

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Faculdade de Letras
Programa de Pós-Graduação em Estudos Linguísticos

Laís Marçal de Oliveira

EXPRESSÕES IDIOMÁTICAS NÃO SÃO PROCESSADAS COMO METÁFORAS

Belo Horizonte
2024

Laís Marçal de Oliveira

EXPRESSÕES IDIOMÁTICAS NÃO SÃO PROCESSADAS COMO METÁFORAS

Dissertação apresentada ao Programa de Pós-Graduação em Estudos Linguísticos da Faculdade de Letras da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Mestre em Estudos Linguísticos.

Área de Pesquisa: Linguística Teórica e Descritiva

Linha de Pesquisa: Processamento e Emergência de Construções Linguísticas

Orientadora: Profa. Dra. Larissa Santos Ciríaco

Belo Horizonte
2024

O48e

Oliveira, Laís Marçal.

Expressões idiomáticas não são processadas como metáforas
[manuscrito] / Laís Marçal Oliveira. – 2024.

1 recurso online (94 f.: il., tabs., grafs., color., p&b.) : pdf.

Orientadora: Larissa dos Santos Ciríaco.

Área de concentração: Linguística Teórica e Descritiva.

Linha de pesquisa: Processamento e Emergência de Construções
Linguísticas.

Dissertação (mestrado) – Universidade Federal de Minas
Gerais, Faculdade de Letras.

Bibliografia: f. 83-85.

Apêndices: f. 86-94.

Exigências do sistema: Adobe Acrobat Reader.

1. Língua portuguesa – Estudo e ensino – Teses. 2. Metáfora –
Teses. 3. Língua portuguesa – Expressões idiomáticas – Teses. 4.
Linguística – Teses. I. Ciríaco, Larissa Santos. II. Universidade
Federal de Minas Gerais. Faculdade de Letras. III. Título.

CDD: 469.5



UNIVERSIDADE FEDERAL DE MINAS GERAIS
FACULDADE DE LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTUDOS LINGUÍSTICOS

FOLHA DE APROVAÇÃO

EXPRESSÕES IDIOMÁTICAS NÃO SÃO PROCESSADAS COMO METÁFORAS

LAÍS MARÇAL DE OLIVEIRA

Dissertação submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em ESTUDOS LINGUÍSTICOS, como requisito para obtenção do grau de Mestre em ESTUDOS LINGUÍSTICOS, área de concentração LINGUÍSTICA TEÓRICA E DESCRITIVA, linha de pesquisa Processamento da Linguagem.

Aprovada em 03 de abril de 2024, pela banca constituída pelos membros:

Prof(a). Larissa Santos Ciriaco - Orientadora

UFMG

Prof(a). Lúcia Monteiro de Barros Fulgêncio

UFMG

Prof(a). Ricardo Augusto de Souza

UFMG

Belo Horizonte, 03 de abril de 2024.



Documento assinado eletronicamente por **Larissa Santos Ciriaco, Membro**, em 09/04/2024, às 08:20, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Lucia Monteiro de Barros Fulgencio, Professor(a)**, em 12/04/2024, às 14:25, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Ricardo Augusto de Souza, Professor do Magistério Superior**, em 23/05/2024, às 21:43, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador 3084557 e o código CRC 44514D26.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

À minha voinha Amazilia (*in memoriam*), grande incentivadora, professora e entusiasta.

AGRADECIMENTOS

Esta dissertação só existe graças à parceria de tantos que passaram pelo meu caminho. Agradeço, primeiramente, à Professora Dra. Larissa Ciríaco pela orientação, desde os tempos da graduação até o mestrado. Muito obrigada pelas contribuições em minha trajetória.

Aos membros da banca Prof.^a Dra. Lúcia Fulgêncio e Prof. Dr. Ricardo de Souza, muito obrigada pela leitura atenta, pelas enriquecedoras contribuições à minha pesquisa e por tornarem o momento da defesa tão leve e ao mesmo tempo tão rico.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) que, por meio da bolsa, permitiu que eu pudesse me dedicar a minha pesquisa e formação. Um viva à Universidade pública e a todos os programas de fomento à pesquisa, à extensão e ao ensino.

Agradeço aos meus professores que tanto me inspiraram e me ensinaram durante todo o meu percurso no campo das Letras e da Linguística. Agradeço também a todos os trabalhadores que, com muita garra, constroem a Faculdade de Letras da UFMG e o POSLIN.

Aos meus colegas de mestrado, meu muito obrigado! Caminhar se tornou mais fácil com a presença, o companheirismo, a alegria e o incentivo de vocês. Agradeço, em especial, aos meus colegas do Laboratório de Psicolinguística por todo apoio: Thaís Sá, Yasmin, Isabella e tantos outros.

Ao querido Gabriel Magno e à Yasmin Guimarães (novamente!) que tanto colaboraram na etapa de programação do teste de leitura autocadenciada. A ajuda de vocês foi fundamental para esta dissertação.

Agradeço também às pessoas do meu coração. À minha avó Amazilia por todo amor, incentivo e alegria que me transmitiu e sempre transmitirá. À minha amada mãe por sempre acreditar em mim e não medir esforços para que eu chegasse onde eu desejasse estar. Aos meus irmãos, meu pai, meus avós, tios e primos por toda força e torcida. Ao meu marido João que se mostrou um verdadeiro companheiro nos momentos em que eu realmente precisava, mostrando-se um grande incentivador dos meus objetivos. A todos vocês meu amor e gratidão.

Por fim, aos voluntários que participaram deste estudo agradeço pela prestatividade e colaboração.

RESUMO

Esta dissertação apresenta um estudo sobre as construções idiomáticas (CI) do PB, contrastando-as com construções metafóricas (CM). As CI's são expressões em que o significado global não se dá a partir do significado de cada item que as compõem. São construções convencionais, amplamente utilizadas pelos usuários da língua. Os estudos linguísticos, em uma perspectiva tradicional, consideram as CI's estruturas periféricas da língua que, por sua natureza idiossincrática, instauram problemas teóricos. A tradição também costuma compreender as CI's como construções metafóricas (Riva, 2009; Xatara, 1995, 1998; e outros). No entanto, seguindo o arrazoado exposto em Fulgêncio e Ciríaco (2018), a hipótese experimental levantada por esta pesquisa é de que as construções idiomáticas não são processadas como metáforas, pois as CM's, especificamente as construções metafóricas novas (CMN), exigem uma postura ativa do usuário para a montagem de significado no momento da interação. Dessa forma, por meio de um teste de leitura autocadenciada, pretende-se responder ao questionamento: as CI's possuem natureza metafórica? Serão analisados os tempos de leitura (TR), nos trechos-alvo e subsequentes, de sentenças experimentais em três níveis de condição: construções idiomáticas (CI), construções metafóricas convencionais (CMC) e construções metafóricas novas (CMN). O objetivo geral deste estudo é, portanto, contribuir para a descrição teórica das construções idiomáticas no PB e para verificação experimental do estatuto processual destas construções. Já os objetivos específicos são: (i) descrever as CI's com base no quadro teórico da linguística baseada no uso; (ii) analisar os resultados do teste de leitura autocadenciada à luz da abordagem supracitada e das teorias da metáfora que a compreendem como um processo cognitivo. A partir da análise dos dados e das discussões, foi possível encontrar diferenças de tempo de reação que apontam para um maior custo cognitivo das construções metafóricas novas em relação às construções idiomáticas no trecho subsequente. Além disso, foram encontradas diferenças de processamento que apontam para o menor custo cognitivo das construções metafóricas convencionais em comparação às construções idiomáticas. O estudo também aponta para as diferenças entre os dois tipos de metáforas analisados: CMC e CMN, sendo o maior tempo de reação para as metáforas novas. Além disso, foi possível descrever as CI's em uma abordagem que comporta seu comportamento irregular, entendendo-as como construções que pareiam uma forma fortemente associada a um sentido idiomático.

Palavras-chave: expressões idiomáticas, metáfora, processamento da linguagem, linguística baseada no uso.

ABSTRACT

This dissertation presents a study on idiomatic constructions (IC) in Brazilian Portuguese (PB), contrasting them with metaphorical constructions (MC). IC's are expressions in which the meaning does not derive from the meaning of each individual item that composes them. They are conventional constructions widely used by language users. Linguistic studies, from a traditional perspective, consider IC's as peripheral structures of the language that, due to their idiosyncratic nature, pose theoretical problems. The tradition also tends to comprehend IC's as metaphorical constructions (Riva, 2009; Xatara, 1995, 1998, and others). However, following the reasoning outlined in Fulgêncio and Ciríaco (2018), the experimental hypothesis raised by this research is that idiomatic constructions are not processed as metaphorical constructions, specifically new metaphorical constructions (NMC), which demand an active engagement from users to construct meaning during interaction. Thus, through a self-paced reading test, the aim is to answer the question: Do IC's have a metaphorical nature? Reading times (RT) in target and subsequent passages of experimental sentences at three condition levels will be analyzed: idiomatic constructions (IC), conventional metaphorical constructions (CMC), and new metaphorical constructions (NMC). The overall objective of this study is, therefore, to contribute to the theoretical description of idiomatic constructions in PB and to experimentally verify the processing status of these constructions. The specific objectives are: (i) to describe ICs based on the theoretical framework of usage-based linguistics; (ii) to analyze the results of the self-paced reading test in light of the aforementioned approach and the theories of metaphor that understand it as a cognitive process. From the analysis of the data and discussions, differences in reading time were found indicating a higher cognitive cost of new metaphorical constructions compared to idiomatic constructions in the subsequent passage. Additionally, processing differences were found indicating a lower cognitive cost of conventional metaphorical constructions compared to idiomatic constructions. The study also points out differences between the two types of analyzed metaphors: CMC and NMC, with longer reading times for new metaphors. Furthermore, it was possible to describe ICs in an approach that accommodates their irregular behavior, understanding them as constructions that pair a form strongly associated with an idiomatic meaning.

Keywords: idiomatic expressions, metaphor, language processing, usage-based linguistics.

LISTA DE FIGURAS

Figura 1 - Trecho com contexto para sentença experimental.	58
Figura 2- Ponto de fixação antes do início da leitura	58
Figura 3 - Lacunas a serem preenchidas pela sentença.	58
Figura 4 - Sequência da leitura autocadenciada.	59
Figura 5 - Pergunta de compreensão	59

LISTA DE TABELAS

Tabela 1 - Construções variando em nível de complexidade e abstração	34
Tabela 2 - Resultados do modelo linear de efeitos mistos no trecho-alvo Erro! Indicador não definido.	
Tabela 3 - Resultados do modelo linear de efeitos mistos no trecho subsequente	66

LISTA DE GRÁFICOS

Gráfico 1 - Média de tempo de reação no trecho-alvo: CI e CM	Erro! Indicador não definido.
Gráfico 2 - Média de tempo de reação no trecho spillover: CI e CM	Erro! Indicador não definido.
Gráfico 3 - Média do tempo de reação no trecho-alvo: CI, CMC e CMN	63
Gráfico 4 - Média do tempo de reação no trecho spillover: CI, CMC e CMN	63
Gráfico 5 - Efeitos de correlação entre as construções no trecho-alvo.....	65
Gráfico 6 - Efeitos de correlação entre as construções no trecho subsequente	66
Gráfico 7 - Análise <i>post-hoc</i> : diferença entre as médias no trecho-alvo.....	67
Gráfico 8 - Análise <i>post-hoc</i> : diferença entre as médias no trecho subsequente	68

Lista de Abreviaturas e Siglas

CEA – Construção de Estrutura Argumental

CENCE ME - Construção, Eficiência, Novas informações, Competição, Expressividade, Memória e aprendizagem impulsionada por Erros.

CI – Construção Idiomática

CM – Construção Metafórica.

CMC – Construção Metafórica Convencional.

CMN – Construção Metafórica Nova

CPF – Cadastro de Pessoa Física

EF – Expressão Fixa

EI – Expressão Idiomática

EM – Expressão Metafórica

EMC – Expressão Metafórica Convencional.

EMN – Expressão Metafórica Nova

GC – Gramática de Construções

LA – Leitura autocadenciada

LBU – Linguística Baseada no Uso

LC – Linguística Cognitiva

PB – Português brasileiro

POSLIN – Programa de Pós-Graduação em Estudos Linguísticos.

SVO – Sujeito Verbo Objeto

TCM – Teoria da Carreira da Metáfora

TR – Tempo de reação

TMC – Teoria da Metáfora Conceptual

UFMG – Universidade Federal de Minas Gerais.

SUMÁRIO

Capítulo 1: Objeto de estudo	10
1.1 Expressões fixas.....	10
1.2 As expressões idiomáticas	12
1.2.1 Idiossincrasia.....	13
1.2.2 Convencionalidade	17
1.2.3 Memorização em bloco	18
1.2.4 Questões levantadas pelas EI's	18
1.3 Justificativa	21
1.4 Objetivos	22
1.4.1 Objetivo geral.....	22
1.4.2 Objetivos específicos.....	22
1.5 Hipóteses.....	22
1.6 Organização da dissertação.....	23
Capítulo 2: Referencial teórico	24
2.1 Linguística Baseada no Uso.....	24
2.1.1 <i>Embodiment</i> : linguagem e cognição corporificada	25
2.1.2 O caráter emergentista e a função comunicativa da linguagem na LBU	26
2.1.3 O papel do <i>input</i> e da frequência.....	27
2.1.4 Composicionalidade e significado	28
2.2 A Gramática de Construções (Goldberg, 1995, 2006, 2019).....	32
2.2.1 A aprendizagem das construções: <i>input</i> , frequência e memória	36
2.2.2 Convencionalidade, entrincheiramento e <i>chunking</i>	39
2.3 A metáfora como processo cognitivo	42
2.3.1 Teoria da Metáfora Conceptual (TMC)	42
2.3.2 Metáforas novas	46
2.4 Processamento linguístico e acesso lexical.....	51
Capítulo 3: Metodologia	55
3.1 Paradigma experimental: leitura autocandenciada.....	55
3.2 Desenho experimental e critério de exclusão.....	56
3.2.1 Procedimentos	57
3.2.2 Materiais.....	59
Capítulo 4: O processamento de expressões idiomáticas	62

4.1 Preparação dos dados.....	62
4.2 Resultados.....	62
4.3 Discussão.....	68
Capítulo 5: Conclusão.....	75
5.1 Limitações do estudo.....	76
5.2 Pesquisas futuras.....	76
REFERÊNCIAS.....	78
APÊNDICES.....	81

Capítulo 1: Objeto de estudo

Este trabalho trata do processamento de construções idiomáticas em contraste com o processamento de construções metafóricas em português brasileiro (PB). As construções idiomáticas são um tipo de expressão fixa (EF). Portanto, é importante apresentar algumas considerações sobre as EF's antes de abordar as especificidades das expressões idiomáticas (EI).

1.1 Expressões fixas

As EF's são construções caracterizadas por serem uma sequência de palavras que representa uma unidade convencionalizada, sendo recuperada, em hipótese, como um todo da memória do usuário, e não obrigatoriamente estando sujeita às regras gerais de formação de sentenças nas línguas (Fulgêncio, 2008). Por exemplo, em uma EF como “só faltava essa”, o pronome “essa” é sempre marcado no feminino, independente do contexto anterior. Ainda, há expressões compostas por verbos que não permitem variação em tempo e pessoa, como “quem me dera” (“? quem me deu”). Além disso, existem EF's compostas por palavras que não aparecem de maneira isolada, a exemplo de “no entanto”. As EF's também variam em grau de “fixidez”, podendo possuir partes fixas, bem como variáveis:

- (1) a. Tomar providências.
- b. Tomei providências.
- c. ? Fazer providências.
- (2) a. em tudo que é *X*.
- b. Procurei em tudo que é *canto*.
- c. Procurei em tudo que é *lugar*.
- d. Procurei em tudo que é *gaveta*.

No exemplo (1), a expressão fixa “tomar providência” é aberta ao acréscimo de morfema de 1ª pessoa do singular, mas não possui um espaço aberto (*slot*) para a troca do verbo “tomar” por “fazer”. Já no exemplo (2), há um *slot* aberto, representado por *X*, que pode ser intercambiado por nomes (2.a-c).

Outro ponto para o entendimento das EF's é o significado que pode ser tanto opaco quanto transparente.

- (3) Sofá-cama
- (4) Gato-pingado.

Em (3), por meio da soma dos itens “sofá + cama” é possível compreender o significado da EF, um móvel multiuso que serve como sofá e também pode virar cama. Já em (4), a soma de “gato + pingado” não transmite a ideia de uma pessoa que compõe um grupo com poucas pessoas que estão frequentando um determinado espaço, pois o significado da EF é opaco. Além das características já apresentadas, as EF’s costumam ser utilizadas de maneira convencionalizada. Os usuários utilizam a mesma forma e costumam rejeitar outras alternativas.

- (5) a. Achados e perdidos.
- b. ? Perdidos e achados.

Em (5), a EF “achados e perdidos” é uma forma convencional de designar um espaço em que se encontram objetos perdidos. Ainda que a EF subverta a ordem lógica (primeiro se perde e depois se acha), os usuários rejeitam outra opção como a (5.b), em que há a troca de posição dos itens lexicais da expressão. Ainda, é possível citar também expressões relacionadas a situações comunicativas específicas como a “até logo!”, usada para se despedir, mesmo que não se tenha a intenção ou previsão de que, em breve, outro encontro aconteça.

As EF’s compartilham das características gerais mencionadas, e alguns grupos possuem traços específicos. Com um intuito de se obter uma análise mais refinada do comportamento dessas expressões, elas são classificadas em subgrupos que determinam as tipologias das EF’s, conforme apontam Fulgêncio (2008) e Fulgêncio e Ciríaco (2018):

a) Colocações (*collocations*): combinações de itens lexicais de maneira co-ocorrente e convencionalizada, em que há relação de dependência léxico-semântica entre um item e outro. Um exemplo desse tipo de EF é “*doido varrido*”, na qual o item “varrido” se combina exclusivamente com o item “doido” (ou louco).

b) Coligações (*colligations*): combinações convencionalizadas de itens lexicais com base em padrões gramaticais. Um exemplo é o caso da regência do verbo “confiar”, que seleciona como objeto um sintagma preposicional com preposição “em”: “confiar em”. Essa preposição é opaca semanticamente, visto que, se comparamos com “desconfiar”, há um outro padrão, um objeto representado por sintagma preposicional com a preposição “de”: “desconfiar de”. Em outras palavras, trata-se de uma combinação convencionalizada que gera um padrão gramatical, sem motivação semântica para o uso, que é arbitrário.

c) Clichês: frases prontas, como “*Você está louco?*”, “*Tirar conclusões precipitadas*”, dentre outras.

d) Fórmulas situacionais: fórmulas fixas e convencionalizadas utilizadas em certos contextos comunicativos como: “*Feliz natal!*”, “*Quanto tempo!*”, ou “*Boa tarde!*”.

e) Binominais: expressões convencionalizadas formadas por itens lexicais da mesma categoria, unidos por preposição ou conjunção, nas quais não é possível trocar a ordem dos elementos, como: “*achados e perdidos*” (*perdidos e achados) e “*cara ou coroa*” (*coroa ou cara).

f) Provérbios: expressão popular convencionalizada não literal utilizada com o intuito moralizante, como “*Não atire pérolas aos porcos*”, “*Mais vale um pássaro na mão [do] que dois voando*”.

g) Expressões idiomáticas (EI’s): expressões convencionalizadas em que o significado é indecomponível, são aprendidas como uma sequência coesa e convencional, com significado idiossincrático e, em hipótese, armazenadas como *chunks*, blocos únicos de sentido. Nessa perspectiva, as EI’s seriam acessadas da memória tal qual itens lexicais, sem processamento interno (Fulgêncio e Ciríaco, 2018). Exemplos desse tipo de EF são: “*dar com os burros n’água*”, “*bater as botas*” e “*ter uma pulga atrás da orelha*”.

Em resumo, as EF’s são sequências de mais de uma palavra que formam um grupo convencional na língua – tanto do ponto de vista semântico, sintático ou discursivo, que possuem, ao menos, um elemento fixo. Pela convencionalidade dessas expressões, elas são, em hipótese, armazenadas em bloco pelos usuários de uma língua. Dentre as tipologias apresentadas, o interesse deste trabalho está nas expressões idiomáticas, a serem abordadas a seguir.

1.2 As expressões idiomáticas

Fillmore e colegas afirmam que “(...) uma expressão ou construção idiomática é algo que o falante de uma língua pode não saber mesmo conhecendo todas as outras propriedades dela.” (*apud* Fulgêncio e Ciríaco, 2018, p.362, **tradução nossa**)¹. Essa constatação evidencia o comportamento idiossincrático das EI’s, a convencionalidade e a importância do aprendizado tal qual um item lexical. Do ponto de vista da forma, as EI’s podem se apresentar de duas

¹No original: “(...) an idiomatic expression or construction is something a language user could fail to know while knowing everything else in language.” (Fillmore e colegas *apud* Fulgêncio e Ciríaco, 2018, p.362)

maneiras: construções idiomáticas esquemáticas/gramaticais e construções idiomáticas substantivas/lexicais (Fulgêncio e Ciríaco, 2018). Nas construções esquemáticas, as EI's possuem *slots* que podem ser preenchidos por itens lexicais que variam parcialmente, como em:

- (6) a. Quanto maior *X* maior *Y*.
 b. Quanto maior *a altura* maior *o tombo*.
 c. Quanto maior *a altura* maior *a queda*.
 d. Quanto maior *a subida* maior *a queda*
- (7) a. Dar uma de *X*.
 b. Dar uma de *bacana*.
 c. Dar uma de *João sem braço*.

Nos exemplos (6) e (7), as letras *X* e *Y* representam variáveis que podem preencher os espaços abertos nas construções esquemáticas. Já em relação às construções substantivas, percebe-se um comportamento menos livre, não há *slots* abertos para quaisquer itens lexicais de uma mesma classe ou grupo, a exemplo de:

- (8) a. Meu vizinho *bateu as botas*.
 b. ??Meu vizinho *bateu as sandálias*.
- (9) a. Ele *quebrou a cabeça* com a prova.
 b. ??Ele *partiu a cabeça* com a prova.

Nos exemplos (8) e (9), as sentenças apresentam construções idiomáticas fechadas, não há espaço para inserção de itens lexicais, ainda que a troca seja feita por palavras do mesmo campo semântico. Ainda, é importante ressaltar que os dois tipos de construções idiomáticas podem ter especificidades quanto ao preenchimento lexical, à sintaxe e à semântica, estes podem se apresentar de maneira idiossincrática, como será apresentado a seguir.

1.2.1 Idiossincrasia

A idiossincrasia se refere ao comportamento atípico dessas expressões em relação às regras gerais da gramática e aos princípios de composicionalidade. Esse fator está presente na composição de algumas EI's em um ou mais níveis. Essas idiossincrasias podem se manifestar no léxico, na semântica, na sintaxe e no discurso. A idiossincrasia semântica é marca das expressões idiomáticas, contudo, elas podem apresentar irregularidades de outros tipos.

Idiossincrasia lexical

Partindo de um ponto de vista lexical, em sentenças que obedecem às regras gerais e ao princípio de composicionalidade, trocas de palavras por outras do mesmo campo semântico são

possíveis (Fulgêncio, 2008). Entretanto, em EI's que apresentam idiosincrasia lexical, não se observa o mesmo princípio, como nos exemplos abaixo apresentados por Fulgêncio (2008):

(10) a. *Quebrar a cabeça.*

b. ? Partir a cabeça.

(11) a. *Mexer os pauzinhos.*

b. ? Mexer os palitos.

(12) a. Beber vinho.

b. Tomar vinho.

Nos exemplos (10) e (11), as EI's "quebrar a cabeça" e "mexer os pauzinhos" não podem ter itens lexicais substituídos por outros com proximidade semântica. No entanto, essa é uma possibilidade bastante possível em expressões regulares da língua como em (12). A troca de "tomar" por "beber" não afeta o sentido pretendido nem gera algum tipo de estranheza. Desse modo, algumas EI's estão sujeitas a idiosincrasias do tipo lexical.

Idiosincrasia semântica

As EI's têm como característica marcante o significado opaco. O significado das expressões não se dá em função da soma do significado dos itens que a compõem. Desse modo, a idiosincrasia semântica marca esse tipo de construção na língua.

(13) Meu irmão *pegou no meu pé* a tarde inteira.

(14) Ele presenciou a baderna, mas *fez vista grossa*.

Na expressão destacada em (13), a composição dos significados das palavras "pegar" e "pé" não resulta no significado da expressão como um todo, qual seja, "importunar alguém". Do mesmo modo, em (14) o significado idiomático "ignorar" não é resultante dos itens lexicais que compõem a EI. Isso acontece, pois há uma associação simbólica entre forma e função, como será apresentado na seção de referencial teórico.

Idiosincrasia sintática

Do ponto de vista sintático, as EI's podem ter comportamentos idiosincráticos que as distanciam dos padrões morfossintáticos aos quais uma sentença regular está submetida. Uma

das formas desse tipo de idiossincrasia se manifestar é pela impossibilidade de movimentação de sintagmas, como em:

- (15) a. *Fazer tempestade em copo d'água.*
 b. ? Em copo d'água, fazer tempestade.

Em (15), o elemento “em copo d'água” não pode ser movido para o início da sentença. É bem comum que os sintagmas, especialmente os de natureza adverbial, tenham certa liberdade em relação à posição na sentença. No entanto, isso não acontece na EI apresentada. Outra evidência que marca a idiossincrasia sintática são as modificações estruturais licenciadas pela relação complemento e objeto. Por exemplo, uma construção transitiva, com um verbo e um objeto direto, licencia uma construção passiva:

- (16) a. O menino comeu a maçã.
 b. A maçã foi comida pelo menino.
- (17) a. João *bateu as botas.*
 b.? As botas foram batidas por João.

Os exemplos em (16.a-b) são construções regulares da língua, que obedecem aos padrões sintáticos. Já em (17.a), há uma EI que, se compreendida como tal, no sentido de “morrer”, não pode ser passivizada. O exemplo (17.b) mostra que a modificação não é possível, mesmo com a composição verbo e objeto direto que licencia a construção passiva.

Ainda, no exemplo a seguir, há uma EI que acumula três formas de comportamento idiossincrático sintático:

- (18) a. Falou e disse!
 b. ? Falamos e dissemos!
 c. ? Disse e falou!

Em (18.a), há uma EI, comumente utilizada quando o interlocutor está em consonância com aquilo que foi dito/escrito pelo enunciador. Nela, percebe-se alguns comportamentos que fogem às regras de formação de sentenças. Em (18.b), não é possível acrescentar morfemas flexionais ao verbo, operação comum de concordância. Já em (18.c), há impossibilidade da inversão dos termos na expressão. Além disso, o verbo “dizer” não tem sua valência levada em consideração na expressão, visto que não possui um complemento como nas sentenças padrão.

É importante salientar que as possibilidades de uma EI se comportar idiossincraticamente do ponto de vista sintático são muitas. Os exemplos aqui apresentados representam algumas das possibilidades que podem se estender para outras EI's além das citadas. Sendo assim, o intuito é revelar esse tipo de comportamento de modo a delimitar a classificação de uma EI.

Idiossincrasia discursiva

Do ponto de vista discursivo, uma EI possui um comportamento idiossincrático quando não há nenhuma evidência, no contexto da enunciação, para que ela possua certo traço morfológico. Em algumas EI's, é possível perceber um comportamento atípico dos pronomes

(19) Você precisa esquecer dos problemas do passado e *partir pra outra*.

(20) a. A mensagem do WhatsApp estava muito esquisita, claramente era um golpe. Não acredito que ela *caiu nessa*.

No exemplo (19), a EI “partir para outra” não possui nenhum referente explícito no discurso, não há motivação para que a palavra “outra” esteja marcada no feminino. Já em (20a), o pronome “essa”, marcado no feminino, não se harmoniza morfológicamente com o antecedente “golpe”, palavra no masculino, de modo que não há nenhuma justificativa para que a palavra “essa” esteja no feminino, que é uma forma marcada. Dessa forma, a expressão “cair nessa”, independente do termo que retoma no discurso, se mantém no feminino. Seguindo regras gerais, a expressão seria como em (20.b), no entanto, o exemplo mostra a impossibilidade de flexão. Dessa forma, os exemplos apresentados revelam que essas expressões são convencionalizadas, apresentam-se dessa forma a despeito do contexto discursivo ou de algum antecedente que possam ter.

Idiossincrasias múltiplas

Uma EI se caracteriza por possuir ao menos o comportamento idiossincrático do tipo semântico. No entanto, uma expressão idiomática pode acumular idiossincrasias. Um exemplo é a expressão “de cabo a rabo”, que significa “inteiramente”, “por completo”:

(21) a. Ela leu a obra *de cabo a rabo* e ainda não conseguia entender.

b. ? Ela leu a obra *de cabo a cauda* e ainda não conseguia entender.

c. ? Ela leu a obra *de rabo a cabo* e ainda não conseguia entender.

Na sentença (21), a EI “de cabo a rabo” acumula comportamentos idiossincráticos dos tipos lexical, semântico e sintático. A idiossincrasia lexical é marcada em (21.b) pela impossibilidade da troca do item “rabo” por “cauda” mesmo com a relação de sinonímia. Em (20.c) o comportamento idiossincrático está na sintaxe, visto que há impossibilidade de mobilidade entre os termos que compõem a EI. Já a idiossincrasia semântica se apresenta no fato da expressão “de cabo a rabo” ser opaca, ou seja, o significado não se dá em função dos itens que a compõem.

Em resumo, é próprio das expressões idiomáticas o caráter idiossincrático, que as distingue de outras expressões regulares da língua. A seguir, será apresentada outra característica presente nas EI's: a convencionalidade.

1.2.2 Convencionalidade

Tagnin (2005) define convencionalidade como algo que é de praxe, comumente utilizado e aceito por uma comunidade. Para a autora, essa convencionalidade pode ser expressa em situações cotidianas, mas também na língua. Desse modo, a autora divide a convencionalidade em níveis, sendo eles: sintático, semântico e pragmático. Os níveis sintático e semântico se relacionam aos comportamentos idiossincráticos de uma EI dos tipos sintático e semântico (Fulgêncio, 2008). Já no nível pragmático, a convencionalidade está relacionada a um comportamento social esperado e ao uso das expressões consolidadas associadas a esse contexto. A autora cita o exemplo das situações em que se espera um agradecimento, sempre relacionadas a expressões como “muito obrigado” ou “muitíssimo obrigado”. Além disso, linguagens não verbais, como um aperto de mão em um velório com o objetivo de prestar condolências, também estão associadas à convencionalidade nesse nível. Ainda, Tagnin aponta que outras estruturas também são marcadas pela convencionalidade pragmática. Outros exemplos são os marcadores conversacionais, a exemplo de “pode repetir, por favor?”, frases feitas, como “divirta-se” e provérbios como “antes só [do] que mal acompanhado”. No caso das EI's, algumas expressões estão associadas a contextos de uso e são adotadas na prática de uso dos falantes de uma língua.

Sendo assim, os tipos de idiossincrasias apresentados (Fulgêncio, 2008) e a convencionalidade (Tagnin, 2005) caracterizam as EI's. Essas duas características são fundamentais para a compreensão de outro elemento que, possivelmente, descreve as EI's: a memorização em bloco, a ser explicada a seguir.

1.2.3 Memorização em bloco

Quando um falante nativo se depara com uma expressão como “*Ele deu com os burros n’água*”, é provável que ele não pensará em uma formulação que envolva o significado do verbo “dar”, nem dos itens “burros” ou “água”. Não será assim que ele chegará à ideia de que “algo deu errado”, “algo não aconteceu conforme o esperado”. A expressão claramente não possui um sentido literal, e o falante não parte do sentido denotativo para construir um sentido derivado. Dessa forma, não parece haver composição do significado item por item. Esse fato evidencia que as EI’s parecem ser memorizadas em bloco, sendo recuperadas da memória de longo prazo como se fosse um item lexical, e, portanto, sem a necessidade de serem interpretadas internamente.

Fulgêncio e Ciríaco (2018), ao discutirem as expressões idiomáticas, também apontam para a memorização em bloco. Segundo as autoras, “O falante nativo não processa uma expressão idiomática item por item. Na verdade, o significado dessas expressões é armazenado na memória do falante/ouvinte, como qualquer outra palavra da língua.” (Fulgêncio e Ciríaco, 2018, p.365, **tradução nossa**)². Ao aprenderem expressões idiomáticas, os falantes repetem esses construtos de maneira sistemática, ou seja, sem alterações do comportamento idiossincrático de determinada expressão. Esse fato parece ser uma evidência empírica da memorização em bloco, pois, se fossem processadas, as alterações iriam ocorrer de maneira livre, e não seria observada a convencionalidade das EI’s ou sua repetição sistematizada. Neste trabalho, essa hipótese da memorização em bloco será checada por meio de metodologia experimental (ver capítulo 3).

1.2.4 Questões levantadas pelas EI’s

Por serem marcadas por essas idiossincrasias, as EI’s são um desafio para a análise linguística com viés composicional. Kennedy (2008), por exemplo, afirma que, nessa concepção, “[...] os componentes da gramática devem ser analisados como módulos autônomos, independentes entre si, no sentido de que são governados por suas próprias regras, que não sofrem influência direta dos outros módulos.” (p.135). Ainda, as abordagens tradicionais postulam uma hierarquia entre esses módulos, tendo a sintaxe primazia sobre os demais, como a fonologia, a morfologia, o léxico, a semântica etc. Desse modo, são perspectivas teóricas que

² No original: “The native speaker does not process the idiom one lexical item at a time. In fact, the meaning of idioms is stored in the speaker/hearer’s memory, as is the meaning of any word of language.” (Fulgêncio e Ciríaco, 2018, p.365)

postulam uma clara separação entre léxico e sintaxe. As expressões idiomáticas são compostas por mais de um item lexical, o que as faz ser um obstáculo para o tratamento via módulo lexical. Em contrapartida, se a análise for feita a partir do módulo sintático, persistem os empecilhos porque as EI's não obedecem às regras gerais de formação de sentenças, como a do princípio da composicionalidade, além de outras idiosincrasias demonstradas.

Outro problema na análise de expressões idiomáticas é a tendência em tratá-las como metáforas, o que traz implicações para o processamento. Esse é um posicionamento adotado por alguns estudos e que se condensa em Xatara (1998):

“(...) em estudos sincrônicos (uma sincronia pura pode corresponder a 60 anos de história da língua segundo Rey-Debove), o obscurecimento de seus elementos formadores manifesta a espessura simbólica e metafórica de uma linguagem que é condensada, presente por toda parte, muitas vezes discretamente. Mas, mesmo se o sentido próprio de seus elementos estiver quase praticamente apagado e o sentido primitivo de uma EI totalmente perdido ou só puder ser restabelecido por um esforço de reflexão ou de explicações etimológicas, o falante continua preservando o sentimento inconsciente de que há uma imagem, ainda que não tenha a menor idéia quanto à sua origem.” (Xatara, 1998, p.155)

Ao assumir a premissa de que as EI's são, na verdade, expressões metafóricas (EM), são ignoradas tanto as evidências experimentais que reportam que a metáfora precisa de um processamento ativo, quanto a sua definição na literatura, que prevê um mapeamento entre domínios conceptuais distintos. Dessa forma, essa não é a postura que assumida neste estudo. Questiona-se o tratamento das EI's como EM's, assim como propõem Fulgêncio e Ciríaco (2018).

A tradição de compreender as EI's como EM's se dá por alguns motivos. A primeira razão trata-se do significado não literal das EI's, também presente nas EM's. Ambas as instanciações compartilham desse caráter; entretanto, alguns pontos chamam a atenção ao se contrapor os dois tipos de expressões. Segundo Lakoff e Johnson (2002), o sistema conceptual tem natureza metafórica. Dessa maneira, não só a linguagem, mas pensamentos e ações são conceituados a partir de nossas percepções sensório-motoras da experiência corpórea. Essa questão é ilustrada pelos autores com o conceito de “discussão” pela metáfora conceptual DISCUSSÃO É GUERRA, que instancia expressões metafóricas convencionais (EMC). Eles afirmam que “Muitas coisas do que fazemos numa discussão são parcialmente estruturadas pelo conceito de guerra. Embora não haja uma batalha física, há uma batalha verbal, que se reflete na estrutura de uma discussão – ataque, defesa, contra-ataque etc.” (Lakoff e Johnson, 2002, p. 47). Dessa maneira, há uma conceptualização de um domínio (discussão) com base em outro (guerra). Por isso, em cada metáfora um é possível identificar um domínio-fonte (por exemplo, guerra) e um domínio-alvo (por exemplo, discussão): aquele tem relação com a experiência

mais concreta e material da vida humana, enquanto este é mais abstrato e imaterial (Ferrari, 2018). Os conceitos metafóricos elencados por Lakoff e Johnson (2002) possuem um fator importante na sua compreensão: a convencionalidade. Desse modo, este estudo considera que expressões metafóricas se dividem entre as expressões metafóricas convencionais (EMC), elencadas pelos autores, e entre expressões metafóricas novas (EMN) e apresentam diferenças do ponto de vista da compreensão.

Na compreensão de metáforas novas (EMN), construídas no discurso e não convencionais, é preciso uma postura ativa tanto de quem enuncia quanto de quem é o interlocutor do enunciado. Isso acontece, pois a necessidade de interação entre os domínios-fonte e alvo exige um processo cognitivo por parte do usuário de uma língua para possibilitar o descolamento de sentido. Esse deslocamento de sentido demanda uma postura ativa dos envolvidos, pois passa pela percepção de uma má-formação semântica para a busca de um sentido plausível. Partindo do pressuposto de que as EI são fixas e recuperadas da memória como as palavras, levanta-se a hipótese de que o processo cognitivo que acontece para EMN's não é o mesmo que ampara o processamento de EI's. Dessa forma, acredita-se que medidas de processamento sejam um fator crucial para a distinção entre EI's e EMN's, como será proposto neste estudo.

Outro fator que também colabora para trabalhos que compreendem EI's como EM's é a motivação de significado. Quando um usuário se defronta com alguma EI e até algumas EF's opacas, é possível imaginar a lógica subjacente a ela. Em sua tese, Fulgêncio (2008) traz como exemplo o cronista Mário Prata e seu livro “Mas será o Benedito?”. Nele, o autor elabora uma série de verbetes, os quais a obra nomeia como “deliberadamente inventados”, em que apresenta o nascimento de diversas expressões idiomáticas e outras EF's, como ditos populares e provérbios. Em um deles, é expresso o seguinte histórico para “aos trancos e barrancos”: “Barranco é aquela ribanceira dos rios e tranco é salto de cavalo. Ou seja, vai-se com dificuldade. Ou ia-se, antigamente.” (Prata, p.26, 2011). O uso de “antigamente” na definição da EI “aos trancos e barrancos” reforça um ponto importante sobre a motivação do significado. É possível que essa EI tenha uma origem, fruto de uma interação entre domínios tal qual a metáfora. Mas isso se deu “antigamente”, como o histórico apresentado por Mário Prata. Não se nega que haja uma motivação metafórica para as EI, mas, como argumentam Fulgêncio e Ciríaco (2018), partir dela para o exame dessas construções seria um desajuste entre o que deve ser uma análise sincrônica e o que deve ser uma análise diacrônica.

Ao se levar em conta a origem dessas expressões idiomáticas, o que está sendo realizada é uma análise diacrônica. Entretanto, na sincronia, será que o falante nativo do PB faz as

operações metafóricas processuais, evidenciadas anteriormente, diante da enunciação de uma EI como “aos trancos e barrancos”? Algumas EI’s podem ser compreendidas pelos usuários da língua ainda que eles não a conheçam. Fulgêncio (2008) cita como exemplo o caso da expressão “memória de elefante” – como o elefante é comumente conhecido por ser um animal com uma boa memória, mesmo se um ouvinte nunca tenha entrado em contato com a expressão, consegue compreender o sentido, deduzir o significado, assim como as metáforas novas. No entanto, isso não é possível com todas as EI’s e, além disso, assim que a primeira exposição acontece, a EI já pode deixar um traço na memória que pode ir se fortalecendo com a frequência de exposição a ela. Com o aprendizado da expressão tal operação de inferência não se fará necessária.

Ainda é possível observar que, em relação à gênese de uma EI, quando se pensa em um contexto subjacente a ela, faz-se o movimento inverso no que diz respeito à compreensão de uma EM. Ocorre que esse movimento parte do significado já conhecido da EI para o significado metafórico, buscando criar uma situação comunicativa que a justifique. Fulgêncio e Ciríaco (2018) apontam que “Na verdade, o oposto acontece em uma expressão realmente metafórica – quando o significado não é conhecido, e o falante/ouvinte deve operar ativamente na produção do significado” (p.369, **tradução nossa**)³. Nesse sentido, uma EM, especialmente uma expressão metafórica nova, se diferencia de uma EI, pois envolve uma atividade cognitiva ativa. Isso porque os participantes do contexto comunicativo precisam realizar uma construção conceptual partindo de um significado literal para uma interpretação distinta para obter um novo significado. Nas EI’s, essa atividade não acontece, visto que são estruturas convencionalizadas, fixas e, e em hipótese, memorizadas como blocos, sendo, conseqüentemente, acessadas no momento do discurso, sem necessidade de processamento e construção do sentido.

Dado esse contexto, em resumo, este estudo pretende responder às seguintes questões: como se dá o processamento das EI’s? Seriam elas processadas como metáforas, corroborando a perspectiva tradicional, ou seriam elas processadas como palavras, sendo apenas recuperadas da memória?

1.3 Justificativa

As expressões idiomáticas são alvo de estudo de alguns trabalhos no campo dos estudos linguísticos, considerando todas suas áreas e vertentes de análise. Há dicionários com essas expressões, como “Dicionário onomasiológico de expressões idiomáticas usuais na língua

³ No original: “In fact, is the opposite of what happens in the case of a real metaphor - when the meaning is not known in advance, and the speaker/hearer has to work actively to produce the meaning.” (Fulgêncio e Ciríaco, 2018, p. 369)

portuguesa do Brasil” (Riva, 2009), descrições das tipologias das EI’s feitas por Tagnin (1989, 2005) e Xatara (1995,1998). Outras publicações abordam os aspectos de ensino-aprendizagem dessas expressões, tanto na perspectiva de aprendizagem de língua materna, quanto de adicional/estrangeira, como as de Xatara (2001), de Pedro (2007), de Nogueira (2008), de Termignoni e Finatto (2017) e outros.

Em muitos desses trabalhos encontramos a ideia de que as EI’s são EM’s, entretanto, partindo da descrição linguística com suporte em teorias que não consideram o processamento. Assim, este estudo se justifica porque parte de uma hipótese inovadora, a de que as EI’s não são EMN’s, e busca encontrar evidências que corroborem o ponto de partida que aqui se assume; e se propõe a buscar evidências experimentais para a hipótese delimitada teoricamente por Fulgêncio e Ciríaco (2018). Além disso, este trabalho se justifica porque se apoia numa abordagem construcionista baseada no uso, perspectiva também ainda inovadora para se tratar do tema.

1.4 Objetivos

1.4.1 Objetivo geral

O objetivo geral deste trabalho é contribuir para a descrição teórica da representação das expressões idiomáticas em português e para a verificação experimental de como são processadas por falantes nativos. Para isso, vamos contrastar o processamento de construções idiomáticas (CI), de construções metafóricas novas (CMN) e construções metafóricas convencionalizadas (CMC).

1.4.2 Objetivos específicos

- Descrever as EI’s no quadro teórico construcionista baseado no uso.
- Levantar EI’s em português.
- Observar e comparar o custo processual entre CI’s, CMN’s e CMC’s,

1.5 Hipóteses

O arrazoado exposto em Fulgêncio e Ciríaco (2018) delimita teoricamente a hipótese de que as expressões idiomáticas não são metáforas. Com base nas autoras, este estudo contrasta três diferentes expressões: construções idiomáticas (CI), construções metafóricas convencionais (CMC) e construções metafóricas (CMN) em um teste de leitura autocadenciada, que mede o tempo de reação (TR) de uma sentença segmentada por trechos. Dado que o tempo

de reação pode ser correlacionado ao custo de processamento, a hipótese experimental desta pesquisa pode ser elaborada nos seguintes termos:

- O tempo de reação das CMN's será maior do que o tempo de reação das CI's, nos trechos alvo e subsequente, ou, ao menos, no trecho subsequente. (*a ser explicado no Capítulo 3*).

1.6 Organização da dissertação

Além da introdução, esta dissertação conta com mais 4 capítulos. No capítulo 2, será apresentado o referencial teórico de Linguística Baseada no Uso (Bybee e Beckner, 2010; Diessel, 2017; Evans e Green, 2006, entre outros), Gramática de Construções (Goldberg, 1995, 2006, 2019), noções sobre metáfora como processo cognitivo (Lakoff e Johnson, 2002; Kövecses, 2002; Gentner e Bowdle, 2008) e um apontamento sobre processamento e acesso lexical (Savage e colegas. 2003; Bybee, 2016). No capítulo 3, será apresentada a metodologia experimental utilizada no estudo: apresentação do paradigma experimental de leitura autocadenciada, procedimentos e desenho experimental. No capítulo 4, serão apresentados os resultados do experimento de leitura autocadenciada que colocou em contraste as construções idiomáticas e as metafóricas, além de uma discussão sobre os resultados encontrados. Por fim, o capítulo 5 apresentará as considerações finais sobre a pesquisa.

Capítulo 2: Referencial teórico

O presente capítulo será dividido em um percurso que engloba pressupostos teóricos e descritivos na seguinte ordem:

- a) Linguística Baseada no Uso
- b) Gramática de Construções
- c) A metáfora como processo cognitivo
- d) Processamento e acesso lexical

2.1 Linguística Baseada no Uso

A Linguística Baseada no Uso (LBU) é uma teoria que visa explicar a língua com base no uso. Ela surge no fim da década de 1980 a partir dos estudos desenvolvidos pela Linguística Cognitiva. À época, a LBU surgia como um contraponto à abordagem tradicional vigente:

“A Linguística Baseada no Uso questiona: por que as línguas têm gramática e como ela surge? Qual forma ela assume em determinada língua e entre as línguas, como os processos de domínio geral criam a gramática? Para além das limitações que a gramática estrutural e gerativa impõe no estudo da língua, a Teoria Baseada no Uso analisa a forma como a experiência e a linguagem impactam diretamente na representação cognitiva da linguagem.” (Bybee, 2023, p.9, **tradução nossa**)⁴

A LBU propõe que a linguagem é uma habilidade que resulta da cognição geral, ou seja, não há um recurso mental exclusivo para essa habilidade, nem adoção de uma visão modular para a representação do conhecimento. Ainda, propõe que o uso tem efeito direto na estrutura da língua. Essa perspectiva não postula separação entre competência (internalista), e performance (dados reais de produção dos falantes) dos estudos (Bybee e Beckner, 2010; Diessel, 2017; Evans e Green, 2006, entre outros).

Com base em estudos das Ciências Cognitivas, especialmente da Psicologia Cognitiva, e da Filosofia e da Ciência da Computação, a LBU busca uma compreensão da linguagem com base em uma perspectiva empírica e psicologicamente plausível (Evans e Green, 2006; Langacker, 2008). Nesse sentido, a representação do conhecimento linguístico não difere da representação de outras estruturas conceituais. Exemplos dessa hipótese são os processos de

⁴No original: “Usage-based linguistics asks the questions, why do languages have grammar and how does it arise? What form and shape does it take within and across languages, and how do domain-general processes create grammar? Moving beyond the limitations that structural and generative grammar put on language study, Usage-Based Theory looks at the way experience with language directly impacts the cognitive representation of language.” (Bybee, 2023, p. 9)

categorização, a aprendizagem estatística inferencial e o uso da memória. Os recursos utilizados pela linguagem, seja do ponto de vista do aprendizado, da produção ou do processamento, são recursos cognitivos compartilhados com outras habilidades não linguísticas. Ocorre que a linguagem, como instrumento de comunicação, recruta recursos e formas de aprendizagem que estão disponíveis a outros domínios cognitivos. Desse modo, as instanciações da língua são fruto de um aprendizado com base no *input* e na categorização (Bybee e Beckner, 2010), dentre outros processos.

A LBU assume alguns pressupostos fundamentais, são eles: a linguagem advém da cognição geral, a cognição é corporificada, a língua emerge do uso e tem função comunicativa, o *input* e a frequência são relevantes, e semântica, pragmática e informação estrutural são interrelacionadas.

2.1.1 Embodiement: linguagem e cognição corporificada

Segundo Bybee (2016), “[...] os fenômenos estruturais observados na gramática das línguas naturais podem ser derivados de processos cognitivos de domínio geral, já que operam em múltiplos casos de uso da língua.” (Bybee, 2016, p.18). Um desses processos, por exemplo, é a repetição, segundo Bybee e Beckner (2010). Em habilidades motoras, como andar de bicicleta, a repetição é uma estratégia que leva ao aprendizado e a uma posterior automação do processo, com sua representação na memória de longo prazo e seu agrupamento em subpartes (*chunks*). O mesmo acontece com a linguagem: a repetição, juntamente a outros processos, leva o usuário a fazer do conhecimento linguístico um conhecimento implícito. Além disso, há outras evidências, como estudos que apontam que rastreamos as co-ocorrências e as sequências de ocorrências, não só de enunciados linguísticos, mas também daqueles de outra natureza. Um exemplo disso são as crianças que conseguem rastrear as ocorrências de exposição e agrupar as co-ocorrências de uma palavra ou expressão com a de um conceito emocional (Nencheva e colegas, 2023).

Portanto, o conhecimento linguístico concebe não só um conhecimento gramatical, mas também um conhecimento de mundo. Isso reflete a perspectiva empirista da teoria, que coloca o corpo como central na experiência humana. Nesse sentido, a Linguística Baseada no Uso entende a cognição como corporificada, pois a experiência, a cognição e a realidade dos usuários estão ancoradas no corpo (Bybee e Beckner, 2010; Ferrari, 2018, Rohrer, 2017). Não à toa costuma-se categorizar experiências e estruturar nossa cognição por meio de conceitos imagéticos como *CONTÊINER* e também pela esquematização de conceitos e palavras por meio de *Frames* (Evans e Green, 2006; Rohrer, 2017). Essas são evidências de que a experiência e a

cognição corporificada são a base do conhecimento linguístico. Dessa maneira, a abordagem compreende a linguagem como uma habilidade cognitiva, fruto de recursos cognitivos gerais, que são influenciados pela experiência corporificada dos usuários da língua.

2.1.2 O caráter emergentista e a função comunicativa da linguagem na LBU

A linguagem é uma habilidade cognitiva que tem função comunicativa. Esse é um pressuposto que pode ser compreendido por capacidades cognitivas gerais da espécie humana: a leitura de intenções e a identificação de padrões. Segundo Tomasello (2000, 2009), a leitura de intenção comunicativa pode ser explicada nos termos de um sujeito compartilhando a atenção com alguém, ambos com foco em uma mesma entidade/referente. Essa é uma capacidade que nos permite compreender um outro ser humano como um agente mental, com crenças sobre o mundo, que podem, inclusive, não se confirmar na realidade. A leitura de intenções recruta habilidades cognitivas gerais que são utilizadas para outros fins e outras práticas da vida humana. Nessa leitura, são incluídos recursos como a atenção compartilhada, capacidade de coordenar a atenção entre dois parceiros sociais, a atenção a referentes fora da interação imediata, direção da atenção para outros referentes distantes utilizando gestos, além do aprendizado das ações intencionais do outro e dos atos comunicativos que advêm dessas intenções. Nesse sentido, é possível perceber que a função comunicativa não envolve simplesmente a transmissão de informações, mas o compartilhamento de um objetivo em comum: a interação social (Croft, 2000). Ainda que, a priori, pareça uma função e um objetivo trivial, não o é. Um grupo de chimpanzés, por exemplo, também consegue seguir um mesmo objeto com o olhar, mas a diferença, ao que parece, está no fato de que esses animais não sabem que estão realizando a mesma tarefa, juntos, ao mesmo tempo, nem há uma leitura de intencionalidade. Diferente de crianças pequenas, que, entre 9 e 12 meses de idade, começam a compartilhar essa intenção comunicativa e também a atenção (Tomasello, 2000, Tomasello e colegas, 2005; Tomasello e Carpenter, 2007). Há, portanto, na leitura de intenção, uma tentativa de manipular os estados intencionais do outro.

Como exemplificação, podemos citar os bebês, que, assim como humanos em outras idades, também utilizam o gesto para se comunicar, apontando para objetos. Se um cuidador conversa com uma criança pequena, aponta para um objeto e diz “pega!”, ela precisa reconhecer, por meio do contexto comum de atenção compartilhada, as pistas para compreender a mensagem pretendida. Além disso, ela precisa entender a intenção do cuidador, o imperativo, e, a partir disso, decidir ou não por atender ao pedido de pegar o objeto, transformando seu estado intencional.

Já a identificação de padrões revela a capacidade que os seres humanos possuem não só de identificar esses padrões, mas também de fazer operações estatísticas sobre eles. Isso inclui desde habilidades sensório-motoras como engatinhar até a identificação de sequências sonoras da língua. Ao identificar padrões, é possível utilizar recursos cognitivos como a categorização de eventos e objetos, identificação de esquemas sensório-motores como ações, análises distribucionais (reconhecimento de sequências) e criação de analogias com base na funcionalidade. Com tudo isso à mão, as crianças conseguem identificar o modo como os adultos utilizam os símbolos linguísticos nos mais diversos contextos. A partir dessa identificação de padrões, possível pela atenção compartilhada, nos mais diferentes contextos, a estrutura abstrata da língua (forma e função) emerge (Tomasello, 2003).

Dado o exposto, é possível notar a importância do contexto real de comunicação para a LBU. Segundo Barlow e Kemmer (2000), “um modelo baseado no uso é aquele em que o sistema linguístico do falante é fundamentalmente baseado em ‘eventos de uso’, instâncias da produção e da compreensão da linguagem” (p.8, **tradução nossa**)⁵. Sendo assim, a experiência dos falantes, nos mais diversos eventos comunicativos, é a responsável pelo aprendizado da língua e também pela produção e compreensão. O *input* recebido nesses contextos diversos juntamente com as habilidades cognitivas de categorização e a memória fazem a estrutura da língua emergir. Em resumo, as unidades da língua não são parte de uma gramática universal, elas são fruto de um processo de exposição (*input*) e de categorização, logo, a língua é motivada pelos eventos de uso.

2.1.3 O papel do *input* e da frequência

Os compromissos da Linguística Baseada no Uso apresentados até então deixam claro a importância da experiência do falante para essa abordagem. Desse modo, a exposição de um falante a uma língua também é fundamental para o aprendizado e para produção ao longo da vida. Assim, o *input* e a frequência são fatores que orientam a abordagem. O *input* pode ser definido como evidências positivas e evidências relevantes (evidências negativas indiretas) encontradas no ambiente em que um usuário está aprendendo uma língua. Nessa visão, não há estímulo pobre, pelo contrário, a LBU aponta para uma riqueza de estímulos. Os *inputs* linguísticos com os quais os sujeitos entram interagem com outros elementos do processo cognitivo para que então os usuários consigam abstrair as regularidades da língua. A LBU,

⁵ No original: “a Usage-Based model is one in which the speaker's linguistic system is fundamentally grounded in 'usage events' instances of a speaker's producing and understanding language” (Barlow e Kemmer, 2000, p.8)

portanto, reconhece a relevância do estímulo para o aprendizado de uma língua, pontuando esse critério como suficiente para explicar como aprendemos uma língua (Tomasello, 2009).

Assim, a exposição que as pessoas têm a uma determinada instância da língua forma a frequência de *input*. Esses dados do uso são categorizados e memorizados, para, assim, os padrões da língua emergirem, pois “(...) se o sistema linguístico é a função do uso da linguagem, então, segue-se que a frequência relativa em que cada palavra ou outra construção é encontrada pelo falante, afetará a natureza desse sistema linguístico.” (Evans e Green, p. 113, 2006)⁶. O usuário da língua, portanto, tem contato com determinada construção e desde a primeira exposição já retém traços dela em sua memória. À medida que determinada estrutura vai sendo testemunhada pelo aprendiz, mais entrincheirada na memória ela ficará. Essa é uma realidade cognitiva que faz com que certas estruturas linguísticas se tornem convencionais. Além das informações linguísticas, o contexto de uso também é registrado na memória, incluindo dimensões pragmáticas. Ainda, o registro de frequência não se relaciona apenas à quantidade de vezes que determinada estrutura foi testemunhada, mas também suas relações de co-ocorrência com outras instâncias, bem como relações de similaridade (Bybee e Beckner, 2010), visto que os usuários são capazes de estabelecer relações e comparações de similaridade entre as instanciações.

Importante ressaltar também que a frequência não é o único fator para explicação do aprendizado e da produção linguística de um falante, outros efeitos e habilidades cognitivas também têm seu papel. A frequência gera importantes efeitos linguísticos que explicam a aprendizagem e a produção, no entanto, não é a única responsável por elas, é um fator que explica muito, mas não explica tudo (Zyzik, 2009). A função comunicativa da linguagem, restrições da memória de trabalho, processos indutivos e a função de uma construção, por exemplo, também podem explicar certas tendências no uso (Perek e Goldberg, 2015).

2.1.4 Composicionalidade e significado

O Princípio da Composicionalidade (conhecido como Princípio de Frege⁷) é um postulado utilizado para explicar os significados das sentenças e expressões produzidas pelos usuários. A formulação desse princípio se apresenta da seguinte maneira: “O sentido de uma expressão complexa é uma função dos sentidos das suas partes e da forma como estão

⁶ No original: “(...)the language system is a function of language use, then it follows that the relative frequency with which particular words or other kinds of constructions are encountered by the speaker will affect the nature of the language system.” (EVANS e GREEN, p. 113, 2006)

⁷ Apesar do nome, não é um consenso nos estudos linguísticos de que o Princípio da Composicionalidade seja um conceito de Frege, pois não há menção explícita ao conceito nas publicações do filósofo.

sintaticamente combinadas” (Partee, 1984 *apud* Goldberg, 2015, p. 419). Desse modo, o significado de uma instanciação se dá pela soma do significado de suas partes/constituintes, a partir de um modelo lógico-matemático que relaciona formas linguísticas a condições de verdade. Em outras palavras, somando o significado de cada item lexical, na ordem que cada um aparece, tem-se o significado do todo. Com base em Dowty (2006), Goldberg (2015) elenca os principais argumentos para a composicionalidade:

- (i) usuários de uma língua produzem e compreender sentenças que as quais ainda não tinham sido expostos;
- (ii) há uma concordância entre os usuários a respeito do significado de sentenças;
- (iii) as sentenças são infinitas e, por isso, não podem ser memorizadas;
- (iv) deve existir algum procedimento para determinar um significado;
- (v) as sentenças são geradas pela gramática da língua;
- (vi) o procedimento para interpretar sentenças deve ser determinado, de alguma maneira, pelas estruturas sintáticas geradas pela gramática juntamente com as palavras.

O apelo desse princípio está na tentativa de se justificar a capacidade de interpretação de sentenças novas, ou seja, a criatividade linguística dos usuários da língua. Há a ideia de que deveria haver um mecanismo de predicação que tornasse possível essa capacidade. No entanto, essa prerrogativa coloca o léxico com um “componente” que carrega o significado de todas as expressões de uma língua. A visão formalista, ao ser confrontada com o uso da língua, encontra problemas, pois muitas instancicações não podem ter seu significado dado a partir dessa função.

Um desafio é o tratamento das expressões idiomáticas, que têm significados que ultrapassam a função proposta pelo princípio, sendo compreendidas como não composicionais. Como tentativa de manter o Princípio da Composicionalidade, é proposto que as sentenças, de modo geral, possuem constituintes (átomos) que são “calculados” por regra algébrica de composicionalidade (trivial); no entanto, as mesmas sentenças podem ter interpretações não triviais (Hodges, 2012). Por exemplo, uma expressão como ‘bater as botas’ pode ter seu significado pela função dos seus itens ‘bater + as + botas’, gerando um sentido literal (trivial), mas também pode possuir o sentido idiomático: morrer (não trivial). Uma proposta para os idiomatismos, com base na teoria de Hodges, é trabalhada por Westerståhl (1999). Nela, o autor sugere que a composicionalidade é uma questão para a semântica, não para sentenças individuais. Com esse pressuposto, Westerståhl (1999) enfatiza que, embora faça sentido perguntar se uma semântica é ou não composicional, não faz sentido fazer a mesma pergunta sobre uma frase específica. Para isso, ele fornece um exemplo em que o significado de uma

frase não idiomática “*kick the dog*” (chutar o cachorro) é uma função dos significados “*kick + the + dog*”, enquanto a expressão idiomática “*kick the bucket*” tem um significado diferente: “morrer”. O autor argumenta que essa diferença de significado não diz nada sobre o Princípio da Composicionalidade, já que o princípio apenas pede que alguma função determine o significado, e não que essa função seja específica. Ele demonstra ainda que redefinir a função para o caso idiomático pode resolver a questão, destacando que a questão da composicionalidade é mais apropriadamente colocada no nível da semântica do que em frases individuais.

Outro ponto de vista considera que os idiomatismos possuem constituintes, visto que alguns podem sofrer modificações de tempo, modo e aspecto verbal, além de movimentações, passivação, entre outros (Goldberg, 2015). Aquelas EI’s que possuem movimentações de constituintes e continuam com o sentido idiomático são consideradas idiomas deformáveis, como no exemplo abaixo:

(28) a. *Pull strings* (fazer uso da influência/conexões pessoais, no sentido idiomático)

b. *You only get a good job if someone pulls strings for you.*

Você só consegue trabalho se alguém “mexer os pauzinhos” para você.

c. *A lot of strings were pulled to get him the part in the movie.*

Muitos pauzinhos foram mexidos para que ele fizesse parte do filme.

O idiomatismo presente em (28.a-c) pode ser traduzido para a expressão idiomática do PB “mexer os pauzinhos”, de sentido semelhante. Em (b), temos a expressão com o verbo na 3ª pessoa do singular. Já em (c) temos “*lot*” quantificando o substantivo “*strings*”, indicando que, de alguma maneira, esse substantivo é interpretado e possui o sentido de “conexões” (Goldberg, 2015). A autora aponta que, para preservar o Princípio da Composicionalidade, poderia ser afirmado que cada palavra que compõe uma expressão idiomática possui uma parte do significado do todo. Porém, essa posição não consegue se sustentar, pois o sentido de “conexões” só é atribuído quando “*strings*” está no plural, nunca no singular. Há uma dependência contextual muito forte, ou seja, não se trata apenas de um significado de constituintes.

Além das questões relacionadas à idiomaticidade, outras estruturas da língua levam a uma reflexão sobre a adoção desse princípio. Em estruturas ambíguas como “O homem viu o

incêndio do prédio”, há duas estruturas subjacentes⁸ com as seguintes paráfrases: (a) um homem estava no prédio e de lá viu um incêndio; e (b) um homem estava em qualquer outro lugar e viu um prédio pegando fogo. Isto posto, a composicionalidade seria aplicada à estrutura subjacente, o que implica em um limite pouco claro em relação a quão diferente uma estrutura superficial pode ser da estrutura subjacente. Sem essa especificação, qualquer estrutura superficial poderia ter, em tese, qualquer significado na língua (Goldberg, 2015).

Outro caso que se mostra um desafio para a aplicação do Princípio da Composicionalidade são as construções de estrutura argumental (CEA). Elas apresentam um padrão esquemático abstrato que independe, em parte, do significado do verbo. Tradicionalmente, assume-se que o verbo é quem determina os papéis dos argumentos de uma sentença. No entanto, na perspectiva construcionista, um padrão abstrato também veicula significado, como nos exemplos a seguir, de Ferrari, (2018):

(29) O urso panda espirrou.

(30) A baleia espirrou água para dentro do navio.

Se o significado fosse dado em termos do verbo, apenas a sentença em (29) seria possível, ou um verbo com novo sentido deveria ser criado para comportar o sentido em (30). Porém, o padrão construcional em que estão situadas as sentenças em (29) e (30) contribui para o significado:

(29') X_{AGIR}

(30') $X_{\text{CAUSAR}} Y \text{ a } \text{MOVER } Z.$

Nesse caso, o significado não é uma função do significado dos itens lexicais, pois o próprio padrão sentencial veicula significado. Além disso, outros fatores contribuem para o significado de uma expressão, como os contextuais e/ou discursivos. No português, sentenças declarativas e interrogativas podem assumir a mesma estrutura, por exemplo: “Ele ligou ontem” ou “Ele ligou ontem?”. A entonação é que determina a interpretação declarativa ou interrogativa. É possível citar também a dependência contextual de certas estruturas como “Não tem ninguém aqui!” (ninguém exceto quem proferiu a frase). Ainda, há uma certa circularidade na definição do Princípio da Composicionalidade. A proposição explica o significado pela

⁸ Em termos gerativistas, de maneira simplificada, a estrutura superficial é a organização das unidades de uma sentença que é interpretada pela interface fonética. Já a estrutura subjacente é a estrutura abstrata que determina a interpretação semântica.

maneira como as sentenças são construídas e a forma como as sentenças são construídas explicam o significado (Goldberg, 2015).

Essas são algumas das evidências para uma revisão do Princípio da Convencionalidade. Tratar esse princípio de maneira tão estrita traz uma série de problemas teóricos, pois a combinação dos constituintes não explica todos os significados sentenciais e nem os pensamentos que são veiculados nessas expressões. O significado no contexto é fundamental na perspectiva adotada nesta pesquisa. As pistas contextuais e o compartilhamento do conhecimento de mundo colaboram para a compreensão de enunciados da língua. Há, inclusive, trabalhos que apontam para uma subespecificação natural da linguagem (*ver Piantadosi e colegas, 2012*), o que evidenciaria ainda mais a importância do contexto. A ideia de subespecificação também pode ser pensada pelo viés do gesto. Ao apontar para algum lugar, no intuito de compartilhar alguma informação com o interlocutor, o gesto por si só não é capaz de revelar todo o significado, é o contexto compartilhado que informa a motivação comunicativa do ato de apontar (Goldberg, 2015).

Desse modo, uma revisão da Composicionalidade no tratamento teórico se faz necessária, compreendendo que o significado está para além de uma estrutura interna de uma expressão, dos itens que a compõem. Além disso, a motivação do significado está no uso, o que o torna flexível e dinâmico. Assim sendo, uma abordagem construcionista baseada no uso propõe uma visão mais flexível para esse princípio. A proposta da Gramática de Construções (Goldberg, 1995, 2006, 2019, *ver seção a seguir*) apresenta um modelo em que, mesmo para expressões idiomáticas, o princípio pode ser mantido se a composicionalidade for considerada sob outra perspectiva.

2.2 A Gramática de Construções (Goldberg, 1995, 2006, 2019)

A Linguística Baseada no Uso adota como modelo de conhecimento uma abordagem construcionista. Foi com Fillmore e colegas (1988), no estudo da construção “*let alone*”⁹, que a abordagem construcionista se fundou. Eles partiram da observação de que as regras sintáticas e de inserção lexical propostas tradicionalmente deixavam de fora uma série de produções da língua por conta da irregularidade. Isso posto, pode-se afirmar sobre a abordagem construcionista: (i) as unidades da língua são um pareamento de forma e função; (ii) a gramática não compreende regras derivacionais e transformacionais, pois são fruto de processos

⁹ Fillmore e colegas (1988) utilizam a expressão idiomática “*let alone*” do inglês para abordar a ideia de regularidade e idiomatismo na gramática. Essa é uma expressão que pode ser usada em uma sentença como “*He can't even run a mile, let alone a marathon*” (Ele não consegue nem correr uma milha, quanto mais uma maratona).

cognitivos de domínio geral, (iii) o conhecimento linguístico é ancorado no uso, (iv) as construções de uma língua formam uma rede de construções e (v) a frequência é um fator importante para explicar o comportamento linguístico dos falantes. Há várias abordagens de caráter construcionista que se aproximam e se distanciam em alguns pressupostos (Fillmore, 1988, Goldberg 1995; 2006; Croft, 2001, entre outros). A abordagem adotada neste estudo é a Gramática de Construções de Goldberg (1995, 2006, 2019).

A GC de Goldberg sofreu algumas adequações ao longo do desenvolvimento das pesquisas e dos resultados experimentais. Em sua obra de 1995, Goldberg define construção da seguinte maneira: “C é uma CONSTRUÇÃO se é um par forma-significado $\langle F_i S_i \rangle$ de tal modo que algum aspecto de F_i ou de S_i não é previsível a partir de algum componente de C ou de outras construções previamente estabelecidas.” (Goldberg, 1995, p.4). Nessa compreensão, uma construção poderia ser definida na condição de que nenhum elemento seja previsível a partir de suas partes ou de outras construções. Essa é uma definição que a própria autora considerou “estrita” (Goldberg, 2019), pois desconsiderava a ideia de redes de construções. Goldberg (2006), então, revê a definição:

“Qualquer padrão linguístico é reconhecido como uma construção se algum aspecto de sua forma ou função não pode ser estritamente previsível das partes que a compõem ou de outras construções reconhecidas como existentes. Além disso, os padrões são armazenados como construções, ainda que totalmente previsíveis, se ocorrem com frequência suficiente.” (Goldberg, 2006, p. 5)

Assim, a definição começa a considerar não só a rede de construções, mas também o papel da frequência de uso. Essa postulação é resultado do reconhecimento de que a língua possui informações redundantes que levam à convencionalidade de certas instanciações de construções. Como exemplo, Goldberg (2006) chama atenção para o fato de que padrões linguísticos bastante gerais instanciam construções convencionais, como no caso do idiomatismo “*go tell it to the mountain*” (algo similar ao idiomatismo “vá e espalhe a palavra” no PB) que reflete o padrão geral “*go VP_{bare}*”. Porém, ao postular na definição que as construções precisam ter frequência suficiente, o papel da memória não é tão relevante como a própria autora compreende posteriormente. Em sua definição mais recente, Goldberg (2019) propõe

“[...] uma compreensão ainda mais inclusiva do que são as construções, motivada por uma melhor apreciação da memória humana, aprendizado e categorização. [...] as construções são compreendidas como *clusters* (aglomerados) emergentes de traços de memória imperfeitos que estão alinhados em nosso espaço conceptual de alta(hiper!) -dimensão, com base no pareamento de forma, função e dimensões contextuais”. (Goldberg, 2019, p. 7)¹⁰

¹⁰ No original: “[...] a still more inclusive understanding of what constructions are, motivated by a better appreciation of human memory, learning, and categorization. Here, as explained in the following chapters,

Desde o primeiro contato com uma construção, traços de memória já serão retidos, alimentando a representação, ou *cluster*, daquele exemplar linguístico, sempre com base no pareamento de forma, função e da dimensão contextual. O *cluster* pode ser definido como uma representação emergente de uma construção, com traços de memória que incluem aspectos como forma, função, contexto, etc. Desse modo, a cada vez que determinada instanciação for testemunhada, o traço será reforçado na memória, se for em um mesmo contexto. Se diferente, um novo traço relativo a um novo contexto ampliará esse *cluster* formando uma rede interrelacionada relativa a essa construção que pode variar dentro de um *continuum*. Tendo isso em vista, a dimensão contextual e a formação do espaço conceptual hiperdimensional, com base no *input* e na frequência, são importantes para a análise das construções idiomáticas e sua memorização em bloco (*chunks*). Essa concepção ajuda a explicar a convencionalidade das expressões e a repetição sistematizada delas, a despeito dos comportamentos idiossincráticos observados. Isso ocorre, pois o usuário da língua se apoia tanto em princípios de criatividade quanto de conservadorismo, bem como de eficiência e acessibilidade (além de outros princípios). Ele tende a usar palavras e estruturas já familiares e empregá-las em novo sentido. Esse fato pode ser observado em EI's, que se apropriam de palavras e estruturas familiares em uma outra função. Assim, com a adoção da construção idiomática pelos falantes, ela se torna convencional, seguindo o Princípio da Idiomaticidade (ver Erman e Warren, 2000; Sinclair, 1991; Wray, 2002).

Desde a primeira definição de 1995 até a mais atual, de 2019, as construções são tidas como unidades básicas da língua, com um modelo de representação uniforme para todas essas unidades. Assim, a Gramática de Construções não postula uma separação estrita entre léxico e sintaxe. Postula ainda que as construções são um pareamento entre forma e função de modo que essas construções podem variar em níveis de complexidade e abstração, incluindo as estruturas regulares da língua e também as estruturas com baixa frequência de uso, ditas periféricas, como é possível observar na tabela a seguir.

Tabela 1 - Construções variando em nível de complexidade e abstração

Construção	Exemplos
Palavra	Iran, another, banana

constructions are understood to be emergent clusters of lossy memory traces that are aligned within our high-(hyper!) dimensional conceptual space on the basis of shared form, function, and contextual dimensions.” (GOLDBERG, 2019, p. 7)

Palavra (parcialmente preenchida)	Pre- <i>N</i> , <i>V</i> -ing
Idiomatismo (preenchido)	Going great guns, give the Devil his due.
Idiomatismo (parcialmente preenchido)	Jog <someone's>, memory, <someone's> for the asking.
Idiomatismo (minimamente preenchido) The Xer the Yer	The more you think about, the less you understand.
Construção ditransitiva: Suj V Obj₁ Obj₂ (não preenchida)	He gave her a fish taco; He baked her a muffin.
Passiva: Suj aux VPpp (PPby) (não preenchida)	The armadillo was hit by a car

Fonte: Goldberg (2013)

Os exemplos apresentados por Goldberg (2013), ainda que em língua inglesa, evidenciam que as construções se encontram em um *continuum* que vai das mais preenchidas e substantivas, como as palavras, até as mais esquemáticas e abstratas como as construções passivas, sempre pareando a forma e a função. A tabela apresenta o exemplo da palavra “banana”, uma construção totalmente preenchida, substantiva. Esquemáticamente, essa palavra se encaixa na construção de forma *N* com a função de nomear seres, entidades, objetos, esquema que também representa a palavra “Iran”. Desse modo, o *continuum* se apresenta com estruturas mais substantivas como palavras específicas (banana, Iran), representações esquemáticas como *N*. Há também padrões compostos de *slots* abertos e partes fixas, como no idiomatismo: “Quanto mais *X*, mais *Y*” (“Quanto mais eu rezo, mais assombração me aparece”, “Quanto mais penso, mais o tempo passa”). As redes de construções também incluem padrões de estrutura argumental como “O menino jogou a bola”, com construção sintática representada por *Suj V Obj*, e que, ainda, possui um padrão mais esquemático: *X*_{AGIR} *Y*, o qual abarca outras instâncias transitivas.

Em resumo, na abordagem construcionista, todas as estruturas da língua podem ser unidades analisáveis como construção. É possível dizer que a Gramática de Construções traz uma abordagem psicologicamente e cognitivamente plausível para a descrição de estruturas linguísticas. Essa posição é derivada de um entendimento de que o conhecimento das estruturas

de uma língua (ou de suas construções) emerge de contextos de uso, instâncias efetivas e reais fruto das interações entre usuários de uma língua.

2.2.1 A aprendizagem das construções: *input*, frequência e memória

A intenção do falante é compreender as mensagens a partir das formas que testemunha e produzir formas a partir da mensagem que deseja transmitir (Goldberg, 2019). Esse processo ocorre em conformidade com as convenções estabelecidas pela comunidade linguística da qual ele faz parte e também segue alguns princípios que Goldberg (2019) desenvolve a partir do acrônimo CENCE ME (Construção, Eficiência, Novas informações, Competição, Expressividade, Memória e aprendizagem impulsionada por Erros), alinhados com a LBU:

- “A. Os falantes equilibram a necessidade de serem Expressivos e Eficientes, ao mesmo tempo em que se conformam com as convenções de suas comunidades linguísticas.
- B. Nossa Memória é vasta, mas imperfeita: traços de memória são retidos, porém são parcialmente abstratos ("imperfeitos").
- C. Memórias "imperfeitas" são alinhadas quando compartilham aspectos relevantes de forma e função, resultando em aglomerados emergentes e sobrepostos de representações: Construções.
- D. Novas informações estão relacionadas a informações antigas, resultando em uma rede ampla de construções.
- E. Durante a produção, várias construções são ativadas e Competem entre si para expressar nossa mensagem pretendida.
- F. Durante a compreensão, discrepâncias entre o que é esperado e o que é observado ajustam nossa rede de construções aprendidas por meio de um aprendizado impulsionado por Erros.” (Goldberg, 2019, p.6, **tradução nossa**)¹¹

Sobre os princípios elencados, é preciso evidenciar que, em alguma medida, podem soar contraditórios, mas não o são. A eficiência é um princípio que corresponde ao uso da língua de maneira mais concisa, comunicar com poucas estruturas. Já a expressividade diz respeito a uma escolha linguística capaz de expressar de maneira compreensível as ideias, crenças e atitudes do falante. Se a expressividade fosse levada ao máximo patamar, as construções da língua poderiam aumentar muito, para designar especificamente cada nível de pensamento, crença ou ideia. Por outro lado, uma eficiência maximizada poderia reduzir muito as formas associadas a comunicações, tornando-as muito gerais e inespecíficas. O usuário, desse modo, mantém um

¹¹ “A. Speakers balance the need to be Expressive and Efficient while conforming to the conventions of their speech communities.

B. Our Memory is vast but imperfect: memory traces are retained but partially abstract (“lossy”).

C. Lossy memories are aligned when they share relevant aspects of form and function, resulting in overlapping, emergent clusters of representations: Constructions.

D. New information is related to old information, resulting in a rich network of constructions.

E. During production, multiple constructions are activated and compete with one another to express our intended message.

F. During comprehension, mismatches between what is expected and what is witnessed fine-tune our network of learned constructions via Error-driven learning.” (Goldberg, 2019, p.6)

equilíbrio, que ora tende à expressividade, ora à eficiência, a depender da leitura da situação de uso em que ele se encontra, da necessidade daquele momento.

Para além disso, a necessidade de se obedecer a convenções da comunidade de fala também instaura mais um fator relevante, mostrando que os usuários de uma língua têm uma percepção conservadora dela. Assim, estruturas que correspondem a padrões construcionais já instaurados na língua, ainda que contemplam os princípios de eficiência e expressividade, não são utilizadas pelos usuários da língua, pois não estão associadas a determinada mensagem pretendida. O exemplo para traduzir esse comportamento apontado por Goldberg (2019) está na construção “? *explain me this*”, que pode ser facilmente compreendida e corresponde a um padrão sintático do inglês bastante produtivo, a construção de duplo objeto. Esse padrão argumental pode ser instaurado por outros verbos como “*tell*”, “*buy*”, “*give*”, “*ask*”, por exemplo. Entretanto, os usuários julgam pouco convencional o uso do verbo “*explain*” nesse padrão construcional e preferem a “*explain this to me*”, uma estrutura que, mesmo sendo mais longa, corresponde a estrutura convencionalizada para expressar a mensagem pretendida. Ainda que pareça contraintuitivo, os usuários preferem uma formulação mais longa, porém convencional, do que uma mais curta. Essa relação é denominada produtividade parcial, ou seja, não é toda e qualquer estrutura que preenche um esquema de uma construção. Em português brasileiro, um estudo apresenta o exemplo do uso do vocábulo “bonito” como adjetivo adverbializado (Furtado da Cunha e Cezário, 2023). O esquema de uma construção utilizando essa palavra seria *V bonito*, sendo instanciado por sentenças como “falar bonito”, “escrever bonito”, “cantar bonito”. Em tese, qualquer advérbio pode modificar um verbo. Contudo, em pesquisa em corpora, as autoras não encontraram instanciações como “? dirigir bonito”, “? cozinhar bonito” e “? pintar bonito”. Em outras palavras, a construção é parcialmente produtiva, e nem todos os verbos são compatíveis.

A pesquisa em corpora apontou, por exemplo, para a inexistência da instanciação “? Pintar bonito”, ainda que a função de “exercer uma ação de maneira bonita” possa se aplicar à ação de pintar. Não há justificativa, no que diz respeito à função comunicativa, para essa restrição. Uma opção que veicula a mensagem pretendida poderia ser “ela fez uma pintura bonita”, uma forma menos eficiente, visto que é mais longa. Embora pareça cognitivamente contraintuitivo escolher uma forma menos eficiente, mais longa, não o é, pois

“Uma explicação mecanicista apoia nossa tendência em produzir formas que nossa comunidade considera aceitáveis, na medida em que formulações parcialmente familiares são mais fáceis de acessar do que formulações totalmente novas. Ou seja, formas mais convencionais são mais eficientes de serem acessadas na memória,

mesmo que sejam menos eficientes de serem produzidas.” (Goldberg, 2019, p.9, **tradução nossa**)¹²

Desse modo, a convencionalidade é muito importante para a GC e também para este estudo (ver capítulo 1). Ela é uma das grandes marcas das construções idiomáticas e um comportamento repetido de maneira sistemática toma proporções na representação linguística de um falante.

Além da adequação do uso da língua em conformidade com a comunidade de fala, aspectos de contexto são importantes no aprendizado de construções. Goldberg (2019) propõe que a maneira como se aprendem palavras serve para explicar a forma como se aprendem outras construções. Com a exposição a uma palavra ou outra construção, logo na primeira vez, traços de forma, função e dimensão contextual são registrados de maneira imperfeita, falha (*lossy*) na memória. O uso de uma palavra não está associado a aprender apenas a forma como se pronuncia e os traços semânticos referentes a ela, por exemplo, mas também a aspectos de contexto. O conhecimento de mundo influencia (e muito) e faz parte do conhecimento linguístico. O clássico caso da palavra “*bachelor*” já revela o poder do conhecimento comum compartilhado no aprendizado linguístico. Em uma definição por traços, a palavra, traduzida como “solteirão”, seria definida como “homem, adulto, não casado”. Entretanto, Fillmore (ver *Fillmore 1975*) aponta para o fato de que se enquadrariam nessa definição sujeitos como o Papa e o Tarzan. No entanto, os usuários não se refeririam ao Papa ou ao Tarzan como “solteirões”, eles sabem muito bem a que contexto a palavra se aplica. Isso ocorre, pois os significados evocam cenas convencionalizadas ou idealizadas, no caso de “*bachelor*”, um homem que passou da idade que se imagina ideal para se casar. Assim, nota-se que o contexto traz nuances de significado bastante específicas e que precisam ser capturadas pelos falantes. As palavras, bem como outras construções, estão associadas de maneira estruturada com ricas informações de forma, função e dimensões contextuais, embasadas nas experiências e no conhecimento de mundo do falante (Goldberg, 2019).

Como enunciado, o primeiro contato com uma construção já pode deixar alguns traços na memória. Essa memória é vasta, rica e estruturada, exemplo disso é a capacidade de utilizar apropriadamente a palavra “*bachelor*”. No entanto, como apresentado na definição mais recente de construção, a memória é imperfeita, falha, e, sendo assim, muitos aspectos não são amplamente registrados. Há informações que chamam mais atenção do que outras:

¹² No original: “A mechanistic explanation supports our tendency to produce forms that our community deems acceptable, insofar as partially familiar formulations are easier to access than wholly novel formulations. That is, more conventional forms are more efficient to access from memory even if they are less eficiente to produce” (Goldberg, 2019, p.9)

“(...) aspectos surpreendentes ou incomuns dos contextos têm maior probabilidade de serem codificados na memória do que aspectos que estão sempre presentes. Aspectos de contextos e palavras que são mais altamente correlacionados (têm maior informação mútua) também são mais provavelmente serão retidos.” (Goldberg, 2019, p. 15, **tradução nossa**)¹³

Pensando no modo como nossa memória retém as informações, é possível refletir sobre as construções idiomáticas. Alguém que nunca tenha testemunhado expressões como “pagar o pato” ou “meter o pau” pode se atentar, de maneira especial, a elas, visto que trazem elementos surpreendentes – uma expressão fora do contexto literal e composicional – com a forma convencional transitiva (SVO). Em um estudo com falantes nativos do português brasileiro de Sobral/CE, ao serem perguntados sobre o significado de expressões idiomáticas relacionadas à violência (como exemplo: “meter o pau (em)”, “quebrar o pau”), os sujeitos respondiam fazendo referência a memórias episódicas. Em outras palavras, eles eram capazes de evocar eventos da própria vida em que testemunharam a expressão (Martins e Monteiro-Plantin, 2012). Esse comportamento pode ser sugestivo de que expressões idiomáticas são salientes para a atenção de quem as testemunha.

Em resumo, os recursos de habilidades cognitivas do domínio cognitivo geral, em especial a memória, bem como o uso, incluindo as dimensões contextuais nele apresentadas, são elementos fundamentais para o aprendizado das construções. O *input* e a frequência também interagem com esses elementos produzindo efeitos linguísticos, que, por consequência, também afetam o uso.

2.2.2 Convencionalidade, entrincheiramento e *chunking*

Uma estrutura é convencional na língua quando é consolidada pelo uso e pela prática por determinada comunidade. Essa informação não implica em uma alta frequência necessariamente. As expressões idiomáticas são estruturas convencionais, pois estão consolidadas no uso e estão associadas a certas situações comunicativas, a certos contextos específicos na prática comunicativa. Ainda, na forma há uma certa estabilidade, que se dá em *continuum* de estruturas menos esquemáticas e substantivas como “bater as botas”, até mais esquemáticas e abstratas como “Dar uma de *X*” – “dar uma de João sem braço”, “dar uma de vítima” – ou “Quanto maior *X*, maior *Y*” – “Quanto maior a subida, maior a queda”, “Quanto maior a altura, maior o tombo”.

¹³ No original: “(...) surprising or unusual aspects of contexts are more likely to be encoded in memory than aspects that are ever present. Aspects of contexts and words that are more highly correlated (have higher mutual information) are also more likely to be retained.” (Goldberg, 2019, p.15)

No entanto, elas não são tão frequentes como outras estruturas regulares na língua. O que poderia explicar, então, sua permanência no uso? Como apontado anteriormente, aspectos incomuns ou chamativos têm maior chance de serem codificados na memória (Goldberg, 2019). Assim, é possível hipotetizar que as EI's podem ser construções que chamam a atenção do aprendiz por serem estruturas mais inusitadas, facilitando o aprendizado. Mas, ainda, há outro fator importante que colabora para que o conjunto dessas expressões seja tão presente na linguagem: a associação de construções a mensagens no contexto (*message-in-context*). As construções estão associadas a informações contextuais e funcionais bastante ricas. Um exemplo desse fenômeno é dado por Goldberg (2019) com palavras como “sofá” e “divã”, ainda que fosse possível, em hipótese, referir-se ao estofado de um analista/terapeuta como sofá, a ideia pretendida não estaria sendo comunicada da melhor maneira. Muitos usuários de uma comunidade linguística compartilham do conhecimento comum de que um sofá de uma casa ou até de uma sala de espera de um consultório tem uma função e evoca uma cena muito diferente de um divã. O mesmo pode ser observado para as expressões idiomáticas. Ainda que se tenha formas mais frequentes de expressar mensagens semelhantes, dizer “Fulano morreu” ou “Fulano bateu as botas” pode ter intenções e contextos muito diferentes, ainda que falem sobre o mesmo evento: morte. Casos de sinonímia completa/perfeita são bastante raros, ainda que existam construções que referenciam o mesmo evento, como no caso apresentado; elas podem estar associadas a contextos comunicativos específicos, registros linguísticos, ou dialetos (Goldberg 1995, 2006, 2019). Em resumo, a permanência das construções idiomáticas revela a associação delas à mensagem, no contexto específico que elas veiculam.

Essa relação com a mensagem no contexto refere-se ao fato de a memória ser acessível por conteúdo. Quando se deseja comunicar alguma mensagem no contexto específico, é feito um movimento de lembrar a construção adequada àquele fim. A informação contextual de uma construção que é memorizada é bastante rica, como mencionado. Nesse sentido, há uma associação da mensagem no contexto a uma construção e, ao produzir, o falante procura essa mensagem e a construção a ela associada.

“(…) nós não proferimos simplesmente quaisquer palavras e construções que estejam mais acessíveis. Como nossa memória é associativa e acessível por conteúdo, nosso conhecimento da linguagem relaciona propriedades formais, propriedades funcionais e aspectos de contexto, e acessamos o conhecimento que serve ao nosso propósito pretendido.” (Goldberg, 2019, p. 60, **tradução nossa**)¹⁴

¹⁴ No original: “(…) we do not simply utter whichever words and constructions happen to be the most accessible. Since our memory is associative and content addressable, our knowledge of language relates formal properties, functional properties, and aspects of context, and we access the knowledge that serves our intended purpose.” (Goldberg, 2019, p. 60)

O atributo associativo da memória faz com que as mensagens pretendidas no contexto fiquem muito associadas a determinadas construções idiomáticas. Nesse sentido, há uma repetição de uma construção idiomática em um contexto específico. O uso e a repetição levam a uma automatização, assim as construções idiomáticas ficam entrincheiradas na memória do falante que as testemunha, ou seja, bastante enraizadas na mente. Mas, assim como afirma Langacker (2008), o entrincheiramento se dá de maneira progressiva, a estrutura se torna mais entrincheirada à medida que o falante a usa e, pelo mesmo motivo, é possível esquecê-la devido a falta do uso. Além disso, o autor propõe que as construções se encontrariam em um *continuum* também sob o aspecto do entrincheiramento, mais ou menos entrincheiradas, a partir da frequência do uso daquele exemplar.

Outro fenômeno que também se relaciona à convencionalidade das construções idiomáticas é o processo de *chunking*. Bybee (2016) reforça que o uso repetido fortalece a representação na memória e, ainda, a formação de um *chunk*, ou, bloco de informação, independentemente de serem partes composicionais ou não. A autora ressalta que o processo de *chunking* (agrupamento) é um processo cognitivo geral, um processo de organização da memória. Costuma-se agrupar as informações de modo a agrupar unidades menores em um todo significativo. Um exemplo é a forma como seccionamos o número do CPF para facilitar a memorização, geralmente separando em três partes com três dígitos cada uma e depois uma com dois dígitos (111.222.333-44, por exemplo). Com a língua não seria diferente, com a exposição a sequências de modo frequente, convencional: com a repetição, aquela estrutura se torna automatizada e, assim, vira um *chunk*. Isso acontece com expressões fixas e convencionais da língua, como “como você (es)tá?” e também com as construções idiomáticas, por exemplo, no caso de “bater as botas”, ao invés de se acessar “bater + as + botas”, em hipótese, acessa-se todo o bloco “bater as botas”, como uma única unidade.

Um estudo sobre a relação entre *chunking* e frequência é muito interessante para explicar o caso das construções idiomáticas. Se a frequência de *token* da expressão “bater as botas” for checada, os dados do “Corpus do Português: Now” apontam 87 ocorrências, incluindo as variações de flexão verbal. Por outro lado, com outra expressão fixa, como “X que vem” (Fulgêncio, 2008), tem uma diferença bem grande. Contando apenas a instanciação “ano que vem”¹⁵, a frequência de token é de 37442. Com base nesses números, poderia ser argumentado que “bater as botas” possui uma frequência suficiente, o que seria um argumento desfavorável à ideia de entrincheiramento e de *chunking* desse tipo de construção. No entanto, Schneider

¹⁵ Outras instanciações comuns de “X que vem” são: semana que vem, mês que vem.

(2016) realizou um estudo que mostra que a frequência de co-ocorrência e probabilidade de co-ocorrência também colaboram para o efeito *chunking* das estruturas. No estudo, a maior probabilidade de co-ocorrência entre estruturas multipalavras, como binômios, em relação a outras com menor probabilidade, diminui o efeito de hesitação nos *chunks*. Ou seja, em uma sentença proferida por um falante, há uma maior tendência de que uma hesitação/fragmentação na fala aconteça em outro ponto da frase que não em uma expressão como “cara a cara”, por exemplo. Essa é uma evidência de que a expressão está em *chunking*, equivalendo a um bloco unívoco de sentido na mente dos falantes.

Desse modo, com a associação das EI’s a determinada situação comunicativa, cria-se uma relação na mente do falante. Ainda que buscas em *corpora* apontem frequências mais baixas, por serem construções especializadas no contexto, é possível compreender que elas estão em *chunking* na memória, visto que “[...] são convencionais no sentido de que foram estabelecidos por meio de repetição no uso, mas não precisam ser de alta frequência.” (Bybee, 2016, p.6). Em resumo, as construções idiomáticas são construções da língua em que há uma associação entre sua forma em bloco com sua mensagem no contexto. Além disso, assim como para as demais construções da língua, as EI’s estão sob o efeito de processos como entrincheiramento e *chunking*, a despeito da frequência observada.

2.3 A metáfora como processo cognitivo

É bastante vigente, ainda, no senso comum, a ideia de que metáfora é um recurso poético. Gilberto Gil dedicou uma canção ao recurso que elucida a posição: “Uma lata existe para conter algo/ Mas quando o poeta diz: Lata/ Pode estar querendo dizer o incontível”¹⁶. Trata-se de uma visão aristotélica que restringia a metáfora à Arte Poética¹⁷. No entanto, ela está presente na linguagem cotidiana e, com o avanço dos estudos linguísticos, passou a ser entendida como um processo cognitivo, como apresentado nesta seção.

2.3.1 Teoria da Metáfora Conceptual (TMC)

Com Lakoff e Johnson, em 1980, o estudo da metáfora tomou contornos para além de aspectos linguísticos e estilísticos. Os autores adotaram uma visão experientialista, na qual “(...) a maior parte do nosso sistema conceptual é de natureza metafórica. (Lakoff e Johnson, 2022, p. 46). O caráter experientialista é postulado, pois é a partir da experiência cotidiana, de

¹⁶ Trecho da canção Metáfora, de Gilberto Gil, lançada em 1982. Letra disponível em: <https://www.lettras.mus.br/gilberto-gil/487564/>

¹⁷ “Arte poética” é uma compilação de anotações do filósofo grego Aristóteles, entre os anos 335 a.C e 323 a.C, com um tratado sobre poesia e arte.

base sensório-motora, que se constrói a metáfora. Já o sistema conceptual é metafórico, pois, segundo os autores, a forma como se pensa, se percebe e se age e os conceitos que governam o pensamento são estruturados a partir da forma como os indivíduos se relacionam e se comportam no mundo. Assim sendo, “a essência da metáfora é compreender e experienciar uma coisa em termos de outra” (Lakoff e Johnson, 2002, p.47-48). Desse modo, um conceito metafórico ocorre quando acontece uma transferência entre domínios. O domínio-fonte, com propriedades mais físicas e que envolve as experiências de base sensório-motoras, estrutura conceptualmente um domínio-alvo, mais abstrato. Forma-se então um conceito metafórico a partir desse mapeamento: DOMÍNIO-ALVO É DOMÍNIO-FONTE, como em TEMPO É DINHEIRO, sendo “dinheiro” o domínio-fonte e “tempo” o domínio-alvo.

“Os dois domínios que participam da metáfora conceptual têm nomes especiais. O domínio conceptual do qual extraímos expressões metafóricas para entender outro domínio conceptual é chamado de domínio-fonte, enquanto o domínio conceptual que é compreendido dessa maneira é o domínio-alvo. Assim, vida, argumentos, amor, teoria, ideias, organizações sociais e outros são domínios-alvo, enquanto viagens, guerra, edifícios, comida, plantas e outros são domínios-fonte. O domínio-alvo é o domínio que tentamos compreender por meio do uso do domínio-fonte.” (Kövecses, 2002, p. 4, **tradução nossa**)¹⁸

Segundo a Teoria da Metáfora Conceptual (TMC), as expressões linguísticas metafóricas são resultado de um processo cognitivo que conceptualiza o mundo por meio de metáforas, ou seja, mapeamentos entre domínios. As metáforas conceptuais se dividem em três tipos: estruturais, orientacionais e ontológicas. Essas são metáforas cujos mapeamentos são convencionalizados, bastante comuns na língua em uso, e que compõem partes dos itens experimentais do teste de leitura autocadenciada aplicado neste estudo (ver capítulo 3).

Metáforas estruturais

Uma metáfora estrutural ocorre quando um conceito alvo complexo é estruturado metaforicamente em termos de outro. Por exemplo, alguns serviços são cobrados ou taxados em relação ao tempo, por exemplo, uma tarifa telefônica, uma recreação para crianças em um shopping, taxas de juros, etc. Há um mapeamento de um domínio-alvo complexo em termos de um domínio-fonte também complexo. Assim, na experiência do cotidiano, o tempo, assim como o dinheiro, é conceptualizado como um recurso limitado e valioso e, assim sendo, pode ser

¹⁸ No original: “The two domains that participate in conceptual metaphor have special names. The conceptual domain from which we draw metaphorical expressions to understand another conceptual domain is called source domain, while the conceptual domain that is understood this way is the target domain. Thus, life, arguments, love, theory, ideas, social organizations, and others are target domains, while journeys, war, buildings, food, plants, and others are source domains. The target domain is the domain that we try to understand through the use of the source domain.” (Kövecses, 2002, p. 4)

“gasto”, “investido”, “desperdiçado”, “poupado”, etc. Nos exemplos a seguir, há um mapeamento do conceito metafórico TEMPO É DINHEIRO.

(31) “Ele abriu os meus olhos para quanto tempo gastei em e-mails ou posts em redes sociais.” (Corpus do Português: NOW)

(32) “(...) é o que costuma acontecer quando fazemos um upgrade de um software: temos de ponderar se as novas funcionalidades oferecidas justificam o tempo que teremos de investir na sua aprendizagem.” (Corpus do Português: NOW)

O conceito de tempo, mais abstrato, é mapeado de maneira convencionalizada a partir do conceito de “dinheiro”. Uma busca rápida em um corpus revela dados como (31) e (32), salientando que é costume conceptualizar, assim como Lakoff e Johnson (2002) afirmam, o tempo como um recurso valioso e limitado. Essa é uma percepção pautada na experiência que se tem dentro uma cultura em que as atividades cotidianas básicas são exercidas em função do tempo que levam para serem cumpridas e do valor financeiro atribuído a elas.

Metáforas orientacionais

Nas metáforas orientacionais, há uma organização de uma série de conceitos em termos de outro, a partir de noções espaciais: para baixo, para cima, para frente, para trás, etc. Esse tipo de metáfora está relacionado à ideia de cognição corporificada: é a associação entre o corpo físico, ambiente e cultura que estabelece o mapeamento de domínios. Um exemplo é TRISTE É PARA BAIXO e FELIZ É PARA CIMA, que tem como base a experiência corporificada de que pessoas tristes possuem uma postura mais encurvada, caída, ao passo que pessoas mais alegres estão mais eretas. Essas metáforas conceptuais podem ser atestadas em expressões linguísticas como nos exemplos a seguir.

(33) “Antes pensava que estávamos no fundo do poço, agora tenho certeza que chegamos na fossa. Estamos sem gestão e sem justiça. (...)” (Corpus do Português: Now)

(34) “Fizemos um contato em Boston, nos Estados Unidos, onde o maestro é uma das estrelas do momento, e nada melhor do que começar a coluna de 2013 em alto astral.” (Corpus do Português: Now)

Os exemplos (33) e (34) apresentam, respectivamente, as metáforas conceptuais TRISTE É PARA ABAIXO, representada pelos trechos “fundo do poço” e “na fossa”, e FELIZ É PARA CIMA, representa pela expressão “alto astral”. Os conceitos mapeados nesse tipo de metáfora conceptual costumam gerir a forma de agir, pensar e estar no mundo: de maneira geral,

avalia-se positivamente o que está em cima/para cima e negativamente o que está para baixo/em baixo (Lakoff e Johnson, 2002; Kövecses, 2002). Essa forma de conceptualizar a experiência se reflete em outras metáforas como STATUS SUPERIOR É PARA CIMA (Ele está no topo da carreira) e STATUS INFERIOR É PARA BAIXO (Ele foi rebaixado de cargo).

Metáforas ontológicas

As metáforas ontológicas são uma forma de experienciar e conceptualizar o mundo a partir do domínio-fonte “substâncias” e “recipientes”. É uma forma de lidar com conceitos-alvo mais abstratos de maneira menos abstrata. Segundo Lakoff e Johnson (2002), as metáforas ontológicas permitem compreender as experiências em termos de objetos e substâncias para que possamos tratá-las como entidades discretas ou de espécie uniforme. Assim, é possível “(...) referir-nos a elas, categorizá-las, agrupá-las e quantificá-las - e, dessa forma, raciocinar sobre elas.” (Lakoff e Johnson, 2002, p. 76). Essa é uma metáfora conceptual que parte da relação com objetos físicos, em especial com o próprio corpo, para conceber outras experiências. De acordo com os autores, ela serve para diversos fins, como: referir-se, quantificar, identificar aspectos, identificar causas, traçar objetivos e motivar ações.

Além disso, Kövecses (2002) aponta para o fato de que, ao conceber essas experiências em termos de objetos e substâncias, não se especifica quais objetos e substâncias são. Os domínios-fonte são de categorias bastante gerais e que não fornecem uma grande explicação sobre o domínio-alvo, visto que esse é o papel das metáforas estruturais. O papel da metáfora ontológica é

“(...) atribuir um status básico em termos de objetos, substâncias, etc., a muitas de nossas experiências. Os tipos de experiências que mais requerem isso são aquelas que não são claramente delineadas, vagas ou abstratas. Por exemplo, não sabemos realmente o que é a mente, mas a concebemos como um objeto (note o uso da palavra "o que" na primeira parte desta frase). Dessa forma, podemos tentar entender mais sobre ela.” (Kövecses, 2002, p. 34, **tradução nossa**).¹⁹

Um exemplo desse tipo de conceptualização está no conceito de mente, a partir do mapeamento MENTE É UMA ENTIDADE, desenvolvida como MENTE É UMA MÁQUINA e MENTE É UM OBJETO QUEBRADIÇO.

¹⁹ No original: “(...) to assign a basic status in terms of objects, substances, etc. to many of our experiences. The kinds of experiences that require this the most are those that are not clearly delineated, vague, or abstract. For example, we do not really know what the mind is, but we conceive of it as an object (note the use of the word what in the first part of this sentence). This way we can attempt to understand more about it.” (Kövecses, 2002, p. 34)

(35) "É uma maneira que eu tenho de me manter ativo. Não quero ficar enferrujado, nem física nem mentalmente (...)" (Corpus do Português: NOW)

(36) "(...) ele é um homem forte que não se encaixa nos formatos convencionais. A minha história é a versão oposta, que ele se desmoronou, que não consegue funcionar e que isto não pode continuar." (Corpus do Português: NOW)

Nos exemplos, é possível perceber a realização das metáforas ontológicas sobre a mente. Em (35), utiliza-se do termo "ficar enferrujado" para se referir ao desejo de manter a mente (e também o corpo) ativa, mapeando o conceito MENTE É UMA MÁQUINA. Em (36), duas metáforas se realizam. Ao utilizar "se desmoronou" para se referir a um estado da mente, é mapeado o conceito metafórico MENTE É OBJETO QUEBRADIÇO e em "não consegue funcionar", MENTE É UMA MÁQUINA novamente aparece. Os exemplos revelam que, assim como afirmam Lakoff e Johnson (2002), a metáfora ontológica é utilizada como tentativa de lidar racionalmente com as experiências mais abstratas.

2.3.2 Metáforas novas

As metáforas novas são aquelas que não se formam a partir de estruturas linguísticas convencionalizadas. No entanto, há alguns posicionamentos divergentes a respeito da existência de metáforas, de fato, novas. Neste estudo, assume-se que os mapeamentos levantados por Lakoff e Johnson (2002) para as metáforas conceptuais são realizados por construções frequentes e convencionalizadas na língua. Os exemplos apresentados na seção anterior foram retirados do *corpus* e se mostram exemplares significativos da língua em uso. Apesar de em pouca quantidade, visto que esta não é uma pesquisa com foco em estudos de *corpora*, eles cumprem a função de demonstrar a convencionalidade das expressões metafóricas.

Segundo Lakoff e Johnson (2002) as metáforas novas se distinguem das metáforas convencionais por não fazerem parte do sistema conceptual ordinário da nossa cultura. Como exemplo citam o mapeamento AMOR É UMA OBRA DE ARTE COLABORATIVA. A partir da visão do que seria "obra" e "arte", é construído um significado que pode caber implicações como: "o amor é trabalho", "o amor exige dedicação", "o amor exige cooperação", de ordem mais literal. Além disso, algumas mais metafóricas também são possíveis como "o amor é a expressão mais profunda do indivíduo", "o amor é uma experiência estética". Nessa construção de significado, alguns aspectos sobre amor serão salientados e outros suprimidos - por exemplo,

ideias de um amor passivo são suprimidas, pois não são coerentes com a ideia de colaboração expressa no conceito metafórico. Além disso, Lakoff e Johnson (2002) salientam que a metáfora evoca aspectos específicos de outros conceitos não diretamente ligados à estrutura mapeada. Por exemplo, o uso de “obra” liga-se ao conceito de “trabalho”, não qualquer trabalho, mas um trabalho que evoque o uso da criatividade e do trabalho em conjunto. Os autores propõem, ainda, que o sentido da metáfora pode ser diferente para cada um, pois depende de fatores culturais e compatibilização com experiências passadas. Em resumo, para Lakoff e Johnson (2002), a metáfora nova é capaz de estabelecer uma nova forma de compreensão da experiência e depende das percepções individuais para compreensão, pois, o conceito de “trabalho”, por exemplo, pode variar de indivíduo para indivíduo.

No entanto, Lakoff e Johnson mencionam que metáforas inteiramente novas são bastante raras. As realizações novas, na verdade, evocam padrões convencionais. Para explicar a ideia, os autores utilizam dois excertos:

(37) “Conte-me a história de sua vida.” (Lakoff e Johnson, 2002 p. 278)

(38) “A vida é uma fábula contada por um idiota, repleta de som e fúria, sem nenhum sentido.” (Lakoff e Johnson, 2002 p. 278)

No exemplo (37), é realizada metáfora convencional A VIDA É UMA HISTÓRIA, já em (38) é atestada a metáfora não convencional A VIDA É UMA FÁBULA CONTADA POR UM IDIOTA. Conceptualizar a vida como uma história compreende entendê-la como uma narrativa, com participantes, partes, etapas, sequenciação linear, causalidade, finalidade, plano, meta. Ou seja, há uma estrutura, uma *gestalt* experiencial, subjacente ao domínio-fonte. Ao conceptualizar a vida como uma fábula contada por um idiota, o usuário deve, primeiro, partir do entendimento da *gestalt* experiencial subjacente do domínio-fonte “história”, visto que “fábula” também possui uma estrutura narrativa. No entanto, essa estrutura é subvertida no trecho “contada como por um idiota”. A subversão acontece, pois, na *gestalt* experiencial de uma narrativa, se espera um narrador coerente, o que não se confirma ao dizer que ela é narrada por um idiota (Lakoff e Johnson, 2002; Lakoff, 2008). Nesse sentido, apesar da subversão, parte-se de um mapeamento já estabelecido em A VIDA É UMA HISTÓRIA para se construir a metáfora A VIDA É UMA FÁBULA CONTADA POR UM IDIOTA. Em resumo, é possível afirmar, segundo os autores, que as metáforas novas tendem a ter mapeamentos metafóricos convencionais subjacentes, que, geralmente, são mapeamentos de ordem mais complexa e abstrata, ou seja, evocam metáforas estruturais.

Essa constatação, porém, não resulta em dizer que as metáforas novas são exatamente equivalentes às outras metáforas convencionais. Elas partem de um mapeamento convencionalizado, mas utilizam alguns recursos para criar imagens não convencionais. Kövecses (2002) ressalta que, nas metáforas novas, há a utilização de estratégias como ampliação, elaboração, questionamento e combinação de conceitos metafóricos convencionais. A ampliação é o uso de novos elementos conceituais no domínio-fonte, por meio de novos elementos linguísticos. O autor exemplifica com um poema de Robert Frost, do qual a metáfora A VIDA É UMA VIAGEM emerge e é ampliada:

(39) “Duas estradas divergiam numa floresta, e eu —
Eu escolhi a menos percorrida.
E isso fez toda a diferença.”²⁰

No fragmento em (39), a o conceito de vida em termos de viagem aparece com o uso de “Duas estradas”, o mapeamento A VIDA É UMA VIAGEM é, no entanto, ampliado, pois a ideia de duas estradas diferentes, dois caminhos, podendo ser escolhidos um ou outro é nova. Há uma conceptualização não usual do domínio-fonte, um acréscimo de um elemento, por meio de uma linguagem não convencional para descrevê-lo. Já no recurso da elaboração, um aspecto do domínio-fonte é trabalhado de maneira pouco usual, como no exemplo abaixo:

(40) “acetileno branco
irradia do meu corpo
liberado sem esforço
perfeitamente direcionado
para o verdadeiro inimigo

rasgando seu corpo até a essência
da existência
queimando sua mentira
deixando-o em um novo
mundo; um homem
transformado.”²¹

²⁰ No original: “Two roads diverged in a wood, and I—
I took the one less travelled by.
And that has made all the difference.”

²¹ No original: “white acetylene
ripples from my body

O poema em (40) apresenta o mapeamento metafórico RAIVA É FLUIDO QUENTE, normalmente esse conceito é mapeado de modo que o domínio-alvo seja o sangue ou o portador da emoção como um contêiner, ressaltando o atributo “quente” (“o sangue ferveu” ou “ela está fervendo de raiva”). No poema, a metáfora é realizada de maneira elaborada com o uso específico de uma substância perigosa “acetileno” na conceptualização do domínio-fonte (Kövecses, 2002). Kövecses (2002) afirma ainda que o entendimento do poema advém do conhecimento do mapeamento metafórico convencional RAIVA É FLUIDO QUENTE, que se dá, provavelmente, de maneira inconsciente. Outro recurso apresentado pelo autor é o questionamento. Nele, a adequação ou validade da metáfora convencional é questionada, ainda que o mapeamento permaneça. Como exemplo, Kövecses (2002) analisa o trecho do poema de Catullus:

(41) “Sóis podem se pôr e retornar novamente,
mas quando nossa breve luz se apaga,
há uma noite perpétua a ser dormida.”

O exemplo (41) apresenta as metáforas convencionais VIDA É DIA e MORTE É NOITE, nessa conceptualização os domínios-fonte “dia” e “noite” são entidades que se sucedem no tempo: o dia dá lugar para a noite, a noite para o dia. No poema, Catulo questiona a metáfora da vida como dia e da morte como noite, sugerindo que a morte não é simplesmente uma passagem para outra fase, mas sim uma noite eterna e definitiva. Há um questionamento do poeta sobre a validade da relação de morte com a noite, pois o domínio-fonte “noite”, na realidade cotidiana, não é perpétuo, há um ciclo entre dias e noites, e é sobre a validade dessa verdade que o poeta faz o questionamento (Kövecses, 2002). A combinação, por fim, como o próprio nome indica, é um recurso em que vários mapeamentos podem aparecer de uma só vez. Como exemplo, Kövecses (2002) utiliza um soneto de Shakespeare.

(42) Em mim, vê o crepúsculo deste dia
Após o naufrágio do sol a Oeste,

effortlessly released
perfectly trained
on the true enemy
raking his body down to the thread
of existence
burning away his lie
leaving him in a new
world; a changed
man.”

Que, pouco a pouco, a noite escura afasta,
A outra face da Morte, que tudo silencia.

Há no quarteto algumas metáforas convencionais: EVENTOS SÃO AÇÕES, A VIDA É UMA POSSE PRECIOSA, MORTE É NOITE e UMA VIDA É LUZ. Por exemplo, no verso “que, pouco a pouco, a noite escura afasta”, é evocado o conceito de que “a vida é um bem precioso”, já que está sujeita a se findar. Há ainda a ideia de que “noite” está retirando a claridade, evocando os mapeamentos MORTE É NOITE e VIDA É LUZ, ao apresentar a ideia de que o eu lírico está morrendo. Apenas nesse verso, foram utilizados alguns conceitos metafóricos convencionais para criar uma metáfora nova. Os exemplos citados, ainda que de fontes literárias, podem ser estendidos a usos ordinários da linguagem. Essa extensão pode ocorrer, pois, na visão de Lakoff e Johnson (2002) e de Kövecses (2002), as metáforas novas são criadas a partir de recursos cognitivos e formas de experienciar e compreender o mundo que são convencionais.

Há, no entanto, posicionamentos teóricos que corroboram uma diferença menos discreta entre metáforas convencionais e metáforas novas. Bowdle e Gentner (2005) propõem que os conceitos metafóricos e seus mapeamentos seguem um caminho evolutivo na língua. A hipótese estabelece que esse caminho se dá em virtude da forma como os mapeamentos são estabelecidos entre os domínios. A chamada Teoria da Carreira da Metáfora propõe que em um primeiro momento as metáforas são mapeadas como comparação e depois são categorizadas. Isso implica dizer que uma metáfora, em seu primeiro uso, vai ser mapeada como uma comparação e à medida que há uma familiarização, uma adoção por uma comunidade linguística, o mapeamento passa a ser via categorização. Em outras palavras, as metáforas novas são processadas, em hipótese, como comparações e as convencionalizadas como categorias.

Desse modo, no caso das metáforas novas, não há uma associação padrão e convencional de uma expressão a uma categoria metafórica. Sendo assim, o domínio-alvo e domínio-fonte são, em hipótese, comparados em sua forma literal, para em seguida ocorrer o alinhamento dos domínios de maneira abstrata. Por outro lado, no caso das metáforas convencionais, há um mapeamento convencionalizado e, por isso, elas são acessadas como uma categoria, a partir da recuperação de significados armazenados (Gentner e Bowdle, 2008). Nesse sentido, “(...) a hipótese da Carreira da Metáfora afirma que o alinhamento entre os conceitos literais alvo e base de uma metáfora pode levar à indução e, eventualmente, à lexicalização de esquemas relacionais de domínio geral, que podem atuar como categorias

metafóricas.” (Bowdle e Gentner, 2005, p.199, **tradução nossa**)²². Isso posto, nesta perspectiva, há uma distinção entre metáfora nova e convencional, em especial em relação ao tempo em que o mapeamento se faz presente em uma língua e a sua ampla adoção por uma comunidade linguística na prática. Em resumo, é uma perspectiva que difere da Teoria da Metáfora Conceptual, visto que compreende a metáfora nova como sendo uma realização mais possível e, ainda, com diferenças de processamento em relação a metáforas convencionais.

Assim, para fins de comparação entre metáfora convencional e metáfora nova, adota-se a perspectiva de que elas são construções distintas. As metáforas convencionais partem de mapeamentos já bem delimitados teoricamente (Lakoff e Johnson, 2002), e, mais importante, com realizações bem representativas e convencionalizadas na língua. As metáforas novas, por sua vez, estabelecem relações entre domínio-fonte e alvo que não são convencionalizadas, nem como forma, nem como mapeamento metafórico, necessitando, portanto, de uma comparação entre domínios. Sendo assim, é possível prever que as expressões metafóricas, em especial as novas, necessitam de um usuário da língua ativo, pois o processamento requer, em ambos os tipos, a construção do sentido no momento da interação.

2.4 Processamento linguístico e acesso lexical

O processamento linguístico pode ser definido como os mecanismos envolvidos na compreensão e na produção de construções da língua. Nesse sentido, ele não é passível de observação direta e, por isso, deve ser estudado por meio de um comportamento (Pasquali, 2015). E é a partir desses comportamentos que, por meio de experimentos, hipóteses sobre o processamento são levantadas. No caso deste estudo, o comportamento será a leitura autocadenciada com a variável de resposta tempo de reação. Desse modo, o custo processual das expressões idiomáticas e das metafóricas, convencionais e novas, podem ser medidos por meio dessa variável. Como hipótese, tem-se que as metáforas novas possuem um custo processual maior do que as expressões idiomáticas, sendo as EI's recuperadas em blocos. A partir do que este estudo hipotetiza, as metáforas novas são processadas e as EI's acessadas, de modo semelhante ao acesso lexical. Cabe, portanto, ressaltar os mecanismos envolvidos no processamento linguístico e no acesso lexical.

Bybee (2016) aponta para mecanismos gerais que constituem parte do processamento linguístico, quais sejam: *chunking*, analogia, similaridade. O *chunking* faz com que sintagmas

²² No original: “(...) the career of metaphor hypothesis states that aligning the literal target and base concepts of a metaphor can lead to the induction and eventual lexicalization of domain-general relational schemas, which can act as metaphoric categories” (Bowdle e Gentner, 2005, p.199)

possam ser analisados como unidades simples. Além disso, o efeito de *chunk*, um recurso de organização geral de nossa memória, colabora para a formação de unidades com muitas palavras, como expressões fixas, idiomáticas e construções, bem como para a recuperação delas como uma unidade. Ainda, em relação ao *chunk*, Bybee (2016) aponta para a tendência a uma perda de composicionalidade e analisabilidade. Ao serem repetidas na mesma forma (contexto e função, idem), estruturas em *chunk* podem perder sua composicionalidade, ou seja, seu significado como função das partes. Podem perder também sua analisabilidade, a capacidade de se identificar as palavras que compõem uma expressão. Entendendo que há uma rede de associações entre construções, dado que a memória é associativa, a perda desses componentes gera efeitos no grau de conexão com essa rede, ou seja, impacta o grau de autonomia de uma determinada expressão e vice versa, impactando o processamento. É importante salientar, no entanto, que medidas de composicionalidade, analisabilidade, bem como de *chunk* estão em *continuum*, não se trata de uma categoria estrita. Assim, é possível apontar o *chunking* como um efeito facilitador do processamento linguístico, visto que há uma alta convencionalidade de expressões em bloco. Por outro lado, estruturas não convencionais, como o caso das metáforas novas, não estão em *chunking* e, assim, podem ter o processamento afetado pelo efeito “surpresa” (Levy, 2007; Smith e Levy, 2013)²³. O conceito de efeito surpresa está relacionado a quão inesperada ou surpreendente uma palavra ou expressão é em um determinado contexto linguístico. Se uma expressão for altamente previsível com base no contexto linguístico em que ocorre, seu valor de surpresa será baixo, indicando que é uma estrutura mais convencionalizada, com maior facilidade de processamento. Por outro lado, expressões que são inesperadas ou improváveis em um determinado contexto terão valores de surpresa mais altos, refletindo a dificuldade adicional de processamento devido a sua natureza não convencional.

Bybee (2016) também cita como fundamentais para o processamento os mecanismos de analogia e similaridade. A analogia é a capacidade de usar um novo item em uma construção. Essa capacidade faz com seja possível fazer associações, em grande parte com base na semelhança semântica ou fonológica. Como exemplo, a autora cita a *prefab*²⁴ “*drives me crazy*” (me deixa louco). Por analogia, poderia ser produzido ou melhor interpretado uma instanciação de “drive me X”, na qual X seja uma locução adjetiva ou adjetivo com valor próximo a “loucura”, do que um exemplar como “drives me happy” (me deixa feliz). Além disso, há sensibilidade à similaridade com enunciados prévios, em especial a sequências frequentes e

²³ Em inglês “*effect of surprisal*”.

²⁴ Neste capítulo, utiliza-se a expressão no inglês *prefabs*, seguindo a tradução de Furtado da Cunha da obra de Bybee (2016). No capítulo 4, passa-se a utilizar a tradução em português: *pré-fabricadas*.

convencionais. Esses mecanismos podem promover um efeito chamado de pré-ativação (*priming*), facilitação de processamento de alguma estrutura a partir da exposição anterior a outra similar, seja sintática, fonológica ou semanticamente. Como exemplo, o experimento de Savage e colegas (2003) demonstrou que crianças de 6 anos tendiam a utilizar sentenças passivas ou ativas na descrição de uma imagem a depender da estrutura da descrição de outra gravura (passiva ou ativa) feita anteriormente pelo pesquisador.

Em resumo, todos esses mecanismos interagem com outros processos cognitivos gerais, como a categorização e a memória, não só como base para o aprendizado linguístico, mas também para o processamento, seja do ponto de vista da produção ou da compreensão. Na sequência, abordaremos a teoria de acesso lexical, a fim de explicar as vias de recuperação de um *chunk* linguístico.

O acesso lexical, por sua vez, diz respeito à habilidade de conseguir identificar qual é a palavra que se está ouvindo, lendo ou produzindo. Na teoria psicolinguística, há três “caminhos” para o acesso lexical, quais sejam: modelo de listagem plena, modelo de decomposição plena e modelo dupla rota. No modelo de listagem plena, todas as palavras são armazenadas de maneira integral, completa, não admitindo, por exemplo, apenas o armazenamento de um radical separado de morfemas flexionais, derivacionais, etc. Desse modo, são armazenados na memória, em exemplo, cada um dos itens lexicais da conjugação de um verbo como estudar: ESTUDO, ESTUDAS, ESTUDA, ESTUDAMOS, etc. Já no modelo de decomposição plena, as palavras são decompostas. Ao ouvir a forma verbal “estudamos”, seriam acessados o radical “stud” mais os morfemas necessários para a sua formação: “stud+a+mos”.

O modelo de dupla rota, por sua vez, admite duas vias: pela listagem plena e pela decomposição. Propostas, como as de Pinker e Prince (1982), sugerem que formas irregulares são armazenadas pela forma plena, ao passo que as regulares na forma decomposicional. Dessa forma, verbos irregulares como “ser” no português teriam suas formas armazenadas de maneira plena: sou, era, fui, por exemplo. Já verbos regulares como “estudar” seriam armazenados pelo radical e compostos como os morfemas necessários. Outro modelo de recuperação em dupla rota é o de Caramazza, Laudanna e Romani (1988 *apud* Medeiros e colegas, 2014). Nele, é proposto que a frequência é um fator que determina qual via: plena ou decomposicional, será a escolhida para o acesso lexical. Os itens lexicais frequentes seriam armazenados na forma plena e os infrequentes pelas regras de composição de afixos. O modelo dual de Schreuder e Baayen (1999), por outro lado, estabelece que itens opacos (como o caso das construções idiomáticas) seriam processados e armazenados via listagem plena, bem como os itens com ortografia

irregular, já os itens frequentes seriam acessados pelo processo decomposicional. Existem estudos que corroboram a ideia de uma dupla rota para o acesso e processamento de palavras no PB. A perspectiva teórica baseada no uso, no entanto, prevê que o falante não precisa escolher entre um acesso composicional ou “inteiriço”, visto que a rede de construções inclui todas essas relações (Bybee e Beckner, 2010).

Capítulo 3: Metodologia

Neste capítulo, será explicado em detalhes o paradigma experimental adotado, bem como as decisões tomadas em relação ao desenho experimental. Este estudo faz parte do projeto de pesquisa “O Papel de Fatores Funcionais no Processamento de Construções de Estrutura Argumental em português brasileiro”, aprovado junto ao Comitê de Ética Em Pesquisa (CEP/UFMG) com parecer Certificado de Apresentação de Apreciação Ética (CAAE) número 60698022.1.0000.5149.

3.1 Paradigma experimental: leitura autocandenciada

A leitura autocandenciada (LA), no inglês *self-paced reading*, é um método experimental que consiste na leitura de uma sentença por segmentos. Essa segmentação pode se dar por palavras ou por sintagmas, a depender do objeto de pesquisa em análise. Seu funcionamento consiste nos seguintes passos: o sujeito lê um segmento da sentença e deve apertar a seta do teclado do computador, ou clicar no botão direito do mouse, para passar para o próximo trecho, atividade que se repete até o fim da sentença. Importante salientar que, ao ativar o próximo o segmento, o anterior desaparece e que é o sujeito experimental quem controla/candencia seu processo de leitura.

Oliveira e colegas (2022) apontam para as vantagens em se utilizar o método experimental de LA. Dentre elas estão o seu pioneirismo em relação às técnicas *online* e a facilidade de se obter dados desse tipo, com o uso de apenas um computador e *softwares* normalmente gratuitos. Entretanto, como os autores apontam, a maior vantagem é que os dados obtidos por essa técnica são eficientes e os registros gerados são corroborados por outros testes com técnicas mais sensíveis como as de monitoramento ocular e de EEG (eletroencefalograma).

Esse método nos oferece uma medida que chamamos de *online*, ou seja, nos oferece dados sobre a reação dos sujeitos pesquisados no momento em que a atividade acontece. Ele fornece tempos de leitura (TR) em relação ao processo cognitivo envolvido, ou seja, no caso da LA, tempos de leitura de cada trecho/segmento de sentença e da sentença como um todo. Esse tempo de reação é importante, pois “(...) estudos seminais e vários outros conduzidos em seguida, evidenciam a relação existente entre tempo de leitura e custo de processamento.” (Oliveira e colegas, 2022, p. 41). O TR fornece dados, geralmente, bem precisos, correspondentes ao estímulo (trecho) e ao início da resposta subsequente do sujeito (ação de passar para próximo trecho ou responder à pergunta de compreensão) (Rayner e Clifton, 2002).

Desse modo, as sentenças atreladas a esse experimento devem ser contrastadas, pois é a partir das diferenças de tempo de reação entre as condições que se deseja analisar, além de sentenças distratoras, que será possível tecer algum tipo de inferência a respeito da hipótese experimental. O experimento deste trabalho com três níveis de condições, divididas com barras para indicar os trechos, como os exemplos a seguir:

(43) Lucas/ **amarrou um bode**/ do colega/ de trabalho.

(44) Pedro/ **gastou muito tempo**/ na reunião/ da manhã.

(45) Sara/ **segue arco-íris**/ na vida/desde cedo.

(46) Marcos/ **foi machucado**/ pelo animal/ ontem.

Os segmentos em negrito são correspondentes aos trechos-alvo. Eles representam o objeto que este trabalho pretende averiguar, sendo (43) expressão idiomática, (44) expressão metafórica convencionalizada e (45) expressão metafórica nova (construída no próprio discurso). Já a sentença (46) é a distratora, que não se configura como parte da análise, mas como constituinte do teste, para que o sujeito não identifique os tipos de sentença que estão sendo de fato avaliadas.

O trecho subsequente ao trecho-alvo, sublinhado, é chamado de *spillover*. Ainda que não seja o segmento que representa o objeto de análise, é possível que efeitos de tempo de resposta do trecho crítico (alvo) sejam encontrados no trecho seguinte (Oliveira e colegas, 2022). Desse modo, o trecho que representa o *spillover* também será considerado nas análises estatísticas.

3.2 Desenho experimental e critério de exclusão

O design experimental deste estudo consiste em um modelo 3x1, ou seja, uma variável independente, a saber, tipo de construção, considerada em 3 níveis, quais sejam: expressão idiomática, expressão metafórica convencional e expressão metafórica nova. Já a variável dependente ou de resposta consiste no tempo de reação (TR) das sentenças experimentais, com atenção aos trechos-alvo e aos trechos *spillover*. O TR será comparado intra-sujeitos e entre os níveis da variável independente em um teste realizado com falantes nativos do PB maiores de 18 anos, tendo o Ensino Médio como faixa de escolarização mínima.

O teste foi composto por 24 itens experimentais, sendo 8 em cada uma das condições, sendo elas: expressões idiomáticas, sentenças do tipo (43); expressões metafóricas convencionalizadas, sentenças do tipo (44); metáforas novas, sentenças do tipo (45). Além

disso, conta com 36 distratores, sentenças do tipo (46), que correspondem a 60% do total de itens do experimento. Além disso, antes das sentenças, os participantes foram expostos a um trecho que contextualiza a sentença experimental que leram em seguida. As sentenças de contexto foram utilizadas para que as metáforas novas pudessem ser processadas e compreendidas, visto que, como não são convencionalizadas, não possuem mapeamentos já conhecidos. Ao final da leitura de cada sentença experimental os participantes responderam a uma pergunta de compreensão (*ver seção Materiais*) sobre a expressão do trecho-alvo: metáfora convencional, metáfora nova ou idiomatismo.

As perguntas de compreensão foram usadas tanto como controle de atenção quanto como critério de exclusão de participantes no experimento: apenas os participantes que tivessem uma taxa de acerto igual ou superior a 85% teriam seus dados considerados na análise. No entanto, nenhum participante foi excluído e todos os 41 sujeitos experimentais tiveram desempenho dentro da margem estabelecida.

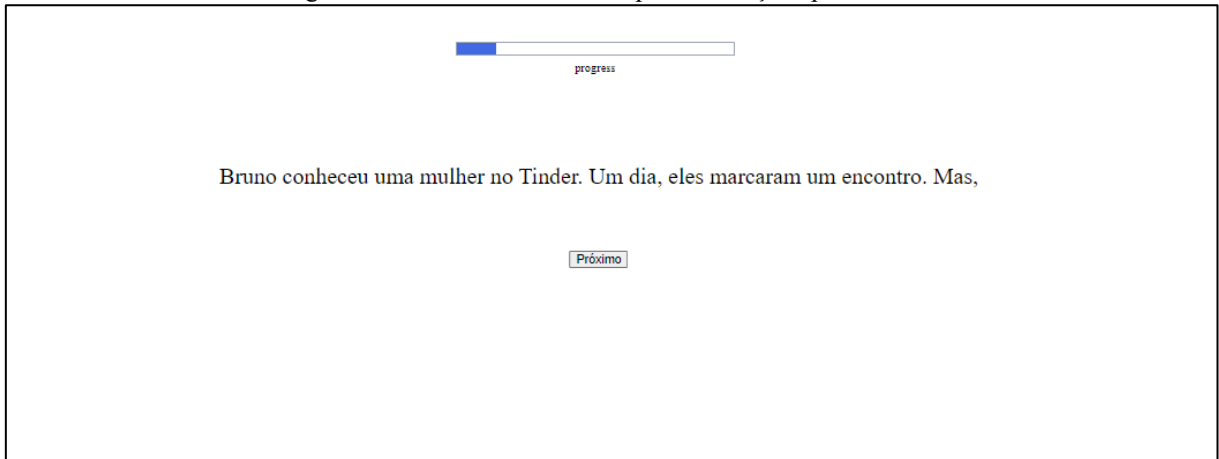
3.2.1 Procedimentos

O experimento foi hospedado e programado no site *PCIBex Farm*, uma plataforma da Universidade da Pensilvânia, que permite criar e rodar experimentos *online* em paradigmas convencionais como o *EyeTracking*, Tarefa *Stroop*, *Masked Priming*, Leitura autocadenciada, entre outros. Depois do experimento programado nessa plataforma, ocorreu o recrutamento dos participantes, 41 ao todo, todos alunos de disciplinas de graduação da Faculdade de Letras da UFMG. Eles acessaram um link com o experimento *online* que poderia ser realizado no computador ou em dispositivos portáteis/móveis (smartphones, tablets, notebooks). Ao acessarem o link, os participantes tiveram acesso ao Termo de Consentimento Livre e Esclarecido²⁵, que deveria ser aceito como requisito para prosseguir, caso concordassem em participar do experimento. Em seguida, responderam ao questionário demográfico, com informações como nome, gênero, idade e escolaridade.

Após realizarem a leitura de alguns itens de treinamento sem aviso prévio, para que se familiarizassem com o experimento, teve início o teste propriamente dito. Primeiramente, os participantes eram expostos a uma frase com um contexto relativo à sentença experimental que iriam ler, como mostra a figura 1.

²⁵ O TCLE está presente no Apêndice 4.

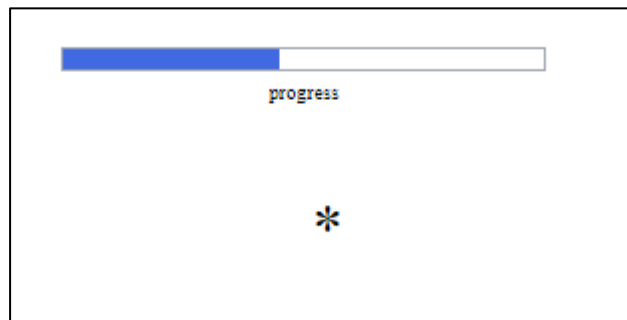
Figura 1 - Trecho com contexto para sentença experimental.



Fonte: PCIbex Farm

Depois da leitura, os participantes deveriam clicar no botão “próximo” para começar a leitura da sentença experimental. Em seguida, uma tela com ponto de fixação aparecia (Figura 2), seguida das lacunas da frase (Figura 3).

Figura 2- Ponto de fixação antes do início da leitura



Fonte: PCIbex Farm

Figura 3 - Lacunas a serem preenchidas pela sentença.



Fonte: PCIbex Farm

Os segmentos foram passados com um clique, no caso de dispositivos com mouse, ou com toques na tela, nos dispositivos com *touchscreen*. Ao passar para o segmento seguinte, o segmento anterior desaparecia, até chegar ao final da sentença, como na figura 4.

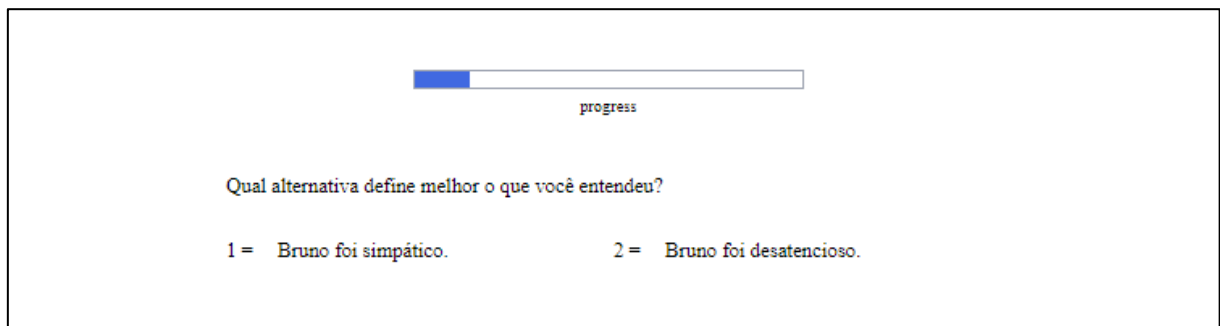
Figura 4 - Sequência da leitura autocadenciada.



Fonte: PCIbex Farm

Após a leitura automonitorada, os participantes respondiam a perguntas de compreensão, exemplificadas na Figura 5, a respeito do trecho-alvo, a fim de servirem como controle de atenção.

Figura 5 - Pergunta de compreensão



Fonte: PCIbex Farm

3.2.2 Materiais

O experimento contou com 24 sentenças experimentais, 8 em cada uma das 3 condições, e com 36 sentenças distratoras. Além disso, o teste contou com sentenças de contexto, antes das sentenças experimentais. e perguntas de compreensão depois da tarefa de leitura autocadenciada.²⁶

Sentenças experimentais

²⁶ As sentenças experimentais divididas em segmentos, bem como as sentenças de contexto e as perguntas de compreensão estão disponíveis no Apêndice 3.

As sentenças foram divididas como nos segmentos apresentados em (28), (29) e (30) e na figura 4, quais sejam:

- (i) *Trecho 1*: sujeito.
- (ii) *Trecho-alvo*: expressão da variável independente do tipo de construção.
- (iii) *Trecho spillover*: de valor adverbial.
- (iv) *Trecho final*: de valor adverbial.

As sentenças foram divididas por trecho, seguindo o formato do paradigma experimental utilizado, e controladas pelo número de sílabas (16 sílabas total, 5 no trecho-alvo, 4 no *spillover*) em cada trecho, com variação entre sentenças de 2 sílabas para mais ou menos. A escolha por número de sílabas se deve à dinâmica de leitura do PB, que, conforme Rayner e colegas (2004), tem cadência silábica. Os participantes do experimento foram expostos a todos os itens experimentais, entretanto, a ordem em que as sentenças apareciam foi aleatorizada de maneira automática via programação.

Além do controle realizado para todas as sentenças, algumas decisões experimentais foram tomadas, individualmente, na seleção dos itens experimentais de cada condição. As sentenças com construção idiomática foram extraídas de levantamentos feitos por Fulgêncio (2008) e do Dicionário Onomaseológico de Expressões Idiomáticas (Riva, 2009). A partir do material dessas fontes, as expressões idiomáticas foram selecionadas²⁷ com checagem no Corpus do Português: NOW e na rede social *X* (antigo *Twitter*). Todas as EI's estavam presentes no *corpus*, porém a frequência (maior ou menor) não foi considerada como um fator para seleção, ou seja, os valores de frequência de ocorrência não foram pareados. A rede social *X* foi utilizada como uma estratégia para averiguação de como sentenças desse tipo eram construídas, e em que contexto estavam sendo usadas, para que os itens fossem feitos de maneira semelhante.

As expressões metafóricas convencionais²⁸ foram retiradas dos mapeamentos e exemplos dados por Lakoff e Johnson (2002). As EMC's foram checadas em corpus, também sem considerar a frequência como fator de seleção de estímulos. Já as metáforas novas²⁹ foram criadas pela pesquisadora, com o auxílio da metodologia do material “New Metaphors: A Creative Toolkit for Generating Ideas and Reframing Problems”³⁰, que trazia alguns conceitos

²⁷ As EI's utilizadas podem ser consultadas no Apêndice 2.

²⁸ As EMC's utilizadas podem ser consultadas no Apêndice 3.

²⁹ As EMN's utilizadas podem ser consultadas no Apêndice 3.

³⁰ Essa é uma ferramenta criada pelo Observatório de Inovação do Setor Público, composta por 33 países. O projeto visa oferecer soluções inovadoras para problemas contemporâneos, com serviço voltado para o setor

e imagens. As imagens deviam ser pareadas com os conceitos de maneira aleatória e mapeadas via comparação metafórica para a criação de um novo conceito metafórico, criando metáforas novas como o exemplo a seguir.

(47) a. Davi alisou o chefe na reunião de equipe.

b. AGRADÁVEL É LISO

Além disso, selecionamos os verbos utilizados nas expressões metafóricas novas por meio da frequência no Corpus do Português: NOW.

Sentenças contextuais e perguntas de compreensão

Os itens experimentais serão antecidos por sentenças com trecho de contextualização a respeito das expressões idiomáticas, metafóricas convencionais e metafóricas novas. Essa decisão foi tomada para que fosse possível checar a atenção e a compreensão das expressões, em especial das metáforas novas, visto que são construídas do discurso e requerem o contexto da situação comunicativa para serem compreendidas. Não há um pareamento silábico entre os contextos apresentados para cada sentença. Há, no entanto, um padrão de construção que consiste em dois períodos simples e mais um conectivo, para transicionar para a sentença experimental, como apresentado na Figura 1: “Bruno conheceu uma mulher no Tinder. Um dia, eles marcaram um encontro. Mas, (...)”.

Depois de lido o contexto, os participantes fizeram a tarefa de leitura autocadenciada. Esta foi seguida de pergunta de compreensão.

(47) Bruno **foi bastante frio** com a mulher no encontro.

(48) Qual alternativa melhor define o que você entendeu?

a. Bruno foi desatencioso.

b. Bruno foi simpático.

A pergunta em (48) é a mesma para todos os itens experimentais e sempre será seguida de duas alternativas relacionadas à expressão analisada. Em (47), como exemplo, temos o trecho-alvo, em negrito, representado pela metáfora convencional TEMPERATURA É SENTIMENTO. As alternativas apareceram em ordem aleatorizada entre os sujeitos em todas as perguntas de compreensão.

Capítulo 4: O processamento de expressões idiomáticas

Este capítulo apresenta os resultados obtidos por meio do teste de leitura autocadenciada e os discute à luz do referencial teórico adotado, seguindo a seguinte ordem:

- a) Preparação dos dados.
- b) Resultados.
- c) Discussão.

4.1 Preparação dos dados

O experimento gerou uma amostra de 3936 observações nos trechos alvo e subsequente. Como mencionado na seção 3.2 (Desenho experimental e critério de exclusão), nenhum participante precisou ser excluído do estudo. No entanto, após análise inicial da distribuição da amostra dos dados de tempo de reação, notou-se a presença de alguns valores atípicos (*outliers*³¹). Alguns estudos adotam critérios entre 1200 e 2000 ms para o corte na leitura de trechos, com uma ou mais palavras (de Oliveira e Correia, 2023; Borges Mota e colegas, 2023), porém este estudo adota como critério o limite de 3000ms, com base:

- i) em um estudo que determina o teto de 2500 ms para a leitura de palavras isoladas (Souza Filho, 2021);
- ii) no fato de, não raro, os dados de tempo de reação não possuírem uma distribuição normal³², manifestando os efeitos justamente na cauda (Baayen e Milin, 2010);
- iii) na adoção do modelo linear de efeitos mistos para análise estatística que, por considerar como efeitos aleatórios participantes e itens, proporciona uma minimização de efeitos dos *outliers* da amostra (Baayen e Milin, 2010).

Portanto, com a adoção do critério de exclusão de *outliers*, ao final, o estudo contou com um total de 3884 observações.

4.2 Resultados

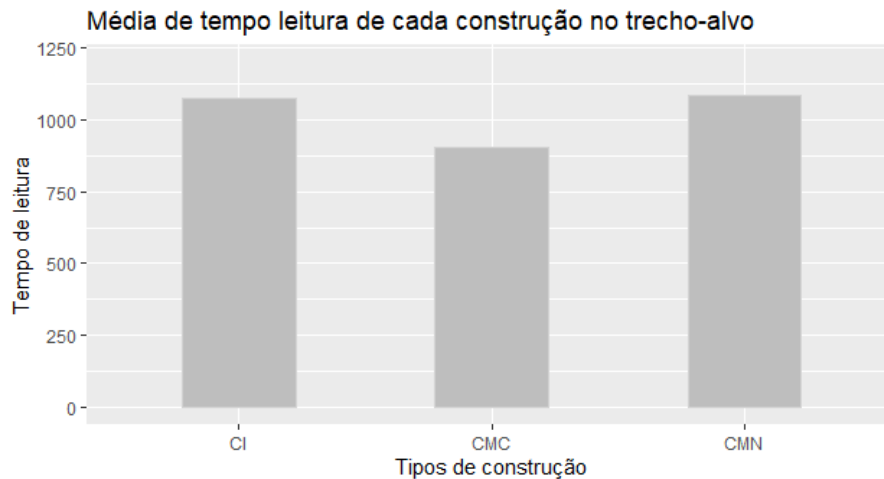
A análise realizada neste trabalho considera as médias dos tempos de leitura no trecho-alvo das construções metafóricas em separado - convencionalizadas (CMC) de um lado e novas (CMN) de outro -, e, como esperado, as CMN's possuem média de tempo de reação maior do

³¹ *Outliers* são dados drasticamente atípicos em relação à amostragem, que podem interferir nas análises estatísticas.

³² A distribuição normal é uma distribuição contínua e simétrica em ambos os lados da média, formando um gráfico com formato de sino, em que o lado direito, por exemplo, é um espelho do esquerdo.

que as CI's. As médias dos tempos de leitura das CMC's e das CMN's foi de 904.1331 e 1086.019 respectivamente, ao passo que a média dos tempos de leitura das CI's foi de 1072.508, como mostra o gráfico 1.

Gráfico 1 - Média do tempo de reação no trecho-alvo: CI, CMC e CMN

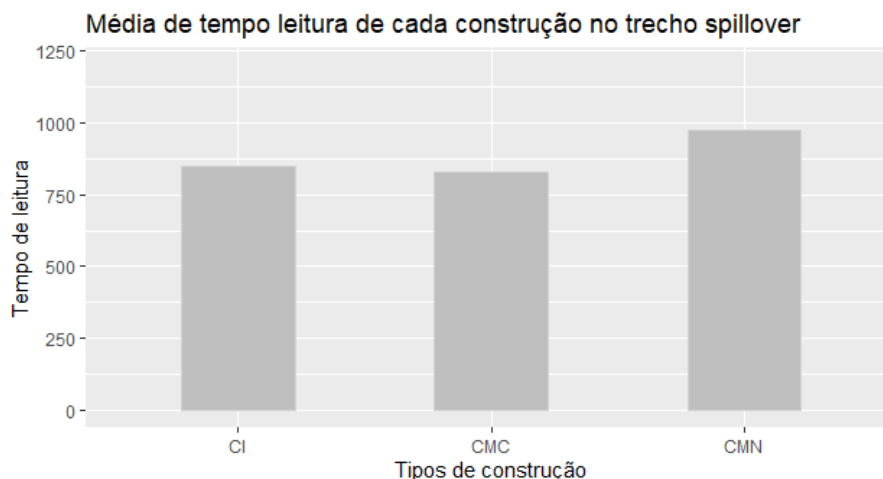


Fonte: a

autora

No trecho subsequente, os resultados revelam que também há diferença entre as médias de tempo de reação, as CMN's possuem maior tempo médio de leitura em comparação às CI's. As médias dos tempos de leitura das CMC's e das CMN's foi de 828.5385 e 975.5399 respectivamente, ao passo que a média dos tempos de leitura das CI's foi de 848.8738, como indica o gráfico 2.

Gráfico 2 - Média do tempo de reação no trecho spillover: CI, CMC e CMN



Fonte: a

autora

Um modelo linear misto (Pacote lme4, versão 1.1-35.1, no RStudio) foi ajustado para determinar se as diferenças nas médias eram estatisticamente significantes. No modelo, a variável de resposta tempo de reação (TR) e os tipos de condição (CI, CMC, CMN) foram incluídas como efeitos fixos, enquanto participantes e itens foram incluídos como efeitos aleatórios. O uso dos participantes como efeito aleatório é importante, pois “(...) ao

construirmos um modelo estatístico (...) seria adequado que esse modelo levasse em conta a variabilidade de cada participante.” (Godoy, 2019). Os itens também foram selecionados como efeitos aleatórios no modelo, a fim de minimizar as diferenças entre eles no que diz respeito à frequência de ocorrência de palavras específicas e familiaridade. Com a inclusão dos dois fatores como efeitos aleatórios, é possível alcançar uma maior generalização dos resultados encontrados. Sendo assim, o modelo foi ajustado, com as características citadas anteriormente, sendo um ajuste feito para o trecho-alvo e outro para o trecho subsequente. Um novo pacote de programação (LmerTest, versão 3.1-3, RStudio - R Core) foi utilizado para que fosse possível associar p-valores ao modelo linear misto.

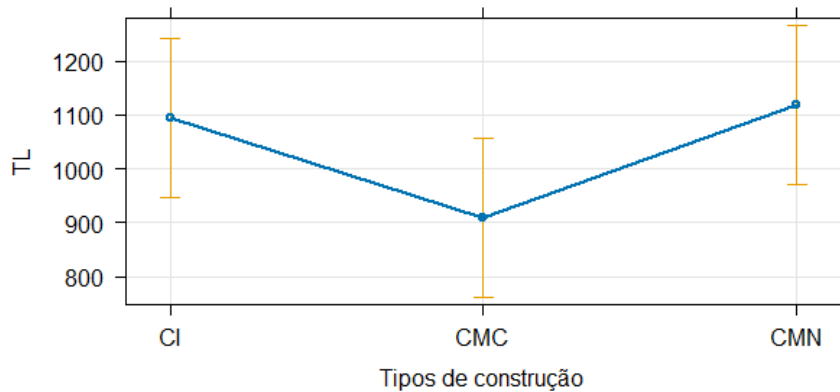
No trecho-alvo, foi identificado que a média de tempo de reação das construções metafóricas convencionais (CMC) é significativamente menor do que a das construções idiomáticas (CI) ($T = -3.674$, $p = 0.00143$), apontando para a influência do tipo de construção no tempo de reação e, portanto, para um maior custo processual das CI's em relação às CMC's. Entre a CI e a CMN não houve significância estatística no trecho-alvo ($T = 0.485$, $p = 0.63249$), conforme tabela 2. O gráfico 3 mostra os efeitos de correlação do modelo no trecho-alvo para os tipos de construção testados.

Tabela 2- Resultados do modelo linear de efeitos mistos no trecho-alvo

<i>Predictors</i>	<i>Estimates</i>	TL	
		<i>CI</i>	<i>p</i>
(Intercept)	1093.66	945.40 – 1241.92	<0.001
tipo cxn [CMC]	-184.56	-283.14 – -85.97	<0.001
tipo cxn [CMN]	24.44	-74.39 – 123.27	0.628
Random Effects			
σ^2	138516.99		
τ_{00} participantes	181996.07		
τ_{00} item	6613.93		
ICC	0.58		
$N_{participantes}$	41		
N_{item}	24		
Observations	952		
Marginal R^2 / Conditional R^2	0.026 / 0.588		

Fonte: a autora

Gráfico 3 - Efeitos de correlação entre as construções no trecho-alvo
Efeitos de correlação no trecho-alvo



Fonte: a autora.

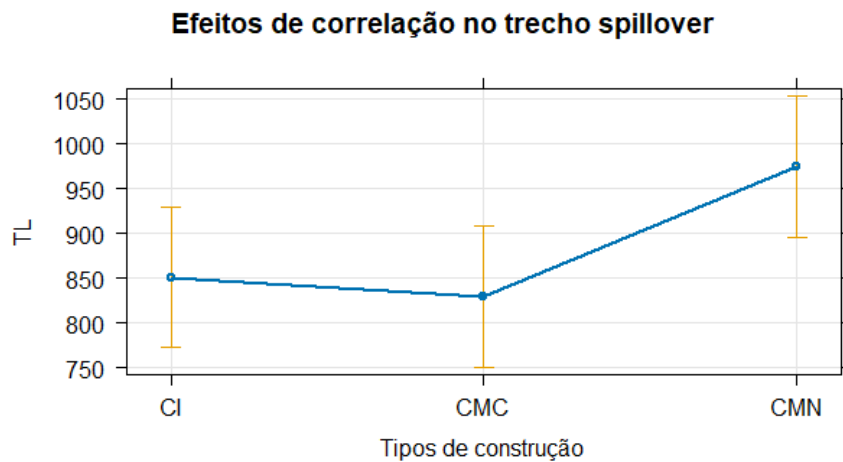
No trecho subsequente, foi ajustado um modelo linear misto seguindo os mesmos parâmetros adotados para o trecho-alvo. Foi identificado no modelo que não houve significância estatística na diferença das médias de tempo de reação da CI e da CMC ($T = -0.519$, $p = 0.60936$). Já entre as construções idiomáticas (CI) e as construções metafóricas novas (CMN) houve significância estatística entre as médias de tempo de reação no trecho subsequente ($T = 2.911$, $p = 0.00835$), conforme tabela 3. Essa diferença revela a interferência do tipo de construção no tempo de reação e, conseqüentemente, a diferença processual entre as duas condições. O gráfico 4 mostra os efeitos de correlação do modelo nas condições testadas nas amostras do trecho subsequente.

Tabela 3 – Resultados do modelo linear de efeitos mistos no trecho subsequente

<i>Predictors</i>	<i>Estimates</i>	TL	
		<i>CI</i>	<i>p</i>
(Intercept)	851.24	772.41 – 930.07	<0.001
tipo cxn [CMC]	-22.02	-105.32 – 61.28	0.604
tipo cxn [CMN]	123.52	40.24 – 206.81	0.004
Random Effects			
σ^2	79038.38		
τ_{00} participantes	29213.88		
τ_{00} item	5260.65		
ICC	0.30		
$N_{participantes}$	41		
N_{item}	24		
Observations	976		
Marginal R^2 / Conditional R^2	0.035 / 0.328		

Fonte: a autora

Gráfico 4 - Efeitos de correlação entre as construções no trecho subsequente



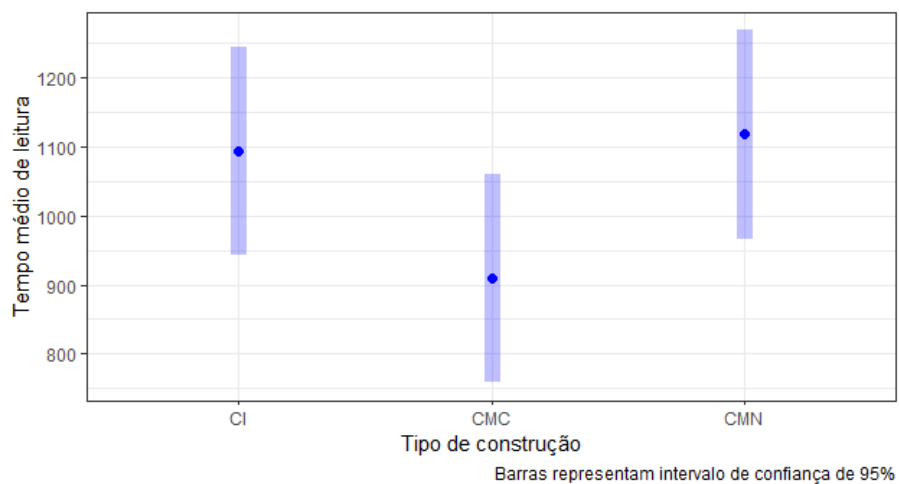
Fonte: a autora.

Além dos modelos apresentados, foi realizada uma análise *post-hoc*³³ utilizando o método Tukey, com ajuste de p-valor para a média do TR e comparação par a par dos tipos de construção testados. A análise foi feita separadamente para cada um dos trechos: alvo e subsequente. Na primeira, foram corrigidos os p-valores do modelo do trecho-alvo,

³³ A análise *post-hoc* é importante para este estudo, em especial, pela possibilidade contrastar os tipos de construção aos pares (CI x CMC; CI x CMN; CMC x CMN). Diferente do modelo linear misto, que compara cada tipo de construção a um intercepto, a análise *post-hoc* contribuiu com uma nova informação: a possibilidade de perceber as diferenças processuais entre as construções metafóricas convencionais e as novas.

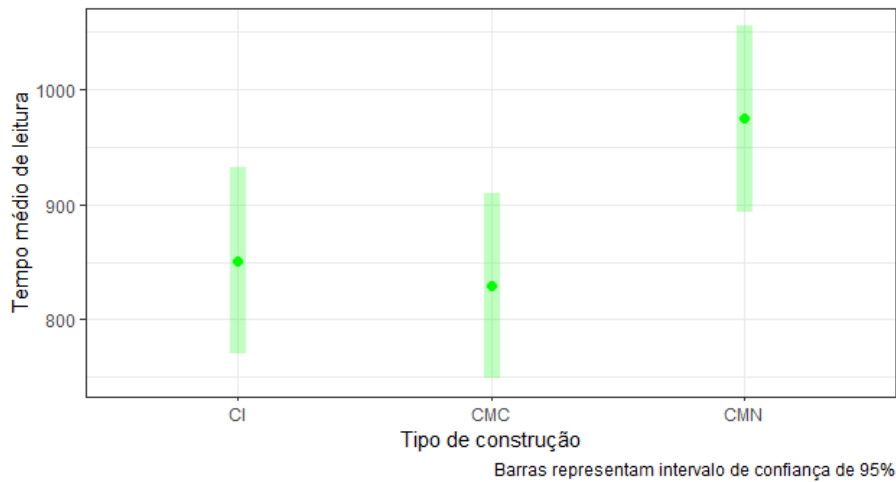
apresentando uma análise par a par. O contraste das CI's com as CMC's apresentou significância estatística ($T = 3.674$, $p = 0.0039$), indicando o maior custo processual das CI's. Entre CI's e CMN's, a análise mostra que a diferença entre as médias de tempo de reação não é significativa ($T = -0.485$, $p = 0.8790$). No contraste entre construções metafóricas convencionais (CMC) e as construções metafóricas novas (CMN), foi demonstrado uma significância estatística que revela uma diferença de processamento entre elas ($T = -4.158$, $p = 0.0012$), sendo as CMN's as construções com maior custo processual no trecho. Os resultados da análise podem ser observados no gráfico 5.

Gráfico 5- Análise *post-hoc*: diferença entre as médias no trecho-alvo



Fonte: a autora

Também foi realizada uma análise *post-hoc* para os trechos subsequentes, seguindo os mesmos critérios: ajuste de p-valores e comparação par a par entre os tipos de construção testados. O contraste entre CI's e CMC's aponta para uma diferença não significativa entre as expressões no trecho subsequente ($T = 0.519$, $p = 0.8631$). Já entre as construções idiomáticas (CI) e as construções metafóricas novas (CMN) há diferença estatística significativa entre as médias de tempo de reação ($T = -2.911$, $p = 0.0218$), evidenciando a natureza processual distinta das construções e o maior custo processual das CMN's. O contraste entre as construções metafóricas convencionais (CMC) e as construções metafóricas novas (CMN) também mostrou diferenças de tempo de reação significativas no trecho subsequente ($T = -3.429$, $p = 0.0068$), revelando que há transbordamento processual das CMN's. Os resultados da análise podem ser observados no gráfico 6.

Gráfico 6 - Análise *post-hoc*: diferença entre as médias no trecho subsequente

Fonte: a autora

Em resumo, os resultados do experimento confirmam a hipótese deste estudo: as construções metafóricas novas (CMN) possuem maior custo cognitivo e não são processadas como construções idiomáticas (CI). Além disso, eles apontam para diferenças importantes entre as construções testadas: CI, CMC e CMN. No trecho-alvo, a comparação entre construções idiomáticas (CI) e construções metafóricas convencionais (CMC), indica que as CI's possuem um maior custo processual em relação às CMC's. Os resultados da análise *post-hoc* apontam também para diferenças de processamento entre CMC e CMN nos dois trechos (alvo e subsequente).

4.3 Discussão

Neste estudo, foram analisados dados do tempo de reação (TR), no trecho-alvo e no trecho subsequente, de três tipos de construção: construção idiomática (CI), construção metafórica convencional (CMC) e construção metafórica nova (CMN). A hipótese deste estudo é de que as construções metafóricas novas possuem um maior custo processual do que as construções idiomáticas novas. Conforme esperado, os resultados obtidos no experimento revelam que há uma diferença significativa de TR entre as CI's e as CMN's no trecho subsequente. Além disso, aponta para diferenças de processamento entre os outros tipos de construção testados.

Na análise do trecho-alvo, foi encontrada uma diferença estatística significativamente menor do tempo de reação das construções metafóricas convencionais (CMC) em relação às construções idiomáticas (CI) o que, conseqüentemente, aponta para um maior custo processual das CI's. A facilidade de processamento das CMC's pode ser explicada em termos de convencionalidade e frequência. As metáforas conceptuais elencadas por Lakoff e Johnson

(2002) e utilizadas como base para a seleção dos itens deste nível de condição (CMC) fazem, segundo os autores, parte do sistema conceptual, organizando pensamentos e conceitos. Desse modo, são maneiras de conceptualizar a experiência e costumam ser adotadas de maneira mais generalizada por uma comunidade de usuários. A frequência, checada em corpus, revela que, dentre os itens utilizados no experimento, as CMC's são as construções mais frequentes (*ver apêndice 1*). Além da frequência ser um fator facilitador do processamento, a similaridade também é (Bybee, 2016). Por exemplo, um dos itens utilizados no experimento foi “Bruno foi bastante frio com a mulher no encontro”, que mapeia o conceito metafórico TEMPERATURA É SENTIMENTO. No exemplo, esse conceito é realizado a partir da expressão “SER frio”, que possui uma frequência por mil de 0,57. No entanto, a metáfora pode ser instanciada por outras construções similares, como “SER caloroso” (0,05 por mil), “SER ardente” (0,01 por mil), “SER seco” (0,25 por mil), dentre outras, estando disponível ao usuário da língua de diversas formas linguísticas. Logo, a existência de diferentes formas linguísticas para um mesmo conceito metafórico cria uma relação de similaridade entre elas. Essa relação de similaridade pode facilitar o processamento de uma construção (Bybee, 2016), visto que há uma representação mais abstrata de um conceito metafórico como TEMPERATURA É SENTIMENTO reforçada na memória.

É possível também dizer que cada conceito metafórico convencionalizado seja um *cluster* bem representado na memória do usuário, pois há várias construções que partem daquele determinado conceito. Sendo a memória associativa e acessada por meio do conteúdo que determinada mensagem deseja veicular (Goldberg, 2019), as instâncias de um conceito metafórico convencional estão mais reforçadas na memória por cada uma de suas instanciações, formando uma rede de construções, e facilitando o processamento das demais. Essa é uma explicação que se associa à Teoria da Carreira da Metáfora, que julga a convencionalidade como um fator fundamental no processamento de metáforas. De acordo com essa concepção, quando uma metáfora é convencional, “o termo base desenvolve uma abstração metafórica clara que pode ser acessada durante a compreensão, um tipo de atalho fica disponível. O ouvinte pode acessar diretamente o sentido abstrato metafórico em vez de ter que derivá-lo alinhando os dois termos literais. (Gentner e Bowdle, 2008, p. 116, **tradução nossa**)³⁴. Com isso, é possível assumir que o menor tempo de reação e, por consequência, menor custo de processamento das CMC's, revela a natureza categorial desse tipo de construção. Os conceitos

³⁴ No original: As the base term develops a clear metaphorical abstraction that can be accessed during comprehension, a kind of short cut becomes available. The listener can access the abstract metaphorical sense directly instead of having to derive it by aligning the two literal terms. (Gentner e Bowdle, 2008, p. 116)

metafóricos convencionais são, portanto, categorias abstratas convencionais associadas a certos padrões linguísticos alvo e, por isso, podem ser acessados mais facilmente.

Ainda no trecho-alvo, as CI's possuem tempo de reação superior às CMC's, como mencionado, indicando um maior custo processual das CI's em relação às CMC's. O fato do maior custo processual das CI's em relação às CMC's ter aparecido já no trecho-alvo é algo esperado no paradigma de leitura autocadenciada, pois é nesse segmento que as expressões analisadas, idiomatismos e metáforas convencionais e novas, estão localizadas. Esse maior custo pode ser explicado pela frequência menor das CI's em relação às CMC's (ver apêndices 1 e 2). Além disso, outros facilitadores, como a relação de similaridade explicada anteriormente para as CMC's (retomar exemplo com "SER frio" dado anteriormente), não estão presentes para as CI's, pelo menos não da mesma forma. Portanto, o tempo de reação maior da CI em relação à CMC mostra que elas são construções de natureza processual distinta. Trata-se de um achado que não estava previsto na hipótese experimental, mas que corrobora o cenário geral esperado para este estudo: as construções idiomáticas têm processamento distinto das construções metafóricas, inclusive das convencionais.

No trecho subsequente, a análise revela que não há diferenças estatísticas significativas entre os tempos de leitura das construções idiomáticas (CI) e das construções metafóricas convencionais (CMC). Esse resultado mostra que não houve efeito de transbordamento (Mitchell, 2018; Oliveira e colegas, 2022) (*spillover*) nas CI's em comparação às CMC's. Esse efeito ocorre quando o processamento incremental "resvala" em outro segmento, ou seja, antes mesmo de ter conseguido terminar o processamento do trecho-alvo, o sujeito passa para o trecho subsequente e o custo processual residual do trecho-alvo recai sobre o segmento seguinte. Esse fenômeno indica que determinada estrutura é mais difícil de ser processada. A ausência do efeito é um indicativo de que não há um processamento que gere algum efeito surpresa, sugerindo que os idiomatismos também são convencionais. Nesse sentido, as CI's parecem estar sob um efeito de *chunking* e são construções da língua, nas quais forma e função, com significado idiomático, estão pareados. Um *chunk* linguístico ocorre quando há um agrupamento de informação reforçado na memória e, com esse entendimento, a hipótese de que as CI's são acessadas como um bloco de sentido é reforçada. Bybee (2016) afirma que esse efeito pode ser observado em estruturas linguísticas, a despeito da frequência maior ou menor que elas tiverem e são bastante comuns em estruturas pré-fabricadas (*prefabs*), das quais as CI's fazem parte. Em resumo, os resultados do contraste da construção idiomática (CI) e da construção metafórica convencional (CMC) no trecho-alvo e também no subsequente sugerem que as CI's estão sob efeito *chunking* na língua. No entanto, isso também parece acontecer com

as CMC's, visto que são estruturas convencionais e, por isso, sujeitas a esse efeito. Tal efeito também colabora para o processamento e se alinha com o fato de que metáforas convencionais são categorias abstratas associadas convencionalmente a um domínio-alvo, e, por isso, processadas mais rapidamente (Gentner e Bowdle, 2008).

Ainda no trecho subsequente, ao contrastar as construções idiomáticas (CI) com as construções metafóricas novas (CMN), os tempos de leitura se mostram significativamente menores nas CI's. Desse modo, houve nas CMN's um efeito *spillover* no contraste com as CI's. Esse resultado corrobora a hipótese deste estudo de que as CMN's possuem maior custo cognitivo do que as CI's e, portanto, possuem natureza processual distinta. O tempo de reação superior das CMN's no trecho subsequente sugere que pode haver um processo de alinhamento conceitual entre os domínios (alvo e fonte), evidenciando a metáfora nova como um processo cognitivo ativo. Esse é um processo de maior custo cognitivo, tendo como justificativa o fato de ter o custo processual da expressão recaindo sobre outro trecho. Desse modo, os resultados corroboram a hipótese da Teoria da Carreira da Metáfora de que as CMN's são processadas via *símile*. Não há, como na CMC, uma categoria abstrata associada ao elemento do domínio-alvo, o que pode levar ao usuário a seguir um processo mais ativo e com maior custo cognitivo. Isso acontece devido a um efeito "surpresa" (Levy, 2007; Smith e Levy, 2013) presente em estruturas como as CMN's. Possivelmente, as CMN's têm custo processual elevado por uma baixa previsibilidade de que um significado literal ocorra em um determinado contexto. Por exemplo, na metáfora nova "esguichar a trama", presente no teste de leitura autocadenciada, a probabilidade de que sentido literal seja o pretendido é baixa em contexto em que alguém conta um segredo/fofoca a uma audiência. Sendo assim, como o usuário deseja ser cooperativo e compreender o seu interlocutor, ele busca um sentido outro, fazendo com que haja uma realocação de recursos durante a desambiguação incremental e probabilística na compreensão de sentenças. A nenhum momento foi avisado ao participante do experimento que ele estaria diante de uma sentença metafórica, assim como na experiência cotidiana não há um aviso prévio sobre a figuratividade de determinado enunciado. É plausível, então, afirmar que o processamento das construções metafóricas novas (CMN) é influenciado pelo fator surpresa que advém da figuratividade e da baixa probabilidade estatística de um significado literal para o conjunto de itens que compõem essas construções. Os três tipos de construção possuem idiosincrasias semânticas, mas são as CMN's que parecem refletir um processamento do que se espera de uma linguagem figurada que envolva a transferência de um domínio-fonte para um domínio-alvo, dado seu maior tempo de reação. Em outras palavras, o processo cognitivo de uma metáfora parece acontecer, pelo menos de forma mais evidente, apenas nas construções

metafóricas novas. A diferença de tempo de reação (TR) entre as CMC's e CMN's, bem como entre CI's e CMN's não permite apontar que essa transferência entre domínios e a comparação entre eles também aconteça com as CI's e as CMC's. Sendo assim, o processamento de uma metáfora não parece ser trivial, mas, ainda que mais custoso, é bem sucedido, ao menos, de compreensão suficientemente boa, visto que os participantes acertaram mais de 85% das perguntas de compreensão a respeito das construções. Em resumo, a comparação entre CI e CMN no trecho subsequente revela que são construções de natureza distinta, corroborando a hipótese deste trabalho: as CI's são diferentes de construções metafóricas novas, especialmente do ponto de vista do processamento, ou seja, construções idiomáticas não são metáforas.

Além disso, os resultados também apresentam diferenças significativas entre construções metafóricas convencionais (CMC) e construções metafóricas novas (CMN) na comparação par a par nos dois trechos (alvo e subsequente). Ainda que haja dois caminhos para o processamento, um para as CMC's e outro para as CMN's, a definição do que é metáfora, de maneira geral, segue a mesma para ambos os casos. São transferências entre domínios conceptuais. Ocorre que, com as CMC's, a convencionalidade tornou-as mais categoriais: no lugar de comparar um domínio mais físico com um mais abstrato, uma categoria abstrata convencionalizada relacionada a um domínio-alvo já existe, facilitando o processamento. As CMN's, por sua vez, não sofreram o mesmo processo de convencionalização e, por isso, precisam partir da transferência de traços mais concretos e físicos (literais) para em seguida transferir esses traços de significado para um domínio mais abstrato, tornando o processamento mais oneroso. Pesquisas corroboram essa diferença entre CMC e CMN a partir do viés de literalidade e abstração que este estudo segue (ver Gentner e Kurtz, 2006). Isso é o que propõe a hipótese da Teoria da Carreira da Metáfora, que, como mencionado, ajuda a explicar a diferença de processamento entre os dois tipos de metáfora, mas não só. A partir dessa hipótese também é possível tecer considerações sobre a natureza não metafórica das construções idiomáticas (CI). Se fossem duas construções de natureza equivalente, não haveria distinção significativa no processamento, do ponto de vista da convencionalidade, entre CMC's e CI's, visto que ambas são construções convencionais do PB. Entretanto, algum processo de mapeamento, presente nas CMC's, quando convencionalizado e frequente, parece ser um facilitador. Como não há um mapeamento prévio nas CI's, ou uma categoria metafórica convencional, as construções são processadas de maneira distinta. Ainda, partindo da definição de transferência entre domínios, o custo processual das CMN's não seria significativamente superior se nas CI's o mesmo processo de símile ocorresse. Desse modo, tanto quando a operação cognitiva metafórica é facilitada pela frequência e pela convencionalidade, quanto

quando não é facilitada, se tratando de um conceito metafórico novo, as construções idiomáticas não são comparáveis às metáforas do ponto de vista do processamento. Assim, a tradição de compreender as CI's como metáforas, ainda que “como um sentimento inconsciente”, como afirma Xatara (1998), para se confirmar, deveria ter as CI's seguindo os mesmos parâmetros de processamento das CMC's ou das CMN's, fato que não é corroborado pelos resultados do experimento de leitura autocadenciada.

Além das questões relacionadas ao processamento de CI's, é possível tecer considerações a respeito do estatuto teórico dessas expressões na abordagem construcionista baseada no uso. Sendo as construções *clusters* emergentes com traços imperfeitos de memória (Goldberg, 2019), há um pareamento entre forma e função em todas as construções da língua. Isso acontece em expressões regulares e composicionais, mas não só. Essa associação parece ser ainda mais forte no caso de expressões idiomáticas, pois há uma forma associada diretamente a um bloco de sentido. Desse modo, as CI's emergem como uma unidade com um nível de construcionalização maior, sendo mais lexicalizada do que as expressões composicionais, visto que é a combinação das palavras, em *chunk*, ou seja, informações agrupadas, e em co-ocorrência, que produz o significado idiomático. A construcionalização das construções idiomáticas é uma hipótese a ser checada em pesquisas futuras. Em outras palavras, não há um padrão abstrato que veicula um significado/função, como no caso de construções de estrutura argumental, por exemplo. Há, na verdade, uma estrutura menos abstrata e mais substantiva. É a combinação da forma, representada por cada palavra que compõe a expressão idiomática, pareada ao significado idiomático que constitui esse tipo de construção na língua.

Ainda, os achados experimentais levam a uma reflexão sobre o Princípio da Composicionalidade, apresentado na seção de Referencial Teórico. A ideia de composicionalidade, na visão tradicional, descartaria a assunção das construções analisadas nesse estudo (CI, CMC e CMN) como interpretada por esse viés. Mas, como apontado por teóricos como Goldberg (2015), esse é um princípio que precisa ser revisto. Assumir que o sentido se dá em função das partes de uma sentença/construção, excluiria construções como as CI's, e as construções metafóricas (CM). No entanto, é possível conjecturar formas de abordagem para o significado. No caso das construções idiomáticas (CI), há uma forte associação entre a combinação de palavras e o significado idiomático, o sentido estrito da composicionalidade não gera o significado de uma sentença idiomática como “bater as botas”. No entanto, além da associação não ocorrer em função do significado de cada item, mas sim da combinação, do bloco, essa é uma expressão que poderia ser analisada também no sentido literal, ou seja, via Princípio da Composicionalidade. Essa é uma possibilidade não só para essa

expressão em específico, mas também para outras expressões idiomáticas e também para as metafóricas. Desse modo, é pertinente apontar que, possivelmente, o contexto também seja importante para a compreensão do significado, bem como as relações de probabilidade que podem gerar efeitos surpresa no processamento ou facilitação pelo processo de *chunking*. No caso das CI's a memória parece estar pronta para acessar o significado idiomático de maneira global. O significado, portanto, é motivado pelo uso, que é flexível e dinâmico, e não pela estrutura interna de uma construção. O mesmo parece acontecer com as CMC's, o significado da expressão se dá pela função não dos itens que a compõem, mas da categoria metafórica abstrata associada ao domínio-alvo. Nas CMN's, o fator contexto também fundamental para o significado, pois é ele quem dita que não se trata de uma expressão convencional e literal, gerando o efeito surpresa. Portanto, a análise dessas construções propicia uma visão distinta sobre a formação do significado e corrobora a necessidade de revisão do Princípio.

Em resumo, os resultados se alinham à hipótese levantada por esse estudo: as expressões idiomáticas não são processadas como expressões metafóricas novas, visto que possuem diferenças em seu tempo de reação, implicando em diferenças na forma como são processadas. Além disso, os resultados apontam para a diferença processual entre CMC's e CMN's, possibilitando tecer conjecturas a respeito do papel da convencionalidade e da frequência na compreensão da língua.

Capítulo 5: Conclusão

Este estudo partiu da hipótese apresentada por Fulgêncio e Ciríaco (2018) de que as expressões idiomáticas não possuem natureza metafórica, buscando encontrar evidências experimentais do que fora delimitado teoricamente pelas autoras. Desse modo, a hipótese experimental desta pesquisa é de que as construções metafóricas novas (CMN) possuem um maior tempo de reação e, portanto, maior custo processual em relação às construções idiomáticas (CI), no trecho-alvo e no trecho subsequente ou ao menos no trecho subsequente. No capítulo 1, o objeto de estudo expressões idiomáticas foi apresentado, descrevendo os pontos que caracterizam essas instanciações da língua: idiosincrasias, convencionalidade e a aposta na memorização em bloco, bem como a sua descrição gramatical. Além disso, foram apresentadas neste capítulo os desafios que as expressões idiomáticas instauram para a teoria linguística e a tradição de compreendê-las como metáfora.

No capítulo 2, com o objetivo de descrever as expressões idiomáticas, bem como as expressões metafóricas, foram apresentadas as teorias que corroboram as análises realizadas. Foram desenvolvidos pressupostos importantes da Linguística Baseada no uso (Bybee e Beckner, 2010; Diessel, 2017; Evans e Green, 2006, entre outros), como frequência, convencionalidade e efeito chunking. Também foram apresentados os pressupostos da Gramática de Construções (Goldberg, 1995, 2006, 2019). O tratamento da metáfora como processo cognitivo (Lakoff e Johnson, 2002; Kövecses, 2002; Gentner e Bowdle, 2008) foi discutido, seguido de breves apontamentos sobre processamento linguístico e acesso lexical (Savage e colegas. 2003; Bybee, 2016).

No capítulo 3, a metodologia experimental de leitura autocadenciada foi apresentada. Discutiui-se a definição do paradigma experimental de leitura autocadenciada, bem como suas vantagens e as medidas ofertadas. Além disso, foi detalhado o desenho experimental, os critérios de exclusão e as escolhas que orientaram o desenvolvimento dos itens experimentais. Foram relatados também os procedimentos envolvidos no experimento realizado pelos participantes.

Os resultados do experimento de leitura autocadenciada foram discutidos no capítulo 4. Foi realizada uma análise descritiva da média de tempo de reação das amostras por trecho (alvo e subsequente) entre as construções idiomáticas e as construções metafóricas, incluindo a metafórica convencional e metafórica nova. Descreveu-se também a comparação das médias de tempo de reação dos três tipos de construção: CI, CMC e CMN, visto que a hipótese tece considerações específicas a respeito do contraste entre CI e CMN. Posteriormente, para a média

das três condições, foi ajustado um modelo estatístico e realizada uma análise *post-hoc* que apontou diferenças estatísticas significativas entre os tipos de construção. Tanto o modelo linear misto quanto a análise *post-hoc* confirmam a diferença processual entre as três condições: CI, CMC e CM. Apontam, ainda, para a natureza processual distinta das CMC's e das CMN's, corroborando a Teoria da Carreira da Metáfora.

Além disso, a diferença entre a construção idiomática (CI) e a construção metafórica nova (CMN) aparece no trecho subsequente, demonstrando um efeito *spillover* das CMN's em relação às CI's. Esse resultado aponta para um maior tempo médio de leitura, e por consequência, um maior custo cognitivo das construções metafóricas novas (CMN) em relação às construções idiomáticas (CI), corroborando hipótese do estudo: CI's não são CMN's do ponto de vista do processamento. Ainda, a diferença de processamento entre CI e CMC, com custo processual maior para CI aparecendo apenas no trecho-alvo, mostra que as CI's também não possuem processamento compatível com a de uma metáfora convencionalizada. Em resumo, as CI's são construções da língua com pareamento de forma e função e significado idiomático e armazenadas como *chunks* da memória, e não processadas como metáforas.

5.1 Limitações do estudo

O estudo possui contribuições para a teoria linguística, em especial para teoria construcionista baseada no uso, pois apresenta discussões sobre uma expressão tida como periférica na tradição linguística e a analisa pelo viés experimental. No entanto, apresenta limitações não só de tempo de pesquisa, mas também metodológicas. As construções idiomáticas (CI) e as construções metafóricas convencionais (CMC) utilizadas nesse experimento não foram selecionadas pela frequência, a pesquisa em *corpus* apenas atestou a existência dessas expressões nos dados do NOW. Seria interessante que os itens experimentais, se possível, fossem pareados por frequência ou, ainda, que um estudo prévio sobre a aceitabilidade e a familiaridade dessas expressões fosse realizado.

5.2 Pesquisas futuras

Esta pesquisa levanta, de maneira geral, questões sobre como a frequência e a convencionalidade afetam o processamento de construções da língua. Uma pesquisa futura sobre a hipótese de construcionalização das expressões idiomáticas pode ser bastante interessante para uma descrição mais aprofundada sob o viés da abordagem construcionista baseada no uso. Estudos sobre o processamento de outras expressões fixas podem ser relevantes

para uma descrição mais aprofundada, sob viés experimental, dos efeitos cognitivos de processos de *chunk* e da convencionalidade na língua. Ainda, uma investigação sobre as diferenças encontradas entre metáforas convencionais e metáforas novas pode ser um caminho para se compreender melhor como ocorrem os processos cognitivos envolvidos nas expressões metafóricas.

REFERÊNCIAS

- KEMMER, S, BARLOW, M. **Introduction: a usage-based conception of language.** (2000). In: Barlow; Kemmer (Ed.). *Usage-based models of language.* (p.7-25). Stanford, CA: CSLI Publications, 2000.
- BAAYEN, R. H.; MILIN, P. **Analyzing reaction times.** *International Journal of Psychological Research*, [S. l.], v. 3, n. 2, p. 12–28, 2010. DOI: 10.21500/20112084.807. Disponível em: <https://revistas.usb.edu.co/index.php/IJPR/article/view/807>. Acesso em: 10 de jan. 2024.
- BORGES MOTA, M.; PINHEIRO DE ANGELI, N.; RICARDO BIN, P.; CASSOL RIGATTI, P.; FREIXIELA ADAMCZYK, P. V. **Efeitos de priming estrutural no processamento de passivas em português brasileiro.** *Letrônica*, [S. l.], v. 16, n. 1, p. e44446, 2023. DOI: 10.15448/1984-4301.2023.1.44446. Disponível em: <https://revistaseletronicas.pucrs.br/index.php/letronica/article/view/44446>. Acesso em: 16 fev. 2024.
- BOWDLE, B; GENTNER D. **The Career of Metaphor.** *Psychological Review*, 112(1), 193–216, 2005.
- BYBEE, J. **Língua, uso e cognição.** Tradução: FURTADO DA CUNHA, M.A, São Paulo: Editora Cortez, 2016.
- BYBEE, J; BECKNER, C. Usage Based Theory. In: HEINE, Bernd; NARROG, Heiko. **The Oxford Handbook of Linguistics Analysis.** Cambridge University Press, 2010, p. 827-855.
- CROFT, W. **Explaining language change: an evolutionary approach.** London: Longman, 2000.
- DAVIES, Mark. **Corpus do Português: News On The Web (2012-2019),** 2006. Disponível em: <<https://www.corpusdoportugues.org/now/>>. Acesso em: 1 dez 2023.
- DE OLIVEIRA, Carolina Aguiar; CORRÊA, Letícia Maria Sicuro. **O processamento de sentenças passivas por adultos com sintomas de Déficit de Atenção e Hiperatividade: um estudo exploratório.** *Veredas-Revista de Estudos Linguísticos*, v. 27, n. 1, 2023.
- DIESSEL, HOLDGER. Usage-Based Linguistics. In: ARONOFF, Mark (org.). **Oxford Research Encyclopedia of Linguistics.** Oxford: Oxford, 2017.
- EVANS, V.; GREEN, M. **Cognitive Linguistics: an introduction.** Edinburgh: Edinburgh University Press, 2006.
- FERRARI, L. **Introdução à Linguística Cognitiva.** Editora Contexto, São Paulo, 2007.
- FILLMORE J., KAY, P. & O’CONNOR, M. C. **Regularity and idiomaticity in grammatical constructions: the case of let alone,** *Language*, 1988.
- FULGÊNCIO, L. **Expressões fixas e idiomatismos do português brasileiro.** 2008. Tese (Doutorado em Linguística e Língua Portuguesa) – Pontifícia Universidade Católica de Minas Gerais, PUC, Belo Horizonte, 2008. 506 p.
- FULGÊNCIO, L.; CIRÍACO, L. The non-metaphoric nature of idioms. In: **Verbs, clauses and constructions: Functional and typological approaches.** Cambridge Scholars Publishing, 2018.
- FURTADO DA CUNHA, M. A.; CEZARIO, M. M. **Conhecimento, criatividade e produtividade sob a perspectiva da linguística funcional centrada no uso.** *ALFA: Revista de Linguística*, São Paulo, v. 67, 2023. Disponível em: <https://periodicos.fclar.unesp.br/alfa/article/view/15041>. Acesso em: 01 dez 2023.
- GENTNER, D.; KURTZ, K. J. **Relations, Objects, and the Composition of Analogies.** *Cognitive Science*, 30(4), 609–642, 2006.
- GENTNER, D.; BOWDLE, B. Metaphor as structure-mapping. In R. W. Gibbs, Jr. (Ed.), **The Cambridge handbook of metaphor and thought.** Cambridge University Press, 2008.
- GODOY, M. C. **Introdução aos modelos lineares mistos para os estudos da linguagem.** PsyArXiv, 2019. Disponível em: <<https://mahayana.me/mlm/>>. Acesso em 01 dez. 2023.

- GOLDBERG, A. **Constructions**: A construction grammar approach to argument structure. Chicago, IL: University of Chicago Press, 1995.
- _____. **Constructions at work**: the nature of generalization in language. Oxford: Oxford University Press, 2006.
- _____. **Explain me this**: creativity, competition, and the partial productivity of constructions. Princeton: Princeton University Press, 2019.
- _____. Constructionist approaches. In: Hoffmann, T.; Trousdale, G. (Ed.). **The Oxford Handbook of Construction Grammar**. Oxford University Press, 2013.
- _____. **Compositionality**. In *The Routledge Handbook of Semantics* (pp. 419-433). Taylor and Francis Inc., 2015.
- HODGES, W. **Formal features of compositionality**. *Journal of Logic, Language and Information*, v. 10, p. 7-28, 2001.
- KENEDY, E. Gerativismo. In: Mário Eduardo Toscano Martelotta. (Org.). **Manual de linguística**. São Paulo, Contexto, 2008. p. 127-140.
- KÖVECSES, Z. **Metaphor**: a practical introduction. Oxford: Oxford University Press, 2002.
- LANGACKER, R. **Cognitive Grammar**. New York: Oxford University Press, 2008.
- LAKOFF, G. Neural metaphor. In R. W. Gibbs, Jr. (Ed.), **The Cambridge handbook of metaphor and thought**. Cambridge University Press, 2008.
- LAKOFF, G.; JOHNSON, M.. **Metáforas da vida cotidiana**. Coordenação de tradução: Mara Sophia Zanotto. São Paulo: Mercado das Letras, 2002.
- MARTINS, V; MONTEIRO-PLANTIN, R. **A expressão de violência no idiomatismo brasileiro**. *Revista Antares*, v. 4, n. 7, jan./jul., 2012.
- MEDEIROS, J. *et al.* **Acesso lexical**: uma rota dupla para o português brasileiro. 2014.
- MITCHELL, Don C. **An evaluation of subject-paced reading tasks and other methods for investigating immediate processes in reading**. *New methods in reading comprehension research*, Kieras, David E. and Marcel A. Just (Eds.). Hillsdale, NJ: Earlbaum. 1984, pp. 69-89.
- NENCHEVA, L; TAMIR DI, LEW-WILLIAMS, C. **Caregiver speech predicts the emergence of children's emotion vocabulary**. *Child development*. 2023 ,
- NOGUEIRA, L.. **A presença das expressões idiomáticas (EIs) na sala de aula de E/LE para brasileiros**. 2008.
- OLIVEIRA *et al.* **Leitura autocadenciada (com e sem labirinto): histórico e reflexões metodológicas**. In: OLIVEIRA, C; SÁ, T. (Org.). **Métodos Experimentais em Psicolinguística**. São Paulo : Pá de Palavra, 2022.
- PASQUALI L. **Psicometria: teoria dos testes na psicologia e na educação**. Petrópolis: Vozes; 2015.
- PEDRO, M. **As expressões idiomáticas no ensino de português como língua estrangeira para estudantes uruguaios**. 2007.
- PEREK, F.; GOLDBERG, A. **Generalizing beyond the input**: The functions of the constructions matter. *Journal of Memory and Language*, 84, 108–127
- PINKER; PRINCE, A. **Regular and irregular morphology and the psychological status of rules of grammar**, In: SUTTON, L.A.; JOHNSON, C.; S HIELDS, R. (Eds.). ANNUAL MEETING OF THE BERKELEY LINGUISTICS SOCIETY. Proc. 17th... Berkeley, CA.: Berkeley Linguistics Society, 1992
- PRATA, M. **Mas será o Benedito?** - dicionário de provérbios, expressões e ditos populares. São Paulo: Editora Planeta, 2011.
- ROHRER, T. Embodiment and Experientialism. In: GEERAERTS, D.; CUYCKENS, H. (Eds.). **Handbook of Cognitive Linguistics**. Oxford: Oxford University Press, 2007. p. 25-47.
- SAVAGE, C.; LIEVEN, H.; THEAKSTON, A.; TOMASELLO, M. **Testing the abstractness of children's linguistic representations**: Lexical and structural priming of syntactic constructions. *Developmental Science*. v. 6, p. 557-567, 2003.

SCHREUDER, R.; BAAYEN, R. H. **Modeling morphological processing**. In: FELDMAN, L. B. (Ed.). *Morphological aspects of language processing*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc., 1995. p. 131-154

SOUZA FILHO, N. S. de. **A Small but Significant Effect: Lexical Prediction in a Selfpaced Reading Study**. 2021.

TOMASELLO, M. **First steps toward a usage-based theory of language acquisition**. *Cognitive Linguistics*, 11(1-2), 61–82, 2000.

_____. **Constructing a language: A usage-based theory of language acquisition**. Harvard University Press, 2003.

_____. *et al.* **Understanding and sharing intentions: The origins of cultural cognition**. *Behavioral and brain sciences*, v. 28, n. 5, p. 675-691, 2005.

_____. **The usage-based theory of language acquisition**. In: *The Cambridge handbook of child language*. Cambridge Univ. Press, p. 69-87, 2009.

_____; CARPENTER, M. **Shared intentionality**. *Developmental science*, v. 10, n. 1, p. 121-125, 2007.

RAYNER, K. *et al.* The effects of frequency and predictability on eye fixations in reading: implications for the EZ Reader model. **Journal of Experimental Psychology: Human Perception and Performance**, v. 30, n. 4, p. 720, 2004.

RAYNER, K., & CLIFTON, C., Jr. *Language processing*. In: MEDIN, D. (Volume Editor) Stevens. *Handbook of Experimental Psychology: Memory and Cognitive Processes*. Third Edition: Volume 2. New York: John Wiley and Sons, Inc. Copyright John Wiley & Sons, Inc. 2002. pp. 261-316.

RIVA, H. **Dicionário onomasiológico de expressões idiomáticas usuais na língua portuguesa no Brasil**. 2009.

TAGNIN, S. **Expressões idiomáticas e convencionais**. Editora Ática, 1989.

_____. **O jeito que a gente diz**. Disal Editora-Bantim, Canato e Guazzelli Editora Ltda, 2005.

TERMIGNONI, S. FINATTO, M. . **Sobre a importância de ensinar expressões idiomáticas**. *Revista de Italianística*, n. 35, p. 112-124, 2017.

XATARA, C. **O resgate das expressões idiomáticas**. *Alfa: revista de linguística*, v. 39, 1995.

_____. **O campo minado das expressões idiomáticas**. *ALFA: Revista de Linguística*, 1998.

_____. **O ensino do léxico: as expressões idiomáticas**. *Trabalhos em linguística aplicada*, v. 37, 2001.

ZYZIK, Eve. **The role of input revisited: Nativist versus usage-based models**. **L2 Journal**, v. 1, n. 1, 2009.

WESTERSTÅHL, D. **Idioms and compositionality**. JFAK. *Essays Dedicated to Johan van Benthem on the Occasion of his 50th Birthday*. Vossiuspers. Amsterdam University Press, , 1999.

APÊNDICES

Apêndice 1: Construções metafóricas: mapeamentos e frequência por mil

Tipo de construção	Sentença	Mapeamento	Frequência/mil
CMC	Pedro gastou muito tempo na reunião da manhã.	TEMPO É DINHEIRO	0,56
	Bruno foi bastante frio com a mulher no encontro.	TEMPERATURA É SENTIMENTO	0,57
	Bruna saiu da rotina nas férias de janeiro.	ESTADO É RECIPIENTE	0
	Mauro foi o alicerce da família desde cedo.	PESSOAS SÃO CONSTRUÇÕES	0,02
	Paulo deu uma ideia excelente no trabalho.	CONDUTO	1,4
	Artur entrou em depressão com o evento familiar.	ESTADO É RECIPIENTE	0,55
	Dayse tem alto renome em seu setor de trabalho.	STATUS É ALTURA	0,02
	Lara estava brilhante na defesa de mestrado.	IDEIAS SÃO LUZ	0,08

CMN	Sara segue arco-íris na vida desde cedo.	SONHO É ARCO-ÍRIS	0
	Davi alisou o chefe na reunião de equipe.	AGRADÁVEL É LISO	0
	Iago afinou a mente em casa pelas manhãs.	MENTE É INSTRUMENTO MUSICAL	0
	Ana aparou o filho na escola pela manhã.	EDUCADO É APARADO	0
	José acumulou ferrão com os anos de trabalho.	CONFIANÇA É FERRÃO	0
	Lucas arranhou a barriga no emprego pela manhã.	SER PREGUIÇOSO É ARRANHAR	0
	Laís colheu manga verde na reunião de trabalho.	FALA É COLHEITA	0
	João esguichou a trama para todos da equipe.	FOFOCA É FONTE	0

Apêndice 2: Construções idiomáticas e frequência de ocorrência

Construção idiomática	Frequência por mil
Lucas amarrrou um bode do colega de trabalho.	0
Paula não deu a mínima para o Dom no evento.	0,39
Juca não pregou o olho na véspera do exame.	0,03
Mara botou panos quentes na discussão dos primos.	0,02
Marcio deu uma colher de chá para Ana no torneio.	0,02
Paulo meteu o bedelho na conversa dos colegas.	0,1
Carla escondeu o ouro do colega de trabalho.	0,01
Carlos moveu céus e terra pela obra da igreja.	0,02

Apêndice 3: Materiais

ITEM	CONSTRUÇÃO	CONTEXTO	TRECHOS				PERGUNTA DE COMPREENSÃO	ALTERNATIVAS* ⁱ	
			INICIAL	ALVO	SUBSEQUENTE	FINAL			
101	CI	Lucas e Pedro eram muito amigos. Um dia, eles estavam conversando e Pedro foi muito mal educado. Depois disso,	Lucas	amarrou um bode	do colega	de trabalho	Qual alternativa define melhor o que você entendeu?	Lucas trabalhou com o colega	Lucas ficou zangado com o colega
102	CMC	Pedro é responsável por um projeto. Ele deveria entregar alguns relatórios até o fim do dia. Porém,	Pedro	gastou muito tempo	na reunião	da manhã	Qual alternativa define melhor o que você entendeu?	Pedro desperdiçou tempo.	Pedro fez bom uso do tempo.
103	CMN	Sara deseja algo novo. Ultimamente, sonha em ser uma influencer de sucesso. Então,	Sara	segue arco-íris	na vida	desde cedo	Qual alternativa define melhor o que você entendeu?	Sara é trabalhadora.	Sara é sonhadora.
104	CI	Dom tinha interesse em Paula. Um dia, eles se encontraram. Mas,	Paula	não deu a mínima	para o Dom	no evento	Qual alternativa define melhor o que você entendeu?	Paula ignorou Dom.	Paula cumprimentou Dom
105	CMC	Bruno conheceu uma mulher no Tinder. Um dia, eles marcaram um encontro. Mas,	Bruno	foi bastante frio	com a mulher	no encontro	Qual alternativa define melhor o que você entendeu?	Bruno foi simpático.	Bruno foi desatencioso.
106	CMN	Davi queria ser promovido. Um dia o chefe o convocou para uma reunião. Então,	Davi	alisou o chefe	na reunião	de equipe	Qual alternativa define melhor o que você entendeu?	Davi bajulou o chefe	Davi encostou no chefe.
107	CI	Juca estava terminando a faculdade. Ele faria a prova da OAB no dia seguinte. Então,	Juca	não pregou o olho	na véspera	do exame	Qual alternativa define melhor o que você entendeu?	Juca não fechou os olhos.	Juca não dormiu.
108	CMC	Bruna tem um dia a dia muito agitado e costuma fazer sempre as mesmas coisas. Ela precisava descansar. Então,	Bruna	saiu da rotina	nas férias	de janeiro	Qual alternativa define melhor o que você entendeu?	Bruna fez algo diferente	Bruna fez o que já sabe.
109	CMN	Iágo queria aprender coisas novas. Todos os dias, antes de ir trabalhar, estudava. Então,	Iágo	afinou a mente	com leitura	pelas manhas	Qual alternativa define melhor o que você entendeu?	Iágo harmonizou algo.	Iágo aprendeu algo

110	CI	Mara estava cuidando dos primos. Eles estavam discutindo muito naquela manhã. Então,	Mara	botou panos quentes	na discussão	dos primos	Qual alternativa define melhor o que você entendeu?	Mara amenizou o conflito entre os primos.	Mara colocou um pano nos primos.
111	CMC	Mauro é muito responsável. Ele começou a trabalhar bem jovem. Por isso,	Mauro	foi o alicerce	da família	desde cedo	Qual alternativa define melhor o que você entendeu?	Mauro deu uma casa para família.	Mauro deu apoio para família.
112	CMN	O filho de Ana brigou com o colega na escola. Ana foi chamada pela diretora. Então,	Ana	apareceu o filho	na escola	pela manhã	Qual alternativa define melhor o que você entendeu?	Ana fez uma correção no filho.	Ana fez um corte no filho.
113	CI	Ana estava disputando um torneio. No dia, ela não foi com o uniforme adequado. Então,	Márcio	deu uma colher de chá	para Ana	no torneio	Qual alternativa define melhor o que você entendeu?	Márcio presenteou Ana.	Márcio facilitou para Ana.
114	CMC	Paulo era muito competente. Ele estava participando de uma reunião de projeto no trabalho. Então,	Paulo	deu uma ideia	excelente	no trabalho	Qual alternativa define melhor o que você entendeu?	Paulo fez uma observação.	Paulo fez uma doação.
115	CMN	José tinha vergonha de falar em público. No trabalho sempre tinha que conduzir reuniões. Então,	José	acumulou ferrão	com os anos	de trabalho	Qual alternativa define melhor o que você entendeu?	José juntou abelhas	José ganhou confiança.
116	CI	Os amigos de Paulo estavam discutindo. Paulo chegou e ouviu a conversa. Então,	Paulo	meteu o bedelho	na conversa	dos colegas	Qual alternativa define melhor o que você entendeu?	Paulo foi intrometido.	Paulo enfiou o trinco.
117	CMC	Artur teve um problema familiar sério. Ele passou a se sentir mal. Então,	Artur	entrou em depressão	com o trauma	familiar	Qual alternativa define melhor o que você entendeu?	Artur caiu num buraco.	Artur ficou abatido.
118	CMN	Lucas tinha muito trabalho para fazer. Só que ele não estava disposto. Então,	Lucas	arranhou a barriga	no emprego	pela manhã	Qual alternativa define melhor o que você entendeu?	Lucas trabalhou com preguiça	Lucas trabalhou esfolado.
119	CI	Carla e seu colega estavam atrás de uma promoção. Ela sabia que haveria uma vaga de gerente disponível. Então,	Carla	escondeu o ouro	do colega	por ambição	Qual alternativa define melhor o que você entendeu?	Carla ocultou uma joia.	Carla dissimulou algo.
120	CMC	Dayse é muito qualificada e competente. Ela é reconhecida na empresa. Por isso,	Dayse	tem alto renome	em seu setor	de trabalho	Qual alternativa define melhor o que você entendeu?	Dayse tem importância.	Dayse tem altura.





121	CMN	Laís foi chamada para uma reunião de projeto. Ela acabou opinando sobre o conteúdo antes da hora. Então,	Laís	colheu manga verde	na reunião	de trabalho	Qual alternativa define melhor o que você entendeu?	Laís pegou fruta.	Laís falou antes da hora.
122	CI	Carlos é uma pessoa muito colaborativa. Um dia, a igreja de sua comunidade entrou em reforma. Então,	Carlos	moveu céus e terra	pela obra	da igreja	Qual alternativa define melhor o que você entendeu?	Carlos foi muito esforçado.	Carlos deslocou algo de lugar.
123	CMC	Lara trabalhou muito em sua pesquisa. Ela se preparou bem para a banca. Mais tarde,	Lara	estava brilhante	na defesa	de mestrado	Qual alternativa define melhor o que você entendeu?	Lara usou purpurina.	Lara fez bem feito.
124	CMN	João ficou sabendo de uma história. Um dia ele se reuniu com a equipe. Então,	João	esguichou a trama	para todos	da equipe	Qual alternativa define melhor o que você entendeu?	João contou um segredo.	João espirrou uma teia.

* As alternativas, assim como a ordem dos itens, foram aleatorizadas. As alternativas em negrito representam a resposta considerada correta.

Apêndice 4: TCLE: Termo de Consentimento Livre e Esclarecido

Figura 6 - Print tela TCLE

progress

Bem-vindo ao experimento

Termo de Consentimento Livre e Esclarecido

Caro participante,

Você está sendo convidado a participar do projeto de pesquisa intitulado "O Papel de Fatores Funcionais no Processamento de Construções de Estrutura Argumental em Português Brasileiro" sediado na FALE/UFMG. Esta pesquisa trata do processamento da linguagem e servirá para entendermos com mais clareza o modo como os falantes do português compreendem a língua. Se você concordar em fazer este teste, sua tarefa consistirá em ler algumas frases e avaliá-las de acordo com sua intuição. Você não estará sendo avaliado nem julgado, apenas nos interessa investigar habilidades que todos os falantes da língua possuem. Portanto, não nos interessa saber se as frases são corretas ou não do ponto de vista da gramática ou da norma culta, mas sim se elas lhe parecem naturais ou não segundo o uso cotidiano que todos os falantes fazem da língua.

Durante a coleta não serão perguntadas questões relacionadas à sua vida pessoal, nem serão tratados assuntos controversos ou delicados. As questões que podem aparecer nos questionários se referem apenas à língua ou sua compreensão dela, mas você também tem o direito de não responder a alguma questão caso ela gere algum constrangimento. Caso você sinta algum desconforto, fique à vontade para interromper a sessão. A participação no estudo é voluntária e você tem a liberdade de se recusar a participar ou interromper a coleta de dados, ou ainda de retirar seu consentimento a qualquer momento, sem que isso lhe cause qualquer prejuízo. Esclarecemos ainda que sua participação não implica nenhum gasto de sua parte ou pagamento da nossa.

Cada sessão dura de 10 a 20 minutos, a depender de sua velocidade individual e das especificidades de cada questionário. A sessão não será gravada, mas os tempos de resposta ficarão registrados no computador de forma codificada, a fim de resguardar sua identidade, que será mantida em sigilo durante todo o processo de coleta e análise de dados. Os resultados desta pesquisa poderão ser apresentados em congressos, artigos e outros textos científicos do gênero, mas as informações individuais dos participantes serão sempre mantidas em sigilo, de forma a garantir-lhe total privacidade.

Agradecemos sua participação e colocamo-nos à disposição para prestar quaisquer esclarecimentos sobre a metodologia empregada neste estudo, tanto antes como depois de sua execução.

Atenciosamente,

Larissa Santos Ciríaco
Profa. Dra. de Linguística da Faculdade de Letras - UFMG

Em caso de dúvidas éticas, entre em contato pelo CEP:

Universidade Federal de Minas Gerais
Av. Presidente Antonio Carlos, 6627, Pampulha
Belo Horizonte - MG - CEP 31270-901
Unidade Administrativa II - 2º Andar - Sala: 2005
Telefone: (031) 3409-4592 - E-mail: coep@prpq.ufmg.br
Horário de atendimento: 09:00 às 11:00 / 14:00 às 16:00

Contato em caso de outras dúvidas:

Laís Marçal de Oliveira
E-mail: lais.marcaloliveira@gmail.com

Larissa Santos Ciríaco
E-mail: laciriac@gmail.com

Clique abaixo se você aceita participar do experimento. Caso não deseje participar, basta fechar esta aba em seu computador.

Li e concordo em participar da pesquisa.

Fonte: PCIBex Farm

Transcrição do termo:

Termo de Consentimento Livre e Esclarecido

Caro participante,

Você está sendo convidado a participar do projeto de pesquisa intitulado “O Papel de Fatores Funcionais no Processamento de Construções de Estrutura Argumental em Português Brasileiro” sediado na FALE/UFMG. Esta pesquisa trata do processamento da linguagem e servirá para entendermos com mais clareza o modo como os falantes do português compreendem a língua. Se você concordar em fazer este teste, sua tarefa consistirá em ler algumas frases e avaliá-las de acordo com sua intuição. Você não estará sendo avaliado nem julgado, apenas nos interessa investigar habilidades que todos os falantes da língua possuem. Portanto, não nos interessa saber se as frases são corretas ou não do ponto de vista da gramática ou da norma culta, mas sim se elas lhe parecem naturais ou não segundo o uso cotidiano que todos os falantes fazem da língua.

Durante a coleta não serão perguntadas questões relacionadas à sua vida pessoal, nem serão tratados assuntos controversos ou delicados. As questões que podem aparecer nos questionários se referem apenas à língua ou sua compreensão dela, mas você também tem o direito de não responder a alguma questão caso ela gere algum constrangimento. Caso você sinta algum desconforto, fique à vontade para interromper a sessão. A participação no estudo é voluntária e você tem a liberdade de se recusar a participar ou interromper a coleta de dados, ou ainda de retirar seu consentimento a qualquer momento, sem que isso lhe cause qualquer prejuízo. Esclarecemos ainda que sua participação não implica nenhum gasto de sua parte ou pagamento da nossa.

Cada sessão dura de 10 a 20 minutos, a depender de sua velocidade individual e das especificidades de cada questionário. A sessão não será gravada, mas os tempos de resposta ficarão registrados no computador de forma codificada, a fim de resguardar sua identidade, que será mantida em sigilo durante todo o processo de coleta e análise de dados. Os resultados desta pesquisa poderão ser apresentados em congressos, artigos e outros textos científicos do gênero, mas as informações individuais dos participantes serão sempre mantidas em sigilo, de forma a garantir-lhe total privacidade.

Agradecemos sua participação e colocamo-nos à disposição para prestar quaisquer esclarecimentos sobre a metodologia empregada neste estudo, tanto antes como depois de sua execução.

Atenciosamente,

Larissa Santos Ciríaco

Profa. Dra. de Linguística da Faculdade de Letras - UFMG

Em caso de dúvidas éticas, entre em contato pelo CEP:

Universidade Federal de Minas Gerais

Av. Presidente Antônio Carlos, 6627, Pampulha

Belo Horizonte - MG - CEP 31270-901

Unidade Administrativa II - 2º Andar - Sala: 2005

Telefone: (031) 3409-4592 - E-mail: coep@prpq.ufmg.br

Horário de atendimento: 09:00 às 11:00 / 14:00 às 16:00

Contato em caso de outras dúvidas:

Laís Marçal de Oliveira

E-mail: lais.marcaloliveira@gmail.com

Larissa Santos Ciríaco

E-mail: laciriaco@gmail.com

Clique abaixo se você aceita participar do experimento. Caso não deseje participar, basta fechar esta aba em seu computador.
