte der Partikelverbbildung mit ein-*. Tübingen: Stauffenburg, 1998. p. 27-103.

WUNDERLICH, D. Predicate composition and argument extension as general options: a study in the interface of semantic and conceptual structure. In: STIEBELS, B.; _____. (Org.). *Lexicon in Focus*. Berlin: Akademie Verlag, 2000. p. 247-269.

AGRADECIMENTOS

Somos grato a Jessé de Sousa Mourão pelos esclarecimentos a respeito de questões de tradução. Agradecemos também aos pareceristas e revisores anônimos pelos valiosos comentários e sugestões sobre versão prévia deste artigo. Todos os erros remanescentes são de nossa responsabilidade.



Recebido em 15/01/2018. Aceito em 26/02/2018.