
LINGUÍSTICA COGNITIVA E ENSINO DE REDAÇÃO CIENTÍFICA: UM ESTUDO DA NARRATIVIDADE E FIGURATIVIDADE DE *ABSTRACTS* COM USO DE TECNOLOGIA

COGNITIVE LINGUISTICS AND THE TEACHING OF SCIENTIFIC WRITING: A STUDY OF NARRATIVITY AND FIGURATIVITY OF ABSTRACTS WITH THE USE OF TECHNOLOGY

Rosana Ferrareto Lourenço Rodrigues¹

RESUMO

Investimento e atenção para o ensino de redação científica nas universidades são evidentes, não só porque alunos precisam produzir, mas também porque a publicação de docentes é medida avaliativa dos programas de (pós)graduação. Muitos modelos teóricos, tais como o CARS (SWALES, 2004), e manuais de normalização, tais como a ABNT, para a boa prática da escrita científica e a publicação bem-sucedida, têm sido aplicados para atender à demanda. Apesar de os cientistas escreverem no formato esperado, muitos não conseguem publicar. Partindo da concepção de escrita como processo de aprendizagem (FLOWER; HAYES, 1981), a proposta deste estudo é acrescentar aos modelos já convencionais constructos da Linguística Cognitiva, tais como narrativa cognitiva, esquemas imagéticos, frames semânticos e integração conceptual (TURNER, 2008). O objetivo é investigar a narratividade e a figuratividade do texto científico, verificando como os cientistas enquadram, desenvolvem e comunicam o processo de pesquisa e os conceitos científicos da sua área de estudo. Analisamos um corpus multidisciplinar de mil abstracts científicos publicados em periódicos de alto impacto na Web of Science. Adotamos a abordagem metodológica de Data Science, a partir do uso de ferramentas tecnológicas, como parsers semânticos e concordanciadores online, para a anotação automática de frames e busca de construções linguísticas baseadas em padrões gramaticais. Além da contribuição teórica, este estudo traz aplicações pedagógicas e tecnológicas, a partir da proposta de um modelo conceptual para a comunicação científica e do uso de ferramentas tecnológicas para a pesquisa linguística e o ensino de redação científica.

Palavras-chave: Linguística cognitiva. Narratividade. Figuratividade. Redação científica. Tecnologia.

ABSTRACT

Investment and attention to the teaching of scientific writing in universities are evident, not only because students need to do research, but also because professors need to publish. It is an evaluative measure of undergraduate and graduate courses. Many conceptual models and manuals, such as CARS (SWALES, 2004) for writing and APA Style for publication, have been used for good writing and publishing practices. Although scientists write in the expected format, many of them are unable to publish. Based on the conception of writing as a learning process (FLOWER; HAYES, 1981), our purpose is to add to the already conventional models, constructs of Cognitive Linguistics, such as cognitive narrative, image schemas, semantic frames and blending (TURNER, 2008). Our aim is to investigate narrativity and figurativity in scientific texts, by verifying how scientists frame, develop and communicate the research process and scientific concepts. We analyzed a multidisciplinary corpus of a thousand scientific abstracts published in high impact journals on the Web of Science. Our methodological approach is Data Science: we have used technological tools, such as a semantic parser and a concordancer, for automatically annotating frames and for searching linguistic constructions based on grammatical patterns. In addition to the theoretical contribution, this study brings pedagogical and technological applications, proposing a conceptual model for science communication, and the use of technological tools for research on Linguistics and for the teaching of scientific writing.

Keywords: : Cognitive Linguistics. Narrativity. Figurativity. Scientific writing. Technology.

1 Doutora em Linguística e Língua Portuguesa – Instituto Federal de Educação, Ciência e Tecnologia de São Paulo.

1. INTRODUÇÃO

Toda universidade oferece, em todos os cursos de graduação e pós-graduação, a disciplina de Metodologia de Pesquisa Científica, cujo escopo envolve a redação de trabalhos acadêmicos, tais como monografias, dissertações, teses, projetos, relatórios e artigos. Registrar o processo e os resultados de uma pesquisa faz parte do fazer científico. Nessa disciplina, não é apenas necessário convencer os cientistas em formação da importância de se fazer ciência com qualidade e da necessidade de comunicar com eficácia. Há de se desenvolver a habilidade de comunicar. Essa habilidade demanda saberes não só cognitivos e linguísticos, para ensinar a pensar com lógica e comunicar com criatividade, mas requer saberes tecnológicos e desenvolvimento de pensamento crítico.

Diante dessa demanda, o câmpus São João da Boa Vista, do Instituto Federal de Educação, Ciência e Tecnologia de São Paulo iniciou, em 2015, um projeto de extensão para a formação de cientistas – o Laboratório de Letramento Técnico-Científico (Laletec)¹. O trabalho desenvolvido nesse projeto está ancorado em três eixos – o da competência informacional, o da metodologia de pesquisa e o da redação científica, coordenados respectivamente por uma bibliotecária, um cientista da computação e uma linguista. Um diferencial do projeto é o uso de ferramentas tecnológicas para a realização de todos os estágios da pesquisa: fontes de informação na web, gerenciadores de referência online, softwares estatísticos, ambientes de escrita semiautomática, entre outros.

Após três anos de trabalho, em 2018, o projeto buscou robustez teórica em ares internacionais. Uma de suas coordenadoras, a linguista, realizou um estágio pós-doutoral em um laboratório de línguas e tecnologia² no Departamento de Ciências Cognitivas de uma universidade estadunidense. Essa experiência de internacionalização transformou o projeto

de extensão em um programa de pesquisa sustentável³, cuja principal investigação científica sobre a Comunicação Científica consiste no estudo dos padrões narrativos e figurativos do texto científico.

Este artigo relata os principais resultados dessa empreitada até o momento e tem como objetivo contribuir teórica e pedagogicamente, por meio de uma análise de *abstracts* científicos, sobre como os cientistas concebem, descrevem e comunicam conceitos e processos de sua área de estudo; e por meio da proposta de um modelo pedagógico para o ensino de redação científica. A partir da aplicação de constructos da Linguística Cognitiva e do uso de ferramentas tecnológicas para a análise linguística e para o ensino de redação científica, nosso intuito é contribuir para a comunicação científica com o desenvolvimento da lógica e da criatividade do cientista.

2. LINGÜÍSTICA COGNITIVA E REDAÇÃO CIENTÍFICA

Nesta seção, abordaremos maneiras pelas quais a teoria da Linguística Cognitiva contribui para a redação científica. Nosso ponto de partida é discutir brevemente sobre até que ponto alguns dos modelos conceptuais convencionais para a escrita e normalização de textos acadêmicos contribuem para uma comunicação científica eficaz. Em seguida, o foco é apresentar os conceitos de narratividade e figuratividade no texto e no processo de escrita científica e propor um modelo pedagógico de ensino dessa habilidade, baseado nos processos de pesquisa e comunicação. Nesse percurso, iremos pontuando como a cognição e o uso da língua estão em jogo no fazer científico, no pensamento e na escrita do cientista.

Para começar, ressaltamos que alguns modelos de escrita científica e de normalização do texto científico são amplamente reconhecidos na literatura. O manual de normalização de trabalhos acadêmicos da

1 <https://sbv.ifsp.edu.br/laletec>

2 <http://www.redhenlab.org>

3 https://www.youtube.com/watch?v=k-nhMj-1Ar8&feature=emb_logo

Figura 1 - Comunicação Científica: cognição/pesquisa & língua/comunicação.



Fonte: Rodrigues (2019, p. 66).

Associação Brasileira de Normas Técnicas (ABNT) é adotado em todo o Brasil. Entre os linguistas professores de línguas, o modelo conceptual CARS (Create a Research Space) não só orienta a organização retórica de gêneros acadêmicos em três movimentos – estabelecimento do território, estabelecimento do nicho e ocupação do nicho – mas também desempenha um papel fundamental nas pedagogias de gênero e ferramentas de escrita. Esses movimentos são acompanhados na normalização proposta pela ABNT para a escrita de resumos científicos – tema, problema (estabelecimento do território), objetivo, metodologia (estabelecimento do nicho) e conclusão (ocupação do nicho). No Laletec, já utilizamos essas diretrizes na produção de material didático e na oferta de cursos online e videoaulas⁴.

O fato é que, embora muitos cientistas escrevam seus trabalhos acadêmicos no formato esperado, grande parte ainda não consegue publicá-los. Essa constatação nos fez levantar questionamentos, tais como: modelar a narrativa de pesquisa subjacente a um artigo científico em uma estrutura imagético-esquemática em três atos – começo (tema, problema, objetivos); meio (metodologia) e fim (resultados e conclusão) – não viabilizaria a escrita de maneira mais eficaz? Além dessa lógica da macroestrutura textual, como estariam estruturados linguisticamente os

novos conceitos e processos científicos a partir de construções lexicais e gramaticais claras e criativas?

Além dessas questões cognitivas e linguísticas envolvendo a escrita científica, temos uma outra preocupação no que tange à comunicação entre cientistas: em ambientes interdisciplinares, em que os cientistas precisam entender conceitos de outras áreas do conhecimento para conduzir pesquisas, escrever e publicar (afinal, publicam artigos em revistas especializadas para que seus pares leiam), que processos usam para essa compreensão? Funcionaria ancorar novas informações no repertório prévio para construir conhecimento? Como?

Essas hipóteses nos levaram a investigar as maneiras pelas quais a Linguística Cognitiva contribui para a redação científica. Uma delas é a narratividade. A outra é a figuratividade. Por isso, a seguir, apresentamos uma breve definição desses conceitos e dos constructos narrativa cognitiva, esquemas imagéticos, frames e integração conceptual (analogia/metáfora), dentro da definição que cunhamos para o conceito “comunicação científica”. É assim que propomos este estudo sobre Comunicação Científica a partir das lentes da Linguística Cognitiva, para fins de ensinar redação científica.

2.1 Comunicação científica

A comunicação científica é uma

4 <http://mooc.ifsp.edu.br/course/redacao-cientifica-com-tecnologia-laletec/intro>;
<https://www.youtube.com/channel/UC48zzdUzvEzvisbdOj3tyMg>.

mescla de como os pesquisadores estruturam sua prática científica e como a expressam ao escrever textos acadêmicos. Está entre o processo de pesquisa invocado na mente do leitor, que interpreta o texto enquanto o lê, e o processo de comunicação evocado pelo material linguístico (Figura 1).

Nesse movimento entre cognição e pesquisa & língua e comunicação, o cientista organiza logicamente seu processo de pesquisa num fio condutor lógico, e registra no texto as etapas do fazer científico, por meio de construções linguísticas que sejam suficientemente claras para comunicar um conceito em dada área do conhecimento ou explicar um processo, que sejam novos e/ou abstratos.

Por exemplo, ao descrever o conceito “buraco negro” e explicar como seu processo funciona no universo, um astrofísico, na revista de ciências estadunidense *Omni*⁵, explica que “se você se encontrasse à beira de um buraco negro e, de alguma maneira milagrosa, estivesse vivo e consciente, seria como se estivesse pendurado na Golden Gate bridge com a metade da população do Canadá agarrada a seus tornozelos” (FITZPATRICK, 2019). Trata-se de uma analogia para definir o buraco negro como uma região do espaço-tempo em que o campo gravitacional é tão intenso que nada pode escapar dela. Organizada narrativamente, a ideia é comunicada por meio desse recurso linguístico – a analogia – que compara uma situação complexa, nova e abstrata (o buraco negro) a uma situação simples, familiar e concreta, que é o fato de se imaginar pendurado em uma ponte gigantesca com um peso enorme sobre si. Figurativamente, e não literalmente, fica evidente a ideia do buraco negro como uma força gravitacional intensa.

A seguir, conceituaremos a narratividade e a figuratividade aplicadas ao processo de comunicação de conceitos e processos científicos.

2.2 Narratividade

As histórias são um instrumento humano para construir sentido através de muitos domínios do conhecimento e tipos de atividade. Este é o conceito de narratividade na Linguística Cognitiva: narrar é um artefato cognitivo humano, inerente ao seu processamento mental. A noção de história denota tanto uma estrutura cognitiva abstrata quanto os traços materiais dessa estrutura refletidos na escrita, na fala ou em qualquer outro canal de representação do pensamento na linguagem (HERMAN, 2003). É também um processo mental de integração conceptual entre muitas histórias que protagonizamos, a partir das quais interagimos socialmente (TURNER, 2003).

As narrativas e suas representações são viabilizadas por esquemas imagéticos e acontecem dentro de cenas comunicativas, denominadas frames. Esquemas imagéticos são uma representação conceptual relativamente abstrata que surge diretamente de nossa interação cotidiana e observação do mundo ao nosso redor. Derivam da experiência sensorial e perceptiva, da nossa experiência corporificada. Por exemplo, nossa fisiologia garante que nosso eixo vertical, que interage com a gravidade, dê sentido a um resultado de como interagimos com nosso ambiente (JOHNSON, 1987 EVANS, 2007). Alguns dos principais esquemas imagéticos utilizados para estruturar o pensamento científico são: ORIGEM-PERCURSO-DESTINO, como instanciação narrativa da pesquisa; CONTÊINER, para a categorização de conceitos; DINÂMICA DE FORÇAS, para a descrição de processos; VERTICALIDADE, para a funcionalidade de resultados, entre outros.

O frame⁶ é uma esquematização da experiência (uma estrutura de conhecimento), representada no nível conceptual e mantida na memória de longo prazo e que relaciona elementos e entidades associados a uma cena,

5 <https://omnimagazine.com>

6 Há uma base de dados para consulta dos principais frames na comunicação humana: https://framenet.icsi.berkeley.edu/fndrupal/framenet_search.

situação ou evento culturalmente incorporado da experiência humana (FILLMORE, 1982 apud EVANS, 2007). Por exemplo, a metáfora “a vida é uma jornada” pode ser pensada para o contexto científico para que se imagine a “ciência como uma jornada”. Nessa instanciamento metafórica narrativa e esquemático-visual (ORIGEM-PERCURSO-DESTINO), há dois frames integrados: o da ciência (cujos elementos são o cientista, a investigação, os materiais e métodos, as descobertas etc) e o da jornada (em que figuram a viagem, a direção, a bagagem, o itinerário etc). Pareadas, essas duas cenas comunicativas evocam experiências sensoriais, culturais e conceptuais, a partir de projeções entre seus elementos componentes, para transmitir uma ideia abstrata, um conceito (a ciência, a vida) de maneira mais clara e concreta e menos arbitrária (a jornada).

Integrações conceptuais acontecem no processo narrativo via esquemas imagéticos, a partir do qual frames são mobilizados. Muitas vezes, esses constructos cognitivos são manifestados linguisticamente por meio de analogias e metáforas. Essas construções linguísticas estão no escopo do que entendemos como figuratividade. Usar palavras de maneira figurativa significa não as usar em seu significado literal, mas de uma maneira que torne uma descrição mais interessante ou impressionante⁷. É um processo criativo que propõe significados paralelos, mas correlacionados.

2.3 Figuratividade

A capacidade humana de tramar histórias mentais que correm paralelamente a uma história em que habitamos - por exemplo, a ciência vista como uma jornada ou a pesquisa vista como uma expedição tão complexa quando a do Santo Graal⁸ – é chamada de *blending* ou integração conceptual. Nesse feito imaginativo, a origem das ideias acontece

7 <https://www.macmillandictionary.com/dictionary/british/figurative>, tradução nossa.

8 <https://super.abril.com.br/mundo-estranho/o-que-e-a-lenda-do-santo-graal/>

na criação de uma terceira história, a partir de duas que se integram – a vivida e aquela imaginada, que buscamos no nosso repertório/memória/experiência para explicar essa em tempo real. O resultado dessa junção cria uma nova história, porque a combinação de traços das outras duas é inovadora (FAUCONNIER; TURNER, 2002; TURNER, 2014).

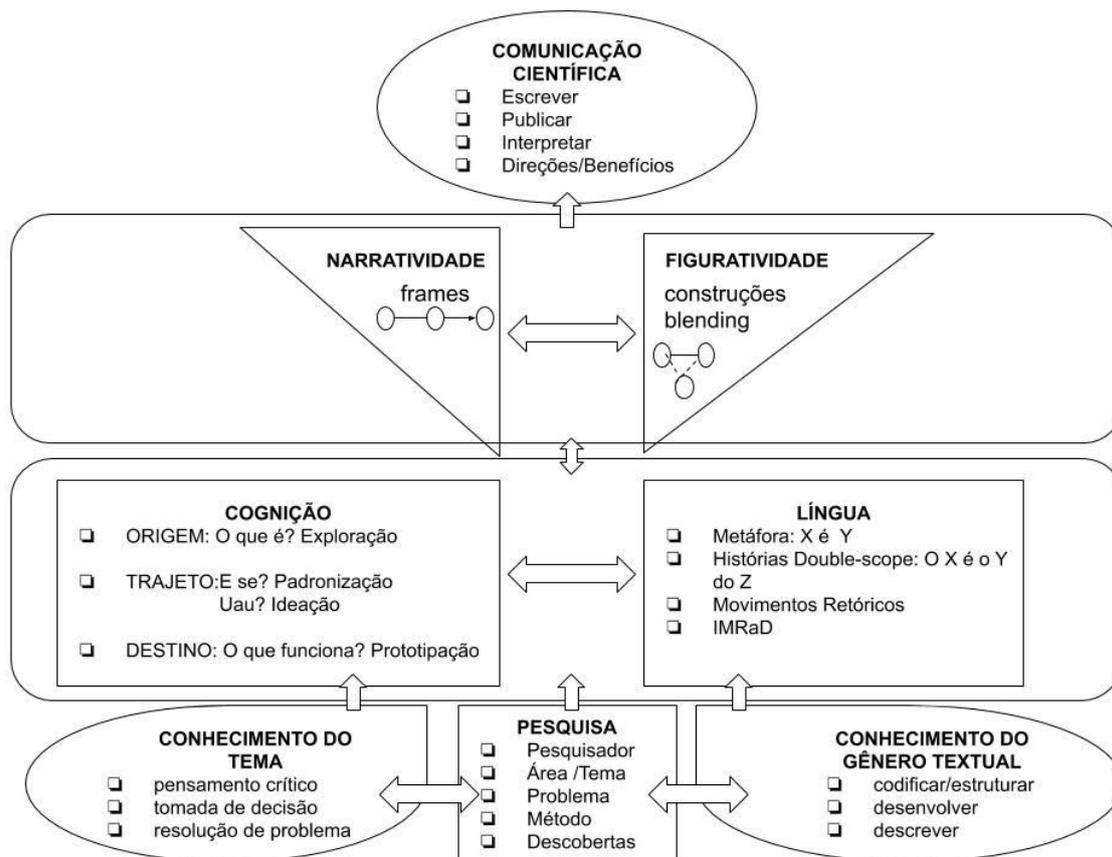
O processo cognitivo da integração conceptual acontece a partir de projeções entre dois domínios – um concreto/fonte e outro abstrato/alvo – e, por isso, é manifestado linguisticamente por meio de analogias e metáforas. Um dos mecanismos para gerenciar ideias complexas é integrar conceptualmente dois tipos de conhecimento: conhecimento do domínio que é o objeto da compreensão e conhecimento de outro domínio distinto usado para interpretar essa compreensão. Este é o cerne do processo metafórico: compreender uma ideia mais complexa/abstrata em termos de uma ideia mais estruturada/concreta. Essa é a definição de metáfora conceptual segundo Lakoff e Johnson (1980). A metáfora é uma analogia condensada, a partir da qual se busca ancorar uma informação nova em uma informação já conhecida para construir significado e criar conhecimento.

Nessa perspectiva linguístico-cognitivista, em que narratividade e figuratividade invocam e evocam o pensamento e a expressão humana, a escrita não poderia ser concebida como produto, mas como processo, cujos estágios de pré, durante e pós-processamento perfazem o ciclo da pesquisa científica e sua comunicação em trabalhos acadêmicos.

2.4 Escrita como processo de aprendizagem

Na comunicação científica, a escrita não é um produto, mas um processo de aprendizagem por meio do qual os cientistas desenvolvem e comunicam conceitos e processos em sua narrativa de pesquisa (RODRIGUES; BAPTISTA, 2019). A comunicação científica escrita, de acordo com

Figura 2 - Modelo para a comunicação científica.



Fonte: Rodrigues (2019, p. 69).

Flower e Hayes (1981), requer do cientista, além de conhecimento profundo do tema de estudo, conhecimento profundo do gênero textual acadêmico.

Nessa concepção, ao ensinar redação científica, propomos no estágio de pré-escrita a utilização dos constructos da Linguística Cognitiva – a narrativa cognitiva, os esquemas imagéticos e os frames para o delineamento das ideias de maneira lógica; e as construções linguísticas que manifestam integrações conceptuais, como as analogias e as metáforas, para a comunicação de conceitos e processos científicos de maneira mais criativa. O intuito é auxiliar o cientista a projetar de maneira mais eficaz o seu texto, de modo que reflita coerentemente seu fazer científico.

Para auxiliar os formadores de cientistas, propomos um modelo pedagógico compatível com essa concepção para que o ensino da escrita científica seja realizado a partir do conhecimento do tema de estudo (cognição/pesquisa) e do gênero textual

acadêmico (língua/comunicação).

2.5 Modelo pedagógico para a comunicação científica

Ao modelar a narratividade e a figuratividade do texto científico para a comunicação científica (Figura 2), partimos dos processos de pesquisa (cognição) e comunicação (língua). Começamos pelo frame da Pesquisa: há um pesquisador dentro de uma área do conhecimento e uma investigação sobre um determinado tema que parte de uma questão. Existe a opção de um método que leve a descobertas. No processo, dois tipos de conhecimento são necessários: conhecimento do tema de pesquisa e conhecimento de gênero textual científico. Cognição e língua entram em jogo para o desenvolvimento do pensamento crítico, tomada de decisão e resolução de problemas, e a escrita começa como processo de aprendizagem nas ações de codificar, estruturar, desenvolver e

descrever estágios e conceitos científicos para potencialmente comunicá-los quando da publicação (RODRIGUES, 2019). Ancoramos esse modelo em ferramentas de visualização do Design Thinking⁹ (LIEDTKA; OGILVIE, 2011).

Uma vez proposto o modelo pedagógico para a comunicação científica, surgiu então a necessidade metodológica de testar todas essas hipóteses contra um *corpus* sistemático maciço de dados ecologicamente válidos, juntamente com ferramentas e práticas para analisar esses dados, que consistem de uma abordagem denominada Data Science¹⁰. Essa abordagem passa pela análise linguística especializada e é confirmada a partir do uso de ferramentas tecnológicas.

A seguir, apresentamos uma amostra dessa análise de dados em que ficam evidentes a narrativa cognitiva, os esquemas imagéticos, os frames, as analogias e as metáforas na comunicação científica. Os dados analisados foram extraídos de um *corpus* de *abstracts* científicos e sua coleta tratada a partir de ferramentas linguísticas tecnológicas.

3. ANÁLISE DE TEXTOS CIENTÍFICOS E O USO DE FERRAMENTAS TECNOLÓGICAS

Esta análise de dados aborda os aspectos cognitivos e linguísticos na pesquisa e na sua comunicação a partir das lentes da Linguística Cognitiva. A justificativa para a escolha desses dados especificamente se baseia no fato de que cientistas que atuam em ambientes interdisciplinares e fazem pesquisa na interface com outras áreas do conhecimento, ou até mesmo entre disciplinas da mesma área, precisam entender conceitos e processos dessas outras áreas/disciplinas para escrever e publicar; na biotecnologia,

9 Para uma descrição completa do modelo, ver Rodrigues (2019).

10 *Data science* é o estudo e a análise de dados, que visa a extração de conhecimento para possíveis tomadas de decisão. Alia *big data* e *machine learning*, além de estatística, economia, engenharia e outros subcampos da computação como: banco de dados e análise de agrupamentos (*cluster analysis*). Disponível em: https://pt.wikipedia.org/wiki/Ci%C3%AAncia_de_dados. Acesso em: 3 abr. 2020.

por exemplo. A amostra de dados a seguir apresenta trechos de um artigo científico da medicina que conversa com as áreas da engenharia e da computação.

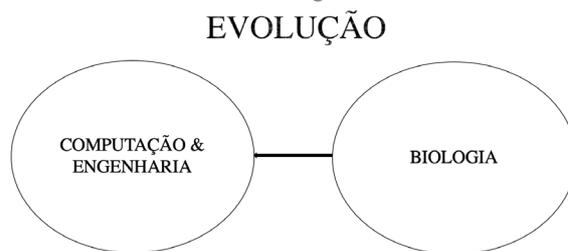
Utilizamos, nas análises, diagramas que representam visualmente as projeções entre dois ou mais domínios. Para representar a metáfora conceptual, duas esferas são utilizadas: a esfera da esquerda representa o domínio-fonte, de onde se buscam elementos concretos para o entendimento da nova informação, abstrata, representada pela esfera da direita, o domínio-alvo (LAKOFF, JOHNSON, 1980).

Para representar processos de integração conceptual, o diagrama tem quatro esferas: a de cima é o espaço genérico, no qual se integram todos elementos dos dois frames (os domínios fonte e alvo), aparentemente incompatíveis, de onde são extraídos os elementos habilitados à construção do novo significado. A esfera debaixo representa o espaço *blending*, onde se encontra o resultado do processo de integração conceptual (FAUCONNIER; TURNER, 2002).

3.1 Análise de dados

Esta análise apresenta como dados três construções linguísticas de um artigo científico da medicina (CALCOTT, 2014), que revela a prevalência da computação e da engenharia integradas conceptualmente com a biologia (Figura 3). Mais especificamente, as construções são sobre genética e revelam a tentativa dos cientistas de mapear o conceito “evolução” em termos de processos mecânicos e tecnológicos.

Figura 3 - Integração conceptual computação & engenharia & biologia.



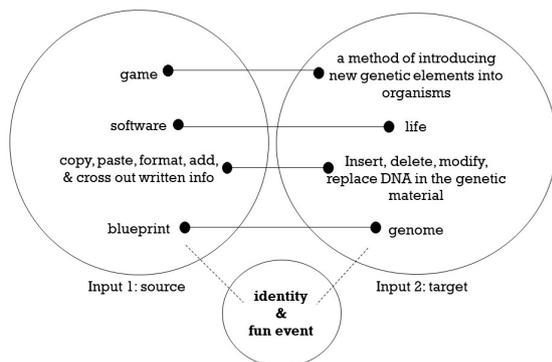
Fonte: Elaboração própria.

O primeiro dado “Genome editing is playing with the software of life” (Editar o genoma é brincar com o software da vida) (CALCOTT, 2014) é uma metáfora – uma construção linguística do tipo [X é Y] em que [X = editar o genoma] e [Y = brincar com o software da vida]. A metáfora conceptual é PESSOAS SÃO MÁQUINAS¹¹.

O genoma é um documento que identifica um organismo vivo e pode ser lido, (re)escrito e até desenhado no processo da evolução. O ciclo de vida é documentado em um projeto em que a vida é como um software. O sistema biológico é um jogo em que novos elementos entram em cena sob o risco de prejudicar os genes, mesmo sabendo-se que, em se ganhando o jogo, o prêmio pode ser a cura de certas doenças genéticas. Valores como identidade e diversão são projetados nessa integração conceptual (Figura 4).

Os esquemas imagéticos são ORIGEM-PERCURSO-DESTINO, uma vez que a edição do genoma é um processo cujo destino é a inserção de material genético em locais desejados específicos; e DINÂMICA DE FORÇAS, porque as “tesouras moleculares” são usadas para causar mudança.

Figura 4 - Integração conceptual de “Editar o genoma é brincar com o software da vida”.



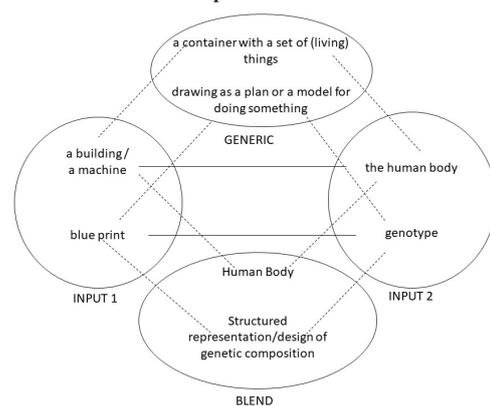
Fonte: Elaboração própria.

O segundo dado “The genotype is the blueprint of the human body” (O genótipo é a planta do corpo humano) (CALCOTT, 2014) é uma analogia – uma construção linguística do tipo [X é Y de Z] em que [X = o genótipo], [Y = a planta] e [Z = o corpo humano]. Há um

elemento inferido: um prédio ou uma máquina [W], que está para o corpo humano [W : Z] assim como a planta está para o genótipo [Y : X]. A metáfora conceptual que estrutura essa analogia e está condensada nela é também PESSOAS SÃO MÁQUINAS. O genótipo do corpo humano é explicado em termos da planta de um prédio ou máquina [W]. Inferências são abstraídas do valor do genótipo a partir do elemento implícito [W]. O corpo humano é uma representação estruturada de sua composição genética (Figura 5).

Em termos de esquemas imagéticos, o corpo humano, assim como um prédio ou uma máquina, é um CONTÊINER de coisas vivas cuja estrutura é representada por uma planta. Além disso é um sistema que deve funcionar, logo funcionalidade é VERTICALIDADE.

Figura 5 - Integração conceptual de “O genótipo é a planta do corpo humano”.



Fonte: Elaboração própria.

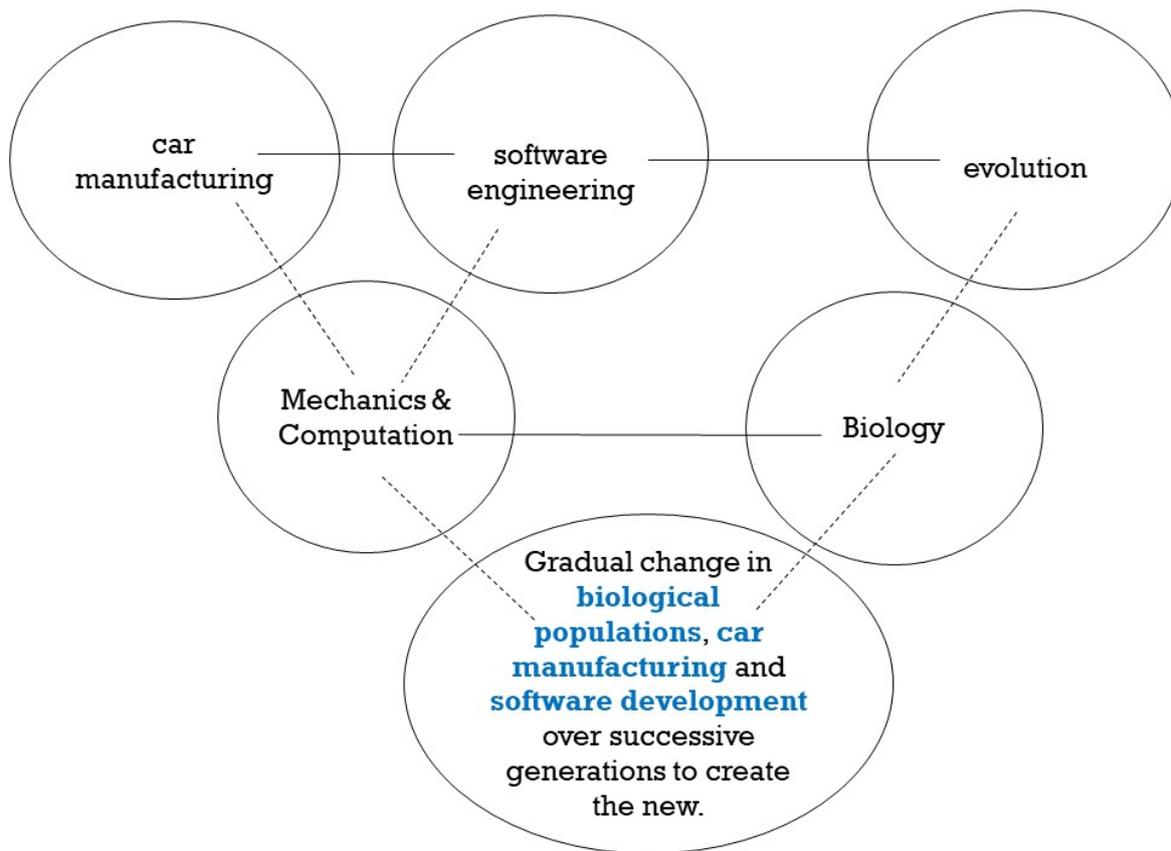
O terceiro dado é também uma analogia, descrita em termos de uma pequena narrativa:

Modificar a fabricação de carros é algo como evoluir. A mudança evolutiva começa copiando um conjunto de recursos de desenvolvimento - em grande parte, mas não apenas, genes. Esse processo de cópia introduz modificações, e esses recursos modificados se transformam em um novo organismo. Além disso, a engenharia de software, assim como a fabricação e evolução de carros, é um processo de copiar, modificar e gerar (CALCOTT, 2014, tradução nossa)¹².

12 Modifying the manufacturing of cars is a closer fit to evolution. Evolutionary change begins by copying a set of developmental resources – largely, but not only, genes. This copying process introduces modifications, and these modified resources then develop into a new organism. Also software engineering, like car manufacturing and evolution, is a copy-modify-and-generate process (CALCOTT, 2014).

11 Há uma base de dados com um inventário de metáforas, de A a Z, e suas respectivas descrições em <https://metaphor.icsi.berkeley.edu/pub/en/index.php/Category:Metaphor>.

Figura 6 - Integração conceitual de “Modificar a fabricação de carros é algo como evoluir”.



Fonte: Elaboração própria.

A metáfora é ALCANÇAR UM OBJETIVO É ADQUIRIR UM OBJETO DESEJADO, interligadas às metáforas EXPERIENCIAR UM ESTADO É POSSUIR UM OBJETO e MUDAR É AQUISIÇÃO.

O processo da evolução é mapeado em termos do processo de fabricação de carros e a engenharia de software integra a mecânica e a computação à biologia. Os frames estruturadores dessa integração são de universos distintos, quase que aparentemente incompatíveis: cultura e tecnologia versus seleção natural e genética geram um mesmo roteiro narrativo com objetos diferentes em uma estrutura emergente, que é a da biodiversidade como um novo produto da natureza (Figura 6).

Nesse roteiro narrativo da fabricação de carros como um processo de evolução, está subjacente o esquema imagético ORIGEM (copiar), PERCURSO (modificar) e DESTINO (gerar). Nesse processo evolutivo da fabricação de carros, bem como na

evolução da espécie humana, há manutenção, mas avanço para um novo produto: uma nova marca de carro, isto é, a biodiversidade. Isso diz respeito a modificar a fabricação de carros em vez de modificar o carro diretamente. A mecânica modifica-no-lugar. A evolução, diferentemente da mecânica, envolve o processo copiar-modificar-gerar. Aqui, há uma compressão de tempo no processo.

Os três dados analisados elucidam o fato de que as narrativas, os esquemas imagéticos, as integrações conceituais entre frames via analogias e metáforas fornecem âncoras para a construção do significado, que gera a nova informação e cria o conhecimento. Os conceitos são evocados por construções linguísticas e se integram a partir delas no texto durante o processo de comunicação.

Esses dados foram filtrados do *corpus* de mil *abstracts*, compilado eletronicamente, anotado por frames automaticamente e tratado estatisticamente para a extração de padrões. Dentre os padrões visualizados,

selecionamos construções linguísticas e as analisamos lexical e sintaticamente também com o uso de uma ferramenta tecnológica. A seguir, apresentamos um breve inventário das principais ferramentas utilizadas nesse processo, bem como o percurso de uso, que nos ajudaram a otimizar a análise linguística.

3.2 O uso de ferramentas tecnológicas na análise linguística

A análise linguística foi realizada a partir da abordagem Data Science, desenvolvida no Red Hen Lab. Data Science, ou Ciência de Dados, consiste no uso de métodos científicos para obter informações úteis a partir de dados de computador, especialmente quantidades massivas de dados¹³.

Primeiramente, coletamos e compilamos eletronicamente um *corpus* de mil *abstracts* científicos, de periódicos de alto impacto, de 10 disciplinas das três grandes áreas do conhecimento na Web of Science¹⁴, entre 2015 e 2016. O *corpus* foi tratado para a modelagem computacional, conforme descrito no diagrama da Basic Text Pipeline¹⁵. Compilar eletronicamente consiste em armazenar os textos em arquivos em formato .txt num ambiente online, onde possam ser processados por softwares.

Em seguida, utilizamos o parser semântico Semafor¹⁶, um software que anota frames automaticamente com base na FrameNet, que é um dicionário online de frames. O próximo passo foi anotar o *corpus* morfológicamente com o uso da ferramenta Stanford Core NLP¹⁷, outro software que etiqueta automaticamente todas as palavras dos textos de acordo com sua classe gramatical (substantivo, verbo, adjetivo etc).

Finalmente, o *corpus* foi inserido

no CQPweb¹⁸, um processador online de consultas de corpus que extrai construções linguísticas a partir de buscas escritas por meio de parâmetros gramaticais do tipo X é Y, para metáforas; e X é Y de Z, para analogias, variáveis que são preenchidas conforme a notação do Penn Treebank Project¹⁹, um dicionário de etiquetas morfológicas para as palavras. Por exemplo, para buscar uma construção linguística como “o menino é o pai do homem” (analogia machadiana), abstraímos a estrutura “o X é o Y do Z”, parametrizada gramaticalmente como “_DT _NN {ser} _DT _NN _IN _NN”, em que _DT é o determinante “o”, _NN é o substantivo (menino, pai, homem) e _IN é a preposição (do). O verbo ser vem entre chaves para que o sistema retorne todas as variações dessa forma infinitiva (é, são, foi, serão etc). A construção linguística em língua natural é traduzida como uma linha de código de programação para que a máquina retorne o dado a partir do comando buscado. Os três dados que compuseram a análise aqui apresentada foram uma amostragem do resultado que obtivemos a partir de buscas feitas por meio desse processo.

4. CONSIDERAÇÕES FINAIS: TEORIA, MODELO E FERRAMENTA

Qual é o significado desses resultados para o estudo da Comunicação Científica e para o ensino de redação científica? Teoricamente, existem implicações para a pesquisa em Linguística Cognitiva, na medida em que é a teoria da integração conceptual, instanciada por narrativas, analogias e metáforas que deixam explícitas na língua o que está saliente na cognição dos cientistas quando eles desenvolvem e comunicam conceitos e processos em sua área de estudo. Por exemplo, a amostra de análise que aqui apresentamos revela que

13 <https://dictionary.cambridge.org/pt/dicionario/ingles/data-science>

14 <http://www.webofknowledge.com>

15 <http://www.redhenlab.org/home/the-cognitive-core-research-topics-in-red-hen/the-barnyard/basic-text-pipeline>

16 <http://www.cs.cmu.edu/~ark/SEMAFOR/>

17 <https://stanfordnlp.github.io/CoreNLP/>

18 <https://cqpweb.lancs.ac.uk>

19 https://www.ling.upenn.edu/courses/Fall_2003/ling001/penn_treebank_pos.html

os cientistas da medicina, nos estudos sobre genética, veem os sistemas biológicos como confiáveis, previsíveis e controláveis, ao compará-los aos processos de engenharia mecânica e engenharia de software. Tais resultados de análise linguística podem ser aplicados didaticamente no ensino de redação científica. Tecnologicamente, orientam o uso e desenvolvimento de ferramentas tecnológicas para a análise linguística.

Em suma, a contribuição do estudo aqui descrito pode ser disposta em termos de teoria, modelo e ferramenta, um tripé que pode ser entendido em termos de território, mapa e veículo. O mapa é a representação do território em que se transita por meio de um veículo. Sendo assim, a contribuição teórica para a pesquisa científica está no território da Linguística Cognitiva, mapeada para a comunicação científica, cuja contribuição é o modelo pedagógico para o ensino de redação científica. As ferramentas de Data Science são o veículo para a análise linguística e completam o ciclo como contribuição tecnológica.

Quais são as perspectivas futuras? Proceder a um tratamento estatístico²⁰ das listas de frequências de frames e construções linguísticas no *corpus* de *abstracts* nos permitirão ver os padrões narrativos e figurativos de artigos de alto impacto publicados em várias áreas do conhecimento. Estes servirão para validar o modelo pedagógico para o ensino de redação científica e poderão prover uma taxonomia para a publicação. O estudo está em andamento e as contribuições concretas, a partir de resultados mais robustos, estão a caminho.

O ponto a que chegamos nessa trajetória de investigação da Comunicação Científica é de transformação de um projeto de extensão, que passou por uma pesquisa internacional de pós-doutorado, e ganhou *status* de programa de pesquisa sustentável na interface Linguística/Educação/Tecnologia.

REFERÊNCIAS

²⁰ https://github.com/BrazilianRedHen/prose_hen/tree/master/statistics/export

CALCOTT, B. *Evolvability and Engineering. Biology and Philosophy*. 2014. 29(3), 293-313.

EVANS, V. *A glossary of cognitive linguistics*. Edinburgh University Press, 2007.

FAUCONNIER, G.; TURNER, M. *The way we think: conceptual blending and the mind's hidden complexities*. New York: Basic, 2002.

FILLMORE, C. Frame Semantics. In: *Linguistics in the morning calm*. Seoul: Hanshin, 1982. p.111-138.

FITZPATRICK, Y. *Science communication: a visual perspective*. Alan Alda Center: Stony Brook University, New York, 2019. Disponível em: <https://www.aldacenter.org/aklc/webinars/visual-perspective>. Acesso em: 2 jun. 2020.

FLOWER, L.; HAYES, J. R. A cognitive process theory of writing. *College composition and communication*, 32, 1981. 365-387. DOI: <http://dx.doi.org/10.2307/356600>.

HERMAN, D. *Narrative theory and the cognitive sciences* (ed.). Standford, California: The University of Chicago Press Books, 2003.

JOHNSON, M. *The body in the mind: the bodily basis of meaning, imagination, and reason*, Chicago: University of Chicago Press, 1987.

LAKOFF, G.; JOHNSON, M. *Metaphors we live by*. Chicago: The University of Chicago Press, 1980.

LIEDTKA, J.; OGILVIE, T. *Designing for growth: a Design Thinking tool kit for managers*. New York: Columbia University Press, 2011.

RODRIGUES, R. F. L. A ciência é uma jornada: um projeto remodelado como programa de pesquisa linguística em Comunicação Científica com uso de Data Science. *Sinergia* (IFSP), v. 20, Edição Especial - Comunicação Científica, Cognition e Persuasão, 2019. p. 56-74. Disponível em: <https://ojs.ifsp.edu.br/index.php/sinergia/issue/view/94>. Acesso em: 2 abr. 2020.

RODRIGUES, R. F. L.; BAPTISTA, A. E. O. B. Design thinking tools for scientific storytelling: a didactic innovation. *Proceedings of the 13th*

annual international technology, education and development conference, INTED 2019. 11th-13th March, 2019. Valencia, Spain. Disponível em: https://iited.org/concreteview_abstract.php?paper_id=72029. Acesso em: 2 jun. 2020.

SWALES, J. **Research genres: exploration and applications**. Cambridge University Press, Cambridge, 2004.

TURNER, M. **The origin of ideas: blending, creativity and the human spark**. Oxford: Oxford University Press, 2014.

TURNER, M. The Mind is an Autocatalytic Vortex. In: **The literary mind**, Volume 24 of REAL: Yearbook of Research in English and American Literature, edited by Jürgen Schlaeger. Tübingen, Germany: Gunter Narr Verlag, 2008.

TURNER, M. Double-Scope Stories. In: **Narrative theory and the cognitive sciences**. David Herman (ed.), The University of Chicago Press Books, 2003.