

Revista Linguística

volume 12 número 1 janeiro 2016

Organizadora da edição:

Lilian Ferrari

Esta edição da Revista Linguística é o terceiro número integralmente dedicado à Linguística Cognitiva. Se os volumes anteriores (2006 e 2010) reuniram trabalhos que representaram incursões iniciais de análise sob a ótica cognitivista, o volume atual apresenta resultados mais sólidos de pesquisa. Assim, os trabalhos aqui publicados refletem, em maior ou menor grau, a considerável expansão que se verificou no campo, envolvendo o aprofundamento teórico, o refinamento dos procedimentos metodológicos e a ampliação do escopo do objeto de investigação.

ISSN: 2238-975X

UFRJ

Comissão Editorial

Editor Responsável

Aniela Improta França, UFRJ/CNPq, Brasil

Comitê Editorial

Aniela Improta França, Pós-Ling, UFRJ | Aleria Lage, Pós-Ling, UFRJ |
Alessandro Boechat de Medeiros, Pós-Ling, UFRJ | Cecília Mollica, Pós-
Ling, UFRJ | Christina Abreu, Pós-Ling, UFRJ | Marcus Antonio Rezende
Maia, Pós-Ling, UFRJ

Conselho Editorial

Andrew Nevins | Universidade Federal do Rio de Janeiro, Brasil
Cilene Rodrigues | Pontifícia Universidade Católica Rio, Brasil
Fernanda Ferreira | Michigan State University, EUA
Gabriel de Ávila Othero | Universidade Federal do Rio Grande do Sul, Brasil
Gabriela Matos | Universidade de Lisboa, Portugal
Kees Hengeveld | Universidade de Amsterdam
Letícia Sicuro Corrêa | Pontifícia Universidade Católica Rio, Brasil
Marcus Maia | Universidade Federal do Rio de Janeiro, Brasil
Lilian Ferrari | Universidade Federal do Rio de Janeiro, Brasil
Luiz Amaral | University of Massachusetts, EUA
Maria Armanda Costa | Universidade de Lisboa, Portugal
Maria Luiza Braga | Universidade Federal do Rio de Janeiro, Brasil
Miriam Lemle | Universidade Federal do Rio de Janeiro, Brasil
Roberto Almeida | Concordia University, Canada
Ruth Elisabeth Vasconcellos Lopes | Universidade Estadual de Campinas, Brasil
W. Leo Wetzels | Universidade Livre de Amsterdam, Holanda

Organizadores da Edição

Aniela Improta França | Universidade Federal do Rio de Janeiro, Brasil
Lilian Ferrari | Universidade Federal do Rio de Janeiro, Brasil

Redação e Assinaturas

Programa de Pós-Graduação em Linguística - Faculdade de Letras/UFRJ
Endereço Postal: Av. Jequitibá, 2151, Ilha do Fundão - Cidade Universitária
CEP: 21941-590 - Rio de Janeiro (RJ) e-mail: posling@letras.ufrj.br

Layout da Capa

Camila Rodrigues | LabLab Design
Endereço eletrônico: <http://www.lablab.com.br>

Design e Diagramação

Helena Gomes Freire
Rafael Laplace | IGEAD
Endereço eletrônico: <http://www.igead.com.br>

Edição

Programa de Pós-Graduação em Linguística - UFRJ

Revista **Linguística**

Volume 12 Número 1 Janeiro 2016

UFRJ

Programa de Pós-Graduação em Linguística
Faculdade de Letras UFRJ

Sumário

- I** **Apresentação**
Por Lilian Ferrari
- 9** **Entrevista - George P. Lakoff (Professor of Cognitive Science and Cognitive Linguistics at the University of California)**
Para Helen de Andrade Abreu (UFRJ/UC Berkeley)
- 17** **Entrevista - Leonard Talmy (Professor Emeritus of Linguistics at the University at Buffalo)**
Para Helen de Andrade Abreu (UFRJ/UC Berkeley)
- 21** **Resenha - Vyvyan Evans. The Language Myth: Why language is not an instinct? Cambridge: Cambridge University Press. 2014. Págs. 304. ISBN 978-1-107-61975-3**
Por Hanna J. Batoréo (Universidade Aberta/ CLUNL)
- 33** **Conformity and beneficiary meanings of Latin *secundum* NP: a corpus-based cognitive analysis**
Por Caterina Guardamagna (Lancaster University)
- 51** **LUDI: um *framework* para desambiguação lexical com base no enriquecimento da semântica de frames**
Por Ely Edison da Silva Matos (UFJF), Maria Margarida Martins Salomão (UFJF)
- 78** **The linguistic-computational modeling of metonymy in the *framenetbrasil* lexical database: a case study**
Por Maucha Andrade Gamonal (UFJF), Tiago Timponi Torrent (UFJF)
- 91** **Metáfora e metonímia em nomes compostos em português: um estudo de construções S-ADJ**
Por Margarida Basilio (PUC-RJ)
- 106** **Uma abordagem construcional para as formações *x-dromo* do português brasileiro**
Por Carlos Alexandre Gonçalves (UFRJ), José Augusto de Oliveira Pires (UFRJ)
- 127** **Formas livres e formas presas: um clássico revisitado com olhar cognitivista**
Por Janderson Lemos de Souza (UNIFESP)
- 143** **Mesclagem conceptual em piadas curtas**
Por Elyssa Soares Marinho (UFRJ), Lilian Ferrari (UFRJ)

157 **O papel dos gestos de apontar na construção da dêixis multimodal: dos usos concretos aos usos abstratos**
Por Maira Avelar (UESB)

173 **A contraparte prosódica em construções de discurso reportado**
Por Patrícia Ribeiro do Valle Coutinho (UFJF), Luiz Fernando Matos Rocha (UFRJ)

APRESENTAÇÃO

Lilian Ferrari (UFRJ)

A Revista Linguística tem testemunhado, em sua trajetória, a crescente vitalidade que se verifica na área, através de volumes que buscam contemplar eixos teóricos norteadores da pesquisa desenvolvida no cenário nacional e internacional. Nos últimos anos, além de publicar contribuições atreladas a vertentes com larga tradição de pesquisa, tais como Funcionalismo, Teoria da Variação e Mudança Linguística, Teoria Gerativa, Psicolinguística e Aquisição de Linguagem, a revista têm incluído contribuições relativas a áreas mais recentemente consolidadas, tais como os Modelos Funcionais Baseados no Uso e a Linguística Cognitiva.

Essa última vertente é foco do presente volume, que constitui o terceiro número da Revista Linguística integralmente dedicado à Linguística Cognitiva. Se os volumes anteriores (2006 e 2010) reuniram trabalhos que representaram incursões iniciais de análise sob a ótica cognitivista, o volume atual apresenta resultados mais sólidos de pesquisa. Assim, os trabalhos aqui publicados refletem, em maior ou menor grau, a considerável expansão que se verificou no campo, envolvendo o aprofundamento teórico, o refinamento dos procedimentos metodológicos e a ampliação do escopo do objeto de investigação.

Em primeiro lugar, considerando-se que o aprofundamento teórico sempre esteve na pauta da Linguística Cognitiva, não deve causar espanto que os arcabouços teóricos fundadores tenham continuado a ser desenvolvidos, como é o caso da Gramática Cognitiva (Langacker, 2008, 2009, Talmy, 2000), Teoria da Metáfora Conceptual (Lakoff, 1987, Lakoff & Johnson, 1999), Teoria dos Espaços Mentais (Fauconnier & Turner, 2002; Turner, 2014), Gramática de Construções (Goldberg, 1995, 2006, 2013), dentre outros. Com quase trinta anos de existência, já é possível fazer uma avaliação dos conhecimentos reunidos a partir de pesquisas na área, bem como vislumbrar novos caminhos a serem percorridos.

Para fazer esse balanço, ninguém melhor do que aqueles que tiveram atuação determinante para a constituição do campo, contribuindo para o desenvolvimento da Semântica Cognitiva e fornecendo novos enquadres para a descrição e análise das relações entre forma e significado. Sendo assim, publicamos neste volume duas entrevistas inéditas de dois *'founding fathers'* da Linguística Cognitiva,

concedidas à Helen de Andrade Abreu, na *University of California, Berkeley*. A primeira entrevista foi realizada com George Lakoff, referência emblemática no que se refere à Teoria da Metáfora Conceptual, cognição corporificada (*embodied cognition*) e aplicação de achados sobre categorização à análise linguística. A segunda entrevista foi realizada com Leonard Talmy, professor emérito da *State University of New York at Buffalo*, que desempenhou papel inestimável na consolidação de uma tipologia semântica associada à codificação linguística do espaço e de eventos relacionados ao movimento, entre outras contribuições relevantes.

Ainda na esteira dos aprofundamentos teóricos, este volume traz a resenha crítica da Profa. Dra. Hanna Batoréo, da Universidade Aberta, Lisboa sobre o livro *The language myth: why language is not an instinct?*, de Vyvyan Evans. Trata-se de uma resenha detalhada e elucidativa dos principais pontos abordados no livro, que discute e questiona, a partir de evidências empíricas, o conceito de inatismo, classicamente relacionado à faculdade da linguagem.

No que se refere ao avanço metodológico, o volume traz estudos que propõem análises baseadas em corpora, em sintonia com a tendência recente de ampliação de escopo e refinamento de dados a serem utilizados na análise linguística. O artigo *Conformity and beneficiary meanings of latin secundum NP: a corpus-based cognitive analysis*, de Caterina Guardamagna, da *University of Liverpool/UK*, enfoca a construção latina [secundum SN], propondo uma análise construcionista baseada em corpus diacrônico.

Na sequência, encontram-se dois artigos vinculados ao Projeto FrameNet Brasil, que constitui um Laboratório de Linguística Computacional, sediado na Universidade Federal de Juiz de Fora, em cooperação com o Projeto FrameNet original, desenvolvido no *International Computer Science Institute*, da *University of Berkeley, California*. Os artigos evidenciam pesquisa pioneira para o português brasileiro, que busca aplicar a ‘*Semântica de Frames*’ à lexicografia prática. O primeiro artigo, intitulado “LUDI: um *framework* para desambiguação lexical com base no enriquecimento da semântica de frames”, de Ely Edison da Silva Matos e Maria Margarida Martins Salomão, debruça-se sobre a desambiguação de Unidades Lexicais, tendo como resultado uma aplicação computacional chamada Framework LUDI (*Lexical Unit Discovery through Inference*), composta por algoritmos e estruturas de dados usados na desambiguação. O artigo seguinte, intitulado *The linguistic-computational modeling of metonymy in the FrameNet Brasil lexical database: a case study*, de Maucha Andrade Gamonal e Tiago Timponi Torrent, propõe uma modelagem linguístico-computacional para formular estratégias teórico-metodológicas para desambiguação metonímica, baseando-se em um estudo de caso com o frame *Teams*.

A metonímia também é enfocada no artigo “Metáfora e metonímia em nomes compostos em português: um estudo de construções S-ADJ”, de Margarida Basílio, que analisa o papel da metáfora e da metonímia na semântica de construções lexicais compostas de forma S-Adj no Português do Brasil. O trabalho discute questões relativas à distinção metáfora/metonímia e à dificuldade de análise de construções lexicais tradicionalmente consideradas como compostos.

O artigo “Uma abordagem construcional para as formações x-dromo do português brasileiro”, de Carlos Alexandre Gonçalves e José Augusto de Oliveira Pires, adota a Morfologia Construcional

para realizar uma descrição detalhada tanto do polo formal quanto do polo semântico da construção morfológica x-dromo e investigar a produtividade da mesma no atual estágio da língua. Ainda no âmbito da morfologia, o artigo “Formas livres e formas presas: um clássico revisitado com olhar cognitivista”, de Janderson Lemos de Souza, revisita a distinção entre formas livres e formas presas, formulada pelo estruturalismo, a partir da perspectiva teórica da Linguística Cognitiva.

O detalhamento de processos cognitivos associados à construção do significado é o foco do artigo “Mesclagem conceptual em piadas curtas”, de Elyssa Soares Marinho e Lilian Ferrari, em que se estabelece uma análise dos elementos linguísticos que contribuem para a emergência de sentido em piadas curtas, do tipo adivinhas. As autoras propõem uma articulação entre o processo de Mudança de *Frame*, já apontado na literatura como responsável pelo efeito cômico (Coulson, 2001), e o processo de Mesclagem Conceptual (Fauconnier & Turner, 2002), demonstrando que as piadas analisadas partem da similaridade fonética entre duas construções para acessar domínios distintos na rede de espaços que configuram a mesclagem conceptual.

Os dois artigos finais representam perspectivas de análise que vêm ganhando terreno na área, contemplando a investigação de aspectos multimodais e suprasegmentais. No artigo “O papel dos gestos de apontar na construção da dêixis multimodal: dos usos concretos aos usos abstratos”, Maíra Avelar discute a inter-relação entre a fala e os gestos de apontar na construção da dêixis multimodal. Já o artigo “A contraparte prosódica em construções de discurso reportado”, de Patrícia Ribeiro do Valle Coutinho e Luiz Fernando Matos Rocha, analisa tendências prosódicas de construções de discurso reportado em amostras de *corpus* de fala espontânea, submetidas ao programa PRAAT (Boersma & Weenink, 2011).

Em linhas gerais, os estudos reunidos neste volume ilustram não apenas esforços de desenvolvimento teórico, refinamento metodológico e expansão do objeto de estudo, como também podem ensejar desdobramentos de pesquisa que contribuam para a expansão quantitativa e qualitativa das pesquisas em Linguística Cognitiva.

entreviSta GEORGE P. LAKOFF

Helen de Andrade Abreu (UFRJ/UC Berkeley/Bolsista da CAPES – Proc. no BEX 6928/15-3)



George Lakoff is one of the most widely recognized and celebrated scholars in the field of Cognitive Linguistics in the world today. And with good reason, because Lakoff not only helped pioneer this field from its earliest beginnings, but also he has guided and shaped its development down to the present day. Currently Professor of Cognitive Science and Cognitive Linguistics at the University of California at Berkeley, where he has taught since 1972; he has previously taught at both Harvard University and University of Michigan. George Lakoff is a man of far ranging intellectual interests who has demonstrated a remarkable talent for applying the concepts of Cognitive Linguistics to other fields of academic discipline, including: Political Science, Literature, Mathematics, Philosophy and Psychology. Within the field of Cognitive Linguistics, he is renowned for his work with: the embodied mind, how metaphorical thought works, and how meaning structures language. A gifted speaker, Professor Lakoff has lectured to audiences around the world and has frequently appeared on television and radio programs where his ability to explain in clear and ordinary language some of the most profound concepts of his academic discipline have allowed him to reach an audience of ordinary citizens as well as academic scholars. A prolific author, Professor Lakoff has written

hundreds of articles and numerous books, including a New York Times best seller entitled *Don't Think of an Elephant*, which applies the concept of framing, from Cognitive Linguistics, to political communication strategies. Many of his works are considered “classics” within the field, and are frequently required reading for students in both undergraduate and graduate Cognitive Linguistics courses at universities all over the world. In addition to his work as scholar and writer, Lakoff has for decades provided leadership and support to a number of professional organizations, including: The Neural Theory of Language Project at the International Computer Science Institute at Berkeley, The International Cognitive Linguistics Association, and The Cognitive Science Society. Professor Lakoff is currently working with Sriniv Narayanan on a new book, soon to be published, which brings together Neuroscience, Neural Computation, Cognitive Linguistics and Experimental Embodied Cognition.

Revista Lingüística: You started your career at the MIT. Can you tell us about how you changed your trajectory and about what made you move from Generative Linguistics to Cognitive Linguistics?

George Lakoff: At the MIT I majored in mathematics and English literature, particularly poetry. Those are two things I love. In my last year, it was the beginning of Chomsky's department at the MIT. And to celebrate the beginning of the department, they got Roman Jakobson from Harvard to come and give a course on poetics. So, my advisor told me, “You have to take this course, but to understand what he is saying you have to take the Introduction to Linguistics course”. So I took the introductory course from a teacher who was one of the first generation of Chomsky's graduate students. And then I took this course from Jakobson, who was really who got me into linguistics. Jakobson, in talking about poetics, went through a lot of phonology. This was a course at the MIT, but I was the only MIT student there. The others were all from Harvard, coming from Slavic literature or something, to study with Jakobson. As I was an MIT student taking the basic linguistics course, I could actually read all the papers he assigned us for homework, and I was the only one to do all his homework. I also went to all his office hours [the time when teachers answer students' questions and try to solve their doubts]. I was also taking a course on Yeats, who was his favorite English poet. Jakobson would assign us homework doing analysis of poems he was teaching, and I was the only one who actually did them. One week he came in and said to me, “I will not be in class this week. Would you like to lecture on your analyses of Yeats?”. I was twenty years old and he asked me to lecture to all Harvard graduate students. So, my first lecture [the way they call classes at the universities in the USA] was substituting for Jakobson, when I was an undergraduate.

Then I took the course by Chomsky the next semester. It was a very strange course, because half was the mathematical foundation of his theory, which was coming out of a mathematical theory of recursive function theory. But I was studying mathematics, so it wasn't a big deal for me. The other people were about seventy students from Harvard and the six new graduate students from the MIT. I sat in the front, and whenever he said something that wasn't clear, or if he made a mistake I would raise my hand and ask, “Don't you mean so and so?” and he didn't like me very much [laugh out loud]. And the second half [of the classes] was on rules of stress in English. So I wrote a paper on the poetry of two great English poets who used special markers to indicate where stress should change. Chomsky didn't know about poetry and didn't care. But I got a good grade.

Then I went to Indiana to go to graduate school in English. And I didn't like it, because English graduate school in 1962 was not what I liked. At one point, I was teaching freshman composition and I got fired. Let me explain how it happened. It was 1962. On the week of the Cuban missile crisis, the assignment was to read an essay by George Orwell on the Spanish Civil War. No one in the class in Indiana had ever heard of the Spanish Civil War. It wouldn't be taught at high schools there. So I explained what the war was about, and that Orwell saw that the Communists were the heroes and the Nazi were the thugs. This explanation happened during the Cuban missile crisis. Someone wrote home to their parents and said that a communist was teaching the class. Then I got fired.

Meanwhile, I was taking the most advanced linguistics classes they had, because I'd come from the MIT where I'd learned all that stuff, and nobody there, even on the faculty, really understood what Chomsky was doing in 1962. So what happened was that while I was taking the class I was also teaching the class. Every class they would go through this thesis they were reading and they wouldn't understand, and the thesis would have mistakes and I would get up and correct their mistakes and explain what was going on. I basically taught the class.

I got to know the chairman of the department because he was interested in discourse, I was too, so I used to go there talk to him. When I got fired, I went there and talked with him, and he said, "I'm here because I got fired from a higher state last year for taking part in a civil rights demonstration". In 1962 that was what America was like. So, at this point, the phone rang, and it was somebody turning down a major fellowship to the Linguistics Department. So the chairman said, "Somebody turned down a fellowship. Would you like to come into the Linguistics Department?" I said, "OK" and I became a linguist [laugh out loud]. Twenty-five years later I found out who turned down the fellowship... Ron Langacker [more laughs]. Life is that way.

So I went into linguistics and I came in as a Chomskyan. A lot of people in the department didn't like Chomsky, or they thought they understood Chomsky and they didn't, and so on. So I got into a lot of trouble in the department, but everybody knew I was pretty smart, I had a fellowship, so it was fine. Then I got married to Robin and I went off to write my thesis at Cambridge, Massachusetts, while she was still at Harvard, and I got a summer job at the MIT making mechanical translation. They gave me the following job: they said, "We've just written a little program that will generate baseball games, because there are simple rules, and we can generate a random sequence of a random baseball game. We'd like you to take Generative Grammar and take the input as a baseball game and give the sentences as output". I said "OK". So I looked at the first sentence. It was about a [baseball] player who later got into the Hall of Fame. The sentence was "Has Shamsky doubled to the left?". How can you handle this in Generative Grammar? You couldn't, because in Generative Grammar you can only look at the symbols and the grammar. "To the left" is a phrase that indicates direction. There are constraints in grammar on directional adverbs that go with verbs of motion and they modify things that move in that direction. But if you say "Shamsky doubled to the left", what moved to the left field is the ball, and not Shamsky. But the ball is not even in the sentence. So, it's in the meaning of the sentence. And that makes you look at the meaning to know what direction adverbs go. And that is not possible in Chomsky's theory. So I said,

“Aha! We need to have semantics built into the theory”. And that was where Generative Semantics came from. That was in 1963, and I wrote a paper called *Toward Generative Semantics* and added some other examples. I showed it to Chomsky and he hated it.

In 1965 I finished my dissertation while I was working at IBM. IBM allowed people to go to Harvard and listen to lectures. So I went to a lecture at Harvard on Computational Linguistics. When I was in Indiana I got bored a lot, so I read everything on Computational Linguistics, which was about three books. [laugh out loud] In 1965, that’s all there was. So, I went to hear this talk, I knew everything about the talk. I raised my hand after the talk, and said, “I know that you’re working on the framework so and so, and within that framework this would make sense, but there are some counter examples”. And I gave him a list of ten counter examples. Susumu Kuno, who was running the project, turned around and looked at me very strangely. After each talk at Harvard, they had a sherry hour. That’s Harvard. So, I went to the sherry hour, and at that time, I had just gotten married and I had almost no money. I had the last part of my fellowship, I was getting my degree and I was going to lose my fellowship. So, I went to the sherry hour, and Susumu started asking me who I was, and bla-bla-bla, and then he said, “Are you interested in getting a job, doing some research?” And I said, “Yes, I am”. So, he had me come to his office. I went to his office, and told him about my dissertation, and so on, and he said, “Well, here is our last project report. Read it and come back. We’ll talk next week”. I read the project report and it was awful. They were doing traditional 1965 Natural Language Processing, not vey good. So, I went back to his office, and I said, “Look, I need the money, but I read your project and there are a lot of things wrong with this”. I showed him all the things I found needed to be worked on. He said, “Would you like a full time job?” And I said, “OK!” [laugh out loud] And I got hired to Harvard.

And I stayed there for four years, until it became politically impossible. Because I had convinced Haj Ross of Generative Semantics. When Chomsky went on leave, Haj Ross was hired as an assistant professor at the MIT, and we taught courses together at Harvard and the MIT, and when Chomsky got back he was furious, that Haj Ross had gone against his theory. And he started attacking me in his writings. And then I answered, because I wasn’t going to take it [laugh out loud]. That was when the Linguistics Wars started. He started attacking me, and I answered by writing a parody of him. He had no sense of humor, and he thought I was copying him. [laugh out loud] But, anyway, he used his influence to make sure that I got fired at Harvard. I had a choice, actually. If I could find another department that would give me tenure there, I could stay. Otherwise, I had to leave. I got an offer from the English department, and I could stay at the English department teaching style, but I didn’t want to stay at the Harvard English department, so I said “No”.

That day I got a call offering me jobs at two universities, Texas and Michigan, because Tom Denver, who I had argued with years before, was on the job market, and he was going to places where no one knew any Generative Linguistics. And he knew that, even if he disagreed with me, I knew Generative Linguistics. So, he said, “Hey, if you’re going to hire me, you need to hire someone who knows some Generative Syntax”. So they called up and they offered me jobs. And then I said, “I have to get something for my wife, who is getting her PhD at Harvard, and so Robin and I were both offered

jobs at the University of Michigan, and so we left for the University of Michigan, and it was very nice. And then I came to Berkeley, as the Chomskyan who didn't believe Chomsky [laughs]. They didn't believe Chomsky, but they needed someone who could teach it, and so I did. I came and taught introduction to syntax, semantics and logic and all those things, in 1972.

In 1975, I got a grant for a summer seminar at Berkeley for eight people who were in philosophy, logic and artificial intelligence, and I had the opportunity to hire some research assistants, and one of them, a graduate student then, and he had previously run rock bands. When he came out there, we formed the Underground Linguistics Institute, and he told a lot of people about this, and 188 people came. He rented out sorority houses and formed communes. The rules were that everybody could give a talk with three days' notice. We would listen to talks six days a week from 10:00 a.m. to 10:00 p.m. and Sundays at the beach. [laughs] So, that summer, I heard talks by Paul Kay on color, Charles Fillmore on Frame Semantics, Eleanor Rosch on Basic Level Categories and Leonard Talmy on Cognitive Primitive Space. And all of them required embodiment. And I got deeply depressed, because I saw that logic was not going to work. I had a terrible depression, and I went hiking with two friends for two weeks. I came back and I could no longer teach what I used to teach.

Then, a couple of years later, I discovered metaphor. In the meantime, I realized that this was about embodiment. Mark Johnson, when I was working on metaphor, happened to come through Berkeley. He was teaching in the Midwest, and he got a job teaching for one quarter at Berkeley at the Philosophy Department. On his way out, he happened to stop at the University of Chicago, where he had been a graduate student and he met Jim McCawley. He asked Jim if anybody at Berkeley was studying metaphor, because he [Johnson] had written a dissertation on the philosophy of metaphor, and Jim said that I was. I got a call from Mark, and it turns out he had an apartment three blocks from my house. He reached the same conclusions that I did about embodiment, and metaphor, and conceptual metaphor from a philosophical point of view. We ended up working together, first writing a paper and then doing *Metaphors We Live By*. And I found out about metaphor because I was teaching a graduate seminar, and one of the students in the class came in, I was reading a paper on metaphor and she said, "My boyfriend said this metaphor that really upset me and maybe you guys can understand it". That's where the LOVE IS A JOURNEY stuff came from. That is what happened. So, from 1975, with Fillmore's Frame Semantics and so on, I didn't understand what it meant, but it came through that it had to do with embodiment. When I did metaphors, it was clear that Frame Semantics and Neural Theory was necessary for color, that we had to have neural embodiment. And it changed philosophy and it changed all kinds of things, and it's in *Metaphors We Live By*.

Revista Linguística: How can we define Cognitive Linguistics? Some people have difficulties drawing a line between Cognitive Linguistics and other areas, like Usage-Based Models. What is the difference between these areas of Linguistics?

George Lakoff: For me, from the beginning, Cognitive Linguistics was neural, because of the color,

it was embodied, because of Eleanor Rosch's work on embodiment, and Len [Leonard] Talmy's embodiment and Chuck [Charles] Fillmore Frame Semantics. So, it had to have embodied cognition, it had to eventually have a neural theory, it had to include frames, conceptual primitives, and basic level and prototype categories. It was defined from the beginning, in 1975, from me, that way.

Ron [Ronald] Langacker had a different view: he was looking at various Native American languages and looking at spacial relations, just as Talmy had here, and he came up with spacial relations, and basic spacial relations for the basis of grammar in general. And he came up with something called Space Grammar, which I thought was a terrible name, because it sounded too "spacey", and I said, "Take Cognitive Grammar". I had previously written a paper called Cognitive Grammar, having to do with showing that you can have a grammar that was a processing grammar, but was also a structural grammar at the same time. I called it Cognitive Grammar because it was both processing and structural, whereas Chomsky had to separate them, and I showed you didn't have to do that. And I wrote this paper with Henry Thomson in 1975.

Langacker was a field worker, and so, he had this notion of a Usage Based Model, which was that you learned language by using it, which isn't false, but it missed all of the embodiment, and all of the neural stuff, and then when I discovered metaphor, he missed that too. But he kept his view of Cognitive Grammar going, and he is still doing it, and his Usage Based Model says if you use things you've just learned, you just somehow learn them and they become conventionalized, and so on. Whereas that, as you know, from the book that I'm doing with Sridhar [Sridhar Narayanan], is not how it works, at all. What happens is that there was a split between me and Langacker on those issues: one, embodiment; two, the neural basis; and three, metaphor. But I decided that, if we were going to have Cognitive Linguistics, we shouldn't be fighting. Langacker was very smart, he was really insightful, his book has great insights into the language, and I respect him and I use a lot of these insights, and I just went beyond doing other things. He didn't want to and he did what he did, and Talmy did what he did, and then, I had known Gilles Fauconnier for a long time, and I can tell you how Mental Spaces started.

It started from a paper I wrote in 1968 taking Possible Worlds Semantics and saying that they had to be cognitive, and they had to be part of your mind, based on McCawley's example, "I dreamed that I was Brigitte Bardot and I kissed me", not "myself". You can't say "I kissed me", but you can in that context, and what you're doing is imagining your mind and transfer to that. There had been a logic invented by David Lewis called Counterpart Theory, and I took his theory, and I wrote a paper called *Counterparts*, on how you would handle these sentences. And I said that Possible Worlds were not worlds, they were part of cognition, and that you have connections across them. And I gave one of the papers to Fauconnier, and he turned it into the Theory of Mental Spaces, very brilliantly. Fauconnier is an absolutely brilliant linguist. I knew him when he was a graduate student, and he was a good friend. So, when we started to form Cognitive Linguistics, those were the people I contacted, Talmy and so on. Chuck [Charles] Fillmore did *not* want to become a Cognitive Linguist at all. He was against it.

We were trying to get to work together on Construction Grammar, because we had the idea that meaning came in there, I had accepted his Frame Semantics, and he didn't accept the metaphor stuff at all. But we accepted Frame Semantics in Construction Grammar. And we were working on a notation for Construction Grammar in my backyard one day, and we had a disagreement about the notation. So, I said, "This notation should reflect how people think, it should reflect cognition", and he said, "No, it shouldn't". And I said, "Why?", and he said, "Construction Grammar is about lexicography. What you want to do is make linguistic descriptions that can be put in dictionaries and in written grammar, so that people can understand it easily and can be taught". And I said, "Wait a minute, language is about cognition, about the way you think, and he said, "No, it isn't. It's about words and how you put them together". And then, we just had two different versions of Construction Grammar. He never adopted Metaphor Theory or Cognitive Primitives or Image Schemas in his theory of Frame Semantics.

Revista Linguística: You mentioned your new book, co-authored with Srini Narayanan. Could you summarize some of its contributions to the field?

George Lakoff: That is a hard question, because every day I rewrite the preface. [laugh out loud] First, I asked Srini how he would describe it, and he said, "It's about the fact that we think with neural structures, neural circuitry, but thought is not in the brain. Thought has to be embodied, has to be part of the body, because the neural circuitry runs throughout your body. If it were just in the brain, you'd have nothing to think about. You couldn't have any meaning at all. Meaning has to come through the body, but not just that. Because it comes through the body, it has to connect to the physical world, because you get feedback in the physical world, and because you connect to other people. It's a physical and a social world, so meaning has to do with neural structure in the brain *and* the body as it reacts to, and functions in, the physical and social world". That is Srini's way of thinking about it, and it's true, but I don't think it's good enough.

I'm trying to think about a better explanation. There are several ways to think about it. One, is that ideas don't flow in the air, they have to come through neural circuitry, and the question is how it is possible. And the book explains it. That is an amazing thing to me. We have a theory of how it works – not everything, but a lot of it – and it shows something that we know from the study of frames and metaphors, that there are learned frames and metaphors. The conceptual system is learned before you learn language, but the same kind of neural mechanism works in language and in thought. They are linked together in a certain way, and we study how they are linked together. This has a lot of consequences. First, the theory we have uses neural computation as a bridge between Cognitive Linguistics and Neuroscience. And what it is doing is *modeling* certain things that a computational system can model. It models the functional computational properties of neural systems. There are metaphors mapping this computational system onto Cognitive Linguistics on the one hand and the neural system on the other, and that makes a bridge between the two. What are the computational properties? Every conceptual structure that you have is a circuit, and now the question is what the circuits look like, and we can model the structures of those circuits. Two, we can give a theory of how the activation of one circuit, one concept, can activate another one. Because they are connected

to each other, how activating certain things relate to other things and give you inferences and also sequential activations. Srinivasan has shown, also, that the motor system has a certain computational structure which is exactly the same as aspect [in language] and the structure of events and actions in the world, so that we could then use that as a model both of how the body works and how thought works and model many aspects of embodiment. We also show how Conceptual Integration could be modeled neutrally and a lot more.

The book is about what happens. What does meaning become? What does thought become? What does politics become? What does language look like when you think about it this way? And it looks very different. It doesn't look at all like traditional Cognitive Linguistics in 1975. It has changed a lot. There are kinds of neural circuitry that you can now characterize and explain certain things, and the biggest thing that can be explained is why you have the sources and targets you have in metaphors. And that is explained in terms of Spike Time Dependent Plasticity, which is a property of change at the individual neuron level. It is astonishing that something that happens at an individual neuron level can explain things like what the sources and targets in metaphors are. We've known for some time that 98% of thought is unconscious, and the reason to that is that it's all in this neural circuitry that you can't have access to consciously.

Out of all these studies, you come out with a very different understanding of what thought is. Thought is neural, it's hierarchical, and it's grounded in things in the body, it works by chemistry, and it has at least the following kinds of functional computational structures, and that explains all sorts of details of language. So, that is what the book is about.

REFERENCES (BIOGRAPHY):

George Lakoff. Blog. In: <http://georgelakoff.com/research/> Access on: January 10, 2016

George P. Lakoff. In: <http://www.linguistics.berkeley.edu/person/21> Access on: January 10, 2016

The elephant! Returns: the “father of framing” offers bold new strategies. In: Chelsie Green Blog. (11/08/2014) <http://www.chelseagreen.com/blogs/the-father-of-framing/> Access on: January 10, 2016.

Recebida 15/02/2016

Aceita 22/02/2016

entreviSta LEONARD TALMY

Helen de Andrade Abreu (UFRJ/UC Berkeley/Bolsista da CAPES – Proc. no BEX 6928/15-3)



Leonard Talmy is Professor Emeritus of Linguistics at the University at Buffalo, State University of New York, where he taught for 15 years and was Director of the Center for Cognitive Science for 14 years. He is now also a Visiting Scholar at the University of California, Berkeley, where he had received his Ph.D. in Linguistics. Over his career, he taught in Hamburg, Rome, and Moscow (the latter two as a Fulbright Fellow) as well as at Stanford, Georgetown and University of California, Berkeley. He did extended research at Stanford on the Language Universals Project, at the UCLA Neuropsychiatric Institute with language-impaired children, and at the University of California at San Diego in cognitive science at the Center for Human Information Processing. And he was the Coordinator of the Cognitive Science Program at the University of California at Berkeley for six years.

His broader research interests cover cognitive linguistics, the properties of conceptual organization, and cognitive theory. His more specific interests within linguistics center on natural-language semantics, including: typologies and universals of semantic structure; the relationship between

semantic structure and formal linguistic structures -- lexical, morphological, and syntactic; and the relation of this material to diachrony, discourse, development, impairment, culture, and evolution. Additional specializations are in American Indian and Yiddish linguistics.

He is the author of a two-volume set with MIT Press (2000): *Toward a Cognitive Semantics* -- volume 1: *Concept Structuring Systems*; volume 2: *Typology and Process in Concept Structuring*.

Previously published articles include *The Relation of Grammar to Cognition*, *Force Dynamics in Language and Cognition*, *How Language Structures Space*, *Fictive Motion in Language and 'Ception'*, *Lexicalization Patterns*, *The Representation of Spatial Structure in Spoken and Signed Languages: a Neural Model*, and *Recombination in the Evolution of Language*. He has also written the Foreword for the edited volume *Methods in Cognitive Linguistics* and the entry on *Cognitive Linguistics* for Elsevier's *Encyclopedia of Language and Linguistics*. And he is currently working on a book for MIT Press titled *The Attention System of Language*.

Virtually all his written work is available on his website, <http://linguistics.buffalo.edu/people/faculty/talmy/talmyweb/index.html> including his 1972 dissertation, his 2000 two-volume set with MIT Press and his articles published and in press since 2000.

He was the recipient of the Gutenberg Research Award for 2012 from the Johannes Gutenberg University of Mainz, Germany, for outstanding contributions to research in the area of linguistics. In 2011, he was honored as one of the three "Founding Fathers" of cognitive linguistics at the 10th Biannual Conference of the International Cognitive Linguistics Association. He was elected a Fellow of the Cognitive Science Society in its 2002 inaugural selection of Fellows (and had been a founding member of the Society). He is included in *Outstanding People of the 20th Century* and in *International Who's Who of Intellectuals, thirteenth edition*.

(Source: <http://linguistics.buffalo.edu/people/faculty/talmy/talmyweb/bioblurb.html>)

Revista Lingüística: Your doctoral dissertation focused on the representation of Motion events in different languages and has influenced the work of many other linguists since then. What were the motivations that led you to pursue this research?

Leonard Talmy: My Motion event typology began when I simply noticed how Spanish and English represented Motion events differently. Now, at that time, the Linguistics department favored working on American Indian languages. So, as part of my graduate student research, I went driving around looking for a Californian Indian language to analyze, and met an Atsugewi speaker with whom I worked. It turned out that that language had a third major way of representing Motion. This was a lucky accident, since this type was one of the rarer in the world. So I found a three-legged stool to base the typology on, and with this basis I could then refine the typology by looking at still other configurations that languages had, such as split systems, mixed systems, minimal systems, and so forth.

Revista Lingüística: In your opinion, what was the impact of your dissertation's findings, and of your initial work, in the formation and consolidation of Cognitive Linguistics?

Leonard Talmy: Well, I'm not a historian of Cognitive Linguistics, but my impression is that there was a gradual build-up of works, first mainly by me and George Lakoff, that provided the initial body that eventually consolidated into Cognitive Linguistics. Ron Langacker came in a bit later and further consolidated the field. I think the work of us three is the main starting impetus of the field.

Of my own work, probably the papers most influential in building up a critical mass for this new perspective on language were those involving spatial structure; figure and ground analysis; my early work on causation, which also showed up in my dissertation, and that eventually led to Force Dynamics; and of course the work on Motion event structure. Also central was a paper of mine called *The Relation of Grammar to Cognition*, which is essentially the semantics of grammar or of closed-class forms, and that I revised as the first chapter of my two-volume set.

Revista Lingüística: Taking into account Cognitive Linguistics' trajectory, how would you evaluate the area nowadays?

Leonard Talmy: Since the earliest days of Cognitive Linguistics, there has been a steady deepening in the analysis of areas it previously addressed, and a steady expansion of the areas that the methods of Cognitive Linguistics have been applied to. For example, we've seen deepening in the analysis of metaphor, blending, the cognitive basis of grammar, and schematic systems such as the ones I've worked on – configurational structure, perspective, attention, force, and cognitive state. And we've seen Cognitive Linguistic analysis extended to gesture and signed language.

A further development within Cognitive Linguistics has been increasing emphasis on empirical methods, which is great, because we need both the empirical and the theoretical sides. What counts as empirical methods includes Psycholinguistic experimentation; corpus research (especially when including a statistical analysis); videographic and audiographic analysis (sometimes frame by frame); maybe the old method of writing down protocols, for example, of children's utterances; and some might include simulations, as in artificial intelligence or computational linguistics. The model I look to is physics, which includes both theoretical physics and experimental physics, interacting with each other. Theoretical physics without input from empirical findings would be empty, producing theories about nothing that exists out there. On the other hand, empirical research without guidance by theories is blind. It wouldn't know what to look at next, nor what to make of its findings. So as long as there is an interaction of both aspects, that's the healthiest state a field can be in, and I include Cognitive Linguistics in that. A current risk, though, is that the emphasis on the empirical side might downplay the importance of the theoretical side, and if that were to happen, that would be an unfortunate development within the field.

Revista Lingüística: Would you tell us about the new book you are writing?

Leonard Talmy: The book I'm completing, which is going to be published by MIT Press, and should appear in 2017, is called *The Targeting System of Language*. Its aim is to unify deixis and anaphora into a single cognitive system that I call *targeting*. In both cases, the speaker wants to refer to something – her *target* – located near or far in either the speech-external or the speech-internal environment. She wants to refer to this target at a certain point in her discourse and get the hearer's attention on it there as well. But how can she bring this about? She can't just directly reach into the hearer's cognition, take hold of his attention, and place it on her selected target at the intended moment.

Language solves this problem with a certain *targeting procedure*. At the intended point in her discourse, the speaker places a *trigger* – a form like *this, that, here, there, now, then*, or a personal pronoun – the set of such triggers is actually quite extensive. This trigger then initiates – or “triggers” – a 3-stage process in the hearer. In the first stage, the trigger directs the hearer to look for certain cues to the target. There are ten categories of such cues, representing ten different sources of information. This part is the mainstay of the analysis – there is a chapter for each cue type. In the second stage, the hearer integrates the cues he has found and uses them to single out the one entity in the environment that most fits those cues. This should be the speaker's intended target. And in the third stage, he takes the concept of this target and maps it onto the trigger, so he can integrate it with the rest of the sentence's meaning.

Recebida 18/04/2016

Aceita 25/04/2016

RESENHA - VYVYAN EVANS. THE LANGUAGE MYTH: WHY LANGUAGE IS NOT AN INSTINCT?¹

Hanna J. Batoréo (Universidade Aberta/ CLUNL, Lisboa, Portugal)

O livro *The Language Myth: Why language is not an instinct?*, da autoria de Vyvyan Evans e publicado pela Cambridge University Press, em 2014, é um bestseller linguístico no mundo inteiro, que, desde o momento do seu lançamento, tem vindo a causar um aceso debate nos meios linguísticos, tanto nos sites da *internet* como ao nível das publicações especializadas da área de Linguística. Logo a seguir à publicação do livro, quando surgiram as primeiras reacções às ideias nele apresentadas, duas grandes revistas mundiais da área – a *Lingua* e a *Language* – convidaram o Autor a responder às acesas críticas que as suas convicções suscitaram. A primeira das revistas publicou as críticas e a respectiva resposta de Evans (Adger 2015a; Behme & Evans 2015; Adger 2015b); a segunda, cuja publicação estava prevista em Dezembro de 2015, foi adiada para Março de 2016, prevendo-se que seja restringida apenas às críticas apresentadas ao livro, sem incluir a inicialmente convidada resposta do Autor, publicada entretanto no site do próprio como uma “resposta aberta” (Evans 2015b; 2016), mantendo-se, assim, o debate presentemente aceso.

O Autor do livro, Vyvyan Evans, que foi considerado uma “estrela do século 21” pelo *21st Century Science Magazine*, é actualmente professor de Linguística na School of Linguistics and English Language em Bangor, País de Gales (Reino Unido). Doutorou-se em Linguística na Universidade de Georgetown (Washington, D. C., nos Estados Unidos), em 2000 e, desde então, publicou uma dúzia de livros e muitos materiais mediáticos (entrevistas, vídeos, etc.) sobre diferentes aspectos da linguagem, a sua evolução, a sua relação com a mente, a importância da linguagem na comunicação humana e construção do sentido. Em finais de 2015, foi publicado o seu livro mais recente (Evans 2015c) – *The Crucible of Language: How Language and Mind Create Meaning*.

A publicação aqui em destaque – *The Language Myth: Why language is not an instinct?* – é um livro complexo que permite várias leituras académicas. Por um lado, pode ser considerado, numa primeira

1 Cambridge: Cambridge University Press. 2014. Págs. 304. ISBN 978-1-107-61975-3

leitura, apenas um livro erudito que analisa os mitos linguísticos, tal como vários outros o fizeram antes dele (cf. Bauer & Trudgill 1998). Por outro lado, no entanto, e explicando a grande popularidade que imediatamente ganhou, bem como a controvérsia e o impacto que sofreu, é considerado um livro único no meio académico, que, ao abordar e debater o maior mito linguístico de sempre – o da linguagem como instinto, inerente à teoria da Gramática Universal de Chomsky –, ousou questionar as ideias estabelecidas há muitas décadas em Linguística pela vertente da Gramática Generativa.

No que diz respeito aos mitos linguísticos em geral, a questão não constitui um tema novo em Linguística, dado que esta temática aparece ciclicamente nas discussões académicas sobre a linguagem, em geral, e, especificamente, sobre várias línguas particulares, debatendo o que se entende pelo uso linguístico, pela natureza e aprendizagem da linguagem, pelo seu desenvolvimento, mudança e variação. As discussões deste tipo propõem-se combater os juízos de valor populares muito disseminados entre os falantes, maioritariamente frutos de preconceito e ignorância, tais como, por exemplo: (a) *A gramática da língua X é mais fácil do que a gramática da língua Y*, (b) *É mais fácil aprender a língua Z do que a língua W*; (c) *Existem línguas sem gramática*; (d) *As pessoas da classe social mais baixa falam mal*; (e) *Todos os falantes da nossa língua materna têm sotaque menos eu*; (f) *A língua X é uma língua lógica e bem organizada, a Z é bela e a W é feia*; (g) *Fala-se melhor língua X na área X1 do que nas áreas X2, X3 ou X4*, etc. Estas constatações, que podem dizer respeito a todas as culturas e a todas as línguas, surgem também nas sociedades que falam o Português, dando origem aos slogans como: (i) *Fala-se mal o Português!* (ii) *Estão a assassinar o Português!* (iii) *O Português falado em Coimbra é o melhor*; (iv) *O Português falado em África é o pretoguês*; etc.

No entanto, sendo *The Language Myth*, em termos latos, um livro que se insere no grupo dos tratados sobre os mitos linguísticos que nos rodeiam desde sempre em todas as circunstâncias do uso da linguagem, o livro é muito mais do que apenas mais uma abordagem sobre um tema conhecido. O seu foco principal está no maior mito linguístico de todos os tempos, o da linguagem como instinto, propagado pela ideia da *Gramática Universal* de Noam Chomsky (p. ex., *Aspects of the Theory of Syntax*, 1965; *Knowledge of Language: Its Nature, Origin, and Use*, 1986), assim como, posteriormente, especificado por Steven Pinker (*The Language Instinct*, 1994). Este foco constitui a originalidade do livro, já que antes dele ninguém ousou questionar de modo tão veemente e acutilante as ideias propostas por um dos pais da Linguística moderna, bem como desenvolvidas e acerrimamente defendidas por uma legião de seguidores no mundo inteiro.

Partindo da perspectiva da Linguística Cognitiva, Vyvyan Evans constrói no seu tratado uma forte, apaixonante e minuciosa refutação dos argumentos chomskianos acerca da natureza, origem e uso linguísticos, criando, assim, um antídoto não só à teoria chomskiana ao nível académico, mas também aos populares livros escolares e universitários, nos quais, há mais de cinquenta anos, se propaga a tese da Gramática Universal e da linguagem como instinto. O Autor argumenta contra estas “verdades eternas” da teoria chomskiana, desmascarando a concepção da linguagem como inata, concebida como um tipo de instinto unicamente humano, pré-programado biologicamente ou, até, como um órgão com dispositivo específico para a aquisição da linguagem (*LAD – Language Acquisition Device*).

Baseando-se na mais recente investigação científica desenvolvida nas Ciências Cognitivas, sobretudo na Psicolinguística Cognitiva e nas Neurociências, Evans demonstra, desenvolve e defende a sua tese de Linguística como sendo uma área que emerge de diversos paradigmas científicos que competem entre si. O Autor contrargumenta a tese da linguagem como instinto, defendendo globalmente as seguintes ideias:

- (i) a linguagem humana está ligada às outras formas de comunicação animal, não sendo, por isso, unicamente humana, nem autónoma em relação às outras capacidades cognitivas;
- (ii) a linguagem, e, por conseguinte, as línguas particulares são objecto de mudança, variação e variabilidade, não se tratando de fenómenos regulares nem lineares;
- (iii) a linguagem é adquirida com base nas capacidades gerais da mente humana e não com base numa pré-programada capacidade universal;
- (iv) a linguagem e a mente por um lado reflectem e, por outro, influenciam o modo como interagimos com o mundo e com os outros.

Segundo Evans, a existência da Gramática Universal de Chomsky nunca chegou a ser cientificamente demonstrada, nem pode ser entendida – conforme seria de esperar – com base na evidência empírica de carácter linguístico, porque, de acordo com o Autor, tal evidência praticamente não existe. Evans defende ainda que, em termos específicos, a tese de linguagem como instinto, nuclear à Gramática Universal, nem pode ser considerada como um pressuposto linguístico; trata-se, antes, de uma hipótese contextualizada num novo ramo de assim chamada *Biolinguística*, o que faz dela uma suposição de carácter mais biológico do que propriamente linguístico.

No que diz respeito à estrutura global do *The Language Myth*, o mito principal nele abordado da linguagem como instinto desdobra-se, ao longo do texto, em seis mitos mais específicos, sendo a cada mito dedicado um capítulo, proporcionando um debate que permite interligar várias áreas das Ciências Cognitivas: Linguística, Biologia, Neurociências, Antropologia, Filosofia, etc., proporcionando uma discussão verdadeiramente interdisciplinar.

O livro é constituído por oito capítulos, dos quais cada um dos seis capítulos do meio começa pela formulação de um mito parcial em forma de pergunta que funciona como ancoragem e ponto de partida para a subsequente discussão. Assim, no primeiro capítulo, o Autor propõe uma revisão da problemática relativa às ligações entre a linguagem e a mente. Os seis mitos abordados e analisados por Evans nos capítulos do meio são os seguintes: (i) A linguagem humana não está ligada aos sistemas de comunicação desenvolvidos por animais? (ii) Os universais da linguagem existem? (iii) A linguagem é inata? (iv) A linguagem constitui um módulo distinto da mente? (v) Existe uma linguagem universal da mente? (vi) O pensamento é independente da linguagem? O último capítulo, que retoma a problemática do primeiro, bem como o debate desencadeado pelos mitos parciais, fecha a discussão, centrando-se, mais uma vez, nas relações entre a linguagem e a mente, “recuperando-as” na sequência da discussão prévia.

As notas apresentadas no livro são pormenorizadas e remetem para uma bibliografia actualizada e científicamente diversificada, mas são reunidas no fim, o que permite uma leitura fluída do texto principal, criando alternativas diversas, mais ou menos especializadas da sua leitura e análise.

O Autor apresenta uma exemplificação linguística proveniente do uso corrente da linguagem, recorrendo simultaneamente à argumentação científica elaborada, o que proporciona um estilo vivo e acutilante.

No primeiro capítulo – *Language and mind rethought* –, o Autor começa por definir o objecto do seu estudo, propondo-se abordar a linguagem, o seu relacionamento com o pensamento e a mente, assim como a sua aquisição e variabilidade. Especificando, Evans propõe-se a abordar a controvérsia, que nasceu em meados do século passado, no seio das ciências do comportamento, por um lado, e nas que estudam o funcionamento do cérebro, por outro. Esta controvérsia, que confronta duas teses diferentes, pode resumir-se à seguinte questão: (i) será a linguagem inata, isto é, trata-se de uma capacidade com que o ser humano nasce à partida ou, pelo contrário, (ii) será que a linguagem emerge do uso linguístico, ancorando-se em capacidades cognitivas mais gerais e não especificamente linguísticas? Evans define a primeira tese como sendo de raiz chomskiana e dominante nas Ciências da Linguagem até a relativamente pouco tempo e classifica-a como “crescentemente instável”, argumentando contra ela com base nos mais recentes resultados da investigação desenvolvida em vários ramos da ciência, que o Autor se compromete desenvolver e aprofundar ao longo do livro. Defende-se, assim, que a linguagem *não* surge do conhecimento da gramática humana programada como inata (proposta pelos inatistas como a Gramática Universal), mas baseia-se nas capacidades gerais da mente humana de cariz cultural sobre as quais se contrói, reflectindo objectivos pró-sociais para a comunicação intersubjectiva dos humanos. Quando adquirimos a linguagem, defende o Autor, fazemo-lo em função do *input* linguístico que recebemos dos nossos pais e outras pessoas que nos rodeiam, o que resulta numa tarefa laboriosa e não isenta de erros, ao longo de um processo demorado, que pode ser penoso e nada tem – ao contrário do que propõem os inatistas – nem de automático nem de fácil. Embora não estejamos equipados especificamente para adquirir a linguagem como uma capacidade linguística distinta das outras capacidades cognitivas, nascemos com uma sólida preparação biológica, dispondo de uma bateria de mecanismos gerais de aprendizagem que nos tornam aptos para adquirir a(s) nossa(s) língua(s) maternal(s).

Contrastando estas duas abordagens distintas, Evans propõe (p. 5) que a discussão a desenvolver no livro contraste *a tese da linguagem como instinto (language-as-instinct thesis)*, defendida pelos inatistas, com a da *linguagem como uso (language-as-use thesis)*, defendida pelo próprio e ancorada nas Ciências Cognitivas. O Autor argumenta que a primeira tese é um *mito*, porque não reflecte a realidade dos factos linguísticos nem dispõe de consubstanciação empírica. Evans contrasta-a com a segunda tese, que entende estar de acordo com a realidade do uso linguístico e em conformidade com a investigação actualizada desenvolvida hoje em dia em vários ramos da ciência. O Autor entende como mito um relato não consubstanciado cientificamente de um fenómeno linguístico que não corresponde ao conhecimento do qual dispomos presentemente sobre o funcionamento real da

linguagem e da mente: “*A myth, for my purposes, is an unproven account of a linguistic phenomenon that appears to be at odds with actual findings relating to language, the mind, and so on.*” (p. 13). O mito resulta, por regra, de uma proposta especulativa, tal como aconteceu no caso da *tese da linguagem como instinto* de Chomsky e os seus seguidores: as ideias nela apresentadas podem ser lógicas e parecer evidentes para os próprios proponentes, mas carecem de evidência actualizada com base em factos linguísticos reais; para se poder fundamentar esta tese como teoria, seria preciso que as propostas iniciais pudessem ser refutadas ou confirmadas, com base na análise fundamentada nas observações detalhadas dos dados linguísticos autênticos.

No que diz respeito à origem e aos primórdios linguísticos, segundo *a tese da linguagem como instinto* (p.24), a própria linguagem surgiu num dado momento pontual da história do Homem, em forma de um estado (quase) perfeito, sendo resultado da mutação genética. Evans argumenta contra esta tese (p. 26) como estando em claro desacordo com as mais recentes descobertas da teoria da evolução, segundo as quais se defende que é preciso uma progressiva adaptação de co-determinação entre as espécies e o meio-ambiente; assim, segundo o Autor, não se pode aceitar que a linguagem tivesse surgido do nada, num único momento da história, mas, antes, defendere que evoluiu progressivamente de diversas tendências e manifestações prévias de comunicação humana, análogas a diversos tipos de comunicação existentes noutros tipos de espécies, sobretudo de diversas formas de proto-linguagem.

Na sequência da introdução global da problemática a abordar apresentada no primeiro capítulo do livro, a partir do segundo capítulo, o Autor dedica-se a discutir um por um os mitos específicos constituintes da *tese da linguagem como instinto*.

Por conseguinte, o capítulo dois – *Is human language unrelated to human communication system?* – aborda a questão do realcionamento entre a linguagem humana, por um lado, e o sistema humano de comunicação na globalidade, por outro. A tese inatista defende que a linguagem deve ser entendida como uma capacidade exclusivamente humana e não comparável nem relacionada com nenhuma outra forma de comunicação que seja não-humana, separando, deste modo, a linguagem humana de qualquer outro modo de comunicação entre outras espécies. Refutando esta tese como um mito, Evans demonstra que, hoje em dia, existem provas científicas em como a linguagem humana esteja ligada a outras formas de comunicação animal – de baleias, abelhas, pássaros (p. ex., estorninhos), símios (p. ex., macacos-vervet), entre outros – não sendo, por isso, exclusiva do Homem. Esta evidência científica permite-nos aprender muito mais acerca da linguagem humana em si e, também, acerca da sua especificidade, procurando entender como, sendo qualitativamente diferente, ela deriva e, posteriormente, se relaciona com os sistemas comunicativos de outras espécies. Na sequência desta evidência, defende-se que a capacidade linguística dos humanos não pode ser considerada como autónoma em relação às outras capacidades cognitivas do Homem, conforme pretendem os inatistas.

No terceiro capítulo – *Are there language universals?* – o debate centra-se sobre a questão dos universais linguísticos. A *tese da linguagem como instinto* postula que a Gramática Universal, com a qual supostamente ficamos equipados à nascença, é constituída por uma série de universais linguísticos, características que constituem uma constante em todas as línguas particulares existentes

no mundo: “Chomsky famously proposed a Universal Grammar, which he dubs «a general principle of linguistic structure on the basis of observation of a single language»: English” (p. 93). Calcula-se que, presentemente, existem entre seis e oito mil idiomas no mundo (p. 65), não sendo fácil o estabelecimento deste número por não ser linear a distinção formal entre uma língua e um dialecto. Segundo esta tese, basta estudar profundamente um único idioma, por exemplo, o Inglês, a fim de perceber como funcionam todas as outras línguas, que se espera preencherem, à partida, os mesmos padrões universais com apenas umas pequenas discrepâncias que possam existir entre eles. Evans refuta esta argumentação com evidência linguística autêntica, proveninete da exemplificação específica de vários idiomas, demonstrando que as línguas variam no tempo e no espaço, emergendo e diversificando-se ao longo da existência da espécie humana e do respectivo uso linguístico e, eventualmente, acabando por morrer ou afastar-se umas das outras. Ao longo das trinta páginas dedicadas a esta temática no capítulo três, o Autor apresenta uma justificação pormenorizada de carácter fonológico (págs.70-71), relativa às partes do discurso (págs.72-73), e, depois, de carácter morfológico (págs. 73-74), sintáctico (págs.74-77) e tipológico (págs. 77-79), acabando por discutir o mito da proto-linguagem (págs. 88-93). Mais especificamente, com base num estudo publicado na *Nature*, em 2011, por Dunn e colaboradores (no artigo intitulado *Evolved structure of language shows lineage specific trends in word-order universals*, *Nature*, 473: 79-82), Evans conclui que a tendência das línguas do mundo é afastarem-se umas das outras por diferenças de ordem cultural e por pressões de usos distintos e diversificados, em vez de ficarem coagidas pelas restrições fixas dos supostos universais. Simbolicamente, a última parte do capítulo 3 é intitulada: “Adieu, Universal Grammar!”

O capítulo quarto – *Is language innate?* – é o núcleo da tese da *linguagem como uso* defendida por Evans, refutando-se o carácter pretensamente inato da linguagem. Começa-se por reconhecer que, do ponto de vista neurobiológico, é inquestionável que os humanos – como nenhuma outra espécie – estejam preparados para adquirir a linguagem, por disporem quer de um aparelho fonador mais sofisticado do que as outras espécies quer da capacidade geral de processamento e armazenamento da informação. Isto não significa, no entanto – defende o Autor – que, à nascença, o ser humano disponha dos fundamentos essenciais da linguagem, comuns a todas as línguas particulares, ou seja, de um conhecimento específico de carácter gramatical, no âmbito da Gramática Universal, conforme pretendem os defensores da *tese da linguagem como instinto*. Evans reconhece, porém, que a tese dos inatistas pode, à partida, parecer bastante atractiva do ponto de vista explicativo, tendo em conta que as crianças adquirem a linguagem com aparente facilidade sem serem ensinadas formalmente nem receberem o *feedback* negativo da parte de quem as rodeia, isto é, sem os adultos corrigirem os seus “erros de aprendizagem”. Apesar de reconhecer estes argumentos, Evans refuta-os, baseando-se na investigação mais recente, desenvolvida na área de Neurobiologia e Psicolinguística Cognitiva, segundo a qual a linguagem não se adquire “aos saltos”, como pretendem os inatistas, mas num processo gradativo, por fases, ao longo dos primeiros anos da idade da criança.

Defende-se, assim, que a aquisição emerge com base nas unidades de aprendizagem provenientes do *input* linguístico que rodeia o falante e não em função de um pré-determinado sistema universal de regras; fundamenta-se, por conseguinte, nas capacidades gerais da mente humana e não numa suposta pré-programada capacidade universal específica. A primeira destas capacidades gerais é a de

reconhecer e definir padrões (*pattern-finding ability*), encontrada igualmente entre os primatas (pág. 118-123). A segunda capacidade tem a ver com o reconhecimento e imitação por parte da criança da intenção comunicativa do outro (*intention-reading ability*), que é determinada culturalmente, o que implica a adaptação de carácter cultural e simbólico de quem adquire a linguagem (págs. 123-124). Evans analisa várias propostas cognitivistas de como se processa a aquisição da linguagem – p. ex., a de Langacker (págs. 123-124), Bybee (125-126) ou Goldberg (126-127) – contrastando-as com a dos inatistas (cf. Bickerton, na sua teoria sobre a formação das línguas crioulas, págs. 128-129). O Autor argumenta a favor da aquisição da linguagem que emerge com base em unidades de uso linguístico chamadas ‘*chunks*’, provenientes do *input* linguístico, às quais o falante se encontra exposto, num processo feito gradativamente e com esforço de aprendizagem (págs. 131-132). Neste processo baseado no uso, dependente da frequência dos itens que ocorrem no *input* e da sua repetição, a criança aprende a linguagem com base em ‘*chunks*’, para, com o tempo e esforço passar, progressivamente, para a fase da abstracção, na qual começa a criar padrões, tornando-se também competente em reconhecer intenções de comunicação determinadas social e culturalmente dos outros falantes. É deste modo – defende o Autor – que os factores determinados pelo uso do *input* linguístico determinam a construção da gramática da criança ao longo do processo da aquisição: “*The child is adept at pattern-finding, and comes equipped with a species-specific ability to recognize communicative intentions, and, moreover, a pro-social desire to communicate. And, finally, usage-based factors guide the construction of a grammar in the mind of the child.*” (pág.132).

No capítulo quinto – *Is language a distinct module in the mind?* – debate-se a questão da modularidade da mente humana, isto é, questiona-se que a linguagem constitua um módulo separado na mente, conforme proposto por Fodor (1983) e, a partir daí, defendido pela tese inatista como uma doutrina inquestionável. A proposta da modularidade da mente surge na sequência de uma longa tradição de vários séculos, desde os tempos escolásticos de São Tomás de Aquino (século XIII), na qual se tem vindo a defender que as funções mentais do ser humano sejam atributos de partes específicas da mente e do cérebro. Esta tradição foi reforçada nos anos cinquenta do século XX pela metáfora largamente utilizada da mente humana como computador (*mind-as-computer perspective*), na qual as diferentes components do *hardware* podiam ser estudadas separadamente como módulos individuais, complementando-se na construção e funcionamento do mecanismo modular do cérebro/ mente, que funciona como um computador. Baseando-se na mais recente evidência empírica proveniente das Neurociências (págs. 145-146), tanto no caso dos falantes normais como nos falantes com desordens da linguagem (*SLI – Specific Language Impairment*), Evans refuta esta abordagem, demonstrando que a aquisição e o desenvolvimento da gramática e do léxico são indissociáveis e não podem ser abordados em termos modulares. O Autor discute igualmente a existência do gene FOXP2 (o suposto gene da gramática; cf. Lai *et al.* 2001) como não sendo exclusivo nem específico da linguagem (págs. 143-144). Defende, ainda, que o desenvolvimento da linguagem e do cérebro são fruto da co-evolução estudada pela Teoria da Complexidade (Gibbs & Orden 2010) na qual qualquer sistema é analisado na sua capacidade de se organizar a si mesmo, ocorrendo ao nível global adaptações às mudanças que podem ocorrer apenas num lugar específico do sistema (pág. 157). Na sequência das provas empíricas hoje em dia disponíveis, chega-se à conclusão que a linguagem é, de facto, diferente dos outros sistemas de comunicação das outras espécies, assim como que a mente humana, tal como

o cérebro, se distinguem por especificidades não verificadas noutras espécies. Isto não significa, no entanto, que se possa defender o seu funcionamento modular ou a existência de qualquer tipo da faculdade da linguagem (postulada pelos inatistas). Evans entende que este mito só continua a existir por se ter institucionalizado na narrativa académica, tornando-se immune à evidência científica mais recente: “(...) *the myth of language modularity – of a language faculty – has become institutionalised via retellings which are now immune to counterevidence*” (págs. 159).

O capítulo sexto – *Is there a universal Mentalese?* – aborda o mito da linguagem da mente / pensamento (*Mentalese*), assim como o seu carácter supostamente universal. A tese inatista pressupõe que exista a linguagem universal interna, que torne o pensamento viável, sendo constituída por símbolos e regras universais da sintaxe mental. Evans refuta este argumento, entendendo que o pressuposto inatista só poderia ser defendido, se se conseguisse comprovar a tese do cérebro como computador, discutida e refutada no capítulo anterior. Para abordar a natureza do mito da linguagem da mente do ponto de vista inatista é preciso – defende o Autor – entender, primeiro, como surge o significado na teoria chomskiana, propondo-se para esta fundamentação três constituintes específicos (págs. 169-172): fundamentação dos símbolos, da interpretação e do carácter nuclear da sintaxe (*syntactocentric nature of Mentalese*). A refutação deste mito (págs. 178-189) é apresentada com base na discussão da ancoragem do significado na experiência humana em geral e, especificamente, na do corpo humano (*embodiment*), na senda da *Teoria da Metáfora* de Lakoff e Johnson (1980), assim como no efeito do *embodiment* sobre o funcionamento do cérebro humano. A experiência humana, como não é dada, precisa de ser aprendida; no caso da linguagem, a experiência provém do *input* linguístico no qual estamos imersos e que nos permite a aquisição / aprendizagem da(s) nossa(s) língua(s) materna(s). Na sequência desta refutação, Evans defende que a oposição do que é conhecido na literatura anglossaxónica como o fenómeno *nature vs. nurture*, isto é, inatismo vs. experiência, não pode ser indevidamente polarizado, visto que as nossas capacidades gerais inatas definem e viabilizam o processo de aprendizagem, que emerge da experiência.

No capítulo sétimo – *Is thought independent of language?* – Evans discute as relações de (in) dependência existentes entre o pensamento e a linguagem. Segundo a teoria inatista, os dois fenómenos são independentes, sem um sofrer a influência do outro. No entanto, existe hoje em dia muita evidência significativa proveniente de vários ramos da ciência em como os paradigmas linguísticos dos idiomas que falamos exercem influência sobre o modo como apreendemos o mundo circundante e como o exprimimos ao nível linguístico. O modo como uma determinada língua particular codifica domínios como espaço, tempo, cor ou género, por exemplo, exerce, de facto, influência sobre o entendimento que os seus falantes têm sobre o mundo e é conhecida como a teoria do *relativismo linguístico* (págs. 193-228). Segundo esta teoria, o modo como construímos o nosso pensamento e nos apercebemos do que nos circunda é influenciado pelas diferenças estruturais que existem entre as línguas particulares que falamos. A teoria do relativismo linguístico foi primeiro defendida por Whorf (1956), sendo depois, reformulada como hipótese de Sapir-Whorf (cf. Sapir 1985) e, posteriormente, desenvolvida como a teoria neo-whorfiana. O relativismo linguístico foi consubstanciado ao longo dos anos por muita evidência linguística, da qual se destacou, primeiro, a discussão sobre a expressão linguística do domínio da cor (desenvolvida a partir do tratado de Berlin & Kay, de 1969), e, depois,

no domínio do espaço (Levison 2003). Evans demonstra que o relativismo linguístico é um fenómeno que tem a ver com o impacto que a língua que é utilizada pelo falante exerce sobre o seu aparelho cognitivo, o que significa que existem diferenças entre as mentes dos falantes de diversas línguas, dada a experiência linguística que cada um dos falantes tem. Segundo a perspectiva defendida pelo autor, a linguagem humana (tal como observada com base no uso efectuado em diferentes idiomas particulares) e a mente humana estão interconectadas, vivendo em simbiose.

Depois de, ao longo dos diferentes capítulos, ter analisado em pormenor os seis mitos componentes do mito nuclear da linguagem como instinto e ter proposto as respectivas alternativas da sua autoria, no oitavo capítulo – *Language and mind regained* – Evans aborda a temática da linguagem e da mente “recuperadas”.

Neste capítulo final, o Autor desenvolve a sua tese alternativa da *linguagem-como-uso*, na qual defende, primeiro, que o ser humano é dotado de uma *inteligência cultural*, graças à qual desenvolve comportamento *cooperativo e pro-social*, que o distingue do comportamento social menos elaborado, observado nas outras espécies. Ao longo da evolução, esta inteligência cultural facilitou, por um lado, o desenvolvimento de comportamentos simbólicos mais avançados, sendo a linguagem humana o seu exemplo paradigmático, e, por outro, uma competência mental adicional, a *inteligência interaccional*.

Em segundo lugar, Evans defende que existem universais ao nível da experiência humana (págs. 250-252), que, por sua vez, exercem restrições sobre o funcionamento da linguagem. Estes universais não se encontram *dentro* da linguagem (como defendia a tese inatista, ao propor a Gramática Universal), mas emergem da experiência humana comum, que – em termos latos – é determinada pelo meio físico em que nos movemos e depende das estruturas neuroanatômicas dos nossos cérebros, comuns a todos os homens.

Por conseguinte, e em terceiro lugar, o Autor caracteriza a *gramática mental* do ser humano (págs.242-249) como sendo constituída por unidades linguísticas de diferentes níveis de complexidade, e que são provenientes do *input* linguístico, formando construções e organizando-se em redes: “*The language-as-use thesis claims that linguistic units – constructions – are organised in an individual’s mind as a network, with more abstract schemas being related to more specific instances of language.*” (pág. 249).

Em suma, o livro *The Language Myth: Why language is not an instinct?*, da autoria de Vyvyan Evans, constitui uma verdadeira pedrada no charco das ideias sobre a origem e natureza da linguagem, cristalizadas com base na hegemonia institucionalizada da teoria chomskiana, estabelecida entre nós ao longo de mais de meio século, sem, no entanto, dispor da base empírica substancial que a comprovasse: “*(...) the language-as-instinct thesis, which today remains the single largest ‘school’ in Anglo-American linguistics, has little empirical basis sustaining itself in an institutional hegemony which, in the view of some, seeks to maintain the status quo, regardless of the actual facts: reputations have been built upon the idea of the language-as-instinct thesis; for these reserchers, the prospect of the Universal Grammar thesis being wrong doesn’t bear thinking about.*” (págs. 140-141).

Os argumentos da *tese da linguagem-como-uso* de Evans constituem uma poderosa alternativa às

verdades adquiridas do “credo linguístico” estabelecido da *tese da linguagem-como-instinto*. Esta alternativa permite uma perspectiva renovada e bem fundamentada no que na ciência mais recente e desenvolvida de modo interdisciplinar se tem descoberto sobre o funcionamento linguístico do Homem, da sua mente e da linguagem humana.

REFERÊNCIAS BIBLIOGRÁFICAS

ADGER, David (2015a). Mythical myths: Comments on Vyvyan Evans’ “The Language Myth”. Review article. *Lingua*, Volume 158, April 2015, 76-80.

ADGER, David (2015b). More misrepresentation: A response to Behme and Evans (2015). Review article. *Lingua*, Volume 162, July 2015, 160-166.

BAUER Laurie & Peter TRUDGILL (eds.), 1998 *Language Myths*. London: Pinguin Books.

BEHME, Christina & Vyvyan EVANS (2015). Leaving the myth behind: A reply to Adger (2015). Review article. *Lingua*, Volume 162, July 2015, 149-159.

BERLIN, Brent & Paul KAY (1969). *Basic Colour Terms: Their Universality and Evolution*. Cambridge University Press.

BICKERTON, Derek (1981). *Roots of Language*. Ann Arbor, MI: Karoma Publishers.

BICKERTON, Derek (1984). The language bioprogram hypothesis. *Behavioral and Brain Sciences*, 7: 173-221.

BYBEE, Joan (2010). *Language, Usage and Cognition*. Cambridge University Press.

CHOMSKY Noam (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

CHOMSKY Noam (1986). *Knowledge of Language: Its Nature, Origin, and Use*. Greenwood Publishing Group.

DUNN, Michael; S. J. GREENHILL; S. C. LEVINSON, and R. D. GRAY (2011). *Evolved structure of language shows lineage specific trends in word-order universals*, *Nature*, 473: 79-82.

EVANS, Vyvyan (s/d). Language in the Mind. The evolution of language and cognitive linguistics. Blog of *Psychology Today*. Consultado a 01/04/2016 em: <https://www.psychologytoday.com/blog/language-in-the-mind>

EVANS, Vyvyan (2014). *The Language Myth: Why language is not an instinct?* Cambridge: Cambridge University Press.

EVANS, Vyvyan (2015a). The structure of Scientific Revolutions. Reflections on radical fundamentalism in language science. April 2015. Consultado a 01/04/2016 em: <https://www.psychologytoday.com/blog/language-in-the-mind/201504/the-structure-scientific-revolutions>

EVANS, Vyvyan (2015b). Challenging ill-conceived dogma. A response. December 2015. Consultado a 01/04/2016 em: <http://www.vyvevans.net/ChallengingIllConceivedDogma.pdf>

EVANS, Vyvyan (2015c). *The Crucible of Language: How Language and Mind Create Meaning*. Cambridge University Press.

EVANS, Vyvyan (2016a). Open Letter, regarding my invited response article on The Language Myth, solicited by *Language*. January 2016. Consultado a 01/04/2016 em: <http://www.vyvevans.net/OpenLetter.pdf>

FODOR, Jerry (1983). *Modularity of Mind*. Cambridge, MA: MIT Press.

GIBBS, Ray W. & G. van ORDEN (2010). Adaptive cognition without massive modularity. *Language & Cognition*, 2: 149-176.

GOLDBERG, Adele (1995). *Contractions: a Construction Grammar Approach to Argument Structure*. University of Chicago Press.

LAI, Cecilia; S. E. FISHER; J. A. HURST; f. VARGHA-KADEM & A. P. MONACO (2001). A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature*, 413: 519-523.

LAKOFF, George & Mark JOHNSON (1980). *Metaphors we Live by*. University of Chicago Press.

LANGACKER, Ronald W. (2008). *Cognitive Grammar: a Basic Introduction*. Oxford University Press.

LEVISON, Stephen C. (2003). *Space in Language and Cognition: Explorations in Cultural Diversity*. Cambridge University Press.

PINKER, Steven (1994). *The Language Instinct*. New York: William Morrow

PINKER, Steven (2002). *The Blank Slate*. New York: Pinguin.

SAPIR, Edward (1985). *Selected Writings in Language, Culture and Personality*. Edited by D. G. Mandelbaum. Berkeley: University of California Press.

“Vyv Evans: Professor of Linguistics at Bangor University UK”. *21st century stars in: 21st century.co.uk*. Consultado a 01/04/2016 em: <http://www.21stcentury.co.uk/21st-century-stars/vyv-evans/>

WHORF, Benjamin (1956). *Language, Thought and Reality: Selected Writings of Benjamin Lee Whorf*. Edited by John B. Carroll. Cambridge, MA: MIT Press.

Recebido em 24/02/2016

Aceito em 01/04/2016

CONFORMITY AND BENEFICIARY MEANINGS OF LATIN *SECUNDUM* NP: A CORPUS-BASED COGNITIVE ANALYSIS

Caterina Guardamagna (*Lancaster University, United Kingdom*)

ABSTRACT

This paper adds to a body of cognitive studies on prepositions by presenting a corpus-based diachronic study of *secundum* NP with meaning of conformity and beneficiary. This paper discusses the possible origin of the conformity sense coming from an expression having the meaning ‘downstream’ and argues that an explanation based on the extraction of schemas from the proto-scene and force dynamics appears to be more plausible. This paper also argues that, although the beneficiary meaning is connected to the spatial meaning via a metaphorical link, force dynamics also plays a role, connecting it directly to the meaning of conformity.

Keywords: preposition, semantic change, metaphor, metonymy, force dynamics

1. INTRODUCTION

To date, no extensive research has been carried out on the Latin prepositional phrase *secundum* NP. The literature consists only of some sketchy observations put forward in Ferrari (1998), Rocha (1998), Heine & Kuteva (2002: 139), Guardamagna (2011) and Trabelsi (2013), mainly focusing on Romance continuers of *secundum* NP which are linked back to their Latin etymological source. This paper contributes to plugging this gap in the literature by providing a cognitive analysis of two meanings of Latin *secundum* NP, namely the conformity meaning (e.g. *secundum naturam* ‘according to nature’) and the beneficiary meaning (*secundum reos* ‘in favour of the defendants’).

This study draws on the *Latin Library* diachronic corpus and spans over a period of 800 years (BC 106 – AD 704). For my research I considered alternate periods and for each period I analysed a

sample of 250 occurrences of *secundum* NP, with the exception of the Early Medieval period, only featuring 140 occurrences of the construction (see table 1, below). The meanings of conformity and beneficiary account for 506 (59%) of the 860 instance sample of *secundum* NP extracted from the 6M prose section of the *Latin Library* corpus via the CQP search engine (Hardie 2012). I analysed this sample of *secundum* NP constructions both qualitatively and quantitatively.

	Classical Latin BC 106-AD 17		Silver Latin 100-258		Late Latin 330-469		Early Medieval 530-704		Total Rows #
	#	%	#	%	#	%	# Normed to 250	%	
Conformity	91	36.4	156	62.4	148	59.2	82	32.8	477
Beneficiary	24	9.2	3	1.2	2	0.8	0	0	29
Total Columns	115	45.6	159	63.3	150	60	82	32.8	506

Table 1. *Secundum* NP with meanings of conformity and beneficiary in the *Latin Library* corpus.

This article is organised as follows. Section 2 presents the proto-scene encoded by the verb *sequor* ‘follow,’ from which *secundum* derives. Section 3 explores the connection of the conformity meaning to the spatial domain via the metaphorical mapping CONFORMITY IS MOTION ALONG. Then, section 4 offers an analysis of the beneficiary meanings of *secundum* NP via the metaphor IN FAVOUR OF IS ON THE SIDE OF. Next, section 5 deals with the role of force dynamics in the development of the two meanings, linking conformity to the proto-scene and beneficiary to conformity. Finally, section 6 concludes the paper. Sections 2 and 4 partly draw on Guardamagna (forth. 2017).

2. THE PROTO-SCENE: *SEQUOR* ‘FOLLOW’

The preposition *secundum* comes from the gerund (Heine and Kuteva 2002: 139) or the gerundive (Ferrari 1998) of the verb *sequor* ‘follow’ (see Poultney 1980: 34, Vineis 1998: 308 and Garnier, 2015). Therefore, its semantics is tightly connected to the “proto-scene” (Tyler and Evans 2007: 3, Evans 2007) encoded by this verb. The basic scenario for the verb *sequor* involves two participants: a trajector and a landmark (Langacker 1987: 217, 231, Taylor, 1993: 153, Talmy 2000a: 37, Croft and Cruse 2004: 56). Both of them are prototypically construed as uniplex, i.e. as point-like, bounded entities (on the notion of plexity see Talmy 2000a: 58-59, 2000b: 458-459, Luraghi 2003: 25, 2006). Prototypically, both the trajector and the landmark are construed as having a front-back orientation, and the trajector is conceptualised as being behind the landmark from the vantage point of the viewer/speaker (see figure 1, below).

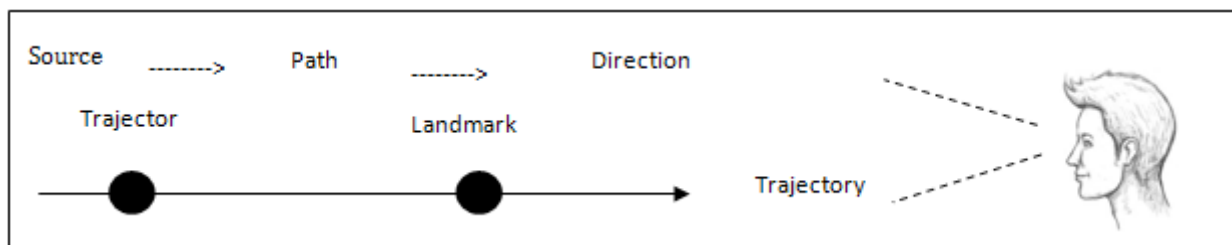


Figure 1. A prototypical motion event (after Luraghi 2014:103).

The movement is initiated by the landmark, with the trajector following (going behind) it. Importantly, both participants move along the same path and in the same direction. An example of this is sentence (1), below.

- (1) *Iam sequ-or te, mater.*
 already follow-IND.PRS.1P.SG.DEP you.ACCmother.VOC
 ‘I am already following you, mother.’ (Plautus, *Aulularia* 4, 7, 16 from Lewis & Short 1879)

The connection between *secundum* NP with conformity meaning and the proto-scene of *sequor* has not been explored in the literature. Section 3, below, deals with a possible source for conformity expressions, namely a construction meaning ‘downstream’. Then, after having presented *secundum* NP expressions with beneficiary meaning in section 4, section 5 explains the connection between *secundum* NP with meaning of conformity and the proto-scene described here.

3. CONFORMITY

The function of conformity is the most frequent meaning of *secundum* NP throughout the history of Latin. It indicates that the trajector’s state of affairs occurs in accordance with (‘in line with’) some properties of the landmark.

- (2) *Ita fin-is bon-orum exist-it, secundum natur-am*
 So end-NOM.F.SG good-GEN.N.PL exist-IND.PRS.3P.SG in.accordance.with nature-ACC.F.SG
viv-ere.
 live-INF.PRS
 ‘Thus arises ‘the end of goods’, namely to live in accordance with nature.’ (Cicero, *De Finibus* 5, 24)
- (3) *Accep-erat a domin-o su-o... disciplin-am*
 Took(PRF)-IND.PLUPRF.3P.SG from God-ABL.M.SG POSS.3P.SG-ABL.M.SG rule-ACC.F.SG
secundum leg-em viv-endi.
 in.accordance.with law-ACC.F.SG live-GEN.GER
 ‘He had received from his Lord... the rule to live according to the law.’ (Tertullian, *Liber Scorpiace* 5, 11)

Analysing the Ancient Greek construction *katà* NP_{accusative} the translational equivalent of *secundum* NP, Luraghi (2003: 200) suggests that the conformity meaning in (4) arises as a metaphorical mapping from the domain of physical movement, namely motion along (‘along-downwards’), exemplified in (5) below.

- (4) *àrsantes* *katà thumón*
 suit:PART.AOR.NOM.PL down mind:ACC
 ‘suiting it according to (my) mind’ (Homer, *Iliad*, 1.136, from Luraghi 2010: 203)
- (5) *potamoùs d’ ètrepse néesthai kàr rhóon*
 river.ACC.PL PTC turn:AOR.3SG flow:INF.PRS.M/P down stream:ACC
 ‘the rivers turned back to flow down the stream’ (Homer, *Iliad*, 12.32-33, from Luraghi 2010: 200)

Specifically, Luraghi (2003: 203) observes that in examples like (5) “where the preposition means ‘downstream’ there is a sense of conformity to the flowing of the water: a downstream motion is a motion that follows the stream (whereas ‘upstream’ would mean ‘contrary to the flow of water’)”.

One might suggest that a similar evolution path may also be posited for Latin. However, the expression *secundum flumen* ‘downstream’ appears only once in a motion-along context in the whole Classical Latin prose and poetry corpus within the *Latin Library* (3,725,672 words), and there are no occurrences of *secundum amnem* ‘according to the river’.

- (6) *Sex legion-esad oppid-um Gergovi-am secundumflumen*
 Six legion-ACC.F.PL to city-ACC.N.SG Gergovia-ACC.F.SG along river.ACC.N.SG
Elaverdux-it.
 Elaverled(PRF)-IND.PRF.3P.SG
 ‘(He) led six legions to the city of Gergovia along the Elaver river.’ (Caesar, *De Bello Gallico* 7)

This low frequency¹ may therefore suggest that the *secundum flumen* construction is unlikely to be the source of a metaphorical mapping CONFORMITY IS MOTION ALONG, connecting the space domain directly with the logical domain. On the one hand, assuming that frequency is a major factor for entrenchment and semantic change, one may speculate that the frequency of the nautical expression *secundum flumen/amnem* ‘downstream’ may have been greater in the spoken language than the written record suggests. On the other hand, frequency is not the only important factor for semantic change; salience is as well. The concept of ontological salience, borrowed from perceptual psychology, indicates that an object “stands out” from its surroundings because of some characteristics which do not conform to the group, conceptualised as background. Therefore, rare constructions like *secundum flumen/amnem* could be perceived as salient by speakers precisely because they are rare and they do not display the most frequently attested characteristics of other related *secundum* NP constructions.

Motion-along constructions are less frequent than location constructions and their landmarks have a force and a direction of their own, in contrast to other non-dynamic constructions in which the landmark is either ontologically static like in (7) or conceptualised like a static container as in (8).

¹ Competing constructions expressing ‘downstream’ and ‘upstream, against the flow’ are nominal ablativus absolutus constructions, namely *secundo flumine* lit. ‘the river being favourable’ (attested twice in the Classical Latin corpus: Caesar, *De Bello Gallico* 7, 58 and 7,60) and *adverso flumine* lit. ‘the river being contrary’ (attested three times in my Classical Latin corpus, specifically in Caesar, *De Bello Civili* 3, 30, *De Bello Gallico* 7, 59-60 and once in Late Latin, namely in Sulpicius Severus, *Chronica* 1, 22). Admittedly, these constructions are also quite infrequent.

- (7) *Tunc ab line-a qu-ae secundum abac-i extrem-am*
 Then from line-ABL.F.SG REL-NOM.F.SG along abacus-GEN.M.SG farthest-ACC.F.SG
part-em demiss-a erit, in interior-em
 part-ACC-F.SG dropped-NOM.F.SG be.IND.FUT.3P.SG toward internal-ACC.F.SG
part-em <ali-a> reced-at.
 part-ACC.3P.SG other-NOM.F.SG go.back-SUBJV.PRS.3P.SG
 ‘Then let another line be drawn... toward the inside of the line from the line which had been drawn along the edge of the abacus (= a flat slab placed above the capital of a column).’
 (Vitruvius, *De Architectura* 3,5,6)
- (8) *An ire-t ad tris legion-es Macedonic-as qu-ae*
 Whether go.SBJV.IMPF-3P.SG to three legion-ACC.F.PL Macedonian-ACC.F.PL REL-NOM.F.PL
iter secundum mar-es super-um faci-unt.
 trip.ACC.N.SG along sea-ACC.N.SG superior-ACC.N.PL do-IND.PRS.3P.PL
 ‘Whether to go to three Macedonian legions which are travelling along the Adriatic’ (Cicero, *Ad Atticum* 16, 8, 2)

Because of their salience, the *secundum flumen/amnem* constructions would then be able to provide image schemas, i.e. embodied structuring patterns for understanding and reasoning (Langacker 1987: 371, 2008: 23), which may then get entrenched in new constructions. This would conform to Schmid’s observation (2010: 120) that “there is a two way relationship between salience and entrenchment”.

However possible, I think that the explanation outlined above does not capture all factors at play. It is in fact also possible to suggest that the image schema of a landmark whose movement is independent from that of the trajector conforming to it is likely to have been inherited directly from the proto-scene of the verb *sequor* ‘follow.’ This additional explanation is explored in section 5.1, below. Before turning to this, section 4 presents the beneficiary meaning of *secundum* NP.

4. BENEFICIARY

The term beneficiary indicates the semantic role of a referent which is advantaged by an event (see Luraghi 2003: 40). This function of *secundum* NP is rare: it is found in Classical Latin and very marginally in Silver and Late Latin, before disappearing altogether in Early Medieval Latin (table 1, above).

- (9) *Summ-ar-ei haec fuit: centumvir-i*
 Highest-NOM.F.SG thing-GEN.F.SG this.NOM.F.SG be.IND.PRF.3P.SG centumvir-NOM.M.PL
dix-erunt d-are ips-os secundum adversari-um
 said(PRF)-IND.PRF.3P.PL give-INF.PRS they-ACC.M.PL in.favour.of opponent-ACC.M.SG
Albuci-i...
 Albucius-GEN.M.SG
 ‘This was the height/limit of the thing: the *centumviri* said that they would judge in favour of Albucius’s opponent...’ (Seneca, *Controversiae et Suasoriae* 7, 7)

Beneficiaries are typically human. By metonymy, also example (10), below, indicating a person's body, can be considered to indicate a human participant.

- (10) *Contra si cui suggest-um fuerit, secundum*
 On.the.other.hand if REL.DAT suggested-ACC.N.SG be.SBJV.IMPF.3P.PL in.favour.of
corp-us fortasse effic-ere remedi-a diabol-i.
 body-ACC.N.SG perhaps carry.out-INF/PRS remedy-ACC.N.PL devil-GEN.M.SG
 'On the contrary, if to someone it was suggested to try, in favour of the body (= for the body's health) the remedies of the devil.' (Augustin, *Sermones* 4, 36)

Beneficiary *secundum* NP also appears in two fixed legal expressions: *litem DARE secundum aliquem* 'judge in favour of someone' (11) and *vindicias secundum libertatem/servitutum DARE/CEDERE* 'to enslave someone, to free someone' (lit. to decide a claim in favour of freedom/slavery) (12) (Adam 1807, Lewis & Short 1879 *vindiciae*, Nettleship 1889, Laffi 2007: 52-53). The expression in (11) is more prototypical than (12) because it encodes a human referent.

- (11) *Secundum telit-emd-o.*
 In.favour.of you.ACC case-ACC.F.SG give-IND.PRS.1P.SG
 'I judge in your favour.' (Valerius Maximus, *Factorum et Dictorum Memorabilium Libri* 2, *De Iure Triumphi*, 2.40)
- (12) *M. Claudi-o client-inegoti-um ded-it,*
 M.Claudius-DAT.M.SG client-DAT.M.SG task-ACC.N.SG gave(PRF)-3P.SG
ut virgin-emin servitut-em ad ser-eret neque
 that girl-ACC.F.SG in slavery-ACC.F.SG claim-SBJV.IMPF.3P.SG nor
ced-eret secundum libertat-em postulant-ibus
 give.in-SBJV.IMPF.3P.SG in.favour.of freedom-ACC.F.SG claimant-DAT.M.PL
vindici-as.
 legal.claim-ACC.F.PL
 'He instructed his client M. Claudius to claim the girl as his slave and not to give in to those who asked for her temporary freedom.' (Livy, *Ab Urbe Condita* 3, 44)

The beneficiary meaning of *secundum* NP may be linked to the proximity meaning 'alongside, near' of spatial *secundum* NP (example 13), and as such it can be explained with the metaphor IN FAVOUR OF IS ON THE SIDE OF outlined by Luraghi in her study on Greek prepositions (2003: 325, see also 2003: 118 fn. 11 and 2014: 115-119).

- (13) *Reliqu-i... Icar-um... secundum arbor-em fod-erunt.*
 Other-NOM.M.PL Icarus-ACC.M.SG near tree-ACC.M.SG buried-IND.PRF.3P.PL
 'The others... buried Icarus near a certain tree.' (Hyginus, *De astronomia* 2, 4, 10)

In the *Latin Library* Classical Latin sub-corpus there is one occurrence of the so-called "behalf beneficiary" (14), i.e. a type of beneficiary in which "the beneficent substitutes for the beneficiary"

(Luraghi 2010a: 97). Luraghi (2010b: 75) suggests that the behalf beneficiary meaning may have originated from prototypical beneficiaries through the following implicature: “acting in someone’s place usually implies acting for his/her benefit”.

- (14) *Qu-iesse-t* *secundum postum-um et natumet*
REL-NOM.M.SG be.SBJV.IMPF-3PSG in.place.ofchild-ACC.M.SG both born-ACC.M.SG and
mortuum here-s *institut-us...*
dead-ACC.M.SG heir-NOM.M.SG nominated-NOM.M.SG
‘[If] a man was appointed as substitute heir to a child (posthumously) who is born and then dies’ (Cicero, *De Oratore* 1, 180)

I suggest that the metaphorical and metonymical explanations provided in this section for the development of the beneficiary *secundum* NP may not be the only ones playing a role in this development. I will elaborate on the role of force dynamics in the development of the conformity and the beneficiary meanings in section 5, below.

5. A FORCE DYNAMICS INTERPRETATION OF CONFORMITY AND BENEFICIARY

The force dynamics model of construal of events (Talmy 1976, 1985, 1988, 2000a: 409-470) casts an interesting light on the semantics of conformity and beneficiary. The model of force dynamics is based on the idea that in many events there are (at least) two entities exerting force: “[t]he agonist is the entity that receives focal attention, and the antagonist is the entity that opposes the agonist, either overcoming the force of the agonist or failing to overcome it” (Evans 2007: 83). Both entities have a natural tendency towards either rest/inaction or motion/action.

In Latin, the evidence that *secundum* NP may have a force dynamic meaning component comes “indirectly” from the fact that both the conformity and the beneficiary functions form a pair of opposites with the preposition *contra* ‘against’.² Section 5.1, below, focuses on the conformity meaning and section 5.2 focuses on the beneficiary meaning.

5.1 FORCE DYNAMICS AND CONFORMITY

Examining collocations³ (McEnery & Wilson 2001: 23, McEnery & Hardie 2011: 122-133) for *secundum* NP and *contra* NP in the *Latin Library* corpus via the CQP web concordancer reveals that in Classical Latin prose both the *secundum* NP and the *contra* NP constructions have the noun *naturam* as their strongest immediate collocate (*secundum*:⁴ log-likelihood = 874.897; *contra*: log-likelihood = 291.398). The direct contrast between *secundum* and *contra* is expressed in (15):

2 However, this opposition does not hold in the spatial domain: just like in English, also in Latin *Mary walks along a path* is not opposed to *Mary walks against a path*, nor does *the house is near the river* oppose to *the house is (located/built) against the river*.

3 The query algorithms are *secundum_PREP* and *contra_PREP*.

4 The expression *secundum naturam* is a direct calque of Greek *katà phúsin* ‘according to Nature’, in Plato and Aristotle and later entered into Roman philosophy (especially Cicero) through Stoicism (Vernon Arnold 2014[1911]: 282, fn. 68). The expression *katà phúsin* contrasts with *parà phúsin* ‘against nature’ (Jerome, *Vulgata, The First Epistle of Paul to the Romans* 26-27). This meaning of *parà* as a malefactive (a type of beneficiary meaning ‘against, contrary to’) is not discussed in Luraghi’s (2003) study of Ancient Greek prepositions.

- (15) *Adpet-enturaudem, qu-aesecundumnatur-am*
 Seek-SBJV.PRS.3P.PL.PASS thereforeREL-NOM.F.PL according.to nature-ACC.F.PL
sint, declin-anturcontrari-a...; quod autem
 are.SUBJV.PRS.3P.PL reject- SBJ.PRS.3P.PL.PASS opposite-ACC.N.PL because in.fact
 [omn-e animal] refug-it id contra
 every-NOM.N.SG animal[N.SG] flee(PRS)-IND.PRS.3P.SG it.ACC.N.SG against
natur-am est...
 nature-ACC.F.Gis
 ‘So, all the things that accord with his/her nature are sought; the ones against nature are rejected...; every animal flees from what is against nature.’ (Cicero, *De Natura Deorum* 3, 3)

In this context, *contra* ‘against’ expresses a force dynamic configuration in which the trajector is the antagonist and the landmark is the agonist. This is particularly clear in contexts in which the agonist is represented by natural forces:

- (16) *Contra vimatque impet-um flumin-is convers-a.*
 Against force.ACC.F.SG and strength-ACC.M.SG river-GEN.N.SG turned.around-F.NOM.SG
 ‘Turned around against the force and the strength of the rivers’ (Caesar, *De Bello Gallico* 4, 17, 5)
- (17) *Ut contra vent-umgreg-em pasc-amus.*
 For against wind-ACC.M.SG herd-ACC.M.SG pasture-IND.PRS.1P.PL
 ‘In order to pasture the herd against the wind’ (Columella, *De Re Rustica* 7,3,2)

Besides natural forces, or the laws of nature itself, a context of opposition is provided also by lexical items having an intrinsic social force, such as laws (*legem* ‘law’, *ius* ‘law’, *foedus* ‘treaty’) and institutions (*rem publicam* ‘the state’, *patriam* ‘fatherland’), or religious force (*auspicia* ‘auspices’), as shown by the strongest immediate collocates of *contra* in table 2, below. Among these, the lexical item *vim* ‘force’ encapsulates the central meaning component. The tenth collocate, *Stoicos* ‘the Stoic philosophers’ indicates a group of people: intended as social agents with their will and intentions, people can also be conceptualised as agonists.

	Immediate Collocate	Log Likelihood
1	<i>naturam</i> ‘nature’	291.398
2	<i>legem</i> ‘law’	265.243
3	<i>rem</i> (publicam) ‘republic’	250.021
4	<i>leges</i> ‘laws’	225.799
5	<i>auspicia</i> ‘auguries’	156.708
6	<i>ius</i> ‘law’	149.06
7	<i>patriam</i> ‘home-country’	129.181
8	<i>foedus</i> ‘treaty’	106.283
9	<i>vim</i> ‘force’	99.423
10	<i>Stoicos</i> ‘the Stoics’	84.236

Table 2. Immediate collocates of *contra* ‘against’ in Classical Latin prose.

The force dynamic configuration observed for *contra naturam* ‘against nature’, i.e. one in which the landmark is the agonist and the trajector is the antagonist, applies also to *secundum* NP expressing conformity, whose collocates are also predominantly law-like. Table 3, below, shows the nominal collocates of *secundum* in Classical Latin prose: *tabulas* means ‘laws’ and *naturam* is understood in philosophical writings as meaning ‘the laws of Nature’.

	Collocate	Log-likelihood
1	<i>naturam</i> ‘nature’	780.065
2	<i>mare</i> ‘sea’	62.288
3	<i>quietem</i> ‘sleep’	62.09
4	<i>tabulas</i> ‘laws’	49.387
5	<i>deos</i> ‘Gods’	48.639
6	<i>proelium</i> ‘battle’	31.492
7	<i>eam</i> ‘this’	17.868
8	<i>ea</i> ‘these things’	5.5

(5 - *deos* ‘Gods’)

Other words which appear as arguments of *secundum* in the conformity construction are legal terms such as *decreta* ‘decrees’, *legem* ‘law’, *ius* ‘law’, *fas* ‘destiny’, *iudicium* ‘judgment’, and words having an element of regularity such as *consuetudinem* ‘habit’. Also these nouns have a semantic component of force, a feature they share with the collocates of *contra*, thus explaining why the two prepositions stand in contrast to one another.

Focusing on *secundum* NP, I argue that its force dynamic meaning component may have been inherited directly from the proto-scene of the verb *sequor* ‘follow’, in which the landmark has the force of imposing on the trajector the path and direction in which it should move.⁶ The contrast between *secundum* ‘in accordance with’ and *contra* ‘against’ is therefore one between compliance or opposition of a trajector (the antagonist) with respect to a landmark conceptualised as the agonist. Figures 2 and 3, below, represent the relationship encoded by *contra* ‘against’ and *secundum* ‘in accordance with’ based on Talmy’s (1988, 2000a, 2000b) force dynamics model. The circle is the agonist and the concave figure represents the antagonist.

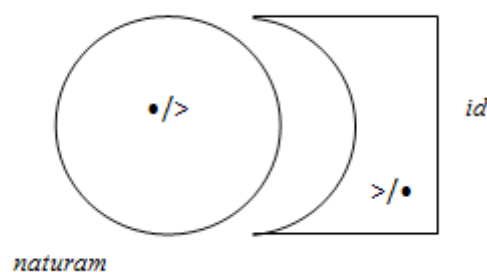


Figure 2. Force dynamics: *contra*. This diagram refers to example (15) above (*id contra naturam est* ‘what is against nature’).

5 I manually examined all the occurrences of *Deos* ‘Gods’, and found only one instance expressing conformity, whereas all the others express ranking.

6 It is possible to see also the constructions meaning ‘downstream’ as inherently encoding a force dynamic element, that is the force of gravity.

In terms of balance of strength between the agonist and the antagonist, in the case of *contra*, the antagonist may be halted by or overcome the agonist. In other words, *contra* simply signals a contrast between the intrinsic force tendencies of antagonist and agonist. The symbols > and • signal opposing tendencies, standing in contrast to one another: if the agonist is >, the antagonist is •, and vice versa.⁷

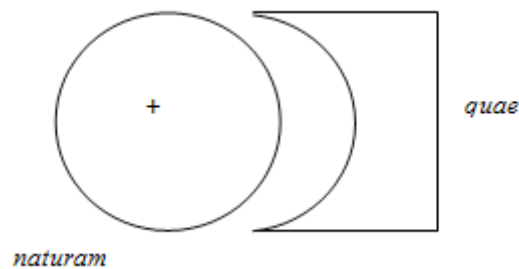


Figure 3. Force dynamics: *secundum* (conformity). This diagram refers to example (15) above (*quae secundum naturam sint* ‘the things which are against nature’).

In contrast to *contra*, *secundum* signals that regardless of the inherent force tendencies of agonist and antagonist (and regardless of whether these are opposed or not), the agonist follows the force of the antagonist (which is represented by the + sign in figure 3, above).

5.2 FORCE DYNAMICS AND BENEFICIARY

Just like the conformity meaning, beneficiary *secundum* ‘in favour of’⁸ also forms a pair with the preposition *contra* ‘against’, encoding a maleficiary, as shown by the contrast between example (18) and (19). A ‘maleficiary’, indicates a referent which is disadvantaged by a state of affairs.

(18) *Secundum te iudicat-um erit.*

In.favour.of.you.ACC judged-NOM.N.SG be.FUT.IND.3P.SG

‘It will be judged in your favour.’ (Aulus Gellius, *Noctes Atticae* 5, 10, 10)

(19) *Si ideo contra emptor-emi iudicat-um est, quod*

If therefore against buyer-ACC.M.SG judged-ACC.N.SG be.IND.PRS.3P.SG because
defu-it, non commit-itur stipulatio.

not.be(PRF)-IND.PRF.3P.SG not commit-IND.PRS.3P.SG.PASS stipulation.NOM.F.SG

‘If the judgment was pronounced against the buyer, because he did not show up [in court], the stipulation is not violated.’ (Ulpian, *Digesta* 21, 2, 55)

The explanation for the emergence of the maleficiary meanings relies on force dynamic interactions. In contexts involving motion, *contra* may indicate counterforce exerted by the agonist (the landmark) against the antagonist (the trajector). This applies to (20)-(21), below, in which there is a hostile context involving physical exertion between two entities.

⁷ Corpus frequencies may shed light on whether one balance of force is more prototypical than the other, but this goes beyond the scope of my study.

⁸ Another Latin preposition meaning ‘in favour of’ is *pro*.

- (20) *Contra Caesar-em... bell-um ger-ere coep-it.*
 Against Caesa-ACC.M.SG war-ACC.N.SG carry.out-INF.PRS start(PRF)-IND.PRF.3P.SG
 ‘He took up arms against Caesar.’ (Caesar, Aulus Hirtius or Gaius Oppius, *De Bello Alexandrino* 24)
- (21) *Ferre contrapatri-am arm-a ill-i*
 Take.INF.PRS.against home.country-ACC.F.SG weapon-ACC.N.PL these-NOM.M.PL
cum Coriolan-o debu-erunt.
 with Coriolan-ABL.M.SG must(PRF)-IND.PRS.3P.PL
 ‘These [friends] had to take up arms against their homeland together with Coriolan.’ (Cicero, *De Amicitia*, 11, 36)

It is possible⁹ that a force configuration of opposition was extended from warfare contexts like (20) and (21) to contexts of legal contention, an example of which is (19), above. Another example within the legal sphere is (22), below, representing the construction *AGERE contra aliquem* ‘to plead against someone’.

- (22) *Cum ager-em contrahomin-em disertissim-um*
 When do-SUBJV.IMP.1P.SG against man-ACC.M.SG eloquent-ACC.M.SG
nostr-ae civitat-is...
 POSS.1P.PL-GEN.F.SG city-GEN.F.SG
 ‘When I was pleading against the most eloquent man of our city...’ (Cicero, *Pro Caecina* 33, 97)

In (22), the trajector-antagonist is the writer, Cicero, raising a law-suit against Cotta, who is the affected landmark-agonist, i.e. the ‘damaged’ party. This force dynamic arrangement is also present in (19), above, in which it may however appear less clear because the passive voice backgrounds the trajector-antagonist (presumably a judge). In (19) the person who is in charge of judging (the agonist), has more power than the antagonist, the buyer, who is construed as negatively affected by the judge’s unfavourable decision and therefore as a damaged party. On the basis of this analysis, I argue that there is a parallelism¹⁰ between physical warfare contexts and legal ones. In both, the antagonist exerts force over the agonist which is affected either by some (potential) physical damage or unfavourable (legal) action. Talmy’s force dynamics model does not encompass a separate symbol representing maleficiaries so the representation of the construction NP₁ *contra* NP₂, where NP₂ fulfils the semantic role of maleficiary, is the same as figure 3 above.

Having clarified a possible force dynamic development for the maleficiary meaning of *contra* ‘against,’ my discussion now turns to the construction NP₁ *secundum* NP₂ (NP₁ ‘in favour of’ NP₂), where NP₂ expresses the beneficiary meaning. All 24 instances of *secundum* NP with the beneficiary meaning in the Classical Latin sub-corpus of the *Latin Library* occur in legal contexts of decision making resulting in someone being favoured (i.e. a beneficiary). As discussed in section 4 above, two instances show the fixed phrase *secundum libertatem/servitutem vindicias DARE/DIRE/DECERNERE* ‘to enslave someone, to free someone’ (lit. to decide a claim in favour of freedom/slavery). The majority of

9 My data does not show diachronic evidence for this extension (legal uses are already attested in the classical period), so the meaning extension is merely stated here as a possibility on the basis of the fact that more abstract meanings tend to be derived from more physical/concrete meanings.

10 Whether this parallelism corresponds to historical development cannot be ascertained on the basis of my corpus data.

instances of beneficiary *secundum* NP (19/24 occurrences) in the *Latin Library* occur with the VPs *IUDICARE* ‘to judge’, *DECERNERE* ‘to judge, vote, declare’, *decretum/iudicium FACERE* ‘to make a decision’, *sententiam DICERE* ‘cast/express a vote’, *ius DARE* ‘to make/give the law’, *DARE* ‘to give (a vote)’, and *DISPUTARE* ‘decide’.

- (23) *Itaque, Lutat-i... secundumtelit-emd-o.*
 So Lutatius-VOC.M.SG in.favour.ofyou.ACC suit-ACC.F.SG give-IND.PRS.1P.SG
 ‘So, Lutatius, I judge in your favour.’ (ValeriusMaximus, *FactorumetDictorumMemorabiliumLibri*, 2, 8, 2)
- (24) *Iudici-um secundumHeracti-um... fact-um er-at...*
 Judgment-ACC.M.SG in.favour.of Heraclius-ACC.M.SG done-ACC.M.SG be-IND.IMPF
 ‘[since] the judgment... had been made in favour of Heraclius’ (Cicero, *In Verrem* 2, 2, 66)

An interesting preliminary observation regards the speakers’ choice of *secundum* NP to encode beneficiary in these contexts instead of the two competitor constructions expressing beneficiary in Latin: the dative ‘of interest’ and the *pro* NP_{ABL} construction. In one of its uses, the dative case expresses the entity to whose advantage or disadvantage a certain state of affairs occurs: in some cases, the reading of beneficiary or maleficiary depends on the semantics of the predicate,¹¹ (*utilis est* ‘is useful’ in example (25), below, while in other cases it is pragmatically determined in context as in example (26), also below).

- (25) *Lex utilisest plebi.*
 Law-NOM.F.SG useful-NOM.F.SG be-3P.SG people.DAT
 ‘This law is useful for the people.’ (Example and glosses after van Langendonck 1998: 234)
- (26) *Lupo est homo homini.*
 Wolf-NOM.M.SG be-3P.SG man-NOM.M.SG man-DAT.M.SG
 ‘Man is a wolf to man.’ (ibid.)

The so called *dativus commodi/incommodi* ‘dative of advantage/disadvantage’ (Pinkster 2015: 855 among others) occurs with verbs of helping/caring (and their opposites); verbs of pleasing, flattering and threatening; verbs of ruling, obeying and serving and verbs of approaching and befalling (ibid.); but not with verbs of decision. It is possible that the dative is not suitable in legal contexts of decision, in which nothing should be left open to interpretation, and the meaning of advantage or disadvantage needs to be unambiguously spelled out.

Two last instances of beneficiary *secundum* NP in the *Latin Library* corpus feature the verb *DARE* ‘give’ in the context of the inheritance of goods, whose allocation is decided by law or a judge. An example of this is (27), below. In these instances, *secundum* NP does not express the recipient of the physical transfer of goods, but the person to whose advantage the decision of transferring the goods is made. In other words, in these contexts *DARE* ‘give’ indicates the transfer of ownership, not possession, two distinct concepts in Roman Law (Sohm et al. 1892: 252). If the dative ‘of interest’

11 The dative can also be taken by adjectives (*utilis/inutilis tibi* ‘useful/useless to you’) or adverbs (e.g. *congruenter naturae vivere* ‘to live in accordance with nature’, Cicero *De Finibus* 3, 9, 30).

was used in this context, its function could be confused with the recipient of DARE ‘give,’ thus suggesting transfer of physical possession instead of the act of granting someone legal rights over some goods. *Secundum* NP, therefore, indicates the person to whose benefit a legal decision regarding conveyance is made.

- (27) *Si intestat-usmortu-us esset, ita secundum*
 If nominated-NOM.M.SG dead-NOM.M.SG be.SUBJV.IMPF.3P.SG so in.favour.of
eum possessi-odar-etur.
 him.ACC.M.SG possession-NOM.F.SG give-SUBJV.IMPF.3P.SG.PASS
 ‘If the person nominated in the will was dead, the possession [of his goods] would be granted to him.’ (Cicero, *In Verrem* 2, 1, 44)

The *pro* NP_{ABL} construction is widely used in legal contexts when the idea of advantage is associated with the idea of protection or defence, as shown in many titles of Cicero’s speeches (e.g. *Pro Sulla* ‘in defence of Sulla’) and in (28).

- (28) ... *et apud popul-um... cum pro Milone dic-eret,*
 And among people-ACC.M.SG as in.defence.ofMilo speak-SUBJV.IMPF.3P.SG
clamor-econvici-o-queiact<at>-us est...
 noise-ABL.M.SG scorn-ABL.N.SG-and thrown-NOM.M.SG be.IND.3P.SG
 ‘In the public meeting he was shaken by noise and abuse as he spoke in defence of Milo.’
 (Cicero, *Ad Familiares* 1.5b.1)

So, since a judgment, unlike a lawyer’s speech, is not pronounced in defence of someone, *pro* NP_{ABL} is also unsuitable to express the role of beneficiary in context of legal choices. In sum, the corpus distribution of *secundum* NP with beneficiary meaning allows us to understand why neither the dative case nor the *pro* NP_{ABL} prepositional phrases are suitable to express a meaning of advantage in legal contexts of decision making. What follows explains why *secundum* NP, unlike the dative ‘of interest’ and the *pro* NP_{ABL} construction, is semantically suitable to express the beneficiary role.

Apart from the fixed expressions *secundum libertatem/servitutum vindicias DARE/DIRE/DECERNERE* ‘to enslave someone, to free someone (lit. to decide a claim in favour of freedom/slavery)’, the argument of *secundum* with beneficiary meaning refers to a human participant. In the context of legal decision making, and even more clearly litigation (as in examples (23)-(24), above), each of the people whose case is being decided have a will, understood as a particular and most preferred desired outcome, in contrast to that of their opponents in a legal case. The beneficiary *secundum* NP expressions can be understood as conformity expressions, and paraphrased as ‘in accordance with the will of NP’ (i.e. the argument of *secundum*), as suggested in Lewis & Short (1879 *secundum*: B2). *Secundum* NP in these contexts is not a modifier of the whole VP, indicating the manner in which the state of affairs is carried out. For instance, if *secundum* NP had scope over *litem do* in (23) and *iudicium factum erat* in (24), *secundum* NP would indicate that the judge made a decision in a way that conforms to the reader’s will or Heraclius’s will. Instead, *secundum* NP modifies the direct object within the predicate, specifically *litem* in (23) and *iudicium* in (24): it is

the judgment which is in line with the reader's or Heraclius's will. The paraphrases in (29) and (30), below, capture this for (23) and (24) respectively.

(29) So, Lutatius, I make a decision/judgment (which is) in accordance with your preferred desire/outcome.

(30) So, the judgment had been made (which is) in accordance with the will of Heraclius.

A judgment matching someone's desired outcome can also be conceptualised as a decision in favour of that person: this may explain the shift in meaning from conformity to beneficiary. This shift is inferential (metonymic) and not metaphorical.

In terms of force dynamics, the conceptualisation underlying the beneficiary meaning of *secundum* NP relates to the semantic frame of compliance (*FrameNet: Compliance*), which was noted in section 3, above, for the meaning of conformity. This frame is characterised by two entities: one entity setting direction and path in the physical world or dictating some rules in the psycho-social world, and the other entity complying with them. However, in the case of beneficiary *secundum* NP there is a difference: the landmark does show a semantic component of force (the person's will), but this force is not stronger than the power exercised by the decision-maker (i.e. the judge in the legal contexts examined) (cf. figure 3, above).

To summarise, this section started off by considering the maleficiary expressions *contra* NP constituting a pair of opposites with the beneficiary *secundum* NP constructions. Both maleficiary and beneficiaries have a force dynamics component. The beneficiary function connects to the semantic frame of compliance, thus showing similarities with the conformity construction. Looking at the scope of *secundum* NP lent further support to the connection between conformity and beneficiary. This link is metonymic in nature (inference). Finally, a contrastive analysis between *secundum* NP and its competitor beneficiary constructions (the bare dative case and pro NP_{ABL}) explained why *secundum* NP is the favourite expression in legal contexts of decision. To conclude, force dynamics, metonymy and the legal contexts of decision making are all factors that may have operated alongside the metaphorical mapping IN FAVOUR OF IS ON THE SIDE OF (described in section 4, above) for the emergence of the beneficiary meaning of *secundum* NP.

6. CONCLUSION

This paper provided a corpus-based cognitive semantic analysis of the conformity and the beneficiary meanings of the *secundum* NP construction through the history of Latin, thus filling a gap in the literature on Latin prepositions. The meaning of conformity is the most frequent over time, reaching a peak in Silver Latin (100-258) whereas the beneficiary meaning is marginal and declining until its disappearance in Early Medieval Latin (from 530).

First, this paper explained the meaning of conformity discussing its possible origin coming from

expressions meaning ‘downstream’ via the metaphor CONFORMITY IS MOTION ALONG. This hypothesis is evaluated by considering frequency, ontological salience and the distinction between written and spoken language. Without completely ruling out the possibility of an extension from an expression meaning ‘downstream’, I argue that an explanation based directly on the extraction of schemas from the proto-scene and force dynamics appears to be more plausible.

Next, this paper considered the beneficiary meaning, ranging from the more to the less prototypical instances. I argued that the beneficiary meaning is connected to the spatial meaning of *secundum* NP via the metaphor IN FAVOUR OF IS ON THE SIDE OF, and that the ‘behalf beneficiary’ indicating substitution may be connected to the prototypical beneficiary via implicature, a process which is metonymic in nature. On top of metaphor and metonymy this paper showed that also force dynamics possibly also play a role, connecting the beneficiary meaning directly to the meaning of conformity.

The evidence that *secundum* NP may have a force dynamic meaning component comes indirectly, namely from the fact that both the conformity and the beneficiary functions form a pair of opposites with the preposition *contra* ‘against’.

I argued that the force dynamic component of the conformity meaning derives directly from the basic scenario of the verb *sequor* ‘follow’, in which the landmark has the force of imposing on the trajector its path and direction of movement. I also showed that the force dynamic component of the beneficiary *secundum* NP is somehow connected to the semantic relation of compliance of the trajector on the path and trajectory imposed by the landmark. However, the landmark’s strength is not greater than the trajector’s. Furthermore, in legal contexts, a metonymical process may take place whereby a judgment conforming to someone’s desired outcome is reinterpreted as a decision in favour of that person. The force dynamics interpretation comes together with metonymy (inference) in explaining the meaning connection between conformity and beneficiary.

To conclude, this paper showed a range of mechanisms which come into play in explaining the development of the conformity and the beneficiary meanings of *secundum* NP. Whilst some metaphorical and metonymical processes had already been suggested in the literature on translational equivalents and Romance continuers of *secundum* NP, the role played by force dynamics had been previously overlooked. Against this background, my paper provided a complete and integrated account of the possible mechanisms and motivations underlying the development of the conformity and beneficiary meanings of Latin *secundum* NP, thus adding to a body of research on prepositions within cognitive linguistics.

ABBREVIATIONS

1P First person; 3P Third person; ABL Ablative; ACC Accusative; AOR Aorist; DAT Dative; F Feminine; FUT Future; GEN Genitive; GER Gerund; IMPF Imperfect; IMPFV Imperfective; IND Indicative; INF Infinitive; M Masculine; M/P Medio/Passive; N Neuter; NOM Nominative; NP Noun

phrase; PASS Passive; PL Plural; PLUPERF Pluperfect; POSS Possessive; PRF Perfect; PRS Present; PTC Participle; REL Relative pronoun; SBJV Subjunctive; SG Singular; VOC Vocative; . Functions of portmanteau morphemes or functions of morphemes whose segmentation is not shown; - Morpheme boundary; (...) Non overt element or sound change.

REFERENCES

ADAM, A. *Roman Antiquities: Or, an Account of the Manners and Customs of the Romans*. New York: Duyckinck, T.A. Ronalds, S.A. Burtin, 1807.

CROFT, W. & CRUSE, D. A. *Cognitive Linguistics*. Cambridge: Cambridge University Press, 2004.

EVANS, V. *A Glossary of Cognitive Linguistics*. Edinburgh: Edinburgh University Press, 2007.

FERRARI, L. A gramaticalização de formas não-finitas como evidência da motivação conceptual do léxico. *Revista de Estudos Linguísticos Veredas* 2/1, 1998, pp. 103-115.

GARNIER, R. Sur les para-participes en *-cundus* du latin. Paper read at the *18th International Colloquium in Latin Linguistics*, Jean Jaurès University, Toulouse, 2015.

GUARDAMAGNA, C. *Secundum*: A cognitive account of the spatio-temporal domains including logical extensions of space. SPEVAK, O. (ed.) *Études de linguistique latine II, Pallas* 103, Spring 2017, (forth. 2017).

GUARDAMAGNA, C. Following the path from 'follow' to 'according to'. Interpreting layering of functions of Italian *secondo*, *seguendo* and *a seconda di* as clues to the grammaticalisation of Latin *sequor* via *secundum*. In KAUFHOLD, K., MCCULLOCH, S. & TOMINC, A. (eds.) *Papers from the Lancaster University Postgraduate Conference in Linguistics and Language Teaching* 5, 2011, pp. 48-76.

HARDIE, A. CQPweb - combining power, flexibility and usability in a corpus analysis tool. *International Journal of Corpus Linguistics* 17/3, pp. 380-409, 2012.

HEINE, B. & KUTEVA, T. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press, 2002.

LAFFI, U. *Colonie e municipi nello stato romano*. Rome: Edizioni di Storia e Letteratura, 2007.

LANGACKER, R.W. *Foundations of Cognitive Grammar. Volume I. Theoretical prerequisites*. Stanford: Stanford University Press, 1987.

- LEWIS, C.T. & SHORT C. *A Latin Dictionary*. Oxford: Clarendon Press, 1879[1962].
- LURAGHI, S. *On the Meaning of Prepositions and Cases*. Amsterdam: Benjamins, 2003.
- LURAGHI, S. “A model for representing polysemy: The Italian preposition *da*”. In GILBERT, J.F.E., GUIMIER, C. & KRAUSE, M. (eds.), *Actes du Colloque “Autour de la préposition”*. Caen: Presses Universitaires de Caen, 2009, pp. 167-178.
- LURAGHI, S. Where do beneficiaries come from and how do they come about? Sources for Beneficiary expressions in Classical Greek and the typology of beneficiary.” In WINTERS, M.E., TISSARI, H. & ALLAN, K. (eds.) *Historical Cognitive Linguistics*. Berlin: de Gruyter, 2010a, pp. 93-132.
- LURAGHI, S. Adverbial phrases. In BALDI, P. & CUZZOLIN, P. (eds.) *New Perspectives on Latin syntax. Volume 2. Constituent syntax: adverbial phrases, adverbs, mood, tenses*. Berlin: Mouton de Gruyter, 2010b, pp. 19-109.
- LURAGHI, S. Plotting diachronic semantic maps. The role of metaphor. In LURAGHI, S. & NARROG, H. (eds.) *Perspectives on Semantic Roles*. Amsterdam: Benjamins, 2014.
- MCENERY, T. & HARDIE, A. *Corpus Linguistics: Method, theory and practice*. Cambridge: Cambridge University Press, 2011.
- MCENERY, T. & WILSON, A. *Corpus Linguistics: An introduction*. Second Edition. Edinburgh: Edinburgh University Press, 2001.
- NETTLESHIP, H. *Contributions to Latin Lexicography*. Oxford: Newby, 1889.
- PINKSTER, H. *Oxford Latin Syntax. Volume 1: The simple clause*. Oxford: Oxford University Press, 2015.
- POULTNEY, J.W. The phonology of *-nd-* and the Latin gerundive. In IZZO, H.J., *Italic and Romance. Linguistic Studies in Honor of Ernst Pulgram*. Amsterdam: Benjamins, 1980.
- ROCHA, L.F. Análise do caráter polissêmico do vocábulo latino SECUNDUS. *Principia Caminhos da Iniciação Científica*, v. 3, Universidade Federal de Juiz de Fora: Edufjf, 1998, pp.173-9,
- SCHMID, H.J. Entrenchment, Salience and Basic Levels. In GEERAERTS, D. & CUYKENS, H. (eds.) *The Oxford Handbook of Cognitive Linguistics*, Oxford: Oxford University Press, 2010, pp. 117-139.

SOHM, R., LEDLIE, J.C. and GRUEBE, B.E. *The institutes of Roman Law*. Oxford: Oxford University Press, 1892.

TALMY, L. Semantic causative types. In SHIBATANI, M. (ed.) *Syntax and Semantics 6: The grammar of causative constructions*. New York: Academic Press, 1976, pp. 43-116.

TALMY, L. Lexicalization patterns: semantic structure in lexical forms. In SHOPEN, T. (ed.) *Language Typology and Syntactic Description. Volume 3. Grammatical categories and the lexicon*. Cambridge: Cambridge University Press, 1985, pp. 57-149.

TALMY, L. Force dynamics in language and cognition. *Cognitive Science* 12, 1988, pp. 49-100.

TALMY, L. *Toward a Cognitive Semantics. Volume I. Concept structuring systems*. Cambridge, MA: MIT Press, 2000a.

TALMY, L. *Toward a Cognitive Semantics. Volume II. Typology and process in concept structuring*. Cambridge, MA: MIT Press, 2000b.

TAYLOR, J. R. Prepositions: patterns of polysemisation and strategies of disambiguation. In ZELINSKY-WIBBELT, C. (ed.) *The Semantics of Prepositions*. Berlin: Mouton de Gruyter, 1993, pp. 151-175.

TRABELSI, H. *La grammaticalisation de la preposition. Étude sémantique et pragmatique de selon, suivant et d'après*. City Unknown: Presses Académiques Francophones, 2013.

TYLER, A. & EVANS, V. *The Semantics of English Prepositions. Spatial scenes, embodied meaning and cognition*. Cambridge: Cambridge University Press, 2007.

VERNON ARNOLD, E. *Romain Stoicism*. London: Routledge, 2014[1911].

VINEIS, E. "Latin". In: GIACALONE RAMAT, A. & RAMAT, P. (eds.) *The Indo-European Languages*. London: Routledge, 1998, pp. 261-321.

The Latin Library corpus. Available at <https://cqpweb.lancs.ac.uk/latinlib/> Last accessed 30/12/2015.

FrameNet: <https://framenet.icsi.berkeley.edu/fndrupal/about/> Last accessed 30/12/2015.

Recebido em 06/01/2016

Aceito em 15/03/2016

LUDI: UM *FRAMEWORK* PARA DESAMBIGUAÇÃO LEXICAL COM BASE NO ENRIQUECIMENTO DA SEMÂNTICA DE FRAMES

Ely Edison da Silva Matos (UFJF)¹, Maria Margarida Martins Salomão (UFJF)²

RESUMO

Enquanto no âmbito da Sintaxe as aplicações em Processamento da Língua Natural estão relativamente bem estabelecidas, no âmbito da Semântica não se observa ainda a mesma maturidade. Este trabalho é uma contribuição para os estudos em Semântica Computacional, com base na Linguística Cognitiva, reunindo ferramentas computacionais e recursos linguísticos. O objeto de estudo é a desambiguação de Unidades Lexicais. O resultado é uma aplicação computacional, chamada *Framework* LUDI (*Lexical Unit Discovery through Inference*). O *framework* pode ser integrado em ferramentas para recuperação de informação, sumarização, *parsers* semânticos e em processos de Etiquetagem de Papéis Semânticos.

PALAVRAS-CHAVES: Desambiguação do Sentido de Lexemas, Semântica Computacional, Semântica de Frames, Léxico Gerativo.

ABSTRACT

While in the field of Syntax, applications in Natural Language Processing are relatively well established, the situation is different for the field of Semantics. This work is a contribution to the studies in Computational Semantics, using ideas from Cognitive Linguistics and bringing together computational tools and linguistics resources. The object of study is disambiguation of Lexical Units. The result is an application, called *Framework* LUDI (*Lexical Unit Discovery through Inference*). The framework can be integrated into information retrieval and summarization tools, semantic parsers and into Semantic Role Labeling process.

KEYWORDS: Lexeme Sense Disambiguation, Computational Semantics, Frame Semantics, Generative Lexicon.

1 Universidade Federal de Juiz de Fora

2 Universidade Federal de Juiz de Fora

1. INTRODUÇÃO

Os maiores sucessos do Processamento da Língua Natural (PLN) estão na área da Sintaxe, apoiados em ideias da Linguística Gerativa. O crescente uso de modelos estatísticos e probabilísticos em PLN expõe claramente, no entanto, que não podemos esperar que a linguagem esteja presa apenas a princípios e parâmetros. Por outro lado, na área da Semântica Computacional, cujo objetivo é a captura do significado de expressões em língua natural (BOS, 2011), o instrumento por excelência tem sido a lógica formal, em especial a Lógica de Primeira Ordem. Mas neste campo a dificuldade não é menor, pois fenômenos como a estrutura de eventos, a aspectualidade, os contextos modais, a quantificação e a resolução de anáforas resistem a uma formalização.

Este trabalho é fruto da tese de doutorado apresentada em Matos (2014), com o objetivo de implementar um *framework* para a desambiguação de itens lexicais. O trabalho pode ser enquadrado na área do PLN, por estar relacionado às técnicas de WSD (*Word Sense Desambiguation*). Porém, ele se afasta do *mainstream* em três aspectos principais: (i) adota-se a filosofia, os conceitos e a terminologia proposta pela Linguística Cognitiva, (ii) apoia-se na Semântica de Frames, uma abordagem que privilegia a compreensão da linguagem em lugar de sua formalização e (iii) é utilizado um Modelo Conexionista Estruturado, sem apelar (diretamente) para uma abordagem baseada em probabilidades.

O resultado do trabalho é corporificado na forma de uma aplicação computacional, chamada Framework LUDI (*Lexical Unit Discovery through Inference*), composta por algoritmos e estruturas de dados usados na desambiguação. O *framework* é uma aplicação de Compreensão da Língua Natural, que pode ser integrada em ferramentas para recuperação de informação, sumarização, *parsers* semânticos e em processos de Etiquetagem de Papéis Semânticos.

O texto foi dividido em cinco seções, além desta Introdução. A seção 2 aborda a questão do significado das expressões linguísticas, além de alguns pressupostos e hipóteses adotadas no trabalho. A seção 3 apresenta aspectos linguísticos e computacionais usados na concepção do *framework* e na sua implementação. A seção 4 descreve o *framework* LUDI, abordando o modelo conexionista e a construção da rede semântica. A seção 5 mostra a aplicação do *framework* no estudo da polissemia do verbo <abrir.v>. Finalmente, a última seção apresenta algumas conclusões e sugestões de trabalhos futuros.

2. A QUESTÃO DO SIGNIFICADO

As seguintes sentenças não devem apresentar dificuldades de compreensão para um leitor ou ouvinte humano:

- i. João abriu a janela.
- ii. João abriu sua boca.
- iii. João abriu o livro.

- iv. João abriu as cortinas.
- v. O cirurgião abriu o peito de João.

Nestes exemplos o significado de <abrir> é função do conhecimento que temos de como as entidades e objetos de diferentes tipos são “abertos” e do contexto da enunciação. A contribuição semântica de <abrir>, portanto, é bastante variável. Porém, se consideramos usos mais figurativos a dependência do contexto se torna ainda mais marcante:

- i. João abriu uma conta no banco.
- ii. João abriu a reunião.
- iii. João está aberto ao diálogo.
- iv. O policial abriu fogo.
- v. “Aquarela do Brasil” abre a coletânea.
- vi. Finalmente ele se abriu com ela.

O significado de <abrir> agora se relaciona com várias ações, eventos, situações e agentes. Apesar desta diversidade, que pode incluir usos idiomáticos, metonímicos e metafóricos, um leitor humano ainda teria pouca ou nenhuma dificuldade para compreensão das sentenças.

A questão muda se considerarmos a legibilidade por máquinas. Desprovidas de um “conhecimento do mundo” e de posse de um contexto limitado (quando este é fornecido), as aplicações computacionais têm muita dificuldade para compreender um texto, se comparadas com os seres humanos. Estas dificuldades podem estar associadas a dois fatos, que se destacam no PLN: o foco em questões relacionadas à Sintaxe (com abordagem gerativista) ou, quando a Semântica é considerada, o uso quase exclusivo de teorias associadas a Semântica Formal (baseada em Lógicas).

Existem alternativas. A abordagem cognitivista da linguagem propõe que o processo de significação, ao invés de ser um processo em que uma “realidade” é “representada”, é um processo em que são evocadas estruturas cognitivas complexas, que se articulam em redes e que categorizam ou analisam a experiência do mundo. Um exemplo desta abordagem é a Semântica da Compreensão (*Understanding Semantics*, ou *U-semantics*) (FILLMORE, 1985). A *U-semantics* destaca algumas limitações das chamadas “condições de verdade” e da categorização baseada em condições “necessárias e suficientes”, pontos chave das semânticas de orientação formalista. Ela estabelece que nenhum aspecto do significado, que seja relevante para a compreensão, deve ser excluído da análise apenas porque as premissas metodológicas da teoria escolhida não consideram o conhecimento de fundo (ZIEM, 2014).

A proposta em Fillmore (1985) é estruturar este conhecimento de fundo através de *frames*. Os *frames* são estruturas cognitivas, formando uma rede, usadas para organizar nossas experiências e, ao mesmo

tempo, funcionar como ferramentas para a descrição e explicação dos significados lexical e gramatical. Assim, a chamada Semântica de Frames (*Frame Semantics*) (FILLMORE, 1982; FILLMORE, 2006), propõe que um item lexical só pode ter seu significado esclarecido quando relativizado a um *background*, expresso pelo *frame*.

No nosso caso, o uso da Semântica de Frames em uma aplicação real (TORRENT et al., 2014), que implementa um dicionário multilingual baseado em *frames*, apresentou uma questão eminentemente prática: se um lexema polissêmico está associado a mais de um *frame*, qual *frame* deve ser exibido ao usuário quando ele consultar o dicionário informando um lexema polissêmico, como no caso de <abrir>?

A resposta a esta questão deveria ser óbvia: devem ser exibidos todos os *frames* associados com aquele lexema. Porém, neste caso, o dicionário baseado em *frames* não se distinguiria dos dicionários comuns, que exibem todos os sentidos de um lexema na mesma entrada ou criam entradas repetidas com o lexema para cada sentido. Caberia ao usuário o processo da desambiguação (o que não deixa de ser paradoxal, pois se ele já conhece o sentido do que procura, o dicionário pouco pode lhe ajudar).

Porém, a questão motivadora do trabalho é outra: se o usuário fornecer “pistas” a respeito do que procura, por exemplo informando mais uma ou duas palavras, não é possível fornecer a ele uma resposta mais precisa? Afinal, se cada palavra pode evocar um ou mais *frames* e os *frames* estão relacionados entre si em uma rede, estas relações poderiam ajudar a estabelecer o *background* proposto pela Semântica de Frames, contra o qual o sentido mais correto (ou mais próximo) da palavra poderia ser descoberto. O processamento no *framework* LUDI é justamente a travessia desta rede, com o propósito de descobrir qual sentido (qual *frame*) está associado ao item lexical e, portanto, promover a desambiguação.

Além dos pressupostos teóricos inerentes à Semântica de Frames, o trabalho adota alguns pressupostos mais específicos:

- a) Os lexemas só tem significação quando associados a um *frame*. Mais claramente, somente Unidades Lexicais (um lexema associado a um *frame*) possuem significação. Esta questão é importante, já que as teorias sobre léxico postulam, quase unanimemente, uma significação para os lexemas independentemente do uso (GEERAERTS, 2009);
- b) O léxico é visto como uma extensão dos modelos cognitivos (no caso, dos *frames*). As Unidades Lexicais são tratadas igualmente como dispositivos cognitivos com uma estrutura interna. Ainda que esta estrutura seja insuficiente para atribuir um significado por si só, é ela que permite a composição da Unidade Lexical em um cenário mais amplo;
- c) O significado de uma sentença é alcançado através de um processo dinâmico de composicionalidade em duas vias: não apenas o significado da sentença está relacionado com a composição dos significados das palavras que a formam, mas o próprio significado das palavras

dentro da sentença é resultado da composicionalidade. Ou seja, enquanto o significado da sentença é função dos *frames* evocados pelos lexemas presentes na sentença, o *frame* escolhido para cada lexema é aquele que melhor se encaixa na composição.

Estes pressupostos fundamentam as seguintes hipóteses básicas do trabalho:

- a) Se o significado é dependente do uso do item lexical, os testes devem ser feitos com dados obtidos a partir da língua em uso, ou seja, dados de corpora;
- b) Se o significado pode ser obtido como produto da composição dos *frames* e esta composição se dá usando as relações existentes entre os *frames*, aumentar o número destas relações pode fornecer uma compreensão melhor do significado;
- c) Se a variação da valência de um verbo é um critério para avaliar a variação do sentido de um lexema (com a definição de um novo *frame*), a interface sintaxe-semântica deve ser considerada no processo de desambiguação;
- d) Se o item lexical possui uma estrutura interna, é necessário o apoio de uma teoria que descreva essa estrutura.

3. ASPECTOS LINGÜÍSTICOS E COMPUTACIONAIS

Esta seção apresenta, de forma bastante resumida, aspectos linguísticos e computacionais que foram considerados na elaboração do trabalho. Estes aspectos ajudam a contextualizar o trabalho nos cenários da Linguística Cognitiva e da Linguística Computacional.

3.1. CATEGORIZAÇÃO E POLISSEMIA

Os lexemas estão associados a conceitos. Se os conceitos não existem de forma isolada na mente, mas estão organizados de alguma forma, o léxico também deve possuir uma estrutura. O léxico não é apenas uma longa lista de lexemas isolados e desconectados, sendo possível reconhecer vários tipos de relações e conexões entre os lexemas que nos ajudam a estabelecer regularidades e reconhecer padrões. Fundamentalmente, podemos considerar que as relações lexicais representam relações entre conceitos e que, inversamente, a categorização de conceitos também deve se refletir na categorização dos lexemas que os evocam.

A Teoria Clássica de conceitos propõe que estes são categorizados com base em definições, que codificam as condições necessárias e suficientes para a sua aplicação. Um conceito pertencerá a uma categoria se apresentar o conjunto de características daquela categoria. Por exemplo, a categoria <pássaro> pode ser definida pelas características: um ANIMAL, que TEM BICO e VOA.

A abordagem cognitivista levanta vários questionamentos contra esta categorização formalista. Em primeiro lugar, os estudos feitos por Rosch e descritos em Lakoff (1987) mostram que as categorias são tanto graduadas em termos de tipicidade quanto estruturadas internamente. Em uma categoria, alguns membros são mais típicos que outros, funcionando como protótipos. Em segundo lugar, as categorias cognitivas possuem limites imprecisos, ou seja, há gradações no pertencimento de um elemento a uma dada categoria. Uma terceira questão diz respeito à completude das condições necessárias e suficientes. Um pássaro empalhado, por não voar, deixa de ser classificado como pássaro? É preciso definir uma nova categoria de pássaros ou ampliar/detalhar as condições de classificação como <pássaro>?

Estas questões se refletem na categorização dos lexemas: um mesmo lexema pode apontar para mais de um conceito. Neste caso, pode se tratar de homonímia ou de polissemia. Embora não exista um consenso na diferenciação dos dois casos (LYONS, 1979), pode-se considerar, simplificada, que na homonímia não há relação semântica entre os conceitos evocados pela palavra - é comum afirmar que na realidade são duas palavras diferentes que calharam de ter as mesmas propriedades formais (CRUSE, 2010). Na polissemia, os sentidos evocados pela palavra estão de alguma forma relacionados. Mas qual a estrutura destas relações?

Dentre as estruturas propostas, destacam-se as categorias radiais, conforme o modelo descrito em Brugman e Lakoff (1988) e Lakoff (1987), o Modelo de Redes Esquemáticas (LANGACKER, 1987) e as Redes Semânticas (SOWA, 2015). Independente da estrutura adotada, não devem ser ignorados os usos metonímicos e metafóricos dos lexemas que, de certa forma, estendem os seus sentidos centrais.

Outro caminho para se estudar a distinção entre sentidos é a análise de dados em corpora (GRIES, 2006). O uso de dados reais permite perceber que a contribuição semântica de uma dada palavra está associada à situação interpretativa no contexto em que ela é usada. Além disso é possível observar as variações de sentido associadas aos padrões sintáticos e a influência de outras palavras em situação de coocorrência. Este cenário apresenta uma grande semelhança com a abordagem proposta pela Semântica de Frames.

3.2. SEMÂNTICA DE FRAMES E FRAMENET

A Semântica de Frames é uma abordagem para estudo do significado lexical baseado nos trabalhos de Fillmore e seus colaboradores ao longo dos últimos 35 anos (FILLMORE, 1977), (FILLMORE, 1982), (FILLMORE, 1985), (FILLMORE, 2006), (PETRUCK, 1996), (FILLMORE; BAKER, 2010), (GAWRON, 2008). A ideia central da Semântica de Frames é que o significado de uma palavra deve ser descrito em relação a um *frame* semântico. Um *frame*, neste contexto, é uma representação esquemática de estruturas conceituais e padrões de crença, práticas, instituições, etc. que provê a base para uma interação significativa em uma comunidade de fala.

Um *frame* representa um sistema de conceitos relacionados entre si de tal forma que, para compreender um deles é necessário compreender toda a estrutura onde ele se situa. O significado de uma expressão

linguística é interpretado de acordo com o *frame* que representa a cena no momento da enunciação. Por exemplo, o lexema <quebrar> pode significar <partir, romper> (“O copo quebrou”) ou <falir> (“A empresa quebrou”), dependendo do *frame* “evocado” e dos elementos linguísticos que ocorrem simultaneamente no mesmo enunciado.

O tratamento da polissemia no contexto da Semântica de Frames envolve a escolha entre manter significados distintos agrupados em uma só acepção (*lumping*) ou separar os significados em diferentes acepções (*splitting*). Embora a Semântica de Frames não direcione esta escolha, a abordagem que considera cada Unidade Lexical (UL) como o pareamento de um lexema com um *frame* possibilita o processo de *splitting*.

Este é um ponto fundamental para o trabalho descrito aqui, uma vez que a melhoria da interpretação por máquinas pressupõe a possibilidade de desambiguação dos lexemas polissêmicos. O fato de cada UL possuir seu próprio sentido, ou seja, evocar um *frame* específico já configura uma restrição para a rede semântica.

Em termos linguísticos, os elementos de um *frame* emergem na valência lexical e construcional das expressões que evocam este *frame* e podem ser considerados, segundo Salomão (2009), funções microtemáticas, postuladas em relação ao *frame* caracterizados por elas. Os elementos do *frame* são importantes no processo de desambiguação de sentenças como em (1) (adaptadas de Salomão (2009)), onde <substituir> evoca diferentes *frames*.

- (1) (a) Dunga _[técnico] substituiu o atacante Robinho jogador no último jogo.
(b) Neymar _[jogador] substituiu o atacante Robinho jogador no último jogo.

As ideias de Fillmore em relação a Semântica de Frames deram origem ao projeto FrameNet. O projeto FrameNet (FILLMORE; JOHNSON; PETRUCK, 2003), uma iniciativa do International Computer Science Institute (ICSI) de Berkeley, Califórnia (EUA), vem, desde 1997, implementando computacionalmente uma base de dados lexical fundada na Semântica de Frames. O objetivo da FrameNet é documentar as possibilidades combinatórias de cada palavra, em cada um dos seus sentidos, ou seja, sua valência sintática e semântica. Para isso, a FrameNet criou uma base de dados com centenas de *frames* (relacionados semanticamente entre si) e associou cada sentido de um lexema (uma Unidade Lexical) a um *frame*. Através da anotação de sentenças obtidas em corpora, são obtidas as valências de cada Unidade Lexical. Estas informações são apresentadas em vários relatórios web.

3.3. COMPOSICIONALIDADE

Um dos pressupostos apresentados para o trabalho é que a interpretação de expressões linguísticas é governada pelo Princípio da Composicionalidade (CRUSE, 2010). Este princípio estabelece que o significado de uma expressão semanticamente complexa é uma função composicional dos significados de seus constituintes semânticos.

Os defensores da composicionalidade direta argumentam que qualquer operação sintática tem uma operação semântica correlacionada e, conseqüentemente, qualquer expressão bem formada em termos sintáticos tem um significado predizível. Este é o chamado Princípio da Composicionalidade Forte.

No entanto, a Linguística Cognitiva mostra que a questão é muito mais complexa: uma língua pode apresentar diversos padrões de combinações de palavras, onde cada combinação (das mesmas palavras) pode apresentar um sentido diferente; o mesmo item lexical pode se referir a conceitos diferentes; a alternância dos verbos pode provocar uma mudança na leitura; existem significados que não são sintaticamente transparentes; existem expressões que são não-composicionais. Além disso, como dito antes, fatores contextuais, conceptuais, discursivos e pragmáticos, além de um “conhecimento do mundo” limitado, podem comprometer seriamente a compreensão do enunciado.

Não obstante estas dificuldades, adotamos como pressuposto que o processo de composicionalidade é fundamental na busca do significado da sentença. Para obter este significado, é essencial olhar além das palavras individualmente, reconhecendo que o significado de uma dada palavra tipicamente inclui conceitos que não são únicos para nenhum item linguístico (DODGE, 2010). Também é preciso reconhecer que o significado não está associado apenas com o lexema, mas também com os padrões construcionais onde estes lexemas ocorrem.

O uso da Semântica de Frames permite que nos afastemos da hipótese forte da composicionalidade, que não admite ambigüidade, sinonímia, homonímia e vagueza. O uso de Frames provê o acesso a um parcial conhecimento de mundo e o entendimento de que a soma do significado das partes pode não ser igual ao significado do todo. A Semântica de Frames possibilita inferências explicadas (ou motivadas) pelo *frame* evocado que, associadas à integração de conceitos através de processos cognitivos, usando as relações entre *frames*, permite uma abordagem mais abrangente para a construção do significado.

Como um recurso teórico complementar à Semântica de Frames, no que se refere especificamente a entidades, bem como para atender a pressuposto que os itens lexicais possuem uma estrutura interna que pode ajudar no processo da composicionalidade, estamos utilizando também a Teoria do Léxico Gerativo.

3.4. TEORIA DO LÉXICO GERATIVO E ONTOLOGIA SIMPLE

A Teoria do Léxico Gerativo (TLG), apresentada em Pustejovsky (1991) e desenvolvida em Pustejovsky (1995), considera o léxico como um sistema complexo, que forma uma parte crucial das línguas naturais e fornece as informações necessárias para o tratamento das propriedades polimórficas dos itens lexicais e sua permeabilidade ao contexto. Assim, as Unidades Lexicais são representadas por estruturas de informação, seguindo um conjunto finito de mecanismos gerativos que possibilitam a descrição do significado no contexto e a descrição da interface sintaxe-semântica (AMARO, 2009).

Um léxico gerativo pode ser caracterizado como um sistema envolvendo pelo menos quatro estruturas:

estrutura argumental (A); estrutura de eventos (E); estrutura qualia (Q) e estrutura de herança lexical (I). Um item lexical (IL) é então representado pela informação nestes níveis distintos por $IL = \langle A, E, Q, I \rangle$. Os mecanismos gerativos conectam estes quatro níveis, provendo a interpretação composicional das palavras em contexto (PUSTEJOVSKY, 1995).

A TLG considera que o significado das palavras está associado à estrutura qualia, na base de quatro fatores gerativos, chamados “papéis qualia”. Cada papel qualia captura como os seres humanos entendem os objetos e as relações no mundo e provê uma explicação mínima para o comportamento linguístico dos itens lexicais. A determinação dos valores de cada papel qualia não é obrigatória; para um dado item lexical, somente os valores dos papéis pertinentes devem ser definidos. (PUSTEJOVSKY, 1995) define os seguintes papéis qualia:

- 1 Formal: valores que diferenciam um dado objeto dentro do seu domínio semântico; tipicamente é a descrição de sua categoria básica.
- 2 Constitutivo: valores que expressam a relação entre um dado objeto e seus constituintes ou suas partes, tais como o material, o peso ou partes características.
- 3 Télico: valores relativos a informação sobre a função ou o propósito do objeto, tais como a intenção de um agente executando uma dada ação ou a função intrínseca do objeto.
- 4 Agentivo: valores que determinam a origem do objeto, tais como seu criador, tipo de origem (natural ou artificial) ou sua causa inicial.

Dado o viés computacional da TLG, os valores dos papéis qualia são especificados através de expressões lógicas, com tipos bem definidos e estruturas relacionais, indicando a associação apropriada com o termo predicador. O sistema de tipos, usado para categorizar os itens lexicais, é assim uma questão importante na TLG. A categorização é feita via herança ortogonal. Neste modelo de herança, ao invés de uma única estrutura organizacional para os tipos, diferentes estruturas são usadas ortogonalmente, baseadas na estrutura qualia. Um item lexical herda informação de acordo com a estrutura qualia que ele carrega. Por exemplo, <livro> é-um <objeto físico> (qualia formal), mas <livro> é-uma <informação> (qualia télico). Os tipos para um dado item lexical podem então ser unificados. Esta estratégia permite a criação de conceitos cada vez mais específicos. Tipos unificados podem ser vistos como estruturados por dimensões ortogonais (ou perspectivas).

Uma implementação computacional que usa (e estende) os conceitos da TLG é a ontologia SIMPLE (LENCI et al., 2000). SIMPLE é um projeto de larga escala, patrocinado pela União Europeia, com o objetivo de desenvolver um léxico semântico computacional, de ampla cobertura e múltiplos propósitos, para doze línguas da Europa. A ontologia SIMPLE foi projetada neste contexto e desenvolvida para garantir a uniformidade e consistência entre os dicionários monolíngues, provendo um modelo comum harmonizado, que codifica tipos semânticos estruturados e um conjunto de templates (análogos aos *frames* semânticos).

O vocabulário básico da SIMPLE se apoia em uma extensão da estrutura qualia, para organizar os tipos semântico/conceptuais como um mecanismo representacional que expresse a multidimensionalidade do significado. Isto fornece ao modelo um alto grau de generalidade, uma vez que os mesmos mecanismos são usados para conceitos de diferentes áreas semânticas (entidades, eventos, nomes abstratos, etc).

Neste trabalho foi utilizada a implementação da ontologia SIMPLE em OWL, chamada SIMPLE-OWL (TORAL; MONACHINI, 2007). Na SIMPLE-OWL os tipos semânticos foram implementados como classes. As relações são modeladas como propriedades de objeto e os traços (*features*) como propriedades de dados. Uma característica importante desta ontologia é que ela consiste tanto de Tipos Simples, que identificam somente o aspecto monodimensional do significado expresso por relações hiperonímicas, quanto de Tipos Unificados, que expressam os aspectos multidimensionais, combinando as relações de subtipo e as dimensões semânticas ortogonais.

3.5. MODELO CONEXIONISTA ESTRUTURADO

Os Modelos Conexionistas são caracterizados como modelos computacionais com habilidades de adaptar, aprender, generalizar, agrupar ou organizar dados, nos quais a operação é baseada em processamento paralelo. Neste modelo, uma rede é formada por um conjunto de unidades de processamento simples, chamadas de nós, células ou neurônios (donde a denominação de Redes Neurais Artificiais). Estas unidades comunicam entre si enviando sinais de uma para outra, através de conexões ponderadas, através da técnica Ativação Propagada (Spread Activation).

Porém, uma questão se levanta: como unidades tão simples podem representar algum conhecimento? A análise de Feldman e Ballard (1982) indica a possibilidade de construção de redes com um nível mais alto de abstração, em que as unidades podem possuir algum tipo de estrutura, necessária para a organização do conhecimento na rede. Este modelo é chamado de Modelo Conexionista Estruturado.

Shastri e Ajjanagadde (1989) apresentam a proposta de adotar o Modelo Conexionista Estruturado como uma abordagem para a implementação de redes semânticas que representem conhecimento e inferências. Este é o modelo usado no trabalho. A integração de recursos lexicais que estão disponíveis em diferentes formatos (os dados da FrameNet estão armazenados em um banco de dados enquanto as ontologias usadas estão descritas em OWL) é feita através de uma rede semântica cujos nós (representando conceitos) são estruturas de traço e a realização de inferências é feita através do uso da técnica de ativação propagada.

4. FRAMEWORK LUDI

O trabalho descrito deu origem a uma aplicação computacional, denominada Framework LUDI, baseado em um recurso lexical (a FrameNet), apoiado em uma visão ontológica (a ontologia SIMPLE e a ontologia LUDI, desenvolvida pelos autores) e implementado como um modelo conexionista estruturado. Esta seção apresenta alguns aspectos relacionados à construção da rede semântica utilizada no framework.

4.1. ENTIDADES, EVENTOS E QUALIDADES

A ontologia LUDI faz a distinção entre três categorias básicas: Entidades, Eventos e Qualidades. Foi adotada a abordagem da ontologia DOLCE (MASOLO et al., 2003), em que eventos são perdurantes (*perdurants*), ou seja, ocorrem ou acontecem no tempo, e entidades são endurantes (*endurants*), ou seja, existem independente do tempo. As qualidades, por sua vez, estão associadas às características das entidades ou eventos.

Ruppenhofer et al. (2010, p.5) reconhece que nomes referentes a tipos artefatuais ou naturais possuem uma estrutura de *frame* mínima, por si próprios. Artefatos geralmente ocorrem com outras expressões, indicando seu subtipo, o material de que é feito, a maneira de produção ou seu propósito e/ou uso, ou seja, sua estrutura qualia. Assim, optamos, no trabalho, pela substituição dos *frames* de Entidades atualmente registrados pela FrameNet, pelas classes da ontologia SIMPLE.

Com relação aos Eventos, foi mantida a modelagem da FrameNet, com a organização dos Frames (e das relações entre eles) segundo uma perspectiva ontológica, utilizando a própria ontologia SIMPLE. É importante ressaltar que as classes de eventos não visam substituir os Frames da FrameNet, mas sim classificá-los.

Para organização do nível estrutural da ontologia LUDI, referente a Qualidades, adotamos uma abordagem baseada nos chamados espaços conceituais (*conceptual spaces*), apresentados em Gärdenfors (2000). Embora fosse necessária uma reestruturação dos *frames* para se trabalhar mais completamente com a ideia de espaços conceituais, optamos por manter os frames atuais, realizando apenas uma simples adaptação das relações entre eles.

4.2 ELEMENTOS DE FRAME

Dentro da hipótese que aumentar a densidade da rede pode melhorar a compreensão do significado, foram feitas duas adições à estrutura da FrameNet:

- a) associação dos Elementos de Frame com uma classe da ontologia SIMPLE (que funciona então como um Tipo Ontológico, na perspectiva da FrameNet);
- b) associação dos Elementos de Frame com outros Frames da própria FrameNet.

A associação (b) necessita de uma justificativa, uma vez que ela não existe na estrutura da FrameNet atual e não foi ainda proposta em trabalhos anteriores, configurando uma extensão de caráter conceitual. A ideia é registrar a referência que o elemento de uma *frame* pode fazer a um outro *frame*. Assim, além de atuarem em termos linguísticos como papéis microtemáticos, específicos para a situação descrita pelo *frame*, os elementos podem desempenhar uma função cognitiva, como evocadores de outros *frames*, a partir do *frame* a que pertencem.

4.3. Integração de Ontologias

A ontologia LUDI faz a integração de duas outras ontologias: uma Ontologia de Frames e a Ontologia SIMPLE. Cada uma atende a um determinado propósito. A Figura 1 ilustra esta integração.

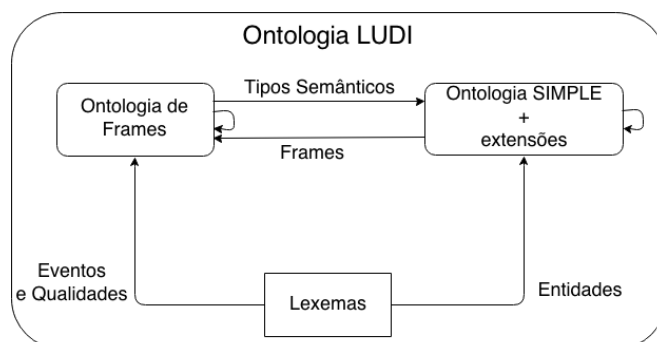


Figura 1 – LUDI - Integração de ontologias.

A Ontologia de Frames não é, na verdade, uma ontologia descrita formalmente, já que há dificuldades em formalizar a estrutura da FrameNet. Esta ontologia consiste na representação de alguns elementos da FrameNet como classes ou propriedades. Ela não contém a totalidade das informações disponibilizadas pela FrameNet, mas apenas aquelas utilizadas nas operações do framework LUDI: os Frames, os Elementos de Frame, as Unidades Lexicais, as Relações entre Frames e as Relações entre Elementos de Frame.

A Ontologia SIMPLE (com algumas extensões) é usada para registrar os tipos semânticos estruturados. Estes tipos possuem relações entre si e são usados basicamente para categorizar os *frames* e os elementos de *frames*. Os lexemas referentes a entidades (tipos naturais e artefatos) são associados a tipos da ontologia SIMPLE. Os tipos da ontologia podem ser associados também a *frames*, quando as relações indicam algum tipo de evento ou qualidade. Mecanismos de inferência são usados para a classificação dos lexemas em relação à estrutura qualia.

4.4. A REDE LUDI

4.4.1. ESTRUTURA DE TRAÇOS

Para a implementação da rede LUDI é realizada uma operação intermediária, que consiste em representar os dados da FrameNet e da SIMPLE formatados como MVAs (Matriz de Valores de Atributos), representadas via uma Estrutura de Traços Tipados (*Typed Features Structure*). A redução a um formato comum facilita a integração das duas fontes de dados.

Em relação à FrameNet, cada Estrutura de Traços representa um *frame*. O tipo da estrutura é o nome do *frame*. Cada traço representa um elemento do *frame*. O valor do traço é uma lista de apontadores (i) para elementos de outros *frames* ou (ii) para as classes da Ontologia SIMPLE, que funcionam como Tipos Ontológicos ou (iii) para outros *frames*. Os lexemas, os lemas e as Unidades Lexicais também são representados por Estruturas de Traços, o que possibilita considerar as Unidades Lexicais como tipos

estruturados com traços próprios (que podem ser unificados com os traços do *frame*, por exemplo).

Em relação a ontologia SIMPLE, cada Estrutura de Traços representa uma classe (um tipo semântico da SIMPLE). O tipo da estrutura é o nome da classe. Cada traço representa uma relação qualia associada com o tipo semântico. O valor do traço é uma lista de apontadores para outros tipos semânticos ou para *frames*.

4.4.2. VALÊNCIAS E FUNÇÕES GRAMATICAIIS

Nesta implementação usamos o termo “valência” no mesmo sentido apresentado por Fillmore (2008): as valências registram as diversas possibilidades combinatórias da interface sintaxe- semântica. As valências indicam as várias configurações dos constituintes associados a uma determinada Unidade Lexical. É, portanto, uma estrutura complexa: ela está associada simultaneamente a uma Unidade Lexical e a vários elementos de *frame*, juntamente com as funções gramaticais atribuídas a cada um destes elementos. Cada valência é representada como uma unidade da rede. Assim, cada *frame* pode ser conectado a diversas valências e cada valência pode ser conectada a diversos elementos de *Frame*. O processo de análise consiste em determinar a valência mais provável para os dados fornecidos e, conseqüentemente, o *frame* cuja evocação é mais provável.

No trabalho foram consideradas as funções gramaticais descritas em Torrent e Ellsworth (2013): Argumento Externo (Ext), Objeto Direto (DObj), Objeto Indireto (IObj) e Dependente (Dep). As funções gramaticais são implementadas como traços dos links entre as valências e os elementos de *frame* e não como unidades da rede. Durante o processo de análise, a ativação ou não de um nó correspondente a uma valência vai depender da compatibilidade entre a função gramatical atribuída aos lexemas e a função gramatical atribuída ao link.

4.4.3. UNIDADES E CONEXÕES DA REDE

A rede LUDI é ativada quando uma análise deve ser realizada. As estruturas de traço constituem as Unidades da rede. A rede é representada como um grafo dirigido, cujos nós recebem uma letra indicando o tipo de Unidade representada. A Tabela 1 apresenta a descrição de cada Unidade.

A Tabela 2 apresenta as conexões (*links*) entre as diversas unidades. O peso atribuído a cada link é sempre 1.0, exceto para os *links* em que o destino é um tipo semântico da ontologia SIMPLE. Neste caso, o peso do *link* é calculado em função da posição do tipo na hierarquia da ontologia.

Símbolo	Unidade	Descrição
W	Word	Palavra: uma das formas que uma palavra pode assumir
X	Lexema	A forma não flexionada da palavra
L	Lema	Um lema é composto de um ou mais lexemas
U	Unidade Lexical Unidade Semântica	Pareamento de um lema com um significado. Na FrameNet o pareamento de um lema com um <i>frame</i> é chamado Unidade Lexical. Na SIMPLE, o pareamento é chamado Unidade Semântica (USem).
F	Frame	Estrutura usada para representar situações, eventos e qualidades
E	Elemento de Frame	Usado para representar os participantes de um <i>frame</i> .
V	Valência	Representa a estrutura argumental de uma Unidade Lexical
T	Tipo Ontológico	Tipo semântico estruturado ontologicamente e baseado nas relações qualia

Tabela 1 - Unidades da Rede LUDI

Conexão	Descrição
XL	Conexão entre o Lexema (fornecido como entrada para a análise) e o Lema correspondente.
LF	Conexão entre o Lema e o Frame, representando o pareamento forma-significado.
FV	Conexão entre o Frame e as Valências consideradas para uma dada análise.
VE	Conexão entre uma Valência e os Elementos de Frame correspondentes.
EE	Conexão entre Elementos de Frame de Frames distintos.
ET	Conexão entre um Elemento de Frame e o Tipo Ontológico associado.
TE	Conexão entre um Tipo Ontológico e um Elemento de Frame (um elemento de um Frame esquemático, representando a conceptualização para um dado Tipo Ontológico).
LU	Conexão entre um Lema e uma Unidade Semântica (cada sentido do Lema, no caso de polissemia).
UT	Conexão entre uma Unidade Semântica e o Tipo Ontológico associado.
UF	Conexão entre uma Unidade Semântica e o Frame evocado por ela.
FT	Conexão entre um Frame e o Tipo Ontológico atribuído ao Frame.
UE	Conexão entre uma Unidade Semântica e um Elemento de Frame (quando a Unidade Semântica puder ser conceptualizada como participante em um Frame).
TT	Conexão entre dois Tipos Ontológicos.

Tabela 2 - Links da Rede LUDI

4.4.4. CONSTRUÇÃO E ATIVAÇÃO DA REDE

O processo de análise é iniciado com a construção da rede completa, chamada **Rede LUDI**, que consiste de todos os nós e relações disponíveis na ontologia SIMPLE, mais os *frames* e os lexemas sob análise. A seguir, a partir da Rede LUDI, é construída uma rede intermediária, chamada **Rede Foco**, em que constam apenas os nós e relações associados aos lexemas analisados. A Ativação Propagada é realizada nesta rede. A rede resultante, composta pelos nós ativados, é chamada **Rede Resultado**.

5. RESULTADOS EXPERIMENTAIS

5.1. O USO DE CORPORA

Uma das hipóteses do trabalho é o uso de corpora, por dois grandes motivos. Em primeiro lugar, ainda que a intuição do analista como usuário falante da língua seja rica, o uso feito pelos outros milhões de falantes é muito mais rico. Em segundo lugar, uma análise superficial de algumas centenas de sentenças é suficiente para concluir que qualquer categorização realizada *a priori* é bastante limitada. Isto não é surpresa, se adotamos como representação das conceptualizações o uso de esquemas

imagéticos (ao invés de uma solução meramente simbólica). A riqueza conceptual de cada esquema, acrescida das possibilidades de combinações e transformações, torna virtualmente impossível uma categorização completa. Assim, o uso de corpora para avaliar a categorização dos *frames* visa também testar a abrangência desta categorização, sem a ilusão de que uma cobertura de 100% possa ser alcançada.

5.2. O ESQUEMA BÁSICO DE <ABRIR.V>

Usando uma abordagem cognitiva, os muitos sentidos de <abrir.v> podem ser projetados a partir da junção do esquema imagético CONTAINER e da noção de dinâmica de forças (*force-dynamics*) (TALMY, 2000). Os sentidos são derivados através das operações de perfilamento, perspectivação e transformações no esquema imagético, bem como sua projeção em domínios mais abstratos, via metáforas. Assim, o *frame* esquemático para esta análise de <abrir.v> (Figura 2) é um CONTAINER sobre cujo portal podem ser aplicadas duas forças (esquema FORCE), através de um Agonista e de um Antagonista.

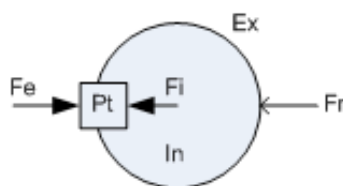


Figura 2 – Esquema para <abrir.v>.

Este é um esquema conceptual básico, servindo para a projeção dos vários sentidos de <abrir.v> apresentados no trabalho. Os elementos do esquema são detalhados na Tabela 3.

Sigla	Nome	Descrição
Ex	Exterior	representa a região externa ao container, incluindo os elementos que podem estar localizados nesta região.
In	Interior	representa a região interna ao container, incluindo os elementos que podem estar localizados nesta região.
Pt	Portal	representa um elemento que funciona como ligação entre as regiões exterior e interior, possibilitando o trânsito de elementos entre essas regiões (de "fora" para "dentro", ou vice-versa). O portal perfila o conceito de "possibilidade de acesso", e pode ser realizado, por exemplo, por uma abertura, uma passagem ou uma barreira.
Fr	Fronteira	representa o elemento que separa as regiões interior e exterior do container, servindo para delimitar o que é considerado interior ao container ou não.
Fe	Força externa	representa a ação de uma força atuando no portal "de fora para dentro" do container.
Fi	Força interna	representa a ação de uma força atuando no portal "de dentro para fora" do container.

Tabela 3 – Elementos do esquema para <abrir.v>

5.3. OS FRAMES DE <ABRIR.V>

A Figura 3 apresenta o *frame* não-lexical ABRIR_ESQUEMA, usado para representar o esquema básico. Os elementos Agente, Causa, Instrumento são definidos mais especificamente como:

- a) Agente: representa uma entidade consciente - dotada de volição e que age com intenção - responsável pelo evento causador de <abrir.v>.

- b) Causa: representa um evento causador de <abrir.v>, que não esteja em relação com uma entidade consciente.
- c) Instrumento: representa o elemento em destaque no evento causador de <abrir.v>, usado pelo Agente ou relacionado a Causa.

I - ABRIR_ESQUEMA		
Elemento de Frame	Tipo	Descrição
Agente	CoreSet	Representa a força associada à ação de <abrir.v>.
Causa		
Instrumento		
Container	Coreset	Representa a entidade que sofre a ação da força e que muda seu estado.
Portal		
Conteúdo	Periférico	Representa os elementos na região interior do container.
Interior	Periférico	Representa a região interior do container.
Exterior	Periférico	Representa a região exterior ao container.
Fronteira	Periférico	Representa a separação entre as regiões interior e exterior do container.

Figura 3 - ABRIR_ESQUEMA

Em relação ao conjunto (Container, Portal) é importante ressaltar a relação parte-todo existente. O Portal é “parte” do Container. Como o padrão metonímico “Parte-pelo-todo” é bastante comum, as sentenças podem lexicalizar tanto o Portal quanto o Container.

A partir do esquema básico, uma rede de *clusters* de sentidos é apresentada na Figura 4. O *cluster* Movimento_corporal é definido para o sentido concreto e corporificado de <abrir.v>. O *cluster* Abrir_físico também possui um sentido concreto. A partir destes *clusters*, os demais são obtidos através de interpretações que consideram o uso de metáforas.

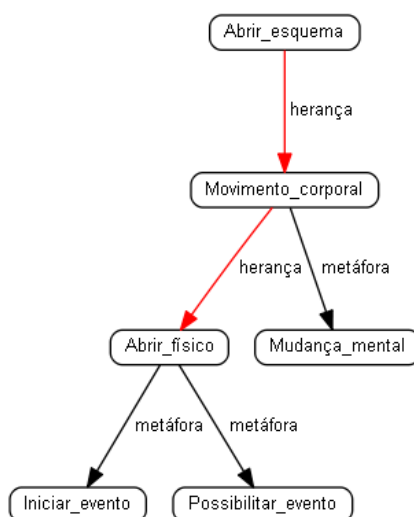


Figura 4 – Cluster de frames herdeiros de ABRIR_ESQUEMA.

A rede de *clusters* é expandida na Figura 5. Basicamente é considerada a alternância causativo-incoativo e são adicionados alguns *frames* já definidos pela FrameNet. A fim de facilitar as referências, a Figura 6 apresenta a numeração dos *clusters* e dos *frames*.

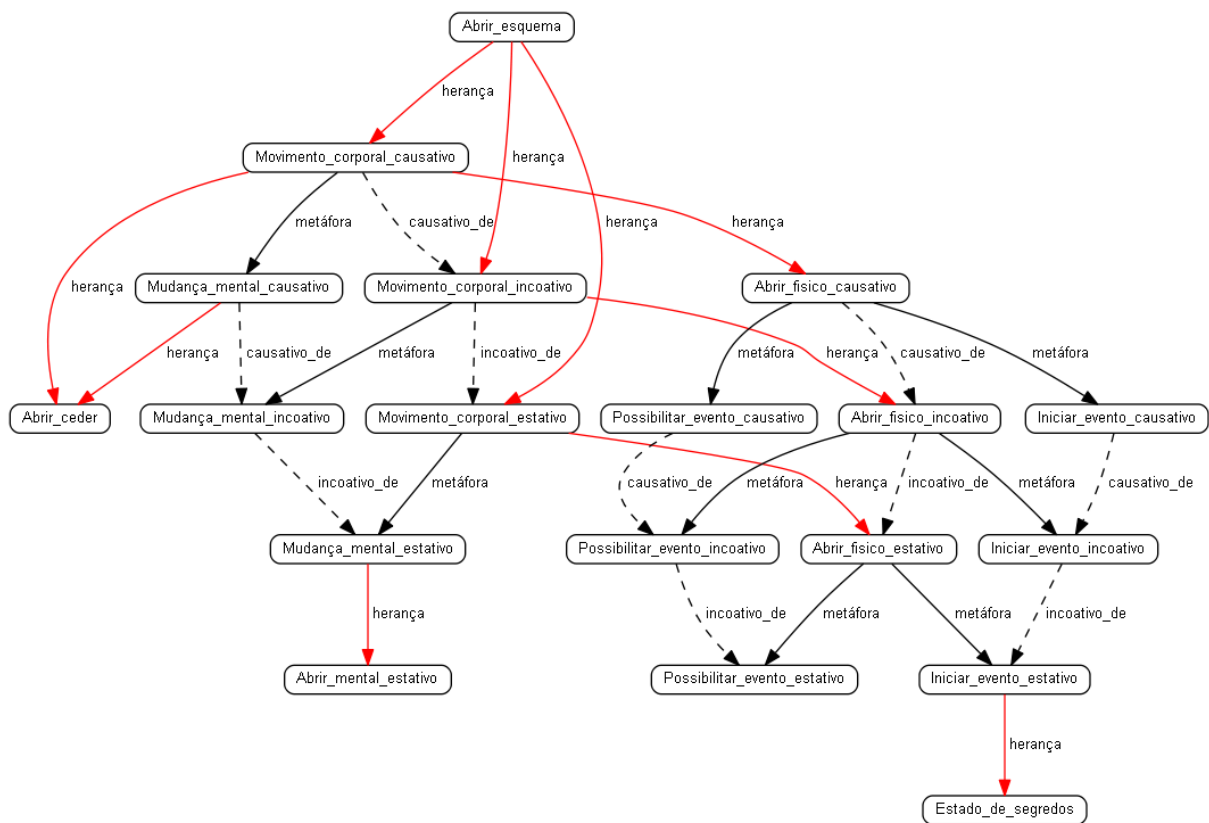


Figura 5 - Rede de frames herdeiros de ABRIR_ESQUEMA

Domínio	Cluster	#Frame	Nome do Frame
Esquema básico	I – Esquema básico	1	Abrir_esquema
Corpo físico	II – Movimento_corporal	1	Movimento_corporal_causativo
		2	Movimento_corporal_incoativo
		3	Movimento_corporal_estativo
Entidades físicas	III – Abrir_físico	1	Abrir_físico_causativo
		2	Abrir_físico_incoativo
		3	Abrir_físico_estativo
Aspecto	IV – Iniciar_evento	1	Iniciar_evento_causativo
		2	Iniciar_evento_incoativo
		3	Iniciar_evento_estativo
		4	Estado_de_segredos
Modalidade	V – Possibilitar_evento	1	Possibilitar_evento_causativo
		2	Possibilitar_evento_incoativo
		3	Possibilitar_evento_estativo
Estado Mental, Emocional ou Psicológico	VI – Mudança_mental	1	Mudança_mental_causativo
		2	Mudança_mental_incoativo
		3	Mudança_mental_estativo
		4	Abrir_mental_estativo
		5	Abrir_ceder

Figura 6 - Frames por clusters

5.4. VALÊNCIAS

Usando a terminologia de Fillmore(2008), a “descrição de valência” de uma dada Unidade Lexical mostra o conjunto de associações dos elementos do *frame* com as Funções Gramaticais e com os

Tipos Sintagmáticos. A Figura 7 mostra algumas valências utilizadas no trabalho. Esta descrição é reconhecidamente incompleta, pois é derivada da análise de algumas poucas dezenas de sentenças. As valências que se diferenciam apenas pela variação dos elementos de *frame* foram agrupadas. Por exemplo, o conjunto (Agente, Causa, Instrumento) com a função “Argumento Externo” é representado como [Agente|Causa|Instrumento].ext. No entanto, cada valência é implementada individualmente.

Frame II.1 – Movimento_corporal_causativo	
<abrir.v>	V1: [Agente Causa Instrumento].ext, Parte_corporal.dobj, Experienciador.dep
<abertura.n>	V1: [Agente Causa Instrumento].ext, Parte_corporal.dep, Experienciador.dep
Frame II.2 – Movimento_corporal_incoativo	
<abrir.v>	V1: Parte_corporal.ext, Experienciador.dep
Frame II.3 – Movimento_corporal_estativo	
<aberto.a>	V1: Parte_corporal.ext, Experienciador.dep
Frame III.1 – Abrir_físico_causativo	
<abrir.v>	V1: [Agente Causa Instrumento].ext, [Container Portal].dobj, [Conteudo Interior Exterior Frenteira].dep V2: [Agente Causa Instrumento].ext, Portal.dobj, Container.dep, [Conteudo Interior Exterior Frenteira].dep
<abertura.n>	V1: [Agente Causa Instrumento].ext, [Container Portal].dep, [Conteudo Interior Exterior Frenteira].dep
Frame III.2 – Abrir_físico_incoativo	
<abrir.v>	V1: [Container Portal].ext, [Conteudo Interior Exterior Frenteira].dep V2: Container.ext, Portal.dep, [Conteudo Interior Exterior Frenteira].dep V3: Container.dep, Portal.ext, [Conteudo Interior Exterior Frenteira].dep
Frame III.3 – Abrir_físico_estativo	
<aberto.a>	V1: [Container Portal].ext, [Conteudo Interior Exterior Frenteira].dep V2: Container.ext, Portal.dep, [Conteudo Interior Exterior Frenteira].dep V3: Container.dep, Portal.ext, [Conteudo Interior Exterior Frenteira].dep
Frame IV.1 – Iniciar_evento_causativo	
<abrir.v>	V1: [Agente Causa Instrumento Evento_inicial].ext, [Evento Entidade].dobj
<abertura.n>	V1: [Agente Causa Instrumento Evento_inicial].ext, [Evento Entidade].dobj
Frame IV.2 – Iniciar_evento_incoativo	
<abrir.v>	V1: [Evento Entidade].ext, Tempo.dep
Frame IV.3 – Iniciar_evento_estativo	
<aberto.a>	V1: [Evento Entidade].ext
Frame IV.4 – Estado_de_segredos	
<aberto.a>	V1: Fenômeno.ext, Conhecedor.dep
Frame V.1 – Possibilitar_evento_causativo	
<abrir.v>	V1: [Agente Causa Instrumento Evento_inicial].ext, Meio.dobj, Evento.iobj
Frame V.2 – Possibilitar_evento_incoativo	
<abrir.v>	V1: Meio.ext, Evento.iobj
Frame V.3 – Possibilitar_evento_estativo	
<aberto.a>	V1: Meio.ext, Evento.iobj
Frame VI.1 – Mudança_mental_causativo	
<abrir.v>	V1: [Agente Causa Instrumento].ext, Parte_corporal.dobj, Evento.iobj, Experienciador.dep
Frame VI.2 – Mudança_mental_incoativo	
<abrir.v>	V1: Parte_corporal.ext, Evento.iobj, Experienciador.dep
Frame VI.3 – Mudança_mental_estativo	
<aberto.a>	V1: Parte_corporal.ext, Evento.iobj, Experienciador.dep
Frame VI.4 – Abrir_ceder	
<abrir_mão.v>	V1: Capitulador.ext, [Força Questão].iobj
Frame VI.5 – Abrir_mental_estativo	
<aberto.a>	V1: Experienciador.ext, Topico.dep

Figura 7 – Valências

5.5. EXEMPLOS DE APLICAÇÃO

Esta seção apresenta alguns exemplos de resultados da aplicação do *framework* LUDI para desambiguação de sentenças com o verbo <abrir.v> (o trabalho completo, em Matos (2014), apresenta

os resultados para as expressões abrir_boca, abrir_olho, abrir_mão, abrir_carteira, abrir_porta, abrir_jornal, abrir_museu, abrir_exposição, abrir_shopping, abrir_caminho, abrir_cerveja, abrir_marcador). Para cada experimento são apresentados: a motivação da escolha da expressão usada; a sentença e os segmentos usados como entrada para os algoritmos (juntamente com as funções gramaticais); a rede resultante, com a inferência final realizada e um comentário sobre o resultado obtido.

5.5.1. ABRIR_PORTA

Motivação: Testar a inferência com entidades físicas conceptualizadas como containers.

Sentença: (...) *quem me abriu a porta do apartamento foi uma empregada* (...)

Segmentos: abrir.target, porta.dobj, empregada.ext

Comentário: A Figura 8 mostra que, apesar do processo de inferência ser direto, são ativadas mais de uma valência. Isto acontece porque <porta> é sempre conceptualizada como uma abertura para um container (uma construção 'building' neste caso).

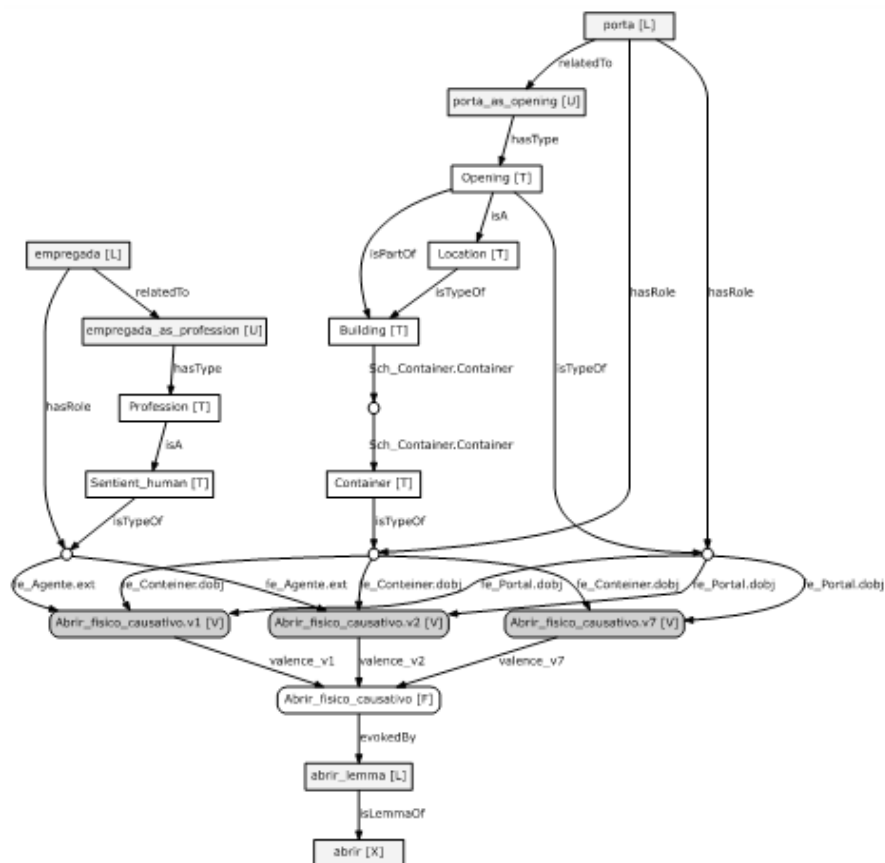


Figura 8 - Abrir_porta

5.5.2. ABRIR_OLHO

Motivação: Testar o cluster VI, onde partes corporais são usadas metaforicamente relacionadas a estados mentais.

Sentença: (...) *Esse curso abriu meus olhos para o lado da informática (...)*

Segmentos: abrir.target, olho.dobj, curso.ext, informatica.iobj

Comentário: Dois fatos se destacam na inferência apresentada na Figura 9:

- A associação de <curso> (uma Unidade Semântica) com o elemento de *frame* EDUCATION_TEACHING.COURSE, mostrando a integração entre a FrameNet e a ontologia SIMPLE.
- Adistinção entre os *frames* VI.1.MUDANÇA_MENTAL_CAUSATIVO e II.1.MOVIMENTO_CORPORAL_CAUSATIVO só foi possível porque, na configuração da rede, as valências cadastradas para II.1.MOVIMENTO_CORPORAL_CAUSATIVO obrigam que o elemento de *frame* EXPERENCIADOR seja expresso, exceto quando o Externo é um AGENTE (que neste caso deve ser o próprio EXPERENCIADOR).

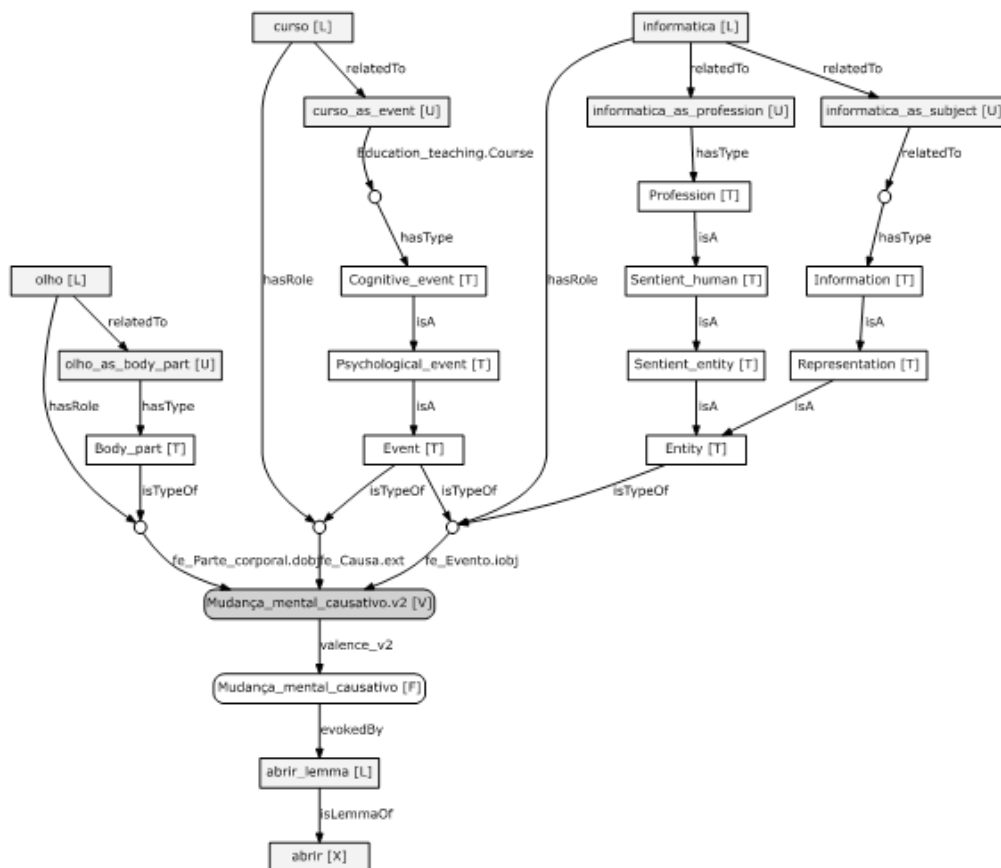


Figura 9 - Abrir_olho

5.5.3. ABRIR_JORNAL

Motivação: Testar o processo de inferência com entidades polissêmicas. No caso, <jornal> tem tanto o sentido de “informação” quanto de um “artefato”.

Sentença: (...) *Evidentemente está pior, basta a gente abrir o jornal e vê o que?* (...)

Segmentos: abrir.target, jornal.dobj, gente.ext

Comentários: Neste caso o processo de inferência (Figura 10) foi praticamente direto, por dois motivos:

- Existe uma relação qualia entre INFORMATION e SEMIOTIC_ARTIFACT, estabelecendo que informações são geralmente contidas em artefatos semióticos. Esta relação permite ao processo de inferência reforçar a ativação de <jornal> como um artefato.
- Uma vez que geralmente contém informações, SEMIOTIC_ARTIFACT é conceptualizado como um contêiner associado com o elemento de *frame* ABRIR_ESQUEMA.CONTEINER. Este é mais um exemplo da integração entre a ontologia e o recurso lexical.

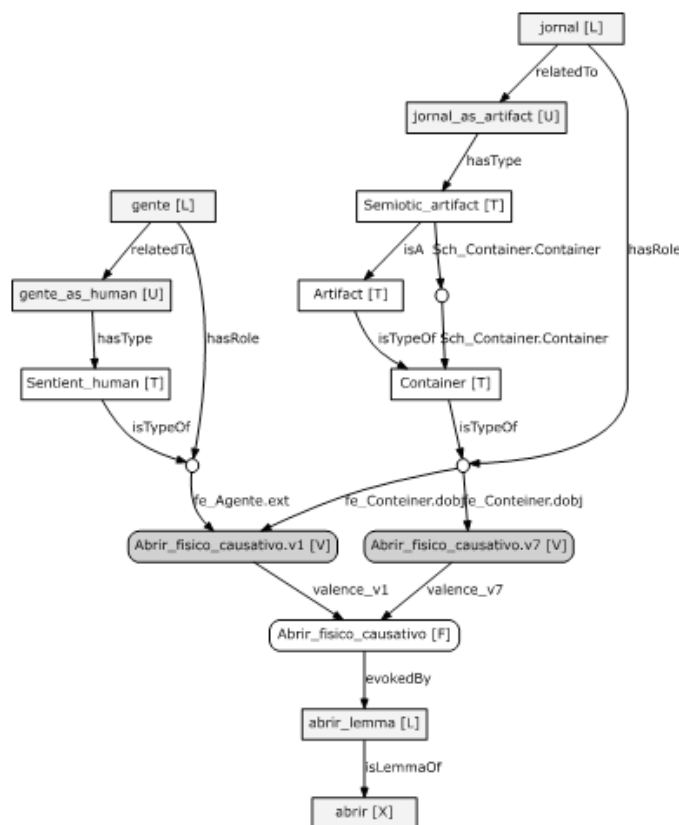


Figura 10 – Abrir_jornal

5.5.4. ABRIR_CERVEJA

Motivação: Testar os algoritmos na resolução de uso metonímicos. Neste caso, o container é conceptualizado pelo seu conteúdo.

Sentença: *Estas são as melhores maneiras de abrir uma cerveja.*

Segmentos: abrir.target, cerveja.dobj

Comentários: A Figura 11 mostra que uma interpretação possível para a sentença é a consideração que <cerveja> é subentendida por um container (a garrafa de cerveja, por exemplo). Foi criada uma valência adicional (V8) e um elemento de *frame* CONTEUDO_AS_CONTEINER, específicos para esta situação, a fim de deixar mais claro o processo metonímico que está ocorrendo.

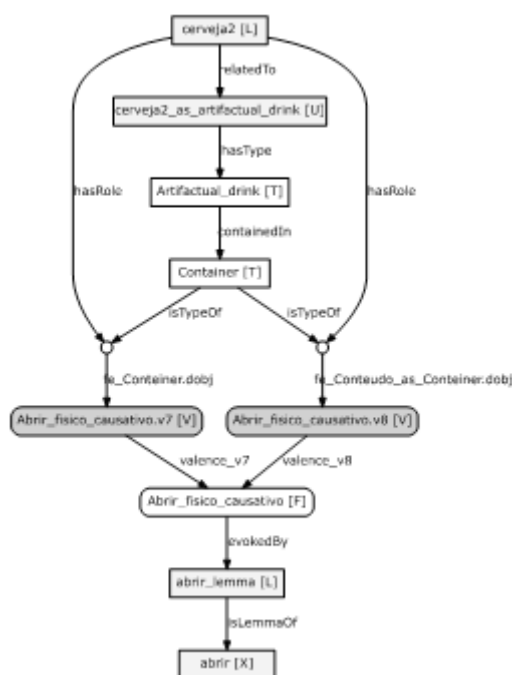


Figura 11 – Abrir_cerveja

6. CONCLUSÃO

O objetivo do trabalho apresentado foi implementar um *framework* para a desambiguação de itens lexicais. O processo de desambiguação pressupõe que dois (ou mais) itens lexicais, que possuem a mesma forma, possuem significados diferentes. Este trabalho lidou, então, com a difícil questão do significado das palavras e da construção/compreensão do significado de sentenças.

O trabalho adotou uma abordagem híbrida, reunindo contribuições da Semântica de Frames (para representação do significado considerando aspectos cognitivos) e da Teoria do Léxico Gerativo (para representação do significado considerando aspectos mais formais). O resultado foi uma aplicação

computacional, chamada Framework LUDI (*Lexical Unit Discovery through Inference*), que pode ser enquadrada na área de Compreensão da Língua Natural.

As contribuições e impactos do trabalho realizado podem ser observados tanto na área da Linguística Cognitiva, quanto na área da Linguística Computacional. Para a Linguística Cognitiva, uma contribuição importante é a demonstração de como pode ser realizado o alinhamento dos papéis temáticos com as relações qualia. A formalização deste alinhamento permite que interpretações de sentenças possam ser geradas dinamicamente, simulando (de maneira ainda bastante limitada, naturalmente) o processo cognitivo humano. Uma vez que a associação entre os elementos de um *frame* e as Entidades que participam do *frame* não está restrita à atribuição de um Tipo Ontológico para o elemento, a construção de uma especificação semântica para o enunciado pode ser feita com base em um número maior de opções.

Para a Linguística Computacional, a contribuição é uma técnica de análise de sentenças baseada em um modelo conexionista estruturado, usando a técnica de Ativação Propagada. Os modelos encontrados na literatura, relativos à desambiguação de *frames*, utilizam predominantemente técnicas baseadas em probabilidade, seja através de Máxima Entropia, seja através de Redes Bayesianas. No framework LUDI usamos um modelo conexionista estruturado, neurologicamente plausível, e um processo de inferência baseado em evidência, não em probabilidade. Consideramos que este tipo de inferência é mais adequado cognitivamente, pelo menos para uma certa classe de situações, em que experiências anteriores possibilitam o estabelecimento de evidências.

A Tabela 4 sintetiza as abordagens escolhidas neste trabalho, em comparação com as escolhas mais comuns em PLN.

PLN	LUDI
Linguística Gerativa	Linguística Cognitiva
Semântica Formal	Semântica de Frames
Forma Lógica	Especificação Semântica
Máquinas de Estados, Gramáticas Livres de Contexto	Modelo conexionista estruturado
Modelos probabilísticos	Modelo Fuzzy

Tabela 4 – Abordagens usadas no framework LUDI

Estas contribuições possuem valor por si mesmas. No entanto, julgamos que a maior contribuição é justamente a demonstração de que conceitos e ideias caros à Linguística Cognitiva podem ser usados em uma implementação computacional. A literatura de viés cognitivista, relacionada aos processos de categorização, conceptualização através de espaços mentais e abordagens gramaticais usando construções, por exemplo, apresenta, via de regra, um sentimento de resistência (quando não de rejeição) a qualquer tipo de formalização. Acreditamos que pesquisas como as realizadas pelo grupo NTL (*Neural Theory of Language*) ou esta, que levou à construção do framework LUDI, mostram que há um caminho viável para uma Linguística Cognitiva Computacional (FELDMAN, 2004).

Este trabalho fornece uma ideia inicial que pode ser expandida em diversas direções. Como exemplo do que pode ser realizado a partir dos resultados alcançados até agora, citamos:

- a) Atualmente as funções gramaticais estão sendo fornecidas como dados de entrada; para uma automatização maior da análise, seria necessário o uso de um *parser* que atribuisse estas funções;
- b) Estamos trabalhando apenas com palavras que evocam Frames (*Frame Evoking Elements*) e Unidades Semânticas associadas a tipos da ontologia SIMPLE; para uma análise completa da sentença, é necessário o cadastro dos *frames* e/ou Esquemas associados às preposições, artigos, quantificadores, etc;
- c) Não foi feita nenhuma abordagem do ponto de vista construcional, mas apenas baseada em valências; o modelo, como está hoje, pode ser facilmente estendido para alguns caso simples de construções;
- d) A ontologia SIMPLE foi utilizada por ser uma ontologia já existente e baseada na TLG; no decorrer do trabalho ficou claro que, para uma aplicação mais genérica, ela precisa ser expandida e possivelmente associada ou alinhada com outras ontologias de topo, como a DOLCE (MASOLO et al., 2003);
- e) Apenas os itens lexicais necessários aos experimentos foram adicionados à ontologia; para uma aplicação mais genérica, seria necessário um léxico mais completo e abrangente.

À guisa de conclusão, acreditamos que a implementação do *framework* LUDI mostra que a Semântica de Frames, enriquecida com a contribuição da Teoria do Léxico Gerativo e o uso de Tipos Ontológicos, possibilita não só a desambiguação lexical, mas se constitui em uma promissora base de pesquisa na área da Compreensão da Língua Natural.

REFERÊNCIAS

AMARO, R. *Computation of verbal predicates in portuguese: relational network, lexical-conceptual structure and context*. Tese (Doutorado), Universidade de Lisboa, 2009.

BOS, J. *A Survey of Computational Semantics: Representation, Inference and Knowledge in Wide-Coverage Text Understanding*. *Language and Linguistics Compass*, v. 5, n. 6, p. 336–366, jun. 2011. ISSN 1749818X.

BRUGMAN, C.; LAKOFF, G. *Cognitive topology and lexical networks*. In S. Small, G. Cottrill, & M. Tannenhaus (Eds.), *Lexical ambiguity resolution* (pp. 477-508), 1988. Palo Alto: Morgan Kaufman.

CRUSE, A. *Meaning in Language: An Introduction to Semantics and Pragmatics*. 3a. ed. Oxford: Oxford University Press, 2010.

DODGE, E. *Constructional and Conceptual Composition*. Tese (Doutorado), University of California at Berkeley, 2010.

FELDMAN, J. A. *Computational cognitive linguistics* in *Proc. 20th Intl. Conf. on Computational Linguistics (Coling 2004)*, Morristown, NJ: Association for Computational Linguistics, 2004, pp. 1114-1118.

FELDMAN, J. A.; BALLARD, D. *Connectionist models and their properties*. *Cognitive science*, v. 254, n. 1982, p. 205–254, 1982.

FILLMORE, C. J. *The need for a frame semantics in linguistics*. In: KARLGREN, H. (Ed.). *Statistical Methods in Linguistics*. Scriptor, 1977.

FILLMORE, C. J. *Frame semantics*. *Linguistics in the morning calm*, Hanshin Publishing Co., Seoul, South Korea, p. 111–137, 1982.

FILLMORE, C. J. *Frames and the semantics of understanding*. *Quaderni di Semantica*, v. 6, n. 2, p. 222–254, 1985.

FILLMORE, C. J. *Frame Semantics*. In: *Cognitive Linguistics: Basic Readings*. Berlin: Walter de Gruyter, 2006. cap. 10, p. 373–400.

FILLMORE, C. J. *Valence issues in Framenet*. In: GUTZ, T.; VOTTELER, K. (Ed.). *Trends in Linguistics, Studies and Monographs :Valency : Theoretical, Descriptive and Cognitive Issues*. Berlin: Mouton de Gruyter, 2008. p. 141–174.

FILLMORE, C. J.; BAKER, C. *A frames approach to semantic analysis*. *The Oxford Handbook of Linguistic Analysis*, n. 1, p. 313–340, 2010.

FILLMORE, C.; JOHNSON, C. R.; PETRUCK, M. R. *Background to Framenet*. *International Journal of Lexicography*, v. 16, n. 3, p. 235–250, 2003.

GÄRDENFORS, P. *Conceptual spaces - the geometry of thought*. MIT Press, 2000. I–X, 1–307 p. ISBN 978-0-262-07199-4.

GAWRON, J. M. *Frame Semantics*. In: MAIENBORN, C.; HEUSINGER, K. von; PORTNER, P. (Ed.). *Semantics: An International Handbook of Natural Language Meaning*. Berlin: Mouton de Gruyter, 2008, (HSK Handbooks of Linguistics and Communication Science Series, v. 23). p. 1–35.

- GEERAERTS, D. *Theories of Lexical Semantics*. Oxford: Oxford University Press, 2009. 368 p. ISBN 9780198700302.
- GRIES, S. *Corpus-based methods and cognitive semantics: The many senses of to run*. In *Linguistics Studies and Monographs*, 2006.
- LAKOFF, G. *Women, Fire, and Dangerous Things*. Chicago: University of Chicago Press, 1987. ISBN 9780226468044.
- LANGACKER, R. W. *Foundations of Cognitive Grammar, Vol. 1*. Stanford: Stanford University Press, 1987.
- LENCI, A. et al. *SIMPLE Work Package 2 - Linguistic Specifications - Deliverable D2.1*, 2000.
- LYONS, J. *Introdução à Linguística teórica*. São Paulo: Companhia Editorial Nacional, 1979.
- MASOLO, C. et al. *Wonderweb deliverabled18, ontology library (final)*.2003.
- MATOS, E. E. *LUDI: Um framework para desambiguação lexical com base no enriquecimento da Semântica de Frames*. 200 p. Tese (Doutorado), Universidade Federal de Juiz de Fora, 2014.
- PETRUCK, M. R. L. *Frame Semantics*. In: VERSCHUEREN, J. et al. (Ed.). *Handbook of Pragmatics*. John Benjamins, 1996.
- PUSTEJOVSKY, J. *The Generative Lexicon*. *Computational Linguistics*, The MIT Press, Cambridge (USA), v. 17, n. 4, p. 409–441, 1991.
- PUSTEJOVSKY, J. *The Generative Lexicon*. Cambridge, MA: MIT Press, 1995.
- RUPPENHOFER, J. et al. *FrameNet II: Extended Theory and Practice*. Berkeley, California: International Computer Science Institute, 2010.
- SALOMÃO, M. M. *FrameNet Brasil: um trabalho em progresso*. *Calidoscópico*, v. 7, n. 3, p. 171–182, dez. 2009. ISSN 16798740.
- SHASTRI, L.; AJJANAGADDE, V. *A connectionist system for rule based reasoning with multi-place predicates and variables*. 1989.
- SOWA, J. *Semantic Networks*. 2015. Disponível em: <http://www.jfsowa.com/pubs/semnet.htm>.

TALMY, L. *Toward a Cognitive Semantics, Vol 1: Concept Structuring Systems*. MIT Press, 2000.

TORAL, A.; MONACHINI, M. *SIMPLE-OWL: a generative lexicon ontology for nlp and the semantic web*. In: Workshop of Cooperative Construction of Linguistic Knowledge Bases, 2007.

TORRENT, T. et al. *Copa 2014 FrameNetBrasil: a frame-based trilingual electronic dictionary for the Football World Cup*. In: COLING (Demos). 2014. p. 10–14.

TORRENT, T. T.; ELLSWORTH, M. *Behind the Labels: Criteria for defining analytical categories in FrameNetBrasil*. *Revista Veredas*, v. 17, n. 1, p. 44–65, 2013.

ZIEM, A. *Frames of Understanding in Text and Discourse: Theoretical foundations and descriptive applications*. John Benjamins Publishing Company, 2014. (Human Cognitive Processing). ISBN 9789027269645.

Recebido em 04/04/2016

Aceito em 15/05/2016

THE LINGUISTIC-COMPUTATIONAL MODELING OF METONYMY IN THE FRAMENETBRASIL LEXICAL DATABASE: A CASE STUDY¹

Maucha Andrade Gamonal¹ (UC Berkeley/UFJF)² e Tiago Timponi Torrent (UFJF)³

ABSTRACT

This work reports on ongoing research aimed at modeling a metonymic relationship in the FrameNet Brasil database. This paper is based on a case study with the `Teams` frame. Both the frame and the corpus consulted are part of a frame-based trilingual (Portuguese – Spanish – English) electronic dictionary covering the soccer, tourism and World Cup domains developed by FrameNet Brasil. The basic infrastructure, analytical categories and methodology used were those developed for FrameNet (Fillmore et al. 2003, Baker et al. 2003, Ruppenhofer et al. 2010), which can be defined as an application of Frame Semantics to practical lexicography.

KEYWORDS: metonymy, Teams Frame, FrameNet Brazil, FrameNet, Frame Semantics.

RESUMO

Este trabalho registra pesquisa em andamento cujo objetivo é modelar uma relação metonímica para a base de dados da FrameNet Brasil. Este artigo é baseado em um estudo de caso com o frame `Teams`. Tanto o frame destacado para análise quanto o corpus consultado são produtos do dicionário eletrônico trilingue (Português - Inglês - Espanhol) baseado em frames semânticos para os domínios do futebol, do turismo e da copa do mundo. A infraestrutura básica, as categorias analíticas e a metodologia utilizadas são as desenvolvidas pela FrameNet (Fillmore et al. 2003, Baker et al. 2003, Ruppenhofer et al. 2010), que pode ser definida como uma aplicação da Semântica de Frames para a lexicografia prática.

1 The m.knob project is funded by CNPq (grant #448990/2014-8) and FAPEMIG (grant # CHE-APQ-00471-15). Authors are thankful to Eve Sweetser for the suggestions made to this paper. The remaining mistakes are our own.

2 Visiting Student Researcher at the Linguistics Department of the University of California Berkeley and Visiting Scholar at the International Computer Science Institute through the Science Without Borders Program. PhD Student at the Federal University of Juiz de Fora, Brazil.

3 Professor at the Federal University of Juiz de Fora, Brazil. Head of the FrameNetBrasil Computational Linguistics Laboratory.

PALAVRAS-CHAVE: metonímia, Frame Teams, FrameNet Brasil, FrameNet, Semântica de Frames.

INTRODUCTION

The purpose of this paper is to present, as a case study, the importance of developing a linguistic-computational model to account for metonymy in the FrameNet Brasil lexical database. The Teams⁴ frame was chosen to show one of the inconsistencies in the lexicographic annotation process due to the lack of a relation modeling the extension or contiguity of meaning via metonymy.

The linguistic-computational modeling carried out in this work aims to devise theoretical-methodological strategies for metonymic disambiguation, combining studies in Cognitive and Computational Linguistics. The work is conducted at FrameNet Brasil, a Computational Linguistics Laboratory at the Federal University of Juiz de Fora that develops lexical and syntactic resources for Brazilian Portuguese and works in cooperation with the Berkeley FrameNet, at the International Computer Science Institute, Berkeley. Both projects use the theoretical assumptions of Frame Semantics.

Considering the increasing demand for research focused on computational tasks and Natural Language Understanding (NLU) (Allen 1995), the purpose of this work is to refine frame to frame and frame element to frame element relation in FrameNet Brasil's lexical database. The study builds on previous research (Gamonal 2013, Peron-Côrrea 2014), which proposes guidelines for the development of the 2014 World Cup FrameNet Brazil Dictionary, a trilingual electronic dictionary (English, Portuguese, Spanish) covering the domains of the World Cup, Football, and Tourism according to Frame Semantics and FrameNet methodology.

The results of the analysis presented here provide background for linguistic and computational solutions adopted by the FrameNet Brasil team in the development of m.knob, a Multilingual Knowledge Base featuring a translation machine and a recommendation system for tourism and sports.

2. TRADITIONAL AND COGNITIVE STUDIES ON METONYMY

In this section, we present a brief review of studies related metonymy, from the traditional view of the phenomenon to the research in Cognitive Linguistics.

2.1 TRADITIONAL STUDIES ON METONYMY

In the majority of Brazilian Portuguese grammars that include the description of metonymy, it is found in the section dedicated to figures of speech or rhetorical figures. This is due mainly to the Aristotelian legacy. Aristotle takes founding role in the systematization of reasoning studies. Interested in discussing logic (the Greek word λογική 'logos'), Aristotle approached human language as a means of analyzing the ability of thought / reason.

⁴ In this paper, Courier font is used to highlight frame names.

As meaning production is a complex process, and the focus of his studies was on logical reasoning, Aristotle decided to separate the study of language in two areas: philosophical-scientific studies and rhetorical-poetic studies. The first would include, for example, the declarative propositions, those supporting inference methods, through which it is possible to achieve the comprehension of new pieces of knowledge from previously given knowledge⁵. In turn, the second would account for issues such as metaphor and metonymy.

Aristotle realized that language itself could not be construed only in regards to truth conditions, i.e., not everything can be judged as true or false. Therefore, all that exceeded the scope of declarative propositions should be conceived as a demand to the field of rhetoric and poetics.

Etymologically, metonymy means name change (from the Greek μετα ‘goal’ – Όνομα ‘onoma’), and the concept has long been interpreted as a kind of artistic ornament, because of the impossibility to literally interpret metonymic expressions. The analytical option of traditional studies has been thus to consider it as part of figurative language, linking its use to strategies for guaranteeing persuasion or the communication of emotions and passions.

2.2. COGNITIVE STUDIES ON METONYMY

Cognitive Linguistics (CL) assigns another epistemological point of view to the metonymic phenomena. Because of its empirical nature, CL includes metonymy as a result of human categorization processes linked to human experiences (Barcelona 2003, Lakoff 1987, Dancynger & Sweetser 2014).

Several studies in CL gathered scientific evidence supporting the centrality of metonymy in human language. In a founding text, Lakoff (1987) shows that metonymy works as a cognitive mechanism to ensure understanding. According to the same theoretical orientation, Radden and Kövecses (1999) consider the source field of metonymic mappings as the vehicle that will provide access to other conceptual entity within the same Idealized Cognitive Model (ICM), while Lakoff and Turner (1989) claim metonymy activates relations between two domains within the same cognitive macro-domain. Croft (1993), in turn, analyzes metonymy as the process of highlighting entities.

Barcelona (2003) points out that metonymy did not get the same attention given to metaphoric phenomena in CL studies; however, several researchers recognize the role of metonymy in the conceptualization processes related to language. Especially, the relationship between metaphor and metonymy shall be investigated rigorously, because, according to this author, metaphorical extension is metonymically structured in many cases.

Considering the studies of Rosch (1978) – discussed by Lakoff (1987) when proposing radial categories –, Barcelona (2003) indicates that categorization by prototypes is itself a metonymic operation, considering that a domain is organized in terms of the attributes of a prominent subdomain.

5

As shown in the classic example: “Every man is mortal. Socrates is a man. Therefore, Socrates is mortal.”

In the description of the general principles of metonymy, Lakoff (1987) highlighted the importance of the background, i.e., the cultural models developed by societies.

According to Lakoff (1987), the metonymy model is a model of how A and B are related in a given conceptual framework:

- i) There is a “target” concept *A* to be understood for some purpose in some context;
- ii) There is a conceptual structure containing both *A* and another concept *B*;
- iii) *B* is either part of *A* or closely associated with it in that conceptual structure. Typically, a choice of *B* will uniquely determine *A*, within that conceptual structure;
- iv) Compared to *A*, *B* is either easier to understand, easier to remember, easier to recognize, or more immediately useful for the given purpose in the given context (Lakoff 1987:84).

Dancygier and Sweetser (2014) organize metonymy into two groups: categorial metonymy and frame metonymy, explaining that, while the first is based on a relationship between a larger category and a smaller category included in the larger one, the latter characterizes a relationship between parts of the same frame. An example discussed by the authors is *the White House*, an important kind of **part-whole** metonymy. In (1), notice that the name of the building is used to refer to the entire executive branch of the US government.

- (1) *The White House* decided to leave its options open⁶

The authors also argued metonymy relates to iconicity. Non-verbal signs at public toilets doors distinguish the space reserved for men, women, disabled people and parents with their babies, for example. This is possible because metonymic icons point to the culturally relevant categories included in each case. For instance, if the sign features a high heel, it designates a restroom reserved for women, and if it features a wheelchair, the space is adapted to people with disabilities.

Note that the CL approach to metonymy is not restricted to a cluster model accounting for the exchange of lexical items, but, rather a conceptual framework in which one cognitive construct stands out from the others, assuming the status of representative for the whole. Thus, the compositional approach to language meaning, according to which the meaning of the whole is a sum of the meanings of its constituent parts does not cover the reality of natural languages, given the existence of pervasive cognitive process such as metonymy, which, based on a reference, explore inference and reveal the cultural aspects involved in language comprehension.

⁶ https://www.washingtonpost.com/world/national-security/how-the-obama-white-house-runs-foreign-policy/2015/08/04/2befb960-2fd7-11e5-8353-1215475949f4_story.html. Accessed in December 1st 2015.

3. FRAME SEMANTICS AND FRAMENET

Frame Semantics (Fillmore 1976, 1977, 1982) is an empirical approach to semantic description and analysis that emphasizes the close relationship between knowledge and experience, and demonstrates the complex networks of meanings involved in such a semantic description.

The frame is the basic analytical unit of Frame Semantics. According to Fillmore:

By the term 'frame' I have in mind any system of concepts related in such a way that to understand any one of them you have to understand the whole structure in which it fits; when one of the things in such a structure is introduced into a text, or into a conversation, all of the others are automatically made available (Fillmore 1982:111).

The theoretical and methodological foundations of Frame Semantics have been applied to the construction of a computational lexicon: the FrameNet project⁷. FrameNet provides a frame-base syntactic and semantic description of the English lexicon, which is grounded on corpus evidence. Also, the database can be useful in the development of resources related to Natural Language Processing (NLP).

According to Fillmore (2008), the main objectives of FrameNet are:

- i) describing lexical units (LUs) in terms of the semantic frames they evoke, and describing those frames;
- ii) defining the frame elements (FEs) of each frame that are essential for a full understanding of the associated situation type;
- iii) extracting from a very large corpus example sentences which contain each LU targeted for analysis;
- iv) selecting from the extracted sentences representative samples that cover the range of combinatorial possibilities, and preparing annotations of them as layered segmentation of the sentences, where the segments are labeled according to the FEs they express, as well as the basic syntactic properties of the phrases bearing the FE,
- v) displaying the results in lexical entries which summarize the discovered combinatorial affordances, both semantic and syntactic, as valence patterns, and creating links from these patterns to the annotated sentences that evidence them, and
- vi) defining a network of frame-to-frame relations and the graphical means of displaying these, that will show how some frames depend on or are elaborations of other frames. (Fillmore 2008: 49,50).

7 <https://framenet.icsi.berkeley.edu/fndrupal/>.

FrameNet methodological proposal has been extended to other languages, in addition to Brazilian Portuguese, such as German, Mandarin, Spanish, Japanese and Swedish. In Brazil, this resource has been developed at the Federal University of Juiz de Fora. All data is made available online.⁸

4. THE COPA 2014 FRAMENET BRASIL PROJECT

The Copa 2014 FrameNet Brasil project (Salomão et al. 2011, 2013) developed a frame-based trilingual (Portuguese – Spanish – English) electronic dictionary to soccer, tourism and World Cup domains⁹. Multilingual challenges involved in the construction of the dictionary are described in Torrent et al. (2014).

Table 1 shows the number of frames described for each domain in each language, lexical units and semantically and syntactically annotated sentences.

128 Trilingual Frames (124 new)	World Cup: 30	
	Soccer: 53	
	Tourism: 45	
1,125 Lexical Items	World Cup: 366	Br-Portuguese: 134
		English: 112
		Spanish: 134
	Soccer: 438	Br-Portuguese: 163
		English: 124
		Spanish: 151
	Tourism: 321	Br-Portuguese: 118
		English: 114
		Spanish: 89
13,451 Annotated Sentences	Br-Portuguese: 7,912	
	English: 3,374	
	Spanish: 2,398	

Focused on human users, the Copa 2014 dictionary was implemented as a web app. The database comprises interconnected framenet-style data – frames, lexical units, frame-to-frame relations, and annotated sentences. The main purpose was to test FrameNet structure as a means of providing semantically accurate word-to-word translations. Four different ways to access the information in the dictionary are available: searching a word, typing a sentence, browsing the list of frames grouped by cognitive domains, and exploring the frame grapher (see Figure 1).

8 <http://www.ufjf.br/framenetbr>.

9 <http://www.dicionariodacopa.com.br>.

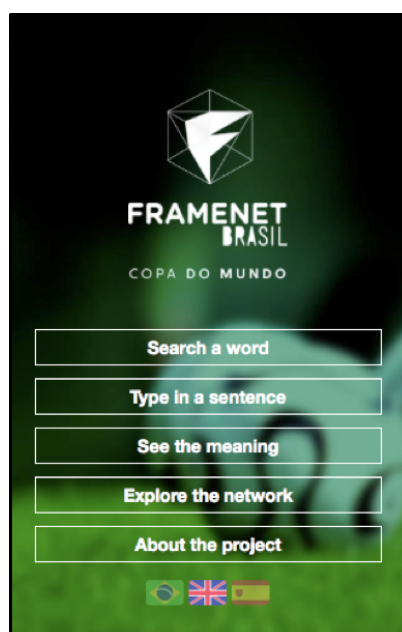


Figure 1: Main menu of the Copa 2014 App

The noun *Brazil* noun is a lexical unit that evokes the *Teams* frame. As expected, all the other names of countries that participated in the 2014 World Cup are also present in the corpora surveyed for building the dictionary: the Copa 2014 FN-Br corpus (1,001,326 tokens in Brazilian Portuguese). In this corpus, *Brasil.n* is used in many metonymies both in the soccer and in the tourism domains, as shown in Figure 2.

Based on occurrences such as the ones presented in Figure 2, the next section discusses how the *country for team* metonymy was accounted for in the dictionary database.

Rio=de=Janeiro capital do Reino=Unido de Portugal , **Brasil** e Algarves . No mesmo ano , uma grande crise financeira a regressar a Portugal . 1822 Príncipe regente do **Brasil** desde o retorno de dom João VI para Portugal , dom coroa- se imperador . 1831 Como primeiro monarca do **Brasil** , dom Pedro=I mostra- se pouco hábil e abdica do economia . As fazendas produtoras se multiplicam e o **Brasil** se torna um grande exportador de café . 1835 Inspirados levante a Revolta=dos=Malês , que quase triunfa . 1865 O **Brasil** se alia ao Uruguai e à Argentina em=torno=dea Tríplice=Aliança mortos e elimina metade da população paraguaia . 1888 O **Brasil** se torna o último país do Novo=Mundo a abolir a escravidão República . 1890 Com a abolição da escravidão , o **Brasil** abre fronteiras para suprir sua demanda por mão=de=obra expande a força militar e centraliza o poder . 1942 O **Brasil** abandona a neutralidade e entra na Segunda=Guerra=Mundial o na inflação que persistiria por décadas . 1958 O **Brasil** ganha sua primeira Copa=do=Mundo , triunfando sobre famílias tem hoje mais=de 1,5 milhão de filiados . 1985 O **Brasil** realiza uma eleição presidencial indireta . Tancredo=Neves República , a primeira mulher a ocupar o cargo no **Brasil** . Sucessora escolhida a=dedo por Lula , ela mantém em=grande=parte as políticas de seu antecessor . 2014 O **Brasil** sedia a Copa=do=Mundo de 2014 ao custo de R\$ 27 bilhões chegaram em 1500 , o território que hoje conhecemos como **Brasil** já era habitado há cerca=de 12 mil anos . Os primeiros encontraram alguns indícios anteriores da presença humana no **Brasil** . Os mais antigos vestígios do gênero na região Amazônica podem ter 12 mil anos . Na época do descobrimento do **Brasil** , é provável que vivessem no território entre 2 milhões pessoas . Cabral e companhia O curso da história do **Brasil** mudou para=sempre em 1500 , quando uma frota de 12 tratar . Os primeiros colonizadores só aportaram no **Brasil** em 1531 . O filósofo francês Jean-Jacques Rousseau idioma tupi-guarani mais tarde se espalhou por todo=o **Brasil** colonial , e ainda é falada por algumas pessoas na itinerantes de aventureiros que exploraram o interior do **Brasil** nos séculos 17 e 18 , pilhando assentamentos indígenas quando=dea chegada dos portugueses , em 1500 , já viviam no **Brasil** entre 2 milhões e 4 milhões de pessoas , divididas

Figure 2: Occurrences of *Brasil.n* in the Copa 2014 FN-Br Corpus

5. TEAMS FRAME AND METONYMIC RECURRENCE

As shown in Table 1, several frames were created to model the soccer and the World Cup experiences in the dictionary. One of them was the `Teams` frame. According to the description in Figure 3, the core frame elements (FE) are `SOCCKER_TEAM` and `COUNTRY`. This methodological decision was due to the fact that the *country for team* metonymy is highly productive in this frame, meaning that it surfaces frequently in the corpus.

However, according to FrameNet methodology, there is an important distinction that must be made between **core** and **peripheral** frame elements: while the first instantiates a conceptually necessary component of a frame, differentiating it from other frames, the latter do not introduce additional, independent or distinct events from the main reported event, being responsible for indicating circumstances such as time and place, for example (Ruppenhofer et al. 2010:19-20).

Note that, in sentences like (2-3) the `COUNTRY` FE represents the same role as the `SOCCKER_TEAM` FE.

- (2) O [Brasil^{COUNTRY}] enfrentará a Alemanha na semifinal.
- (3) [Brazil^{COUNTRY}] will play Germany in the semifinals.

Because of the high prevalence of structures like the ones in (2-3) in the corpus, the `COUNTRY` FE was included as a core FE to mark its occurrence to indicate the teams rather than only the countries they represent. However, this decision is not lexically consistent because nouns like *team.n*, *squad.n*, *seleção.n* and *equipe.n*, which may be accompanied by genitives, adjectives or prepositional phrases that designate the name of the countries, also appear in the corpus, as shown in (4-6).

- (4) [Brazil's^{COUNTRY}] [squad^{TEAM}]
- (5) The [Brazilian^{COUNTRY}] [squad^{TEAM}]
- (6) [Equipe^{TEAM}] [do Brasil^{COUNTRY}].

Teams

[Lexical Unit Index](#)

Definição [Definition]:

A **Soccer_team** that represents a **Country** in The World Cup Tournament and is formed by a group of 23 players, being 11 in the official group and 12 in the reserve group. The official group must contain 1 goalkeeper and 10 line players.

Elementos de Frame [Frame Elements]:

Nuclear [Core]:

Country Place of origin represented by the soccer team.

Soccer_team Group of players that plays the game.

Não-Nucleares [Non-Core]:

Depictive A quality or characteristic of the **Soccer_team** or **Country**.

Figure 3: The Teams frame

The metonymy is productive at the frame level, since all the names of countries involved in the World Cup are LUs that evoke this frame (see Figure 4). Therefore, the decision to include COUNTRY as a core FE was based on the absence of an alternative to properly indicate the *country for team* metonymy in the lexical database.

Unidade Lexical [Lexical Unit]	Status da LU [LU Status]	Rel Entrada Lexical [Lexical Entry Report]
Algeria.n	Created	Entrada Lexical [Lexical entry]
Argentina.n	Created	Entrada Lexical [Lexical entry]
Australia.n	Created	Entrada Lexical [Lexical entry]
Belgium.n	Created	Entrada Lexical [Lexical entry]
Bosnia-Herzegovina.n	Created	Entrada Lexical [Lexical entry]
Brazil.n	Created	Entrada Lexical [Lexical entry]
Cameroon.n	Created	Entrada Lexical [Lexical entry]
Chile.n	Created	Entrada Lexical [Lexical entry]
Colombia.n	Created	Entrada Lexical [Lexical entry]
Costa Rica.n	Created	Entrada Lexical [Lexical entry]
Croatia.n	Created	Entrada Lexical [Lexical entry]
Ecuador.n	Created	Entrada Lexical [Lexical entry]
England.n	Created	Entrada Lexical [Lexical entry]

Figure 4: LUs in the Teams frame

As discussed in section two, metonymy is a role to role relation, in which the mapping between source and target entails a relationship between roles within the frame while metaphors are mappings between frames in different cognitive domains.

FrameNet recognizes that frame elements are not independent of each other, they are related to the frame, required by it, and interrelated in different ways. There are three types of frame element relations in FrameNet: *coreset*, *requires*, and *excludes*.

The **coreset** relation occurs when a frame has FEs that act like sets, so that the presence of any member of the set is sufficient to satisfy the semantic valence of the predicator. For instance, DIRECTION, GOAL, PATH, and SOURCE are all core FEs in the TRAVEL frame. As an example, in (7), only the PATH FE is expressed, and the sentence is semantically complete, even without the other core FEs.

In contrast, in **excludes**-relation, if a FE is expressed, another will not be expressed. For instance, in (8), the GOAL FE in the Travel frame excludes the AREA FE. Hence, the presence of both FEs would render the sentence ungrammatical. In the opposite direction, the **Requires** relation happens when the occurrence of one core FE requires that another core FE also occur. For instance, in (9), if ITEM occurs, then GOAL is required.

(7) [We^{TRAVELER}] traveled^{TARGET} [across France^{PATH}].

(8) *[Pedro^{TRAVELER}] journeyed^{TARGET} [around the Caucasus^{AREA}] [to Europe^{GOAL}] [last month^{TIME}].

(9) [He^{AGENT}] attached^{TARGET} [the message^{ITEM}] [to the wall^{GOAL}].

In order to mark the metonymic relationship between *country* and *team*, and many other **whole-part** metonymies, for example, the addition of a FE-to-FE metonymic relation in the FrameNet Brasil database can prove to be appropriate, because the relations recognized by FrameNet are not specific to metonymies. This inclusion would enable peripheral FEs to take on core status without leading to theoretical inconsistencies.

The proposed intraframe relation, which will be referred to here as **metonymic substitution**, could be added to the list of FE-to-FE mappings to allow non-core FEs to substitute core FEs in sentences evoking a given frame.

The kind of metonymic mapping discussed for the `Teams` frame is not restricted to the soccer domain. Common and proper nouns referring to places engage in a variety of metonymies. Such entities can be described in a specific frame. Nevertheless, it is necessary to lexicographically validate such additions to the database. For instance, in the `People_by_origin` frame (Figure 5), one finds both `PERSON` and `ORIGIN` as core FEs.

In LUs such as *german.n*, FrameNet proposes an analysis in which both core FEs are incorporated to the LU, as shown in Figure 6, since a *german* is, at the same time, a `PERSON` defined in regards to some `ORIGIN`. Again, the metonymic substitution relation could provide a more theoretically accurate model for cases like that.

People_by_origin

[Lexical Unit Index](#)

Definition:

This frame contains words for individuals, i.e. humans, with respect to their `Origin`. The `Origin` is generally incorporated, but may occasionally be specified separately. The `Person` is conceived of as independent of other specific individuals with whom they have relationships and independent of their participation in any particular activity. They may have an `Age`, `Description`, `Persistent_characteristic`, or `Ethnicity`.

I'm a proud BRIT.

FEs:

Core:

`Origin (Ori)`

The `Origin` is the place where the `Person` was born or lived a salient part of her or his life. In this frame, it is generally incorporated in the lexical item, but may be separately specified as in the following:
`FOREIGNERS from Japan` get short shrift around here.

`Person (Person)`

Semantic Type: Human

The `Person` is the human being whose `Origin` is specified.

Figure 5: The `People_by_origin` frame

german.n

Frame Element	Core Type
Age	Peripheral
Context_of_acquaintance	Peripheral
Descriptor	Peripheral
Ethnicity	Peripheral
Origin	Core
Persistent_characteristic	Peripheral
Person	Core

[Turn Colors Off](#)

- added
 1. Following this debacle , the Ottomans then allied themselves to Germany in the World War I , losing more territory with the defeat of the **GERMANS** in that war .
 2. Later the **GERMANS** came in force and occupied many of the islands .
 3. Today Jamaican population is a mixture of African and English , with Spanish , Indian , and a smattering of Portuguese Jews , **GERMANS** , Welsh , and Scots .

Figure 6: Example annotation for *german.n*

Since the purposes of this work are both to improve the adequacy of computational decisions in the FrameNet Brasil database, and, at the same time, to implement CL theory on metonymy, another relation must be proposed to improve the system capability to perform inferences, as well as to define the types of constraints to which metonymies are subject: a frame-to-frame – or interframe – relation referred to as **metonymic grounding**.

Let's return to the discussion concerning the *Teams* frame. In the context of the World Cup, the FE defined to indicate the team's origin was *COUNTRY*. Nevertheless, because teams are always headquartered at – and, therefore represent – a location, the deployment of place names to metonymically refer to teams is very productive. Hence, the system would gain in the capability to perform inferences if a relation between the incorporated FE denoting a location and a frame referring to locations is modeled. In other words, the system would be able to infer that Germany is a soccer team, for example, if the *COUNTRY* FE, and the LUs incorporating it, are mapped to a frame referring specifically to countries, in this case, the *Political_locales* frame.

However, there are issues to be verified, for example: are metonymies based on general words such as *country.n*, *city.n* and *territory.n* as productive as those involving proper nouns? Again, a proper corpus study and lexicographic validation will be necessary, because the results of these analyses will interfere in the way the frame and the relations involving it and its FEs are defined. Let us consider sentences (10-12).

- (10) **Brazil** lost to **Germany** in the World Cup semifinal match by six goals.
- (11) The host **country** lost to **Germany** by a 7 to 1 score.
- (12) ??? **Brazil** lost to **the country**.

Sentences (10) and (11) show two different ways to refer to probably the most infamous part of the Brazilian soccer squad history. For everyone minimally initiated in soccer, the interpretation of (10) and (11) are trivial. However, in (12), the substitution of the proper noun Germany by the corresponding common noun country, without any adjective that differentiates it from other countries, render the sentence odd.

Thus, since (12) is not likely to be found in corpora, while (10) and (11) are, the semantic and syntactic valences provided by lexicographic annotation validate such uses and show their constraints.

6. CONCLUSIONS

Cognitive Linguistics embraces different concepts on metonymy that reflect different approaches to this topic. There is linguistic motivation to include metonymic relations in the FrameNet Brasil database, and the lexicographical relevance of the case presented in this paper is in accordance with the FrameNet aim to provide lexical databases grounded on a network of cognitive domains. Hence, bringing together Frame Semantics – as a theory that presents the lexicon not as simple lists of words, but as a system of related concepts – and the studies on Metonymy is crucial to the progress of this and other works focused on accurately representing human languages in computational models.

The intraframe relation proposed in this work – metonymic substitution – is an appropriate means of showing which metonymies are lexicographically attested in a given language, while interframe relation – metonymic grounding – could provide constraints and enable the system to perform inferences.

The next steps in this research include increasing the corpus analyses in order to identify lexicographic bases for the proposition of new instances of these two relations.

REFERENCES

ALLEN, J. *Natural Language Understanding*. Menlo Park, CA: Benajmins Cummings Publishing, 1995.

BARCELONA, A. *Metaphor and metonymy at the crossroads: a cognitive perspective*. Berlin: Mouton de Gruyter, 2003.

BAKER, C.; FILLMORE, C. J., AND CRONIN, B. *The Structure of the FrameNet Database*, vol. 16, Google ScholarBibTexRTFTaggedXML, 2003, pp. 281-296.

CROFT, W. The role of domains in the interpretation of metaphors and metonymies. *Cognitive Linguistics* 4, 1993, pp.335-370.

DANCYGIER, B, & SWEETSER, E. *Figurative language*. Cambridge, UK: Cambridge University Press, 2014.

FILLMORE, C.J. Scenes and frames semantics. In: ZAMPOLLI, A. (ed.). *Linguistic Structures Processing: Fundamental Studies in Computer Science*, n. 59. Amsterdam: North Holland Publishing, 1977.

FILLMORE, C.J. Frame semantics and the nature of language, *Annals of the New York Academy of Sciences: Conference on the Origin and Development of Language and Speech*, v. 280, 1976, pp. 20-32.

- FILLMORE, C. J. Frame semantics. *Linguistics in the morning calm*, ed. by The Linguistic Society of Korea, Seoul: Hanshin Publishing Co, 1982, pp.111-137.
- FILLMORE, C.J. Border Conflicts: FrameNet Meets Construction Grammar, in *Proceedings of the XIII {EURALEX} International Congress*, Barcelona, 2008, pp. 49-68.
- GAMONAL, M. A. *Copa 2014 FrameNet Brasil: diretrizes para a constituição de um dicionário eletrônico trilingue a partir da análise de frames da experiência turística*. M. A. Dissertation in Linguistics. Universidade Federal de Juiz de Fora, Juiz de Fora, 2013.
- LAKOFF G. *Women, Fire, and Dangerous Things*. Chicago: University of Chicago Press, 1987.
- LAKOFF, G. & M. TURNER. *More than Cool Reason: A Field Guide to Poetic Metaphor*. Chicago & London: The University of Chicago Press, 1989.
- PERON-CORRÊA, S. R. *Copa 2014 FrameNet Brasil: frames secundários em unidades lexicais evocadoras da experiência turística em português e em espanhol*. Dissertation in Linguistics. Universidade Federal de Juiz de Fora, Juiz de Fora, 2014.
- RADDEN, G. & Z.KOVECSES. Towards a Theory of Metonymy. In: PANTHER K-U. & RADDEN, G. (eds.). *Metonymy in Language and Thought*. Amsterdam & Philadelphia: John Benjamins, 1999, pp.17-59.
- RUPPENHOFER, J; ELLSWORTH, M; PETRUCK, M.; CHRISTOPHER R. JOHNSON; SCHEFFCZYK. J. *FrameNet II: Extended theory and practice*. Berkeley, California: International Computer Science Institute, 2010.
- SALOMÃO, M. M. M.; TORRENT, T. T.; CAMPOS, F. C. A.; BRAGA, R. M. M. & VIEIRA, M. B. *Copa 2014 Framenet Brasil*. Projeto apresentado ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), no âmbito da Chamada Universal no 14/2011. Juiz de Fora, 2011.
- TORRENT, T. T.; ELLSWORTH, M. Behind the labels: criteria for defining analytical categories in FrameNetBrasil. *Veredas*, v.17, n.1, 2013, pp. 44–65.
- TORRENT, T. T. SALOMÃO, M. M. M.; CAMPOS, F. C.; BRAGA, R. M.; MATOS, E. E.; GAMONAL, M. A.; GONÇALVES, J.; GOMES, D. S.; SOUZA, B. C. P. & PERON-CORREA, S. R. *Copa 2014 FrameNetBrasil: a frame-based trilingual electronic dictionary for the Football World Cup*. In: *Proceedings of the 25th International Conference on Computational Linguistics (COLING 2014) – System Demonstrations*. Dublin. 2014, pp. 10-14.

Recebido em 31/03/2016

Aceito em 30/04/2016

METÁFORA E METONÍMIA EM NOMES COMPOSTOS EM PORTUGUÊS: UM ESTUDO DE CONSTRUÇÕES S-ADJ

Margarida Basilio (PUC-RJ)¹

RESUMO

O trabalho se concentra na análise do aspecto semântico de construções lexicais compostas de forma S-Adj no Português do Brasil, com ênfase no papel da metonímia e da metáfora nestas construções, dentro de uma visão cognitiva do léxico. Analisam-se compostos metafóricos e metonímicos semanticamente exocêntricos e endocêntricos e discutem-se questões relativas à distinção metáfora/metonímia e à dificuldade de análise de construções lexicais tradicionalmente consideradas como compostos.

PALAVRAS-CHAVE: semântica lexical – metáfora – metonímia – compostos S-Adj – Língua Portuguesa

ABSTRACT

The work is concerned with the semantic analysis of Portuguese compounds formed by a noun followed by an adjective. Taking a cognitive approach to the study of the lexicon, the author emphasizes the role of metaphor and metonymy in these lexical constructions. Metaphoric and metonymic compounds in Portuguese are focused in the analysis, which includes both exocentric and endocentric compounds. The analysis is permeated by a discussion of the main questions involved in the metaphor/metonymy distinction and in the difficulties in the analysis of lexical constructions traditionally called compounds.

KEYWORDS: lexical semantics – conceptual metaphor – conceptual metonymy – N-Adj compounds – Portuguese

¹ Professora Emérita - Pontifícia Universidade Católica do Rio de Janeiro – Departamento de Letras – Programa de Pós-Graduação em Estudos da Linguagem. marbas@centroin.com.br

INTRODUÇÃO

Este trabalho é parte de uma pesquisa mais abrangente, que tem como objetivo investigar o papel da metáfora e da metonímia conceptual, assim como o carácter emaranhado dos dois tipos de figuração em construções lexicais do Português.² O foco específico do artigo é a análise da relação semântica adjetivo/substantivo em formações compostas S-Adj no português do Brasil, com ênfase no papel da metáfora e da metonímia conceptuais e seus entrelaçamentos em construções lexicais compostas dentro de uma perspectiva da Linguística Cognitiva.

Inicialmente, o trabalho aborda rapidamente conceitos e procedimentos preliminares, como a questão do composto e a definição dos conceitos de metáfora e metonímia adotados no trabalho, além de especificar as fontes de dados. Em seguida, alguns exemplos de compostos metafóricos são analisados, sendo determinados dois tipos de formação metafórica. O trabalho prossegue com a análise de compostos metonímicos e, em sua terceira parte, analisa diferentes tipos de compostos em português, o que aponta para a necessidade de se investigar mais as características da possível unidade “composto” como elemento demarcador para um melhor entendimento da fluidez das construções lexicais à luz das proposições da Gramática Cognitiva.

2. CONCEITOS E PROCEDIMENTOS PRELIMINARES

Tradicionalmente consideramos como compostos palavras que consistem de duas ou mais palavras ou lexemas, embora as definições possam variar, na dependência de diferentes arcabouços teóricos de perspectivas analíticas (Lieber&Stekauer 2011:4-5). Por outro lado, é bastante complexa a questão do conceito de palavra.³ Em consequência, um dos maiores problemas na investigação de compostos é a dificuldade de estabelecer quando temos ou não uma palavra composta (Fabb 1998, Dressler, Kastovsky & Pfeiffer 2005, Benczes 2006, Lieber 2010, Rio-Torto 2013, etc.). O trabalho não se detém nesta relevante questão tradicional, mas esclarece alguns pontos relevantes da discussão e especifica critérios adotados na seleção dos dados, na esteira de proposições anteriores, mas tendo em vista a perspectiva da Gramática Cognitiva, em que se assumem pressupostos teóricos tais como o contínuo léxico-gramática, a noção de categorias não discretas, a metáfora e a metonímia conceptuais, a relevância do conhecimento enciclopédico como parte do significado lexical, e assim por diante.

Conforme observado em Basilio (2012), dentre outros, a dificuldade de distinção entre compostos e outras construções lexicais é diretamente relacionada ao conceito de palavra composta, visto que a noção de composição incide sobre palavras, sendo estas consideradas tanto do ponto de vista lexical quanto do ponto de vista gramatical (morfológico), sem esquecermos outros aspectos, como o fonológico e o gráfico, que não serão aqui abordados.

Neste trabalho, em que o maior interesse está no aspecto semântico e, mais especificamente, no papel de mecanismos de figuração nas construções que normalmente consideramos como compostos, os

2 O trabalho se vincula diretamente ao projeto “Domínios indefinidos: um estudo da metáfora e metonímia em compostos e expressões lexicais”, financiado pelo CNPq (bolsa de produtividade, proc. Nº 306150/2014-0).

3 Para um maior detalhamento desta discussão no Português e seu impacto na questão dos compostos, v. Basilio 2000.

dados centrais são os chamados compostos lexicais, isto é, construções que podem ser consideradas como unidades lexicais. Ou seja, o trabalho pressupõe uma distinção entre construções lexicais, que, embora constituídas com estruturas normalmente abordadas no nível sintático, servem a fins lexicais de denominação consolidada, em oposição a construções que ocorrem no âmbito dos enunciados linguísticos nos atos de fala.

Mais especificamente, conforme esclarecido com mais detalhe no trabalho mencionado, expressões S-Adj de função denotativa em que o segundo termo é um adjetivo denominal em geral podem ser consideradas como expressões que eventualmente ocorrem nos enunciados, não tendo atingido um nível de estabilidade ou consolidação como construções lexicais compostas. Este seria o caso de sequências como *renda familiar*, *expressão linguística*, etc., que apresentam um alto nível de transparência.

Ao lado desses, podemos considerar casos como *merenda escolar* e código penal, que, embora mantendo as características gerais dos exemplos acima, apresentam um nível mais significativo de consolidação no léxico, assim como uma diminuição no grau de composicionalidade, de modo que especificações de conhecimento enciclopédico se fazem mais necessárias e relevantes. Em ambos os casos, entretanto, temos construções semanticamente endocêntricas, no sentido de que o adjetivo apenas restringe ou caracteriza o significado genérico do substantivo.

Um terceiro caso que ilustra o caráter de gradualidade nessas expressões é o *de pessoa jurídica*, em que, ainda que o adjetivo possa restringir o significado do substantivo, o significado do conjunto já não pode ser analisado como formado pelo mero acréscimo do adjetivo especificador a um substantivo. Isto é, não se trata de uma pessoa adjetivada como pertinente a algo considerado como jurídico, como um indiciado ou um procurador, mas um conceito abstrato que decorre de um complexo sistema de relações institucionalizadas. Neste caso, já é mais difícil dizer que *pessoa jurídica* se refere a uma pessoa; na realidade a evocação mais imediata da sequência traria uma resposta negativa. Assim, em *pessoa jurídica* não é possível inferir o significado do todo em função do significado das partes. Ou seja, a interpretação específica agrega-se à expressão lexicalizada. Trata-se, portanto, de um composto ou expressão lexical.⁴

Assim, não causa surpresa o fato de que esta última construção não admite a substituição de um dos termos por sinônimos (?pessoa judicial, ?gente jurídica, etc.) nem elementos interferentes (?pessoa especialmente jurídica, ?pessoa sociojurídica, etc.), critérios normalmente utilizados para a caracterização de uma sequência como composto.⁵

Os dados a serem analisados neste trabalho foram retirados de um corpus de construções lexicais coletadas ao longo dos últimos anos no jornal O GLOBO, acrescido de uma pequena quantidade de dados variados encontrados recentemente em diferentes meios de comunicação, além de construções reconhecidas como

4 De acordo com Langacker (2002:292), as unidades geralmente consideradas como itens lexicais se reúnem no espaço da especificidade. Para o autor, o léxico é simplesmente o inventário das expressões fixas numa língua.

5 Para um maior detalhamento de critérios e argumentos, v. Basilio (2012), em que, no entanto, algumas posições são diferentes das adotadas no presente trabalho.

compostos no Português Brasileiro, tais como, por exemplo, construções arroladas como compostos em gramáticas do Português, dicionários ou artigos sobre formação de palavras em português.

O artigo tem por base compostos S-Adj, considerados como compostos lexicais em termos dos critérios referidos, usuais na literatura, tais como a discrepância semântica, a rigidez da construção e a resistência à coordenação, entendendo-se, no entanto, que não se trata de critérios rígidos e que a própria noção de composto está sendo utilizada para denominar um tipo de construção cuja relevância numa abordagem cognitiva ainda está para ser considerada com mais detalhe, dada a noção da continuidade léxico-gramática.

Observe-se, ainda, que, tendo em vista que nosso foco é fundamentalmente a análise semântica, a parte numérica referente aos dados não será considerada relevante. Todos os dados analisados no trabalho, no entanto, são de conhecimento geral, embora alguns possam não estar dicionarizados, seja porque relativamente recentes, apesar de conhecidos e reconhecidos, seja em virtude das questões que envolvem a análise de compostos lexicais.

Compostos são frequentemente classificados como endocêntricos ou exocêntricos. Nos compostos endocêntricos, o composto é um hipônimo da cabeça, como em *obra-prima*; compostos exocêntricos, por outro lado, não apresentam cabeça (Booij 2007). Os compostos a serem aqui analisados são semanticamente exocêntricos em sua grande maioria; de fato, uma das preocupações do trabalho é a análise da figuração na exocentricidade semântica de compostos lexicais. O trabalho focaliza fundamentalmente compostos metafóricos e compostos metonímicos formados por uma sequência S-Adj em que o adjetivo incide sobre o substantivo, como, por exemplo, em *carta-branca* e *pé-sujo*.

Para os fins deste trabalho, assumo a distinção usual entre metonímia conceptual, mecanismo cognitivo em que fonte e alvo estão no âmbito do mesmo domínio conceptual, e metáfora conceptual, mecanismo cognitivo em que fonte e alvo estão em domínios conceptuais distintos. (Radden e Kovecses 1999, Barcelona 2003, Benczes 2006, etc.). No momento, deixo de lado a questão crucial relacionada à interpretação do termo *domínio* na distinção entre metáfora e metonímia, embora minha interpretação se aproxime da noção de modelo cognitivo idealizado (ICM). Na verdade, em convergência com a visão de Benczes (2011), um dos pontos de interesse do trabalho é obter maior conhecimento sobre a interação entre metonímia e metáfora em compostos com a finalidade de contribuir para a discussão com dados analisados, dentro do pressuposto de que a estabilidade da relação entre os dois membros do composto constitui um ambiente promissor para a investigação desta questão crucial.

Neste trabalho, então, a metonímia será fundamentalmente entendida como um processo cognitivo em que uma entidade ou (sub)domínio conceptual, a fonte, provê acesso mental a outra entidade conceptual ou (sub)domínio, o alvo, dentro do mesmo domínio cognitivo. O trabalho também incorpora a visão de metonímia de Langacker (2000, 2009) como um ponto de referência. Por exemplo, em Português o substantivo *cego*, que se origina no adjetivo *cego*, tem na propriedade “sem visão” um ponto de referência para denotar um indivíduo que é cego. A metonímia “Propriedade por Categoria” é clara em praticamente todos os casos de conversão de adjetivo para substantivo em Português.

Em relação à metáfora, nosso foco é o mecanismo cognitivo por meio do qual uma entidade conceptual ou domínio projeta alguma(s) de suas propriedades relevantes numa entidade alvo que pertence a outro domínio cognitivo, de sorte que isto nos leva a encarar o alvo de um modo que incorpora a entidade fonte. Esta visão da metáfora é, portanto, compatível com a visão de Steen (2013:57), quando este diz que “Deliberate metaphor truly is a means of understanding one thing in terms of something else”. Assim, por exemplo, se dizemos de um adulto que ele é criança, todos entenderão que estamos considerando este adulto como psicologicamente infantil.

Duas observações iniciais se fazem necessárias para informar a análise dos exemplos de construções compostas em Português.

A primeira é que metáforas lexicais em compostos sempre envolvem metonímia, pelo fato de que quando uma palavra é usada como fonte para estabelecer uma conexão com o alvo conceptual, nem todas as propriedades da palavra são evocadas, mas apenas aquelas pertinentes à conexão pretendida, de modo que conexões metafóricas são baseadas no conceito de metonímia como ponto de referência. No exemplo acima, o uso de “criança” nos leva a atribuir ao adulto assim qualificado atitudes psicologicamente infantis, tais como dependência ou falta de responsabilidade, mas propriedades relativas a condições legais ou físicas de crianças, por exemplo, são irrelevantes no caso. Do mesmo modo, no composto *reunião-relâmpago*, apenas as propriedades correspondentes a rápido, súbito ou inesperado são selecionadas.

A segunda observação é que a metáfora lexical não é equivalente à metáfora no discurso, porque quando usamos uma palavra como fonte para projetar algum(ns) dos seus significados potenciais na interpretação da palavra-alvo, esta projeção não é proposicional, sendo, antes, deliberadamente incorporada no poder evocativo da palavra-alvo. Isto aponta para a relevância da investigação da metáfora nos compostos, já que compostos são construções formadas por duas ou mais palavras, mas são palavras, isto é, unidades lexicais básicas, no dizer de Langacker, e com função de denotação, o que significa que o poder expressivo da metáfora na construção não pode ser questionado, em virtude da função denotativa/evocativa das palavras, compostas ou não.

A análise se concentra em compostos metonímicos e metafóricos parcialmente motivados, isto é, compostos cuja motivação é parcialmente reconhecida quando o falante conhece seu significado global. Por exemplo, uma vez saibamos o que é *carne-seca*, não é difícil estabelecer a conexão entre o composto e seu significado. Estas construções são, portanto, distintas dos compostos em que uma motivação, mesmo eventualmente existente, é complexa demais para o entendimento do falante médio, como, por exemplo, em *amor-perfeito*. Entretanto, incluímos na análise alguns casos de compostos apresentando algo de classificatório, que se revelaram de interesse como contraponto.

Naturalmente, em qualquer circunstância a interpretação de compostos só é possível dentro do pressuposto do conhecimento enciclopédico. Entendemos o conhecimento enciclopédico como equivalente ao conhecimento do mundo e incorporamos a visão de Langacker de que apenas uma semântica linguística de caráter enciclopédico é viável.

3. ANÁLISE DE COMPOSTOS S-ADJ

Passamos então à análise de dados específicos. A situação padrão – embora não a única - em compostos S-Adj do Português é aquela em que o segundo elemento, o adjetivo, especifica o primeiro, um substantivo. Esta é a situação que abordaremos neste trabalho.

3.1 COMPOSTOS METAFÓRICOS

A situação mais frequente que encontramos em compostos metafóricos é a de compostos semanticamente exocêntricos, de que nos ocuparemos na próxima seção. Entretanto, também existem compostos metafóricos não exocêntricos, como veremos posteriormente.

3.1.1 COMPOSTOS METAFÓRICOS EXOCÊNTRICOS

Começaremos a análise de compostos metafóricos exocêntricos com o composto *secretária-eletrônica*, usado para denotar um instrumento eletrônico que faz as vezes de telefone e gravador de mensagens. Neste composto o primeiro elemento é duplamente metonímico, já que *secretária* denota uma pessoa, mais especificamente uma mulher, por meio de uma profissão ou cargo funcional (*Papel funcional por Indivíduo*)⁶; e, naturalmente, uma secretária exerce outras tarefas além de atender o telefone (*Parte pelo Todo*).

Mas o que é peculiar a esta construção é o efeito do adjetivo *eletrônica* no substantivo *secretária*, cuja evocação central é transferida do domínio conceptual de pessoas, profissões e cargos para o domínio dos instrumentos eletrônicos, por meio de uma similaridade funcionalmente parcial, além de hiperbólica⁷.

Mais especificamente, o que podemos notar neste tipo de construção é que, embora o adjetivo incida sobre o substantivo, ele não o modifica no sentido usual do termo, mas o transforma, ou mesmo aniquila, inserindo-o num novo plano, o da construção S-Adj, que emerge num outro domínio, o dos instrumentos eletrônicos, ou seja, o domínio especificado pelo adjetivo. O aspecto mais interessante dessas construções reside, pois, no uso de uma hipérbole disfarçada, através de um mecanismo convencional de seleção.⁸

Esta construção lexical é produtiva em português em seu nível médio de esquematicidade. Ou seja, não apenas o esquema S-Adj é produtivo na formação de compostos, mas também, especificamente, o esquema S-eletrônico, possivelmente na base do modelo inicial, o obsoleto *cérebro eletrônico* para a denotação de computadores⁹.

6 Aqui e nas próximas páginas, especificaremos entre parênteses as metonímias mais usuais.

7 Entende-se hiperbólico por “desmesurado” ou “excessivo” (Houaiss 2009). Nas formações de que tratamos nesta seção, o termo se refere ao fato de que há um claro exagero nas denominações de mecanismos eletrônicos, em que o primeiro termo enfatiza o lado humano, implicitamente estabelecendo uma associação que insinua a possibilidade de se substituir (com vantagem) o humano pelo instrumento eletrônico.

8 Croft&Cruse (2004:47).

9 Observe-se que, independentemente dos compostos, o termo *eletrônico* pode também adjetivar substantivos sem que a sequência decorrente seja considerada necessariamente um composto.

Outros exemplos do esquema incluem *porteiro-eletrônico*. Observamos, em ambas as construções, a transferência de domínio, do biológico para o instrumental, e a força expressiva. No primeiro caso, a escolha do termo *cérebro* na denominação tem o objetivo de insinuar que a máquina a ser apresentada ao público seria comparável ao cérebro humano, o que também transparece na expressão *Inteligência Artificial*. No segundo, mais humilde, paralelo a *secretária-eletrônica*, percebemos a insinuação de que o instrumento poderia substituir o indivíduo no exercício da função.

Assim, não constitui surpresa o fato de que a área em que este tipo de construção surge com mais frequência é a do marketing, em que a função hiperbólica se relaciona à necessidade de uma transferência de domínio significativa. O mesmo fenômeno acontece no caso que analisamos a seguir.

Consideremos agora *olho-mágico*, um exemplo também analisado, embora com outros objetivos, em Basilio (2006). O composto *olho-mágico* denota um pequeno orifício em portas de apartamentos com o propósito de possibilitar a quem está dentro ver quem está batendo a campainha sem se expor ou ser visto. Neste composto, temos uma transferência do domínio do corpo para o domínio da arquitetura domiciliar, através da semelhança funcional: em ambos os casos, uma abertura propicia a visão. O fato de que o mecanismo arquitetônico torna a pessoa que está dentro invisível para quem está do lado de fora é qualificado no composto pelo adjetivo “mágico”. Observe-se que, do mesmo modo que no exemplo anterior, o adjetivo incide sobre o substantivo, mas não o qualifica, antes o transforma num objeto em outro domínio cognitivo. No caso, o olho natural, que é fonte da visão, dá lugar à abertura mágica que, ao contrário do olho comum, permite a visão de quem está além da porta, mas mantendo a invisibilidade do usuário de ambos os olhos, o natural e o metonímico-metafórico, já que também se transfere ao mecanismo, metonimicamente, a visão de quem dele fizer uso.

Neste caso a metonímia está envolvida com a metáfora, para além do fato geral, anteriormente mencionado, de que a metáfora impõe ao domínio alvo apenas alguns aspectos do potencial evocativo do domínio-fonte.

Mais especificamente, em *olho-mágico* temos as noções de uma abertura e visão transferidas do domínio físico, o olho, para o domínio arquitetônico, o mecanismo na porta. Mas ninguém pensa em cores, conexões cerebrais, lágrimas e assim por diante em conexão com o pequeno orifício nas portas dos apartamentos. Ou seja, a metáfora consiste na transferência de alguma zona ativa do elemento-fonte a ser incorporada na interpretação do elemento alvo. O papel do adjetivo é menos direto que no primeiro exemplo, visto que há uma atribuição do caráter mágico ao fato da invisibilidade, o que é certamente o caso da capa de Harry Potter, mas puramente hiperbólico e de fins comerciais no caso do mecanismo nas portas.

É importante observar que o mesmo fenômeno acontece nos dois compostos: tanto em *secretária-eletrônica* quando em *olho-mágico* a metáfora resulta do efeito do adjetivo sobre o substantivo; é este efeito que concretiza a transferência do todo resultante para um domínio distinto do domínio do substantivo, constituindo-se um composto semanticamente exocêntrico, que, no entanto, mantém a

concordância flexional característica dos compostos endocêntricos. Em *secretária-eletrônica*, o efeito do adjetivo sobre o substantivo não o modifica no sentido habitualmente utilizado para a ação de um adjetivo sobre um substantivo; na verdade, a natureza da mudança é radical, de agente humano para instrumento eletrônico. No segundo exemplo, ainda que *mágico* seja mais expressivo que objetivo, o que se sugere como mágico é o mecanismo. Em ambos os casos, naturalmente, a mensagem subjacente é que mecanismos eletrônicos ou arquitetados são funcionalmente intercambiáveis (ou melhores) que o correlato físico ou humano.

Estes compostos, em que o entranhamento entre metonímia e metáfora é tão diferente, são ambos instâncias de metáfora deliberada. Na verdade, mesmo em casos convencionais, como o são todos os itens lexicais, eu diria que eles o são no momento de sua formação. E, naturalmente, tanto a propaganda política quanto a comercial fazem constante uso da metáfora deliberada para seus propósitos de persuasão. No caso, ambos os exemplos apresentam uma motivação de caráter comercial.

Vejam agora o composto *retrato-falado*, que denota uma imagem de alguém, feita a partir de um relato descritivo, em geral por testemunhas numa situação de investigação policial. No composto, não há metonímia relevante em *retrato*: a palavra evoca a noção de representação de um rosto, mas não evoca o rosto, antes evoca sua representação. O aspecto peculiar deste composto é o uso do adjetivo *falado* para modificar *retrato*: a combinação provoca, uma certa reação de desconforto, quase de incompatibilidade, à primeira vista.

De fato, a metáfora no composto está na transferência do domínio visual para o domínio linguisticamente relatado no processo de construção da imagem: *falado* se refere a uma fonte de dados linguística e não visual para a composição da face a ser representada, em contradição com nosso conhecimento enciclopédico de como são feitos os retratos. Isto é, o adjetivo retira *retrato* do domínio da fotografia e, ainda mais crucialmente, coloca-o fora do domínio do rosto reproduzido visualmente, o que dá força expressiva ao composto. Novamente, vemos a situação do adjetivo que, ainda que incidindo sobre o substantivo, não o modifica, propriamente, mas, ao contrário, o descaracteriza, transportando-o para um novo domínio. Esta transferência, no entanto, é parcial, como nos demais casos aqui estudados, em que a semântica dos componentes é sempre fundamental, como background evocativo em relação ao significado emergente. Em *retrato-falado*, saímos do domínio habitual de representação facial em termos de fontes visuais para o domínio da fonte falada; mas o fato da representação não estabelece propriamente uma suposta contradição como vemos nos exemplos “alleged affair” e “fake gun”, que, segundo Grady, Oakley & Coulson (1997:121) correspondem a construções privativas.¹⁰

Um exemplo adicional de composto N-Adj é *braço-direito*. Na sequência *braço direito*, *direito* especifica *braço*, de modo que a sequência se refere ao braço que temos do lado direito. No composto, entretanto, o conhecimento enciclopédico do papel fundamental de nosso braço direito em atividades físicas motiva a metáfora que transfere a relevância física do braço direito para os mais variados tipos de ação para a esfera das relações de trabalho, com base na semelhança funcional. Há também neste caso uma dupla metonímia envolvida, juntamente com a metáfora, na medida em que o composto

10 V. Gibbs, R. & Steen, G. (1997)

braço-direito metonimicamente denota uma pessoa (*Todo pela Parte*) e a pessoa é referida em termos de um papel social (*Papel social por Indivíduo*).

Passamos agora ao composto *batata-quente*. Em *batata-quente*, temos uma construção que indica algo difícil, de que alguém pretende se livrar o mais rápido possível. Nesta construção, o adjetivo qualifica o substantivo, sendo os efeitos do *todo* denotado por *batata-quente* os motivadores para a evocação principal da expressão. Neste caso, temos, portanto, uma situação de metonímia combinada com metáfora: nomeia-se a situação difícil de que queremos nos livrar à semelhança de algo que nos queima (*Propriedade por Categoria?*) É possível que a construção possa ser melhor entendida no contexto de “ter nas mãos uma batata-quente”, expressão por si metafórica; neste caso, a zona ativa é a da condição, embora a possibilidade de jogar-se uma batata para outra pessoa ajude a motivar o primeiro elemento.

O último exemplo que analisaremos nesta seção é *patinho-feio*. No Dicionário UNESP do Português Contemporâneo, ou Borba (2004), *patinho feio* é definido como “quem é desprezado; enjeitado. O exemplo fornecido é “Sempre fomos tratadas como o patinho feio”.

A expressão deriva do personagem-título do conto de Hans Christian Andersen, “O patinho feio”. No caso, temos uma transferência do domínio literário, digamos, já que o patinho feio é um personagem fictício, para o domínio do mundo real. Na história infantil, o patinho é considerado feio junto aos outros patos e em vários lugares em seu percurso, sendo reabilitado apenas no final, quando se descobre que ele era, de fato, um cisne. Mas a transferência é feita a partir da zona ativa negativa, de modo que *patinho feio* corresponde a quem é julgado feio e vítima de maus-tratos. Neste caso, temos um composto de significado óbvio na superfície, mas o significado pleno só pode ser atingido pelo conhecimento enciclopédico do conteúdo do conto.

Resumindo, analisamos nesta primeira parte do trabalho os compostos *retrato-falado*, *secretária-eletrônica*, *olho-mágico*, *braço-direito*, *batata-quente* e *patinho-feio*. Todos os exemplos envolvem metáfora, isto é, transferência de um domínio fonte para um domínio alvo. Em *retrato-falado*, os dados básicos para a reprodução de uma face são transferidos do domínio visual para o linguístico; em *olho-mágico*, o adjetivo incide sobre o substantivo *olho*, assim transferindo o olho físico para um mecanismo supostamente mágico de visão com invisibilidade. Em *secretária-eletrônica*, o adjetivo realiza a transferência de um rótulo ocupacional do domínio humano para o domínio dos instrumentos eletrônicos; em *braço-direito*, a metáfora corresponde à transferência da sequência *braço direito* do domínio do corpo para o domínio da articulação de funcionalidade, confiança e dependência em relações de trabalho entre pessoas; em *batata-quente*, temos uma transferência do domínio físico (desconforto por excesso de calor) para o psicológico (desconforto pela dificuldade do problema), constituindo a metáfora, que se mescla à metonímia representada pela zona ativa (quente); e, finalmente, em *patinho-feio*, temos uma projeção do espaço da ficção para o espaço da vida real.

Existe uma diferença entre os exemplos, configurando dois tipos semanticamente diferentes de construção em compostos metafóricos exocêntricos S-Adj: no primeiro, o efeito do adjetivo sobre o

substantivo cria a metáfora que faz o composto, como em *secretária-eletrônica*, *olho-mágico* e *retrato-falado*; no outro tipo, a modificação objetiva do substantivo pelo adjetivo forma uma expressão denotativa que é fonte da metáfora que cria o composto, como em *braço-direito*, *batata-quente* e *patinho-feio*.

Em todos os casos, no entanto, a transferência metafórica vem do mapeamento da construção S-Adj como um todo para um domínio distinto.

3.1.2 COMPOSTOS METAFÓRICOS ENDOCÊNTRICOS

Os casos que analisamos até agora são de compostos metafóricos exocêntricos. Entretanto, existem também compostos metafóricos semanticamente endocêntricos, isto é, compostos que correspondem a uma instância do que é denotado no substantivo; nestes casos, a metáfora se origina ou se situa no adjetivo. Encontramos em nossos corpus o caso de *baleia-assassina*.

Em *baleia-assassina*, temos, por um lado, um composto classificatório, dado que se trata de um tipo de baleia. Entretanto, o caráter expressivo do adjetivo “assassina” pode ser interpretado como uma transferência metafórica do domínio humano para o domínio animal, uma vez que o conceito de assassinato não se aplica normalmente a animais que podem matar, e menos ainda como um instrumento de classificação.

3.2 COMPOSTOS METONÍMICOS

Do mesmo modo que nos metafóricos, a situação mais frequente que encontramos em compostos metonímicos é a de compostos semanticamente exocêntricos, de que nos ocuparemos na próxima seção. Entretanto, também existem compostos metonímicos endocêntricos, que serão tratados na seção subsequente.

3.2.1 COMPOSTOS METONÍMICOS EXOCÊNTRICOS

Começaremos a análise de compostos metonímicos exocêntricos com o composto *ferro-velho*, nome que damos a sucata, ou, mais especificamente, referência genérica a objetos de base ferrosa em processo de deterioração ou de pouco valor. Neste composto, *velho* incide sobre *ferro* e o qualifica, mas o sentido mais comum, pelo menos em textos jornalísticos, é o metonímico, que corresponde ao local em que se negocia (muitas vezes, ilegalmente) com este tipo de material. Temos, portanto, a metonímia *Produto negociado por Local*.

Outro exemplo de composto em que a construção se refere ao local é *pé-sujo*, denominação dada a determinados tipos de botequim de caráter popular e informal no Rio de Janeiro. Neste caso, a metonímia é dupla, sendo que *pé-sujo* qualifica o tipo de frequentador do botequim, pela metonímia *Propriedade por Categoria*, e o frequentador, então, provê a denominação do tipo de botequim, pela metonímia *Local por Indivíduo*.

Compostos metonímicos que designam indivíduos por propriedades ou condições são também frequentes, desde os tradicionais como *bóia-fria*, que designa trabalhadores do campo pelo seu tipo de refeição e *pé-rapado*, que designa indivíduos pobres pela condição de seus pés, até o recente *ficha-suja*, que designa políticos que não podem apresentar um atestado de bons antecedentes e, portanto, não são considerados elegíveis. Ainda no campo profissional, *relações-públicas* faz referência aos profissionais que exercem esta ocupação pela metonímia *Ação por Indivíduo*.

Outro exemplo de composto metonímico, consolidado mas de referência a uma ocasião específica, é *cara-pintada*, composto que denominou os estudantes participantes das passeatas que levaram ao impeachment do então presidente Collor na década de noventa. Na época, os estudantes em passeatas de protesto pintaram o rosto de verde e amarelo, motivação para a denominação *cara-pintada* para os participantes das passeatas. Temos, então, uma variante da metonímia *Propriedade por Categoria*, na medida em que se trata de uma condição ocasional e não permanente como os tradicionais *cara-pálida* ou *pele-vermelha* dos filmes de faroeste americanos.

Na verdade, a metonímia é um processo largamente utilizado na formação de palavras em casos de conversão adjetivo/substantivo, nos quais o adjetivo é fonte para um substantivo que denota um indivíduo pela propriedade expressa pelo adjetivo. Alguns exemplos são *cego*, *careca*, *pobre*, etc.. Como se vê, o mesmo mecanismo cognitivo geral está em jogo nas construções compostas do parágrafo anterior.

Sintetizando nossos resultados até agora, vimos que existe um contínuo no nível de composicionalidade na formação de compostos S-Adj em português, desde a expressão que pode ser formada no discurso, embora haja consolidação na evocação de determinados conceitos - por exemplo, a sequência *batata frita* evoca palitos de tamanho convencional ou rodela de batata fritas, mas não uma batata inteira¹¹ – até os compostos rigidamente consolidados, como *pé-rapado*.

3.2.2 COMPOSTOS METONÍMICOS ENDOCÊNTRICOS

Conforme observamos anteriormente, compostos metonímicos e metafóricos são, em geral, semanticamente exocêntricos. Entretanto, podemos encontrar alguns compostos metonímicos semanticamente endocêntricos, isto é, compostos em que o todo constitui uma instância do que é denotado no primeiro elemento, sendo a metonímia originária do segundo elemento. Em nosso corpus encontramos um exemplo, o composto *casa-grande*.

Em *casa-grande*, denominação para a casa dos donos de uma fazenda na época da escravidão, em oposição à senzala, lugar reservado aos escravos, temos um composto parcialmente composicional, na medida em que as casas dos donos de fazenda eram, por óbvio, grandes; e, por outro lado, adjetivos como *grande* e *pequeno* são relativos. No caso, temos uma metonímia *Tamanho por Qualidade*, sendo implícito o conhecimento enciclopédico sobre a questão do tamanho na avaliação da qualidade

11 Do mesmo modo, banana frita pressupõe o fatiamento e cenoura frita provoca estranheza, o que indica que o conhecimento enciclopédico interfere não apenas no entendimento das formações, mas na maior ou menor aceitação das construções.

de locais de moradia. Embora *casa-grande* como denominação não se oponha a *casa pequena*, mas a senzala, no nível lexical, temos a manutenção da metonímia conceptual na atualidade em pares opostos como *apartamento/apartamentinho* e *casarão/casinha/casebre*, em que o primeiro item é associado a riqueza, fausto ou importância social, em oposição ao segundo, que evoca simplicidade e/ou dificuldade financeira.

3.3 OUTROS EXEMPLOS

Nesta última parte do trabalho, vamos analisar alguns outros compostos de nosso corpus, com o objetivo de ilustrar a dificuldade de estabelecer determinados grupos preferenciais na composição semântica e verificar se, em termos objetivos, a noção de contínuo deveria ser entendida em termos absolutos. Os compostos a serem aqui analisados são: *chapa-branca*, *ato-falho*, *cinema-mudo*, *Ano Novo* e *medida provisória*.

No composto *chapa-branca*, temos, como em muitos outros casos, duas construções compostas, envolvendo metonímia e metáfora. Primeiramente, *chapa-branca* é a denominação objetiva e classificatória da “placa de cor branca de um veículo do serviço público oficial” (Borba 2004). Assim, *chapa-branca* metonimicamente aponta para carros do serviço público oficial. Já o uso e entendimento de *chapa-branca* como uma característica de algo é de caráter metonímico e pejorativa, indicando fatos, propostas, etc. como oriundos de indivíduos ou grupos envolvidos com interesses específicos em diferentes esferas do poder. Em outras palavras, a expressão metonímica referente a carros oficiais sai da esfera automobilística concreta para o domínio da qualificação de esferas do poder.

Passamos à análise de *ato-falho*, termo da psicanálise que passou a ser usado na linguagem cotidiana. O ato falho pode ser definido como um erro na fala supostamente causado pelo inconsciente. Em geral, o ato falho poderia ser considerado como o chamado “lapsus linguae”, ou seja, um lapso ou troca de palavras que muitas vezes acontece, e geralmente envolve a semelhança linguística. Na construção, o adjetivo incide sobre ato, caracterizando-o em termos da falha, ou seja, da emissão diferente do que se pretendia. Na interpretação psicanalítica, o ato-falho corresponde à projeção da entidade conceptual fonte do domínio do inconsciente para o domínio dos atos de fala. Assim, um ato-falho concreto poderia ser considerado como uma metáfora não-intencional ou, mais frequentemente, definitivamente indesejada, já que desvela o que é reprimido. A construção, entretanto, não é uma metáfora do ponto de vista linguístico: trata-se apenas de um composto em que na sequência *ato falho* o adjetivo *falho* caracteriza um certo tipo de ato através de um certo tipo de falha. Assim, na expressão de caráter terminológico, existe apenas a fixação das respectivas zonas ativas nos termos “ato” e “falho”, dentre as inúmeras possibilidades de evocação de cada um dos termos.

Podemos considerar o mesmo tipo de transferência de domínio (humano para não-humano) no caso de *cinema-mudo*, no pressuposto de que *mudo* é um adjetivo que caracteriza pessoas pela impossibilidade de falar. Tendo em vista que a linguagem é uma propriedade dos seres humanos, caracterizar algo como mudo poderia corresponder a um uso metafórico do adjetivo. Entretanto, no Dicionário da Língua

Portuguesa Contemporânea, *cinema mudo* se opõe a cinema falado, sendo o primeiro caracterizado pela ausência de som interno à projeção; e um dos sentidos do adjetivo *mudo* se refere apenas à ausência de som, caso em que temos em *cinema mudo* apenas um composto classificatório.

Passamos a *Ano Novo*, expressão que se refere a um novo ano que começa no dia 1º de Janeiro, após 31 de Dezembro do ano anterior. Trata-se de uma expressão solidificada para referência ao novo ano que se inicia, mas o adjetivo, incidindo sobre o substantivo, não é de caráter classificatório em termos de tipo de ano; e não há metáfora óbvia envolvida. Observa-se, no entanto, que a expressão poderia ser considerada como um composto, de acordo com os critérios tradicionais, na medida em que os termos não podem ser substituídos, os termos não podem ser coordenados com outros e não há possibilidade de interferência.

A expressão *medida provisória*, de formação mais recente e grande frequência de uso, de certo modo surgiu para substituir o decreto-lei; trata-se de um projeto de lei que pode entrar em vigor de imediato, mas depende da aprovação do Congresso para sua manutenção. Como muitos outros casos, a expressão é transparente em seus termos, mas difere das possibilidades de interpretação que teríamos na sequência análoga não lexicalizada, conforme ilustrado em frases como “Estas medidas provisórias serão confirmadas quando o coordenador chegar” ou “tomei suas medidas, mas são medidas provisórias, porque você pode engordar ou emagrecer até o casamento”. Como muitos outros termos, *medida provisória* apresenta rigidez de construção e impossibilidade de elementos interferentes, em virtude de sua oficialização como denominação de um determinado procedimento institucional. Isto, naturalmente, levanta a questão dos compostos do ponto de vista enciclopédico, a qual abarca, aliás, também a questão dos nomes próprios, a saber, a natureza dos critérios de distinção de construções compostas.

Nos exemplos arrolados nesta última seção, o leitor deve ter notado que não há uma sistematização. Os exemplos foram incluídos para ilustrar a dificuldade tanto da decisão de se considerar uma sequência S-Adj como composto quanto da discussão em relação à metáfora e à metonímia, muitas vezes emaranhadas em compostos, situação em geral denominada de “metaftonímia” (Goossens 2003), além da questão relativa ao caráter endocêntrico ou exocêntrico do composto, conforme se privilegiem os aspectos sintáticos ou semânticos da construção lexical envolvida.

Assim, ao concluir esta análise preliminar da situação dos compostos numa abordagem cognitiva, observamos que, dentre as inúmeras questões levantadas nos breves comentários aqui apresentados, ressaltam na relevância e na urgência a questão do emaranhamento de metáfora e metonímia, quer nos seus conceitos, quer em sua aplicação na análise de construções; a questão da pertinência ou não de considerarmos compostos lexicais como entidades relevantes de demarcação no contínuo léxico-gramática e, sobretudo, a questão da natureza semântica dos processos de composição.

REFERÊNCIAS

BARCELONA, A. Clarifying and applying the notions of metaphor and metonymy within cognitive linguistics: an update. In: DIRVEN, R. e PÖRING, R. *Metaphor and Metonymy in Comparison and Contrast*. Berlin: Mouton de Gruyter, 2003.

BASILIO, M. Em torno da palavra como unidade lexical: palavras e composições. *Veredas*, v.4, n.2. Ed. UFJF, Juiz de Fora, 2000.

BASILIO, M. Construções Nominais no Léxico e na Fala: um estudo de construções envolvendo S e Adj no Português do Brasil. In: SELLA, A.F., CORBARI, C.C. e BIDARRA, J. (orgs) *Pesquisas sobre Léxico: Reflexões Teóricas e Aplicação*. Campinas: Pontes, 2012.

BASILIO, M. Metaphor and Metonymy in Word-Formation. In: ZANOTTO, M.S., NARDI, M.I. e VEREZA, S. *Essays on Metaphor in Language and Thought*. D.E.L.T.A., v.22 – Especial, 2006.

BENEZES, R. Putting the notion of “domain” back into metonymy: Evidence from compounds. In: BENCZES, R., BARCELONA, A. e IBÁÑEZ, F. *Defining Metonymy in Cognitive Linguistics*. Amsterdam/Philadelphia: John Benjamins, 2011.

BOOIJ, G. *The Grammar of Words*. Oxford: Oxford University Press, 2007.

BORBA, F.S. (org) *Dicionário UNESP do Português Contemporâneo*. São Paulo: Editora UNESP, 2004.

CASTELEIRO, J.M. *Dicionário da Língua Portuguesa Contemporânea*. Academia das Ciências de Lisboa. Lisboa: Editorial Verbo, 2001.

CROFT, W. e CRUSE, D.A. *Cognitive Linguistics*. Cambridge: Cambridge University Press, 2004.

DRESSLER, Wolfgang et al. (eds) *Morphology and its Demarcations*. Amsterdam/Philadelphia: John Benjamins, 2005.

FABB, N. Compounding. In: Spencer, A. e Zwicky, A. (eds.) *The Handbook of Morphology*. Oxford: Blackwell, 1998.

GIBBS Jr, R.W. e STEEN, G.J. (eds.) *Metaphor in Cognitive Linguistics*. Amsterdam/Philadelphia: John Benjamins, 2001.

GOOSSENS, L. Metaphonymy: the interaction of metaphor and metonymy in expressions for

linguistic action. In: DIRVEN, R. e PÖRING, R. (eds.) *Metaphor and Metonymy in Comparison and Contrast*. Berlin: Mouton de Gruyter, 2003.

LANGACKER, R. *Grammar and Conceptualization*. Berlin: Mouton de Gruyter, 2000.

LANGACKER, R. *Concept, Image and Symbol – The Cognitive Basis of Grammar*. Berlin: Mouton de Gruyter, 2002.

LANGACKER, R. Metonymic Grammar. In: PANTHER K., THORNBURG L. e BARCELONA A. (eds) *Metonymy and Metaphor in Grammar*. Amsterdam/Philadelphia: John Benjamins, 2009.

LIEBER, R. *Introducing Morphology*. Cambridge: Cambridge University Press, 2010.

LIEBER, R. e STEKAUER, P.(eds) *The Oxford Handbook of Compounding*. Oxford, Oxford University Press, 2011.

RADDEN, G. e KOVECSES, Z. Toward a Theory of Metonymy. In: PANTHER, K. e RADDEN, G. (eds) *Metonymy in Language and Thought*. Amsterdam/Philadelphia, John Benjamins, 1999.

RIO-TORTO, G. et al. *Gramática Derivacional do Português*. Coimbra, Imprensa da Universidade de Coimbra, 2013.

STEEN, G. When is metaphor deliberate? In: Johannesson, N. e Minugh, D. (eds) *Selected Papers from the 2008 Stockholm Metaphor Festival*. Stockholm, Stockholm University, 2013.

Recebido em 31/01/2016

Aceito em 12/04/2016

UMA ABORDAGEM CONSTRUCIONAL PARA AS FORMAÇÕES X-DROMO DO PORTUGUÊS BRASILEIRO

Carlos Alexandre Gonçalves (UFRJ)¹ e José Augusto de Oliveira Pires (UFRJ)²

RESUMO

O presente trabalho tem por objetivo analisar o comportamento do formativo *-dromo* ('bumbódromo', 'sambódromo', 'camelódromo', 'fumódromo') no português brasileiro contemporâneo tendo por base o modelo de Morfologia Construcional proposto por Booij (2005, 2010). Levando-se em consideração que a literatura tradicional não fornece um estudo satisfatório e aprofundado acerca do formativo em questão, o intuito é (a) realizar uma análise mais detalhada sobre as transformações e os desdobramentos pelos quais o formativo passou no português e (b) investigar a produtividade do mesmo no atual estágio da língua a partir de um modelo que possibilite uma descrição mais adequada tanto do polo formal quanto do polo semântico de construções morfológicas.

PALAVRAS-CHAVE: Gramática das Construções. Morfologia. Formações *X-dromo*.

ABSTRACT

This paper aims to analyze the behavior of the formative *-dromo* ('bumbódromo', 'sambódromo', 'camelódromo', 'fumódromo') in contemporary Brazilian Portuguese, based on the model of constructional morphology (BOOIJ, 2005, 2010). Taking into account that the traditional literature does not have a satisfactory study about this morpheme, the intent of this work is (a) to perform a more detailed analysis of the changes and developments that the formative undergoes in Brazilian Portuguese and (b) to investigate the productivity of these formations in the current language from a model that will enable a more accurate description of both the formal and semantics poles of morphological constructions.

KEYWORDS: Construction Grammar. Morphology. *X-dromo* formations.

1 Professor Associado IV do Departamento de Letras Vernáculas da UFRJ, Doutor em Linguística e Pesquisador do CNPq e da FAPERJ. E-mail: carlexandre@bol.com.br.

2 Doutorando em Letras Vernáculas e professor substituto do Colégio Pedro II. E-mail: augustopires2010@gmail.com.

1. CONSIDERAÇÕES INICIAIS

O presente trabalho analisa o comportamento das formações *X-dromo* ('autódromo', 'sambódromo', 'fumódromo') no português brasileiro contemporâneo, tendo por base o modelo de Morfologia Construcional desenvolvido por Booij em vários artigos (2005, 2007, 2009), até culminar no livro *Construction Morphology* (BOOIJ, 2010). Considerando que a literatura tradicional não possui um estudo satisfatório e aprofundado acerca do formativo em questão, nosso intuito é (a) realizar uma análise mais detalhada sobre as transformações e os desdobramentos pelos quais o formativo passou no português; (b) checar a produtividade do mesmo nos dias de hoje; (c) observar as diferenças entre as construções mais antigas em relação às mais novas; e (d) conferir se as formações ainda devem ser consideradas casos de composição neoclássica. O objetivo maior do texto é representar a relação entre os polos formal e semântico das construções *X-dromo*, com base em Booij (2010), a fim de conferir o verdadeiro estatuto do elemento à direita, se radical ou sufixo.

Para alcançar tais metas, dividamos o trabalho da seguinte maneira: a primeira parte é destinada à revisão da literatura sobre *-dromo*. Nessa seção, o objetivo é verificar a origem dos dados, assim como averiguar de que maneira o formativo é tratado nas mais variadas abordagens: (a) dicionários etimológicos, como Cunha (1986), e eletrônicos, a exemplo de Aurélio (2004) e Houaiss (2009); (b) algumas gramáticas tradicionais – Bechara (2004), Cunha & Cintra (2008); e (c) manuais de morfologia do português – Basílio (2007), Laroca (2005) e Sandmann (1987, 1988).

A etapa seguinte se destina à abordagem da Morfologia Construcional, de Geert Booij. O principal propósito da seção é apresentar o arcabouço teórico, observando as principais motivações para a utilização do mesmo na análise do formativo em questão. Tendo em vista que, para Booij (2005), as fronteiras entre composição e derivação não são claramente delimitadas (ao contrário, são maleáveis de ambos os lados, havendo, na verdade, um *continuum*), essa perspectiva é de extrema importância para indagar qual o verdadeiro estatuto do formativo *-dromo* no processo de formação de palavras no atual estágio da língua: se ainda se comporta como radical neoclássico, se constitui sufixo ou se constitui elemento com propriedades desses dois principais tipos morfológicos.

A terceira parte é destinada à aplicação da Morfologia Construcional (doravante MC) às formações terminadas em *-dromo*. Com o instrumental da MC, procuramos demonstrar (a) como a fronteira entre composição e derivação é questionável, uma vez que há a possibilidade da transformação do *status* morfológico de elementos ao longo dos anos, (b) a possibilidade de criação de padrões construcionais que, por sua vez, colaboram para a alta produtividade do até então chamado “radical erudito” (neoclássico), hoje um neossufixo. Nessa transição, as construções se apresentam, em sua borda direita, com uma nova formatação na língua: *-ódromo*. Assumimos, portanto, que houve tanto mudança formal quanto semântica nas construções, que passaram de compostas e opacas (*X-dromo*, em que o elemento à esquerda não constitui palavra) a derivadas e transparentes (*X-ódromo*, com X recebendo, agora, etiqueta lexical e se atualizando como palavra). É o que se observa, por exemplo, no

confronto entre ‘antídromo’ (“que se enrola em direção oposta à normal) e ‘masturbódromo’ (“local utilizado para a prática da masturbação”).

Os dados que compõem o *corpus* foram recolhidos nas mais variadas situações de uso da língua, durante o ano de 2015: conversas espontâneas em que nos envolvemos ou presenciamos, aulas, conferências e palestras, programas de TV e rádio (sobretudo entrevistas), jornais e revistas de circulação nacional, *outdoors*, nomes de lojas e produtos comerciais (oniônimos). Além disso, um rastreamento mais sistemático foi implementado por intermédio da ferramenta de busca eletrônica *Google*, bem como de *sites* de redes sociais, tais quais *Facebook* e *Twitter*. O *corpus* contém, hoje, cerca de 150 formações.

2. REFERÊNCIAS A *-DROMO* NA LITERATURA SOBRE O PORTUGUÊS

Dicionários e gramáticas tradicionais classificam *-dromo* como radical erudito que participa do processo de composição; por sua vez, em alguns manuais de morfologia, já é possível notar outra classificação para o formativo. Abordemos, em primeiro lugar, as definições mais usuais para, então, apresentar as conceituações mais recentes.

Ao procurarmos a definição no dicionário etimológico de Cunha (1986), constatamos que *-dromo* é analisado como um *elemento de composição*, sendo oriundo do grego **drómos**, “ação de correr, corrida, lugar de corrida”. Como exemplo de formação que não veicula tal sentido, tem-se ‘craspedódromo’, adjetivo relacionado à área da botânica, cujo significado é um qualificativo para formas “em que as nervuras secundárias são secantes em relação à margem da folha (diz-se de nervação)” (CUNHA, 1986: 441).

Em se tratando dos dicionários eletrônicos consultados, no *Houaiss* (2009) constatamos que a definição é praticamente a mesma: “elemento de composição; pospositivo do grego **drómos**, ou “ação de correr, lugar para corrida, corrida” (HOUAISS, 2009: 561), com ‘aeródromo’ sendo definido como “espaço delimitado, em terra, provido de relativa infraestrutura para o pouso e decolagem de aeronaves destinadas ao transporte de passageiros ou de cargas diversas” (HOUAISS, 2009: 562).

Muito embora o dicionário *Aurélio* (2004) traga definições um pouco diferentes, podemos verificar que o conceito central não fica muito distante dos já mencionados; a saber: “elemento de composição; ação de correr’, ‘corrida’; ‘deslocamento rápido em veículo’; ‘lugar para correr’; ‘pista’; ‘local de’; ‘fluxo’, ‘corrente’” (AURÉLIO, 2004: 345). O exemplo apresentado na obra é ‘autódromo’, “conjunto de pistas e edifícios (instalações para administração, arquibancadas, controle, oficinas de reparos, etc.), para corrida de automóveis” (AURÉLIO, 2004: 345). Vale a ressalva de que no dicionário *Aurélio* já existe a acepção *lugar de* isoladamente, indicando uma modificação na definição inicial.

Nas gramáticas tradicionais, também podemos perceber que as conceituações são basicamente as mesmas. Em Cunha & Cintra (2008: 125), *-dromo* é definido como “lugar para correr” e exemplificado com o vocábulo ‘velódromo’, “local destinado a corridas ciclísticas, dotado de pistas, instalações

para o público etc” (CUNHA & CINTRA, 2008: 126). Além disso, os autores afirmam que *-dromo* “funciona, preferentemente, como segundo elemento da composição” (CUNHA & CINTRA, 2008:126), como se observa em ‘hipódromo’.

Por sua vez, Sandmann (1988) descreve o formativo não mais como elemento de composição e sequer como *-dromo*, mas como elemento de derivação e *-ódromo*. Justifica tal transição relacionando-o a um sufixo em razão de “se prestar à formação de novas palavras em série e porque não ocorre livremente na frase” (SANDMANN, 1988: 47). Com essa visão diferenciada, o autor pondera que o significado inicial –“ação de correr, corrida, lugar de corrida” – não pode ser relacionado diretamente com o significado atual das novas formações. Tal é o caso de ‘camelódromo’, que se refere, no Município do Rio de Janeiro (Centro da Cidade), a um lugar de concentração de comércio popular, e ‘amoródromo’, “lugar para a prática do amor livre”. Nessas novas formações, a ideia de locativo se mantém, sem, entretanto, fazer referência à corrida.

Laroca (2005) aborda o elemento *-dromo* como um *neossufixo*, isto é, um novo elemento derivacional do português. Mantém, todavia a estrutura *-dromo*, diferentemente de Sandmann, que já incorpora a vogal <ó> ao formativo. À semelhança do autor, Laroca também constata a transformação no sentido: de “curso, corrida, marcha, condutibilidade, passou a designar o local (apropriado) para acontecer determinado fato ou evento” (LAROCA, 2005: 75), como em ‘namoródromo’, ‘papódromo’ e ‘beijódromo’.

Em um *corpus* contendo, aproximadamente, 120 palavras, Pires (2014) observa que as encontradas em dicionários eletrônicos somam em torno de 30. Por sua vez, a outra parte advém justamente da ferramenta de busca eletrônica *Google* e das redes sociais. O autor constata que estas respondem pela maior parte do *corpus*, constituído, majoritariamente, de palavras (ainda) não dicionarizadas. Os vocábulos terminados com o formativo aparecem em situações reais de uso, em que a necessidade e a inventividade dos falantes permitem a criação sistemática de novas palavras.

Como se vê, há divergências no que diz respeito ao entendimento desse “radical erudito”. Além disso, vale ressaltar que as palavras dicionarizadas e encontradas nas gramáticas tradicionais são as mais antigas na língua: são vocábulos que datam dos séculos XVII, XVIII, XIX e do começo do século XX. Com isso, há uma maior tendência para que sejam vinculadas ao processo de composição. Já os exemplos dos manuais de morfologia do português (LAROCA, 2005) e análises mais recentes (SANDMANN, 1988; PIRES, 2014) são, basicamente, da segunda metade do século XX e início do século XXI. As formas obtidas a partir de situações conversacionais variadas, da ferramenta eletrônica de busca *Google* e das redes sociais (*Facebook*, *Twitter*) evidenciam a possibilidade de modificação tanto do estatuto morfológico de *-dromo* quanto de seus usos e funções.

3. A MORFOLOGIA CONSTRUCIONAL

A expressão “Morfologia Construcional” não é nova na área. Em 1987, Corbin propõe um modelo de

análise morfológica que denomina de construcional. Essa abordagem, no entanto, diferentemente da de Booij (2005), está inserida num quadro linguístico gerativista de inflexão lexicalista. Diferentemente da proposta de Corbin, o modelo de Booij (2005, 2007, 2010) se inscreve no paradigma da Linguística Cognitiva e adapta a abordagem construcionista de autores como Goldberg (1995), voltados para a sintaxe, à descrição de fatos morfológicos.

Em Basilio (2011: 02), entende-se léxico como um espaço de formas simbólicas, isto é, formas que se associam a conceitos. Para ela, as unidades lexicais, “cujas possibilidades de evocação são infinitas, dependendo de circunstância, podem envolver desde a história da língua e a história dos falantes (...) até relações entre formas e suas potenciais evocações”. Convergiendo com essa definição, Booij (2005, 2007) entende que formas morfológicamente complexas são esquemas que propiciam diversas instanciações, pois, tal como Basilio (2011), abraça a ideia de que (a) itens evocam conceitos e (b) essa evocação é situada linguística e socioculturalmente. As instanciações, no entanto, não serão exatamente iguais, pois são submetidas a ajustes focais e estão sujeitas a outras operações de língua em uso, como seleção, focalização, perspectização e relevância (LANGACKER, 1987).

Inaugurando um paradigma que chama de Morfologia Construcional (*Construction Morphology*), Booij (2005) mostra que as unidades linguísticas são estruturas simbólicas convencionais, não havendo, por isso mesmo, diferença considerável entre palavras derivadas³ (Xy: ‘lingu-ista’, ‘roqueiro’), compostos (XY: ‘seca-pimenteira’, ‘criado-mudo’), expressões semiabertas (pé-de-X_s, “árvore”: ‘pé de coco’, ‘pé de jaca’) e expressões idiomáticas (DAR uma X_v-da, que significa “fazer X_v rapidamente”, como em ‘dei uma piscada’, ‘dava uma passeada’): “todas essas unidades, que são complexas, podem, igualmente, ser analisadas, em suas estruturas de formação, por meio de esquemas construcionais” (GONÇALVES & ALMEIDA, 2014: 110).

Booij (2010) postula que a semelhança estrutural entre composição e derivação pode ser expressa por meio de esquemas de formação de palavras que generalizam conjuntos de formas existentes e podem ser usados para criar novas palavras. Por exemplo, em português, palavras complexas são formadas, muito frequentemente, por meio de composição, sufixação e prefixação. Esses padrões morfológicos podem ser ilustrados pelos seguintes esquemas, que Gonçalves & Almeida (2014) adaptaram para o português:

(01) (a) composição: $[[X]_x [Y]_y]_s$

(b) prefixação: $[X[Y]_y]_y$

(c) sufixação: $[[X]_x Y]_y$

Nesses esquemas, as variáveis X e Y, em maiúsculas, representam sequências fonológicas e os subscritos _x e _y, em minúsculas, categorias lexicais. O esquema geral dos compostos, em (a), expressa a generalização de que a composição, independentemente da etiquetagem lexical de seus constituintes, sempre forma substantivos em português (daí o rótulo _s, após o último colchete). O esquema da

³ Nesses exemplos, X e Y, em maiúsculas, representam formas livres. Por sua vez, y, minúsculo, representa um afixo, forma presa que não receber etiqueta lexical.

prefixação, em (b), expressa que essa operação é neutra categorialmente, sendo a classe gramatical das palavras prefixadas idêntica à de sua base, que constitui a cabeça. Para Booij (2005: 13), “a diferença entre composição e derivação está no fato de, na derivação, um dos constituintes não ter etiqueta lexical, uma vez que não corresponde a uma palavra”. No caso da sufixação (esquema em (c)), no entanto, o elemento preso porta informação sintática e constitui cabeça categorial (SCALISE *et al.*, 2009), por determinar a classe gramatical do produto. Neste texto, defendemos que, além desses três esquemas gerais, o português apresenta um outro: o da composição neoclássica, cuja representação é $[X Y]_S$, a exemplo de $[[\text{cardio}][\text{pata}]_S$, $[[\text{pedó}][\text{filo}]_S$ e $[[\text{orto}][\text{rexia}]_S$.

Booij (2010) mostra que uma abordagem construcional possibilita tratar mais satisfatoriamente a relação entre semântica, sintaxe, morfologia e léxico, observando melhor as semelhanças de formação nos níveis da palavra e da frase. Assim, a Morfologia Construcional constitui enfoque bem mais integrado para a morfologia, pois esquemas morfológicos podem ser interpretados como padrões sintáticos gramaticais ou expressões idiomáticas no nível da palavra, com uma posição fixa (lexicalmente preenchida) e outra aberta(s), representada(s) por variável(is). Tal fato aponta para a constatação de que não há diferença considerável entre expressões perifrásticas (como $[IR_{[pres]} X_V]_{[inf]}$], que expressa a noção de futuro, a exemplo de ‘vou fazer’), e idiomas construcionais (como $[PAGAR X_S-inho]$], “vexame por deixar X (parte do corpo) à mostra”, como em ‘pagar peitinho’, ‘pagar pintinho’)⁴ e palavras morfológicamente complexas, sejam elas derivadas ou compostas⁵.

Ao se considerar a existência de idiomas construcionais, a concepção do léxico como lista de palavras e expressões convencionalizadas e fixas na língua (DI SCIULLO & WILLIAMS, 1987) é posta em xeque. De acordo com Booij (2007), o léxico pode ser estendido como constituído por idiomas construcionais parcialmente especificados que representam o pareamento da estrutura formal com a semântica das palavras complexas. O esquema de que participa o sufixo denominal *-ista*, em (02) a seguir, é abstraído das instanciações em (03) e pode ser interpretado como exemplo de construção idiomática no nível da palavra (uma construção morfológica, portanto): substantivos deverbiais terminados em *-ista* significam, genericamente, “que trabalha com/prática X”, significado esse ligado à instanciação mais especificada do esquema básico da sufixação, $[X]_x Y]_y$:

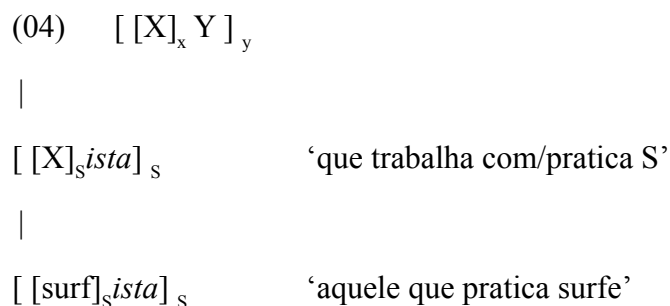
(02) $[X]_S ista]_S$ – ‘que trabalha/prática S’

(03) skate – skatista surfe – surfista dente – dentista
 loja – lojista greve – grevista cópia – copista
 lobby – lobista piano – pianista

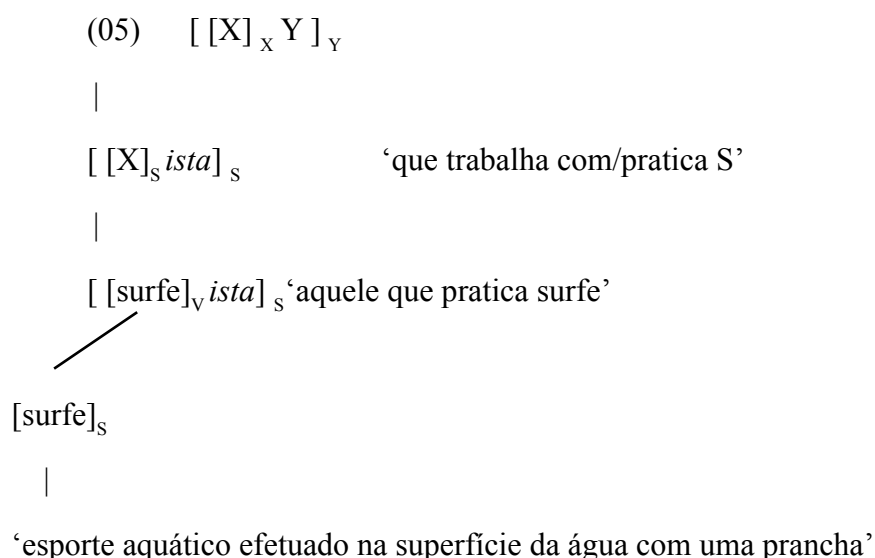
Como na Gramática das Construções, a relação entre o esquema mais abstrato e as instanciações individuais é representada por meio de uma árvore, em que construções mais específicas herdaram propriedades de construções dominantes ou mais gerais. Pode-se exemplificar essa cadeia, em (04), a seguir, com ‘surfista’:

4 O esquema $[PAGAR X_S-inho]$ parece ter origem na expressão idiomática ‘pagar mico’ (“passar por situação vexatória”, “cometer uma gafe”) e concentrou na forma verbal ‘pagar’ o significado de “vexame”, passando a se combinar com diminutivos *X-inho* em referência a uma parte íntima do corpo que, por descuido, ficou descomposta (‘pagar cofrinho’, ‘pagar xaninha’).

5 Essa proposta espelha, para o nível da morfologia, o princípio central da Gramática de Construções: a ideia de que a totalidade do conhecimento linguístico do falante pode ser apreendida de maneira uniforme, por meio de pareamentos convencionais de forma e significado.



Como exemplificado em (04), cada nó inferior da árvore herda propriedades dos nós dominantes e, por isso mesmo, nós mais baixos portam informações redundantes. As construções X_s-ista instanciam o esquema abstrato da sufixação ($[[X]_x Y]_y$), já que o elemento à direita é um afixo e, portanto, não deve ser listado no léxico por não constituir forma livre na língua. ‘Surfista’, por sua vez, herda do nó imediatamente dominante a semântica de agente. É importante ressaltar, ainda, que ‘surfista’ também herda propriedades de sua base, também ligada à árvore⁶:



A segunda linha das árvores em (04) e (05) generaliza a estrutura e o significado dos substantivos denominais em *-ista*. Novas palavras podem ser criadas por meio do que Booij (2010) denomina “unificação” de um esquema com um item lexical⁷. Gonçalves & Almeida (2014) consideram mais interessante nomear esse mecanismo de “compatibilização”. Nesse caso, a ideia é que o item combine suas propriedades lexicais com as propriedades semântico-gramaticais da construção. Dessa forma,

6 Observe que no primeiro nível da árvore aparece o esquema geral da sufixação. Considerando o modelo de Booij (2005, 2010), trata-se, nesse caso propriamente dito, não de uma construção gramatical, nos termos, por exemplo, de Goldberg (1995), pois não há referência ao polo semântico. Nesse primeiro nível, o esquema, praticamente um molde esquelético, informa apenas as contrapartes fonológica (X, Y) e sintática (especificação da classe, em minúsculas, x, y) dos processos mais gerais de formação de palavras, como os exemplificados em (01): composição, prefixação e sufixação.

7 Como bem lembra o(a) parecerista anônimo(a), a quem agradecemos, o termo “unificação” advém dos primeiros estudos construcionistas de Charles Fillmore, na década de 1980 e “Booij emprega um termo já consagrado na GC, a partir do qual, inclusive, se batiza o modelo conhecido como Unification Construction Grammar. Nesse sentido, a ideia de que a unificação é a “operação utilizada para criar expressões linguísticas bem formadas” é uma premissa fundamental da Gramática de Construções”, válida tanto para palavras novas quanto para sentenças inéditas, e não exatamente uma particularidade da Morfologia Construcional.

é instaurada uma relação bidirecional da construção para o item e do item para a construção. Evidência disso é a compatibilização da palavra portuguesa ‘passeador’, utilizada em referência a pessoas que recebem remuneração para levar animais de estimação, geralmente cachorros, para andar nas ruas. Nesse caso, a unificação do verbo ‘passear’ com o esquema de deverbais em *-dor* resulta no constructo ‘passeador’ (“aquele que passeia (com cachorros) profissionalmente”). Na proposta de Booij (2010), unificação (leia-se compatibilização) é a operação utilizada para criar expressões linguísticas bem formadas.

4. RELAÇÕES DE HERANÇA ENTRE CONSTRUÇÕES

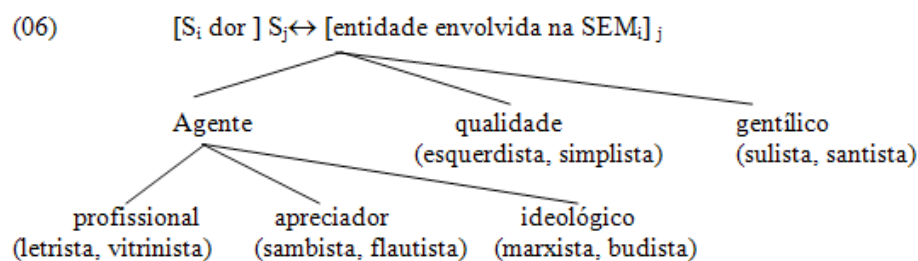
Na Gramática das Construções, entende-se por herança qualquer característica formal ou semântica que esteja na construção básica e se transfira para a construção decorrente. Na proposta de Goldberg (1995), há quatro tipos de herança:

- I. por polissemia (quando há relação entre um sentido específico de uma construção e alguma extensão desse sentido em outra);
- II. por extensão metafórica (quando duas construções se relacionam por meio de mapeamento metafórico);
- III. por subparte (quando parte de uma construção existe independentemente, constituindo outra construção); e, por fim,
- IV. por instanciação (quando uma construção instancia outra, apresentando mais elementos especificados).

Esses quatro tipos de herança, postulados para construções sintáticas, também caracterizam construções morfológicas. Exemplifiquemos o primeiro tipo de herança, por polissemia, com as formações denominais em *-ista*, já representadas conforme o modelo. Booij (2010) afirma que, na formação de palavras, a polissemia pode proporcionar evidências para diferentes níveis de generalização e graus de abstração em uma rede integrada de construções. Para o autor, uma abordagem polissêmica deve partir de um significado prototípico como ponto de partida para os demais significados existentes.

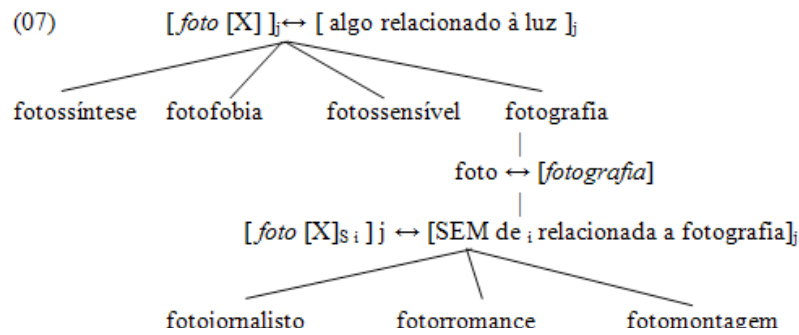
Nas formações *X-ista*, o significado prototípico é o de agente (BASILIO, 2004). Nas formações terminadas em *-ista*, o papel de agente é normalmente associado a uma profissão (‘dentista’, ‘desenhista’, ‘vitrinista’); entretanto, a prática da atividade pode ser prazerosa, vindo daí a ideia de “apreciador”, como em ‘surfista’, ‘jazzista’, ‘sambista’. Na maior parte dos casos, a própria palavra remete às duas acepções de agente, o profissional e o apreciador, a exemplo de ‘pianista’. A prática de uma atividade também pode se dar por algum tipo de adesão ideológica, vindo daí a ideia de “partidário” ou “sectário”, como em ‘umbandista’, ‘gerativista’, ‘esquerdista’. Há formações *X-ista* tipicamente adjetivas, como ‘intimista’ e ‘futurista’, além de gentílicos como ‘paulista’, e ‘nortista’. A reconceptualização da noção de agente modifica esse significado primário e desenvolve um esquema caracterizado por *ligação por polissemia*.

A polissemia não é propriedade da palavra individual, mas do esquema construcional para substantivos denominais em *-ista*, o que leva à criação de um subesquema para agentes. Nesse caso, subesquemas podem ser interpretados como extensões metafóricas ou metonímicas e, por isso, nomes denominais em *-ista* apresentam interpretações que devem ser representadas por subesquemas em uma rede. O esquema geral sanciona algumas opções e os subesquemas expressam quais delas são usadas produtivamente na formação de novas palavras. Em (06), o polo semântico da construção é caracterizado por uma especificação geral o suficiente para abrigar as diversas possibilidades de significação dos nomes em *-ista*, ou seja, estamos, nos termos de Soares da Silva (2006), “puxando o significado para cima”. Na formalização a seguir, SEM, que em Booij (2010) é inespecificado, é aqui interpretado como o *frame* evocado pela palavra-fonte, pois somente desse modo podemos entender que, a partir do mesmo esquema, formam-se subesquemas com semânticas distintas, mas relacionadas.

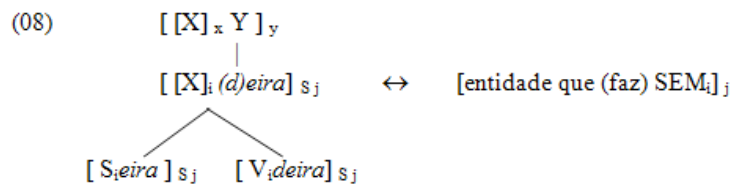


A ligação por extensão metafórica, segunda apontada por Goldberg (1995), constrói-se a partir da projeção do sentido da construção primitiva para outro domínio na nova construção. As formações aumentativas ilustram bem esse tipo de herança, pois várias palavras *X-ão* têm motivação metafórica à medida que a transferência de imagem do domínio-fonte é mantida no domínio-alvo, a exemplo de ‘abelhão’ (“óculos de sol tão grandes que se assemelham aos olhos de uma abelha”) e ‘cebolão’ (“relógio grande demais, semelhante a uma cebola”).

Na ligação por subparte, uma construção corresponde a um pedaço de outra, constituindo porção independente da construção de que se origina. Bom exemplo desse tipo de herança na morfologia do português é a recomposição (CUNHA & CINTRA, 2008; MONTEIRO, 1989). Nessa operação, um radical neoclássico se ressemantiza por metonímia e, ao se fixar em determinada borda da palavra, passa a ser utilizado em referência a uma construção de que era constituinte (GONÇALVES, 2011a). Tal é o caso de *foto-*, cujo significado etimológico (“luz, radiação magnética”), encontrado em ‘fotofobia’ e ‘fotossíntese’, por exemplo, não se atualiza em palavras como ‘fotomontagem’ e ‘fotoestúdio’. Nesses casos, *foto-* equivale a ‘fotografia’, isto é, constitui parte dessa construção – é, literalmente um “pedaço” de ‘fotografia’. O uso de *foto-* como afixoide decorre, portanto, de uma construção por sub-parte, já que esse constituinte compacta o significado do todo e leva essa acepção “zipada” para novas formações, como se vê na representação abaixo:



Para exemplificar o último tipo de herança, por instanciação, considere-se a rede construcional em (08), a seguir, para nomes instrumentais X-(d)eira, em que SEM, como vimos, abrevia semântica e é usado em referência ao *frame* ativado pela palavra-fonte:



Em (08), os esquemas $[V_i deira]_{s_j}$ e $[S_i eira]_{s_j}$ são instanciações do esquema mais básico, $[[X]_i (d)eira]_{s_j} \leftrightarrow [entidade envolvida na SEM_i]_j$, pois atualizam a posição de [X] com informação detalhada sobre a classe da palavra-base, que, por receber a indexação (i), faz parte do léxico tanto quanto o produto (j). Esse tipo de representação possibilita relacionar nomes de instrumentos, como, como ‘iogurteira’ e ‘omeleteira’, de um lado, e ‘batedeira’, e ‘geladeira’, de outro, cuja única diferença é a classe da base (substantivo e verbo, respectivamente). Em modelos baseados em regra – que sustentam hipóteses como a da base unitária (ARONOFF, 1976), em que *inputs* de processos morfológicos não podem ser sintaticamente diferentes –, dados como esses são extremamente problemáticos.

5. APLICAÇÃO DA MORFOLOGIA CONSTRUCIONAL ÀS FORMAS X-DROMO

Fundamentado em sua origem do grego *drómos*, cuja definição é “ação de correr, lugar para corrida, corrida”, o formativo *-dromo* passou por diversas transformações ao longo dos séculos. O elemento em questão adentrou o português por via erudita, sendo ‘hipódromo’, “lugar para as corridas de cavalos”, a primeira palavra dicionarizada. Em Cunha (1986), ‘hipódromo’ data do século XVII, mais especificamente do ano de 1677.

É a partir do século XIX que *-dromo* começa a ser mais utilizado, havendo, na língua, um número maior de palavras com essa terminação. Tal aumento decorre da incorporação de muitas formações no período da nomenclatura científica internacional (final do século XIX e início do século XX), particularmente no domínio da botânica. Um exemplo relacionado à botânica é ‘campilódromo’, cujo significado, de acordo com Houaiss (2009: 441), é a parte da folha “em que diversas nervuras primárias partem de um único ponto na base da folha, formando arcos acentuados e convergindo em direção ao ápice”.

O que se observa, nesse período, é uma tendência a se optar por formativos mais eruditos. Em outras palavras, significa dizer que a utilização de bases gregas e latinas foi feita de modo intencional, com o objetivo de unificar as áreas técnico-científicas e filosófico-literárias cunhando termos a partir de elementos das línguas clássicas. Considerando o exemplo ‘campilódromo’, quando desmembramos as bases, verificamos que *campilo* é oriundo do grego *kampylos*, significando *curvo*, e, uma vez anexado a *-dromo*, nomeia tecnicamente uma pequena parte da folha. São inúmeros os exemplos de nomes em *-dromo* relacionados à botânica, todos definidos de acordo com o dicionário Aurélio:

(09) Acródromo – (1) que ou o que corre para cima; (2) em que duas ou mais nervuras primárias, ou secundárias muito desenvolvidas, partindo da base da folha, se arqueiam e convergem em direção ao ápice.

Actinódromo – em que três ou mais nervuras primárias divergem radialmente a partir da região basal da folha.

Anádromo – (1) que, vivendo no mar, se dirige para o rio na época da desova, ou, que, vivendo em água doce, nada contra a correnteza do rio até a sua cabeceira para a reprodução; (2) em que as nervuras ímpares estão localizadas na face superior e as pares, na inferior (diz-se de nervação das frondes de pteridófitas).

Axonódromo – Em que as nervuras terciárias se dispõem paralelamente às secundárias, das quais partem.

Camptódromo – em que as nervuras secundárias são tangentes à margem da folha.

Catádromo – (1) que, vivendo em água doce, se dirige para o mar na época da desova; (2) Em que as nervuras ímpares estão localizadas na face inferior e as pares, na superior (diz-se da nervação das frondes de pteridófitas)

Outros nomes técnicos incorporados à língua não são necessariamente da botânica, mas, igualmente, são bastante difíceis de interpretar, uma vez que as condições de analisabilidade (BYBEE, 2010) são precárias, considerando-se a opacidade das bases:

(10) Períldromo – Galeria ou espaço coberto em torno de um edifício.

Palíndromo – diz-se de ou frase ou palavra que se pode ler, indiferentemente, da esquerda para a direita ou vice-versa.

Plasmódromo – (1) relativo aos plasmódromos; (2) espécime dos plasmódromos; (3) subfiló de seres unicelulares da classe dos mastigóforos, cujas spp. apresentam núcleo de um único tipo e são providas de pseudópodes [Este agrupamento não tem correspondência nas mais recentes classificações do reino protista.]

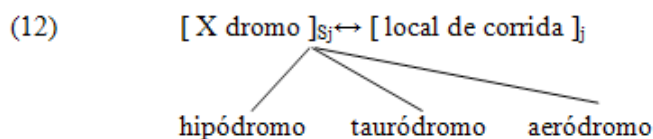
Pródromo – (1) o que antecede a (algo); precursor, prenúncio, antecedente; (2) espécie de prefácio; introdução, preâmbulo; (3) sinal ou sintoma inicial de doença; propatia (ger. us. no plural, reunindo, neste caso, os sinais e sintomas); (4) as primícias de um escritor, os primeiros escritos.

Tomando por base apenas esse conjunto de formas, assumimos, com base no instrumental fornecido pela Morfologia Construcional, que a chamada composição neoclássica⁸ também pode ser modelada por esquemas construcionais semelhantes aos da derivação e composição. No nosso entendimento, como os dois elementos constituem formas presas e, por isso, não recebem etiqueta lexical, podem ser genericamente referenciados como X e Y, em maiúsculas, já que não são afixos, e, por não constarem do léxico, não são indexados (ou seja, não recebem os símbolos i e j, subscritos). Nesse caso, o produto é sempre previsível em termos categoriais: como estamos falando de um vocabulário técnico-científico, o produto é sempre um substantivo⁹:

(11) **Esquema geral da composição neoclássica: [X Y]_s**

Especificação das formações com *-dromo*: [X *dromo*]_s

No caso das formações em *-dromo*, um terceiro conjunto de formações neoclássicas segue o modelo de poucas palavras já existentes na língua – a mais antiga, ‘hipódromo’ (1667), além de ‘tauródromo’ (“lugar de correr dos touros”, 1789), e ‘aeródromo’ (“lugar em que planam aeronaves”, 1878). Aqui, também as bases são presas, mas já é possível associar as formas a um significado genérico de “lugar de corrida”, o que possibilita parear o polo formal da construção com o polosemântico:



Observe-se que o elemento à esquerda continua sem etiqueta lexical, uma vez que não constitui palavra. Por analogia, outras unidades lexicais com base presa foram criadas, aumentando, assim, o volume de formações *X-dromo* remetendo a um local de corrida. Tal é o caso, entre outras, de ‘autódromo’ (“lugar de corridas de carros”), ‘canódromo’ (“local para corrida de cães”) e ‘velódromo’ (“lugar de competição de velocistas, ou seja, condutores de bicicletas”). Ainda assim, os elementos

8 De acordo com Gonçalves (2011b), uma categoria bastante expressiva no conjunto de formativos de classificação problemática é a dos neoclássicos. Nesse grupo, estão os elementos de origem grega ou latina que geralmente não aparecem como forma livre na língua, como mostram os seguintes: *bio-*, *antropo-*, *eco-*, *-metro*, *-nomo*, *-teca*. (origem grega) e *ego-*, *retro-*, *sócio-*, *-cida*, *-fero*, *-cola* (origem latina). De acordo com a definição de Gonçalves (2011b), os neoclássicos são formas de origem grega ou latina que não foram totalmente assimilados na língua que os abrigou. Assim, esses itens não se resumem tão somente a formativos de origem clássica (grega ou latina), mas possuem configuração própria frente aos constituintes propriamente nativos. Elementos dessa natureza são recorrentes na formação de internacionalismos, o que significa que se mostram produtivos no vocabulário técnico-científico e filosófico-literário universal. A terminologia própria da área da ciência, da tecnologia e das artes em geral revela o emprego de formas inspiradas tanto no grego quanto no latim na formação de novos itens lexicais.

9 Novamente aqui, como estamos fazendo referência a um tipo básico de processo morfológico, a contraparte semântica é omitida. Trata-se de nossa proposta de representação para o modelo da composição neoclássica, para nós um processo de formação de palavras diferente da composição em termos formais, como o esquema sugere.

a que *-dromo* se adjunge são radicais (e não palavras), o que nos autoriza afirmar que tais formações ainda são compostos neoclássicos¹⁰.

Nos séculos XX e XXI, aparecem diversas palavras *X-dromo* cunhadas em conformidade com (12). A maioria perde o caráter científico internacional, típico das formações neoclássicas, e experimenta usos menos formais, alguns até bastante populares, como se vê em (13), a seguir:

(13) Jegódromo – pista de eventos em Campina Grande reservada para a corrida de jegues.

Bodódromo – pista para a corrida de bodes.

Burródromo – local reservado para a corrida de burros.

Boiódromo – centro de eventos onde se realizam rodeios de bois.

Cavalódromo – lugar reservado para corrida de cavalos em São José do Cerrito (SC).

A diferença entre as formações em (13) e as mais antigas na língua, no entanto, não reside apenas no caráter mais popular e menos técnico das nomeações. Está em jogo, também, o tipo morfológico com que *-dromo* se combina: embora continue sendo um radical (uma forma presa, portanto), tem livre curso na língua, uma vez que se atualiza como palavra e, por isso, aparece no léxico. Nesse caso, pode receber etiqueta lexical, pois porta informação categorial.

Certamente o gatilho para a grande proliferação de formas *X-dromo* em português foi ‘sambódromo’, “local destinado aos desfiles das escolas de samba, composta por uma passarela e uma área de dispersão das agremiações” (AURÉLIO, 2004: 1134). Na verdade, ‘sambódromo’ é o nome popular dado à Passarela Professor Darcy Ribeiro, localizada na Avenida Marquês de Sapucaí, nos bairros Centro e Cidade Nova, no município do Rio de Janeiro, em 1984. Ao que tudo indica, quem cunhou o termo foi o então governador fluminense Leonel Brizola, que, com o projeto de autoria do arquiteto Oscar Niemeyer, “dotou a cidade de um equipamento urbano permanente para a exibição do tradicional espetáculo do desfile das escolas de samba” (<https://pt.wikipedia.org/wiki/Sambódromo>. Acesso em 28/01/16).

Outra criação em *-dromo* que merece destaque é ‘camelódromo’, que, de acordo com Moreira (2005), constitui denominação popular dada, na década de 1980, pelo então prefeito Marcelo Alencar para os mercados populares. Criado “com a finalidade de atender a crise de emprego, organizar o comércio ambulante que saturava as áreas da Central do Brasil e da Rua Uruguaiana no centro da cidade” (MOREIRA, 2005: 01), o camelódromo da Uruguaiana logo deu origem a diversos outros, tanto no Rio quanto em outros estados, sendo o termo, hoje, usado em referência a qualquer concentração de camelôs (comércio ambulante). Sem dúvida alguma, a popularidade e a frequência de uso dessas

¹⁰ Concordamos com Bauer (2005: 105), quando ele afirma que “o rótulo ‘composto neoclássico’ se mostra inadequado, uma vez que um composto neoclássico não é um composto (de acordo com leitura normal da palavra), sendo mais um problema terminológico do que um problema de substância”.

formações engatilhou uma produção em série, levando à criação de formas como as listadas em (14), a seguir, em que a noção de “corrida” já não se manifesta:

(14) Fumódromo – área destinada exclusivamente ao uso de cigarros, cigarrilhas, charutos, cachimbos ou de qualquer outro produto fumífero, derivado ou não do tabaco, devidamente isolada e com arejamento conveniente.

Alcoódromo – lugar destinado para a cura de ressacas provenientes de malestar causado pela ingestão de bebidas alcoólicas.

Mamódromo – Local que se encontra, em alguns shoppings no Rio e em São Paulo, reservados para quem está amamentando o bebê.

Urinódromo – muro de azulejo, com água e algum tipo de detergente escorrendo o tempo todo, onde os homens podem urinar protegidos por uma tapadeira (espécie de biombo, que deixa à mostra apenas cabeças e pés).

Os dados revelam, em relação à forma, que *-dromo* passa a incorporar sistematicamente a vogal média posterior aberta, [ɔ], o que nos leva a pensar na sua reformatação como *-ódromo*, certamente motivada não apenas pelas formações mais antigas e transparentes, todas com essa vogal, mas também pelas mais novas, mesmo aquelas terminadas em outros segmentos. Como se anexa a radicais, as vogais temáticas das palavras-base são sempre suprimidas, como se vê em (15), a seguir. Além disso, bases atemáticas são sempre caracterizadas pela epêntese consonantal de /z/, consoante de ligação por excelência na língua (KEHDI, 1989). Não faz sentido analisar a vogal <ó> como elemento de ligação, pois as palavras em (16), a seguir, teriam uma vogal e uma consoante de ligação simultaneamente, o que obviamente é contraintuitivo. Preferimos interpretar a sequência morfológica como *-ódromo*, seguindo, portanto, Sandmann (1987) e Laroca (2005).

- | | | |
|------|----------------------|-------------------|
| (15) | samba – sambódromo | cagar – cagódromo |
| | bicha – bichódromo | fumar – fumódromo |
| (16) | café – cafezódromo | pai – paizódromo |
| | urubu – urubuzódromo | gay – gayzódromo |

Em decorrência dessas considerações, a alta produtividade das formações *X-ódromo* no português contemporâneo nos leva a inferir que houve mudança do esquema da composição neoclássica para um esquema que mais se assemelha ao da sufixação. São várias as razões para isso:

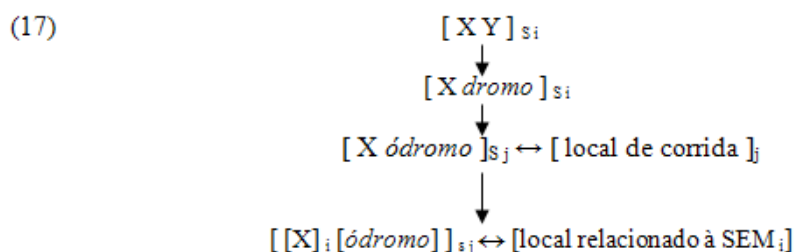
- o elemento à esquerda deixa de ser um radical neoclássico e passa a ser uma forma com livre curso, que, portanto, possui etiqueta lexical;
- a forma à direita passa a ser sempre *-ódromo*, ao contrário do que se observa em palavras mais antigas, com elementos diversos antes

de *-dromo*: ‘catádro’ (“que, vivendo em água doce, se dirige para o mar na época da desova”), ‘perídro’ (“galeria ou espaço coberto em torno de um edifício”), e ‘palíndro’ (“frase ou palavra que se pode ler, indiferentemente, da esquerda para a direita ou vice-versa”);

(c) as formas mais novas espelham o comportamento formal e semântico das mais antigas e de uso mais geral (‘hipódro’, ‘autódro’, ‘velódro’);

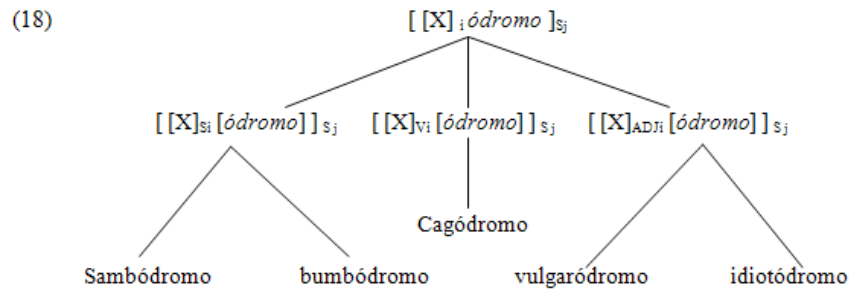
(d) *X-ódro* passa a significar apenas “lugar”, não necessariamente remetendo à noção de corrida.

Em termos representacionais, temos o seguinte: primeiramente, temos o esquema genérico da composição neoclássica no primeiro nó de (17), a seguir, cuja instanciação, no segundo nó, incorpora a partícula *-dromo*, mas sem qualquer referência ao polo semântico. Na terceira linha, temos o esquema da composição neoclássica *X-ódro* fazendo alusão a locais de corrida, como ‘hipódro’. Observe que X, a forma à esquerda, não contém etiqueta lexical nem é indexada. No nó terminal, X recebe a indexação, o que sinaliza que a base faz parte do léxico. No pólo semântico, tem-se a ideia de um lugar relacionado ao *frame* (SEM) projetado pela palavra-matriz. Em (17), portanto, tem-se o percurso histórico das formações *X-(ó)dro*¹¹:

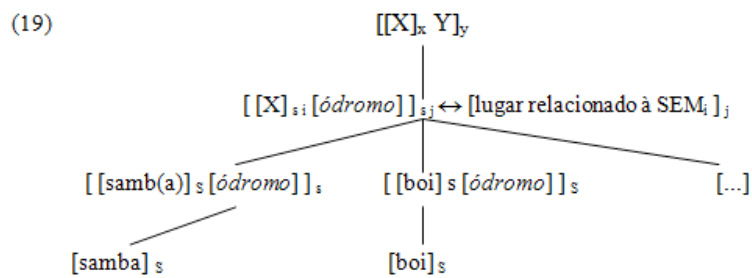


Deve-se considerar que o elemento X pode estar relacionado a três diferentes tipos de base: substantiva, verbal e adjetiva. A partir do *corpus* levantado, pôde-se constatar que o *input*, em sua maioria, é formado por substantivos, a exemplo do que se observa em ‘samba’. Com o intuito de melhor visualizar as três possibilidades, representamos cada uma delas no esquema a seguir. Vale ressaltar que os subscritos (i) e (j) são indicativos de que tanto a base quanto o produto fazem parte do léxico.

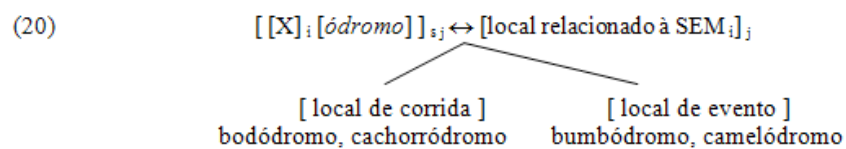
11 No primeiro nível da formalização, temos o esquema abstrato da composição neoclássica que, de modo algum, constitui padrão fossilizado na língua, sendo, nos dias de hoje, tão produtiva quanto os demais processos: composição, prefixação, sufixação etc. Evidência disso são as recentes construções ‘ortorexia’ (“distúrbio alimentar caracterizado pelo consumo de produtos sem química”), ‘pedófilo’ (“pessoa com atração sexual por crianças”) e ‘heterofóbico’ (“aversão a heterossexuais”). Na segunda linha, aparecem as formas com *-dromo* existentes na língua para as quais o falante não consegue abstrair qualquer significado (daí a falta de especificação do polo semântico), tais como ‘palíndromo’ e ‘anádro’. Por fim, as duas últimas linhas generalizam os significados mais gerais, que seguem o percurso histórico das formações *X-dromo* na língua.



Levando-se em consideração as esquematizações desde o nível mais abstrato até as instanciações individuais, em que os nós mais baixos herdariam propriedades dos nós dominantes e, por isso mesmo, contêm informações redundantes, as construções *X-ódromo* hoje instanciam o esquema mais abstrato da sufixação ($[X]_i Y_j$). Atendo-se à representação mais recorrente na formação, isto é, tendo como base um substantivo, demonstramos, a seguir, como ‘sambódromo’ e ‘boiódromo’ herdaram as características do nó imediatamente dominante da semântica de lugar, tendo em vista a possibilidade de outras formações serem contempladas:



No que diz respeito à contraparte semântica, a noção genérica é de ‘local relacionado à SEM_i’, tendo, portanto, uma ampliação no sentido original, que fazia alusão a “ação de correr, lugar para corrida, corrida” (HOUAISS, 2009). A partir das instanciações individuais, a contraparte semântica nos possibilitaria destacar, para alguns exemplos, as acepções ‘local de corrida’ e de ‘local de evento’:



A partir dessas noções, que, na Gramática das Construções, de Goldberg (1995), seriam heranças **por instanciação**, os exemplos acima possibilitam entender que, seja ‘local de corrida’, seja ‘local de evento’, a noção genérica de ‘lugar’ aparece em todos os dados. De certa forma, a própria expansão do significado em relação ao original ajuda na percepção da existência de uma modificação no estatuto morfológico do formativo.

Ao levarmos em consideração, especificamente, a acepção ‘local de evento’, o *corpus* nos permite perceber que há especificação de significado em exemplos como ‘bambódromo’, ‘porcódromo’ e

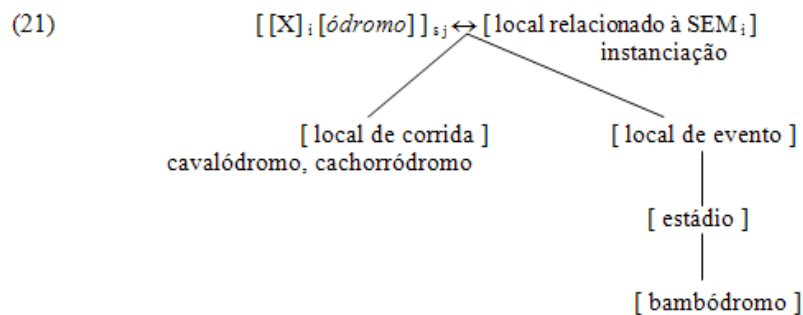
‘peixódromo’. Sem contextualização, os dados seriam entendidos como “local que se destina para a prática de corrida de bambis, porcos e peixes”, respectivamente. Ponderando-se, todavia, o contexto de onde esses vocábulos foram retirados, a intenção foi (re)nomear depreciativamente estádios de futebol. Para o primeiro – ‘bambódromo’ – há alusão pejorativa ao estádio do time do São Paulo F. C., o Cícero Pompeu de Toledo, popularmente conhecido como Morumbi. O contexto para a criação do termo alusivo ao animal teve a participação da equipe do Sport Club Corinthians Paulista, partindo de um jogador. O volante Vampeta, ao passar por uma sorveteria, avistou dois jogadores do São Paulo, Kaká e Júlio Baptista, e, em um tom jocoso, chamou a dupla de ‘bambi’, símbolo da homossexualidade masculina. Alguns dias depois, o jogador do Corinthians voltou a repetir o mesmo termo em uma entrevista coletiva, fato que teve repercussão considerável. Daí a origem de *bambi* como sendo uma alusão aos torcedores saopaulinos (metonímia) e, por conseguinte, ‘bambódromo’ para o Morumbi.

Nos casos de ‘porcódromo’ e ‘peixódromo’, também há alusão a estádios. Em ambos os exemplos, levamos em conta que os processos atuantes na formação assemelham-se aos atuantes na palavra ‘bambódromo’. Sem contexto, os novos itens lexicais também seriam entendidos como locativos genéricos (não como estádios de futebol).

‘Porcódromo’ constitui referência ao estádio da equipe da Sociedade Esportiva Palmeiras, cujos torcedores são conhecidos como *porcos*. Tal denominação se deu em razão de uma divergência entre os Presidentes do Palmeiras e do Corinthians na década de 60. Após o alvinegro – Corinthians – empatar no campeonato paulista com a equipe do São Bento em 1x1, o time retornou à capital e uma tragédia ocorreu com dois de seus principais jogadores: falecimento. Requisitando à Federação Paulista de Futebol que ambos fossem substituídos, o presidente corinthiano recebe a notícia de que a substituição só seria possível mediante aprovação unânime de todos os outros presidentes que faziam parte da Federação. Somente um dirigente se posicionou contra a substituição dos jogadores falecidos: o do Palmeiras. A negativa por parte do dirigente fez com que o cartola alvinegro chamasse os palmeirenses de “porcos”. O jogo subsequente do Corinthians era justamente contra a equipe alviverde, circunstância que levou os primeiros a soltarem um porco no gramado do Morumbi provocando um coro uníssono de “porco! porco!” como forma de zombar do adversário. O tom zombeteiro perdurou até o ano de 1986, quando a equipe palmeirense goleou por 5x1 a equipe Corinthians na semifinal do campeonato paulista. Desde então, faz-se uma referência aos torcedores do Palmeiras como *porcos* e, por conseguinte, ‘porcódromo’ é o seu estádio, o Palestra Itália. A curiosidade se dá pelo fato de os próprios torcedores do Palmeiras terem adotado a “mascote” de modo positivo. Tal fato se comprova, inclusive, a partir da criação de uma torcida organizada com a nomenclatura de *pork’s*.

Por último, ‘peixódromo’ faz alusão, também por metonímia, à torcida do Santos F. C. Nessa situação, no entanto, há um fato curioso: a mascote do time é uma baleia, mais especificamente uma orca. Ainda que este seja o representante animal oficial, a equipe é tida como *peixe*. Tal fato remonta à primeira metade dos anos 1990, quando a equipe da baixada santista viajava para a capital. Quando jogava em São Paulo, a torcida adversária costumava chamar os torcedores de ‘peixes podres’, ‘peixeiros’. No ano de 1933, quando o Santos recebeu a equipe da Floresta – o time precursor do atual São Paulo

F. C. – na Vila Belmiro – estádio do Santos –, uma vez mais a torcida adversária começou a gritar a mesma ofensa. Entretanto, os próprios torcedores santistas responderam que ‘eram peixeiros com muito prazer, com muita honra’. Desde então, o apelido foi assumido e a equipe virou o ‘time do peixe’. Ao se levar em consideração que o time é conhecido justamente por essa referência, percebe-se que há relação em que a totalidade – a torcida – passa a ser designada pelo animal representativo dessa torcida: o peixe. Desse modo, o estádio do Santos F. C. passa a ser conhecido por ‘peixódromo’. Em termos de esquematização, teríamos o seguinte:



Com isso, percebemos que, na contraparte semântica, a ideia de locativo se mantém; todavia, há possibilidade de expansão no que diz respeito a esse lugar. Com isso, especificamente para esses três vocábulos – ‘bambódromo’, ‘porcódromo’ e ‘peixódromo’ – o local de evento estabelece relações que permitem fazer alusão a lugares mais particulares: estádios de futebol.

5. CONSIDERAÇÕES FINAIS

Diante das considerações feitas ao longo do texto, concluímos que a Morfologia Construcional apresenta-se como um arcabouço apropriado para um estudo mais sistemático acerca do comportamento de formativos de difícil categorização, como o aqui analisado, *-ódromo*. Desse modo, composição e derivação não são processos claramente distintos; possuem fronteiras maleáveis e os processos de formação de palavras podem se modificar ao longo dos anos, bem como o estatuto dos formativos. No nosso entendimento, é isso que vem acontecendo com as formações examinadas no artigo.

Em termos de mudança, temos o seguinte para as formações *X-dromo*: primeiramente, constituem instanciação ([X *dromo*]_{sj}) do esquema genérico da composição neoclássica que defendemos no texto, [X Y]_{sj}, que expressa a generalização de serem opacas, mas sistematicamente categorizadas como substantivas, dada principal função desse processo: nomear termos de áreas de especialidade (Botânica, no caso de *-dromo*). A partir de ‘hipódromo’, novas formas *X-ódromo* foram criadas para cunhar locais de corrida. Nesse caso, a forma à esquerda não contém etiqueta lexical nem é indexada, pois as bases são presas e, por isso mesmo, não recebem informação categorial, mas, nessa fase, já há um pareamento mais sistemático entre o polo formal e o polo semântico: [X *ódromo*]_{sj} ↔ [local de corrida]_j. Nos dias de hoje, X recebe a indexação, o que sinaliza que a base faz parte do léxico. No polo semântico, tem-se a ideia de um lugar relacionado ao *frame* (SEM) projetado pela palavra-matriz, como em ‘camelódromo’ e ‘boiolódromo’. Deve-se considerar que o elemento X pode estar

relacionado a três diferentes tipos de base: substantiva ('sambódromo'), verbal ('fumódromo') e adjetiva ('vulgaródromo'). Os esquemas [V_iódromo]_{s_j}, [S_iódromo]_{s_j} e [Adj_iódromo]_{s_j} são instanciações do esquema mais básico, [[X]_iódromo]_{s_j} ↔ [local relacionado à SEM_i], pois atualizam a posição de [X] com informação detalhada sobre a classe da palavra-base, que, por receber a indexação (i), faz parte do léxico tanto quanto o produto (j).

Assumimos, portanto, que houve tanto mudança formal quanto semântica nas construções, que passaram de compostas e opacas (X-dromo, em que o elemento à esquerda não constitui palavra) a derivadas e transparentes ([X-ódromo, com X recebendo, agora, etiqueta lexical e se atualizando como palavra).

REFERÊNCIAS:

ARONOFF, M. *Word formation in generative grammar*. Linguistic Inquiry Monograph 1. Cambridge, MA: MIT Press, 1976.

AURÉLIO, Buarque de Holanda. *Novo Dicionário Eletrônico Aurélio*. Versão 5.0. 3ª ed. Atualizada. 1ª. impressão Editora Positivo. Positivo informática Ltda, 2004.

BASILIO, M. *Teoria Lexical*. 8ª edição. São Paulo: Ática, 2004.

BASÍLIO, M. Abordagem gerativa e abordagem cognitiva na formação de palavras: considerações preliminares. *Linguística* (Rio de Janeiro), v. 6, p. 11-26, 2011.

BAUER, L. The Borderline between Derivation and Compounding. In: W. Dressler et al. (eds.). *Morphology and its Demarcations*. Amsterdam / Philadelphia: John Benjamins Publishing Company, pp. 97-108, 2005.

BECHARA, E. *Moderna Gramática Portuguesa*. 37ª ed. ver. e ampl. 14ª reimpr. Rio de Janeiro: Lucerna, 2004.

BOOIJ, G. *Compounding and Derivation. Evidence for Construction Morphology*. In: W. Dressler et al. (eds.). *Morphology and its Demarcations*. Amsterdam / Philadelphia: John Benjamins Publishing Company, 2005, pp. 109-131.

BOOIJ, G. Construction morphology and the lexicon. In: MONTERMINI, F.; BOYÉ, G.; HATHOUT, N. (eds.). *Selected proceedings of the 5th Décembrettes: Morphology in Toulouse*. Somerville: Cascadilla Press, pp. 34-44, 2007.

BOOIJ, Geert. *Construction morphology*. Oxford: Oxford University Press, 2010.

- CORBIN, D. *Morphologie dérivationnelle et structuration du lexique*. Tubinga: Max Niemeyer Verlag, 1987.
- CUNHA, A. G. *Dicionário Etimológico Nova Fronteira da língua portuguesa*. 2ª ed. Rio de Janeiro: Nova Fronteira, 1986.
- CUNHA, Celso & CINTRA, Luís F. Lindley. *Nova gramática do português contemporâneo*. 5ª ed. Rio de Janeiro: Lexikon, 2008.
- DI SCIULLO, A. & WILLIAMS, E. *On definition of word*. Cambridge: Cambridge University Press, 1987.
- GONÇALVES, C. A. V. Compostos neoclássicos: estrutura e formação. *Revista Virtual de Estudos da Linguagem*, Vol. 9, número 5, p. 6-39, nov. de 2011.
- GONÇALVES, C. A. V. Composição e derivação: polos prototípicos de um continuum? Pequeno estudo de casos. *Domínios de Lingu@gem*, 5, p. 62-89, nov. 2011.
- GONÇALVES, C. A. V. & ALMEIDA, M. L. L. *Morfologia construcional: principais ideias, aplicação ao português e extensões necessárias*. Alfa (ILCSE/UNESP), v. 58, n. 1: 165-193 2014.
- GOLDBERG, A. *Constructions: a construction grammar approach to argument structure*. Chicago; London: The University of Chicago Press, 1995.
- HOUAISS, Antônio. *Dicionário eletrônico Houaiss da Língua Portuguesa*. Versão monusuário 3.0. Objetiva, 2009.
- KEHDI, V. *Morfemas do português*. São Paulo: Ática.
- LANGACKER, R. W. *Foundations of cognitive grammar*. v. I: theoretical prerequisites. Stanford: University Press, 1987.
- LAROCA, M. N. C. *Manual de morfologia do português*. 4ª edição, revisada e ampliada, Campinas, SP: Pontes, Juiz de Fora MG, 2005.
- MONTEIRO, J. L. *Morfologia portuguesa*. São Paulo: Ática, 1989.
- MOREIRA, Ângela. Mercados populares ou camelódromos: nascimento e variações de um objeto arquitetônico. *Anais II Colóquio Internacional sobre o comércio e cidade*. Rio de Janeiro, UFRJ, 2005.

PIRES, J. A. O. *O estatuto morfológico do formativo -dromo no português brasileiro*. Dissertação (Mestrado em Letras Vernáculas). Rio de Janeiro: UFRJ, 2014.

SANDMANN, A. J. Novidades do “front” da formação de palavras. *Letras*, Curitiba, 36 (1), p. 54-68, 1987.

SANDMANN, A. J. *Formação de palavras no português brasileiro contemporâneo*. Curitiba: Scientia & Labor, 1988.

SCALISE, S. *et alii*. Exocentricidade na composição. *Gengo Kenkyu* 135, p. 49-84, 2009.

SOARES da SILVA, A. *O mundo dos Sentidos - Polissemia, semântica e cognição*. Coimbra: Almeida, 2006.

Recebido em 31/01/2016

Aceito em 02/05/2016

FORMAS LIVRES E FORMAS PRESAS: UM CLÁSSICO REVISITADO COM OLHAR COGNITIVISTA

Janderson Lemos de Souza (UNIFESP)¹

RESUMO

Este artigo retoma a distinção entre formas presas e formas livres para abordá-la segundo a linguística cognitiva, especificamente pela (i) suspensão do compromisso com o princípio da sincronia; (ii) identificação dos processos cognitivos que estruturam os fenômenos formais; e (iii) afinidade com modelos baseados no uso. O artigo combina tais fundamentos da teoria com fenômenos morfológicos do português, tendo em vista caracterizar a distinção entre formas livres e presas, não mais como fato relativo a produtos, mas como fato relativo a processos. Assim, são consideradas formas presas não somente as que não existem na língua e, portanto, não podem ocorrer sozinhas, mas sobretudo as que o falante já não associa a outras formas graças à inatividade de um processo ou que o falante não tem como rastrear a partir de outras apesar da atividade de um processo.

PALAVRAS-CHAVE: Formas livres. Formas presas. Morfologia. Semântica. Linguística cognitiva.

ABSTRACT

This article approaches the distinction between bound forms and free forms according to cognitive linguistics, especially in what comes to (i) its non-synchronic character; (ii) the pursuit of cognitive processes behind formal phenomena; and (iii) its compatibility with usage-based models. The article combines such bases with morphological facts of Portuguese, with a view to characterizing the distinction between free and bound forms, not as a fact related to products, but as a fact related to processes. Thus, not only the non-existing forms, which consequently cannot be spoken alone, but mainly those which the language-user does not link to others due to the inactivity of a process or that the language-user cannot retrieve from others despite the activity of a process are postulated to be bound forms.

KEYWORDS: Free forms. Bound forms. Morphology. Semantics. Cognitive linguistics.

¹ Professor Adjunto do Departamento de Letras da Universidade Federal de São Paulo (UNIFESP).

INTRODUÇÃO

A distinção entre formas livres e formas presas entrou para a história da linguística por obra de Bloomfield (1933, p. 160): “Uma forma linguística que nunca é dita sozinha é uma *forma presa*; todas as outras (como, por exemplo, *João correu* ou *correr* ou *correndo*) são *formas livres*.”² Tal distinção deu ensejo a outra distinção clássica, proposta por Câmara Jr. (1989, p. 88), que “(...) nos leva a ampliar a classificação de Bloomfield com uma categoria de formas, que não são livres, mas também não são presas, senão apenas dependentes daquelas a que se adjungem nas frases.”

Neste artigo, pretendo revisitar essas distinções, formuladas pelo estruturalismo, escola centrada nas formas linguísticas, com um olhar comprometido com a linguística cognitiva, escola cujo pressuposto básico é o de que a gramática é semanticamente motivada. O objetivo é caracterizar formas livres como cognitivamente disponíveis e formas presas como cognitivamente indisponíveis a partir do compromisso com o modelo baseado no uso em Bybee (2010).

Pretendo, ao longo do artigo, sustentar o entendimento de que formas livres e presas são as que, respectivamente, estão cognitivamente disponíveis e indisponíveis por obra da frequência de uso, e não as que existem ou não existem num dado recorte sincrônico, o que me leva a considerar como presas não somente as formas que não foram transmitidas por aquisição.

Porque cada instância de uso da língua impacta a representação, a variação e a gradiência têm uma representação direta no sistema do usuário da língua. Num modelo exemplar, todas as variantes são representadas na memória como nuvens de exemplares. Tais nuvens podem mudar gradativamente, representando as mudanças por que a língua passa à medida que é usada. Assim, postula-se que a mudança ocorra mais no uso da língua que no processo de aquisição.³ (Bybee, 2010, p. 9)

Os dados com que pretendo ilustrar a pertinência de deslocar entre formas livres e presas dos produtos para os processos são de caráter morfológico. A utilização desses dados pressupõe um léxico em que “Palavras são vistas como construções, e o significado lexical é uma rede intrincada de molduras conectadas”⁴ (Fauconnier & Turner, 2006, p. 303).

ANÁLISE DO PROBLEMA

Em Lemos de Souza (2015), identifico, na história da linguística, em geral, e da morfologia, em particular, o convívio entre a atenção dedicada à dimensão ativa (a dos processos estruturantes, nos termos de Castilho, 2010) e a atenção dedicada à dimensão passiva (a dos produtos estruturados,

2 “A linguistic form which is never spoken alone is a bound form; all others (as, for instance, John ran or run or running) are free forms.”

3 “Because each instance of language use impacts representation, variation and gradience have a direct representation in the language-user’s system. In an exemplar model, all variants are represented in memory as exemplar clusters. Such clusters can change gradually, representing the changes that language undergoes as it is used. Thus change is postulated to occur as language is used rather than in the acquisition process.” Opto por traduzir cluster como nuvem de exemplares acompanhando Cristóforo-Silva (2002), que revisita a clássica definição de alofonia segundo a teoria de exemplares proposta por Bybee (2001).

4 “Words are themselves viewed as constructions, and lexical meaning is an intricated web of connected frames.”

nos termos do mesmo autor). Atribuo esse convívio no campo dos analistas ao fato de, no campo dos objetos, as línguas naturais, coexistirem produtos de processos ativos e produtos de processos inativos, razão pela qual, “(...) se estabelecermos de início e por princípio uma diferença entre formas já feitas e processos de formação, aí passaremos a ter alguma possibilidade de estudar com mais detalhes os processos (...)” (Basilio, 1987, p. 24).

A concepção de processos não é exclusiva da linguística cognitiva. Em tese, a proposta acima poderia ser acolhida por qualquer teoria. Neste artigo, o processo que considero central é o de estabelecer relações entre formas, que pressupõe, mais que a existência das formas, o uso das formas.

Vejam, por exemplo, formações em –nte:

(1) *refrigerante, conservante, fertilizante, inteligente, detergente e dente*

Escolhi formações com diferentes graus de transparência/opacidade pois, como se verá adiante, essa é uma questão relevante no quadro teórico aqui adotado. As três primeiras ilustram o que tanto Bloomfield (1933) quanto Câmara Jr. (1989) reconheceriam como formadas a partir de formas livres: respectivamente, *refrigerar*, *conservar* e *fertilizar*. As duas seguintes, o que ambos os autores reconheceriam como formadas a partir de formas presas. Somente um recuo etimológico permite rastrear os verbos *intelligĕre* e *detergĕre*. Já a última forma seria considerada como simples por escolas formalistas, tamanha a opacidade. Escolhi uma formação que compartilha raiz com o verbo *edĕre* porque também é ele a base do exemplo usado por Câmara Jr. – *comer* – para ilustrar a importância de distinguir entre sincronia e diacronia.

Semanticamente, os diferentes graus de transparência/opacidade morfológica se revelam irrelevantes já que a todas as formações se pode atribuir a interpretação “aquele/aquilo que X, em que X é a base”: respectivamente, que refrigera, que conserva, que fertiliza, que entende, que limpa, que come. Refiro-me aos significados etimológicos, e não aos significados com que tais palavras são usadas atualmente, exatamente porque as de maior opacidade reduzem a analiticidade da relação base-afixo.⁵

Outro ponto de interface entre semântica e morfologia é a relação entre complexidade morfológica e frequência de uso:

Hay 2001, 2002 discute os efeitos da frequência relativa sobre palavras morfológicamente complexas. A frequência relativa diz respeito à frequência de uma palavra complexa comparada com a da base que ela contém. Costuma ser o caso de uma palavra mais complexa ou derivada ser menos frequente que a base mais simples a partir da qual é formada, como a teoria da marcação preveria. Assim, *aliciar* é mais frequente que

5 Mantém-se afastada a falácia etimológica contra a qual Lyons (1968, p. 4-6) previne: tomar o significado etimológico como verdadeiro ou natural. Pretende-se invocar o caminho da etimologia à pragmática tal como proposto por Sweetser (1990), em que a recuperação do étimo ilumina as bases experienciais de que parte a polissemia.

aliciamento; eterno é mais frequente que eternamente; sutiã é mais frequente que sem-sutiã. No entanto, também há casos em que o oposto é verdadeiro: diagonalmente é mais frequente que diagonal; depreciação é mais frequente que depreciar e frequentemente é mais frequente que frequente. Hay demonstra, por meio de vários experimentos, que as palavras derivadas que são mais frequentes que suas bases são menos composicionais ou menos transparentes semanticamente que as palavras complexas que são menos frequentes que suas bases.⁶ (Bybee, 2010, p. 46)

Essa citação propicia algumas observações importantes. A primeira é que a frequência de uso das palavras citadas só pode ser medida não apenas em que cada língua, como, acredito, em cada gênero num dado momento. A segunda é que os experimentos empregados por Hay (2001, 2002) estão sendo assumidos tanto por Bybee (2010) quanto por mim. A terceira é que a linguística cognitiva cada vez mais adota métodos empíricos que vão além do levantamento de dados em uso real, tendo em vista, por exemplo, refinar o que se deve entender por alta ou baixa frequência de uso, medição tão relevante para as asserções teóricas em Bybee (2010) e aqui. A quarta é que “(...) palavras morfológicamente complexas podem variar em frequência ou força de representação e cada uma pode ter seus próprios graus de composicionalidade e analiticidade, dependendo do quão fortemente cada palavra é associada a outras instâncias de seus componentes”⁷ (Bybee, 2010, 32). A quarta observação reforça a posição da autora segundo a qual composicionalidade e analiticidade são propriedades gradientes.

Como se vê, a conclusão a que Hay (2001, 2002) chega é que as palavras morfológicamente complexas mais usadas que suas bases são as menos composicionais ou transparentes. A discussão sobre composicionalidade e transparência mereceria um artigo. Aqui me limito a observar que a conclusão de Hay (2001, 2002) remete a Langacker (1987), para quem o *status de unidade* é proporcional a menor analiticidade. Estudos sobre processos não concatenativos de formação de palavras (cf. Gonçalves, 2006), no entanto, indicam que o status de unidade é compatível com transparência, ou não seria possível ao falante mapear a estrutura prosódica envolvida nos cruzamentos vocabulares e substituições sublexicais, principalmente.

DISCUSSÃO DO PROBLEMA

Dados do português brasileiro nos permitem endossar a tese de insensibilidade da frequência à complexidade morfológica. Tomemos alguns pares a título de exemplos:

6 “Hay 2001, 2002 discusses the effects of relative frequency on morphologically complex words. Relative frequency refers to the frequency of a complex word as compared to the base that it contains. It is often the case that the more complex or derived word is less frequent than the simpler base from which it derives, as the theory of markedness relations would predict. Thus entice is more frequent than enticement; eternal is more frequent than eternally; top is more frequent than topless. However, there are also cases where the reverse is true: diagonally is more frequent than diagonal; abasement is more frequent than abase and frequently is more frequent than frequent. Hay demonstrates through several experiments that the derived words that are more frequent than their bases are less compositional or less semantically transparent than complex words that are less frequent than their bases.”

7 “(...) morphologically complex words can vary in frequency or strength of representation and each can have its own degrees of compositionality and analysability, depending upon how strongly each word is connected to other instances of its component parts.”

(2) *assistir – assistência*
gerir – gerência
emergir – emergência
urgir – urgência

insistir – insistência
arder – ardência
viger – vigência
permanecer – permanência

A relação V – N já foi, em capítulo anterior do lexicalismo gerativo, considerada como nominalização (cf. Basilio, 1980), o que, em grande medida, guarda o espírito do que aqui chamo de relações entre formas como processo que, nos termos de Castilho (2010), pode ser ativado, desativado ou reativado, e é a desativação desse processo o que destitui de formas existentes a condição de livres. A concepção de nominalização empregada neste artigo, no entanto, é a de formação de um substantivo a partir de um verbo (cf. Basilio 1987, 2009; Lemos de Souza, 2010, 2012), sem que isso implique adesão a uma grande classe dos nomes em detrimento da distinção entre substantivos e adjetivos, pelos motivos expostos por Basilio (2008).

Os pares em (2) permitem constatar que *assistência* é mais frequente que *assistir*, *gerência* mais que *gerir*, *emergência* mais que *emergir*, *urgência* mais que *urgir*. Já *insistir* e *insistência* parecem gozar da mesma frequência de uso.⁸ Os demais pares, em que o verbo é da 2ª conjugação, e não da 3ª, não permitem ao falante apostar, com certeza, num verbo em –ir a partir do substantivo deverbal quando este é mais usado que o verbo. Daí ser comum o falante culto apostar em *vigir* a partir de *vigência*, por exemplo. A inexistência de nominalizações em –ência contribui para essa insegurança no rastreamento dos verbos correspondentes a substantivos em –ência.

(3) “O envolvimento da categoria no fortalecimento de suas entidades é fundamental neste momento em que tramitam no Congresso Nacional duas Propostas de Emenda Constitucional prevendo o retorno da exigência do diploma de curso superior em Jornalismo como requisito para o exercício da profissão. Considerado um dos pilares da regulamentação profissional dos jornalistas, após vigir por 40 anos no Brasil este requisito foi derrubado pelo STF que, no ano passado, acatou recurso impetrado por entidades patronais.” (*FENAJ convoca jornalistas a participarem do processo eleitoral da entidade*, 29/11/1999, Sindicato dos Jornalistas Profissionais do Estado de São Paulo, www.sjisp.org.br)

Até aqui, estou considerando pares verbo – substantivos deverbais por três motivos. Primeiro, porque a nominalização ocupa lugar de destaque na linguística desde Chomsky (1970), quando a teoria gerativa migrou do modelo padrão ao modelo padrão estendido. Segundo, porque tenho desenvolvido reflexões sobre a nominalização de verbos do ponto de vista da linguística cognitiva (Lemos de 2010, 2011, 2012, 2015). Terceiro, porque a relação entre verbos e os substantivos formados a partir deles ilustra a relação entre menor e maior complexidade morfológica de que trata Hay (2001, 2002) do ponto de vista de um modelo baseado no uso.

⁸ Importante ressaltar que asserções sobre frequência de uso devem ser fundamentadas em análises de corpora, por sua vez sensíveis aos gêneros textuais reunidos. Da mesma forma, carece de precisão em Bybee (2010) a distinção entre alta e baixa frequência. O quanto deve ser considerado alto ou baixo assim como a verificação empírica num corpus escapam ao caráter teórico deste artigo, e não do alcance do problema aqui abordado.

Ressalto, no entanto, que a relação de maior ou menor complexidade se pode encontrar na esfera da flexão e que a distinção entre frequências de uso se estende a categorias gramaticais, não se aplica apenas a formas. Por exemplo, um tempo verbal pode ser mais ou menos usado que outro, e o tempo mais usado pode recrutar uma categoria de formas, que, por isso mesmo, terá sua frequência de uso inflada por participar desse tempo.

Entendo que é o que se verifica na relação entre tempos compostos e participípios passados. Novamente, o grau de regularidade morfológica do verbo pode inibir o uso de um tempo em que a irregularidade se revela, como o presente do indicativo, e favorecer ao uso de um tempo em que a irregularidade se neutraliza. Por exemplo, *caber* tem a forma *caibo* menos usada que a forma *cabido*, não somente porque *caibo* revela a irregularidade enquanto *cabido* a neutraliza, mas também porque todos os tempos compostos em que se flexione o verbo recrutam a forma *cabido*.

Isso é pertinente ao tema analisado e contribui para a tese defendida porque, da mesma forma que o falante culto aposta em *vigir* pela maior frequência de alguns substantivos que a dos verbos a partir dos quais se formam (cf. *viger* – *vigência*), o falante culto aposta em *tolhir* pela maior frequência do participípio que a do infinitivo (cf. *tolher* – *tolhido*). A maior frequência do participípio, por sua vez, atribuo à alta frequência dos tempos compostos, necessariamente formados por participípios.

(4) “Presidente do TCE-AM é acusado de tolhir atividades de auditor”
(A Crítica, 30/3/2016, www.acritica.com)

Quando mencionei que a inexistência de nominalizações em –íncia dificulta o rastreamento de verbo da 2ª ou da 3ª conjugação a partir dos substantivos em –ência nos casos em que o substantivo deverbal é mais usado que o verbo-base, apontei o uso de *vigir*. Agora, cabe mencionar que a inexistência de participípios em –edo, por outro lado, dificulta o rastreamento de verbo da 2ª ou da 3ª conjugação a partir dos participípios em –ido nos casos em que o participípio é mais usado que o infinitivo e apontar o uso de *tolhir*. A recapitulação dos verbos que formam os pares (2) e dos verbos que formam os pares em (5) permite constatar que, a despeito do padrão derivacional a que se vinculem, o padrão flexional é de neutralização de todo infinitivo em –er ou –ir em participípios em –ido. A rede é realmente intrincada (cf. Fauconnier & Turner, 2006, acima).

Isso posto, retorno ao estabelecimento de relações entre formas por obra da frequência de uso, levado em consideração o padrão morfológico envolvido. Ao pensar em *arder* e *viger*, me ocorreu que os pares acima são posteriores a um padrão desativado na língua, aquele em que o verbo é em –er e o substantivo é em –or:

(5) <i>arder</i> – <i>ardor</i>	<i>temer</i> – <i>temor</i>
<i>feder</i> – <i>fedor</i>	<i>tremer</i> – <i>tremor</i>
<i>ferver</i> – <i>fervor</i>	<i>valer</i> – <i>valor</i>
<i>pende</i> – <i>pendor</i>	<i>viger</i> – <i>vigor</i>
<i>saber</i> – <i>sabor</i>	

Trago os pares em (5) para defender que a consciência dos processos oferece mais segurança no

rastreamento de formas a partir de formas que o conhecimento de produto a produto. No quadro do lexicalismo gerativo, Jackendoff (1975) cuidou do menor custo de aquisição de formas que ofereçam maior grau de redundância morfológica ou semântica, em oposição ao maior custo de aquisição de formas que ofereçam menor grau de redundância morfológica ou semântica. O cálculo do custo remete diretamente à distinção entre o arbitrário absoluto e o arbitrário relativo segundo Saussure (1916, segunda parte, cap. VI), que, por sua vez, remete ao princípio da motivação em Lakoff (1987) e ao princípio da motivação maximizada em Goldberg (1995).

Trata-se da mesma questão a receber diferentes respostas teóricas.

A ênfase dada ao léxico pela gramática cognitiva equivale à de certas abordagens gerativas, como, por exemplo, a da Linguística Léxico-Funcional, e o tratamento ao léxico assemelha-se, em vários aspectos, ao da teoria da entrada plena de Jackendoff por meio de regras de redundância (1975). (...) Finalmente, sua ênfase no uso da língua e sua visão de significado lexical apresentam muitos paralelos com as do segundo Wittgenstein (1953).⁹ (Langacker, 1987, p. 5)

A contribuição que identifiquei em Bybee (2010) a essa questão diz respeito à influência da frequência de uso na consciência da redundância ou na motivação, que, já em Jackendoff (1975), se caracteriza como fenômeno cognitivo.

Tendo em vista a frequência de uso, pode-se constatar que a forma *assistir* de fato usada não é a que se relaciona com *assistência*, o que gera um quadro lacunar de relações verbo-substantivo:

V	S
<i>assistir</i>	-
-	<i>assistência</i>

Aliás, o cada vez menor uso do *assistir* que se relaciona com *assistência* torna expletivo o papel distintivo que a preposição *a* desempenha na gramática de poucos usuários da língua, o que contribui para a gradativa supressão da preposição na nuvem de exemplares do *assistir* frequente.

Combinando os pares em (2) com os pares em (5), podemos afirmar que a relação entre as formas é mais viva naqueles que neles, o que pode ser atribuído ao fato de que o processo de normalização em –(nc)ia está ativo enquanto o processo de nominalização em –or se tornou inativo. Passamos, então, a trabalhar com dois fatores: a atividade ou inatividade dos processos, em atenção à observação de Basilio (1987), e o mapeamento de formas mais pela atividade dos processos que pelo conhecimento

⁹ "The emphasis on lexicon in cognitive grammar parallels that of certain generative approaches, e.g. lexical-functional grammar (...), and its treatment is similar in various respects to Jackendoff's full-entry proposal employing lexical redundancy rules (1975). (...) Finally, its emphasis on language use and its view of lexical meaning show many parallels to the later Wittgenstein (1953.)"

de produto a produto, visto que não é a complexidade morfológica dos produtos que interfere na frequência de uso, em atenção a Bybee (2010).

A confluência dos dois fatores deixa *viger* na situação que aqui pretendo caracterizar como de indisponibilidade cognitiva. Se *vigência* já não leva necessariamente a um verbo da 2ª conjugação, o processo que associa verbos em –er a substantivos em –or está morto na língua. Nos termos em que retomo a distinção entre forma livre e forma presa neste artigo, trata-se de uma forma presa, uma vez que nem a forma *vigência* nem a forma *vigor*, estas livres, contribuem para sua identificação.

A maior frequência de *gerência*, *emergência* e *urgência* tende a relegar, respectivamente, *gerir*, *emergir* e *urgir* à mesma condição de forma presa, sobretudo em se tratando de verbos que se caracterizam pela forte restrição formal de serem ensinados como unipessoais segundo a gramática tradicional. Essa restrição se desfaz nos substantivos deverbais na medida em que “(...) a nominalização do verbo nos permite expressar a noção verbal em si, sem as amarras dos requisitos gramaticais do verbo” (Basilio, 2009, p. 42).

Uma evidência do não reconhecimento das relações entre formas, processo aqui enfatizado como prioritário à existência das formas, é o acréscimo de camada morfológica ao substantivo. Retomemos os pares em –er / –or para ilustrar tal evidência. A relação entre *arder* e *ardor*, *feder* e *fedor*, *ferver* e *fervor*, *pende* e *pendor*, *treme* e *tremor* parece viva à consciência dos falantes enquanto a relação entre *saber* e *sabor* e *viger* e *vigor* não. Não me parece dever-se ao acaso que *sabor* e *vigor* – porque formadas a partir de base presa nos termos defendidos neste artigo – motivem a formação de verbos que lhes correspondam, respectivamente, *saborear* e *vigorar*. São esses os verbos que formam pares com *sabor* e *vigor* na língua em uso.

Mantendo o foco em processos, a língua preserva os pares verbo-substantivo a despeito da opacificação causada pelo desuso dos verbos graças à recursividade que existe entre a formação de substantivos a partir de verbos e a formação de verbos a partir de substantivos. A citação de Basilio (2009) acima trata do primeiro desses processos. Já “A formação de verbos a partir de substantivos tem o objetivo de aproveitar a noção expressa pelo substantivo para designar a ação ou processo a ser expresso pelo verbo” (Basilio, 2009, p. 32). Tal processo promove a substituição de verbos mais básicos da segunda e da terceira conjugações por verbos mais complexos da primeira conjugação, uma vez que “As estruturas mais produtivas de formação de verbos por sufixação são as correspondentes à adição de –izar, –ar e –ear a substantivos e adjetivos” (Basilio, 2009, p. 35).

Os verbos *saborear* e *vigorar* são produtos de tal processo, que supre a lacuna deixada pelos verbos mais básicos, *saber* e *viger*, tomados como formas presas nos termos deste artigo. São eles os que ocorrem a falantes mais jovens a quem se pergunte qual o verbo correspondente a *sabor* e *vigor*, respectivamente, e não os que exigiriam um recuo nas camadas morfológicas, o que caracteriza a mudança linguística como sucessivas reorganizações cognitivas.

A mudança linguística não é apenas um fenômeno periférico que pode ser amarrado a uma teoria sincrônica; sincronia e diacronia têm de ser vistas como um todo integrado. A mudança é tanto uma janela para representações cognitivas quanto um criador de padrões linguísticos. Ademais, se virmos a língua da maneira descrita neste livro, como variável e possuindo categorias gradientes, então a mudança se torna uma parte constitutiva do quadro completo.¹⁰ (Bybee, 2010, p. 105)

Entre os pares em (2), a atividade do processo de formação de substantivos a partir de verbos está em tensão com a imprevisibilidade do verbo a partir do substantivo, o que pode explicar que *gerência* esteja evocando mais relação com *gerenciar* que com *gerir*. Os demais substantivos dentre os pares, não atuam como bases para verbos denominais, o que pode ser considerado como evidência da força do processo.

Trata-se de um fator, e não do fator, ou uma língua natural não seria um sistema complexo, nos termos de Castilho (2010). O fato de *dissidência* e *dissidente* serem mais frequentes que *dissidir*, por exemplo, não conduz à formação de *dissidiar*. O verbo, em vez de ser substituído por outro formado camadas morfológicas acima, como *gerenciar* e *tangenciar* em detrimento de *gerir* e *tanger*, respectivamente, é substituído por outro sem nenhuma relação morfológica, como *divergir* ou *discordar*.

Esse ponto dá margem a duas observações importantes. Uma diz respeito à pertinência da distinção entre morfologia (*menino – menina*) e léxico (*homem – mulher*). Distinção, e não cisão, como nos pressupostos da linguística cognitiva! Este artigo se atém à relação entre frequência de uso e padrões morfológicos, não se estende a padrões lexicais. Outra é que a formação de um verbo camadas acima está sendo associado à perda da relação entre o substantivo e o verbo mais básico, o que não exclui nem a hipótese de, ativa a relação, outro se formar nem a hipótese de, inativa a relação, outro verbo não se formar, sem que se invalide a tese de que quanto menos ativa a relação, maior a tendência a que outro verbo se forme.

Ofereço exemplos para assegurar a clareza sobre a segunda observação incidental. Mesmo ativa a relação entre *agir* e *ação* – alta a frequência de *agir* e ativo o padrão V-S(ção) –, forma-se *acionar*. Mesmo inativa a relação entre *dissidir* e *dissidência* – quase absoluto o desuso de *dissidir* ainda que ativo o padrão V-S(ncia) – não se forma *dissidenciar*. No entanto, dado que, em consequência da desativação do padrão –er / –or, *sabor* não remete a *saber*, a despeito de serem formas igualmente frequentes, e que *vigor* não remete a *viger* (na mesma rede em que *vigência* não exclui *vigir* como forma possível), a despeito de serem formas com frequências diferentes, a frequência de *sabor* e *vigor* favorece à formação de *saborear* e *vigorar*, respectivamente, para suprirem a lacuna V (concebida como semântica, e não formal), ainda que o verbo seja mais complexo morfológicamente que o substantivo, e não menos, como no quadro anterior ao da mudança linguística.

10 “Language change is not just a peripheral phenomenon that can be tacked on to a synchronic theory; synchrony and diachrony have to be viewed as an integrated whole. Change is both a window into cognitive representations and a creator of linguistic patterns. Moreover, if we view language in the manner described in this book, as both variable and having gradient categories, then change becomes an integral part of the complete picture.”

Pares infinitivo / particípio também podem servir de dados para este esboço de generalização. Tomemos os seguintes pares:

(6) *tingir – tinto*
pingir – pinto

implicar – implícito
explicar – explícito

Apenas o primeiro par se mantém relacionado na consciência dos falantes. Os demais passaram à condição de dados históricos. Tendo em vista que a relação entre infinitivos e particípios é tão paradigmática quanto a relação entre verbos e substantivos, uma forma participial na origem não percebida como tal no curso da mudança linguística motivará a formação de um infinitivo a partir dela, que, por sua vez, motivará a formação do particípio correspondente. Essa asserção se justifica na medida em que não há *tintar* a partir de *tinto* mas há *pintar* a partir de *pinto*, *implicitar* a partir de *implícito* e *explicitar* a partir de *explícito*.

O Wikcionário já registra *tintar* na língua portuguesa:

Verbo[editar]

tin.tar, transitivo

1. o mesmo que [tingir](#)

2. **cobrir** com tênue camada de **tinta** para **impressão** (a forma tipográfica, pedra ou chapa litográfica)

(<https://pt.wiktionary.org/wiki/tintar>)

A afirmação “o mesmo que tingir” não encontra respaldo em princípio considerado organizador da linguagem:

O Princípio da Não Sinonímia: Se duas construções são sintaticamente distintas, elas têm de ser semanticamente ou pragmaticamente distintas (...). Aspectos pragmáticos das construções envolvem especificidades da estrutura informacional, o que inclui tópico e foco, assim como aspectos estilísticos adicionais da construção, como registro (...).

Corolário A: Se duas construções são sintaticamente distintas e S(semanticamente)-sinônimas, então elas não podem ser P(pragmaticamente) sinônimas.

Corolário B: Se duas construções são sintaticamente distintas e P-sinônimas, então elas não podem ser S-sinônimas.¹¹ (Goldberg, 1995, p. 67)

Goldberg (1995) associa não sinonímia a distinção sintática e a emprega em sua formulação de uma

11 “*The Principle of No Synonymy*: If two constructions are syntactically distinct, they must be semantically or pragmatically distinct (...). Pragmatic aspects of constructions involve particulars of information structure, including topic and focus, and additionally stylistic aspects of the construction such as register (...). *Corollary A*: If two constructions are syntactically distinct and S(semanticly)-synonymous, then they must not be P(ragmatically)-synonymous. *Corollary B*: If two constructions are syntactically distinct and P-synonymous, then they must not be S-synonymous.”

gramática das construções. Neste artigo, associo não sinonímia a distinção morfológica e a emprego em consonância com um modelo baseado no uso. A diferença é estritamente teórica e bastante relevante para o desenho de um quadro epistemológico. Nesse quadro, convém lembrar a dissolução entre semântica e pragmática, que está na base da linguística cognitiva e, a meu ver, se encontra em retrocesso esquemático na formulação do princípio por Goldberg (1995).

Já a afirmação “cobrir com (...) tinta” reforça que “A formação de verbos a partir de substantivos tem o objetivo de aproveitar a noção expressa pelo substantivo (...)” (Basilio, 2008, p. 32). Aproveitar o significado do substantivo implica ser formado depois do substantivo na história da língua.

Do ponto de vista defendido neste artigo, a formação de *tintar* resulta do afrouxamento da relação entre *tingir* e *tinto*, somado aos “propósitos expressivos” (Langacker, 1987, 449) que exigem o aproveitamento do significado de *tinta* no âmbito pragmático. A relação entre *pingir* e *pinto*, por sua vez, se restringe a *Pinto* como sobrenome na história externa da língua. Quando da ocupação da Península Ibérica por mouros, esses receberam tal sobrenome por terem cor mais escura que a dos habitantes anteriores.

O que se pode afirmar sobre todas as formas que convivem na língua em estado de dicionário, ou seja, registradas na língua escrita e apartadas do uso, é que, diferentemente do que se poderia esperar, não são sinônimas. Uma vez inserido o crucial fator uso, a mudança da língua se desloca para a cognição dos falantes:

(...) não é propriamente na estrutura das línguas, ainda que ela possa ser condicionante, mas nos utentes das mesmas que encontraremos respostas para a questão. Mais claramente ainda, a mudança não tem a sua origem dentro da língua, mas no uso da língua, isto é, em factores externos à estrutura da língua. (Soares da Silva, 2006, p. 88)

Recapitulemos *gerir* e *gerenciar*, *saber* e *saborear*, *viger* e *vigorar*, *pingir* e *pintar*, *implicar* e *explicitar*, *explicar* e *explicitar*. Os casos em que há um verbo mais básico e um verbo mais complexo em torno do mesmo substantivo revelam não sinonímia exatamente porque, primeiro, os verbos têm idades diferentes na história interna da língua e, segundo, o verbo mais complexo aproveita a noção expressa pelo substantivo, como observado por Basilio (2009), enquanto o mais básico não. Tendo em vista que “(...) a polissemia é um fenómeno de motivação, que introduz uma certa redundância no léxico mental (...)” (Soares da Silva, 2006, p. 49), o verbo mais complexo se forma a partir de um dos significados ativados por um verbo polissêmico, razão pela qual jamais poderia ser sinônimo do verbo mais básico.

A formação de *valorar* e *valorizar* a partir de *valor* ainda que mantida a relação com *valer* pode ser atribuída a esse mecanismo em que um substantivo polissêmico ativa um ou mais de um significado com maior frequência de uso e sobre esse(s) significado(s) mais frequente(s) se dá o aproveitamento

semântico de que trata Basilio (2009) na formação de verbos a partir de substantivos. Esse mecanismo corrobora a não sinonímia entre o verbo mais básico e o verbo mais complexo.

Neste ponto, me afasto um pouco de Hay (2001, 2002) e de Bybee (2010). Ambas as autoras tomam como mais ou menos frequente uma forma que guarde com outra relação derivacional. Já tive a oportunidade de estender a relação formal à esfera da flexão e a frequência de uso ao âmbito das categorias gramaticais. Agora, desloco a comparação entre frequências de uso do grau de analiticidade para a polissemia.

Primeiro, porque o verbo mais básico e o verbo mais complexo deflagram percursos polissêmicos próprios, dissociados um do outro, o que contribui ainda mais para a não sinonímia. A opacificação do verbo mais básico, por sua vez, chega ao ponto de não transmissão de uma geração a outra, o que gera a situação caracterizada pelas formações em *-nte* no início deste artigo, que as escolas formalistas, comprometidas com o princípio da sincronia, conseguem descrever mas não explicar. A foto sincrônica permite identificar as lacunas formais, sem associá-las ao papel do uso como determinante da mudança linguística, nos termos de Bybee (2010), muito menos entender a dinâmica entre uso e processos cognitivos da qual as lacunas formais são consequências, e não causas.

Segundo, porque a polissemia me parece a chave para aqueles casos em que se forma um verbo mais complexo mesmo quando o verbo mais básico goza de alta frequência e o padrão morfológico está ativo na língua. O exemplo que ofereci foi *agir – ação – acionar*. Depois de trazer a polissemia à discussão, acrescentei *valer – valor – valorizar / valorar*.

Um dos significados produzidos por *ação* diz respeito ao ato de agir, qualquer agir, o que caracteriza esse significado como verbal (cf. Basilio, 1980, 2004). Outro significado produzido por *ação* diz respeito a um tipo de agir na organização social, cultural e histórica que temos, o agir juridicamente, e a forma *ação* é mobilizada para nomear o produto desse agir específico, o processo judicial. A forma *acionar*, por sua vez, se distingue da forma *agir* porque lhe é privativo o significado específico em questão. Daí o uso de *acionar*, não de *agir*, com diferentes grades argumentais, consequentemente: alguém pode acionar alguém, mas não agir alguém.

O significado prototípico de *valer* é ter valor. Isso faz com que o verbo seja inacusativo. A organização social, cultural e histórica leva os indivíduos a atribuir valor, o que motiva a formação de um verbo transitivo. Só que *valor* também é polissêmico: pode-se tratar de propriedade ou de juízo, o que me parece motivar a formação de *valorar* e de *valorizar*, respectivamente. O verbo que indica atribuição de juízo de valor, *valorar*, por sua vez, se insere no campo da subjetivização (cf. Soares da Silva, 2006), o que deflagra um percurso polissêmico próprio e ratifica o princípio da não sinonímia.

Outros exemplos podem ilustrar o papel da figuratividade da configuração da polissemia, tal como concebida pela linguística cognitiva, e ajustar no deslocamento do foco que proponho da frequência de uso de formas para a frequência de uso de significados indicados pelas formas. Desse modo,

a insensibilidade da frequência de uso ao grau de complexidade morfológica das formas dá lugar à sensibilidade da frequência de uso à polissemia das formas, assim como a frequência de uso de certos significados em relação à de outros é associada à frequência de uso de categorias gramaticais, resultando na caracterização de formas livres como cognitivamente disponíveis e na de formas presas como cognitivamente indisponíveis.

CONCLUSÃO

A distinção entre formas livres e formas presas, estabelecida pelo estruturalismo norte-americano, é clássica em morfologia. Teorias formalistas como aquela pressupõem que a semântica tem papel interpretativo, subsidiário à combinação de formas. A recapitulação da distinção neste artigo se dá à luz da linguística cognitiva, teoria que pressupõe que a semântica motiva a gramática, o que desloca as formas do lugar de ponto de partida ao de ponto de chegada.

Substituída a lente teórica, a distinção entre formas livres e presas deixa de ser uma questão de existência ou inexistência de formas num determinado recorte sincrônico e passa a ser uma questão de atividade ou inatividade do processo de estabelecimento de relações entre formas, no qual interferem a frequência de uso de significados associados a formas, e não das formas; o percurso polissêmico, e não a configuração morfológica, das formas; e a frequência de uso de padrões gramaticais, que podem ser ativados ou desativados no curso de mudança da língua.

Nesse quadro, que reputo mais fiel à intrincada rede de significados lexicais, formas existentes que gozam de baixa frequência de uso são gradativamente opacificadas à consciência dos falantes, a ponto de merecer o status de formas presas. Enquanto relegadas à condição de registro escrito ou restrito a algum gênero, o acesso a tais formas se dá pela via do letramento, e não da aquisição da linguagem. Uma vez ampliado o lapso temporal, tais formas já não são transmitidas e chegam ao ponto que as teorias formalistas capturam, o das lacunas formais.

A redefinição de formas livres e formas presas tal como proposta neste artigo se alinha à concepção de formação de palavras como fenômeno semântico com repercussão morfológica, e não como fenômeno morfológico com repercussão semântica. Tal concepção interdita afirmar, por exemplo, que o adjetivo *acadêmico* deriva do substantivo *academia*. Primeiro, é preciso reconhecer a polissemia de *academia* e o uso de *acadêmico* vinculado ao significado relativo a universidade, e não ao lugar onde se pratica exercício. Com isso, afixos, formas presas, deixam de ser considerados formas que se adjungem a formas e passam a ser considerados formas que indicam qual significado em uso é selecionado da base. Diga-se o mesmo dos sufixos *-or* e *-ivo* em relação a diferentes significados que compõem a polissemia de *executar*, e assim por diante. A concepção tem alcance geral ao fenômeno da formação de palavras.

Igual primazia tem a polissemia na redefinição de formas livres e presas. Bybee (2010) assume de Hay (2001, 2002) a comparação entre frequências de uso de formas e avança na discussão quanto à gradência da analiticidade e da composicionalidade. Já neste artigo, desloco a comparação para

frequências de uso dos significados que constituem a polissemia das formas e subscrevo que “A mudança é tanto uma janela para representações cognitivas quanto um criador de padrões linguísticos” (Bybee, 2010, p. 9).

Os padrões enfocados são a formação de substantivos a partir de verbos e a formação de verbos a partir de substantivos. A desativação do primeiro motiva a ativação do segundo e gera pares V-S em que o verbo é morfologicamente mais complexo que o substantivo (*vigorar – vigor* em vez de *viger – vigor*, *gerenciar – gerência* em vez de *gerir – gerência* etc.), preservada a hipótese de formação de novos substantivos a partir dos verbos mais complexos (*vigorar – vigoração*, *gerenciar – gerenciamento* etc.), de formação de novo verbo a partir de tais substantivos (*vigoração – vigoracionar*, *gerenciamento – gerencionamentar* etc.), o que se entende por recursividade, só que em bases semânticas, e não formais.

Os verbos mais básicos, por sua vez, vão sendo gradativamente retirados de circulação e, por isso mesmo, assumem o status de formas presas nos termos deste artigo. A mudança linguística, por sua vez, compensa a perda dessas formas e sua polissemia pela formação de outras, o que pressupõe também o alcance da frequência de uso, já estendida das formas aos significados, aos padrões gramaticais envolvidos.

REFERÊNCIAS

BASILIO, Margarida. *Estruturas lexicais do português*. Petrópolis: Vozes, 1980.

BASILIO, Margarida. *Teoria lexical*. São Paulo, Ática, 1987.

BASILIO, Margarida. Polissemia sistemática em substantivos deverbais. In: Roberta Pires de Oliveira & Apóstolo Nicolacópulos (Org.). *Semantics: Lexicon, Grammar and Use*, nº 47. Florianópolis: Editora da UFSC, 2004.

BASILIO, Margarida. Substantivação plena e substantivação precária. *Diadorim*, vol. 4, 2008.

BASILIO, Margarida. *Formação e classes de palavras no português do Brasil*. São Paulo: Contexto, 2009.

BLOOMFIELD, Leonard. *Language*. Chicago: The University of Chicago Press, reimpressão de 1997 [1933].

BYBEE, Joan. *Phonology and language use*. Cambridge Studies in Linguistics 94, Cambridge University Press, 2001.

BYBEE, Joan. *Language, usage, and cognition*. Cambridge, Cambridge University Press, 2010.

- CASTILHO, Ataliba T. de. *Nova gramática do português brasileiro*. São Paulo, Contexto, 2010.
- CHOMSKY, Noam. Remarks on nominalization. In: *Readings in English transformational grammar*. Waltham: Ginn & Co, 1970.
- CRISTÓFARO-SILVA, Thaís. Descartando fonemas: a representação mental na fonologia de uso. In: Dermeval da Hora & Gisela Collischonn (Org.). *Teoria linguística: fonologia e outros temas*. UFPB, Editora Universitária, 2002.
- FAUCONNIER, Gilles & TURNER, Mark. Conceptual integration networks. In: Dirk Geeraerts (Ed.). *Cognitive linguistics: basic readings*. Berlin: Mouton de Gruyter, 2006.
- GOLDBERG, Adele. *Constructions: a construction approach to argument structure*. Chicago: The University of Chicago Press, 1995.
- GONÇALVES, Carlos Alexandre. Usos morfológicos: processos marginais de formação de palavras em português. *Gragoatá*, vol. 21, 2006.
- HAY, Jennifer. Lexical frequency in morphology: is everything relative? *Linguistics* 39, 2001.
- HAY, Jennifer. From speech perception to morphology: affixordering revisited. *Language* 78, 2002.
- JACKENDOFF, Ray. Morphological and semantic regularities in the lexicon. In: William Bright *et alii* (Ed.). *Language: Journal of the Linguistic Society of America*, vol. 51, nº 3. Baltimore: Waverly Press Incorporate, 1975.
- LANGACKER, Ronald. *Foundations of cognitive grammar*. Vol. I: theoretical prerequisites. Stanford: Stanford University Press, 1987.
- LEMOS DE SOUZA, Janderson. *A distribuição semântica dos substantivos deverbais em -ção e -mento no português do Brasil: uma abordagem cognitiva*. UFRJ, tese de doutorado, 2010.
- LEMOS DE SOUZA, Janderson. Restrições sintáticas em torno dos substantivos deverbais no português do Brasil. In: Augusto Soares da Silva; Amadeu Torres; Miguel Gonçalves (Org.). *Línguas pluricêntricas: variação linguística e dimensões sociocognitivas*. Braga: ALETHEIA – Associação Científica e Cultural, 2011.
- LEMOS DE SOUZA, Janderson. Alinhamento, balanceamento, direção e suspensão: constatações sobre os substantivos deverbais no português brasileiro. In: *Linguística*, v. 8, nº 1, 2012.
- LEMOS DE SOUZA, Chunks em moldes: hipótese de diálogo entre Bybee e Booij. In: *Cadernos do*

NEMP, v. 6, 2015.

LYONS, John. *Introduction to theoretical linguistics*. Cambridge: Cambridge University Press, 1968.

SAUSSURE, Ferdinand de. *Curso de lingüística geral*. 20 ed. Tradução: Antônio Chelini, José Paulo Paes e Izidoro Blikstein. São Paulo, Cultrix, 1997 [1916].

SOARES DA SILVA, Augusto. *O mundo dos sentidos em português: polissemia, semântica e cognição*. Coimbra: Almedina, 2006.

SWEETSER, Eve. *From etymology to pragmatics: metaphorical and cultural aspects of semantic structure*. Cambridge studies in linguistics 54, Cambridge University Press, 1990.

Recebido em 30/03/2016

Aceito em 16/05/2016

MESCLAGEM CONCEPTUAL EM PIADAS CURTAS

Elyssa Soares Marinho (UFRJ)¹ e Lilian Ferrari (UFRJ)²

RESUMO

Neste artigo, buscou-se identificar e analisar os elementos linguísticos que ativam o efeito cômico e constroem cognitivamente o significado em piadas curtas, a partir da Teoria da Mesclagem Conceptual (FAUCONIER & TURNER, 2002) e da noção de Mudança de *Frame* (COULSON, 2001). Os resultados demonstram que as piadas curtas, especificamente as adivinhas, evocam dois domínios diferentes cujo elemento em comum é a similaridade fonética entre duas construções.

PALAVRAS - CHAVE: Mesclagem Conceptual; mudança de *Frame*; piadas curtas

ABSTRACT

In this paper, we sought to identify and analyze the linguistic elements that contribute to comic effects by prompting meaning construction in short jokes. The research is based on Conceptual Blending Theory (FAUCONIER & TURNER, 2002), as well as on the notion of Frame-shifting (COULSON, 2001). The results demonstrate that short jokes, specifically riddles, evoke two different domains whose common element is the phonetic similarity between two constructions.

KEYWORDS: Conceptual Blending; Frame-Shifting; short jokes

INTRODUÇÃO

A Linguística Cognitiva se baseia na hipótese de que a linguagem humana é um instrumento de organização e processamento mental que concebe o significado como construção cognitiva num movimento contínuo de categorização do mundo a partir de modelos compartilhados e de crenças

1 Universidade Federal do Rio de Janeiro. lyssamarinho@hotmail.com

2 Universidade Federal do Rio de Janeiro.

socioculturais (FERRARI, 2011). A mente humana é capaz de elaborar e construir o significado a partir de habilidades imaginativas.

Fauconnier & Turner (2002) defendem que o que está por trás da forma não é apenas algo já definido ou pronto, mas o retrato de como a mente humana é poderosa e inventiva na construção do significado. Alguns dos significados mais simples produzidos pela mente são frutos de operações inconscientes, complexas e misteriosas que aparecem na investigação da linguagem criada e reproduzida pelo ser humano.

Essas operações mentais são altamente imaginativas e produzem nossa compreensão sobre a identidade, as similaridades ou diferenças de expressões e entidades. A capacidade cognitiva do ser humano em construir e reconstruir o significado é governada por princípios coerentes que são restringidos pelas nossas atividades sociais, culturais, físicas e cognitivas.

Sendo assim, tomando o aspecto imaginativo e criativo da mente humana, este artigo tem como objetivo identificar e investigar os elementos linguísticos que ativam o efeito cômico de uma piada curta. A partir da hipótese de que as piadas ativam Mudança de *Frames* associadas a processos de Mesclagem Conceptual, evidenciando grande flexibilidade no processo de interpretação, pretende-se investigar as marcas linguísticas que fornecem ao ouvinte/leitor um deslocamento de *frames*, e representar os mecanismos de Mesclagem Conceptual que, por sua vez, são responsáveis pela emergência de sentido na piada, produzindo um efeito essencial para a compreensão do humor.

A partir dos objetivos propostos nesta seção, este artigo apresenta a Teoria dos Espaços Mentais e da Mesclagem Conceptual (FAUCONNIER & TURNER, 2002) que fornecem o embasamento teórico para a investigação proposta. O processo de Mesclagem Conceptual é construído a partir da integração de espaços mentais cujos conteúdos relacionam uma rede de conceitos advindos de esquemas conceptuais como os *Frames* e os Modelos Cognitivos Idealizados. (FAUCONNIER, 1994, 1997)

Em seguida, apresenta-se o processo de Mudança de *Frames* nas piadas, estudado por Coulson (2001). A autora ressalta que as piadas são construídas para violar as expectativas do ouvinte, explorando inferências decorrentes do conhecimento de cenários típicos, e em seguida, promovendo a alteração dessas inferências iniciais devido à Mudança de *frame*, evidenciando a flexibilidade no processo de interpretação.

A análise é composta por quatro piadas curtas, do tipo adivinhas ou charadas. A investigação do caminho cognitivo da interpretação envolve etapas importantes como a identificação do contexto da piada, o recrutamento semântico das construções, a ativação de *frames*, entre outros que levarão o ouvinte/leitor ao *punchline*³ da piada, evidenciada na estrutura emergente, onde o cômico se consolida. Portanto, o modelo diagramático da Mesclagem Conceptual é aplicado, revelando cada etapa da construção da comicidade.

3 *punchline* é a segunda parte da piada em que se faz uma reinterpretação do cenário oferecido pela primeira parte, instaurando o cômico. Tal reinterpretação é viabilizada por diferentes regras cognitivas. (RITCHIE, 1999)

Na última seção, apresentam-se os resultados da análise, demonstrando que na emergência de sentidos em piadas curtas, a sequência fonética idêntica ou similar funciona como gatilho para o efeito cômico, além de ser um processo construído *online* a partir do qual o ouvinte/leitor ativa, de maneira seletiva e simultânea, elementos advindos de seus conhecimentos prévios para alcançar o efeito cômico.

2. ESPAÇOS MENTAIS, *FRAMES* E MCIS

A principal teoria que fundamenta esta investigação é a Teoria dos Espaços Mentais e, mais especificamente, a Teoria da Mesclagem Conceptual (FAUCONNIER, & TURNER, 2002), que se concentram nos aspectos criativos da mente humana, descrevendo, em seu modelo, os movimentos cognitivos efetuados na busca da construção de sentidos.

Após os anos 70, com a mudança de um enfoque objetivista nos estudos da linguagem para um enfoque cognitivista na construção do significado, Fauconnier (1994, 1997) criou um modelo de estudo rico sobre as conexões de domínios cognitivos nas línguas naturais, com uma pesquisa interdisciplinar. A Teoria dos Espaços Mentais oferece um modelo de conexão entre semântica e cognição que permite focar questionamentos importantes em uma gama de problemas linguísticos como a opacidade das questões referenciais, as escolhas de construções gramaticais, de tempo verbal ou formas pronominais.

Assim, a construção de sentido ocorre na medida em que o discurso é desenvolvido, envolvendo diferentes conexões cognitivas entre domínios chamados de espaços mentais. Fauconnier & Sweetser (1996) apontam que a cognição humana é configurada contextualmente; por isso, é possível analisar quais conexões cognitivas entre domínios permitem, por exemplo, o uso de determinada palavra ou expressão que pertence a um espaço como um gatilho (*trigger*) para identificar ou referir-se a uma entidade alvo (*target*) localizada em outro espaço mental. Dessa forma, a análise do significado emergente é feita a partir do exame das diferentes conexões que nossa mente pode fazer tomando os elementos linguísticos produzidos no discurso.

A construção de sentido, então, é um processo mental complexo que ocorre entre domínios quando pensamos e nos comunicamos. Os espaços mentais incluem o conhecimento cognitivo e conceptual que o ser humano possui a partir das experiências culturais e sociais adquiridas ao longo de sua vida. A estrutura de um espaço mental herda informações de esquemas conceptuais e *frames*. Sendo assim, os elementos desses espaços se encaixam em modelos cognitivos idealizados (MCI) que são importados do conhecimento prévio durante um dado discurso.

O conhecimento adquirido ao longo das experiências culturais e sociais dos seres humanos é armazenado em estruturas permanentes na memória e forma uma rede conceptual, denominada *frame*. Fillmore (1982) afirma que o significado das palavras está subordinado ao *frame* em que elas são acessadas, de modo que a interpretação das palavras requer o acesso às estruturas do conhecimento que relacionam os elementos e itens lexicais referentes a uma experiência humana ou domínio social específico naquele contexto.

Ao considerar, por exemplo, o evento de “compra e venda”, é possível identificar nesse *frame* o comprador, o vendedor, o produto, o custo, a forma de pagamento, etc. Na frase “Maria comprou uma blusa de Ana”, pode-se identificar cada elemento num determinado espaço mental, sendo Maria, a compradora; Ana, a vendedora e blusa, o produto. Tais elementos são projetados a partir do *frame* de “compra e venda” preenchendo a configuração apropriada dos papéis e valores de cada um.

A partir da noção de *frame*, Lakoff (1987) menciona que os espaços mentais são estruturados por Modelos Cognitivos Idealizados (MCIs) que organizam o conhecimento adquirido a partir das experiências cotidianas, culturais e sociais. Os MCIs podem advir também de um conhecimento compartilhado em determinada comunidade de fala, possibilitando que, em um dado evento enunciativo, os interlocutores troquem informações implícitas, que só poderão ser acessadas pelos indivíduos que compartilham os mesmos *frames*.

MCI, portanto, é configurado como uma estrutura que armazena o conhecimento adquirido de maneira mais complexa e organizada do que *frames* e pode ser estruturado por três princípios: a estrutura proposicional, esquemas imagéticos e projeções metafóricas e metonímicas (FERRARI, 2011).

A estrutura proposicional apresenta o mesmo conceito da noção de *frame* mencionada anteriormente. Esquemas Imagéticos são fundamentados por experiências sensório-motora e espacial dos seres humanos. A estrutura conceptual dos MCIs apresenta esquemas imagéticos dos tipos contêiner, parte-todo, frente-trás, etc. As projeções metafóricas e metonímicas mostram que os MCIs são fonte de correspondências metafóricas como em *estados são locais*, no exemplo “Ele ficou nas nuvens com o prêmio” ou metonímicas que mostram uma *relação de contiguidade*, como no exemplo “Li Machado de Assis” (entende-se a obra de Machado de Assis).

Assim, as noções de *Frame* e MCI assumem um papel de destaque na construção de sentido, pois explicam o motivo pelo qual, em uma dada expressão, a interpretação envolve mais informação do que é possível visualizar a partir dos elementos linguísticos. Os MCIs, então, relacionam-se com os espaços mentais a partir de projeções entre domínios, alimentando-os de informações. Na construção de sentido há, ainda, a possibilidade de ocorrer Mesclagem Conceptual (*Blending*), que integra parcialmente os elementos de espaços mentais específicos.

3. MESCLAGEM CONCEPTUAL E A CONSTRUÇÃO DO SIGNIFICADO

A Mesclagem Conceptual é uma operação cognitiva que consiste na integração parcial de elementos de dois ou mais domínios em uma estrutura denominada espaço-mescla. Essa noção foi introduzida por Fauconnier (1997) que demonstrou ser este um processo dinâmico que pode ocorrer repetidamente em uma mesma rede conceptual. Para ilustrar o processo de mesclagem, utilizam-se diagramas que representam os espaços mentais com círculos e seus elementos com pontos dentro desses círculos. As projeções inter-domínios são representadas por linhas e as estruturas de MCI e *frames* são representadas tanto por retângulos fora dos círculos, como por ícones dentro dos mesmos, podendo, ainda, ser facultadas nas representações diagramáticas. (FAUCCONNIER, & TURNER, 2002)

Os espaços mentais que fornecem os elementos para o processo de mesclagem são chamados de *Input*. É preciso, no mínimo, dois *inputs* em uma mesclagem. A projeção entre os *inputs* conectam as contrapartes, ou seja, os elementos que possuem traços em comum. Essas conexões geram, assim, um terceiro espaço conhecido como espaço genérico. Esse espaço reflete a estrutura abstrata em comum entre os inputs, contendo elementos genéricos de suas contrapartes.

Assim, um quarto espaço surge, o espaço-mescla. Os espaços de *input* projetam alguns de seus elementos neste único espaço formando uma mescla. Tais elementos, advindos dos *inputs*, podem possuir contrapartes ou não e, ainda, podem ser projetados separadamente ou fundidos em um só. O espaço-mescla forma uma estrutura emergente que não havia nos espaços de *input*. Essa estrutura é construída por *composição* dos elementos que se relacionam na mescla; por *completamento*, que traz uma estrutura adicional à mescla, complementando-a com informações que não foram adicionadas nos *inputs*, advindas de algum *frame*; e por *elaboração*, que pode desenvolver a mescla a partir de um trabalho imaginativo regido por princípios próprios. A estrutura emergente pode ser representada por um retângulo no interior do círculo do espaço mescla. Essa operação está representada na figura a seguir:

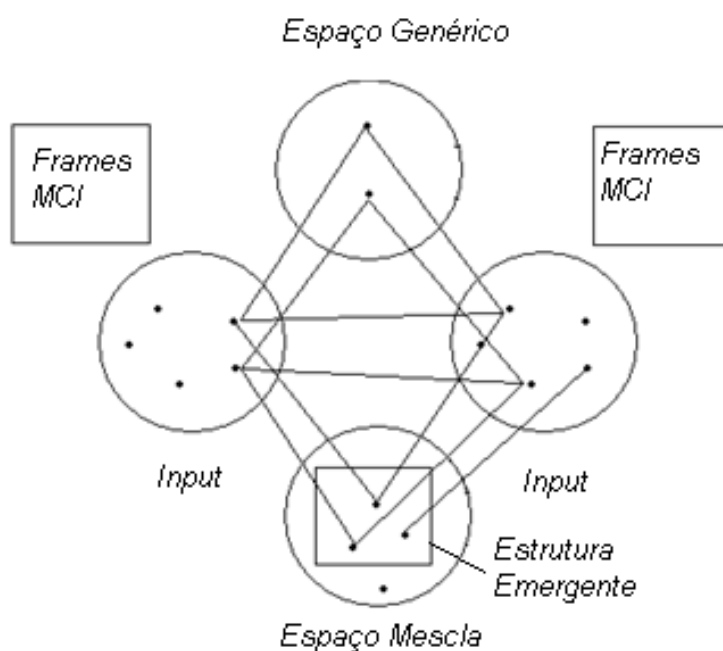


Figura 1 - Processo de Mesclagem Conceptual

Fauconnier & Turner (2002) utilizam uma analogia com a biologia evolucionária para explicar a construção do significado na Mesclagem Conceptual:

In crucial respects, the construction of meaning is like the evolution of species. It has coherent principles that operate all the time in an extremely rich mental and cultural world. Many, many, many new integrations are attempted and explored in an individual's backstage cognition

and in interchange by members of culture, and most of them never go anywhere. (FAUCONNIER, & TURNER, 2002, p. 309)

A capacidade cognitiva do ser humano em construir e reconstruir o significado é governada por princípios coerentes que são restringidos pelas nossas atividades sociais, culturais, físicas e cognitivas. Muitas possibilidades de significações e integrações conceptuais são possíveis na cognição humana, entretanto, nem todas estas possibilidades evoluem. Quando a integração conceptual ocorre efetivamente, produz na mescla uma estrutura emergente mais rica e elaborada que as estruturas presentes nos *inputs*, podendo ser reutilizada pela cognição, caracterizando a Mesclagem Conceptual como um processo dinâmico e flexível na construção do conhecimento.

Associado à Mesclagem Conceptual, há o processo de Mudança de *Frames* que corrobora para emergência de sentidos em piadas, como será abordado a seguir.

4. MUDANÇA DE *FRAMES* E PIADAS

Em estudo mais amplo sobre a atuação de processos de Mudança de *Frame* (*frame-shifting*) e Mesclagem Conceptual na construção do significado, Coulson (2001) inclui a análise de piadas. A autora ressalta que as piadas são construídas para violar as expectativas do ouvinte, explorando inferências decorrentes do conhecimento de cenários típicos, e em seguida, promovendo a alteração dessas inferências iniciais devido à Mudança de *Frame*. Além disso, as piadas podem ativar a alteração de *frames* associados a processos de Mesclagem Conceptual, mostrando a flexibilidade na interpretação desse gênero.

O princípio central da abordagem desenvolvida em sua pesquisa é que a construção do significado não é uma simples manipulação das representações pré-existentes na memória, mas um processo ativo no qual falantes integram continuamente aspectos perceptuais e conceptuais com as informações disponíveis na memória de longo prazo. Assim, suas análises mostram que na dinâmica da construção dos sentidos, podemos promover Mudanças de *Frame* que revelam a natureza das informações recrutadas em nosso conhecimento prévio. Trata-se de um tipo de revisão conceptual em que há uma operação de reanálise semântica no processo de reorganização de uma informação já existente em um novo *frame*.

Acredita-se que o significado de um dado item lexical depende do contexto em que ele está inserido; por isso, a mensagem reflete o *frame* motivado pela linguagem e a sua reanálise é a chave para a interpretação de significados não convencionais:

The contribution of an appropriate word meaning depends upon the context in which it appears. Moreover, the message-level meaning is influenced by the particular words that occur in the sentence. Lexical reanalysis can trigger pragmatic reanalysis that results in substantial

alteration to the message-level representation. Similarly, the choice of a new frame may change how we interpret the meanings of previously encountered words. (COULSON, 2001, p. 69)

Uma das piadas analisadas em seu trabalho requer que o ouvinte faça uma busca nas informações contidas em sua memória para reinterpretar o que ouviu inicialmente: “Quando pedi ao atendente do bar algo gelado e cheio de rum, ele recomendou sua esposa” (*When I asked the bartender for something cold and full of rum, He recommended his wife*) (COULSON, 2001, p.57). A autora explica que a informação dada na frase “algo gelado e cheio de rum” é reinterpretada quando o ouvinte se depara com a segunda parte da piada “ele recomendou sua esposa”. Primeiramente, o ouvinte cria um cenário típico de um bar no qual “algo gelado e cheio de rum” preenche as características de uma bebida encontrada em bares. Porém, na segunda frase da piada, o objeto recomendado é a esposa do atendente. Assim, é necessária a criação de um novo cenário evidenciado pela correspondência entre o termo “gelado” e a resultante interpretação de “frígida”. A correferência entre “algo gelado” e “cheio de rum” com “sua esposa” requer uma revisão da suposição de que o objeto recomendado será preenchido pela bebida. Além disso, como as características “gelado” e “cheio de rum” são favoravelmente aplicáveis à bebida, mas não a uma pessoa, a fala do atendente pode ser reinterpretada como um insulto.

Portanto, a Mudança de *Frame* parece ser motivada por uma violação das restrições de preenchimento das lacunas em um determinado *frame*. Ao invés de falhar na interpretação desse tipo de exemplo, os interlocutores resolvem esses casos criando novos *frames* nos quais essas lacunas podem ser preenchidas. A piada exposta acima é um exemplo de reanálise semântica num âmbito mais geral, demandada pelo fato de que o significado do enunciado não é computado a partir das representações linguísticas, mas motivado por eles.

5. ANÁLISE

Vários estudiosos, como Freud (1969), Bergson (1980), Raskin (1985) e Possenti (1998), já exemplificaram que o material linguístico da piada é justamente o caráter polissêmico das palavras, seus usos ambíguos que evidenciam o duplo-sentido e o trocadilho em todo enunciado de humor. Essas expressões linguísticas podem acessar múltiplos referentes através de conexões estruturais que apresentam papéis sociais e valores culturais.

Assim, tem-se na piada, o reflexo de diversas manifestações culturais e ideológicas de cada comunidade de fala. Por ser um gênero textual presente nas mais variadas culturas e esferas comunicativas, lhe é conferido certo caráter universal. Além disso, trata-se de um material autêntico de uso corrente em uma determinada comunidade de fala, que revela questões controversas inerentes ao ser humano, como preconceito, racismo e indiferença.

A análise que segue é composta por quatro piadas curtas, do tipo adivinhas e charadas, que evocam dois domínios diferentes cujo elemento em comum é a similaridade da sequência fonética de duas ou mais construções. A piada (1) exemplifica essa similaridade:

(1) Qual o estado do Brasil que queria ser carro?

- Sergipe

A piada acima se mostra bastante simples e rápida e, para além das questões que envolvem a oralidade e a escrita, observa-se que o efeito cômico está na percepção da homofonia entre a palavra Sergipe e a construção “ser jipe”. Trata-se de um exemplo claro de Mesclagem Conceptual entre esses dois termos que são ativados a partir da pergunta “Qual estado do Brasil que queria ser carro?”. O *frame* de geografia do Brasil que, entre vários outros elementos, contém também os nomes dos estados, capitais e cidades, e o *frame* de carros, em que se encontram as marcas, diferentes modelos, cores, acessórios.

Estes, então, compartilham apenas uma característica em comum: a sequência fonética idêntica de seus elementos. Em função dessa característica, um quarto espaço é projetado: o espaço-mescla, de onde se apreende a Estrutura Emergente dessa piada, que constata a similaridade fonética entre “Sergipe”, o estado, e a construção “ser jipe” – que inclui o carro. O efeito cômico é apreendido no fato de que a resposta da charada é fornecida por uma mesma expressão que integra dois domínios que, a priori, são independentes e, também, pelo fato de que há homofonias não usualmente percebidas em português.

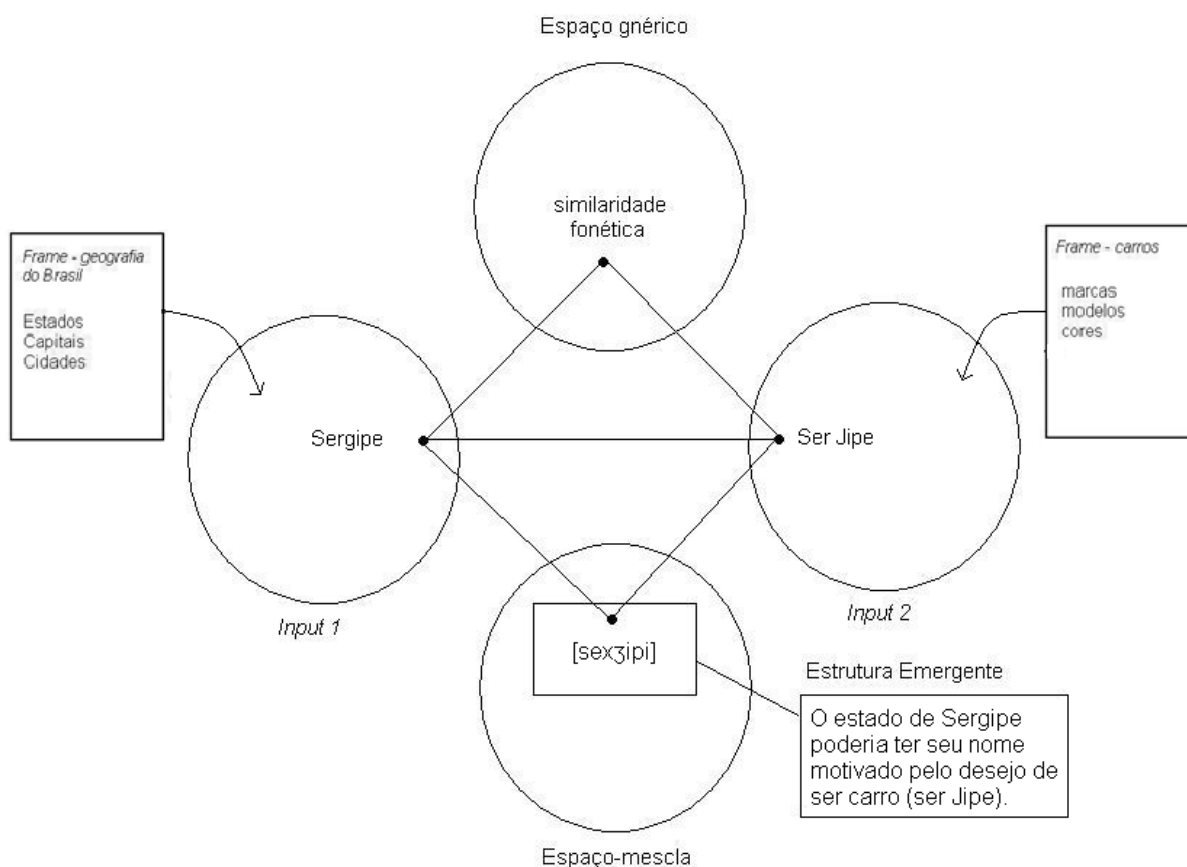


Figura 2: Representação diagramática da piada (1)

Na piada (2), a seguir, tem-se mais um caso de similaridade fonética:

(2) Quem é a mãe do mingau?

- A mãezena

E quem é a avó do mingau?

- A véia Quaker

A piada (2) também é construída a partir de charadas como em (1). Duas perguntas interligadas levam o ouvinte/leitor a acessar o *frame* de relações familiares. Considerando os efeitos prototípicos estudados por Lakoff (1987) da palavra mãe, pode-se considerar que seu conceito se baseia na combinação de diferentes *frames* individuais e, quando há convergência de todos esses modelos, há o conceito ideal de mãe. Em (2), apenas um deles, o modelo de nascimento – a pessoa que dá à luz – se aplica ao contexto da piada, pois se considera a “mãe do mingau” a farinha que origina o mingau. Estendendo o modelo de nascimento para a palavra avó, tem-se aquela que deu à luz a mãe, sendo a progenitora da criadora do mingau, ou seja, a “avó do mingau” (que é a aveia Quaker, um outro tipo de farinha). Percebe-se, também que, há uma função pragmática que estabelece uma relação de identidade entre avó e velha, no *input* 1, assumindo que, neste contexto, “avó é velha (veia)”.

Assim, evidencia-se um processo de Mesclagem Conceptual entre dois domínios distintos. O primeiro *input* apresenta uma estrutura referente às relações familiares, e o segundo *input* apresenta uma estrutura referente às marcas de farinhas para mingau. A resposta das charadas aparece na Estrutura Emergente da mesclagem, revelando elemento cômico da piada e a presença de sequências fonéticas parcialmente semelhantes em português. Há a compressão das palavras advindas dos *inputs* 1 e 2, criando novos termos, “Mãezena” e “A véia Quaker”, evidenciando uma fusão na mescla. As expressões “mãezena” e “a véia Quaker” são o gatilho para que o ouvinte/leitor acesse seu conhecimento prévio sobre marcas de farinha para mingau. Sendo assim, a similaridade fonética entre mãe e Maizena permite que a relação de causa-efeito entre Maizena e mingau seja concebida a partir do *frame* de relações de parentesco. O mesmo ocorre em relação à aveia Quaker e mingau.

Nesta piada, o conhecimento prévio assume grande destaque uma vez que se não ocorrer a identificação da analogia entre as respostas das piadas e as marcas Maizena e Aveia Quaker não haverá comicidade; portanto, se o ouvinte/leitor não acessar o *frame* de marcas de alimentos por não compartilhar tal conhecimento com o seu interlocutor, não haverá a emergência do sentido.

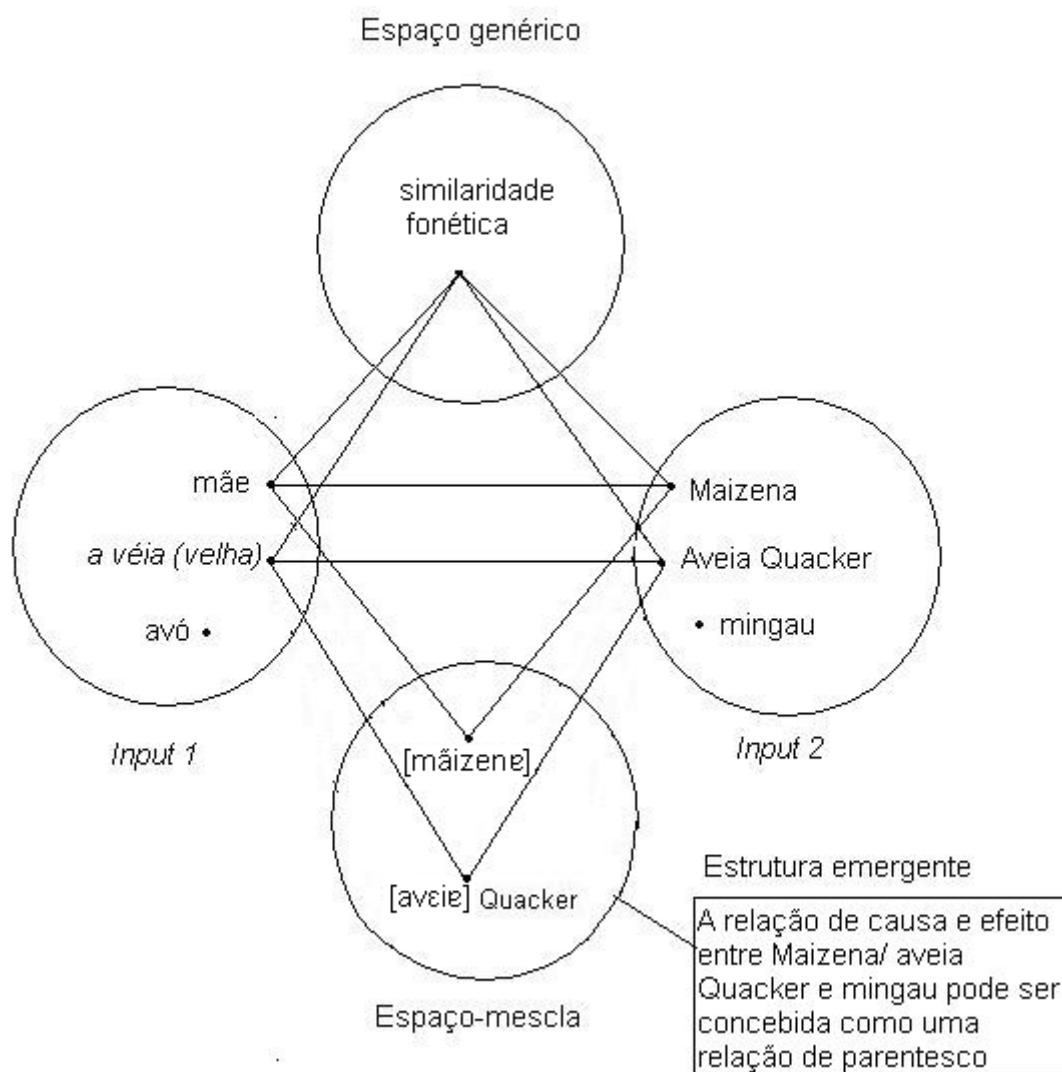


Figura 3: Representação diagramática da piada (2)

As piadas (1) e (2) mostram que a compressão que surge na estrutura emergente do espaço mescla é o gatilho para o efeito cômico. Existe uma composição dos elementos advindos dos domínios, assim como ocorre em (3) e (4), a seguir.

(3) O que o passarinho falou para a passarinha?

-Quer danoninho?

Ferrari (2011) lembra que a construção da estrutura emergente por composição cria relações que podem estar ou não disponíveis nos domínios anteriores à mescla. Em (3), a expressão “Danoninho” evoca dois domínios: um tipo de iogurte (*input 1*) e, também uma “cantada” do passarinho “quer da(r) no ninho?” (*input 2*).

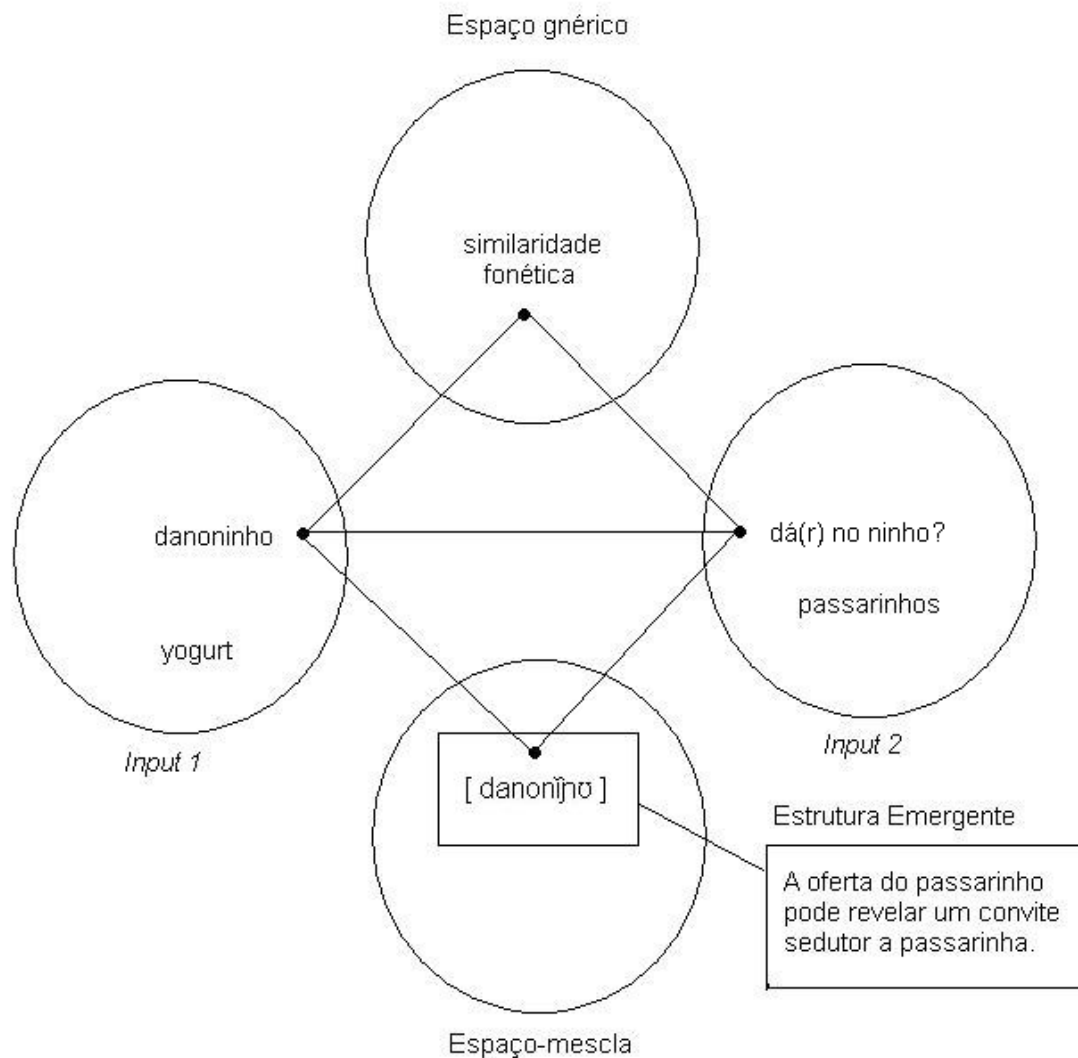


Figura 4: Representação diagramática da piada (3)

(4) O que o Seno respondeu quando o cosseno bateu na porta do banheiro?

- Tangente

Em (4), a mescla é alimentada pelo *input 1* que trata de uma relação social específica: bater na porta do banheiro para saber se está ocupado. A resposta default para esta situação seria “tem gente” quando este está ocupado; já o *input 2* fornece um elemento que constitui um *frame* matemático: seno, cosseno e tangente. Assim, na mescla há a compressão de “tem gente” e “tangente” resultando na resposta da charada que evoca o cômico, por apontar a similaridade fonética entre as duas expressões.

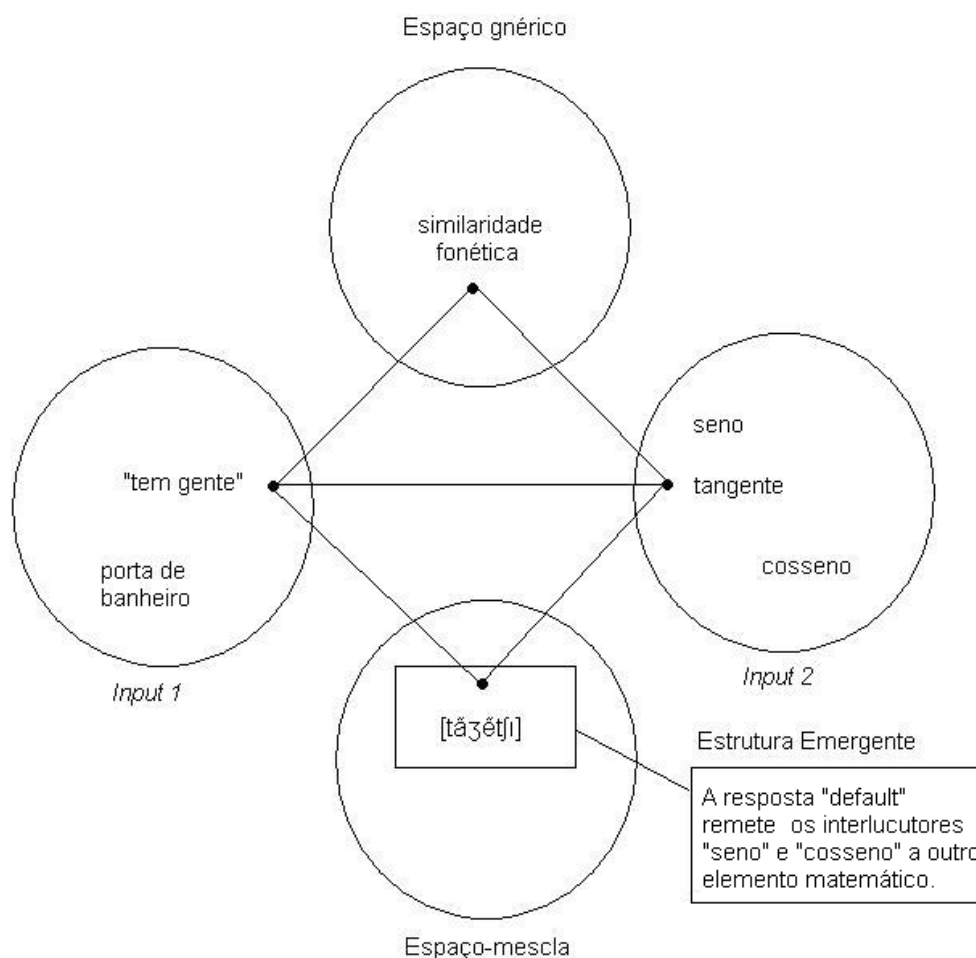


Figura 5: Representação diagramática da piada (4)

O novo conceito deflagrado pela resposta da adivinha, tanto em (1) quanto em (3), foi construído a partir dos elementos fornecidos pelos *inputs* 1 e 2 e nesta projeção, o material linguístico é preservado, permitindo que a mescla desenvolva uma estrutura emergente em que há apenas uma expressão que comprime elementos advindos dos dois *inputs* iniciais. Assim, a estrutura emergente fornece uma construção cuja ambiguidade é, então, o gatilho para o efeito cômico.

Nas piadas (2) e (4), percebe-se que além da Mesclagem Conceptual, há também uma Mesclagem Formal que cria uma nova construção na estrutura emergente a partir de uma compressão do material linguístico advindo dos dois *inputs*. Fauconnier & Turner (2002) afirmam que as formas das representações linguísticas são elementos mentais como quaisquer outros, podendo ser integradas, se alinhando à mesclagem da estrutura conceitual a que a forma está atrelada.

As análises demonstradas exemplificam como o processo de Mesclagem Conceptual permite interpretar essas adivinhas sob um ponto de vista imaginativo, tornando-se a chave para evocar o cômico. Afinal, como poderíamos responder a tais charadas sem esses mecanismos imaginativos: Em que situação, conceitos inanimados como seno, cosseno bateriam na porta de um banheiro? Passarinho fala? Se não, como ele propôs uma aventura romântica para a passarinha? Por acaso mingau tem mãe ou avó?

O que se propõe nesta análise é fornecer a resposta para perguntas como estas a partir das redes de integração conceptual que criam mesclagens que, mesmo não sendo verdadeiras no mundo real, são úteis para a produção do cômico por chamarem atenção para similaridades linguísticas que normalmente passam despercebidas. Tal processo não requer que acreditemos que senos e cossenos andam por aí, ou passarinhos falam, mas nos permite criar uma saída real para entidades não humanas, já que se constrói uma rede de eventos cuja estrutura se baseia em mecanismos imaginativos.

Além disso, não se pode esquecer que as piadas possuem uma lógica interna em que os ouvintes/leitores vão construindo a cena cômica em estágios. Na medida em que o conteúdo da piada é apresentado, ativamos *frames* relacionados diretamente às expressões linguísticas e à situação comunicativa como um todo.

6. CONCLUSÃO

Este artigo objetivou investigar os elementos linguísticos que ativam o efeito cômico e constroem cognitivamente o significado de piadas curtas. A hipótese de que as piadas ativam Mudança de *Frames* associadas a processos de Mesclagem Conceptual foi comprovada ao percebermos que os espaços mentais são domínios alimentados e estruturados por *frames* que, por sua vez, são motivados por marcas linguísticas das piadas curtas que funcionam como gatilhos para o ouvinte/leitor construir uma estrutura referencial que pode ou não representar entidades no mundo. Além disso, a sequência fonética idêntica ou similar funciona como o gatilho para o efeito cômico.

O processo de Mesclagem Conceptual que ocorre em piadas curtas apresenta como característica genérica a sequência fonética, que permite a projeção entre as contrapartes dos elementos contidos nos domínios iniciais. As adivinhas e charadas possuem uma pergunta inicial que ativa o conhecimento prévio do ouvinte/leitor construindo os espaços mentais. Os *frames* e MCIs são de fundamental importância para a construção emergente do efeito de humor, pois são eles que fornecem e estruturam os elementos dos *inputs*.

Ocorre, então, a projeção parcial entre domínios que interliga as contrapartes por analogia. Essas contrapartes se projetam na mescla, produzindo uma estrutura emergente cujos elementos ali formados não existem nos *inputs* iniciais. No processo, ocorre uma compressão desses elementos e a estrutura emergente concebe a semelhança fonética como o gatilho para o efeito cômico.

Portanto, as piadas (1) e (3) em que não há uma Mesclagem na forma, a resposta das adivinhas é projetada na mescla e fornece um novo conceito que recruta um *frame* familiar do ouvinte/leitor. Trata-se de uma complementação que ocorre na estrutura emergente e caracteriza a sequência fonética idêntica dos elementos comprimidos na mescla, como gatilho para o cômico. Quando há uma Mesclagem na Forma, além da Mesclagem Conceptual, uma nova construção é criada, como em “Mãezena” e “avéia Quaker” na piada (2), em “tangente” da piada (4), o que caracteriza a similaridade da sequência fonética dos elementos comprimidos na mescla como gatilho para o efeito cômico.

Nos termos de Fauconnier & Turner (2002), a conexão entre os elementos dos espaços mentais não ocorre sem uma razão. Fazemos tais conexões para compreendermos os conceitos em escala humana, nos tornando seres criativos e imaginativos. Por isso, a compressão feita a partir do processo de Mesclagem Conceptual é um aspecto importante da nossa capacidade cognitiva em criar novos sentidos e pode ser evidenciado nas piadas analisadas neste artigo.

REFERÊNCIAS

BERGSON, H. *O Riso: Ensaio sobre a significação do cômico*. Trad. Nathanael C. Caixeiro. São Paulo: Zahar Editores, 1980.

COULSON, S. *Semantic Leaps: Frame-Shifting and Conceptual Blending in Meaning Construction*. New York: Cambridge University Press, 2001.

FAUCONNIER, G. *Mental spaces*. Cambridge: Cambridge University Press, 1994.

_____. *Mappings in thought and language*. Cambridge: Cambridge University Press, 1997.

FAUCONNIER, G. & E. SWEETSER. *Spaces, Worlds, and Grammar*. Chicago: University of Chicago Press, 1996.

FAUCONNIER, G & TURNER, M. *The way we think*. New Yorker: Basic Books, 2002.

FERRARI, L. *Introdução à Linguística Cognitiva*. São Paulo: Contexto, 2011.

FILLMORE, C. Frame Semantics. In: Linguistic Society of Korea (Ed.) *Linguistics in the morning calm*. Seoul: Hanshin Publishing, 1982. pp 111-137.

FREUD, S. *Os chistes e sua relação com o inconsciente*. Rio de Janeiro: Imago, 1969.

LAKOFF, G. *Women, fire and dangerous things: what categories reveal about the mind*. Chicago: University of Chicago Press, 1987.

_____; JOHNSON, M. *Metaphors we live by*. Chicago: University of Chicago Press, 1980.

POSSENTI, S. *Os humores da língua: análises linguísticas de piadas*. Campinas: Mercado de Letras, 1998.

RASKIN, V. *Semantic mechanisms of humor*. Dordrecht: D. Reidel Publishing Company, 1985.

RITCHIE, G. *Logic and reasoning in jokes*. European Journal of Humor Research 2 (1), p. 50 -60, 2004.

Recebido em 23/04/2016

Aceito em 30/05/2016

O PAPEL DOS GESTOS DE APONTAR NA CONSTRUÇÃO DA DÊIXIS MULTIMODAL: DOS USOS CONCRETOS AOS USOS ABSTRATOS

Máira Avelar (UESB)¹

RESUMO

Neste artigo, investigamos a inter-relação entre a fala e os gestos de apontar na construção da dêixis multimodal. Nossa hipótese é de que a ocorrência mais prototípica dos gestos de apontar se relacionaria a referentes concretos presentes na cena enunciativa. Primeiramente, abordamos as considerações feitas por Kendon sobre as diferentes descrições icônicas dos gestos de apontar, bem como as considerações de McNeill e colegas sobre a dêixis abstrata. Analisamos, quantitativa e qualitativamente, dois trechos de 50 minutos de sessões legislativas dos deputados-pastores Feliciano e Cunha. Os resultados demonstraram que a hipótese não se confirmou, pois os referentes da maioria expressiva dos gestos eram abstratos.

PALAVRAS-CHAVE: Dêixis multimodal. Gestos de apontar. Inter-relação gestos-fala

ABSTRACT

We investigate the interrelation between speech and pointing gestures in the multimodal deixis construction. Our hypothesis is that the pointing gesture most prototypical occurrence would be related to concrete referents from the enunciative scene. Firstly, we approached Kendon's considerations about the different iconic descriptions of pointing gestures, as well as McNeill and colleagues' considerations about abstract deixis. We analyze, quantitatively and qualitatively, two 50-minute samples of legislative sessions from the so-called "deputy-pastors" Feliciano and Cunha. Results showed that the hypothesis cannot be confirmed, because most part of the referents were abstract.

KEYWORDS: Multimodal deixis. Pointing gestures. Speech-gestures interrelation.

¹ Professora Adjunta do Departamento de Estudos Linguísticos e Literários (DELL) e do Programa de Pós-graduação em Linguística (PP-GLin) da Universidade Estadual do Sudoeste da Bahia (UESB). mairavelar@uesb.edu.br

INTRODUÇÃO

A palavra “dêixis”, em grego, significa “ação de mostrar, indicar, assinalar”. Do ponto de vista linguístico, os dêiticos, enquanto uma categoria de signos, só existem na realidade enunciativa. Logo, eles “só existem na rede de indivíduos que a enunciação cria e em relação ao ‘aqui-agora’ do locutor. Por exemplo: o ‘eu’ o ‘isto’ e o ‘amanhã’ da descrição gramatical são somente os nomes metalinguísticos de ‘eu’, ‘isto’ e ‘amanhã’ produzido pela enunciação” (BENVENISTE, 1966, p.253). Dessa forma, só conseguimos estabelecer um referente para as expressões dêiticas na situação de enunciação em que são utilizadas, pois são expressões que apontam para um “fora do texto”.

Em consonância com as ideias de Benveniste, Kendon (2004, p. 222), baseando-se nos estudos realizados por Bühler (1990), afirma que a dêixis se refere à relação entre os enunciados e as condições de tempo e espaço em que ocorrem. Do ponto de vista operacional, as expressões dêiticas são frequentemente associadas aos gestos de apontar, uma vez que, em última análise, essas são as “ações não-linguísticas” responsáveis pelo estabelecimento da inter-relação referencial entre o enunciado e as circunstâncias espaço-temporais de sua ocorrência. Ainda segundo Kendon (2004), os gestos de apontar (*pointing gestures*) ou gestos dêiticos constituem uma classe separada de gestos, sendo que poucos estudos foram desenvolvidos sobre essa categoria.

Levando em consideração essa escassez de estudos sobre os gestos dêiticos, pretendemos sistematizar algumas considerações sobre a inter-relação entre esses gestos e a fala. Delineamos, então, a seguinte hipótese de trabalho: considerando a natureza dos dêiticos, de apontar para fora do enunciado, prototipicamente, gestos dêiticos coocorreriam com palavras dêiticas e apontariam, para objetos, pessoas ou lugares concretos, relativos à situação mais imediata de comunicação. Haveria também contextos, menos prototípicos, em que gestos dêiticos coocorreriam com palavras não-dêiticas e apontariam para objetos, pessoas, lugares e até mesmo ideias abstratas.

1. REFERENCIAL TEÓRICO-METODOLÓGICO

1.1 GESTOS DÊITICOS OU GESTOS DE APONTAR: CARACTERIZAÇÃO E CATEGORIZAÇÃO

Os gestos de apontar, sejam eles realizados com as mãos ou com outras partes do corpo (como, por exemplo, a cabeça, o lábio inferior ou os olhos em movimento) possuem um padrão comum de movimento, que consiste numa trajetória em linha reta e bem definida, havendo uma tendência de que o movimento realizado atinja sua extensão máxima e seja mantido de maneira breve nessa posição. Dessa forma, excetuando-se os casos de objetos em movimento, o gesto de apontar parece ser realizado numa direção específica e direcionado a um alvo específico (KENDON, 2004, p. 200).

Sendo assim, numa perspectiva icônica (MCNEILL; CASSELL; LEVY, 1993, p.5), os gestos de apontar – sejam eles pertencentes ao nível concreto ou abstrato – podem ser esquematizados da

mesma maneira: há “um ponto de origem, ou *origo* (BÜHLER, 1990), um alvo, e uma linha icônica ou trajetória ligando o *origo* ao alvo”². Essa esquematização pode ser visualizada na figura 1. No esquema em questão, o gesto de apontar conecta o *locus* da origem àquele do referente. Nesse sentido, o gesto de apontar apresenta uma dupla natureza, dêitica e icônica, uma vez que representa a orientação espacial de um determinado objeto ou entidade no espaço, dada pelo falante (MCNEILL; CASSELL; LEVY, 1993).

Origo —→ Objeto referente

Numa perspectiva operacional, ao abordar os gestos manuais de apontar, Kendon (2004, p.223) estabelece duas maneiras distintas de realizar essa ação: apontar com o dedo indicador, relativa à ideia de singularidade do objeto referido, e apontar com a mão aberta, relativa ao “status simbólico, conceptual ou exemplar” do objeto apontado. Foram estabelecidos sete tipos de combinações entre formato de mão e combinação de antebraço, brevemente descritas e contextualizadas a seguir (KENDON, 2004, p. 205-222):

- a) Dedo indicador estendido pronado (palma para baixo): utilizado para apontar um único objeto individual ou tópico conversacional.
- b) Dedo indicador estendido neutro (palma vertical): utilizado para distinguir um objeto adicional, que possui alguma relação com o objeto individuado, ou para fazer um comentário sobre o objeto individuado, ou para demonstrar que o objeto indicado é condição ou causa de outra coisa.
- c) Polegar estendido (orientação de antebraço variável): utilizado para apontar para objetos situados ao lado ou atrás do falante ou apontar quando a localização ou a identidade precisa do objeto apontado não precisa ser especificada ou não é o foco do discurso.
- c) Mão aberta neutra (palma vertical): utilizado para se referir não a uma localização específica, que seria individuada pelo dedo indicador, mas a algo que pode ser encontrado numa determinada localização já referida anteriormente no discurso e que está sendo repetida.
- e) Mão aberta supinada (palma para cima): trata-se de uma das famílias gestuais descritas por Kendon (2004, 248-283), cuja função pragmática central consiste em “apresentar”. No contexto dêitico, o objeto apontado não é apenas individuado como tópico ou como uma localização associada a outro foco discursivo, mas como se estivesse sendo apresentado ao interlocutor como algo que deve ser olhado com atenção ou inspecionado de uma maneira particular.
- f) Mão aberta oblíqua: utilizada para indicar um objeto quando algum comentário sobre ele está sendo feito a um terceiro, sendo que,

2 Todas as citações diretas em língua inglesa foram traduzidas de maneira livre por nós. Optamos por inserir as traduções, de modo a uniformizar o texto em uma só língua.

frequentemente, o objeto apontado é uma pessoa e o comentário feito é negativo. Esses gestos, direcionam, então, a atenção de uma terceira pessoa ao alvo da crítica e desempenham um papel de tornar a crítica pública.

g) Mão aberta pronada (palmas para fora): utilizada para se referir à extensão espacial de um objeto ou quando vários objetos estão sendo considerados como um conjunto. Sendo assim, o que está em foco não é um objeto a ser distinto dos outros pela localização espacial, mas o que pode ser visto naquele local.

Cada uma das combinações descritas pode ser ilustrada sucintamente por meio da figura a seguir, bem como por meio dos exemplos analisados no *corpus* – cf. subseção 3.4.



Do ponto de vista multimodal, é relevante observar que, enquanto para Kendon (2004), as combinações que contêm o dedo indicador estendido são frequentemente associadas a palavras dêiticas – especialmente aos dêiticos espaciais –, as combinações que contêm a mão aberta são menos frequentemente associadas a palavras dêiticas, pois, de maneira geral, o dedo indicador estendido tem a função de especificar um objeto, enquanto a mão aberta, em quaisquer de suas combinações, indica objetos que não constituem o tópico conversacional, mas algo que está relacionado ao tópico em questão, seja o exemplar de uma categoria ou a localização de uma atividade em discussão, relacionada a algo que aconteceu ou a algo que deve ser olhado com atenção, porque está relacionado ao tópico principal.

1.2. GESTOS DE APONTAR CONCRETOS E ABSTRATOS

Do ponto de vista espacial, os gestos de apontar podem se referir tanto ao espaço físico partilhado pelos interlocutores, e, conseqüentemente, a um objeto concreto presente no ambiente, quanto à

estruturação de ideias do falante. Neste último caso, a utilização dos dêiticos pode servir para mapear a localização de personagens e objetos numa narrativa (MCNEILL; CASSELL, LEVY, 1993), como também para contrastar metaforicamente posições ideológicas no discurso (MIRANDA; MENDES, 2015). Sendo assim, haveria uma utilização dêitica prototípica, referente à localização do espaço físico, e uma utilização dêitica não-prototípica, correspondente à localização metafórica de pessoas, objetos ou ideias.

Em relação à dêixis abstrata, o ato de apontar para um espaço fisicamente vazio, cria, na verdade, um referente orientado espacialmente no contexto discursivo. Entretanto, os gestos de apontar apresentam diferentes valores semióticos, relacionados, sobretudo, à estruturação de três diferentes níveis discursivos³ (MCNEILL; CASSELL, LEVY, 1993):

- a) o nível metadiscursivo, em que o falante aborda questões relativas ao próprio ato de narrar/ descrever/ argumentar. Por exemplo: “Vou te contar a história de um filme que vi”. Esse nível é, portanto, relativo à representação ou à introdução do evento/argumento e cria pontos de junção entre os eventos/argumentos expostos.
- b) o nível discursivo, em que o falante relata/expõe uma sequência de eventos/argumentos. Por exemplo: “O menino subiu a escada e depois foi para o quarto”. Esse nível é, portanto, relativo à narração/descrição/argumentação propriamente dita.
- c) o nível paradiscursivo, em que o falante se distancia da situação narrativa/descriptiva/argumentativa e estabelece contato direto, na posição social de falante, com o ouvinte. Por exemplo: Após narrar uma sequência de eventos, o falante endereça-se diretamente ao ouvinte e diz: “Entendeu?”. Esse nível é, portanto, relativo à interlocução entre falante e ouvinte.

Segundo McNeill e colegas (1993), no nível discursivo, o gesto de apontar pode ter função referencial: “o falante descreve uma localização ou o movimento de algo em direção a essa localização e estabelece um ponto no espaço no espaço gestual para representar esse *locus*”. Dessa maneira, o *origo* é adjacente ao evento. Ainda no nível discursivo, os gestos de apontar podem “estabelecer cadeias correferências onde referências sucessivas estão ligadas em virtude de ocuparem o mesmo *locus* no espaço”. Por exemplo: ao se referir ao mesmo personagem sucessivamente, o falante utiliza o mesmo gesto e aponta para o mesmo ponto no espaço.

Já no nível metadiscursivo, segundo os autores supracitados (1993), o ato de apontar para o espaço físico corresponde a apontar aspectos da narração/ descrição/argumentação que está sendo desenvolvida. Em outras palavras: “o *origo* para o apontar metanarrativo é o *locus* do falante próximo da história ou ao elemento estrutural da história concebido metaforicamente como um objeto”. Por exemplo: ao dizer “na primeira cena”, o falante aponta para um espaço central, assinalando um segmento da

³ Na verdade, os autores estabelecem essa distinção para o nível narrativo. Porém, como eles próprios assinalam na conclusão de seu artigo, essa distinção pode ser aplicável a outros tipos textuais.

história a ser narrada. Ao contrário do que ocorre nos outros dois níveis, em que um *locus* é indicado no espaço gestual neutro, no nível paradiscursivo, o gesto de apontar é direcionado diretamente ao interlocutor.

Do ponto de vista metodológico, pretendemos aliar as considerações feitas por Kendon (2004), a respeito do núcleo pragmático correspondente a cada ação de apontar -seja ela realizada com o dedo indicador, com o dedão ou com a mão aberta -com as considerações feitas por McNeill, Cassell e Levy (1993) sobre as inter-relação entre gestos dêiticos e níveis discursivos. Os procedimentos metodológicos para a análise quantitativa e qualitativa dos dados serão descritos na seção seguinte.

2. METODOLOGIA

2.1 SELEÇÃO E DESCRIÇÃO DO *CORPUS*

A fim de testar a hipótese inicial, em que foi estabelecida uma relação prototípica entre a ocorrência de palavras e gestos dêiticos referindo-se a contextos concretos, bem como de verificar o contexto de ocorrência dos diferentes gestos de apontar, foram selecionadas as amostras finais, de 50 minutos cada, provenientes de dois vídeos, exibidos em 2013 e 2014, de sessões legislativas de dois deputados autodenominados “deputados pastores”: Marco Feliciano, do Partido Socialista Cristão (PSC) e presidente da Comissão de Direitos Humanos e Minorias na época, e Eduardo Cunha do Partido , atual presidente da Câmara dos Deputados, eleito líder do PMDB na Câmara na época.

O vídeo de Eduardo Cunha corresponde à sessão plenária do dia 25/06/2013, em que o deputado, em nome do PMDB, posiciona-se contrariamente à aprovação da Proposta de Emenda Constitucional (PEC) 37, em que se propunha que o poder de investigação criminal se restringisse às polícias militar e civil, retirando o poder de investigação de órgãos como, por exemplo, o Ministério Público. Já vídeo de Marco Feliciano corresponde a um pronunciamento realizado na Sessão Legislativa do dia 14/07/2014, em que o deputado tece considerações sobre a entrevista concedida pelo autor de novelas Manoel Carlos no jornal “Estadão”, de grande circulação nacional. Feliciano posiciona-se contrariamente à exibição de um beijo homoafetivo na novela “Em Família”, em exibição na época, e ressalta a importância da família tradicional brasileira.

PROCEDIMENTOS METODOLÓGICOS

2.2.1 PROCEDIMENTO DE CATEGORIZAÇÃO DOS GESTOS

O primeiro passo para realizar uma análise de gestos manuais consiste em distinguir esses gestos dos demais movimentos realizados com as mãos (MÜLLER, 2014). Gestos constituem, então, “movimentos distintos de esforço identificável das mãos e antebraços, ou seja, o curso (*stroke*) dos gestos” (CIENKI, 2005, p. 425). A realização dos gestos compreende três fases, em que é empreendida uma “excursão de movimento”, denominada unidade gestual (KENDON, 2004, p. 110):

- a) a preparação: fase opcional, em que os membros se movem a partir de uma posição de relaxamento ou descanso;
- b) o curso (*stroke*): fase obrigatória, em que a expressão gestual é realizada, havendo a manifestação clara de movimentos dinâmicos que demandam esforço e foco de energia. Nesta fase, considerada como o ápice do gesto, as mãos tendem a descrever formas e completar padrões de movimento ou cessarem brevemente o movimento, em que os membros são mantidos parados, antes de relaxarem e retornarem à posição inicial;
- c) a retração: fase opcional, em que há uma retração do movimento para a posição inicial de relaxamento ou descanso.

A partir da marcação dos *strokes* correspondentes aos gestos dêiticos, foram realizados testes de frequência de ocorrência comparando as amostras selecionadas para análise – cf. seção 3. A fim de identificar os gestos dêiticos de maneira confiável, a velocidade do vídeo foi diminuída para 40%, os movimentos das mãos foram visualizados sem o som – cf. indicações metodológicas de Bressemer (2013) – e os *strokes* foram identificados dentro de cada unidade ou frase gestual.

2.2.2 PROCEDIMENTOS DE ANÁLISE MULTIMODAL DOS DADOS

As análises das amostras foram realizadas por meio da ferramenta ELAN 4.9.1, que permite que sejam criadas trilhas de análise correspondentes a cada um dos aspectos a serem analisados. Segue uma explicação mais detalhada sobre cada uma das trilhas:

- a) Transcrição (*Transcription*) e Tradução (*Translation*): estas trilhas, que correspondem à primeira e à segunda trilha do *template*, contêm o texto do vídeo transcrito ortograficamente e traduzido livremente para o inglês, respectivamente.
- b) Gesto de apontar (s/n) (*Pointing gesture (y/n)*) e Palavra dêitica (s/n) (*Deictical word (y/n)*): nestas trilhas, que correspondem à terceira e à quinta trilha, respectivamente, é possível selecionar as opções “gesto sim” (*gesture yes*), “palavra dêitica sim” (*deictical word yes*) e “gesto não” (*gesture no*), “palavra dêitica não” (*deictical word no*). Considerando a inter-relação entre fala e gestos na emergência da dêixis multimodal, hipotetizamos três situações: a coocorrência entre fala e gestos, a ocorrência de gestos sem a ocorrência de palavra dêitica e a ocorrência de palavra dêitica sem a ocorrência de gestos.
- c) Gesto (c/a) (*Gesture (c/a)*): no caso de ocorrência de gestos, é necessário selecionar, nesta trilha – que corresponde à quarta trilha de

análise –, se o gesto se refere a uma entidade concreta ou a uma entidade abstrata.

d) Palavra (*Word*): nesta trilha, que corresponde à sexta trilha de análise, é possível escrever a palavra: seja a palavra dêitica ou a palavra que acompanha o gesto abstrato. Dessa maneira, é possível refinarmos a análise dos dados e observarmos, de maneira mais específica, quantas e quais são as palavras utilizadas no contexto dêitico, seja verbal ou verbo-gestual.

e) Descrição icônica do gesto (*Gesture iconic description*): nesta trilha, que corresponde à sétima trilha de análise, seguimos a lógica de categorização de gestos estabelecida por Kendon (2004) e Müller (2004) para os gestos de mão aberta supinada, categorizada por ambos como “Mão aberta com as palmas para cima” ou “Palm-up open hand (PUOH)” e criamos, então, uma categorização para as sete diferentes combinações dos gestos de apontar que leva em consideração: a orientação das palmas – horizontal, vertical ou para fora – e o formato da mão ou dedo – mão aberta, dedo indicador ou polegar. A fim de manter uma padronização internacional de classificação, foram utilizadas as iniciais em inglês para a codificação realizada: *palm down* (PD), *palm vertical* (PV), *palm up* (PU), *palm oblique* (PO) *palm away* (PA) para as palmas vertical, para cima, oblíqua e para fora, respectivamente, *index finger* (IF) para dedo indicador, *open hand* (OH) para mão aberta e *thumb* (THUMB) para polegar. A codificação está sistematizada a seguir:

Dedo indicador estendido pronado (palma para baixo): PDIF

Dedo indicador estendido neutro (palma vertical): PVIF

Polegar estendido (orientação de antebraço variável): THUMB

Mão aberta neutra (palma vertical): PVOH

Mão aberta supinada (palma para cima): PUOH

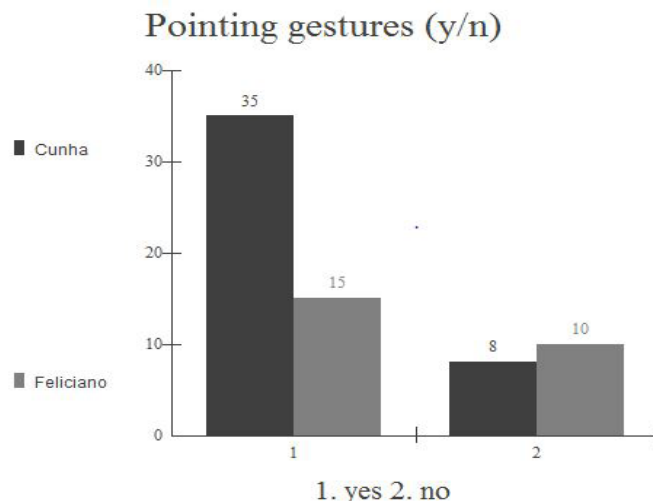
Mão aberta oblíqua: POOH

Mão aberta pronada (palmas para fora): PAOH

Após a categorização dos dados de ambas as amostras, procedemos à análise dos dados. Os resultados serão apresentados na seção seguinte.

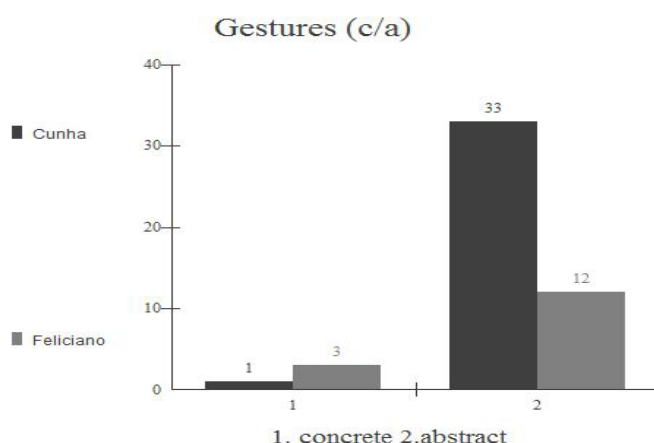
3. RESULTADOS E DISCUSSÃO

3.1 GESTOS DE APONTAR (S/N)/ POINTING GESTURES (Y/N)



Conforme é possível visualizar no gráfico, em ambas as amostras, os deputados realizam um número expressivamente maior de gestos de apontar, sendo que Cunha realiza um número maior desses gestos (35 ocorrências) do que Feliciano (15 ocorrências). Aplicando o Teste-G⁴, foi possível constatar que a diferença é estatisticamente significativa: numa tabela 2x2, o resultado de $p = 0,05$. Contudo, nos casos categorizados como “pointing gesture no”, é importante ressaltar que, no caso de Feliciano, ele realiza outros gestos – como, por exemplo, o gesto do soco – quando não realiza o gesto de apontar, enquanto Cunha alterna entre realizar outros gestos e não gesticular. Verificamos, então, se os gestos de apontar coocorriam ou não com palavras dêiticas, conforme os resultados apresentados a seguir.

3.2 GESTO (C/A)/ GESTURE (C/A)

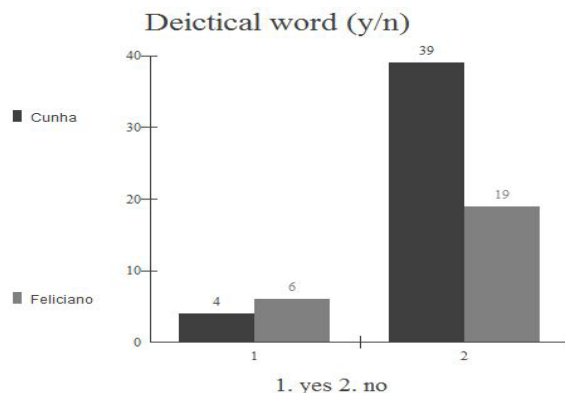


A visualização dos dados permite concluir que, ao contrário do que prevíamos, os gestos dêiticos apontam, na expressiva maioria dos casos, para referentes que não estão presentes na situação comunicativa imediata: no caso de Cunha, ele realiza apenas um gesto dêítico que remete a pessoas, objetos ou situações concretas, enquanto, nas demais ocorrências (33), os gestos realizados

4 Teste não-paramétrico, aplicado a dados nominais, semelhante ao teste Qui-quadrado, aplicado a duas amostras independentes.

possuem referentes abstratos. Os dados relativos a Feliciano apresentam resultados análogos: em apenas 3 ocorrências, os gestos dêiticos apontam para referentes concretos, enquanto nas demais (12 ocorrências), eles apontam para referentes abstratos mencionados na fala. Do ponto de vista estatístico, a diferença também se revelou como sendo significativa, pois o resultado do Teste-G, numa tabela 2x2, foi de $p = 0,05$.

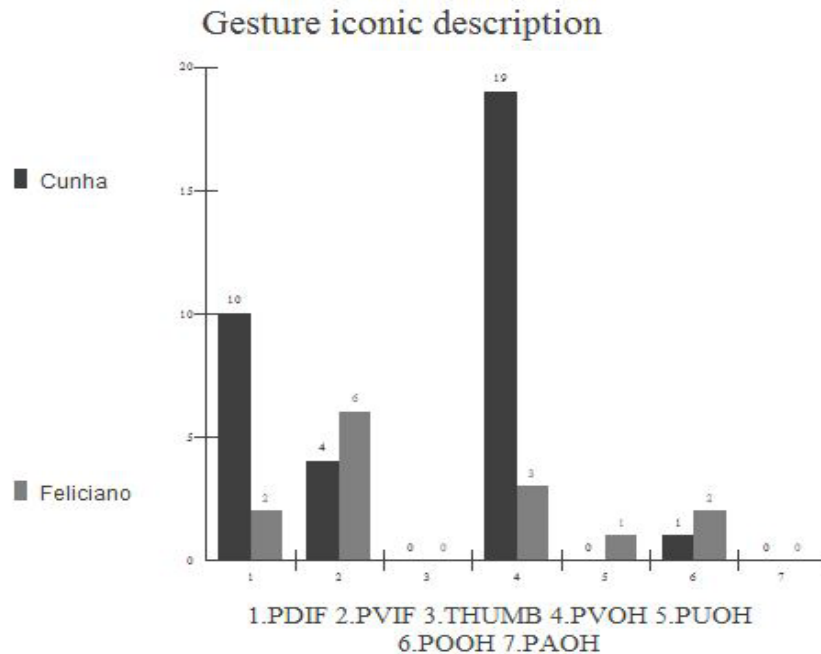
3.3 PALAVRA DÊITICA (S/N)/ DEICTICAL WORD (Y/N)



Os resultados encontrados no gráfico em questão corroboram aqueles encontrados no gráfico anterior, pois, novamente, na expressiva maioria dos casos, os gestos de apontar são acompanhados por palavras não-dêiticas, o que vai de encontro, mais uma vez à nossa hipótese de trabalho. No caso de Cunha, palavras não-dêiticas correspondem a 39 das ocorrências (x 4 ocorrências de palavras dêiticas), enquanto no caso de Feliciano, palavras não-dêiticas correspondem a 19 das ocorrências (x 6 ocorrências de palavras dêiticas). Em outras palavras, os sujeitos apontam para o espaço, mas não apontam para objetos concretos ou presentes na cena enunciativa mais imediata. Entretanto, do ponto de vista estatístico, a diferença não é significativa, pois, numa tabela 2x2 o resultado do Teste-G é de $p = 0,10$.

De maneira geral, o resultado encontrado corrobora a afirmação feita por McNeill, Cassell e Levy (1993, p. 9), de que “o espaço físico torna-se um espaço de possibilidade referenciais”. Por fim, é importante salientar que não há uma correspondência exata dos dados contidos nos gráficos 2 e 3 pelo fato de algumas palavras dêiticas ocorrerem desacompanhadas de gestos – no caso de Cunha – ou acompanhadas por outros gestos que não foram quantificados, como, por exemplo, o gesto de soco – no caso de Feliciano. Sendo assim, é possível afirmar que Cunha realiza mais gestos dêiticos mas, não necessariamente, realiza maior número de gestos no total. Além disso, é possível hipotetizar que Feliciano realiza gestos mais variados que Cunha, pois a amostra inclui gestos não-dêiticos.

3.4 DESCRIÇÃO ICÔNICA DOS GESTOS/ GESTURE ICONIC DESCRIPTION⁵



Em relação à descrição icônica dos gestos, dois dados que, inicialmente, atraem nossa atenção dizem respeito à ausência de ocorrências de gestos de apontar com o dedo (THUMB) e de “mão aberta, palmas para cima” (PAOH). A nosso ver, o principal fator motivador da não-ocorrência desses gestos se deve ao fato de eles individuarem objetos concretos, situados atrás do locutor ou naquele lugar específico apontado. Como a maioria dos gestos, em ambas as amostras, se refere a contextos abstratos, isso desfavoreceria a ocorrência dos gestos em questão.

Em relação aos demais gestos, é possível observar uma distribuição bastante diferente entre as duas amostras: no caso dos gestos de apontar com o dedo (IF), Cunha utiliza o gesto de “apontar com o dedo, palmas para baixo” (PDIF) em 10 ocorrências, enquanto Feliciano o utiliza apenas em 2 ocorrências. Segundo Kendon (2004), esses gestos serviriam tanto para individuar objetos, quanto o tópico conversacional. No caso de Cunha, ele aponta entidades no espaço, sendo que os gestos servem não apenas para aponta-los, mas para enfatizá-los -como na sequência exemplificada, em que ele repete o mesmo gesto a cada referente dito. Já no caso de Feliciano, ele usa esses gestos para se referir concretamente ao interlocutor do discurso. Ambos os exemplos podem ser visualizados a seguir:

⁵ Por questões de coerência e de espaço, optamos por comentar algumas ocorrências da trilha “Words”/Palavras” juntamente com a descrição icônica dos gestos.



“O **nosso** país é um país conservador” “Nós (1) **queremos** (2) **combater** (3) a **criminalidade**

Já no caso dos gestos de “apontar com o dedo, palma vertical” (PVIF), Feliciano realiza maior número desses gestos (6 ocorrências), enquanto Cunha realiza 4 ocorrências. Segundo Kendon (2004), esses gestos seriam utilizados para fazer comentários sobre um objeto ou para apontar para um objeto adicional. No caso de Feliciano, ele da família brasileira, usando sempre o “nós” – como pode ser visualizado no exemplo a seguir. Cunha, por sua vez, ao especificar como queremos combater a criminalidade, cita diversos órgãos que poderiam fazê-lo e, ao citá-los utiliza os gestos em questão. Mais uma vez, os gestos cumprem também o papel de enfatizar os referentes.



“(…)mostrando que, de fato, **somos conservadores**” “Com o **Ministério Público**, com a **polícia**,”
com **todos**”.

No caso dos gestos de apontar com a mão aberta (OH), os gestos de “mão aberta, palma vertical” (PVOH) constituem a maioria das ocorrências na primeira amostra (19 ocorrências), enquanto correspondem a apenas 3 ocorrências na segunda. Segundo Kendon (2004), esses gestos seriam utilizados para apontar um referente já mencionado antes no discurso. Entretanto, no caso de Feliciano, eles são utilizados para se individuar pessoas ou entidades como, por exemplo, Deus. No caso de Cunha, ele utiliza uma longa sequência de gestos repetitivos para, de certa maneira, reiterar e resumir a posição contrária tomada pelo PMDB a respeito de restringir o poder de investigação da corrupção apenas pela Polícia Federal:



“E que **Deus abençoe** a todo Brasil” “O **PMDB** (1) **espera** (2) que, **no futuro**(3) (...)”

Por fim, os gestos de “mão aberta, palmas para cima” (PUOH) apresentam apenas uma (1) ocorrência, realizada por Feliciano e nenhuma (0) ocorrência, no caso de Cunha. Segundo Kendon (2004) e Müller (2004), esses gestos correspondem à função pragmática e “apresentar”. Feliciano esse gesto para apresentar, de maneira resumida, os fatos: “o que aconteceu nada mais foi que a população brasileira mostrando que, da fato, somos conservadores”. Acreditamos que, por se tratar do trecho final dos discursos, a ideias e referentes já tenham sido, em sua maioria, apresentados anteriormente, razão para a baixa ocorrência desses gestos.



“A **população brasileira**”

Constatamos também um baixo número de ocorrências dos gestos “mão aberta, palmas oblíquas” (POOH): apenas uma (1) ocorrência, no caso de Cunha, e duas (2), no caso de Feliciano. Segundo Kendon (2004), esses gestos são utilizados quando está sendo feito um comentário negativo sobre uma terceira pessoa, para a qual se aponta. No caso de Feliciano, ele está criticando a novela “Em Família”, de fato fazendo um comentário negativo sobre ela e diferenciando o “nós” x “eles”: “todos os dias, entravam dentro da nossa casa e expunham todo tipo de coisa que eu, como pai de família, não gostaria que minhas filhas pudessem olhar”. Cunha também aborda uma terceira pessoa em seu discurso: o PMDB. Porém, ressalta aspectos positivos do partido que liderava na Câmara dos Deputados:



“**Entravam dentro da nossa casa**” “**Levou os dois lados** hoje para o debate”

De maneira geral, os resultados encontrados demonstram que a maior parte dos gestos realizados por Cunha concentram-se no primeiro (PDIF) e no quarto (PVOH) grupo de descrição icônica. Já no caso de Feliciano, as ocorrências se apresentam de maneira mais distribuída, sendo que o maior número delas corresponde ao segundo grupo de descrição icônica (PVIF). Esses resultados demonstram que, somadas as ocorrências, ambos os deputados fazem maior uso dos gestos de apontar com o dedo, o que poderia demonstrar a prototipicidade desse tipo de gesto. Por fim, é válido ressaltar que os resultados são estatisticamente significativos, pois o resultado do Teste G é de $p = 0,05$, numa tabela 7x2.

CONSIDERAÇÕES FINAIS

Tal como ocorre com a dêixis verbal, a dêixis gestual é atualizada de acordo com o “aqui-agora” da enunciação, pois o mesmo gesto de apontar pode ser ocupado por diferentes referentes. Mostra-se relevante, inclusive, o estabelecimento de diferentes gestos de apontar, que não dizem respeito a uma mera característica individual do discurso, mas a padrões de movimento que podem variar sistematicamente, pois podem ser associados a diferentes funções pragmáticas exercidas ao longo da estruturação discursiva.

Ao analisarmos os três níveis de estruturação propostos por McNeill e colegas (1993), em ambas as amostras, é possível verificar que os gestos de apontar são realizados expressivamente no nível discursivo – em que a argumentação propriamente dita se desenvolve –, e, em menor grau, no nível metadiscursivo – em que são estabelecidos pontos de junção entre os diferentes argumentos, muitas vezes acompanhados por conjunções como, por exemplo, “então”.

O fato de a expressiva maioria dos gestos também se remeterem a entidades abstratas remete à metáfora conceptual básica “IDEIAS SÃO OBJETOS” que, no contexto específico de análise, pode ser desdobrada na metáfora “POSIÇÃO REFERENCIAL É POSIÇÃO ESPACIAL”. Mais do que isso, como afirmado pelos próprios McNeill, Cassell e Levy (1993, p.17, grifos dos autores): “Um gesto de apontar *cria* um referente no espaço”. Assim, “os gestos de apontar criam uma transposição mental do espaço do objeto no espaço abstrato”. Portanto, “graças ao trabalho criativo do gesto, *o espaço não é, de fato, vazio!*”.

Por fim, gostaríamos de ressaltar que, devido às limitações de tempo, espaço e escopo de desenvolvimento deste artigo, algumas questões relevantes não puderam ser abordadas. Questões como: a trajetória dos gestos, a qualidade e precisão dos movimentos (*movement sharpness*), o número de articulações envolvidas na realização do gesto e, sobretudo, a repetição dos gestos de apontar estão sendo devidamente endereçadas no trabalho “The role of gesture strokes in the emergence of multimodal metaphors: an analysis of political-religious discourses”⁶, ainda em construção, que conta com um *corpus* ampliado. Em última análise, hipotetizamos que esses recursos funcionam como indicadores de ênfase, mobilizando emoções no discurso. Essas e outras questões sobre o estudo dos gestos, ainda tão pouco explorado no Brasil, ocupam lugar central em nossas pesquisas.

REFERÊNCIAS

AYRES, M., AYRES JÚNIOR, M.; AYRES, D.L.; SANTOS, A.A *BioEstat: Aplicações estatísticas nas áreas das ciências biomédicas*. Versão 5.3. Ong Mamiraua. Belém, PA, 2007. Disponível em: <http://www.mamiraua.org.br/pt-br/downloads/programas/bioestat-versao-53/>, Acesso em: jan/2015.

BENVENISTE, E. La nature des pronoms. In: *Problèmes de linguistique générale 1*. Paris: Gallimard, 1966, p. 251-257.

BRESSEM, J. A linguistic perspective on the notation of form features in gestures. In: MÜLLER, C. et al (eds.) *Body – Language – Communication: an international handbook on multimodality in human interaction*. v. 38, n. 1. Berlin/Boston: De Gruyter Mouton, 2013, p. 1079-1098.

BÜHLER, K. *Theory of Language: The representational function of language*. Amsterdam: John Benjamins, 1990.

CIENKI, A. Image Schemas and Gestures. In: HAMPE, B; GRADY, J. (Eds.). *From Perception to Meaning: Image Schemas in Cognitive Linguistics*. Berlin, New York: Mouton de Gruyter, 2005, p. 421-442.

CUNHA, E. *Sessão Plenária da Câmara dos Deputados do dia 25/06/2014*. YouTube, 13 de agosto de 2014. Disponível em: <https://www.youtube.com/watch?v=Ku8pTFWqxqE>. Acesso em: jan/2015.

FELICIANO, M. *Sessão Plenária da Câmara dos Deputados do dia 13/07/14*. YouTube, 13 de agosto de 2014. Disponível em: <https://www.youtube.com/watch?v=Erbl-nwNW8w>. Acesso em: jan/2015.

KENDON, A. *Gesture: visible action as utterance*. Cambridge: Cambridge University Press, 2004.

MCNEILL, D.; CASSELL, J; LEVY, E.T. Abstract deixis. In: *Semiotica*, v. 95, n.1. Berlin: Walter de Gruyter, 1993, p. 5-19.

MIRANDA, M.A.; MENDES, P. The role of gestures in the construction of multimodal metaphors: analysis of a political-electoral debate. In: *Revista Brasileira de Linguística Aplicada*, v.15, n.2, 2015, p. 343-376.

MÜLLER, C. Gesture as “deliberate expressive movements”. In: SEYFEDDINIPUR, M.; GULLBERG, M. (eds.) *From gesture in conversation to visible action as utterance: essays in honor of Adam Kendon*. Amsterdam: John Benjamins, 2014.

MÜLLER, C. Forms and uses of the Palm Up Open Hand: a case of a gesture family? In: MÜLLER,

C.; POSNER, R. *The semantics and pragmatics of everyday gestures*. Berlin: Weller, 2004, p. 233-256.

SLOETDJES, H., & WITTENBURGH, P. *ELAN*. Version 4.8.1, retrieved 20 November 2014 from <http://tla.mpi.nl/tools/tla-tools/elan/> by Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands, 2008.

Recebido em 28/02/2016

Aceito em 07/04/2016

A CONTRAPARTE PROSÓDICA EM CONSTRUÇÕES DE DISCURSO REPORTADO

Patrícia Ribeiro do Valle Coutinho (UFJF)¹ e Luiz Fernando Matos Rocha (UFJF)²

RESUMO

Este artigo defende que aspectos prosódicos desempenham um papel no pareamento de construções instanciadas na fala. Em conformidade com as bases teóricas da Gramática das Construções (GOLDBERG 1995, 2006), argumenta-se que construções são coleções linguísticas sistemáticas de pareamentos forma-função aprendidos na linguagem cotidiana e que podem ser mapeadas em dados contextualizados de fala. O objetivo é descrever e analisar tendências prosódicas de construções de discurso reportado em amostras de *corpus* de fala espontânea do português carioca culto brasileiro (NURC-RJ). As ocorrências são submetidas ao programa PRAAT (BOERSMA & WEENINK, 2011), por meio do qual informações prosódicas são analisadas e correlacionadas.

PALAVRAS-CHAVE: Gramática das Construções. Prosódia. Discurso Reportado.

ABSTRACT

This article claims that prosodic aspects have a bearing on construction pairings instanced in the spoken variety of language. According to Construction Grammar theory (GOLDBERG 1995, 2006), it is argued that constructions are linguistic form-meaning pairings acquired in everyday language and can be found in contextualized speech data. The goal is to describe and analyze prosodic tendencies in reported speech constructions of spontaneous speech corpus samples of Brazilian Portuguese (Rio de Janeiro – standard dialect NURC/RJ). Occurrences are subjected to PRAAT program (BOERSMA & WEENINK, 2011), by which prosodic information is analyzed and correlated.

KEYWORDS: Construction Grammar; Prosody; Reported Speech.

1 Doutora em Linguística pela Universidade Federal de Juiz de Fora - Programa de Pós-Graduação em Linguística - Rua José Lourenço Kelmer, s/n - Bairro Martelos. patriciarvcoutinho@yahoo.com.br

2 Doutor em Linguística - Universidade Federal do Rio de Janeiro. luiz.rocha@ufjf.edu.br

INTRODUÇÃO

O objetivo específico do presente artigo é reivindicar a participação da prosódia na constituição do sentido e na organização da cena interacional. Assumimos o ponto de vista de que a forma linguística, o significant, do qual fazem parte os correlatos prosódicos, é uma pista para a construção do significado, e que dados contextuais dinâmicos moldam a compreensão do sentido.

Assim, como um importante recurso linguístico de estruturação da comunicação, elegemos o discurso reportado (doravante DR) da fala espontânea do dialeto carioca do português brasileiro (doravante PB), para nos servir de objeto nesta pesquisa. A escolha também foi influenciada pelo empreendimento iniciado por Rocha (2004), tese de doutorado de autoria do orientador deste trabalho, ao abordar a construção de DR na fala espontânea e levantar importantes intuições sobre o funcionamento prosódico desse fenômeno discursivo, o que propiciou curiosas perguntas de pesquisa.

As questões que surgiram podem ser ilustradas com: há situações em que interpretamos uma sequência linguística a partir do que sinaliza a entoação da fala, mesmo que as palavras, ou seja, o conteúdo lexical, indiquem o contrário? E, especificamente, quando um discurso é reportado, em quanto da entoação nos apoiamos para julgar e avaliar a fala do outro?

Justificamos a pertinência deste trabalho a partir das contribuições que se podem oferecer tanto para a área dos estudos em prosódia, quanto para a área dos estudos em Gramática das Construções (doravante GC). A descrição prosódica a que nos propomos poderá incrementar o acervo de estudos brasileiros que visam ao tratamento dos aspectos suprasegmentais do PB. Além disso, acreditamos que nossa empreitada em conciliar estudos sociocognitivistas – cuja tradição é associada à modalidade escrita – e estudos prosodistas pode ampliar o horizonte teórico da GC.

Definimos como *corpus* o NURC-RJ (Projeto da Norma Urbana Oral Culta do Rio de Janeiro), disponível em <http://www.lettras.ufrj.br/nurc-rj/> e escolhemos dois diálogos entre falantes. Para a análise, submetemos as ocorrências de DR ao programa PRAAT (BOERSMA & WEENINK, 2011).

O presente artigo se estrutura inicialmente com os fundamentos teóricos gerais sobre GC. Em seguida, é apresentado nosso objeto de estudo: o DR e sua tipologia. Apresentamos, na seção seguinte, os procedimentos metodológicos da pesquisa, bem como as categorias de análise utilizadas, a explanação acerca do *corpus* e do programa computacional para leitura acústica. Mais adiante, resumizamos os resultados e discussões e, por fim, na seção 5, tecemos as considerações finais.

2.REFERENCIAL TEÓRICO

2.1.A GRAMÁTICA DAS CONSTRUÇÕES

Goldberg (1995) escreveu uma tese sobre os tipos de sentenças simples das gramáticas, postulando que esses tipos de sentenças são instâncias de construções – correspondências entre forma e sentido que existem independentemente de verbos particulares. Realiza-se um tratamento da natureza do

significado verbal e da sua relação com o significado sentencial, negando que a sintaxe e a semântica da cláusula sejam projetadas exclusivamente a partir de especificações do verbo principal. Essa perspectiva não categoriza sentidos implausíveis aos verbos, caso eles não ocorram em ambientes usuais. Isso porque os verbos se integram com a semântica da construção e são associados a *frames*.

As construções de estrutura argumental envolvem cenas *gestálticas* básicas, como alguém causando o movimento de algo para algum lugar, alguém transferindo algo para outra pessoa, alguém se movendo para algum lugar, alguém causando a mudança de estado de alguma coisa etc.

Qualquer padrão linguístico é reconhecido como uma construção quando alguns aspectos da sua forma ou função não são estritamente previsíveis a partir de suas partes composicionais ou de outras construções reconhecidas. Padrões são armazenados como construções mesmo se são totalmente previsíveis quando eles ocorrem com suficiente frequência.

Historicamente, o grande interesse da pesquisa construcionista foi mostrar que não existe uma separação estrita entre léxico e sintaxe, semântica e pragmática. O enfoque construcionista envolve todos os níveis de análise linguística, desde morfemas até unidades discursivas como os gêneros. Tal abordagem foca principalmente no papel da semântica, da pragmática e da sintaxe. Embora existam estudos que atentem para o fator prosódico em construções linguísticas do inglês, esse tipo de análise mantém-se inédito para o português brasileiro – o que representa o propósito deste trabalho em um empreendimento recente e desafiador sob a ótica da Linguística Cognitiva, de um modo em geral.

2.2.O DISCURSO REPORTADO

Um dos indícios primários na marcação do Discurso Reportado é o léxico-sintático. Se os enunciados são construídos com verbos *dicendi* ou com outra expressão que indique citação, diferentes vozes são introduzidas e serão interpretadas. Falantes podem ancorar deiticamente seus enunciados fazendo referências pessoais, locativas e temporais à situação reportada.

Com ou sem marcadores reportativos ou expressões dêiticas, o DR pode ainda ser marcado na escrita por recursos gráficos; e na fala, pela entoação. Citar consiste em adotar enunciados emanados de outra pessoa, para atingir objetivos argumentativos, por exemplo.

2.2.1.DISCURSO DIRETO (DD)

Tradicionalmente, o DR é categorizado como direto ou indireto (MAYES, 1990; SAKITA, 2002; HOLT & CLIFT, 2007). No DD, há independência sintática, presença de propriedades expressivas, além de delimitação precisa entre as fronteiras. Na concepção rotineira do ouvinte, são colocadas as palavras exatas do falante original. O reportador cita a fala a partir da perspectiva do reportado, emprestando sua voz para ele. Simples exemplos são colocados abaixo:

- a) A menina, em tom zangado, disse:
- Não gosto disso.

O DD pode ocorrer sem verbo *dicendi*, ficando para as aspas o papel de sinalizar na escrita a fala do outro.

- b) A menina, em tom zangado: “não gosto disso”.

O DD com verbo *dicendi*, como poderá ser checado nos dados, é o prototípico entre as construções encontradas, isto é, em termos de frequência; já o DD sem verbo tende a ocorrer em sequências narrativas, em que se prevê facilmente a presença do verbo.

2.2.2.DISCURSO INDIRETO (DI)

No DI, não há independência sintática nem a presença de propriedades expressivas, e sim a existência de complementizador (“que” ou “se”). São as palavras do reportador que transportam o que foi dito ou escrito pelo falante ou escritor original. O reportador interpreta o discurso do reportado e o cita com suas próprias palavras.

Quem reporta não assume plenamente a perspectiva do reportado. Exemplo:

- c) A menina disse, em tom zangado, que não gostava daquilo.

Nesse caso, tornam-se necessárias as modificações quanto às referências dêiticas de tempo verbal, pronomes e advérbios para se fazer a paráfrase do enunciado direto.

2.2.3.DISCURSO INDIRETO LIVRE (DIL)

O DI pode ser denominado livre, quando se misturam marcas atribuíveis à voz do narrador e à voz de outros. Há uma maior liberdade, em que o narrador insere outra fala sem usar as marcas do DD.

- d) A menina perambulava pela sala zangada. Não gosto disso.
Mas ninguém a ouvia.

Nesse caso, não há uma introdução *dicendi* para a fala “não gosto disso”; e nem uma marca de conclusão da citação, com a voz do reportador sendo retomada: “mas ninguém a ouvia”.

Lembramos que tradicionalmente a divisão e a classificação dos tipos de DR são realizadas exclusivamente a partir da modalidade escrita da língua. Existe, então, uma tradição grafocêntrica para se marcar e delimitar o DR. O que veremos, em nossa análise de dados de fala, é que a prosódia, pareada no jogo construcional, sinaliza a função do travessão e das aspas.

3. PROCEDIMENTOS METODOLÓGICOS

3.1. CATEGORIAS DE ANÁLISE

Seguem alguns esclarecimentos acerca de cada uma das categorias empregadas em nossa análise.

Construção de discurso reportado

As construções de DR instanciam o fenômeno discursivo buscado. Ao ouvir cada um dos diálogos completos, identificamos atentamente as ocorrências em que o falante reportava a voz ou o pensamento de outrem ou dele mesmo. Destacamos que, principalmente, nos casos de DIL, acabamos contando com a ajuda dos arquivos de transcrição do *corpus*. Ao consultar esse arquivo, não só confirmamos as ocorrências ouvidas, como também solucionamos algumas dúvidas, dado que o DIL era sempre marcado com as aspas.

Dividimos a construção de DR em duas unidades: a unidade 1 ou unidade introdutora, e a unidade 2 ou unidade reportada. A primeira é aquela que serve como anunciadora da outra voz, seja com o emprego do verbo *dicendi* ou de outro tipo de expressão, como no caso do DIL. A segunda é a fala reportada, isto é, a outra voz que foi inserida no discurso para diversos fins.

Ênfase

A ênfase, dentro de um processo de comunicação, pode ser definida como o ato de acentuar, ressaltar, focalizar, ou pôr em evidência um determinado item no texto. Trata-se de uma proeminência dada a uma sílaba, a uma palavra ou a um sintagma que destaca a informação, colaborando na construção do sentido e sinalizando aspectos que podem ir além da informação textual.

Pausa

A pausa tem uma função aerodinâmica que permite ao falante respirar em momentos oportunos. Esses momentos tendem a ocorrer no final do que chamamos sintagmas entoacionais. Então, a pausa age como “segmentadora” do discurso e pode até ser usada depois de sílabas quando se silaba uma palavra.

Neste estudo, marcamos a pausa quando ocorria o silêncio numa duração mínima de 150 ms. O momento relevante da construção em que foi observada a presença ou não de pausa foi na transição da unidade introdutora para a unidade reportada.

Tom de fronteira

Os tons de fronteira associam-se aos limites de constituintes e caracterizam a modulação melódica do fim de um domínio prosódico. Esse tom é indicado por % (H% e L%)³.

3 Conforme a anotação de Pierrehumbert (1980), o símbolo de porcentagem (%) é empregado ao lado de cada unidade tonal H (*high*/

Frequência Fundamental (F0)

A frequência fundamental (F0) é o número de ciclos, por segundo, de vibrações das pregas vocais. Um Hertz (Hz) equivale a um ciclo por segundo. A frequência de uma voz está relacionada com o tamanho da laringe. Por essa razão, os homens tendem a ter uma F0 mais baixa e, conseqüentemente, mais grave; e as mulheres, mais alta e mais aguda.

F0 média

Para obtermos a F0 média, recorremos a um recurso do próprio programa computacional. Seleccionamos, na tela do PRAAT (BOERSMA & WEENINK, 2011), o trecho da unidade analisada e na opção “pitch”, clicamos em “get pitch”. Ao realizar esse processo, abre-se uma janela com a medida da F0 média em Hz do trecho selecionado.

Gama de Variação (GV)

Costuma-se verificar a organização de uma sequência de mais de dois L's (tons baixos) ou de mais de dois H's (tons altos) sobre uma mesma linha. A diferença entre o valor máximo de F0 da frase e seu mínimo é designada GV e corresponde a todo espaço tonal utilizado pelo falante na produção da unidade. Analisaremos a GV das unidades 1 e 2, com o intuito de checar como o falante usa o espaço tonal ao reportar discursos.

3.2.O CORPUS E O PROGRAMA DE ANÁLISE ACÚSTICA

Os dois diálogos escolhidos para a análise são os inquéritos 147 e 369 do NURC-RJ. Conforme ouvimos os dois diálogos, percebemos que as duas falantes do inquérito 147 não se conheciam; já os dois falantes do inquérito 369 são marido e mulher. As duas falantes do primeiro inquérito serão nomeadas AI e AII; e os dois falantes do segundo inquérito serão aqui reconhecidos como BI e BII.

O programa usado na análise acústica é o PRAAT. Ele oferece recursos técnicos para manipulação, análise, síntese, produção e reprodução das ondas acústicas. O programa foi desenvolvido por Paul Boersma e David Weenink, no Instituto de Ciências Fonéticas da Universidade de Amsterdam, em 1992. Desde então, inúmeras revisões e atualizações do programa foram realizadas.

Analisamos, para o presente estudo, 91 construções de DR, produzidas pelos quatro falantes. Uma vez que esse trabalho surge, principalmente, do empenho em se levar adiante o que Rocha (2004) iniciou, ao apontar para aspectos entoacionais envolvidos na produção de DR, é parte integrante desta análise considerar as ocorrências analisadas e seus tipos construcionais respectivos.

alto) ou L (low/baixo). Forma-se um único tom, que é associado às fronteiras dos domínios prosódicos e não à marcação de proeminência.

4. RESULTADOS E DISCUSSÃO

Abaixo, segue a tabela por meio da qual se comparam as ocorrências dos quatro falantes, divididas nos quatro tipos construcionais:

Falante	Discurso Direto com verbo <i>dicendi</i>	Discurso Direto sem verbo <i>dicendi</i>	D i s c u r s o Indireto	D i s c u r s o Indireto Livre
AI	17	4	4	6
AII	5	2	0	5
BI	24	3	8	4
BII	7	1	1	0
Total do tipo	53	10	13	15

Tabela 1: Comparação - tipos construcionais

Passemos a discutir o que encontramos especificamente para cada tipo de construção.

4.1.DISCURSO DIRETO COM VERBO *DICENDI*

eu ainda disse ó quem me roubou foi esse individuo (BII8)

Correspondendo a 53 ocorrências, das 91 analisadas, esse tipo de construção pode ser considerado o prototípico, em função do critério frequência. Somente a falante AII não teve a maioria de suas ocorrências de DD com verbo *dicendi*.

Dos aspectos prosódicos analisados, os mais recorrentes nesse tipo de construção foram o tom L% na unidade introdutora e a ênfase na unidade reportada. Das 49 construções com ênfase na unidade reportada, 29 são do tipo DD com verbo *dicendi*. Isso nos permite inferir que a ênfase tende a vir depois do prefácio *dicendi* prototípico. Além disso, entendemos que a descida é o movimento padrão para encaixar o DD com verbo *dicendi*.

4.2.DISCURSO DIRETO SEM VERBO *DICENDI*

quando ele entrou eu hã (AII2)

Esse tipo de construção foi o que apresentou menos ocorrências. Um fato curioso é que das dez ocorrências, oito apresentaram ênfase. Parece-nos que, ao omitir o verbo *dicendi*, torna-se importante ressaltar alguma parte da unidade reportada. A outra característica mais recorrente foi o tom L% na unidade 1, ocorrendo em todas as construções de DD sem verbo.

Comparando os falantes que produziram todos os tipos de DR, realizamos a seguinte observação quanto à presença de pausa: das quatro ocorrências desse tipo de construção produzidas por AI, há apenas um exemplo em que não há pausa entre as duas unidades. BI, por sua vez, fez a pausa em uma

das três construções desse tipo. Pode ser que falantes utilizem a pausa para ajudar a sinalizar que, mesmo sem o verbo *dicendi*, há outra voz sendo introduzida no discurso.

4.3. DISCURSO INDIRETO

ela dizia na minha cara que o desenho tava feio (AI20)

Dentre os quatro falantes, somente AII não produziu nenhuma construção de DI. Para esse tipo construcional, as características mais recorrentes foram o tom L% e as medidas tidas como padrão, isto é, médias de F0 mais altas nas unidades introdutoras e GVs mais altas nas unidades reportadas.

Uma característica que buscamos analisar nesse tipo construcional é a presença ou não de pausa. Assim, poderíamos discutir a hipótese de que há um contínuo entre a unidade 1 e a unidade 2 quando se trata de DI.

4.4. DISCURSO INDIRETO LIVRE

de modo que ele sabe tudo sabe a corrente sabe que tem buraco ali tem um buraco não entra(BI33)

De todos os falantes, BII não produziu nenhuma construção com DIL, embora esse tenha sido o segundo tipo mais produzido no total. Assim como nos outros, o tom L% foi característica recorrente, seguido da pausa e da ênfase.

Gostaríamos de chamar a atenção para o fato de que nas construções de DIL de AI e BI, a F0 média é mais alta na unidade reportada em 50% dos casos. Isso vai ao encontro da ideia de que, quando não há um introdutor *dicendi*, a F0 do DR é mais alta.

Observamos, com a análise dos dados, que o DIL pode ser introduzido das mais variadas formas. Alguns exemplos de unidades que serviram para introduzir o DIL:

Falante AI:

“a gente na fila”, “aquela coisa eterna”, “eis que chegamos”, “vô pro paraná vô pra são paulo vô pra acolá”, “minha mãe era orgulhosíssima sabe”, “é uma malícia incrível”.

Falante AII:

“ele ficou tão tranquilo”, “avançam avançam”, “então a gente procura”, “sai dali”, “ela ficou maravilhada”.

Falante BI:

“daqui a pouco já é uma amiga que telefona encontro uma amiga na”, “até a minha empregada ficou tão impressionada”, “de modo que ele sabe tudo sabe a corrente sabe que tem buraco”, “há um ano atrás o exame era exatamente a mesma coisa”.

Essa variedade de “introdutores” diz respeito, justamente, ao fato de que no DIL não temos uma expressão prevista para introduzir o DR, pelo menos não do ponto de vista sintático e pragmático.

As expressões que precedem o DIL, de alguma maneira, contextualizam a cena em que alguém reporta uma fala. Essas cenas, mesmo que não explicitem a fórmula *dicendi*, permitem-nos conceber um falante que diz alguma coisa, pré-enquadrando a possibilidade de uma fala reportada em uma narrativa oral. Normalmente, esse tipo de introdutor, além de anunciar o DIL, traz informações circunstanciais sobre o falante reportado ou mesmo sobre a cena narrada.

Esse pré-enquadre narrativo está representado com o caso de “aquela coisa eterna”. Além de AI reportar o que a mãe disse, subentende-se que se trata de um discurso sempre proferido por ela. Podemos simular a construção da falante com a fórmula *dicendi* exemplificada em “minha mãe sempre diz”. Ao ouvir a proclamação da fala reportada, percebemos uma voz cansada e entediada.

4.5. ARTICULAÇÃO TEÓRICO-ANALÍTICA

Apresentamos agora de que maneira nossos resultados corroboram a reivindicação principal de que a prosódia organiza a cena interacional e sinaliza informações para a construção do sentido que, muitas vezes, não são recuperadas via sintaxe ou semântica.

Inicialmente destacamos que analisar dados de fala espontânea é, de fato, encarar a complexidade da “chama” (MIRANDA, 2010, p. 59). Quando atribuímos tendências prosódicas para construção de DR, apostamos fortemente na inserção da contraparte prosódica no pólo formal das construções linguísticas.

Grande parte dos estudiosos reconhecem padrões sentenciais em termos de prosódia. O desafio deste trabalho ultrapassou esse reconhecimento, visto que colocou em foco construções produzidas em contextos absolutamente naturais. Isso demonstrou a complexidade da fala em detrimento dos padrões *default* estereotipados. Não é o caso de se desconsiderar a importância de perfis prosódicos de sentenças declarativas, interrogativas e exclamativas, por exemplo, mas quando nos deparamos com a fala real e espontânea, subtipos de sentenças consolidam um espectro altamente rico e variado, bem diferente dos padrões já conhecidos na literatura.

Pensemos, então, que a construção-mãe para os casos de DR seria o tipo DD com verbo *dicendi* “dizer”. Um exemplo do nosso *corpus* seria:

“ela olhava e dizia assim eu não ACREDITO isso é mentira eu não ACREDITO” (BI5)

As letras maiúsculas, mantidas em “acredito”, são para marcar as ênfases. Esse tipo de construção foi o que pareceu mostrar o ambiente mais propício para se enfatizar partes do discurso, uma vez que a citação é direta. Temos, assim, para esse tipo construcional:

DD com verbo <i>dicendi</i>
Aspectos sintáticos: [SUJ disse/dizia OBJETO ORACIONAL]
Aspectos semânticos: alguém diz algo que outro disse
Aspectos pragmáticos: construção mais mimética, com mudança de <i>footing</i> ⁴ etc
Aspectos prosódicos: A unidade introdutora tende a se encerrar com um tom de fronteira L%. Quando comparadas as unidades 1 e 2, as medidas de F0 média tendem a ser mais altas na unidade introdutora, e a GV, na unidade reportada. O sintagma entoacional que a fala reportada pode receber ênfase.

Tabela 2: descrição do tipo construcional DD com verbo *dicendi*

A partir dessa construção-mãe, temos a rede de construções de DR com mais os outros três tipos que se ligam por uma relação de herança. A construção DD sem verbo *dicendi* pode ser considerada uma sub-parte da construção-mãe. Abaixo, um exemplo retirado do nosso *corpus*:

“*ai nós presidente do brasil mas cumé que pode*” (BI31)

Lembrando que, nesse caso, houve uma pausa entre o “*ai nós*” e a fala reportada. O apagamento do verbo *dicendi* se percebe em ambientes de sequências de DRs, em que a construção do sentido não é prejudicada por essa omissão. Além disso, a pausa se verificou frequente como um recurso prosódico que sinaliza, apesar da ausência do verbo *dicendi*, a fala do outro. Segue uma descrição desse tipo construcional:

DD sem verbo <i>dicendi</i>
Aspectos sintáticos: [SUJ OBJETO ORACIONAL]
Aspectos semânticos: alguém diz algo que outro disse
Aspectos pragmáticos: construção mais mimética, com mudança de <i>footing</i> etc
Aspectos prosódicos: A unidade introdutora tende a se encerrar com um tom de fronteira L%. Quando comparadas as unidades 1 e 2, as medidas de F0 média tendem a ser mais altas na unidade introdutora, e a GV, na unidade reportada. Tendência em existir uma pausa entre a unidade introdutora e a unidade reportada.

Tabela 3: descrição do tipo construcional DD sem verbo *dicendi*

O outro tipo construcional herdado é o DI. Um exemplo encontrado em nossos dados é:

4 Esse conceito envolve mudanças na atuação do participante. Estamos nos referindo ao alinhamento] da postura e da projeção de um interlocutor com o outro, consigo próprio e com o discurso construído, negociando sua participação na interação (GOFFMAN, 1998). Mudamos de *footing*, por exemplo, quando reportamos nossa própria fala ou quando reportamos a fala do outro, mesmo que em um ou em outro caso sejam nossas próprias ideias que estejam sendo articuladas.

“e dizendo que pra semi-final não foi conseguida a venda da entrada”(BII1)

Dentre os quatro tipos de DR (DD com verbo *dicendi*, DD sem verbo *dicendi*, DI e DIL), o DI foi o terceiro em número de ocorrências, todas em terceira pessoa. Em termos prosódicos, pudemos confirmar a tendência do DI em não haver pausa entre as unidades 1 e 2. Abaixo, o que podemos descrever sobre esse tipo construcional:

Discurso Indireto
Aspectos sintáticos: [SUJ disse/dizia OBJETO ORACIONAL]
Aspectos semânticos: alguém diz algo que outro disse
Aspectos pragmáticos: construção menos mimética, sem mudança de <i>footing</i> etc
Aspectos prosódicos: A unidade introdutora tende a se encerrar com um tom de fronteira L%. Quando comparadas as unidades 1 e 2, as medidas de F0 média tendem a ser mais altas na unidade introdutora, e a GV, na unidade reportada. Tendência em não existir pausa entre a unidade introdutora e a unidade reportada.

Tabela 4: descrição do tipo construcional DI

E o último tipo construcional, o DIL, é o que apresenta a maior divergência sintática da construção-mãe. E é nesse tipo em que acreditamos haver a maior contribuição do contexto conversacional, bem como da prosódia. O exemplo trazido dos nossos dados é:

“Avançam avançam ah da onde você é?” (AII4)

Nesse exemplo, percebemos a mudança de *footing* quando a falante tenta reproduzir o modo como as garotas faziam a pergunta aos rapazes. Embora não apresente o prefácio *dicendi* prototípico, que seria “elas diziam” por exemplo, esse tipo construcional se mostrou eficiente e produtivo na língua. Abaixo, sua tabela descritiva:

Discurso Indireto Livre
Aspectos sintáticos: [X OBJETO ORACIONAL]
Aspectos semânticos: alguma expressão anuncia algo que outro disse
Aspectos pragmáticos: construção mais mimética, com mudança de <i>footing</i> etc
Aspectos prosódicos: A unidade introdutora tende a se encerrar com um tom de fronteira L%. Quando comparadas as unidades 1 e 2, a GV tende a ser mais alta na unidade reportada. A F0 média da unidade reportada tende a ser mais alta do que a F0 média da unidade que a precede.

Tabela 5: descrição do tipo construcional DIL

Lembrando que “X” é representado por uma variedade infinita de expressões que podem servir como “introdutoras” do DR. É interessante notar que esse é o único tipo construcional que não “respeita” a fusão dos papéis semânticos e dos papéis argumentais / sintáticos (Princípio da Coerência Semântica).

Nos três tipos anteriores, temos o papel argumental sujeito se fundindo com o papel semântico de alguém que fala o que o outro diz. Nesse caso, não temos esses papéis na construção. O DIL é, linguisticamente, anunciado de várias formas e mesmo que o interpretemos, não temos o sujeito/reportador expresso. Cabe ao discurso fornecê-lo.

O pontapé inicial dado por Rocha (2004) nos permitiu trilhar um caminho que sugeriu a possibilidade de a prosódia atuar decisivamente no pólo formal das construções de DR. Dado o caráter intuitivo da análise do autor, tivemos a ideia de reeleger o DR como objeto de estudo, no sentido de tornar mais criteriosa a busca de evidências externas no mapeamento dos perfis entoacionais das construções DR. Assim, pudemos verificar, com o auxílio de um programa de análise acústica, o PRAAT, a plausibilidade das intuições acerca do fenômeno discursivo de reportar vozes.

Quando trouxemos à tona os tipos construcionais formulados por Rocha (2004) confirmamos algumas hipóteses e negamos outras. Vale lembrar que, mesmo confirmadas, apoiamo-nos em tendências entoacionais para o DR e não em limites categóricos. Em vista da quantidade de dados gerados, consideramos o padrão como um resultado ainda distante de alcançar – se é que ele pode ser considerado plausível no terreno da fala espontânea, dada a sua complexidade e variedade de expressão.

Confirmamos, por exemplo, a hipótese de que a transição entre unidade introdutora e unidade reportada no DI é marcada por um contínuo, sendo minimizados os casos em que se nota a presença de pausa. Nos termos de Rocha, encontramos a referência a uma distinção tonal que sinalizaria uma diferença entre DD e DI. Essa distinção tonal entre unidade introdutora e unidade reportada ocorreria somente no primeiro caso, e não no segundo.

Ainda que tenhamos encontrado somente uma ocorrência de DD com verbo *dicendi* em 2ª pessoa nos nossos dados, comparando 1ª e 3ª pessoas, não parece haver a diferenciação do modo como sugeriu Rocha (2004) (sendo mais marcada a citação conforme passa da 1ª para a 3ª pessoa). Houve, inclusive, um caso de uma falante que produziu mais ênfases em construções de primeira pessoa. Nesse caso, o caráter mais mimético também está associado aos casos de autocitação. Enfim, de alguma forma a prosódia sinaliza as diferenças quanto às pessoas do discurso. Vale mencionar que, para efeito de comparação, não submetemos ao PRAAT os dados de Rocha, em função de sua baixa qualidade acústica, o que comprometeria a análise.

Contudo, não se pode deixar de considerar as diferenças significativas de molduras comunicativas que encontramos no nosso *corpus*, quando o comparamos com o *corpus* de Rocha. As construções de DR do presente trabalho são inseridas em narrativas que falantes produzem em uma moldura de diálogo informal, a partir de temas que os documentadores sugerem, para compilação de um *corpus* linguístico. Assim, a fim de se alcançar certa “naturalidade”, os falantes tratam de suas vidas, incluindo narrativas sobre memórias da infância, de viagens etc. Já as construções analisadas em Rocha (2004) foram produzidas por participantes do *reality show Big Brother Brasil* em 2002, que foram deslocados de seus cotidianos particulares para um palco de representações improvisadas,

espiado pelas câmeras de TV. Esses falantes estão imbuídos do propósito de reportar outros discursos na tentativa de manter suas próprias faces e serem bem quistos pelo público. Afinal, estavam em um jogo que valia um prêmio alto em dinheiro.

O que Jansen, Gregory e Brenier (2000) apresentaram para dados do inglês quando defendem que o DD é o mais proeminente dentre os tipos de DR também se aplica aos dados de PB analisados nesse estudo. Essa aproximação de resultados advém do fato de que grande parte das nossas ocorrências com ênfase nas unidades reportadas enquadram-se no tipo DD. Das 49 construções com ênfase na unidade reportada, 37 foram DD, incluindo as construções com e sem verbo *dicendi*.

Günthner (1998) também pode ser lembrado, uma vez que nossos resultados corroboram sua assertiva de que o DR, quando não precedido por introdutores, apresenta uma F0 mais alta. Os casos de DIL, isto é, em que não há um introdutor *dicendi* - seja com a presença do verbo ou da estrutura sintática introdutora -, pelo menos na metade das ocorrências, geraram uma F0 média mais alta na unidade reportada.

Ainda podemos registrar que nossos resultados também vão ao encontro do que Maia Rocha (2011) propõe para a unidade informacional do introdutor locutivo – em que se anunciam os casos de DR. Confirmamos que essa unidade é delimitada por um perfil melódico descendente e por uma F0 média mais alta do que a da unidade posterior, exceto no DIL, que justamente não apresenta a unidade do introdutor locutivo.

Sintetizamos os principais resultados para cada tipo construcional. Destaca-se que as tendências prosódicas relacionadas ao tom de fronteira da unidade 1 e à GV da unidade 2 mantêm-se para os três primeiros tipos construcionais:

- DD com verbo *dicendi*: trata-se do tipo construcional que assume o prefácio *dicendi* prototípico. E é nessa construção que a unidade reportada tende a receber uma ou mais ênfases. Assim, quando se quer enfatizar a voz do outro, o falante a anuncia no modo prototípico;

- DD sem verbo *dicendi*: as construções desse tipo tendem a apresentar uma pausa entre a unidade introdutora e a unidade reportada. O uso da pausa, nesse momento, parece, assim, estar relacionado ao apagamento do verbo *dicendi*;

- DI: com relação a esse tipo construcional, confirmou-se a tendência em não haver pausa separando as duas unidades. É o caso em que o reportador realiza uma paráfrase e os devidos ajustes dêiticos e, então, não faz pausa para marcar a distinção de falas, possivelmente por conta da presença do complementizador integrando as duas unidades;

- DIL: distinguindo-se dos demais tipos construcionais, essa foi a construção que apresentou, em grande parte dos dados, a F0 média mais alta na unidade reportada e não na unidade “introdutora”,

que é a tendência para os outros três tipos. Acredita-se que, na mesclagem de vozes, sem introdução expressa, o falante eleva a F0 para marcar o discurso do outro.

Com relação às outras contrapartes construcionais dos diferentes tipos de DR, em síntese, apresenta-se:

- Aspectos sintáticos: o tipo construcional que mais se distingue sintaticamente é o DIL;
- Aspectos semânticos: o tipo construcional que mais se distingue semanticamente também é o DIL;
- Aspectos pragmáticos: o tipo construcional em que parece ser menos nítida a mudança de *footing* é o DI.

Não se mostrou uma tarefa fácil determinar essas características, visto que nossos dados são de fala espontânea – imprevisíveis e incontroláveis. Todavia, não podemos deixar de retomar e defender a importante tese de que a fala não deve ser tratada apenas como escrita e de que, conseqüentemente, a noção de construção gramatical precisa ser enriquecida com informações assumidas pela prosódia. Por isso, preocupa-nos registrar as minuciosas descrições que os dados permitiram extrair.

5. CONSIDERAÇÕES FINAIS

Ao longo deste trabalho, podemos observar que o arcabouço teórico da Gramática das Construções (GOLDBERG 1995; 2006) é fortemente respaldado por evidências linguísticas de modalidade escrita. Quando nos deparamos com os estudos dessa autora, notamos que seus exemplos, ao contrário deste estudo, são de sentenças artificialmente criadas que instanciam, por exemplo, construções de movimento causado, ditransitivas, resultativas, etc. Dessa forma, sua teoria tem grande peso nos domínios sintático e semântico em termos de pareamento entre forma e significado. Não há menção ao polo fonológico no âmbito da construção gramatical, trabalho este iniciado por Rocha (2004) e desdobrado por esta tese.

Os resultados do presente trabalho giram em torno da aplicação de mais um nível no pólo formal de construções faladas do português brasileiro. Estamos nos referindo, com isso, ao nível prosódico que compõe a contraparte formal das construções. Pudemos evidenciar que a prosódia favorece a interpretação, por exemplo, de casos de DIL, em que a semântica e a sintaxe são distintas dos outros moldes de DR, mas trata-se ainda de um caso de DR. Coligada à pragmática, a prosódia pode ser decisiva nesses casos. Além disso, os tipos construcionais de DD sem verbo *dicendi* são exemplares ideais da não anterioridade da sintaxe no pareamento construcional. Novamente, a prosódia e a pragmática constituem caminhos eficientes que o falante percorre para introduzir a voz do outro.

O presente trabalho encerra-se, portanto, apostando em uma nova versão teórica da Gramática das Construções, que contemple as construções da fala como faz com as construções da escrita.

REFERÊNCIAS

- BOERSMA, P.; WEENINK, D. *Praat: Doing Phonetics by Computer*. Versão 5.3.01, 2011.
- GOFFMAN, E. Footing. In: RIBEIRO, B. T.; GARCEZ, P. M. (org). *Sociolingüística interacional: antropologia, lingüística e sociologia em análise do discurso*. Porto Alegre: AGE, 1998. p. 70-97.
- GOLDBERG, A. *A construction grammar approach to argument structure*. Chicago: The University of Chicago Press, 1995.
- _____. *Constructions at work*. New York: Oxford University Press, 2006.
- GÜNTNER, S. *Polyphony and the „layering of voices“ in reported dialogues: an analysis of the use of prosodic devices in everyday speech*. Konstanz: Inlist3, 1998.
- HOLT, E.; CLIFF, R (Eds.) *Reporting talk: reported speech in interaction*. Cambridge: Cambridge University Press, 2007.
- JANSEN, W.; GREGORY, M.; BRENIER, J.M. Prosodic correlates of directly reported speech: evidence from conversational speech. In: *proceedings of the ISCA workshop on prosody in speech recognition and understanding*. USA, 2001.
- MAIA ROCHA, B. *A unidade informacional de introdutor locutivo no português brasileiro: uma análise baseada em corpus*. Dissertação (Mestrado em Linguística). Belo Horizonte: UFMG, 2011.
- MAYES, P. Quotation in spoken English. *Studies in language* 14:325-363. 1990.
- MIRANDA, N. S. O caráter partilhado da construção da significação. *Revista Veredas*, v. 5, n. 1, Juiz de Fora, 2001. p. 57-81.
- PIERREHUMBERT, J. *The Phonology and phonetics of english intonation*. 1980. 402f. Tese (Doutorado em Linguística) – Department of linguistics and philosophy, Massachusetts Institute of technology, Indiana University Linguistics Club, 1980.
- ROCHA, L. F. M. *A construção da mimesis no reality show: uma abordagem sociocognitivista para o discurso reportado*. Tese (Doutorado em Linguística). Rio de Janeiro: UFRJ, 2004.
- SAKITA, T. I. *Reporting discourse, tense, and cognition*. Elsevier, 2002.

Recebido em 31/01/2016

Aceito em 12/04/2016