



Programa de Pós-Graduação em
LINGUÍSTICA

**EXPRESSÕES IDIOMÁTICAS DO PORTUGUÊS SOB A PERSPECTIVA DA
GRAMÁTICA DISCURSIVO-FUNCIONAL**



SÃO CARLOS

2025



**EXPRESSÕES IDIOMÁTICAS DO PORTUGUÊS SOB A PERSPECTIVA DA
GRAMÁTICA DISCURSIVO-FUNCIONAL**

DIOGO OLIVEIRA DA SILVA

Bolsista: CAPES

Tese de doutorado apresentada ao Programa de Pós-Graduação em Linguística da Universidade Federal de São Carlos, como parte dos requisitos para a obtenção do Título de Doutor em Linguística (Área de concentração: Descrição, análise e processamento automático de línguas naturais).

Orientadora: Prof.^a Dr.^a. Joceli Catarina Stassi Sé

Programa de Pós-Graduação em Linguística/ UFSCar

São Carlos – São Paulo – Brasil

2025

SILVA, Diogo Oliveira da*.

Expressões idiomáticas do português soba perspectiva da Gramática Discursivo-Funcional.
São Carlos: UFSCar, 2025.

193 páginas.

Tese de Doutorado - Universidade Federal de São Carlos, Centro de Educação e Ciências Humanas, Programa de Pós-Graduação em Linguística, 2025.

Orientadora: Profa. Dra. Joceli Catarina Stassi Sé (UFSCar)

1. Expressões idiomáticas; 2. Motivação; 3. Composicionalidade;
4. Gramática Discursivo-Funcional.

*UFSCar Universidade Federal de São Carlos. Programa de Pós-Graduação em Linguística (PPGL) em nível de doutorado. E-mail: diogosilva@estudante.ufscar.br



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Educação e Ciências Humanas
Programa de pós graduação em Linguística

Folha de Aprovação

Defesa de Tese de Doutorado do candidato Diogo Oliveira da Silva, realizada em 08/08/2025.

Comissão Julgadora:

Profa. Dra. Joceli Catarina Stassi Sé (UFSCar)

Profa. Dra. Talita Storti Garcia (UNESP)

Prof. Dr. Michel Gustavo Fontes (UFMS)

Prof. Dr. Edson Rosa Francisco de Souza (UNESP)

Prof. Dr. Oto Araujo Vale (UFSCar)

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa de Pós-Graduação em Linguística.

RESUMO: O objetivo geral desta tese é investigar, à luz da Gramática Discursivo-Funcional (GDF) (Hengeveld; Mackenzie, 2008) e dos estudos de Keizer (2016, 2018), o estatuto semântico-cognitivo de expressões idiomáticas (EIs) do português, utilizando, como princípios norteadores, a motivação, para uma análise cognitiva, e a composicionalidade, para uma análise morfossintática, a fim de definir graus de transparência e de opacidade das EIs, com a finalidade de propor uma tipologia de EIs para o português. Para alcançar esse objetivo, aplica-se, com base em Keizer (2016, 2018), um conjunto de quatorze critérios de variação morfossintática, quais sejam: substituição lexical, modificação adjetiva, quantificação, número, definitude, passivização, nominalização, referência anafórica, referência catafórica, inserção de advérbios, inserção de pronomes, inversão da ordem preferida, elisão de um termo fixo e possibilidade de conversão para modo imperativo. A partir de uma compilação inicial de EIs já dicionarizadas (Riva, 2009; Silva, 1997; Almeida, 2022), fez-se o levantamento de dados, que levou em conta o *Corpus* de Referência do Português Contemporâneo (2016), registros escritos da *web* e dados construídos. Tendo em vista esse universo de investigação, busca-se, como objetivos específicos: (i) verificar como se dá o processo de esforço extra cognitivo do falante, envolvido na reanálise das EIs; (ii) averiguar as EIs conforme o alinhamento entre os níveis Interpessoal, Representacional e Morfossintático da GDF; e (iii) determinar como ocorre a violação da relação de um-para-um entre forma e significado dos elementos constituintes das EIs (Leufkens, 2015). Após a aplicação dos critérios, a realização da análise cognitiva e morfossintática das ocorrências, a verificação de seu grau de transparência e do alinhamento entre os níveis do modelo, foi possível identificar, no português, além dos quatro tipos de EIs encontradas no inglês por Keizer (2016), uma nova tipologia, compondo o seguinte conjunto: (i) desmotivadas e não-decomponíveis; (ii) motivadas e não-decomponíveis; (iii) motivadas e decomponíveis; (iv) expressões *AD-HOC*, com uso específico; e o último tipo, (v) parcialmente motivadas por metáforas conceituais e decomponíveis. Além da identificação desse quinto tipo de EI no português, esta pesquisa também agrega à discussão de Keizer (2016) os casos de EIs *AD-HOC*, envolvendo descontinuidade e dependência de contexto e atesta a aplicabilidade do conjunto de primitivos CombPIF – Combinação de quadros parcialmente instanciados (Keizer, 2016), agora com resultados em EIs de variedades lusófonas do português, o que pode contribuir para a ampliação de horizontes de investigação sobre a interpretação metafórica das EIs à luz da GDF.

PALAVRAS-CHAVE: Expressões idiomáticas; Motivação; Composicionalidade; Gramática Discursivo-Funcional.

ABSTRACT: The general aim of this thesis is to investigate, in light of the Functional Discourse Grammar (FDG) (Hengeveld; Mackenzie, 2008) and studies by Keizer (2016, 2018), the semantic-cognitive status of idiomatic expressions (IEs) in Portuguese, using, as guiding principles, motivation, for a cognitive analysis, and compositionality, for a morphosyntactic analysis, in order to define degrees of transparency and opacity of IEs, with the purpose of proposing a typology of IEs for Portuguese. To achieve this aim, based on Keizer (2016, 2018), a set of fourteen morphosyntactic variation criteria is applied, namely: lexical substitution, adjectival modification, quantification, number, definiteness, passivization, nominalization, anaphoric reference, cataphoric reference, insertion of adverbs, insertion of pronouns, inversion of the preferred order, elision of a fixed term and possibility of conversion to imperative mode. From an initial compilation of IEs already included in dictionaries (Riva, 2009; Silva, 1997; Almeida, 2022), data collection took place, which took into account the Reference Corpus of Contemporary Portuguese (2016), written records from the web and constructed data. In view of this universe of research, the specific aims sought are: (i) to verify how the speaker's extra cognitive effort process functions, involved in the reanalysis of IEs; (ii) to check the IEs according to the alignment among the Interpersonal, Representational and Morphosyntactic levels of the FDG; and (iii) to determine how the violation of the one-to-one relationship between form and meaning of the constituent elements of the IEs occurs (Leufkens, 2015). After applying the criteria, performing the cognitive and morphosyntactic analysis of the occurrences, verifying their degree of transparency and the alignment among the levels of the model, it was possible to identify, in Portuguese, in addition to the four types of IEs found in English by Keizer (2016), a new typology, comprising the following set: (i) unmotivated and non-decomposable; (ii) motivated and non-decomposable; (iii) motivated and decomposable; (iv) AD-HOC expressions, with specific use; and the last type, (v) partially motivated by conceptual metaphors and decomposable ones. In addition to identifying this fifth type of IE in Portuguese, this research also adds to Keizer's (2016) discussion of cases of AD-HOC IEs, involving discontinuity and context dependence, and attests to the applicability of the set of primitives CombPIF – Combination of partially instantiated frames (Keizer, 2016), now with results in IEs of Lusophone varieties of Portuguese, which can contribute to expanding research horizons on the metaphorical interpretation of IEs in light of the FDG.

KEYWORDS: Idioms; Motivation; Compositionality; Functional Discourse Grammar.

SUMÁRIO

0.1: Introdução.....	16
0.1.1: Universo de investigação.....	17
0.1.2: Objetivos de pesquisa.....	19
0.1.3: Proposta teórico-metodológica.....	22
0.1.4: Problemáticas e hipóteses.....	23
0.1.5: Metodologia da pesquisa.....	24
0.1.6: Resultados esperados.....	26
0.1.7: Justificativa.....	26
0.1.8: Organização da tese.....	27
CAPÍTULO 01: AS EXPRESSÕES IDIOMÁTICAS NA LITERATURA LINGUÍSTICA	29
1.1: EIs: definições gerais.....	31
1.2: Idiomatismos na Gramática Construcional e no Modelo de linguagem baseada no uso...36	
1.3: Expressões cristalizadas e interpretação metafórica.....	39
1.4: Aprendizagem das EIs.....	52
CAPÍTULO 02: A GRAMÁTICA DISCURSIVO-FUNCIONAL.....	54
2.1: Nível Interpessoal (NI).....	60
2.2: Nível Representacional (NR).....	65
2.3: Nível Morfossintático (NM).....	68
2.3.1: Ordenação no NM.....	70
2.4: Nível Fonológico (NF).....	73
2.5: EIs para a GDF.....	76
2.5.1: Duas visões opostas (estrutura interna).....	78
2.5.1.1: Visão não-composicional (baseada na forma).....	78
2.5.1.2: Visão composicional (baseada no significado).....	79
2.5.2: Uma análise da GDF sobre as EIs.....	80
2.5.3: Três tipos de EIs na GDF com base nas variações formais motivadas por aspectos semântico-cognitivos.....	81

2.5.4: Transparência linguística.....	88
2.5.5: Transparência e opacidade das EIs.....	90
2.5.6: Como representar EIs no Léxico pensando na distinção Léxico x Gramática.....	92
2.5.7: Referências anafórica e catafórica para a GDF.....	98

CAPÍTULO 03: FUNDAMENTOS TEÓRICO-METODOLÓGICOS.....101

3.1: Metodologia de pesquisa.....	101
3.1.1: Problemáticas e hipóteses da tese.....	101
3.1.2: Critérios e fatores de análise.....	104
3.2: Banco de dados de EIs.....	107
3.2.1: Dicionário onomasiológico das EIs.....	107
3.2.2: Dicionário sobre a origem das palavras e expressões da língua portuguesa.....	109
3.2.3: Dicionário de locuções e expressões da língua portuguesa.....	111
3.2.4: Outros bancos de dados de EIs: consulta secundária.....	112
3.3: Materiais de busca de EIs: <i>corpora</i> de análise.....	115
3.3.1: <i>Corpus</i> de Referência do Português Contemporâneo (CRPC), dados da <i>web</i> e dados construídos.....	115
3.4: Análise quantitativo-qualitativa.....	117

CAPÍTULO 04: ANÁLISE DAS EIS DO PORTUGUÊS.....126

4.1: Análise de dados.....	126
4.1.1: EIs do Tipo 1.....	126
4.1.2: EIs do Tipo 2.....	133
4.1.3: EIs do Tipo 3.....	140
4.1.4: EIs do Tipo 4.....	147
4.1.5: EIs do Tipo 5.....	154
4.2: Resultados finais.....	166
Considerações finais.....	179
Referências Bibliográficas.....	187

Lista de quadros:

Quadro 1: Definição geral das EIs (Elaboração própria, 2025).....	34
Quadro 2: Comportamento morfossintático dos três tipos de EIs (Keizer, 2016, p. 1001) Adaptado para português (Elaboração própria, 2025).....	84
Quadro 3: Ativação semântica das EIs (Elaboração própria, 2025).....	95
Quadro 4: Fundamentação metodológica da pesquisa (Elaboração própria, 2025).....	106
Quadro 5: Variedades abrangidas pelo CRPC e número de <i>tokens</i> por registro (CLUL, 2016).....	115
Quadro 6: EIs selecionadas (Elaboração própria, 2025).....	119
Quadro 7: Comportamento morfossintático das EIs em português (Elaboração própria, 2025).....	167
Quadro 8: Representação do alinhamento das EIs selecionadas (Elaboração própria, 2025).	173
Quadro 9: Fundamentação metodológica da pesquisa - perguntas e respostas finais (Elaboração própria, 2025).....	183

Lista de figuras:

Figura 1: Estrutura simbólica de uma construção (Croft, 2007, p. 472) Adaptado para português (Elaboração própria, 2025).....	37
Figura 2: Esquema representativo da estrutura prototípica (Teixeira, 2020, p. 29).....	42
Figura 3: Estrutura prototípica da EI <i>descascar o abacaxi</i> (Elaboração própria, 2025).....	43
Figura 4: Nova arquitetura da GDF (Hengeveld et al., no prelo).....	56
Figura 5: Hierarquia no Nível Interpessoal (Hengeveld; Mackenzie, 2010, p. 18).....	62
Figura 6: Hierarquia no Nível Representacional (Hengeveld; Mackenzie, 2010, p. 21)	65
Figura 7: Hierarquia no Nível Morfossintático (Hengeveld; Mackenzie, 2010, p. 26).....	68
Figura 8: Hierarquia no Nível Fonológico (Hengeveld; Mackenzie, 2010, p. 30).....	73
Figura 9: Graduação dos tipos de EIs (Elaboração própria, 2025).....	96
Figura 10: Descontinuidade na EI <i>Fazer tempestade num copo de água</i> (Elaboração própria, 2025).....	164

Lista de siglas e abreviaturas:

A - Actor

A - Adjetivo

A – Ato Discursivo

ADMIR -Admirativo

Adp- Sintagma Adposicional

Advp – Sintagma Adverbial

Aff - Afixo

Ang – Angola

Ap – Sintagma Adjetival

BADAF – Base de Dados Fraseológica do Português Europeu

Bra – Brasil

C – Conteúdo Comunicado

Cl - Oração

CLUL – Centro de Linguística da Universidade de Lisboa

CombPIF – Combination of Partially Instantiated Frames

con- Contraste

COP - Cópula

CRPC – *Corpus* de Referência do Português Contemporâneo

CV – Cabo Verde

DECL – Declarativa

DEF – Definido

DS – Distância semântica

e – Estado-de-Coisas

El – Expressão idiomática

emph - Ênfase

ep – Episódio

EXPL - Explicação

f – Propriedade Configuracional

F – Ilocução

F – Pé
Foc - Foco
GB – Guiné-Bissau
GDF – Gramática Discursivo-Funcional
GF – Gramática Funcional
Gw – Palavra Gramatical
Iloc – Ilocucionário
IMPER – Imperativa
INDEF - Indefinido
INTER – Interrogativa
IP – Sintagma Entonacional
l – Lugar
L2 – Segunda língua
Le – Expressão Linguística
Loc – Locativo
m - Plural
M – Movimento
Moç – Moçambique
Mot - Motivação
M. ULN – Modelo de Usuário de Língua Natural
NF – Nível Fonológico
NI – Nível Interpessoal
NIL – Sem data registrada (ficha técnica dos excertos do CRPC)
NM – Nível Morfossintático
Np– Sintagma Nominal
NR – Nível Representacional
Num – Numeral
N/A – Não aplicado/aplicável
O – Objeto
p – Conteúdo Proposicional
P_A – Ouvinte
past - Passado
P^{centro} – Posição Centro da Oração

P^F – Posição Final
P^{F-1} – Posição Penúltima
P^I – Posição Inicial
P^{I+1} – Posição Pós-Inicial
P^M – Posição Medial
P^{M+1} – Posição Pós-Medial
PMW – *per million words*/palavras por milhão
pp – Frase Fonológica
P^{pre} – Posição Pré-Oracional
P^{pos} – Posição Pós-Oracional
prs - Presente
P_s - Falante
PT – Portugal
pw – Palavra Fonológica
q – Quantidade
r - Razão
R – Subato Referencial
Ref - Referência
s - Sílabas
SA – Subato
SG - Singular
Sub – Subordinação
subj - Sujeito
t - Tempo
T – Subato Atributivo
To-Pr – São Tomé e Príncipe
U - Undergoer
U - Enunciado
V – Verbo
Vp – Sintagma Verbal
Xp - Sintagma
Xs - Raiz
Xw - Palavra

x- Individuo

0.1: Introdução

Esta tese investiga expressões idiomáticas (EIs) do português, tais como: *descascar o abacaxi*, *procurar sarna para se coçar*, e *fazer tempestade em copo d'água*, compiladas a partir da consulta a dicionários de EIs (Riva, 2009; Silva, 1997; e Almeida, 2022). O universo de investigação compreende contextos reais de fala e de escrita e leva em conta o *Corpus* de Referência do Português Contemporâneo (2016), registros escritos da *web* e dados construídos. A análise segue uma perspectiva funcionalista, baseada nos trabalhos de Keizer (2016, 2018) sobre *idioms* do inglês e no modelo teórico da Gramática Discursivo-Funcional (GDF) de Hengeveld e Mackenzie (2008).

Expressões idiomáticas são aqui definidas, com base em Keizer (2016, 2018), como expressões semifixas, compostas por um verbo e um sintagma nominal ou preposicional, que apresentam significados idiomáticos os quais, muitas vezes, não correspondem ao sentido literal de suas partes, como em *to take the cake* (ser o pior), *to bring home the bacon* (sustentar a casa) e *to open the floodgates* (conceder exceções).

Keizer (2016) atesta que as EIs podem ser analisadas também sob o aporte teórico da GDF, pois há uma relação sistemática entre funções pragmáticas e semântico-cognitivas que influenciam o comportamento morfossintático dessas construções e sua representação na gramática das línguas, o que se busca também comprovar aqui com a aplicação da proposta da autora ao português.

Importante ressaltar que o objetivo desta investigação não é de cunho sociolinguístico nem comparativo, já que não se busca comparar usos entre as variedades do português, mesmo que o *corpus* de língua falada – *Corpus* de Referência do Português Contemporâneo - forneça dados que permitam mapear o perfil dos falantes, tais como contextualização histórica em determinados registros, idade, sexo, escolaridade, etc., de diferentes países falantes de português como língua oficial, tais como Angola, Brasil, Portugal, São Tomé e Príncipe, Cabo Verde, Moçambique, Guiné-Bissau e Timor Leste, além de registros de Goa e Macau (CLUL, 2016), países em que o português não é o idioma oficial.

Assim, ao identificar que duas EIs possuem a mesma significação idiomática, são dicionarizadas, podem ser encontradas em variedades diferentes e ocorrer em diversos contextos de uso (debates políticos, reportagens, diálogos informais carregados de sarcasmo, ironia, etc.) entende-se que tais contextos podem resultar em novas variações morfossintáticas e novas significações idiomáticas, o que não invalida sua significação idiomática dicionarizada e também não anula as demais formas encontradas entre as variedades do português. Dessa forma, não se consideram, como

únicas ou homogêneas, as variações morfossintáticas encontradas, por exemplo, em *untar as mãos e molhar a mão* (propina), e sim entende-se que ambas as EIs podem sofrer diferentes processos de variação morfossintática em usos específicos, sem perder seu valor idiomático dicionarizado.

0.1.1: Universo de investigação

As EIs são tratadas no campo da Fraseologia como unidades lexicais fraseológicas idiomatizadas que, em conjunto, perdem o sentido literal e adquirem um novo sentido semântico, como ocorre com provérbios e fórmulas rotineiras (Xatara, 1998).

Considerando-se, assim, esses aspectos conceituais, o universo de investigação desta pesquisa compreende duas etapas: (i) seleção de EIs no português a partir de dicionários (Riva, 2009; Silva, 1997; e Almeida, 2022), tratados aqui como banco de dados que atestam que determinada expressão é considerada idiomática através do seu registro formal, frequência e significação idiomática; e (ii) levantamento de EIs nos *corpora*: o *Corpus* de Referência do Português Contemporâneo, doravante CRPC, a principal fonte de busca e dados da *web*, a segunda fonte de busca, acessada conforme a falta de ocorrências no CRPC, ambos tratados aqui como *corpora* em que os dados aparecem em situações de uso, em que identificamos os excertos de inserção, contextualização e significação idiomática efetivamente ativados. Além disso, utilizam-se dados construídos, com fins de discutir possibilidades de **variação morfossintática**¹ em ocorrências em que EIs aparecem somente da forma consagrada nos *corpora* e nos interessa aferir possibilidades de variação.

Dentre os dicionários que serviram como banco de dados para consulta e compilação de EIs consagradas, mencionamos, como principal fonte, o Dicionário Onomasiológico de expressões idiomáticas usuais no português do Brasil (Riva, 2009), e como fontes secundárias, as obras de Silva (1997), que versa sobre a origem das palavras e expressões da língua portuguesa, e o Dicionário aberto de calão e expressões idiomáticas da língua portuguesa de Almeida (2022), obras nas quais encontramos definições do objeto de estudo (EIs) e suas respectivas classificações e exemplificações de usos, como, por exemplo, a definição do termo *abacaxi* e a origem sócio-histórica da expressão *levar um puxão de orelhas*, respectivamente expostas em (0-1) e (0-2), encontradas em Silva (1997):

- (0-1) Abacaxi: do tupi, *ibacati*, pela composição *ibá*, fruto, e *cati*, cheirar muito. Apesar de ser fruto saboroso, logo virou metáfora de coisa difícil. Deixar o abacaxi nas mãos de alguém é passar-lhe tarefa incômoda. Incumbi-lo de descascá-lo é pior ainda. (Silva, 1997, p. 11)

¹ Utilizamos essa terminologia presente no trabalho de Keizer (2016), entendendo que toda variação morfossintática é o reflexo codificado das operações da Formulação, seja de natureza pragmático-discursiva e/ou semântica.

- (0-2) Levou um puxão de orelhas: A origem desta frase, expressão que significa repreender, está ligada a antigas tradições populares, que a recolheram de usos e costumes nem sempre vagos, já que inspirados também em documentos jurídicos. As Ordenações Afonsinas prescrevem que os ladrões tenham as orelhas cortadas. O grande navegador português Vasco da Gama (1469-1524) relatou o corte de 800 delas. E Gomes Freire de Andrade, *o conde de Bobadela* (1685-1763), governador e capitão-geral do Rio de Janeiro, Minas Gerais e São Paulo, personagem do filme *Xica da Silva*, de Carlos Diegues, recebeu 7.800 delas. Depois as orelhas deixaram de ser cortadas e foram somente puxadas. Por fim, tudo virou apenas metáfora de admoestação. (Silva, 1997, p. 15)

O CRPC é organizado pelo Centro de Linguística da Universidade de Lisboa e disponibilizado via plataforma digital de acesso público (CLUL, 2016). Essa plataforma permite a filtragem de ocorrências por palavra, apresenta sua classe gramatical e número de ocorrências, além de conter a descrição do perfil dos participantes (informantes) e o *background* das cenas comunicacionais, com dados que nos auxiliam a compreender fatores contextuais importantes para a análise e uso das EIs.

Os dados buscados na *web* foram levantados com o intuito de complementar a pesquisa com exemplos que constavam nos dicionários como EIs, mas não foram encontrados no CRPC.

Já os dados construídos foram elaborados para atestar a funcionalidade (ou não) de casos de variação morfossintática, a título de discussão com contra exemplos, como é o caso da substituição lexical de *água* por *café* na EI *fazer tempestade num copo de água/fazer tempestade num copo de café*, em que se nota não haver registros desse tipo de EI nos dicionários, não constituindo uma EI comprovadamente, não sendo possível levantá-la no CRPC, nem em buscas na *web*, uma vez que não se aponta tal uso (*café*) como viável.

Salienta-se que esta pesquisa não objetiva investigar a origem etimológica de cada expressão, nem as razões para as variações lexicais que possam ocorrer entre EIs de diferentes variedades. Traduções entre o português e o inglês também estão fora do escopo da pesquisa, embora tenham sido encontrados dados que permitam comparações entre as expressões nas duas línguas.

Foi delimitada a quantidade de 35 EIs para análise, oriundas de um grupo de 60 EIs previamente coletadas em dados da *web* obtidas com o filtro “expressões idiomáticas usuais e comuns no português”. Após esse primeiro momento de coleta, alguns aspectos foram observados, entre eles: quais EIs estavam registradas nos dicionários citados acima, e qual sua relevância em frequência nos *corpora* CRPC e na própria *web*, para assim refinar a amostra e definir quais EIs seriam selecionadas para a investigação final aqui proposta.

Também foi utilizado, como critério de seleção, a presença de forma e significado propenso à variação morfossintática, o que auxiliou na compreensão da motivação e da composicionalidade das EIs. Esse processo teve como base a leitura e interpretação dos *idioms* que Keizer (2016) também

selecionou em sua proposta, pois, embora as línguas envolvidas sejam línguas distintas (inglês e português), essa comparação inicial entre EIs nos trouxe certa clareza para delimitar o objeto de estudo em uma quantidade satisfatória.

Um exemplo prático desse refinamento na coleta da amostra final de EIs, executado com base na seleção vista em Keizer (2016), foi comparar EIs com a mesma forma e significado idiomático distinto nas duas línguas - *to kick the bucket* (morrer) e *chutar o balde* (desistência, impaciência), assim como EIs com o mesmo valor idiomático e diferentes formas – *to grasp the nettle e descascar o abacaxi* (enfrentar/resolver um problema). Esse embasamento auxiliou na compreensão de quais EIs do português poderiam ser mais ou menos produtivas, assim como sua apresentação no CRPC em casos de variação morfossintática encontrados durante a investigação dos dados.

0.1.2: Objetivos de pesquisa

O objetivo geral da pesquisa é investigar, com base nos preceitos teóricos da GDF (Hengeveld; Mackenzie (2008), Keizer (2016, 2018)), o estatuto semântico-cognitivo das EIs encontradas em variedades lusófonas do português, utilizando, como princípios norteadores, o conceito de motivação, para uma análise cognitiva, e o de composicionalidade, para uma análise morfossintática², a fim de definir graus de transparência e opacidade das EIs, para propor uma tipologia dessas expressões no português, aplicando a proposta de Keizer (2016).

Assim, com base em Keizer (2016), faz-se importante conceituar ambos os princípios norteadores e como eles se relacionam: primeiramente, a **motivação** é compreendida como a facilidade com que o falante identifica o sentido da expressão e seu respectivo uso em caráter conotativo. Entende-se que os processos cognitivos que levam o falante a compreender, reconhecer, associar, imaginar, pensar e perceber uma construção como idiomática perpassam por duas leituras: literal e conotativa.

Para compreender como o falante diferencia ambas as leituras, apoiamo-nos nos conceitos de interpretação metafórica (Lakoff; Johnson, 1980), cristalização das EIs (Vale, 2001), metáfora conceitual (Langlotz, 2006), metáfora cultural (Kövecses, 2010), metáfora convencional (Vega Moreno, 2007), entre outros estudos que investigam a experiência de mundo do falante sobre as EIs, a sua exposição a esse fenômeno na língua, a repetição e o uso das EIs, a codificação e a decodificação dessas expressões, o processo de cristalização, as metáforas universais, etc.

² A etapa de análise morfossintática apresentará os resultados da variação na forma das EIs, que aqui é entendida como a codificação das operações dos níveis de Formulação.

O segundo princípio norteador é a **composicionalidade**, entendida como a verificação dos aspectos específicos do significado idiomático que podem ser relacionados às partes componentes das EIs, atuando na ligação significativa das partes da expressão da cena literal com a cena idiomática. A análise morfossintática aqui proposta das partes componentes das EIs está pautada em dois momentos de investigação: primeiro, na definição dicionarizada da forma consagrada, através dos registros dos bancos de dados de EIs, e segundo, na investigação da forma em uso, verificada no CRPC e na *web*, através da aplicação de critérios de variação morfossintática, detalhados a seguir, com o intuito de definir diferentes tipos de EIs no português, de acordo com seus respectivos graus de transparência e opacidade.

É neste momento que ambos os princípios norteadores (motivação e composicionalidade) se relacionam: para avaliar se uma EI é transparente ou opaca, é necessário entender seu grau de motivação e o seu grau de composicionalidade. Por exemplo, se uma EI é desmotivada (opaca) e não-decomponível (não é possível depreender o significado da EI pelas partes componentes), significa que possui um grau baixo de motivação, i.e. o falante não reconhece seu significado idiomático com facilidade, além de não conseguir depreender a contribuição do significado literal das partes componentes para o significado idiomático.

Keizer (2016) utiliza a EI *to kick the bucket* (morrer) para ilustrar o tipo não-motivado e não-decomponível: é muito difícil associar a cena literal - *chutar o balde* e o significado idiomático - *morrer*, pois não há uma conexão clara entre ambas as leituras literal e idiomática, assim como é difícil compreender como os termos *to kick*, *the* e *bucket* contribuem individualmente para a formação do significado idiomático de morte. Esse tipo de EI possui um comportamento morfossintático mais restrito, o que impossibilita resultados satisfatórios de variação morfossintática, como por exemplo, modificação adjetiva (*to kick the empty bucket**).

Dessa forma, para que seja possível alcançar o objetivo geral, que compreende uma análise semântico-cognitiva dessas expressões, elenca-se três objetivos específicos: (i) verificar como se dá os processos de esforço extra cognitivo do falante, envolvidos na reanálise das EIs *AD-HOC*; (ii) averiguar as EIs conforme o alinhamento entre os níveis Interpessoal, Representacional e Morfossintático da GDF; e (iii) determinar como ocorre a violação da relação de um-para-um entre forma e significado dos elementos constituintes das EIs (Leufkens, 2015).

O primeiro objetivo específico diz respeito a verificar como se dá os processos de esforço extra-cognitivo do falante para reanalisar novas expressões *AD-HOC*, com base no processo de mudança de não-decomponibilidade para decomponibilidade das expressões (Keizer, 2018), processo em que o falante reanalisa o novo elemento lexical da expressão (e.g. *to bite the bullet/to gnaw the*

bullet) com base em sua contribuição nas cenas literal e idiomática para a expressão previamente convencionalizada.

O segundo objetivo específico relaciona-se a averiguar o alinhamento das representações das EIs, conforme as regras de formulação da GDF, considerando-se os três níveis mais altos da teoria. Nessa etapa, verifica-se o comportamento de cada EI nos níveis Interpessoal (NI), Representacional (NR) e no Morfossintático (NM). Neste trabalho, o Nível Fonológico (NF) não será analisado, mesmo que Keizer (2016) o tenha incluído em sua análise, pois entende-se que os resultados encontrados nos níveis NI, NR e NM foram suficientes para compreendermos os aspectos formais e funcionais do fenômeno aqui investigado. Deixamos essa proposta como investigação futura.

Após a verificação de quais camadas dos níveis da GDF são determinantes para a análise das EIs, e.g. se a EI representa um Subato (S) no NI, se é um Estado-de-Coisas (e) no NR, etc., averigua-se a existência de um possível padrão de alinhamento entre esses três níveis. Tal objetivo foi motivado pela existência de um alinhamento padrão das EIs, conforme Keizer (2016), que obteve tais resultados após comparar diferentes tipos de EIs no inglês. Caso haja um padrão de alinhamento a partir da análise de cada nível, ou seja, relação de um-para-um entre os níveis de Formulação e de Codificação, trataremos essa relação como transparente.

Um exemplo desse tipo de alinhamento acontece com a EI da língua inglesa *to kick the bucket*: ela é representada por um Subato Atributivo (T) no NI, um Estado-de-Coisas (e) no NR e três Palavras no NM (uma Palavra Verbal (Vw), uma Palavra Nominal (Nw) e uma Palavra Gramatical (Gw)). Assim, temos como resultados de alinhamento: transparência entre NI e NR (relação um-para-um) e opacidade entre NI e NM (relação um-para-três) e entre NR e NM (relação um-para-três).

Tal relação entre termos argumentais (sujeitos, verbos e objetos) e sua respectiva representação em diferentes níveis leva em conta a natureza idiomática das EIs. *To kick the bucket* significa morrer. Conforme veremos mais adiante, o argumento *bucket* seria visto como paciente do verbo *to kick* (alguém chuta o balde), de acordo com a cena literal, mas não há essa relação entre sujeito-verbo-objeto na significação idiomática – alguém morre.

Já o terceiro objetivo específico, também conectado com a dimensão da transparência e da opacidade, diz respeito à aplicação de um dos critérios de violação da relação um-para-um entre forma e significado, conforme trabalho de Leufkens (2015). Nesse âmbito, busca-se determinar como ocorre os casos de descontinuidade - divisão de uma unidade formal, resultando em um novo alinhamento não-paralelo entre os níveis, fenômeno encontrado nas ocorrências de referências anafórica e catafórica, esta de caráter inovador para esse tipo de investigação, identificadas nas camadas do

Movimento e também do Discurso (Stassi Sé, 2012), devido à extensão de Movimentos realizados em troca de turno entre falantes. Assim como Leufkens (2015), enfoca-se aqui em fenômenos de não-transparência gramatical, e não-transparência lexical, tais como homonímia, sinonímia e polissemia.

Como se pode ver, o conceito de transparência e de opacidade é bastante caro à pesquisa, uma vez que é investigado como os graus de motivação e de composicionalidade influenciam na definição dos graus de transparência e opacidade das EIs, fazendo transparecer uma relação implicacional importante para a caracterização dessas expressões: quanto maior o grau de motivação e composicionalidade de uma EI, maior será seu grau de transparência em termos semântico-cognitivos.

0.1.3: Proposta teórico-metodológica

Tratados como fonte de inspiração para esta tese, os trabalhos de Keizer (2016, 2018) investigam as EIs no inglês, extraídas dos *corpora COCA* “*The corpus of contemporary American English*” (1990-2015) e “*NOW Corpus*” (*News on the Web*) (2010), que incluem diferentes gêneros divididos em ocorrências de língua falada e escrita. Além disso, a autora utiliza um apêndice referente às EIs encontradas e a suas definições, com o intuito de compreender o comportamento formal e funcional dessas expressões com base na motivação e na composicionalidade, e em sua representação na GDF.

A autora classifica as EIs do inglês em diferentes tipos: (i) EIs não-motivadas, opacas e semanticamente não-decomponíveis, e.g. *to kick the bucket* (Keizer, 2016); (ii) EIs motivadas, transparentes e semanticamente não-decomponíveis, e.g. *to hit the ceiling* (Keizer, 2016); (iii) EIs motivadas, mais transparentes e decomponíveis, e.g. *to spill the beans* (Keizer, 2016); e (iv) novas EIs *AD-HOC*, e.g. *to gnaw the bullet* (Keizer, 2018). A partir do que é proposto por Keizer (2016, 2018), este trabalho busca analisar quais tipos de EIs ocorrem no português.

A autora realizou oito testes formais em sua análise, quais sejam: modificação adjetiva, substituição lexical, passivização, nominalização, definitude, número, quantificação e anaforicidade, para analisar o estatuto semântico-cognitivo das EIs, a fim de compreender como a motivação e composicionalidade auxiliam no processo de definição de graus de transparência e opacidade das EIs.

Com base em outros autores, como Nunberg et al. (1994) e Langlotz (2006), a autora definiu os quatro tipos de expressões citados acima e, ao estratificar tais tipos na análise de níveis e camadas

da teoria, propôs a inclusão de um novo conjunto de primitivos, chamado CombPIF – *Combination of Partially Instantiated Frames* - Combinação de quadros parcialmente instanciados, como parte integrante do quadro geral da GDF.

Tal proposta fundamenta-se no funcionamento das EIs como expressões semifixas parcialmente imprevisíveis (lexicalmente) e parcialmente produtivas (gramaticalmente), e que, por serem formuladas nos quadros interpessoais e representacionais da teoria, e codificadas em modelos morfossintáticos, devem ser armazenadas no léxico, sabendo que os lexemas separados de cada EI já estão armazenados no léxico, com base na relação forma-significado, enquanto unidade isolada, como entradas parcialmente “seladas” (instanciadas), associadas ao significado convencional enquanto EI, como por exemplo, *descascar o abacaxi* vista em (0-1).

Para este trabalho, mantém-se os oito testes formais da autora, denominados aqui de critérios de variação morfossintática, e adicionam-se a eles outros seis critérios, que se mostraram relevantes para a análise das ocorrências do português, quais sejam: inserção de pronomes, inserção de advérbios, referência catafórica, possibilidade de conversão para modo imperativo, elisão de um termo fixo da EI e inversão da ordem preferida.

Destaca-se que, nas ocorrências em que as EIs apresentam algum tipo de alteração com relação à forma consagrada, é considerada a forma alterada, conforme a aparição no CRPC, nosso *corpus* principal e primeira fonte de busca; o mesmo ocorre com os excertos da *web*, nossa segunda fonte de busca. Já nos casos de expressões não encontradas nos *corpora*, tem-se em mãos apenas sua forma consagrada e convencionalizada nos dicionários como referência, fazendo-se necessária a adição de elementos de variação morfossintática nas EIs nas discussões dos dados, constituindo o que chamamos aqui de “dados construídos”, nossa terceira “fonte”, se assim podemos denominá-la. Cumpre ressaltar que esse procedimento metodológico é aqui replicado conforme Keizer (2016), que realizou testes em dados construídos, com base inferencial, quando os mesmos não foram encontrados com variações morfossintáticas nos *corpora*, e.g. *to kick the empty bucket*.

0.1.4: Problemáticas e hipóteses

Os objetivos foram arrolados a fim de responder às seguintes problemáticas: (i) É possível verificar uma relação sistemática entre funções pragmáticas e semântico-cognitivas que influenciam o comportamento morfossintático de EIs e sua representação na gramática das línguas? Em caso afirmativo, (ii) Como se dá a caracterização dessas expressões no português, sob a óptica da GDF,

uma teoria funcional-estrutural (Butler, 2005), que leva em consideração a função, mas tem como foco primordial fenômenos que manifestam distinção na codificação, ou seja, na estrutura das unidades linguísticas? (iii) Como se avalia a transparência e a opacidade dessas expressões em situações de uso, considerando-se o significado que assumem e o papel dos Participantes na compreensão do sentido em diferentes contextos (Leufkens, 2015; Keizer, 2016, 2018)? (iv) Como as interpretações metafóricas de EIs podem ajudar no processo de ativação semântica (Langlotz, 2006) e de composicionalidade de suas partes?

Como primeira hipótese a essas problemáticas, partimos do pressuposto de que se observam no português, funções pragmáticas e semântico-cognitivas impactando as características morfossintáticas dessas expressões, sendo possível encontrar, à luz da GDF, uma sistematização para as EIs no português, conforme os quatro tipos de *idioms* que Keizer (2016, 2018) classificou no inglês, tendo como norte de investigação os princípios de motivação e composicionalidade.

Como segunda hipótese, acredita-se que as EIs do português podem apresentar resultados de variação morfossintática similares aos vistos no trabalho da autora, além de outros resultados diferentes, uma vez que o português pode apresentar algumas especificidades com relação à estruturação de EIs. A terceira hipótese, levantada durante as investigações, relaciona-se à ocorrência de violação da relação entre forma e significado em EIs, o que também poderia ser avaliado e sistematizado dentro do modelo da teoria da GDF, levando em conta os trabalhos de Leufkens (2015) sobre transparência e opacidade em diferentes línguas e de Keizer (2016, 2018) sobre a descrição de EIs no inglês.

A quarta e última hipótese diz respeito à possibilidade de encontrar algum padrão de alinhamento entre os níveis da GDF para as EIs no português, conforme análise de Keizer para o inglês (2016). Cumpre esclarecer que tal verificação não objetiva determinar um padrão absoluto para definir transparência e opacidade em EIs, pois entende-se que pode haver diferentes situações de variação morfossintática que ainda não foram verificadas em outras línguas, mas sim identificar alguma tendência para o português, o que seria uma contribuição para a descrição dessas expressões à luz do modelo teórico da GDF.

0.1.5: Metodologia da pesquisa

A metodologia de investigação está organizada em três etapas de análise: a primeira compreende uma **análise cognitiva das EIs**, organizada a partir do seguinte método: (i) seleção e compilação de EIs dicionarizadas, em sua forma consagrada, para verificação de seu significado e

usos possíveis; (ii) levantamento dessas EIs no CRPC e em registros escritos da *web* e interpretação do contexto de inserção dessas expressões nos excertos em que se inserem nos *corpora*³; e (iii) interpretação metafórica das EIs.

Sendo assim, após a definição da EI em sua forma consagrada, os processos de interpretação do contexto de inserção e a interpretação metafórica levam em conta tanto a forma consagrada, como também os elementos em variação nos excertos. Para fins da classificação dos tipos de EIs e da observação de seu alinhamento entre os níveis e camadas da GDF, leva-se em conta somente a forma consagrada.

A segunda etapa diz respeito à **análise morfossintática das EIs**, baseada na aplicação de quatorze critérios de variação morfossintática, sendo os oito primeiros baseados em Keizer (2016, 2018) e o restante nas especificidades do português. Cumpre esclarecer que partimos da hipótese de que a semântica da EI influencia seu comportamento morfossintático e a aplicação desses critérios revela como se dá essa sistematização.

Keizer (2016, 2018) nomeia essa aplicação de critério de “testes”, contudo, nesta pesquisa, preferimos considerá-los como critérios de análise. Assim, os critérios utilizados são os que seguem: substituição lexical, modificação adjetiva, quantificação, número, definitude, passivização, nominalização, referência anafórica, referência catafórica, inserção de advérbios, inserção de pronomes, inversão da ordem preferida, elisão de um termo fixo e possibilidade de conversão para modo imperativo. Nessa etapa, seguindo a metodologia aplicada por Keizer (2016, 2018), considera-se a forma consagrada, conforme padrão dicionarizado da EI, e adicionam-se elementos lexicais (adjetivos, verbos, advérbios, números, artigos, etc.), altera-se a ordem preferida das EIs, retira-se elementos fixos, entre outras estratégias de testagem, a fim de analisar o comportamento morfossintático das expressões. O foco, nesse momento do estudo, é analisar os graus de composicionalidade das EIs, i.e., o quanto uma EI é rígida ou flexível, e como a variação morfossintática perpassa entre as leituras literal e conotativa da expressão.

A terceira etapa é a **análise das EIs de acordo com os níveis e camadas da GDF**. Aqui será apresentada a representação formal das EIs de acordo com os três níveis mais altos da teoria - NI, NR e NM, a fim de compreender o funcionamento pragmático, semântico e morfossintático das EIs. Também foram elaborados critérios de análise para essa etapa: (i) investigação de aspectos pragmáticos e camadas envolvidas no NI; (ii) investigação de aspectos semânticos e camadas envolvidas no NR; e (iii) investigação de aspectos morfossintáticos e camadas envolvidas no NM.

³ Os dados construídos são elaborados, quando necessário, para discussão, no momento da aplicação dos critérios morfossintáticos, no caso de EIs que aparecem nos *corpora* apenas da forma consagrada.

Ao final dessa etapa, busca-se discutir se há algum padrão de alinhamento entre os resultados obtidos em cada nível.

Resumidamente, segue a ordem de organização da metodologia de pesquisa: (i) apresentação do excerto; (ii) análise cognitiva das EIs; (iii) análise morfosintática das EIs; (iv) análise das EIs de acordo com os níveis e camadas da GDF. Ordenamos a apresentação das EIs, de seus respectivos excertos e das três etapas de análise por tipos de EI, a começar pelo tipo um, até o tipo cinco.

0.1.6: Resultados esperados

As etapas de análise cognitiva e de análise morfosintática das EIs auxiliam a definir os cinco tipos de EIs encontradas no português, de acordo com os graus de transparência e de opacidade que apresentam, conforme segue: (i) desmotivadas e não-decomponíveis; (ii) motivadas e não-decomponíveis; (iii) motivadas e decomponíveis; (iv) expressões *AD-HOC*, com uso específico; e o último tipo, (v) parcialmente motivadas por metáforas conceituais e decomponíveis.

Essas duas etapas também auxiliam na compreensão do esforço extra-cognitivo do falante para reanálise de EIs e na investigação dos casos de violação da relação um-para-um entre forma e significado dos elementos constituintes das EIs, identificando casos de descontinuidade específicos das EIs do português.

Já a terceira etapa de análise, que leva em conta os níveis e camadas da GDF, serve à verificação do padrão de alinhamento entre os níveis de Formulação (NI e NR) e o nível de Codificação (NM), o que define o tipo de relação, transparente ou opaca, entre os níveis do modelo e auxilia na caracterização dessas expressões no português, mostrada em detalhe no capítulo de análise.

0.1.7: Justificativa

O objeto de estudo aqui discutido, as EIs, tem sido alvo de pesquisas realizadas pelo autor desde a graduação, em que foi desenvolvido um trabalho de conclusão de curso (da Silva, 2023) em que foram investigados casos de variação entre Fraseologia e Parafraseologia, e.g. *quem não tem cão caça como gato* e *quem não tem cão caça com gato*, sendo a primeira a forma consagrada, tida como uma unidade fraseológica e a segunda, a forma desviante, compreendida como uma unidade parafraseológica.

Esse primeiro contato levou o autor a conhecer o universo das unidades fraseológicas, o aporte

teórico das definições formais e funcionais de provérbios, de fórmulas rotineiras e de EIs vistos em Xatara (1998), Monteiro-Plantin (2011), Corpas Pastor (1996), além de dicionários de EIs (Riva, 2009) e em outros trabalhos acadêmicos sobre o universo da Fraseologia.

Compreende-se também que as EIs constituem o universo sociocultural e linguístico da comunidade falante, principalmente em termos geográficos, e expressam pluralidade linguística e cultural, o que as tornam um alvo de investigação por si só relevante nos estudos linguísticos.

Por fim, se faz importante ressaltar que este estudo se pauta na contribuição de dados reais de EIs encontradas em variedades lusófonas do português e pode servir: para insumo no ensino de Fraseologia (também reconhecida como uma disciplina autônoma dentro da Linguística (Klare, 1986)); para a elaboração de materiais para alimentação de banco de dados; para elaboração de ferramentas de ensino e pesquisa através da Inteligência Artificial; para alimentar reconhecedores automáticos de fala; e para confecção de novos dicionários fraseológicos, pois o vasto campo da Fraseologia brasileira ainda carece de avanços, materiais e dicionários atualizados (Almeida, 2022).

0.1.8: Organização da tese

Este trabalho está organizado da seguinte maneira: o primeiro capítulo traz o aporte teórico aprofundado sobre a Fraseologia e reúne definições das EIs sob a ótica da Gramática Tradicional (Bechara, 2009), da Gramática de Construções (Croft, 2007), da Linguagem baseada no uso (Bybee, 2010), do Cognitivismo (Vega Moreno (2007) e Langlotz (2006)) e da Léxico-Gramática (Vale (2001) e Xatara (1998)). Ainda no primeiro capítulo, discutimos as definições de interpretação metafórica (Lakoff e Johnson, 1980), convenção metafórica (Kövecses, 2010) e dicionarização das EIs (Serpa (1982) e Riva (2009)).

O segundo capítulo discorre sobre o modelo teórico utilizado nesta pesquisa: a Gramática Discursivo-Funcional (Hengeveld; Mackenzie, 2008), com atenção especial às definições dos Componentes Conceitual e Contextual da teoria e um breve apontamento para os fenômenos das referências anafórica e catafórica para o modelo. Além disso, discorreremos sobre a literatura que discute a transparência e a opacidade na GDF a partir de estudos como os de Pezatti et al. (2017) e Leufkens (2015), e apresentamos a abordagem funcional das EIs, proposta por Keizer (2016, 2018).

Já o terceiro capítulo, apresenta os procedimentos metodológicos, as etapas de pesquisa e os critérios de análise. Além disso, são apresentados os dicionários de EIs de Silva (1997), Riva (2009) e Almeida (2022) e outros trabalhos sobre dicionários de EIs (aqui chamados de bancos de dados), os *corpora* utilizados para levantamento das ocorrências em contextos reais de fala e escrita, e a

análise quantitativo-qualitativa utilizada para a seleção das 35 EIs. Cumpre esclarecer que o grupo de EIs selecionadas para esta pesquisa não se deu a partir de recorte de algum campo semântico específico ou de presença de léxico específico, mas sim a partir da reunião de um grupo de EIs consagradas no português, primeiramente filtradas através de sua usualidade e popularidade, conforme pesquisa inicial na *web*, e depois selecionadas após consulta aos dicionários. Em um segundo momento, levou-se em conta, para a análise das ocorrências, a frequência das EIs selecionadas no CRPC e na *web* e a aplicação teórica sobre motivação e composicionalidade de Keizer (2016, 2018) para os *idioms* em inglês.

A análise e a discussão dos dados estão no quarto capítulo, e são apresentados de acordo com a organização proposta na metodologia de pesquisa: (i) apresentação do excerto; (ii) análise cognitiva das EIs; (iii) análise morfossintática das EIs; (iv) análise das EIs de acordo com os níveis e camadas da GDF. Ordenamos a apresentação das EIs, de seus respectivos excertos e das três etapas de análise por tipos de EI, a começar pelo tipo um, até o tipo cinco.

Por fim, as considerações finais fecham o trabalho e, na sequência, finalizando o texto da tese, as referências bibliográficas utilizadas na pesquisa são apresentadas.

CAPÍTULO 01: AS EXPRESSÕES IDIOMÁTICAS NA LITERATURA LINGUÍSTICA

O campo da Fraseologia tem como objeto de estudo as unidades fraseológicas, também chamadas de fraseologismos ou expressões pluriverbais cristalizadas, que são compreendidas como unidades lexicais idiomatizadas que, atuando em conjunto - como um bloco de pensamento - perdem o sentido literal e adquirem um novo sentido idiomático, conotativo, tornando-se uma unidade indissolúvel a partir do seu uso repetitivo, conforme Bally (1951). Exemplos desses tipos de unidades, que representam um conhecimento passado através de gerações, são os provérbios, expressões de cunho reflexivo e orientador quanto à conduta humana, quanto ao que seria ético e moralmente correto aos homens, como, por exemplo, *a pressa é inimiga da perfeição*, *a mentira tem perna curta*, e as expressões idiomáticas (EIs), foco deste estudo, e.g. *descascar o abacaxi*.

Klare (1986) entende que a definição de unidades fraseológicas como bloco de pensamento também pode ser compreendida como a representação de uma palavra, ou seja: uma unidade fraseológica, composta por mais de uma unidade lexical, poderia ser compreendida como uma só palavra. Para a autora, no campo da Linguística, tal investigação teria, assim, seu lugar na Fraseologia, entendida como uma disciplina autônoma que investiga o fraseoléxico de uma língua.

Tal disciplina, para Vinogradov (1946 *apud* Pastore, 2009), tem como objetivo estudar as normas e leis que condicionam a falta de liberdade das palavras e dos seus significados nos esquemas de combinação fixa, tanto em seu estado atual, como também em seu processo histórico, o que o levou à compreensão de três grandes grupos de fraseologismos: (i) unidades fraseológicas, semanticamente indivisíveis, parcialmente motivadas pela significação das partes, e.g. *mais vale um pássaro na mão do que dois voando*; (ii) combinações fraseológicas, semanticamente divisíveis, com um dos elementos realizando o significado da expressão sozinho, e os demais elementos podendo variar, e.g. *recusar categoricamente/definitivamente/terminantemente*, em que o segundo vocábulo “acompanha quase sempre o primeiro, mas é passível de substituição sinônima; assim, poderíamos ter “recusar terminantemente” ou “recusar definitivamente”” (Vinogradov, 1946 *apud* Pastore, 2009, p. 18) e (iii) aderências fraseológicas, semanticamente desmotivadas, e.g. *chover canivetes*.

Fraseologia é um fenômeno linguístico que merece atenção especial, e, ao pensar em sua inserção na educação, concordamos com Riva (2009) sobre a importância desse ramo da Linguística na prática escolar, como ferramenta produtiva nos estudos de regras gramaticais, lexicais, semânticas e também no ensino e aprendizagem de língua materna e estrangeira.

Sobre a Fraseologia do português do Brasil, Riva (2009) menciona a obra de João Ribeiro

(1908) intitulada *Frases feitas: estudo conjectural de locuções, ditados e provérbios* como um marco inicial no país, além de nomes consagrados, como Xatara (*O campo minado das expressões idiomáticas*, 1998; *Tipologia das expressões idiomáticas*, 1998), Biderman (Lexicologia e Lexicografia, 1985; *Unidades complexas do léxico*, 2005) e Vale (*Sintaxe, léxico e expressões idiomáticas*, 1998; *Avanços da linguística de corpus no Brasil*, 2008).

Sobre dicionários brasileiros de fraseologismos, cita-se o *Tesouro da Fraseologia Brasileira* – de Antenor Nascentes (1987), e o *Dicionário Brasileiro de provérbios, locuções e ditos curiosos* – de Magalhães Júnior (1974), que trabalha com locuções populares, provérbios e ditos comparados com formas idênticas ou aproximadas, além dos dicionários brasileiros de Riva (2009) e Silva (1997), utilizados como banco de dados para consulta e compilação das EIs desta pesquisa e que foram detalhados no terceiro capítulo.

Um dos grandes desafios nesse campo está em delimitar e classificar tais fenômenos linguísticos em grupos e categorias, e em elaborar seu processo de dicionarização e atualização dos usos, que, conforme apontado por Sabino (2011), carece de um consenso entre os estudiosos da área, até mesmo no processo de denominação dos fraseologismos:

Ao consultar a literatura específica, é possível perceber que não há limites rígidos capazes de estabelecerem e delimitarem, com precisão, a diversidade de fraseologismos existentes. Fica evidente que ainda não há muito consenso sobre quais unidades são objetos de estudos da fraseologia, nem tampouco sobre as denominações que elas devem receber, quer dizer, não há uma concordância de ideias entre os teóricos no que diz respeito aos termos genéricos que nomeiam os diversos tipos de combinações de palavras. (Sabino, 2011, p. 385)

Fraseologismos, lexias complexas, sequências combinatórias cristalizadas, expressões pluriverbais, unidades pluriverbais lexicalizadas, são alguns dos vários nomes que as unidades fraseológicas recebem em diferentes trabalhos do campo linguístico e que resultam em diferentes materialidades linguísticas fraseológicas, como provérbios e EIs.

Tendo contextualizado o campo da Fraseologia, nas próximas três subseções, abordaremos diferentes definições e perspectivas sobre EIs, como as da Gramática normativa (Bechara, 2009), da Gramática de Construções (Croft, 2007), da Linguagem baseada no uso (Bybee, 2010), Léxico-Gramática (Xatara, 1998), do Cognitivismo (Kövecses (2010) e Lakoff e Johnson (1980)) e as baseadas em processos de dicionarização (Serpa (1982) e Riva (2009)). Tais abordagens sobre EIs são importantes para uma compreensão mais abrangente sobre o universo de investigação linguística desse fenômeno, e embora não tomemos todas essas definições como fundamento para o capítulo de análise, enxergamos tais estudos como bibliografia fundamental sobre o tema.

1.1: EIs: definições gerais

Xatara (1998) define EI como lexia complexa, com formato de uma unidade locucional ou frasal, de caráter indecomponível, pois constitui uma combinatória fechada. Para a autora, a EI apresenta distribuição única ou muito restrita, tendo seu valor semântico conotativo, porque sua interpretação semântica corresponde a pelo menos um primeiro nível de abstração, calculada a partir da soma de seus elementos, sem considerar os seus significados individuais. Além disso, a EI, conforme Xatara (1998) é também considerada cristalizada, já que a sua significação é estável devido à frequência de uso na comunidade.

Fonseca (2013) também destaca o valor expressivo das EIs ao compreender sua natureza como parte dos costumes da comunidade linguística, tendo forte empregabilidade na oralidade, de acordo com Náray-Szabó (2008). Essas expressões são notadas em diferentes cenários pragmático-discursivos, tais como: (i) evocativos, i.e. refletem a realidade de modo mais ou menos objetivo, e.g. *estar com a cabeça nas nuvens*; (ii) avaliativos, que descrevem impressões expressas da realidade, e.g. *custou o olho da cara*; e também (iii) dialógicos, que descrevem uma situação particular, e.g. *levou um puxão de orelhas*.

Tal empregabilidade reforça o desejo de o falante ser expressivo, comunicando-se de maneira mais persuasiva, pitoresca, irônica, cômica, por meio dessas “combinações inusitadas que são unidades funcionais significativas do discurso, constituídas por sequências estereotipadas de lexemas” (Riva, 2009, p. 23). O autor ainda faz um adendo ao distinguir EIs de expressões terminológicas, sendo estas expressões denotativas e relativamente cristalizadas, e aquelas, conotativas e relativamente cristalizadas. No seu dicionário, o autor ancora sua definição de EI com base em Biderman (2001), Corbin (1983) e Tagnin (1988):

Concordamos com Biderman (2001) quando ela nos diz que os idiomatismos são combinatórias de lexemas que o uso consagrou numa determinada seqüência, ou seja, desconsiderando suas partes como unidades semânticas e cujo significado não se dá na simples somatória dessas partes. Os significados literais podem, em boa parte dos casos, ajudar a entender o significado conotativo da expressão. [...] pode-se concluir que a EI não é apenas um aglomerado de idiosincrasias lexicais (escolhas pessoais), mas combinações convencionais de relações sintático-semânticas regulares, aceitas e reconhecidas por um grupo (Corbin, 1983; Tagnin, 1988 apud Riva, 2009, p. 21).

Xatara (1998) delimita o estatuto das EIs ao afirmar que estas não são:

1. locuções, pois estas encerram uma forma funcional de organização dos elementos disponíveis da língua (por exemplo, “vez ou outra” ou “à medida que”) e não uma maneira de exprimir algo com a implicação de uma retórica e uma estilística e com a suposição de uma figura, como a metáfora;
2. combinatórias usuais ou colocações, pois essas são unidades linguísticas convencionais, de sentido denotativo, que se caracterizam pela co-ocorrência léxicosintática de seus elementos, por exemplo, “diametralmente oposto”;
3. perífrases verbais, cujos argumentos não têm restrição, de mera função nominativa, que correspondem ao modelo SV + SN, sem transformação dos significados dos componentes, por exemplo, “fez um acordo (contrato)” ou “dar ajuda (apoio)”;
4. ditados, por exemplo, “achado não é roubado” que são elementos não conotados, ou provérbios, ricos em imagens, por exemplo, “água mole em pedra dura tanto bate até que fura”. A formulação arcaizante e os enunciados concisos e fechados desses dois tipos de fraseologismos confere-lhes uma autoridade que depende da “sabedoria dos antigos”, isto é, tanto ditados quanto provérbios enunciam verdades eternas na forma de simples constatações;
5. fórmulas situacionais ou sentenças usadas em situações de comunicação específicas (“Não seja infantil!” “Você tá brincando...” [Tagnin, 1987; Xatara, 1994]);
6. clichês ou chavões, ou seja, criações literárias que se banalizaram pelo uso como, por exemplo, “o astro do dia” ou “na primavera da vida”;
7. sintagmas terminológicos, unidades lexicais complexas que correspondem a um conceito restrito a uma determinada área científica ou técnica, por exemplo, “unidade de terapia intensiva”. (Xatara, 1998 apud Pastore, 2009, p. 25-26)

Contrariamente ao que foi discutido até aqui, Bechara (2009) entende **idiomatismos** como toda a maneira de dizer. Conforme o estudioso, idiomatismos são aceitos na linguagem culta, mas não podem ser analisados com base nos princípios gerais da Gramática, pois são **anomalias** de linguagem, compostas por sintagmas complexos que podem ser constituídos por mais de dois elementos lexicais.

O autor destaca a importância de não se definir idiomatismo como um fenômeno particular de uma língua, pois ninguém conhece todos os outros idiomas em todos os seus modos de falar, dado importante para casos de EIs que se assemelham na forma, mas com significados distintos em línguas diferentes, como por exemplo, a expressão *to break a leg* – desejo de boa sorte em uma performance ou apresentação (Simões, 2018) em inglês, e que sua tradução literal em português brasileiro – *quebrar a(s) perna(s)*, é entendida como frustração, dificuldade em concluir determinada atividade.

O dicionário de EIs português-inglês e inglês-português de Serpa (1982) traz vários exemplos de expressões e traduções entre as duas línguas, com várias expressões muito semelhantes entre forma e significado, como *to cost an arm and a leg* e *custar um olho da cara*, i.e. algo muito caro, de alto valor financeiro, e outras com significado semelhante e formas um pouco distintas, como *to pay the piper* e *pagar o pato*, i.e. assumir as consequências.

Importante ressaltar que a grande maioria das EIs surge da necessidade da comunidade falante criar recursos linguísticos para interação comunicativa para expressar diferentes necessidades, sentimentos, emoções e percepções. Um exemplo desse reflexo é a representação do alto custo em inglês e em português com o uso de uma parte do corpo na EI, em *to cost an arm and a leg* e *custar um olho da cara*, ambas as expressões descrevem uma situação corriqueira, envolvendo o comércio de itens e a percepção do falante em relação ao sentimento do alto custo de um item ser algo equivalente a uma parte do corpo, vital à sua existência.

Drew e Holt (1996) realizaram um estudo sobre o uso de EIs em conversas cotidianas e destacaram seu papel na transição ou encerramento de tópicos, contextos em que o seu uso foi identificado como um recurso aproximador e orientador do tópico discursivo, além de realçar percepções dos falantes, como reclamação e desacordo.

Tal julgamento de valores também é destacado no trabalho de Marques e Teletin (2011), em que as autoras destacam a contribuição do valor expressivo das expressões para a constituição de estereótipos culturais de determinada comunidade:

O locutor actualiza o sentido virtual da expressão, integrando-a num enunciado concreto através do qual realiza um acto determinado de discurso, pondo em destaque uma certa intencionalidade comunicativa (informar, ironizar, etc.) numa situação concreta de comunicação. Os estereótipos veiculados por estas expressões estão presentes na língua e, na maioria dos casos, manifestam um julgamento negativo. (Marques; Teletin, 2011, p. 346)

Kövecses (2010) reforça que a experiência de vida dos falantes de determinada língua reflete-se na língua e, conseqüentemente, cada cultura o faz de determinado modo. Por exemplo, as convenções metafóricas de sensações como “felicidade é para cima” e “tristeza é para baixo” foram encontradas em diferentes culturas e línguas, como português, inglês, húngaro e chinês. Tais convenções metafóricas são refletidas na língua, como na expressão *João está tão para baixo hoje*, entendida como *João* está triste, desanimado, angustiado ou preocupado.

Outros autores como Monteiro-Plantin (2011), Zuluaga (1980) e Corpas Pastor (1996) discorrem sobre definições formais e funcionais envolvendo a idiomaticidade, delimitando tais definições em quatro parâmetros norteadores: pluriverbalidade, idiomaticidade, fixação e frequência. O quadro 1 resume como se dá a abordagem desses quatro parâmetros, com base nos três autores citados e em Xatara (1998):

Característica	Descrição	Exemplo
Pluriverbalidade ou polilexicalidade	São sequências linguísticas, constituídas por, pelo menos, dois elementos que se apresentam de forma mais ou menos fixa. São divididas em lexia simples: corresponde apenas a uma unidade lexical (Xatara, 1998);	Cachorro, gato, cada, cadeira, telefone, etc. (Elaboração própria, 2025).
	Lexia composta: corresponde à integração semântica de unidades lexicais.	Arco-íris, Bem-te-vi, Super-herói, Anti-inflamatório, etc. (Monteiro-Plantin, 2011).
	Lexia complexa: corresponde a uma sequência de lexicalização em diferentes graus, podendo ser rígida ou variável.	De modo que, Complexo industrial, À medida que, etc. (Monteiro-Plantin, 2011).
	Lexia textual: corresponde a um enunciado ou um texto (Monteiro-Plantin, 2011).	Água mole em pedra dura, tanto bate até que fura, Cavalo dado não se olha os dentes, etc. (Elaboração própria, 2025).
Idiomaticidade	Caráter não literal sobre o sentido da expressão, na qual os elementos assumem novos sentidos de acordo com o conjunto como um todo, alcançando um novo referencial e sentido (Xatara, 1998).	Descascar o abacaxi (resolver um problema), acertar na mosca/lata (acertar algo difícil com precisão), etc. (Riva, 2009).
Fixação na comunidade falante	É a utilização combinatória previamente feita que certas expressões apresentam num discurso, tratando como nível interno - fixação estrutural sintática, e nível externo – aceitação dessas expressões e frequência no discurso. Para Zuluaga (1980), as EIs são fixas devido a sua representação em bloco - são um produto de processos de repetição na diacronia da língua. Corpas Pastor (1996) apresenta ainda no nível externo dois tipos de fixação: pragmática, presente nas fórmulas de cortesia com o emprego em determinadas situações sociais, e a fixação posicional, onde há preferências de uso de certas expressões sobre outras, determinando posições dentro de um texto.	Estar no bico do urubu (forma interna estrutural)/Estar no bico do gavião (forma desviante); Estar no bico do urubu (forma fixa externa com utilização pragmática de sentido pejorativo/agressivo) (Elaboração própria, 2025).
Frequência e convencionalidade (cristalização)	É a frequência e uso das expressões, além da estabilidade de sua forma, o que marca seu poder de expressão e influência na comunidade falante, fator que a consagra e a caracteriza como EI (Monteiro-Plantin, 2011).	Ter a paciência de Jó, Abandonar o barco, fazer vista grossa, Estar de saco cheio, etc. são alguns exemplos de expressões usuais no português do Brasil (Riva, 2009).

Quadro 1: Definição geral das EIs (Elaboração própria, 2025)

Conforme se observa no quadro 1, existem características que são específicas das EIs, e, como defendido também por Biderman (2001), a cristalização é uma delas e está intrinsecamente ligada à aceitação e ao uso da comunidade falante. A idiomaticidade da expressão contribui para sua

cristalização, pois favorece a convencionalidade da expressão enquanto unidade fraseológica utilizada para determinado fim, com significado selado específico. A exemplo, dificilmente se observa a expressão *vestir o paletó de madeira* em outros contextos que não seja o de velório, morte, luto ou similares.

Pastore (2009), citando o trabalho de Xatara (1994), reconhece que o caráter rígido das EIs pode aceitar certos tipos de modificação, como de gênero (*ser surdo/surda como uma porta*), superlativo/aumentativo/diminutivo (*tirar o cavalo/cavalinho da chuva*), de número, singular ou plural (*levar um puxão/uns puxões de orelha*), de acréscimo de artigos ou pronomes (*ver a/aquela/nossa/sua luz no fim do túnel*), de inserção de advérbios (*acertar novamente na mosca*), negação (*não colocar a carroça na frente dos bois*), de permuta lexical (*descer/baixar o pau em alguém*), de tempo verbal (eu acerto/acertarei/acertei na mosca) e de modo verbal (*ela pegou/que pegue o touro à unha*).

Há casos também em que os novos valores idiomáticos expressos dentro desses blocos de pensamento contribuem também para a abstração semântica dos elementos, o que aumenta sua empregabilidade em novos usos, estabelecendo mudança semântica (Lopes, 2007). Um exemplo disso é a expressão *começar com o pé direito*, uma antiga superstição do Império Romano para evitar o agouro (Silva, 1997). Tal expressão tomou maiores proporções, e passou a ser empregada em demais ações, como *entrar com o pé direito*, *acordar com o pé direito*, e quase que automaticamente, o pé esquerdo virou sinônimo de azar, má sorte.

Ainda sobre os estereótipos culturais incorporados nas EIs, Schapira (1999) destaca os estereótipos do pensamento enquanto representação conceitual de determinada tradição cultural presente em uma comunidade. Para a autora, as EIs adquiriram fixidez verbal na língua e, através da repetição, tornaram-se esquemas conceituais que não podem ser modificados pelo falante. Esses estereótipos do pensamento traduzem-se em estereótipos linguísticos. Lyons (1979) também reforça a compreensão das EIs como estereotipadas, aprendidas como um total analisável e indecomponível.

Simões (2018) reforça a função das EIs em descrever o mundo real através desses estereótipos que recriam a realidade, constituindo, assim, uma imagem da identidade cultural inerente aos fatos sócio-históricos e culturais de cada comunidade. Simões (2018) entende que as EIs são **idiosincrasias** que individualizam determinada comunidade em suas vivências enquanto comunidade social. Por exemplo, a expressão *para inglês ver*, que significa fazer algo superficial, apenas para mascarar a realidade, usada no Brasil, em Angola e também em Portugal, tem suas origens em terras brasileiras durante a primeira metade do século XIX (Silva, 1997).

A etimologia da expressão está ligada à época em que a Inglaterra estabeleceu um prazo para

o fim do tráfico dos escravos em todos os países com os quais possuía relações comerciais, e no Brasil, os prazos não foram cumpridos e foi promulgada a Lei Feijó, em 1831, que declarava livres todos os africanos desembarcados em solo brasileiro, mas que na prática, não era assim.

Vejamos a seguir uma pequena amostra de três EIs usuais na variedade brasileira, extraídas do dicionário de Riva (2009), em que se observam as características apontadas pelo quadro 1:

- (1-1) ***Molhar a mão***: dar dinheiro para alguém em troca de favores escusos [orig. sup.: alusão ao dinheiro ilícito retido nas mãos] Ex: Prefiro pagar multa a *molhar a mão* de policiais. (www.desvelar.blogger.com.br/2004_11_01_archive.html; acesso em 06/04/05) (Riva, 2009, p. 134)⁴
- (1-2) ***Estar com todo o gás***: estar disposto, com energia [orig. sup.: alusão à imagem daquilo que é movido a gás] ♦ Ela *está com todo o gás*: já fez de quase tudo na vida e permanece como testemunha viva da história da indústria... (quatorrodas.abril.com.br/classicos/grandesbrasileiros/1104_kombi.shtml; acesso em 10/09/04) (Riva, 2009, p. 112)
- (1-3) ***Tirar o pé do acelerador***: levar uma vida mais calma [orig.: alusão ao ato de deixar de acelerar o veículo] Ex: A ordem é *tirar o pé do acelerador* e dar um tempo na correria. Senão, daqui a pouco você vai ficar no maior estresse. (www2.uol.com.br/simbolo/atrevida/0500/h_sagitario.html); (acesso em 14/04/05) (Riva, 2009, p.)

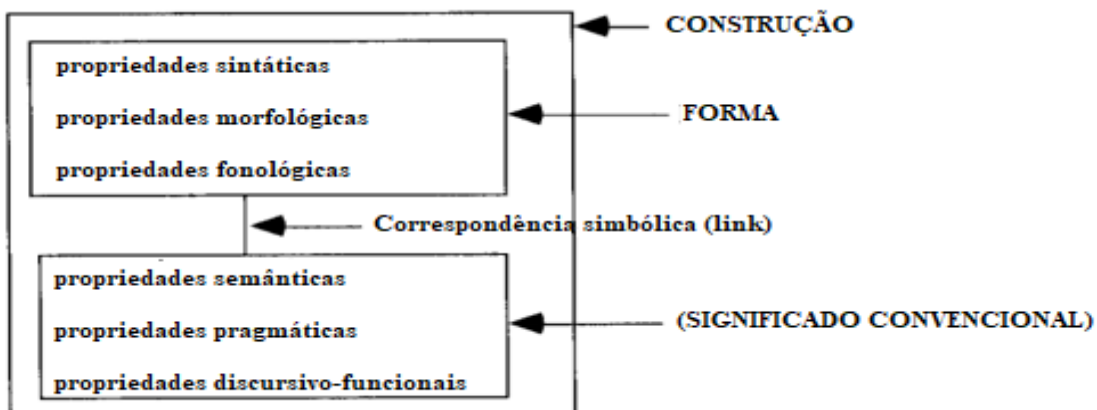
Nos três exemplos, as expressões são lexias complexas, apresentam valor idiomatizado (conotativo), apresentam certo grau de fixação dos elementos e usos, e, conforme o autor, são expressões convencionalizadas na comunidade.

1.2: Idiomatismos na Gramática Construcional e no Modelo de linguagem baseado no uso

Para Croft (2007), as EIs são um grande desafio para o modelo componencial da Gramática de Construções – teoria que entende a forma básica de uma estrutura sintática como uma construção, representando um emparelhamento de uma estrutura gramatical complexa com o seu significado, ambos organizados em uma rede.

As estruturas maiores que uma palavra são compreendidas de acordo com as regras gerais dos componentes gramaticais e suas interfaces – propriedades fonológicas, sintáticas e morfológicas da forma, que possuem uma ligação correspondente com as propriedades semânticas, pragmáticas e discursivo-funcionais do significado convencional, conforme figura 1:

⁴ Note que a contagem dos exemplos é reiniciada a cada capítulo da tese e está delimitada da seguinte forma: número do capítulo e número do exemplo, conforme exemplo (1-1) - capítulo 1, exemplo 1.



**Figura 1: Estrutura simbólica de uma construção (Croft, 2007, p. 472)
Adaptado para português (Elaboração própria, 2025)**

Para Croft (2007), as EIs são expressões linguísticas semântica e sintaticamente idiossincráticas, com extensão maior que uma palavra, e que não podem simplesmente ser inseridas no léxico sem um mecanismo especial. O autor cita, como exemplo, a expressão lexicalmente idiossincrática *kith and kin* (familiares e amigos), na qual tais elementos não são encontrados em nenhum outro lugar, classificando esses tipos como irregulares, assim como outras expressões que são sintaticamente idiossincráticas, e.g. *all of a sudden* (de repente), ou semanticamente idiossincráticas, e.g. *tickle the ivories* (tocar piano).

Tais idiossincrasias requerem o acionamento de múltiplos componentes - pragmáticos, semânticos, sintáticos, morfológicos e fonológicos - em uma leitura complexa dos elementos que, ora seguem regras gramaticais, ora parecem não pertencer ao léxico da língua. Para Croft (2007), as EIs desafiam a separabilidade dos níveis sintáticos, como se os elementos da expressão fossem pedaços que não podem ser definidos individualmente, representando um conjunto complexo que aciona diferentes níveis em sua forma e significado.

Fillmore et al. (1988) distinguem diferentes tipos de EIs com base no léxico e em regras sintáticas regulares da língua, e também com base no que é familiar e no que é não-familiar ao ouvinte, em termos dos elementos e sua disposição na expressão.

Assim como visto na seção anterior, os autores também compreendem a composição das EIs com as seguintes propriedades: (i) convencionalidade – o seu significado não pode ser compreendido pela interpretação de seus elementos isolados; (ii) inflexibilidade- há uma fixação ou um número limitado de combinação sintática; (iii) figuração – leitura metafórica do seu significado idiomático; (iv) proverbialidade – descrevem situações cotidianas ou de interesse social particular; (v)

informalidade – usos coloquiais ou informais; e (vi) afetividade – implicam instâncias afetivas ao objeto denotado e implicam valor expressivo em seus usos.

Os autores delimitam, assim, três diferentes formas de se distinguir as EIs: (i) expressões codificadas e expressões decodificadas: estas com significado imprevisível ao falante com base no conhecimento de gramática da língua, e.g. *to kick the bucket* (morrer), e aquelas com algum sentido proveniente das partes componentes da expressão, e.g. *to spill the beans* (revelar um segredo); (ii) EIs substantivas e EIs formais – as substantivas são lexicalmente preenchidas, como por exemplo *to hit the nail on the head* (acertar com precisão), e as formais, lexicalmente abertas, e.g. *the faster we run, the slower they run* (expressões de escalas correlacionadas); e (iii) expressões gramaticais – têm estruturas gramaticais devidamente preenchidas de acordo com as regras gerais da língua, e.g. *to kick the bucket*, e expressões extragramaticais – têm estruturas anômalas, e.g. *so far so good* (tudo bem até o momento).

Bybee (2010), sob a perspectiva do Modelo de linguagem baseado no uso, que é um modelo de representação gramatical em que o uso determina a representação gramatical com base especificamente na frequência de uso da forma e significado, vê nas EIs construções metafóricas que requerem uma representação cognitiva própria, pois possuem significado imprevisível, e.g. *to pull the strings*, que significa influência de poder para manipulação de determinadas ações em benefício a alguém. Tal significado imprevisível e idiomático está relacionado ao seu processo de convencionalização na comunidade falante, sendo um tipo de expressão pré-fabricada de sentido não-literar.

A autora usa o termo *prefab* (pré-fabricada) para reforçar o caráter de expressão multi-palavra convencionalizada das EIs, na qual sua convencionalização está ligada ao seu processo de uso e repetição, o que as tornam construções parcialmente esquemáticas, e o falante da língua realiza sua interpretação através de conexões entre significado literal e metafórico, de acordo com a ligação dos *chunkies* da expressão.

Segundo a autora, a alta frequência de uma expressão numa determinada língua faz com que tal construção perca sua estrutura interna em termos de identificabilidade das partes constituintes. Essas construções passam a ter um significado especial e outras propriedades formais. Para Bybee (2010) esse processo é natural às línguas:

A língua é também um fenômeno que aparentemente apresenta uma estrutura e padrão regulares, ao mesmo tempo que apresenta variações consideráveis a todos os níveis: as línguas diferem-se umas das outras, embora sejam patentemente moldadas pelos mesmos princípios; construções comparáveis em línguas diferentes servem funções semelhantes e baseiam-se em princípios semelhantes, mas diferem-se umas das outras de maneiras

especificáveis; as expressões dentro de uma língua se diferenciam entre si, embora ainda exibam os mesmos padrões estruturais; as línguas mudam com o tempo, mas de maneira bastante regular. (Bybee, 2010, p. 1. Tradução nossa)⁵

A autora ainda discorre sobre a autonomia das EIs, que aumenta conforme sua frequência de uso, e sobre sua estrutura metafórica, que ganha força conforme sua estrutura interna de significado literal perde força. Essa autonomia pode ser mensurada através do acesso direto à repetição das EIs em suas conexões complexas, i.e., a exposição do falante a essas expressões; a redução fonética, que está ligada aos casos de contrações na língua; e a interpretação da expressão como idiomática, conforme sua conexão com o contexto de uso. Por exemplo, em casos de expressões em que ambos os sentidos (literal e idiomático) são possíveis e naturais na língua, o contexto de inserção direciona o falante ao sentido correto.

Para Cacciari e Glucksberg (1995), a linguagem figurativa tem um papel importante na interpretação dos significados idiomático e literal: este como ação concreta, e aquele, como ação abstrata, referente aos estados e eventos mentais. Os autores trazem, como exemplo, a expressão *to carry coal do Newcastle*, i.e. ação inútil, na qual ocorre a ativação dos dois significados – literal e idiomático, através de inferência baseada em conceitos culturais: Newcastle já foi reconhecida como uma das produtoras e exportadoras de carvão do mundo.

1.3: Expressões cristalizadas e interpretação metafórica

Na proposta tipológica de Vale (2001) sobre as EIs verbais, sob a ótica Valencial (Tesnière, (1959); Chafe (1979); Busse e Vilela (1986); Borba (1996) e também sob a teoria da Léxico-Gramática (Gross, 1975, 1976), nos interessa destacar as abordagens psicolinguística e cognitivista para interpretação das expressões cristalizadas, estas entendidas como expressões verbais de valor idiomático, constituídas por um sujeito fixo ou livre; um verbo, entendido como o centro da expressão; e complementos fixos ou livres, que apresentam comportamento morfossintático similar ao de um sintagma simples, mas comportamento semântico a partir da soma das partes constituintes.

A significação metafórica da expressão contribui para sua cristalização e a diferença da significação componencial, o que requer ao falante um contexto de uso que o faça diferenciar essas

⁵ Texto original: Language is also a phenomenon that exhibits apparent structure and regularity of patterning while at the same time showing considerable variation at all levels: languages differ from one another while still being patently shaped by the same principles; comparable constructions in different languages serve similar functions and are based on similar principles, yet differ from one another in specifiable ways; utterances within a language differ from one another while still exhibiting the same structural patterns; languages change over time, but in fairly regular ways. (Bybee, 2010, p. 1)

significações. Uma das questões centrais da pesquisa diz respeito a como o falante diferencia esses diferentes cenários - leituras literal e conotativa - diante de uma EI.

O autor destaca a dificuldade de simplesmente responder a essa questão pela leitura metafórica de uma expressão, pois há diversos casos que é possível uma interpretação não-metafórica de uma expressão idiomática, e.g. *bater as botas, carregar o piano, descascar o abacaxi, levar um puxão de orelha, etc.* Daí entram as abordagens citadas como parte de nosso interesse. A abordagem psicolinguística procura compreender os processos de produção e compreensão dessas expressões, e a abordagem cognitivista, persegue os processos cognitivos do falante para compreensão das expressões cristalizadas.

Na abordagem psicolinguística, o autor faz referência a duas hipóteses: a interpretação de uma expressão cristalizada passa por uma leitura literal dos seus componentes (Glass, 1983); e, de forma inversa, haver somente uma leitura conotativa dos componentes (Gibbs, 1985).

Na abordagem cognitivista, Lakoff (1986) entende que as metáforas e metonímias presentes nas EIs passam por processo psicolinguístico e cognitivo de diferentes estágios de sensações humanas, que se refletem no processo de criação e decodificação das expressões. Como exemplo disso, pode-se pensar sobre a sensação de raiva de um sujeito que sofre um ato agressivo; depois experimenta uma série de efeitos; pode se manter no controle; perder o controle; e revidar contra quem o provocou (Lakoff, 1986 apud Vale, 2001, p. 66). Tal processo resulta em expressões que descrevem o sentimento de raiva: *chutar o pau da barraca, rodar a baiana, perder as estribeiras, ter pavio curto, etc.*

Assim como Lakoff (1986), outros trabalhos, como de Ortíz-Alvaréz (2000) e Tristán Pérez (1988) reforçam o grande número de EIs regidas por interpretação metafórica ou metonímica⁶ e processos de comparação. Pastore (2009) discorre sobre o processo de metáfora na língua:

Uma metáfora surge por um processo de denominação de novos objetos, fenômenos etc., sob a base de palavras já existentes na língua, ou por redenominação de outros objetos e fenômenos. As imagens, que são o cerne de qualquer metáfora, podem desaparecer com o tempo e o que acaba ficando na língua é um termo que ninguém mais associa com a origem real. É, pois, por meio de um processo de metaforização que uma característica ou propriedade inerente a um objeto passa a caracterizar outro, que pode pertencer a uma classe totalmente diferente. (Pastore, 2009, p. 19)

⁶ Lakoff e Johnson (1980) defendem que a metonímia (representação da parte pelo todo) também faz parte da vida cotidiana da comunidade e tem comportamento metafórico, representado através de sua materialidade linguística, traços culturais de uma comunidade e sua maneira de representar o mundo.

Para Lakoff e Johnson (1980), a metáfora faz parte da vida cotidiana da comunidade, e o nosso sistema conceitual é, em boa parte, metafórico, já que nossa maneira de se relacionar com o mundo e de definir nossa realidade através da linguagem passa por processos de natureza metafórica. Relacionamos as expressões linguísticas com o que experienciamos no mundo e, assim, conceitualizamos essas vivências em processos de pensamento com base metafórica, de modo estruturado e sistemático, assim como se dá o seu registro na língua – um sistema estruturado e sistemático.

Desse modo, os autores entendem que é possível usar as expressões linguísticas metafóricas para compreender a natureza metafórica das atividades humanas, como por exemplo, a expressão *Time is money* (tempo é dinheiro), que, em várias culturas, significa que o tempo é valioso, e assim como o dinheiro, usado como pagamento pelo trabalho.

Tal conceito também foi incorporado em outras relações de tempo e serviços, como taxas diárias de juros, chamadas telefônicas, etc. A partir desse processo de pensamento metafórico, agimos como se o tempo fosse um bem valioso e limitado, tal como o dinheiro, e assim, concebemos novas práticas cotidianas vivenciadas por essa relação tempo-dinheiro: gasto, poupado, bem investido, desperdiçado, etc.

Tal processo reflete-se e materializa-se na língua e, citando a metáfora do canal de Reddy (1979), as ideias e significados são objetos, as expressões linguísticas são recipientes e a comunicação é um envio. O falante coloca ideias dentro das palavras e as envia para seu interlocutor, que retira as ideias-objetos das palavras-recipientes.

Com base nessa tríade, Lakoff e Johnson (1980) entendem que as palavras e sentenças têm significados em si mesmas, e, em determinados casos, são independentes do contexto ou do falante, como, por exemplo, a ideia de localização espacial e suas representações culturais com nossas ações e sensações cotidianas: felicidade e saúde serem para cima (base física que representa que homens e mamíferos levantam-se na posição vertical quando acordam); e tristeza, morte e doença, para baixo (base física que representa que homens e mamíferos morrem ou ficam deitados quando doentes/mortos). Os autores ainda reforçam que, cada cultura tem sua base experiencial própria, evitando fazer generalizações acerca dos significados das metáforas e suas representações em expressões linguísticas metafóricas.

A própria ideia do funeral e dos diferentes estilos culturais de sua realização, como enterros, velórios com um ambiente de festa, simbolizando uma celebração de uma nova etapa, ou clima de tristeza pela perda de um ente querido, ou o ato de enterrar o corpo no chão, ou de manter em dispositivos suspensos, alocados na horizontal ou vertical, ou ainda de serem cremados, são exemplos

de representações culturais distintas sobre o mesmo evento ou sentimento, pois os valores culturais são coerentes com a estrutura metafórica dos conceitos fundamentais que regem cada cultura, e, dentro de uma mesma cultura, e suas respectivas subculturas, podem haver divergências sobre a mesma metáfora. Isso ocorre porque essas diferentes subculturas, integrantes de uma cultura dominante, partilham certos valores básicos, mas podem dar-lhes diferentes prioridades.

Ainda sobre a organização conceitual do significado das palavras, Lakoff (1986), com base na teoria do protótipo, presente nos estudos de psicologia de Rosch (1975), adotou o conceito de organização prototípica, no qual as palavras nem sempre têm os usos dotados do mesmo estatuto conceitual, mas usos mais prototípicos e outros mais periféricos em relação a um centro prototípico, que é descrito como a significação componencial, primária e denotativa da palavra. Já os significados periféricos são os sentidos correlacionados, metafóricos e conotativos.

Com base na organização prototípica de Lakoff (1986), Teixeira (2020) utiliza como exemplo o conceito *vermelho*, tendo como sentido prototípico a cor vermelha (V1), e como sentidos periféricos, pessoa ruborizada (V2), cartão de expulsão no futebol (V3), simpatizantes de movimento político de esquerda (V4), adeptos de clubes de futebol, como Benfica-Portugal ou Liverpool-Inglaterra (V5), etc.

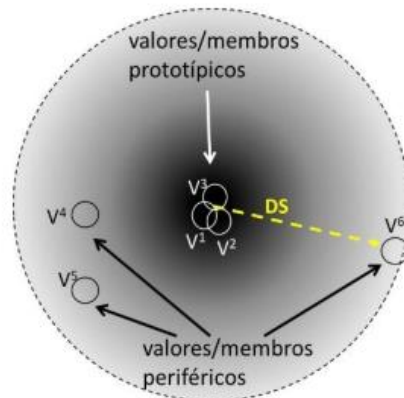


Figura 2: Esquema representativo da estrutura prototípica (Teixeira, 2020, p. 29)

Podemos ver na estrutura radial do conceito *vermelho* que os sentidos correlacionados englobam diferentes relações semântico-cognitivas, como a metonímia (clube/partido político pela cor) e a metáfora (cor vermelha representando raiva, vergonha, sinal de alerta, sinal de erro, etc.). A proposta da estrutura prototípica não pretende separar os significados conotativo e denotativo em dois planos distintos de significação, mas sim entender a relação de distância semântica (DS) entre os valores, e como o falante correlaciona ambos e os interpreta.

Lakoff (1986) utiliza, como exemplo, a radialidade do conceito *mãe*, tendo como sentidos centrais: genitora, fêmea que contribui com material genético, fêmea adulta que cria, educa, esposa, ancestral fêmea, e como sentidos mais periféricos, mãe de aluguel, mãe de leite, mãe adotiva, mãe solteira e madrasta.

Essa abordagem dos trabalhos de Lakoff (1986,1995) e Lakoff e Johnson (1980) interessa ao processo de reflexão sobre EIs proposto nesta fundamentação teórica, pois acreditamos que a interpretação metafórica das EIs pode ser disposta em esquemas de estrutura prototípica, em que o falante reconhece a leitura idiomática de uma expressão através da distância semântica dos seus elementos em conjunto, e sua relação se estabelece não apenas com o contexto de uso relacionado a sua comunidade e cultura, mas também com o sistema conceitual metafórico.

Tomemos como exemplo a expressão *descascar o abacaxi*, que apresenta, conforme vimos anteriormente (Silva, 1997), significação cristalizada e dicionarizada, conotando o significado de resolução/enfrentamento de um problema. Nesse caso, temos como sentido componencial, mais próximo ao centro prototípico, a ação literal de descascar a fruta (V1), e como sentido idiomático, resolver um problema (V2), este mais periférico na estrutura radial; já a distância semântica (DS) é o percurso que o falante faz, correlacionando V1 e V2, o contexto de uso e seu conhecimento idiomático:

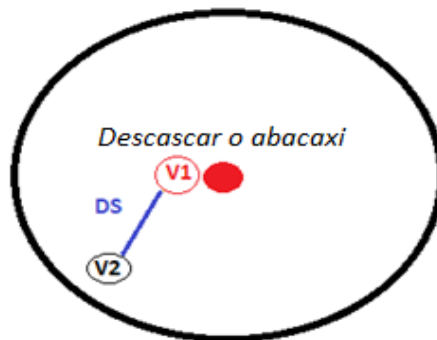


Figura 3: Estrutura prototípica da expressão idiomática *descascar o abacaxi* (Elaboração própria, 2025)

Temos acima uma exemplificação de uma EI do português encontrada em registros no Brasil, Portugal e países africanos, conforme consulta no CRPC e na *web*, que nos fornece diferentes reflexões sobre a abordagem cognitivista de Lakoff (1986, 1995): é uma expressão cuja leitura literal é possível (V1), e, com isso, em um contexto específico, como uma narrativa de uma receita, poderia direcionar sua significação para V1, e em um contexto distante da receita, como a necessidade de se resolver um problema numa empresa ou ambiente doméstico, direcionaria o falante para V2.

Durante esse percurso, a base experiencial do falante, enquanto pertencente a uma comunidade falante de português, em determinada época (no caso, 2025, por exemplo), entenderia a significação V2 e a associaria o conceito *abacaxi* com problema, conforme vimos em seu verbete, no dicionário de Silva (1997)⁷.

Possivelmente em outra época, ou mesmo em alguma variedade lusófona, esse cenário não seria (tão ou facilmente) possível, como vimos na expressão *para inglês ver*, oriunda de uma época e cenário específicos. A mesma base experiencial e cultural do falante de português o faria entender a expressão *fazer tempestade num copo de água* diretamente relacionada à sua significação metafórica (no caso, V1), devido à impossibilidade, ou em outros casos, a uma grande improbabilidade da realização de sua significação componencial (no caso, V2).

Goldberg (1995) reforça o caráter das EIs como composicionais devido a essa representação como construção de base semântica independente, que possui regras de interpretação semântica exclusiva para esse tipo de construção. Essas regras composicionais gerais de interpretação semântica estão associadas a estruturas sintáticas esquemáticas, que, no caso das EIs, funcionam como um bloco de pensamento e ativam um nível de abstração semântico-cognitiva no processo de codificação e decodificação por parte do falante, que avalia os diferentes sentidos – prototípico ou correlacionado daquela EI.

Kövecses (2010) traz um importante conceito de representação mental e de conceitualização metafórica na língua: a sua categorização em quadros, esquemas ou modelos, que são representações mentais estruturadas de uma organização coerente da experiência humana. Tais quadros de representação mental variam de cultura para cultura, gerando interpretações diferentes (construções alternativas) ou similares. As similaridades se agrupam em construções que compartilham algum tipo de semelhança familiar, o que contribui para a construção e disseminação de metáforas universais.

O autor entende a complexidade do termo, devido ao elevado número de línguas faladas no mundo e suas particularidades culturais, mas traz apontamentos de Alverson (1994) e de vários outros estudos sobre metáforas conceituais partilhadas em um grande número de línguas, como o conceito metafórico de *Tempo é espaço*, similar ao que vimos anteriormente *Tempo é dinheiro*.

Sobre as construções alternativas, Kövecses (2010) discorre sobre como diversas línguas partilham, modificam ou emprestam construções metafóricas entre si, algo que também se reflete nas EIs, como *to take the cake* (variedade americana) e *to take the biscuit* (variedade britânica), fruto de diferentes experiências de vida em diferentes contextos socioculturais para cada comunidade falante.

⁷ Conforme o exemplo que está na p. 17 da introdução desta tese.

O trabalho do autor nos auxilia aqui a sustentar nossa proposta de trabalhar com diferentes variedades lusófonas sem nos ater a representações únicas entre forma e significado de EIs em variedades lusófonas do português, pois entende-se que há essa partilha de construções metafóricas entre as variedades da língua, o que pode se refletir em construções similares ou alternativas.

Os registros das expressões nos dicionários tomados como nossos bancos de dados de EIs auxiliam a compreender essa conexão entre forma e significado dessas expressões, considerando suas variações, como por exemplo a expressão *acertar na mosca* e as variações *acertar na lata*, *acertar no olho do mosquito* e *acertar na muche* (termo do francês – mosca, inseto), esta última mais presente na variedade portuguesa (Almeida, 2022), ambas com a mesma significação idiomática: acertar com precisão.

O trabalho de Carneiro (2016) sobre expressões cristalizadas com *ser* e *estar* no português do Brasil sob a perspectiva da Léxico-Gramática (Gross, 1968, 1975) traz um ponto de grande valia para este trabalho: a consideração da substituição lexical (a autora trata como ruptura paradigmática), da inversão da ordem preferida e da elisão de um termo fixo do predicado na caracterização de EIs. A autora segue a mesma definição de expressão cristalizada que vimos em Vale (2001) e reforça que tais expressões são tratadas como um bloco, o que traz implicações sobre a fixação dos seus elementos constituintes.

Essa propriedade de fixação dos elementos, conforme vimos em outros autores anteriormente, permite alguns tipos de variação (Xatara, (1994, 1998) e Oliveira; Xatara (2008)), tais como: de tempo e de modo verbais; de gênero; de quantidade - superlativo/aumentativo/diminutivo; de número - singular ou plural; de usos de artigos; de pronomes; de inserção de advérbios; de inserção de negação; e de permuta lexical. Carneiro (2016) reforça que há casos em que a ruptura paradigmática não é possível, pois a substituição lexical de uma palavra por outra equivalente em seu eixo paradigmático desconfiguraria a cristalização da expressão, como por exemplo, *ser cara de pau/madeira**, *descascar o abacaxi/ananás/kiwi*/tangerina** (frutas do tipo ácida).

Outra propriedade de fixação é a impossibilidade de inversão dos elementos quando há a necessidade de manutenção da ordem preferida e cristalizada da expressão, como por exemplo *ser maior e vacinado/ser vacinado e maior**, *tirar o cavalo da chuva/tirar da chuva o cavalo**. A última propriedade de fixação é a impossibilidade de elisão do adjetivo, o que tomamos aqui como elisão de um termo fixo do predicado, quando este possuir mais de um complemento fixo, como em *estar de rabo preso/estar de rabo**, *fazer tempestade num copo d'água/fazer tempestade num copo**. Em nossos critérios de variação morfossintática, utilizamos essas propriedades de fixação com o intuito de compreender o grau de transparência e opacidade das EIs no português.

Bakker e Hengeveld (2012) entendem as EIs como expressões de significação “fixa”, com significado específico, que não é totalmente previsível pela soma das partes componentes e que, em certa medida, permitem variações, como de tempo e de modo. Essas variações delimitam diferentes graus de transparência e de fixação entre as EIs, fato relacionado com a composicionalidade dessas expressões.

Os autores utilizam os exemplos *to go down like a lead balloon* (descer como um balão de chumbo, i.e. apresentar uma ideia sem sucesso, sem aceitação das pessoas) e *to throw the sponge* (jogar a esponja, i.e. desistir, abandonar uma ideia ou situação) como exemplos de expressões que apresentam diferentes graus de denotação e conotação, sendo a última, uma expressão que apresenta maior identificabilidade de uma interpretação literal, o que ainda é distante de sua interpretação idiomática.

A interpretação metafórica, conforme vimos em Lakoff (1986) e em Lakoff e Johnson (1980), é um fenômeno de suma importância para o processo criativo das EIs e também para sua interpretação não-literal, pois a linguagem figurativa está presente na fala de poetas e crianças, que criam diversas frases conotativas que não são idiomatismos, mas que possuem valor metafórico. Contudo, Bakker e Hengeveld (2012) entendem que a metáfora não é o único meio de se explicar a complexidade das EIs, já que parte delas nem sempre apresenta caráter metafórico.

Em alguns casos, a origem da cena literal da expressão explica sua cena idiomática, como vimos no exemplo *para inglês ver*, e no exemplo dos autores *to bite the bullet*, que se originou no contexto militar, quando os médicos das forças armadas tinham que realizar uma operação dolorosa ou de alto risco em soldados e não havia anestésicos ou analgésicos para a dor e eles davam aos soldados balas de armamento para serem mordidas ou postas na boca, para que os soldados controlassem as dores. Tal expressão se espalhou na comunidade, cristalizou-se e tornou-se uma EI.

Para Bakker e Hengeveld (2012), as EIs contribuem para a abstração dos elementos e também para o enriquecimento na língua, uma vez que a partir de uma EI, novas expressões surgem, seja pela mistura de expressões, seja por novas significações e usos por parte de seus constituintes armazenados no léxico mental do falante:

As expressões idiomáticas constituem um tema intrigante na Linguística. Seus significados são muitas vezes imprevisíveis e não podem ser derivados de suas palavras constituintes; e sua forma sintática pode ser fixa ou cristalizada. Ao mesmo tempo, uma vez que muitas vezes se originam de metáforas, as expressões idiomáticas estão intimamente envolvidas com a mudança linguística (Bakker; Hengeveld, 2012, p. 278. Tradução nossa)⁸

⁸ Texto original: Idioms constitute an intriguing subject in linguistics. Their meanings are often unpredictable and cannot be derived from the constituent words; their syntactic form can be fixed or fossilised. At the same time, since they often

Os autores ainda reforçam que a mudança linguística implicada pelas EIs é fruto da inovação criativa do falante, assim como a mistura de EIs, ampliação de vocabulário por parte dos componentes, novas leituras metafóricas que se armazenam no léxico da língua. Tal situação também abrange os provérbios, estes tratados como expressões completamente fixas na forma e significado; e citações de personalidades famosas ou obras literárias, que também se cristalizam na comunidade falante através do seu uso diário.

Langlotz (2006), autor que tem como base de seus estudos os trabalhos de Nunberg et al. (1994), tendo sua obra como parte integrante do arcabouço teórico dos trabalhos de Keizer (2016, 2018), avalia a complexidade do conjunto de etapas no processo de ativação semântica de construções idiomáticas, tomando como exemplo a expressão *to grasp the nettle*.

O autor divide esse processo em duas grandes frentes de análise: significado literal (veículo padrão) e idiomático (tópico lexicalizado como alvo). O polo semântico (predicação), pertencente à análise do significado literal, é organizado da seguinte maneira: (i) significado literal composto pelas partes constituintes da expressão (*to grasp the nettle* - agarrar a urtiga); (ii) domínio conceitual dentro do qual a cena literal está incluída (cenário conceitual de agarrar a urtiga); (iii) subestrutura conceitual do primeiro constituinte (agente que agarra); (iv) subestrutura conceitual do segundo constituinte (urtiga como paciente); (v) esquema construcional que integra os constituintes (Sujeito - Verbo - Objeto).

Já o polo fonológico, também pertencente à análise do significado literal, é organizado da seguinte forma: (i) pólos fonológicos associados à composição literal dos componentes da expressão (agarrar a urtiga); (ii) polo fonológico associado ao primeiro constituinte (agarrar); (iii) polo fonológico associado ao segundo constituinte (urtiga).

Partindo para a análise do significado idiomático, o autor delimita três etapas de análise: (i) composição do significado idiomático (enfrentar um problema); (ii) cenário conceitual dentro do qual o significado idiomático está inserido (enfrentamento/resolução de problemas); e (iii) subestrutura conceitual saliente dos componentes (agarrar - enfrentar e urtiga - problema). Assim, é possível estabelecer correspondências entre as duas significações, resultando em uma unidade simbólica de configuração idiomática (Langlotz, 2006, p. 104).

O autor também reforça o que vimos em Lakoff e Johnson (1980) e Lakoff (1986) sobre metáforas conceituais: o papel da composição metafórica como fator motivacional no processo de

originate from metaphors, idioms are intimately involved with language change. (Bakker; Hengeveld, 2012, p. 278)

criação e codificação de EIs, e.g. *Ideas are objects, The mind is a container*, correspondendo a sub-metáforas de a mente ser a parte superior de um recipiente, este, o corpo (Langlotz, 2006, p. 122). Contudo, o autor também faz uma ressalva e admite expressões que não são constituídas por processos metafóricos.

Para o autor, tais expressões funcionam apenas como imagens únicas, baseadas em imagens metafóricas. Como exemplo, as metáforas de imagem, que estão subjacentes às seguintes expressões: *a couch potato* (pessoa preguiçosa), *a pretty kettle of fish* (situação embarçosa), nas quais a associação do significado idiomático com o significado literal é baseada nas imagens associadas com as fontes metafóricas (batata, sofá, chaleira de peixe) e as imagens esquemáticas correspondentes associadas às conceituações alvo da idiomaticidade (postura física da pessoa, situação caótica).

Para Langlotz (2006), há várias EIs que são tratadas como *blended metaphorical idioms*, ou cenas literais híbridas, fruto de evocações mentais construídas com base em subestruturas conceituais sem conexão com experiências baseadas na cena literal, e.g. *to have butterflies in one's stomach* (ansiedade e nervosismo).

Com isso, o autor delimita padrões de transparência das EIs com base em suas respectivas motivações de bases conceituais e de composicionalidade, das quais destacamos: (i) motivação global, referente à dimensão paradigmática *top-down*: a expressão pode ser motivada como um todo pela relação entre cena literal e idiomática; (ii) motivação constitucional, referente à dimensão paradigmática *top-down*: os constituintes são semanticamente autônomos porque possuem significados figurativos lexicalizados, motivados metaforicamente ou metonimicamente; (iii) composicionalidade, referente à dimensão sintagmática *bottom-up*: a composicionalidade das EIs pode ser parcial (o significado literal de um constituinte contribui diretamente para o significado idiomático), literal (os constituintes podem ser integrados para construir uma cena literal significativa), ou não-literal (os constituintes não podem ser integrados para construir uma cena literal significativa).

Com esses parâmetros alinhados, o autor caracteriza assim, quatro tipos de EIs aplicados em exemplos do inglês que seguem a estrutura V+NP (PP)⁹. Os quatro primeiros tipos foram agrupados na categoria de expressões com significado literal, completamente composicional e inconspícuo (tipos nucleares) e os quatro últimos na categoria expressões com significado literal conspícuo ou

⁹ V = verbo; NP = Sintagma Nominal e PP = Sintagma Preposicionado.

não-composicional (tipos marginais):

- (i) EIs com composicionalidade literal, com motivação global e isomorfismo figurativo-literal, com elevado grau de transparência (mas não tratado como completamente transparentes, pois incluem EIs que são motivadas por imagens metafóricas únicas, como *to spill the beans*), nas quais a constituição de base metafórica é diretamente aplicada aos constituintes, e.g. *to grasp the nettle* (i.e. urtiga como planta perigosa e a ação de agarrá-la com enfrentamento);
- (ii) EIs com composicionalidade literal, motivação global e constitucional, e isomorfismo figurativo-literal, e.g. *to swallow the bitter pill* (aceitar uma situação desagradável), na qual vemos que os constituintes possuem significados figurativos lexicalizados autônomos que também aparecem fora do contexto idiomático, com significado motivado metaforicamente, o que contribui para sua alta transparência (engolir como aceitar e pílula amarga como algo desagradável), com uma ativação conceitual rica e claramente estruturada (Langlotz, 2006, p. 129);
- (iii) EIs com composicionalidade literal e motivação global, e.g. *go round in circles* (conversar sobre determinado assunto sem conclusões ou objetividade), em que se observa o reflexo da cena literal na significação idiomática, mas em uma clara relação semântica analisável, esta com base em metonímia: andar em círculos implica em estagnação, mas não há um mapeamento direto entre os constituintes;
- (iv) EIs com composicionalidade literal, mas sem motivação ou isomorfismo, e.g. *to kick the bucket*, nesse caso não há conexão entre as cenas literal e idiomática, sendo EIs opacas, sem ativações metafóricas ou metonímicas entre seus constituintes, como chutar o balde e morrer, nas quais falta motivação semântica que conecte essas significações ou que direcione a alguma metáfora conceitual entre a ação de chutar o balde e o evento de morte;
- (v) EIs com misturas metafóricas, e.g. *to cook the books* (fraudar materiais contábeis), são expressões que possuem significação literal tratada como não experiencial na realidade, inviável, irreal, hiperbólica ou mesmo impossível. Tais cenas são imaginativas e sua relação com a significação idiomática se dá por meio de uma metáfora conceitual (livros como registros e cozinhar como ação que altera a forma física);

- (vi) EIs parcialmente composicionais, e.g. *to rain cats and dogs* (chuvas fortes), são expressões que possuem conexão com um dos constituintes da cena literal (chover), mas há sinais de violação (verbo *chover* é intransitivo) como resultado de mistura conceitual;
- (vii) EIs literalmente não-composicionais e construcionalmente idiossincráticas, e.g. *to shoot the breeze* (gastar tempo conversando sobre coisas não importantes): são EIs opacas que podem ser reconhecidas como construções altamente complexas com ativação da cena literal existente apenas em um elevado grau de abstração mental, como em atirar na brisa como forma pouco importante de atirar;
- (viii) EIs literalmente não-composicionais com morfemas *cranberry* (morfemas únicos ou também conhecido como termos fossilizados, que não possuem significado independente, mas servem para diferenciar palavras), e.g. *to blow the gaff* (revelar um segredo), em que a cena literal também é altamente imaginativa, e o morfema *cranberry gaff* (arpão) dificulta uma construção de sentido literal, o seu significado idiomático pode adquirir motivação pela ação do verbo *soprar*. O autor traz exemplos de pré e pós modificação do termo *gaff* que sugere sua significação idiomática (segredo) motivado pela presença do verbo (Langlotz, 2006, p. 132).

Tais apontamentos e categorizações de Langlotz (2006), assim como o que é apontado por Nunberg et al. (1994), contribuíram para as definições que Keizer (2016, 2018) realizou em seus estudos, base de nosso trabalho, e que serão apresentados no segundo capítulo.

Destaca-se como contribuição para nossa pesquisa, as definições de metáfora conceitual e ativação semântica das EIs. A ativação semântica nos auxiliou na compreensão do processo de análise dos significados literal (veículo) e idiomático (alvo), enquanto os apontamentos sobre metáfora conceitual nos ajudaram a compreender os processos de criação e de codificação das EIs, bem como observar as imagens metafóricas no processo de criação de imagens esquemáticas correspondentes, associadas às conceituações alvo da idiomaticidade. O autor também utilizou os conceitos de motivação e de composicionalidade para delimitar tipos de EIs, algo também refletido no trabalho de Keizer (2016). Essas contribuições do autor nos auxiliaram na formalização da proposta de um novo tipo de EI encontrada em variedades lusófonas do português: as expressões parcialmente motivadas por metáforas conceituais e decomponíveis.

Reforçamos a necessidade de se observar a definição cognitiva das EIs, com base no que vimos nas expressões com misturas metafóricas, que possuem significação literal tratada como não

experiencial na realidade, inviável, irreal, hiperbólica ou mesmo impossível. Vimos que tais cenas são imaginativas, e que sua relação com a significação idiomática se dá por meio de uma metáfora conceitual. Poderíamos simplesmente utilizar essa mesma denominação que Langlotz (2006) utilizou e tratá-la como um tipo de EI, mas propõe-se uma nova denominação, seguindo o que foi proposto por Keizer (2016, 2018).

Outro trabalho que dialoga com o conceito de metáforas na formação de expressões idiomáticas é a pesquisa de Vega Moreno (2007) sobre analisabilidade das EIs, termo que também remete à composicionalidade, pois também se refere ao modo como as partes constituintes contribuem para o significado idiomático. Uma das definições de EI apontada pela autora é o conceito de *abnormally decomposable idioms* (expressões idiomáticas anormalmente decomponíveis), que possui significado idiomático licenciado por uma metáfora convencional, em que um termo da EI ativa essa abstração metafórica convencionalizada na comunidade falante.

Para a autora, a analisabilidade das expressões está relacionada à percepção do falante em reconhecer o significado literal e o idiomático, e em verificar se a EI é ou não familiar, algo que se reflete em como o falante lida com variações sintáticas aceitáveis/acessíveis ou não. Tome-se como exemplos as sentenças em inglês *My father is **glued** to the computer for hours* (Meu pai está colado no computador por horas), na qual temos uma metáfora única no elemento em destaque, e *Since I work at university, I feel I am **swimming with sharks*** (Como eu trabalho na universidade, sinto que estou nadando com tubarões), com uma EI em destaque.

Em ambos os casos, o falante: (i) toma conceitos codificados literais como ponto de partida; (ii) segue um caminho de menor esforço; (iii) considera um subconjunto de suposições enciclopédicas, fazendo inferências com relevância seletiva em um processo orientado; (iv) procede à construção do conceito *AD-HOC* (*glued*); (v) após ajuste pragmático no nível frasal, com base no significado composicional (*to swim with sharks*); (vi) Tem seu próprio endereço conceitual vinculado a informações sobre estrutura sintática (VP [V [PP P NP]])¹⁰ e fonológica; (vii) a informação é vinculada a uma entrada enciclopédica especificando as suposições sobre a situação denotada (os tubarões são perigosos e imprevisíveis, se alguém nada com tubarões, sente que o perigo está à espreita, se alguém nada com tubarões, sente-se inseguro, sob ameaça, etc.).

Tais processos também dialogam com o que vimos em Langlotz (2006) no processo de ativação semântica entre cena literal (veículo) e cena idiomática (alvo), trabalho que contribui para a

¹⁰ VP – Sintagma Verbal; V – Verbo; PP – Sintagma Preposicionado; P – Sintagma; NP – Sintagma Nominal.

literatura das EIs e auxilia no mapeamento semântico e cognitivo desse fenômeno em diferentes graus de transparência e opacidade.

A seguir, são apresentados alguns aspectos sobre o processo de aprendizagem das EIs.

1.4: Aprendizagem das EIs

A aprendizagem de EIs por falantes não-nativos de uma dada língua pode ser uma tarefa um pouco mais difícil, devido à falta de familiaridade com essas construções na língua estrangeira. Já quando o falante encontra casos semelhantes em ambas as línguas, sua interpretação e uso podem ser facilitados e mais frequentes, gerando assim, expansão do vocabulário, e, ao mesmo tempo, interpretações equivocadas, como visto anteriormente, na expressão *quebrar as pernas* em inglês e português.

Irujo (1986) denomina de *idiomphobia* o receio de os falantes desenvolverem habilidades comunicativas com EIs em língua estrangeira, devido às dificuldades nos processos de transferência positiva e interferência de sua língua materna em relação à língua estrangeira.

Cieslicka (2010) realizou um experimento voluntário com trinta estudantes falantes nativos de polonês e aprendizes de inglês como segunda língua (L2) em nível avançado, com mais de doze anos de exposição à língua, todos estudantes no programa de mestrado na Universidade de Adam Mickiewicz, em Poznań, Polônia. O teste consistia em analisar o grau de ativação da cena literal no processo de efeito de preparação, em que os participantes analisaram EIs do inglês: dezesseis EIs que os participantes demonstraram alguma familiaridade foram apresentadas com um termo final retirado, e a tarefa dos participantes era analisar diferentes termos que possivelmente completariam a EI. Eles receberam uma pequena listagem de expressões como por exemplo *to kick the bucket* (sem o termo *bucket*), e uma tabela com diferentes palavras para completá-la: um termo semanticamente relacionado à EI (*water*) e outro não (*heat*), um termo fonologicamente relacionado (*budget*) e outro não (*lime*).

A proposta da investigação era entender o papel do significado literal provocado pelo conjunto de palavras para cada EI e se existiria alguma diferença nos padrões de ativação semântica e fonológica entre EIs decomponíveis e não-decomponíveis. Dentre os resultados, destaca-se a influência da ativação da cena literal das EIs, com as palavras semanticamente relacionadas apresentando tempo de reação mais rápido do que com a ativação de palavras fonologicamente relacionadas.

Tal resultado reforça os estudos de Cacciari e Tabossi (1988), que identificaram que o significado literal da última palavra de uma EI é ativado imediatamente após sua inserção e permanece ativo por cerca de 300 milissegundos na mente do falante; após esse intervalo de tempo, o significado idiomático é ativado através de elementos-chave, como *bucket* em *to kick the bucket*. Tais apontamentos fogem um pouco à ideia de Gibbs (1980), que entende que uma EI tem sua interpretação conotativa reconhecida ou recuperada de modo mais rápido do que sua significação literal, assumindo assim que o significado idiomático é entendido desde o começo, ignorando o significado literal.

Concluindo, neste capítulo foram apresentados conceitos e definições sobre o objeto de estudo aqui focado (EIs), com base em uma literatura consagrada de diferentes vertentes teóricas. Tais contribuições nos serviram de base para uma compreensão aprofundada sobre as EIs, abarcando: suas propriedades formais e funcionais; os processos psicolinguísticos e cognitivos de sua criação e decodificação; sua dicionarização; a base experiencial do falante e os processos de interpretação metafórica e de seu sistema conceitual metafórico; sua convencionalidade e frequência de uso; sua empregabilidade na oralidade; seu funcionamento como um bloco de pensamento; a existência de casos de variação; seu valor expressivo como estereótipos culturais e metáforas universais; além da apresentação de reflexões sobre a Fraseologia e sobre a aprendizagem das EIs que, no conjunto do referencial discutido, moldam seu funcionamento e caracterização na literatura sobre o tema.

Seguimos para o segundo capítulo, sobre a Gramática Discursivo-Funcional (Hengeveld; Mackenzie, 2008). Nele também retomaremos o conceito das EIs na GDF, com o aporte teórico das obras de Keizer (2016, 2018) e Dik (1981, 1997) e outros autores.

CAPÍTULO 02: A GRAMÁTICA DISCURSIVO-FUNCIONAL

A Gramática Discursivo-Funcional é definida como um Componente Gramatical de um modelo de interação verbal disposto em uma organização *top-down* – partindo da intenção comunicativa do falante até a articulação da Expressão Linguística, na qual o Componente Gramatical se relaciona ao Componente Conceitual, referente às intenções comunicativas e conceitualizações extralinguísticas; ao Componente Contextual, que se refere ao contexto real do evento de fala e a relação social entre os participantes; e ao Componente Articulatório (ou Componente de Saída), responsável pelas expressões acústicas, ortográficas, visuoespaciais ou linguísticas baseadas na informação corrente.

Hengeveld e Mackenzie (2008) destacam dois processos de construção de enunciados nessa organização *top-down*: Formulação e Codificação. A Formulação baseia-se nas regras que determinam aquilo que constitui as representações semânticas e pragmáticas subjacentes válidas em uma língua, com base na interação entre três processos: seleção de moldes apropriados nos níveis Interpessoal e Representacional; a inserção de lexemas apropriados nesses moldes, e a aplicação de operadores que simbolizam as distinções gramaticais na língua.

Já a Codificação refere-se às regras que convertem as representações semânticas e pragmáticas em representações morfossintáticas e fonológicas, o que também envolve três processos internos: seleção de padrões apropriados nos níveis Morfossintático e Fonológico; a inserção de morfemas gramaticais livres e vinculados, e a aplicação de operadores morfossintáticos e fonológicos no processo de articulação de saída.

Para os autores, a GDF é um modelo gramatical de intenções codificadas e de conceitualizações, sendo considerada uma teoria que busca entender como as unidades linguísticas são estruturadas em termos referentes ao mundo que elas descrevem e às intenções comunicativas envolvidas em sua produção, em uma implementação dinâmica de uma gramática que permite ao analista entender a sequência de etapas para a investigação de diversos fenômenos nas línguas.

A GDF pode ser vista, conforme Butler (2003), como uma gramática estrutural–funcional, pois sua proposta é a análise da língua moldada pelo uso, e entende a língua como um sistema que deve ser descrito e correlacionado com funções do discurso. Butler e Hengeveld (2009) reforçam essa proposta ao entender que a essência das teorias funcionalistas está em analisar as propriedades do sistema linguístico e sua conexão com os fatores cognitivos, socioculturais e históricos da língua.

Butler (2003) entende que a GDF ocupa uma posição intermediária entre a Gramática Sistêmico-Funcional (Halliday; Mathiessen, 2004), que está mais próxima do funcionalismo radical, e a Gramática de Papel e Referência (Van Valin, 2001), que está mais próxima do formalismo radical. Stassi Sé (2012) discorre brevemente sobre as principais escolas funcionalistas:

As vertentes europeias são representadas pelas seguintes perspectivas (i) o trabalho da *Escola de Praga* (LUELSDORFF 1994, ČERMAK 1995); (ii) a *Gramática Funcional* (DIK, 1997); (iii) a *Gramática Sistêmico-Funcional* (HALLIDAY; MATTHIESSEN 1999, 2004); (iv) a *Gramática da Palavra* (HUDSON 2007); (v) a *Gramática Discursivo-Funcional* (HENGEVELD; MACKENZIE, 2008); e (vi) o *Modelo de Construção Lexical* (IBÁÑEZ; USÓN, 2008). Já as vertentes americanas compreendem: (i) o trabalho da *Escola Colúmbia* (CONTINI-MORAVA; GOLDBERG, 1995; CONTINI MORAVA, KIRSNER; RODRÍGUEZ-BACHIOLE 2004, HUFFMAN 2001); (ii) a *Gramática de Papel e Referência* (VAN VALIN, 2005); e (iii) uma série de abordagens que podem ser compreendidas em um mesmo conjunto por serem “teorias da linguagem baseadas no uso” (GIVÓN 1995, HOPPER 1998, BARLOW AND KEMMER, 2000, BYBEE; HOPPER, 2001 e BYBEE 2006) (Stassi Sé, 2012, p. 27-28)

Em comparação à vertente precursora, a Gramática Funcional de Dik (1981), essa nova versão se torna uma teoria inovadora do ponto de vista analítico e operacional, pois, de acordo com Hengeveld e Mackenzie (2008), a GDF abarca uma nova visão sobre as unidades a serem analisadas funcionalmente, porções oracionais maiores e analisadas em diferentes níveis e camadas, situadas nos campos das teorias funcionalmente orientadas, e que também tem como intuito a criação de um sistema de representação formal.

São apontadas como principais diferenças entre GF e GDF os seguintes avanços: (i) organização *top-down*; (ii) Ato Discursivo como unidade básica de análise; (iii) inclusão de representações morfossintáticas e fonológicas, e (iv) relação sistemática com os Componentes Conceitual, Contextual e de Saída.

A teoria tem como unidade de análise linguística o Ato Discursivo, que, conforme os autores, é mais um fator que diferencia a GDF de outras teorias linguísticas (como por exemplo, a Gramática Funcional) que tem a Oração como unidade de análise:

Os Atos Discursivos combinam-se em estruturas discursivas maiores, como Movimentos. Essas estruturas maiores respondem pelas unidades maiores do que a Oração. Por outro lado, os Atos Discursivos podem se manifestar na língua como orações, mas também como fragmentos de orações totalmente gramaticais, frases ou palavras. O último ponto é crucial: o modelo gramatical deve ser capaz de mapear a unidade do Ato Discursivo em unidades de

vários tipos. (Hengeveld; Mackenzie, 2008, p. 4¹¹. Tradução nossa)

Vejamos a figura 4, que mostra a nova arquitetura da GDF (no prelo¹²):

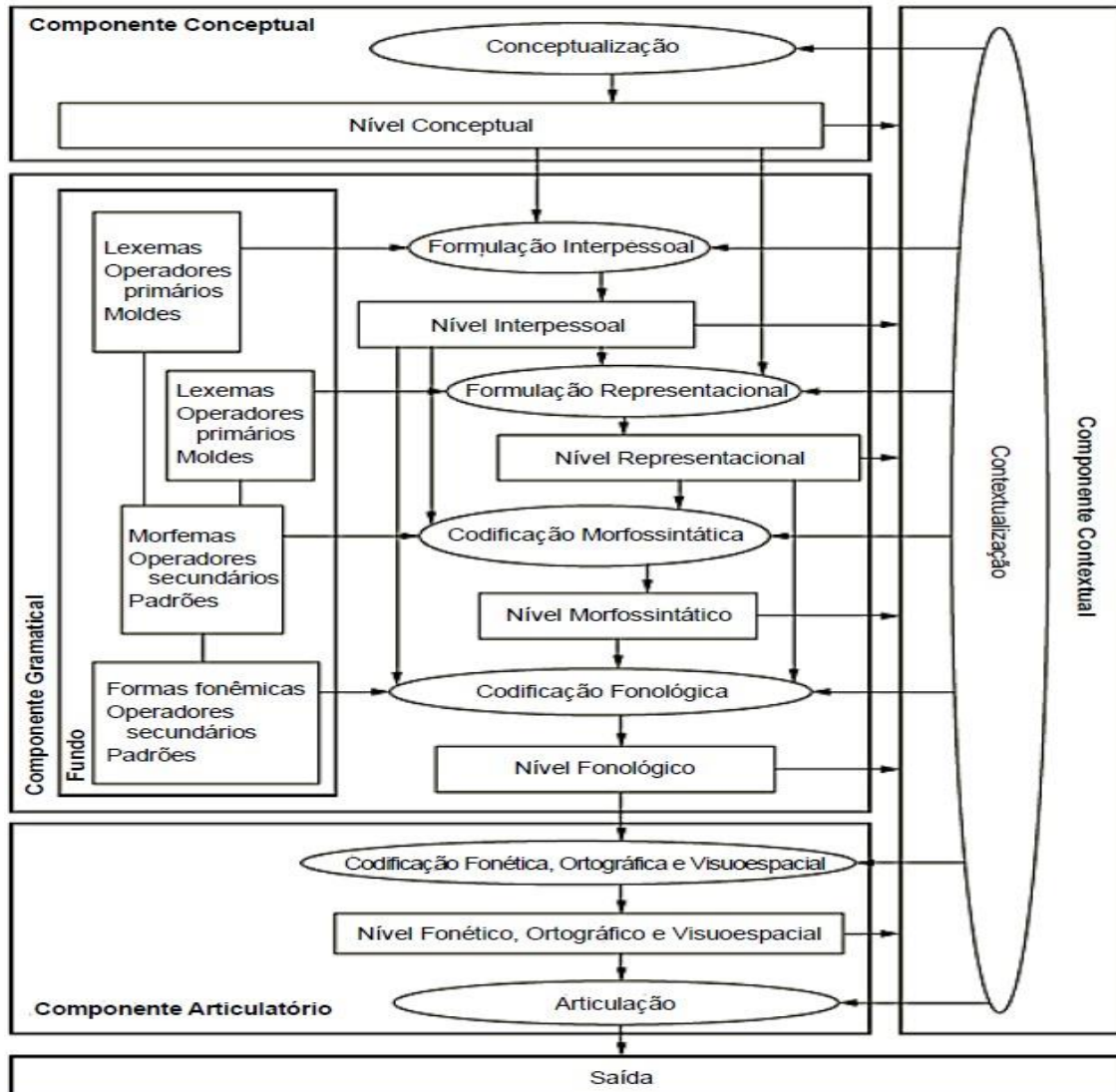


Figura 4: Nova arquitetura da GDF (Hengeveld et al., no prelo)

¹¹ Texto original: Discourse Acts combine into larger discourse structures, such as Moves. These larger structures account for the units larger than the individual clause discussed above. On the other hand, Discourse Acts may be manifested in language as clauses, but also as fully grammatical clause fragments, phrases or words. The latter point is a crucial one: it requires the grammatical model to be capable of mapping the unit of Discourse Act onto morphosyntactic units of various kinds. (Hengeveld; Mackenzie, 2008, p. 4)

¹² O Anexo 1 contém o quadro com arquitetura anterior da GDF, encontrada na obra de 2008 (Hengeveld; Mackenzie, 2008, p. 13).

Os círculos contêm operações, as caixas contêm os primitivos, os retângulos contêm os níveis de representação, e as setas, as relações entre os elementos da arquitetura da teoria. Tal configuração da teoria teve influência nos trabalhos de Levelt (1989) e em sua distinção de três módulos no processo de produção discursiva: Conceitualizador, Formulador e Articulador, que correspondem respectivamente aos Componentes Conceitual, Gramatical e Articulatório (ou de Saída).

Todos os quatro níveis de representação da GDF são hierarquicamente organizados em camadas, com sua estrutura geral organizada da seguinte forma: $(\pi v_1: [head(v_1) \Phi]: [\sigma(v_1) \Phi]) \Phi$, em que v_1 é uma variável da camada relevante, *head* é o núcleo, que toma a variável como seu argumento, e pode ser restringido por um modificador σ , que toma a variável como seu argumento. Já a camada pode ser especificada por um operador π , que se aplica à própria unidade, e carregar uma função Φ , que se aplica a uma mesma unidade ou a diferentes unidades da mesma camada. O núcleo e os modificadores representam estratégias lexicais, e os operadores e função representam estratégias gramaticais.

O Componente Conceitual é tido como a força motriz, onde se concentra o material interativo e ideacional do discurso sob análise. A intenção comunicativa do falante pode ser de natureza pragmática, pertencente ao Nível Interpessoal (NI), ou de natureza semântica, especificada no Nível Representacional (NR), ambas como partes do processo de Formulação. É através do Componente Conceitual que ocorre a conversão de informação conceitual pré-linguística em representações pragmáticas e semânticas linguisticamente relevantes e permitidas na gramática de uma língua.

É importante destacar que esse componente não contém todas as informações conceituais requeridas para a produção de Expressões Linguísticas, mas contém aspectos cognitivos diretamente relacionados com a intenção comunicativa do falante. Por exemplo, quando o falante deseja informar ao seu interlocutor que um amigo está doente (*João está doente*), há a possibilidade de inserir informações adicionais, como mitigação (*Eu receio que João esteja doente*), discurso indireto (*Maria me contou que o João está doente*), etc.

Nesse cenário, outras informações conceituais, que fazem parte do conhecimento particular dos participantes, são ativadas no momento do discurso, como o conhecimento de quem é João, conceito do estado de enfermidade, relação dos falantes com João, etc.

O Componente Contextual abarca as representações de natureza contextual do discurso, ou seja, o contexto multifacetado em que o falante está inserido, tais como aspectos socioculturais, a sequência dos enunciados em uso, informações de longo prazo e a relação dos participantes do

discurso. Tais aspectos auxiliam na compreensão das escolhas do falante, como formalidade/informalidade, encadeamento de narrativas, anáfora, catáfora, etc. A situação informacional tem um papel importante na produção de Expressões Linguísticas, como o uso de dêiticos, localização de eventos e entidades no tempo e espaço.

Os níveis de representação da Gramática alimentam o Componente Contextual com informações relevantes, tais como referentes e entidades discursivas previamente introduzidas no discurso, avaliação dos antecedentes, participantes do discurso e sua influência na composição dos Atos Discursivos subsequentes.

Vejamos em (2-1) um exemplo extraído do trabalho de Keizer (2015), traduzido em (2-1'), no qual podemos observar a descrição dos tempos verbais – presente (avistamento do animal no tempo de fala), passado (fala sobre o mesmo evento em um tempo posterior) e futuro (predicação sobre um possível avistamento futuro do animal), a localização espacial dos eventos e a interação dos participantes:

(2- 1) A: Look! A deer!

B: Where?

A: There - it's hiding behind those bushes, Now it's gone. Did you see it?

B: No, I'm afraid I missed it. But I'm sure we'll see another one. (Keizer, 2015, p. 26)

(2- 1') A: Olhe! Um veado!

B: Onde?

A: Lá – ele está escondido atrás daqueles arbustos. Agora se foi. Você o viu?

B: Não, Eu acho que o perdi. Mas tenho certeza de que veremos outro. (Tradução nossa)

O Componente Articulatório é responsável pela conversão final das estruturas do Componente Gramatical em estruturas acústicas, provenientes de movimentos respiratórios, das estruturas laríngeas e supralaríngeas, visuoespaciais, ou de sinais, tais como movimentos corporais, e também a produção de expressões ortográficas. Sua principal função é a tradução de informação digital em informação analógica, sendo, dessa forma, o Componente responsável pela conjuntura do andamento de fala, como ritmo, tempo, pausas, ritmos de escrita, elevações de tons, etc.

Os primitivos da esfera da Formulação contém os moldes (*frames*) nos quais se enquadram as possíveis combinações de elementos do NI – natureza expressiva, comunicativa e função retórica dos Atos Discursivos, e NR – valências quantitativas e qualitativas das combinações das categorias semânticas disponíveis, de acordo com o inventário de cada língua - o conjunto de Lexemas – unidades independentes que possuem forma fonêmica e denotam entidades de natureza pragmática e semântica em uma dada língua. Através do acesso a esse conjunto de Lexemas, o Falante seleciona quais Lexemas utilizar para, assim, dimensioná-los nos possíveis moldes de combinações pragmática e semântica de acordo com sua intenção comunicativa.

Já os primitivos do Nível Morfossintático (NM) são organizados em padrões (*templates*) para as Expressões Linguísticas, Orações, Sintagmas e Palavras, que são armazenadas no inventário de uma língua, assim como os Morfemas gramaticais, tais como afixos, partículas e auxiliares. Os primitivos do Nível Fonológico (NF) são organizados em padrões para os Enunciados, Sintagmas Entonacionais, Frases Fonológicas, Palavras Fonológicas, Pés e Sílabas, assim como as formas supletivas e os operadores fonológicos, como elevação e descida da entonação no processo de fala.

Hengeveld e Mackenzie (2008) entendem que o usuário da língua é um conhecedor tanto das unidades funcionais, como também das unidades formais, estando capacitado a realizar diferentes combinações de acordo com seu propósito comunicativo. Para os autores, esse conhecimento do falante apresenta um elevado grau de estabilidade, podendo assim ser comparado entre as línguas, revelando as tendências universais na estrutura linguística.

A GDF se debruça a investigar a relação entre a dimensão individual e a psicológica do falante, e a conexão dessas dimensões com a dimensão social em que esse falante está inserido. Tal proposta tem uma forte orientação tipológica como direção teórica para aplicação linguística, relacionando as relações textuais analisadas nos níveis. A exemplo disso, no NI estão os Atos Discursivos (A) agrupados em Movimentos (M); no NR estão os Estado-de-Coisas (e) agrupados em Episódios (ep), e assim por diante.

Hengeveld e Mackenzie (2008) entendem que a abordagem tipológica da GDF, estruturada em níveis e camadas, pode ajudar o linguista a compreender fenômenos linguísticos complexos dentro da uma perspectiva minuciosamente estruturada e ordenada:

A GDF procura fornecer uma estrutura para os trabalhos tipológico e linguístico contrastivo e de contato linguístico que é neutro em relação a qualquer tipo de linguagem específica. Ao mesmo tempo, a GDF pode ser empregada para a descrição de línguas individuais (mas

sempre tendo em vista as implicações para outras línguas), bem como o crescimento e declínio das línguas em contextos de aquisição e desgaste. (Hengeveld; Mackenzie, 2008, p. 42. Tradução nossa)¹³

Vejamos o detalhamento dos quatro níveis de representação da teoria, com exemplificações de uso provenientes de diferentes línguas no mundo. Começemos pelo nível mais alto: NI.

2.1: Nível Interpessoal (NI)

O NI é o nível **acional** responsável pelas representações pragmáticas referentes às intenções comunicativas do falante, e também é responsável pelas distinções de Formulação na interação, baseadas em noções pragmáticas, retóricas e de conhecimentos e sentimentos do falante.

Na representação do NI, como a camada mais alta identificada na teoria, está o **Movimento** (M para *Move*) – definido como a menor unidade livre do discurso (Kroon, 1995 *apud* Hengeveld; Mackenzie, 2008, p. 50), visto como uma contribuição autônoma para a interação contínua, constituído de uma ação introdutória ou modificadora de tópicos discursivos. Conforme exemplo (2-2), em que há um Movimento (M_I) composto por dois Atos Discursivos – (A_J) e (A_I):

(2-2) Watch out, because there is a bull in the field.

(M_I: [(A_I: –watch out– (A_I)) (A_J: –there is a bull in the field– (A_J))_{Expl}] (M_I)) (Hengeveld; Mackenzie, 2008, p. 43)

Cuidado, porque há um touro no pasto. (Tradução nossa)

O **Ato Discursivo** (A para *Discourse Act*) é entendido como a unidade básica do discurso, definido como a menor unidade linguística presente no processo comunicativo, envolvendo os participantes do discurso – **Falante** (P₁ para *Speaker*) e **Ouvinte** (P₂ de *Adresse*); uma **Ilocução** (F

¹³ Texto original: FDG seeks to provide a framework for typological and language-contrastive and language-contact work that is neutral with respect to any specific language type. At the same time, it can be employed for the description of individual languages (but always with a view to the implications for other languages), as well as the growth and decline of languages in contexts of acquisition and attrition. (Hengeveld; Mackenzie, 2008, p. 42)

para *Illocution*), que é a finalidade do ato verbal, e um **Conteúdo Comunicado** (C para *Communicated Content*), que é o conteúdo que o **Falante** deseja evocar ao **Ouvinte**).

Esse **Conteúdo Comunicado** é formado por **Subatos** (S para *Subact*), estes sendo formas de ação comunicativa do Falante, podendo ser de **Atribuição** (T para *Ascriptive Subact*), que é a tentativa de o Falante evocar uma propriedade, e/ou de **Referência** (R para *Referential Subact*), que é uma tentativa do Falante de evocar um referente como, por exemplo, o Subato Referencial *um carro azul*, que evoca as propriedades *carro* e *azul*, sendo *azul* a propriedade que restringe a aplicabilidade da propriedade *carro* (Hengeveld, Mackenzie, 2008). Apenas o Falante e a Ilocução são obrigatoriamente necessários para a execução de um Ato.

Hengeveld e Mackenzie (2008) discorrem sobre diferentes tipos de Atos manifestados em um Ato Discursivo, de acordo com sua funcionalidade: (i) Expressivos, (ii) de Conteúdo e (iii) Interativos. Os **Atos Expressivos** e os de **Conteúdo** estão relacionados à transmissão de emoção, sentimentos, informações pragmáticas e semânticas. Os **Atos Interativos** operam na manutenção verbal, sendo subordinados ao uso interacional entre Falante e Ouvinte e à relação comunicativa construída pelos participantes, a partir da qual o Falante busca chamar a atenção do Ouvinte. Para os autores, os Atos Interativos são normalmente compostos por elementos lexicais disponíveis em um conjunto restrito de situações, tais como formas de monitoramento da interação verbal e outras formas de execução do discurso em andamento.

Os **Atos Discursivos** podem expressar **funções retóricas**, utilizadas como estratégias do Falante para influenciar o Ouvinte a aceitar seus propósitos e intenções, tais como Orientação i.e. uma introdução de um referente importante para interpretação do Ato Discursivo, como em (2-3); Esclarecimento, que funciona como uma adição de uma informação para a correta interpretação da intenção comunicativa, como em (2-4); Concessão, que consiste na preservação da face do Falante ou de outrem ao apresentar uma reconsideração do Conteúdo Comunicado, como em (2-5); Motivação, utilizada para justificar uma Ilocução Declarativa, como em (2-6); Correção, que atua como uma autocorreção, como em (2-7); e *Aside*, que introduz uma informação relativa não restritiva direcionada ao Sujeito introduzido na Oração principal, como em (2-8):

(2-3) Orientação: *o mais novo*, esse estava sempre com, com a mãe (Moç86: Chuva: 30) (Pezatti, 2014, p. 83)

(2-4) Esclarecimento: dói muito para nós, *as senhoras* (Moç97: Mocidade: 56) (Pezatti, 2014, p. 84)

- (2-5) Concessão: e a gente teve junto *mesmo não tando no mesmo lado do do... da política*. (Moç97: Mocidade: 56) (Pezatti, 2014, p. 84)
- (2-6) Motivação: reconciliação eu não digo não é?, *porque até agora abandalam-me* (Ang97: JG: 113) (Pezatti, 2014, p. 84)
- (2-7) Correção: Eu gostaria de devolver o livro a sua mãe (*quero dizer*) a sua irmã. (Hengeveld; Mackenzie, 2008 apud Stassi Sé, 2012, p. 51)
- (2-8) *Aside*: Os alunos, *que trabalharam bastante*, passaram no exame. (Hengeveld; Mackenzie, 2008 apud Stassi Sé, 2012, p. 52)

A **Ilocução** pode indicar diversas intenções comunicativas do Falante, ao buscar associar as propriedades lexicais e formais do Ato. As Ilocuções podem ser expressões relacionadas, interjeições, verbos performativos, ilocuções abstratas ou performativos implícitos, conforme esquema: Ato_{Iloc} → DECL/INTER/ADMIR/IMPER, etc.+ F + D + C. Tome-se como exemplo, uma Ilocução Declarativa, que é uma espécie de instrução e/ou transmissão de informação pragmática do Falante para o Ouvinte:

- (2-9) [...] a água vem de um rio... (Bra80: Fazenda) (Pezatti, 2010, p. 15)

Em sua representação formal, as siglas são representadas por letras maiúsculas. Segue abaixo o modelo de relação hierárquica do NI:

(π M1: [Movimento
(π A1:[Ato Discursivo
(π F1: ILL (F1): Σ (F1))	Ilocução
(π P1: ... (P1): Σ (P1))S	Falante
(π P2: ... (P2): Σ (P2))A	Ouvinte
(π C1:[Conteúdo Comunicado
(π T1: [...] (T1): Σ (T1)) Φ	Subato de Atribuição
(π R1: [...] (R1): Σ (R1)) Φ	Subato de Referência
] (C1): Σ (C1)) Φ	Conteúdo Comunicado
] (A1): Σ (A1)) Φ	Ato Discursivo
] (M1): Σ (M1))	Movimento

Figura 5: Hierarquia no Nível Interpessoal (Hengeveld; Mackenzie, 2010, p. 18)

Quando temos os excertos reais das Expressões Linguísticas, a representação sequencial de camadas é dada em ordem alfabética, e quando é colocada uma representação geral, utiliza-se a ordem numérica, como em (M₁: [(A₁) (A₂)_{Expl}] (M₁)), esta referente ao exemplo (2-2) (M_I: [(A_I: –watch out– (A_I) (A_J: –there is a bull in the field– (A_J))_{Expl}] (M_I)) (Hengeveld; Mackenzie, 2008). Como podemos ver na figura acima, o uso de **colchetes** demarca a relação interna dos elementos não-hierárquicos e que, em conjunto, são subordinados à camada superior.

No trabalho de Stassi Sé (2012) é proposta a inclusão de uma nova camada no NI da teoria: a camada do **Discurso**¹⁴, que engloba porções textuais maiores, entre Movimentos sequenciados em um mesmo ou em diferentes turnos de fala, para a análise de fenômenos orientados discursivamente e motivados por funções interacionais. Vejamos abaixo o excerto extraído do CRPC (2016), em que a autora identificou a função interacional de Resgate na construção iniciada pela conjunção *como*:

(2-10) L1- e aos treze anos lembra-se de algum drama, eh, como é que eram as noites, a sua sobrevivência e como é que conseguiu até hoje manter-se?

L2 -> bem, eh, naquel[...], na[...], naquele tempo, portanto, **como eu já antes tivera dito**, não é, eu dormia hoje aqui, amanhã ali, às vezes, eh, portanto, dormia mesmo na rua assim ao ar livre. e mais tarde eu verifiquei que estava, estava, estava mesmo a caminhar muito mal, não é, estava que nem um menino de rua e no entanto eu tive de fazer o esforço, poder trabalhar um bocadinho para ver se conseguisse um tecto para mim (Ang97:JovemGaspar)

Essa nova camada visa proporcionar à teoria recursos que possibilitem a análise de Movimentos dependentes pragmaticamente, que estejam em uma relação subordinada discursiva, conectados por conjunções que os conecte aos Movimentos em funções interacionais de Resgate (evocação de uma informação anteriormente partilhada, resgatada pela conjunção *como*), Salvaguarda (preservação da face do falante em meio a alguma insegurança em relação ao conteúdo em discurso, manifestada pela conjunção *se*), Adendo (inserção de uma nova informação manifestada pelas conjunções *embora (que)*, *mesmo (que)* e *apesar de (que)*) e Transição (mudança de assunto expressa pela conjunção *porque*). Essa camada será acessada em casos de referência anafórica e catafórica, identificados em porções textuais maiores entre turnos de fala dos participantes no quarto capítulo.

¹⁴ Tal inclusão na esfera da Formulação também exige uma nova inclusão na esfera da Codificação no NM – a camada do **Texto**, conforme será visto mais adiante no NM.

Além das propriedades de interação de natureza retórica, temos as de natureza **pragmática**, i.e., a modelagem das mensagens do falante em relação às expectativas sobre o ouvinte. No Nível Interpessoal, essas estratégias interpessoais, gramaticais e lexicais constituem os Operadores e Modificadores, além das Funções Pragmáticas presentes no NI.

Os Operadores Interpessoais representam estratégias gramaticais que especificam uma camada e se aplicam a uma unidade em si mesma, constituindo traços abstratos e manifestados na esfera da Codificação como **Palavras Gramaticais**. No NI, os Operadores ocorrem na interação entre Falante e Ouvinte, em posições periféricas, podendo assumir uma posição absoluta ou relativa. Já a **Ênfase** é uma estratégia do Falante em atingir seu propósito comunicativo, através de uma intensificação lexical ou gramatical em algum constituinte ou na Expressão Linguística inteira.

O Operador de **Polidez** indica a intenção do falante em ser educado, estando presente nos Subatos, antes do Núcleo especificado. O Operador **Aproximação** representa um termo que indica aproximação da real intenção do Falante, posicionado antes do núcleo que especifica. O Operador de **Mitigação** é usado para delimitar o compromisso do Falante em relação ao que é dito, para preservação de sua face.

Os Operadores Interpessoais também podem atuar como **Funções pragmáticas**, como de **Foco** - posição absoluta final, onde o termo é codificado como foco da sentença, podendo também ser marcado como **Foco-Ser**, onde o Falante fornece ao ouvinte o Operador-ser (verbo) para marcar o constituinte como Foco, e também como **Contraste**, que marca o desejo do Falante em destacar diferenças entre os Conteúdos Comunicados ou entre informações previamente enunciadas, como por exemplo, o uso de *however* (no entanto, porém), para marcar contraste no exemplo no Ato Discursivo subsequente - *Rangers lost* (Rangers perdeu):

(2-11) Celtic won; *however*, Rangers lost. (Hengeveld; Mackenzie, 2008, p. 60)

Celtic venceu; no entanto, Rangers perdeu. (Tradução nossa)

Os **modificadores interpessoais**, opcionais e restritos ao Núcleo de uma camada, são constituintes lexicais que refletem o papel de uma unidade linguística na interação entre os participantes e modificam o Ato Discursivo, podendo escopar uma camada toda, dependendo de sua posição na sentença, como, por exemplo, um Ato Discursivo, Ilocuções, Conteúdo Comunicado e Subatos.

Sua atuação modificadora pode ocorrer em diferentes funções e camadas, como Modo (*lentamente, à parte*), Conclusão (*portanto, pois*), Sequenciação (*primeiro, segundo*), Exemplificação (*por exemplo*), Atitude (*felizmente, infelizmente*), Divergência (*contudo, no entanto*) e Ênfase (*caramba* em caso de raiva, espanto), e.g., em (2-12) no Ato Discursivo com ilocução imperativa *Responda-me caramba!*:

(2-12) Answer me *dammit!* (Hengeveld; Mackenzie, 2008, p. 59)

Responda-me caramba! (Tradução nossa)

Finalizam-se aqui as descrições do NI, também reconhecido como o nível da evocação, e seguimos agora para o segundo nível da esfera da Formulação, o Nível Representacional.

2.2: Nível Representacional (NR)

Este nível está relacionado com os aspectos semânticos das unidades linguísticas, sendo responsável pela designação, sendo tratado como o nível da **denotação**. O termo “semântica” fica restrito aos meios pelos quais uma língua se relaciona com os mundos possíveis que ela descreve. Para Hengeveld e Mackenzie (2008), as camadas de maior relevância no Nível Representacional são definidas em termos das categorias semânticas que elas designam, e da compreensão de como as categorias semânticas auxiliam no processo de descrição de um fenômeno na língua.

Em sua representação formal, temos suas siglas representadas por letras minúsculas. Segue, abaixo, a relação hierárquica no Nível Representacional:

(π p1:	Conteúdo Proposicional
(π ep1:	Episódio
(π e1:	Estado-de-Coisas
(π fl:	Propriedade Configuracional
[
(π fl: \blacklozenge (fl): [σ (fl) Φ])	Propriedade Lexical
(π x1: \blacklozenge (x1): [σ (x1) Φ]) Φ	Indivíduo

] (f1): [σ (f1)Φ]	Propriedade Configuracional
(e1)Φ]: [σ (e1)Φ]	Estado-de-Coisas
(ep1): [σ (ep1)Φ]	Episódio
(p1): [σ (p1)Φ]	Conteúdo Proposicional

Figura 6: Hierarquia no Nível Representacional (Hengeveld; Mackenzie, 2010, p. 21)

Hengeveld e Mackenzie (2008), com base nos trabalhos de Lyons (1979), delimitam quatro diferentes categorias semânticas organizadas em ordens de entidades: a **primeira ordem** refere-se aos **Indivíduos (x)**, que podem ser localizados no espaço e podem ser avaliados em termos de existência, como por exemplo, *cadeira, mesa*. O **Estado-de-Coisas (e)** é uma entidade de **segunda ordem**, que pode ser localizado no tempo e espaço e ser avaliado em termos de realidade, e.g., *um encontro*.

O **Conteúdo Proposicional (p)** é uma entidade de **terceira ordem** e está relacionado com as construções mentais, como desejos, crenças e conhecimentos do falante. Essas construções mentais podem ser factuais (relacionadas ao mundo real) e não factuais (mundo imaginário), podendo ser qualificadas por atitudes proposicionais (como certeza ou dúvida), ou também por inferências, conhecimento comum partilhado, ou evidências. Vejamos o exemplo:

(2-13) *Jenny acreditava que/esperava que fosse para casa porque talvez sua mãe fosse visitá-la.*
(Hengeveld; Mackenzie, 2010, p. 13-14)

Hengeveld e Mackenzie (2008) também discorrem sobre a existência de subclasses de categorias semânticas, tais como **quantidade (q)**, **tempo (t)**, **modo (m)**, **razão (r)** e **lugar (l)**: *litro, semana, maneira, razão e topo*.

A camada de **Estado-de-Coisas (e)** está relacionada a eventos e estados que podem ser localizados no tempo e no espaço e também ser avaliados em termos de seu estatuto de realidade. Pode-se dizer que o Estado-de-Coisas “(não) ocorre”, “(não) acontece” ou “(não) é o caso” em algum ponto ou intervalo de tempo. Vejamos no exemplo (2-14), a combinação do tempo absoluto (presente no **Episódio (ep)**) com o tempo relativo (característica do Estado-de-Coisas):

(2-14) *Ontem Sheila saiu antes de jantar.* (Hengeveld; Mackenzie, 2010, p. 15)

A marcação de tempo absoluto (ep) dada pelo advérbio *ontem* é válida também para os dois Estado-de-Coisas (e) do exemplo acima, na medida em que as duas ações fazem parte do mesmo **Episódio** (ep), separadas e especificadas em sua relação temporal pelo lexema *antes* (locução conjuntiva).

A **Propriedade Configuracional (f)** é moldada por categorias semânticas que estabelecem uma relação não-hierárquica entre si. Essas categorias semânticas podem ser compostas de **Indivíduos** (objetos concretos que podem ser localizados no espaço), **Propriedades Lexicais** – que não têm existência independente e só podem ser avaliadas em termos de seu funcionamento em conjunto a outros termos da entidade, localização, tempo, modo, razão e quantidade. Vejamos o exemplo a seguir, no qual os dois Estados-de-Coisas (e) se relacionam e constituem (f):

(2-15) *The heavy rainfall (e) caused (f) a lot of damage (e).* (Hengeveld; Mackenzie, 2008, p. 181)

As fortes chuvas (e) causaram (f) muitos estragos (e). (Tradução nossa)

Vejamos o exemplo (2-16) com o enunciado *The man went to the countryside* (*O homem foi para o campo*), no qual temos o Estado-de-Coisas (e_i) caracterizado pela Propriedade Configuracional (f_i), i.e, Indivíduo ir para o campo. O Indivíduo (x_i) é caracterizado pela Propriedade (f_k) *man* e está localizado no lugar (l_i), este que por sua vez é caracterizado pela Propriedade (f_i) *countryside*.

(2-16) (e_i: [(f_i: [(f_j: go (f_j)) (x_i: [(f_k: man (f_k)) (x_i)φ])φ] (l_i: [(f_i: countryside (f_i)) (l_i)φ])φ] (f_i)) (e_i)φ])
(Hengeveld; Mackenzie, 2008, p. 140)

O homem foi para o campo (Tradução nossa)

Os modificadores de NR podem atuar, por exemplo, na camada do Conteúdo Proposicional (p), expressando atitudes proposicionais, como graus de comprometimento com o estado racional de (p), ou a sua especificação de fonte, como por exemplo: *Probably/evidently/hopefully/undoubtedly Sheila is ill* (Hengeveld; Mackenzie, 2008).

Já os operadores de NR podem expressar marcas de evidencialidade do Conteúdo Proposicional, como no exemplo (2-17) traduzido para o português como *Você cantou* (o Falante não

viu, mas escutou alguém dizer) da língua Maricopa, que diferencia evidências sensoriais visuais e não-visuais:

- (2-17) Mashvar-ʔa.
 2SG.sing-nonvis
 ‘You sang (I heard it). (Hengeveld; Mackenzie, 2008, p.155)
 Você cantou (Eu ouvi isso) (Tradução nossa)

Após finalizar a apresentação dos Níveis Interpessoal e Representacional, que são responsáveis pelo processo de Formulação na teoria, seguimos para a esfera da Codificação, a começar pelo Nível Morfossintático.

2.3: Nível Morfossintático (NM)

É o nível responsável pelos aspectos estruturais da unidade linguística, por sua codificação e pelas distinções interpessoais e representacionais. A representação formal das camadas desse nível é feita com o uso da primeira letra maiúscula, seguida pela segunda letra em formato minúsculo. Segue abaixo a relação hierárquica entre as camadas desse terceiro nível:

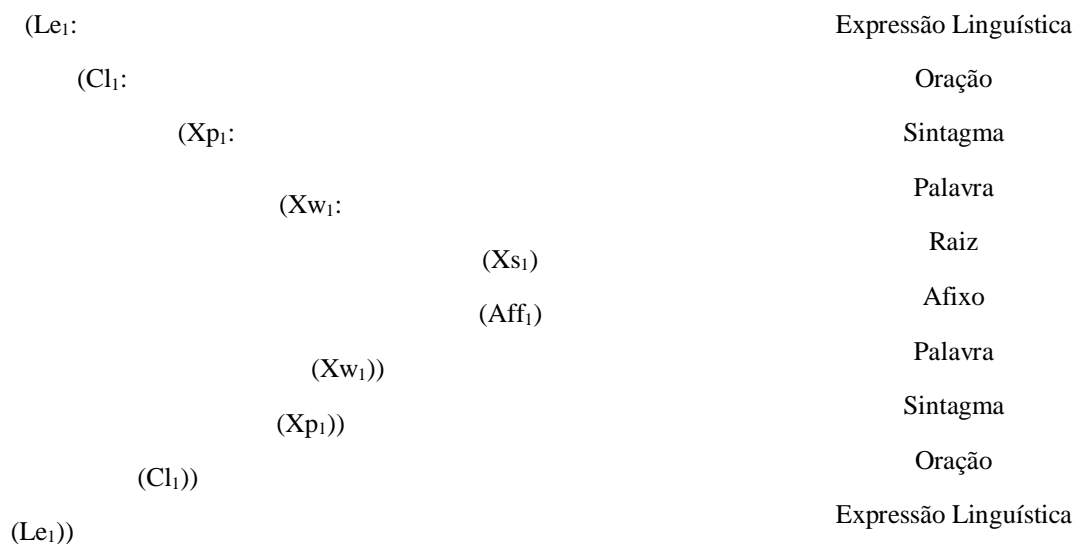


Figura 7: Hierarquia no Nível Morfossintático (Hengeveld; Mackenzie, 2010, p. 18)

Como o material de entrada do NM é proveniente dos níveis do processo de Formulação, não é permitida a adição ou subtração de nenhum material fornecido no processo anterior. A **Expressão Linguística (Le)** é o conjunto de ao menos uma unidade morfossintática, independente, que pode ser formada pela combinação de unidades presentes nas **Orações (Cl)**, que são grupos de **Palavras** ou Sintagmas, atuando como molde para estabelecer a ordem entre seus constituintes, **Sintagmas (Xp)**, que são a configuração em sequência de Palavras, Sintagmas e Orações encaixadas, possuem um núcleo lexical – pode ser verbal (Vp), nominal (Np), adjetival (Ap), adverbial (Advp) e adposicional (Adp), ou **Palavras (Xw)**, que representam a configuração sequenciada de morfemas, formada por **Raiz (Xs)** e **Afixos (Aff)**. Vejamos o exemplo (2-18):

(2-18) (Cl_i: [(Np_i: (Nw_i: he (Nw_i)) (Np_i))_{Subj} (Vp_i: [(Vw_i: can (Vw_i)) (Gw_i: not (Gw_i)) (Vp_j: (Vw_j: /swIm/ (Vw_j)) (Vp_j))] (Vp_i))] (Cl_i)) (Hengeveld; Mackenzie, 2008, p. 289)

Ele não pode nadar (Tradução nossa)

A Oração Cl_i é composta por um Sintagma Nominal (Np_i) com o sujeito *he* (ele) representado por uma Palavra Nominal (Nw_i), e dois Sintagmas Verbais (Vp) com suas respectivas Palavras Verbais (Vw) *can* (poder), esta acompanhada pela Palavra Gramatical (Gw_i) que representa negação (*not*), e o verbo *swim* (nadar).

Considerando a camada do Discurso no NI, aqui também consideramos a camada do Texto (Te) acima da camada (Le), pois há a necessidade de codificar as funções interacionais por meio das conjunções lexicais, utilizadas na subordinação adverbial, levando em conta que a relação entre Expressões Linguísticas constitui a camada do Texto (Stassi Sé, 2012).

Na camada das Palavras (Xw), tais unidades podem ser consideradas simples, constituídas de apenas um morfema (e.g. *feliz*), ou complexas, com mais de um morfema, como em palavras compostas (e.g. *gasômetro*, *girassol*, *passatempo*, *beija-flor*), compreendidas como a junção de dois Lexemas, e Palavras derivadas pela junção de afixos que adicionam um aspecto independente de significado (e.g. *in-feliz* – não feliz, *re-fazer* – fazer novamente).

É importante ressaltar que, para a GDF, há uma diferença entre Lexema e Palavra: uma única Palavra no NM pode corresponder a mais de um Lexema no NR. Há casos em determinadas línguas que não há essa distinção, e também existem Palavras Gramaticais que não apresentam Lexemas correspondentes, pois representam operadores, funções dos níveis NI e NR, ou também representam

suporte ou elementos vazios (*dummies*) (Hengeveld; Mackenzie, 2008). Os *dummies* são elementos que não correspondem a nenhuma unidade do NI ou do NR, mas são introduzidos no NM para preenchimento obrigatório de posições, como o pronome *it* no inglês em *It rained*. Nesse caso, *it* ocupa a função de sujeito, sem nenhuma representação nos níveis de Formulação.

Reforçando a diferença entre Palavra e Lexema na GDF, Keizer (2015) traz a EI *to kick the bucket*, entendida como uma expressão multi-palavra, como exemplo de incompatibilidade entre os dois termos. Essa EI é representada no NR como uma única entidade, listada no léxico pelo seu significado idiomático (morrer), uma vez que suas partes componentes não contribuem para o significado idiomático. Já no NM, sua representação é feita por três Palavras:

(2-19) He kicked the bucket (=he died)

NR: (p₁: (past ep₁: (e₁: [(f₁: kick-the-bucket_v) (x₁)_U][°])^{ep})^p)

NM: (Cl₁: [(Np₁: (Nw₁: Peter)) (Vp₁: (Vw₁: kick_{past})) (Np₂: [(Gw₁: the) (Nw₂: bucket)]^{Np})]^{Cl})
(Keizer, 2015, p. 236)

Ele chutou o balde. (= ele morreu) (Tradução nossa)

2.3.1: Ordenação no NM

Muitas operações no NM são motivadas funcionalmente por princípios ordenadores de **iconicidade**, **integridade de domínio** e **preservação das relações de escopo**, de acordo com sua relação com os dois níveis anteriores, o que contribui para o paralelismo entre as estruturas, o que fortalece a transparência e a facilidade de interpretação das estruturas linguísticas.

A **iconicidade** é o resultado da relação direta entre a ordem das unidades do NI e do NR e a ordem linear em que são expressas no NM, como o exemplo (2-20), na qual temos dois Atos Discursivos no NI codificados como duas Orações:

(2-20) The students worked very hard and they all passed the exam.

NI: (A₁) (A₂)

NM: (Cl₁) (Cl₂) (Keizer, 2015, p. 174)

Os estudantes trabalharam duro e todos passaram no exame. (Tradução nossa)

A **integridade de domínio** refere-se às unidades de informação que pertencem juntas aos níveis Interpessoal e Representacional e são dispostas próximas umas das outras na Codificação,

como no exemplo (2-21), onde temos uma oração restritiva colocada imediatamente após o antecedente, formando uma única unidade semântica no NR (Indivíduo x_1), e que no NM é expresso como um Sintagma Nominal (Np_1):

(2-21) *The students who worked very hard* all passed the exam.

NR: (x_1)

NM: (Np_1) (Keizer, 2015, p. 174)

Os estudantes que trabalharam duro passaram no exame. (Tradução nossa)

A **preservação das relações de escopo**, ou **estabilidade funcional**, é definida como a tendência de unidades com certas especificações interpessoais ou representacionais serem dispostas na mesma posição uma em relação às outras, como por exemplo no inglês, advérbios de lugar normalmente precedem advérbios de tempo:

(2-22) I saw Jane in London last Monday. (Keizer, 2015, p. 174)

Eu vi a Jane em Londres na última segunda-feira. (Tradução nossa)

Contudo, os três princípios podem sofrer **violações** em sua ordenação, como por exemplo, a violação do princípio de iconicidade na expressão linguística *Os estudantes, **que trabalharam duro**, passaram no exame*, na qual vemos que a oração relativa em destaque, que corresponde a um Ato Discursivo dependente, interrompe a Oração principal (Ato Discursivo nuclear).

O princípio de integridade de domínio no NM pode também sofrer alterações com base em estratégias de comunicação do falante, como por exemplo, a ilocução declarativa em (2-23), traduzida na mesma ordem como *I am now going to town*, na qual o advérbio de tempo *now* (Advp) interrompe a construção de Vp *am going*:

(2-23) I am now going to town. (Hengeveld; Mackenzie, 2008, p. 285)

Eu estou agora indo para a cidade (Tradução nossa)

Os princípios de ordenação dos constituintes no NM representam uma etapa de fundamental importância na compreensão de como o falante codifica e alinha seu discurso, com base nas funções pragmáticas, semânticas e sintáticas da língua. Essa ordenação sintática dos constituintes visa manifestar as relações entre funções pragmáticas e semânticas de cada língua, como um mecanismo de expressão usado pelo falante para expressar determinadas intenções comunicativas, estruturando Atos Discursivos e Expressões Linguísticas de determinada maneira.

A GDF trata dessa ordenação dos constituintes levando em consideração a diferença entre constituintes hierárquicos, i.e. pertencem a camadas diferentes, baseados em considerações de escopo, havendo orientação descendente de relação hierárquica; e não-hierárquicos, que atuam na mesma camada e têm estatuto dentro da Oração, denominados predicções nucleares, baseados no alinhamento.

Para Hengeveld e Mackenzie (2008), há quatro posições básicas para entender a linearização dos constituintes, que são: **Posição Inicial (P^I)**, **Posição Secundária (P²)**, **Posição Medial (P^M)** e **Posição Final (P^F)**, nas quais as posições periféricas (Inicial e Final) são psicologicamente mais salientes, reservadas para constituintes hierárquicos (Funções, Operadores e Modificadores – são colocados centripetamente, da margem para o centro), e a Posição Medial reservada para constituintes não hierárquicos (predicado e seus argumentos). Há a possibilidade do aparecimento de outras posições relativas – P^{I+1} (Pós Inicial), P^{M+1} (Pós Medial) e P^{F-1} (Posição Penúltima) quando as demais posições absolutas já estiverem preenchidas.

As posições marginais, que estão fora da Oração, são parte da Expressão Linguística, são tratadas como constituintes extraoracionais e, conforme Hengeveld e Mackenzie (2008), ocupam as posições P^{pre} (Pré-Oracional), P^{centro} (Oração) e P^{pos} (Pós-Oracional), classificadas para distinguir P^I e P^F da camada da Expressão Linguística e da Oração.

As Expressões Linguísticas podem ser constituídas de duas ou mais Orações, que podem relacionar-se de modo independente, numa relação de **coordenação**, como em *João comprou dois cadernos (Cl₁) e Maria comprou três lápis (Cl₂)*. A relação de **listagem** acontece quando há um número coordenado de Xps, como em *João comprou: um livro (Xp₁), três cadernos (Xp₂), quatro lápis (Xp₃) e um apontador (Xp₄)*.

A **equiordenação** estabelece a relação de mútua dependência entre duas unidades do mesmo tipo (Orações ou Sintagmas), como em *Tanto homens (Xp₁), quanto mulheres (Xp₂)*. Já a relação de **cossubordinação** ocorre quando há a combinação de duas Orações, sendo apenas uma independente, como no exemplo (2-24):

(2-24) As for the Beatles being boring, you should read Lennon's biography... (*Internet*)

(Le₁: [^{dep}(Cl₁) (Cl₂)]^{Le}) (Keizer, 2015, p. 183)

(2-24') Quanto aos Beatles serem chatos, você deveria ler a biografia de Lennon... (Tradução nossa)

Como podemos ver em (2-24), a primeira Oração (Cl₁) *As for the Beatles being boring* é a Oração dependente, e (Cl₂) *you should read Lennon's biography...* pode ser usada de modo independente.

Tendo discutido os principais aspectos do NM, seguimos, na próxima subseção, para o último nível do modelo teórico da GDF.

2.4: Nível Fonológico (NF)

O último nível é responsável também pela codificação, recebendo *input* dos outros três níveis e fornecendo *input* ao Componente Articulatorio, traduzindo essas representações em material fonológico, tais como representações acústicas, escritas, visuoespaciais ou de sinais. A representação formal das camadas é feita com duas letras maiúsculas em tamanho pequeno (*small capitals*).

Nesse nível, operam **padrões prosódicos**, **sequências segmentais**, que configuram os morfemas ou marcadores de posição, além de **Operadores terciários** que dão efeitos finais ao *input* no Componente Articulatorio. O NF também recebe material do NM, em casos em que processos fonológicos são sensíveis às características morfossintáticas de uma Expressão Linguística, como formas supletivas (e.g. verbos com conjugação irregular – *ser/era, to write/wrote*).

Esse é o nível digital em paralelo com os níveis anteriores da teoria, contendo representações em fonemas que são baseados em posições fonológicas binárias, tais como segmentos fonológicos individuais, contrastes prosódicos, etc. Vejamos sua relação hierárquica:

(π U1: [Enunciado
(π IP1: [Sintagma Entonacional
(π PP1: [Frase Fonológica
(π PW1: [Palavra Fonológica
(π F1: [Pé
(π S1) ^N	Sílaba
] (F1))	Pé
] (PW1))	Palavra Fonológica

] (PP1))	Frase Fonológica
] (IP1))	Sintagma Entonacional
] (U1))	Enunciado

Figura 8: Hierarquia no Nível Fonológico (Hengeveld; Mackenzie, 2010, p. 22)

Para Hengeveld e Mackenzie (2008), o **Enunciado** (U) é a maior camada do NF, distinguindo-se por se separar de outros Enunciados por meio de uma pausa substancial, que é utilizada pelo Falante com maior frequência para separar Enunciados do que para separar Frases Entonacionais (IP), sendo um tipo de pausa não compreendida como hesitação por parte do Ouvinte.

O **Sintagma Entonacional** (IP) é composta por um Núcleo (ponto tonal localizado em uma ou mais Sílabas, operando como marcação de queda ou elevação, importante para a interpretação do Sintagma), e sua separação ocorre por meio de pausas menores (em comparação aos Enunciados), com ritmos e durações adicionais que se refletem no Componente de Saída.

A **Frase Fonológica** (PP) pode desempenhar diferentes funções, a depender do tipo de língua, como por exemplo, nas línguas de acento, a Frase Fonológica contém uma Sílabas destacada por ser mais forte, marcando uma queda ou alçamento dentro do Sintagma Entonacional, ou em línguas tonais, a Frase Fonológica acontece no processo de *sândi*.

Já a **Palavra Fonológica** (PW) compõe uma parte da estrutura fonológica, dispondo de uma característica criteriosa relacionada ao número de segmentos, características prosódicas ou presente no domínio de regras fonológicas. Essas Palavras Fonológicas são divididas em **Sílabas** (s), estas que são agrupadas em **Pés** (F) em línguas acentuais. As Sílabas indicam, por meio de **operadores**, a presença de acento (em línguas acentuais), de tom acentual, não acentual ou de acento tonal (Pezatti, 2014).

Os **primitivos do NF** são padrões prosódicos que se aplicam a cada camada, e.g. tons específicos da língua que marcam a divisão de Sintagmas Entonacionais em Frases Fonológicas; um inventário de sequências segmentais com configurações particulares de morfemas ou *placeholders* (espaços reservados), e.g. morfemas acentuados em classes verbais específicas; e **operadores terciários**, e.g. elevação ou queda na Frase Fonológica. Vejamos em (2-25) e (2-25') um exemplo em inglês *Celtic won and Rangers lost* representadas em dois e em um Sintagma Entonacional, respectivamente:

- (2-25) Celtic won. And Rangers lost.
 NF: ((_{U_i}: [(_{IP_i}: /'seltIk'wʌn/ (_{IP_i})) (_{IP_j}: /ænd'reindʒəz' lɒst/ (_{IP_j}))] (_{U_i})) (Hengeveld; Mackenzie, 2008, p. 432)
 Celtic venceu. E o Rangers perdeu. (Tradução nossa)
- (2-25') Celtic won and Rangers lost.
 NF: ((_{U_i}: (_{IP_i}: /'seltIk'wʌnænd'reindʒəz'lɒst/ (_{IP_i})) (_{U_i})) (Hengeveld; Mackenzie, 2008, p. 432)
 Celtic venceu e o Rangers perdeu. (Tradução nossa)

Enunciados, Sintagmas Entonacionais e Frases Fonológicas podem conter operadores de queda (f) e elevação (r), que combinados podem produzir tons complexos (queda-elevação ou elevação-queda). Vejamos os exemplos (2-26) e (2-27) com a expressão em inglês *sit down*:

- (2-26) Sit down.
 NI: (M₁: (A₁: [(F₁: IMP) (P₁)_S (P₂)_A (C₁: (T₁))] A) M)
 NF: (_{U_i}: (f _{IP₁}: (pp1/ 'sit 'daʊn /) ^{IP}) ^U)
- (2-27) Sit down!
 NI: (M₁: (**emph** A₁: [(F₁: IMP) (P₁)_S (P₂)_A (C₁: (T₁))] A) M)
 NF: (_{U_i}: (f _{IP₁}: (f _{PP₁}/ 'sit 'daʊn /) ^{IP}) ^U) (Keizer, 2015, p. 264)
 Sente-se! (Tradução nossa)

Em (2-26) o operador f atua na camada do Sintagma Entonacional (IP₁), expressando um padrão entonacional para a declaração imperativa, pedindo educadamente ao interlocutor para sentar-se. Já em (2-27) a presença do operador de Ênfase atuando na camada do Ato Discursivo (A₁) no NI resulta em um novo padrão entonacional de ordem, com a adição de um novo operador (f) na camada da Frase Fonológica (PP₁).

Conforme já explicitado, não adentraremos no NF em nosso capítulo de análise linguística, pois a proposta aqui restringe-se a investigar o comportamento morfossintático das EIs e suas motivações de natureza pragmática e semântica. Seguimos agora para as definições das EIs com base nos trabalhos de Keizer (2016, 2018).

2.5: EIs para a GDF

Com base na distinção entre Léxico e Gramática da GDF, Keizer (2016) trata as EIs como unidades que possuem forma e função heterogêneas, tanto em sua composição (tipo de sintagma ou oração, número e posições abertas), como em seu comportamento morfossintático e semântico. Assim como Xatara (1998) e Monteiro-Plantin (2011), a autora também compreende as EIs como um bloco de pensamento, uma unidade convencionalizada que possui significação enquanto um conjunto, em uma leitura não literal dos seus componentes.

Para a autora, as EIs podem ser caracterizadas como elementos lexicais secundários, pois representam uma combinação de elementos que se comportam como um único bloco, atuando como elementos descritivos, mais restritos, como itens de conteúdo secundário.

Tal caráter idiomático levou a autora a destacar dois princípios funcionais, com base em Langlotz (2006) e Nunberg et al. (1994), para analisar as EIs do inglês: a **motivação** e a **composicionalidade**. A motivação diz respeito à identificação que o falante faz da expressão e seu respectivo uso em caráter conotativo. Já a composicionalidade diz respeito aos aspectos específicos do significado idiomático que podem ser relacionados às suas partes componentes, sendo a ligação significativa das partes da expressão da cena literal com a cena idiomática.

Tais expressões variam em diferentes tipos, de acordo com seu grau de motivação e de composicionalidade, resultando em **quatro tipos** classificados pela autora: **três** deles apontados como parte da literatura relativamente recente (Nunberg et al., 1994; Langlotz, 2006) e publicados no trabalho de 2016, e um quarto tipo (*AD-HOC idioms*), especificado pela autora em trabalho posterior (Keizer, 2018). Esse trabalho posterior da pesquisadora será abordado em detalhes na subseção 2.5.6, com foco nas delimitações sobre o tipo *AD-HOC* de EI e os casos de referência anafórica.

Keizer (2016) analisa as EIs do inglês como **expressões formadas por um verbo e um sintagma nominal/preposicional**, como *to kick the bucket, to grasp the nettle, to spill the beans*, entendidas como **construções semifixas** de caráter parcialmente imprevisível (lexicalmente) e parcialmente produtivo (gramaticalmente), pelo viés da GDF, delimitando as distinções entre Léxico e Gramática provenientes da teoria. Ela propõe uma classificação consistente e cognitivamente plausível sobre as variações dentro das classes de EIs, com base em critérios semânticos e sintáticos. Conforme a autora aponta, tais variações no comportamento morfossintático não são totalmente arbitrárias e sim, **funcionalmente motivadas** e, portanto, podem ser delimitadas em diferentes tipos de EI.

A vantagem de se utilizar a teoria da GDF nesse caso é que grande parte da variação de comportamento formal das expressões pode ser explicada através dos vínculos das propriedades formais com o estatuto semântico-cognitivo de cada expressão.

Com o propósito de compreender como a GDF pode ser responsável pelo armazenamento dessas expressões no Léxico e também explicar os fenômenos de variação de forma e comportamento das EIs, a autora propõe um novo primitivo chamado *Combination of Partially Instantiated Frames* (Combinação de Quadros parcialmente instanciados), pois entende que a entrada lexical de uma EI deve assumir a forma de uma definição de significado que está associada a uma combinação semifixa de quadros interpessoais e representacionais e modelos morfossintáticos que especificam as restrições funcionais e formais em expressões de modo individual, enquanto, ao mesmo tempo, explica os diferentes graus de flexibilidade sintática (*slots*) exibida pelos três principais tipos de EI.

A autora discorre sobre as dificuldades e desafios de definir as EIs em seu caráter formal, por serem expressões fixas (Barkema, 1996; Moon, 1998; e outros) e também apresentarem novas visões flexíveis das EIs, entendidas como expressões multi-palavras que, em termos de significado, parecem funcionar como um bloco unificado, mas que sintaticamente, sofrem variações que compõem novas conjunturas em termos de forma (frase, oração, número e tipos de posições abertas, etc.), além de variar em grau de convencionalidade, composicionalidade, transparência, colocabilidade e flexibilidade morfossintática.

O trabalho da autora traz uma gama de definições das gramáticas tradicionais e gerativas, das décadas de 60 e 70, nas quais os *idioms* eram tratados como expressões fixas, **não composicionais**, em que o significado das partes componentes não contribuem para a compreensão do significado idiomático, conforme visto em Katz e Postal (1963), Fraser (1970), Chomsky (1980) e outros, e ao mesmo tempo **não-produtivas**, que não permitiam passivização, topicalização ou relativização (Wood, 1986). Novas visões sobre variação em diferentes graus de idiomaticidade surgiram, principalmente com Nunberg et al. (1994), que delimitaram seis propriedades presentes nos processos de variação das EIs:

- (i) **convencionalidade** (única propriedade obrigatória das EIs, conforme os autores): o significado e o uso de uma EI não é inteiramente previsto ou compreendido com base no significado convencional dos componentes quando analisados isoladamente;
- (ii) **Inflexibilidade**: as EIs tipicamente apresentam um comportamento morfossintático restrito;

- (iii) **Figuração:** metáforas ou hipérboles;
- (iv) **Proverbialidade:** uso comum e recorrente em situações de interação e interesse social;
- (v) **Informalidade:** caráter coloquial e informal dos usos;
- (vi) **Afetividade:** implicação afetiva sobre o que é denotado, referente ao valor expressivo das EIs.

Com o entendimento dos processos de variação das EIs (Nunberg et al. (1994)), duas visões opostas tomaram forma, baseadas na relação forma-significado, e foram discutidas no trabalho de Keizer (2016). Começemos pela visão não-composicional.

2.5.1: Duas visões opostas (estrutura interna)

2.5.1.1: Visão não-composicional (baseada na forma)

Essa visão remete à impossibilidade de prever o significado da EI com base no significado literal das partes componentes, fato que contribui para sua opacidade e não-decomponibilidade: as EIs são tratadas como unidades lexicais não-motivadas que devem ser armazenadas no léxico (mental) como unidades semântica e sintaticamente únicas, com apenas uma forma convencional.

Essa visão, porém, não consegue dar conta de explicar certas expressões que sofrem variações sintáticas, como a passivização em (2-28) e (2-29), em que atenta-se para o fato de, no nível semântico, operar como uma unidade não-decomponível, mas no nível sintático as EIs poderem variar seu comportamento como um sintagma regular:

(2-28) *Matthew kicked the bucket* (forma convencionalizada)/*The bucket was kicked by Matthew**/*The bucket that Matthew kicked was empty**/*Matthew kicked the empty bucket/several buckets** (Keizer, 2016, p. 896)

(2-28') Matthew chutou o balde/O balde foi chutado por Matthew/O balde que Matthew chutou estava vazio/Matthew chutou o balde vazio/vários baldes. (Tradução nossa)

(2-28'') *Matthew died*/Matthew morreu. (Tradução nossa).

(2-29) *Edith spilled the beans* (forma convencionalizada)/*The beans were spilled by Edith*. (Keizer, 2016, p. 895)

(2-29') Edith espalhou os feijões/Os feijões foram espalhados por Edith. (Tradução nossa).

(2-29'') Edith revelou o segredo/O segredo foi revelado por Edith. (Tradução nossa)

2.5.1.2: Visão composicional (baseada no significado)

Ainda com base em Nunberg et al. (1994), a autora trata das três principais propriedades semânticas que compõem as EIs: (i) **relativa convencionalidade**, i.e. discrepância entre o significado idiomático e o significado literal de suas partes; (ii) **opacidade/transparência**, referente à facilidade com que a **motivação** para o uso pode ser recuperada pelo falante; e (iii) **composicionalidade**, que se refere ao grau em que o significado frasal/literal pode ser conhecido com base no significado das partes idiomáticas. Vejamos abaixo exemplos de uso, baseados em três classes de expressões:

(i) *to kick the bucket* (morrer)– **convencionalizada**, já que o falante a reconhece como expressão idiomática, **opaca (não-motivada)**, pois é difícil entender porque a expressão tem tal significado idiomático, e **não-composicional**, o que significa que é difícil reconhecer a contribuição individual das partes componentes da EI para o significado geral;

(ii) *to hit the ceiling* (ficar furioso(a)/nervoso(a)) - **convencionalizada, transparente**, motivada, já que é possível ver a motivação para o significado idiomático da expressão através de uma associação metafórica: ficar com raiva e **não-composicional**;

(iii) *to spill the beans* (revelar um segredo) – **convencionalizada, transparente e composicional**, pois é possível reconhecer a contribuição individual das partes componentes da EI para a sua significação idiomática- *spill* – espalhar/revelar/divulgar e *beans*/segredo/informação.

2.5.2: Uma análise da GDF sobre as EIs

Para a análise de EIs ser possível na teoria, é necessário que somente uma relação sistemática seja encontrada entre as funções semântico-cognitivas ou pragmáticas de uma EI e que seu comportamento formal seja representado na gramática.

Para Dik (1997), os principais problemas ao analisar as EIs pelo viés da GF eram: (i) como atribuir um significado unificado a uma expressão complexa, ao mesmo tempo em que se permite que essa expressão seja usada em uma variedade de tipos de construção gramatical? e (ii) como podemos atribuir funções aos elementos da EI, da mesma forma que fazemos com as expressões livres? A solução encontrada pelo autor foi analisá-las pelo quadro de predicados. No lado lexical, seriam especificadas todas as características formais e semânticas imprevisíveis (armazenadas como itens lexicais complexos na forma e parcialmente fixos associados a uma definição de significado), e no lado gramatical, os processos produtivos, conforme vemos no exemplo abaixo:

(2-30) $\text{kick}_v(x_1: \langle \text{anim} \rangle (x_1))_{\text{Ag}} (x_2)_{\text{Pat}} = \text{df hit}_v [(x_1)_{\text{Ag}} (x_2)_{\text{Pat}}]: (x_3: \{ \text{foot, leg, hoof, paw, ...} \} (x_3))_{\text{Instr}}$

(2-30') $\text{kick}_v(x_1: \langle \text{hum} \rangle (x_1))_{\text{Proc}} (\text{the bucket})_{\text{Pat}} = \text{df die}_v (x_1)_{\text{Proc}}$ - (Keizer, 2016, p. 990)

Dik (1997) entende a EI *kick the bucket* como uma única entrada no Léxico. Conforme podemos ver em (2-30'), o segundo argumento não é variável, e sim “pronto” (o balde), e o primeiro argumento se restringe a “entidades humanas”, que recebe a nova função de Processado, pois a EI designa um processo pelo qual (x_1) passa: morrer (predicado “selado” que não permite modificações).

Para Keizer (2016), os pontos fracos dessa aplicação são: (i) falta de explicação das restrições e diferenças dos comportamentos pragmático, semântico, morfossintático e fonológico das expressões; (ii) atribuição deficiente de funções semânticas nos quadros de predicados, por exemplo: *to kick the bucket*, *bucket* é visto como paciente do verbo *kick*, o que reflete na cena literal, mas não há o evento de chute no significado idiomático; (iii) a função semântica “Processado” opera como argumento de *to kick the bucket*, e não apenas de chutar (*to kick*), pois ou o verbo opera no seu significado literal, com dois argumentos - agente e paciente, ou a expressão completa *to kick the bucket* funciona como uma unidade idiomática, com apenas um argumento – *to die* (morrer).

2.5.3: Três tipos de EIs na GDF com base nas variações formais motivadas por aspectos semântico-cognitivos

Após a aplicação de oito testes formais¹⁵ nas EIs do inglês, Keizer (2016) definiu três tipos de EIs com base em aspectos semântico-cognitivos e seus respectivos comportamentos formais. Tais aspectos das EIs levaram a autora a definir tipos diferentes de EIs, pensando em seus respectivos graus de motivação e composicionalidade. Vejamos abaixo os oito testes formais, a começar pela modificação:

1. **Modificação:** adjetivo na cena literal, e.g. *to spill the **refried** beans*, ou na cena idiomática, e.g. *to spill the **legal** beans*. Há uma definição mais detalhada dos modificadores de componentes individuais que envolvem três parâmetros:

1(a) **Representacional vs Interpessoal:** pré-modificadores representacionais têm a função semântica de atribuir uma propriedade a uma entidade, e.g. *to touch the **raw** nerve* (provocar alguém), e pré-modificadores interpessoais têm função pragmática ou retórica, e.g. *to touch the **proverbial** nerve*;

1 (b) **Congruente vs incongruente:** pré-modificadores congruentes modificam o significado literal da EI, e.g. *to spill the **refried** beans*, na qual vemos o adjetivo *refried* (frito) modificar uma característica física dos feijões; já os incongruentes se aplicam ao significado não-literal das EIs, e.g. *that rocked the **local** boat* (perturbação ao ambiente ou situação local), com adjetivos que modificam a EI em seu significado metafórico/conotativo, e.g. *to spill the **gossipy** beans* (espalhar os feijões fofoqueiros);

¹⁵ Oito testes formais da autora: modificação adjetiva, substituição, passivização, nominalização, definitude, número, quantificação e anaforicidade.

1 (c) **Relacionados à EI** (significado da EI, como em *legal beans*) x **não-relacionados à EI**: relacionados a situações externas, como em *If your hemlock isn't doing that, it most likely has kicked the botanical bucket*, (se a sua cicuta não está funcionando assim, provavelmente ela chutou o balde botânico).

Abaixo são apresentadas formas de relacionar os parâmetros de modificação:

- (i) Representacional, congruente e relacionados à EI: modificação designa uma propriedade na cena literal, contribui para o significado geral da EI e é possível em EIs motivadas decomponíveis, e.g. *he had touched the raw nerve*, e não-decomponíveis, como em *negociators went around in ever-decreasing circles* (andar em círculos cada vez menores);
- (ii) Representacional, incongruente e relacionados à EI: os modificadores fornecem uma informação extra relacionada à situação real de fala, e.g. *Spain voluntarily swallows the bitter economic pill* (engolir a pílula amarga econômica);
- (iii) Representacional, incongruente e não-relacionados à EI: ocorrem em EIs não-decomponíveis motivadas, como em *Ian gritted his bigoted teeth* (Ian cerrou os seus dentes fanáticos), e não-decomponíveis não-motivadas, como em *popped his ducal clogs* (estourou os tamancos ducais), nas quais há uma manipulação na cena literal para provocar humor;
- (iv) Interpessoal, incongruente e não-relacionados à EI: modificadores de função interpessoal que operam como um meta-comentário, o que indica que o falante está ciente de que a descrição é idiomática e convencional, como em *to smoke the old peace pipe with Congress and the executive branch* (fumar o velho cachimbo da paz com o Congresso e o poder executivo, i.e. finalizar uma briga/discussão). Esse tipo de modificação é comum a todos os tipos de EIs e escopam a expressão como um todo;
- (v) Interpessoal, incongruente e não-relacionados à EI: Uso dos expressivos (*fuck, bloody, damned*) com função interpessoal de intensificar o significado sem adicionar nenhuma informação descritiva, e.g. *kick the damned bucket*.

Os três últimos tipos de modificação (expressivos, meta-comentários e os modificadores representacional, incongruente e não-relacionados à EI) não fazem parte do rol de critérios de classificação das EIs por não apresentarem evidência de que a EI é **composicional**.

2. **Substituição:** algumas EIs permitem substituição de itens lexicais por outros itens semanticamente relacionados, como sinonímia, hiponímia e metonímia. As expressões não-motivadas, como *to kick the bucket/to kick the **pail***, e as expressões motivadas decomponíveis, e.g. *to spill the beans/to spill the **peas*** não permitem substituição, pois o significado idiomático depende inteiramente da combinação fixa dos elementos.

Já as expressões motivadas não-decomponíveis permitem algum tipo de substituição, como *to hit the ceiling/to hit the **roof***, pois os falantes conseguem ainda recuperar o significado idiomático através do significado literal das partes. A autora cita que tais possibilidades de substituição dependem da interpretação do falante e o contexto inserido;

3. **Passivização:** EIs decomponíveis são os únicas permitem passivização, com em *the executives tried to show that the **hatchet was buried**/to bury the hatchet* (encerrar uma briga/confusão);

4. **Nominalização:** todos as EIs motivadas permitem nominalização (e.g. *a final **burying** of the hatchet*);

5. **Definitude/Definitividade:** todos as EIs motivadas permitem definitude/definitividade, e em alguns casos já vem marcada na forma convencionalizada da EI, como *to spill **the** beans* e o excerto “*Vinterberg is at his best when grasping **a** nettle*” (*to grasp **the** nettle*);

6. **Número:** todas as EIs motivadas permitem manipulação de Número, como em *we were grasping **nettles** – big **ones**/we would all bury **our** hatchets together*;

7. **Quantificação:** todos as EIs motivadas permitem manipulação de quantificação, como por exemplo, *touching **several** nerves/to bury **all** hatchets*;

8. **Anaforicidade:** é aplicada em EIs motivadas decomponíveis. Essa relação pode ser estabelecida com a parte componente de uma EI, e.g. *grasping nettles – big ones*, ou a expressão como um todo (ou parte dela) pode ser usado para se referir a uma entidade discursiva previamente introduzida, como em *grasping these nettles*.

Após a testagem formal das expressões, a autora define, assim, três tipos de EIs nesse primeiro trabalho: **Tipo 1 (menos produtivas):** EIs não-motivadas (opacas) e semanticamente não-decomponíveis, e.g. *to kick the bucket, to pop one's clogs*; **Tipo 2:** EIs motivadas (transparentes) e semanticamente não-decomponíveis, e.g. *to hit the ceiling, to grit one's teeth*, e **Tipo 3 (mais produtivas):** expressões idiomáticas motivadas (+ transparentes) e decomponíveis (e.g. *to spill the beans, to grasp the nettle*). O quarto tipo – EIs AD-HOC é definido no trabalho de 2018.

A decomponibilidade vista no tipo 3 (permite uma gama maior de manipulação no comportamento morfossintático) reforça que a composicionalidade semântica desempenha um papel mais importante que a transparência, conforme podemos ver no quadro 2:

	Tipo 1	Tipo 2	Tipo 3
modificação (relacionada a EI)			
- congruente	–	±	+
- incongruente, nível EI	±	+	+
- incongruente, nível NP	–	–	+
substituição	–	±	–
passivização, nominalização	–	–	+
definitude, número, quantificação	–	–	±
anaforicidade	–	–	±

Quadro 2: Comportamento morfossintático dos três tipos de EIs (Keizer, 2016, p. 1001)

Adaptado para português (Elaboração própria, 2025)

Sobre a representação na GDF, a autora entende que as características interpessoais e representacionais são refletidas no comportamento morfossintático e fonológico da teoria. No NI é feita a análise dos aspectos pragmáticos, como a referenciabilidade e identificabilidade das partes componentes da EI (diferenças entre número e tipos de Subatos).

Já o **NR** é responsável pela análise dos aspectos semânticos, que fornecem informações sobre a valência quantitativa e qualitativa do Estado-de-Coisas designado (diferenças entre número de Indivíduos e Propriedades, tempo, aspecto e modalidade).

No processo de Codificação, o **NM** abrange a análise da forma e ordem sintática dos componentes, e o **NF** analisa os aspectos fonológicos das expressões (traços segmentais e suprasegmentais). Vejamos agora a representação formal de cada um dos tipos de EIs:

1. **Tipo 1 (menos produtivas):** EIs não-motivadas (opacas) e semanticamente não-decomponíveis (*to kick the bucket, to pop one's clogs*) serão analisadas como itens lexicais complexos:

(2-31) He *kicked the bucket*.

NI: (A₁: [(F₁: DECL (F₁) (P₁)_S (P₂)_A (C₁: [(T₁)_{FOC} (+ id R₁)] (C₁))] (A₁)) (Keizer, 2016, p. 1002)

No NI, *to Kick the bucket* é representado como um Subato Atributivo (T₁) com Função Foco (maioria dos casos) e a unidade *bucket* não é um Subato Referencial separado, o que explica a não-possibilidade de definitude/definitividade, anaforicidade e nominalização.

No NR, *to kick the bucket* é representado como um único predicado verbal, com um único argumento Undergoer (correspondendo ao Sujeito paciente), e *bucket* não é visto como uma unidade representacional separada (não funciona como argumento de *to kick*), o que não permite número, quantificação, ou passivização. A representação se dá como segue:

NR: (p₁: (past ep₁: (e₁: (f₁: [(f₂: **kick_the_bucket**_v (f₂)) (1x₁)_U] (f₁)) (e₁)) (ep₁)) (p₁)) (Keizer, 2016, p. 1002)

Já no NM, *to kick the bucket* é representado em três palavras: *kick* como verbo encabeçando um Sintagma Verbal (Vp₁) flexionado no tempo passado, *the* uma Palavra Gramatical (Gw₁) e *bucket* como uma Palavra Nominal (Nw₂). Segue a representação:

NM: (Cl₁: [(Np₁: (Nw₁: he (Nw₁))_{Subj} (Np₁)) (Vp₁: (Vw₁: **kick-past** (Vw₁)) (Vp₁)) (Np₂: [(Gw₁: **the** (Gw₁)) (Nw₂: **bucket** (Nw₂))] (Np₂))] (Cl₁)) (Keizer, 2016, p. 1003)

No que se refere ao NF, *to kick the bucket* se comporta de modo previsível e desde que não haja função pragmática, não há destaque prosódico, mantendo o padrão prosódico de uma Declarativa. Segue a representação:

NF: (U₁: (fIP₁: (PP₁: /hi/ (PP₁)) (PP₂: /kɪktðə'bʌkɪt/ (PP₂)) (IP₁)) (U₁)) (Keizer, 2016, p. 1003)

2. **Tipo 2:** EIs motivadas (transparentes) e semanticamente não-decomponíveis (*to hit the ceiling*, *to grit one's teeth*) serão analisadas como dois itens lexicais designando uma Propriedade:

(2-32) He *hit the ceiling*.

NI: (A₁: [(F₁: DECL (F₁) (P₁)_S (P₂)_A (C₁: [(T₁)_{FOC} (+ id R₁)] (C₁))] (Keizer, 2016, p. 1004)

No Nível Interpessoal, *to hit the ceiling* é representado como o tipo 1: Subato Atributivo (T₁) com Função Foco e a unidade *ceiling* não é um Subato Referencial separado.

No NR, *to hit the ceiling*, *ceiling* é uma unidade separada (x₁) funcionando como argumento Undergoer do predicado verbal (f₃) *hit*. Já *hit the ceiling* funciona como um predicado complexo de um lugar (f₂) com um argumento Undergoer (x₂). Segue a representação:

NR: (p₁: (_{past} ep₁: (e₁: (f₁: [(f₂: [(f₃: hit_V (f₃)) (1x₁: ceiling_N (x₁))_U] (f₂)) (1x₂)_U] (f₁)) (e₁)) (ep₁))(p₁)) (Keizer, 2016, p. 1004)

Já no que concerne ao NM, o argumento (x₁) *ceiling* pode ser modificado por adjetivos que contribuam para o significado da EI como um todo, e a posição do operador de (x₁) “é selada”, pois o operador de número 1 é um recurso invariável e não serve para nenhuma função comunicativa, o que pode ser explicado por características idiossincráticas e não arbitrarias das EIs, como no caso, uma sala/quarto tem apenas um teto (denotação), o que não permitiria quantificadores.

NM: (Cl₁: [(Np₁: (Nw₁: he (Nw₁)) (Np₁)) (Vp₁: (Vw₁: **hit-past** (Vw₁)) (Vp₁)) (Np₂: [(Gw₁: **the** (Gw₁)) (Nw₂: **ceiling** (Nw₂))] (Np₂))] (Cl₁)) (Keizer, 2016, p. 1004)

To hit the ceiling é representado em três Palavras no NM: *hit* como verbo (Vw₁) encabeçando um Sintagma Verbal (Vp₁) flexionado no tempo passado (past), *the* uma Palavra Gramatical (Gw₁) e *ceiling* como uma Palavra Nominal (Nw₂).

No NF, *to hit the ceiling* se comporta de modo previsível e desde que não haja função pragmática, não há destaque prosódico, mantendo o padrão prosódico de uma Declarativa:

NF: (U₁: (fIP₁: (PP₁: /hɪ/ (PP₁)) (PP₂: /hɪtðə'si:lɪŋ/ (PP₂)) (IP₁)) (U₁)) (Keizer, 2016, p. 1004)

3. **Tipo 3 (mais produtivas):** EIs motivadas (+ transparentes) e decomponíveis (*to spill the beans*, *to grasp the nettle*) serão analisadas como (quase) completamente composicionais.

(2-33) *He spilled the beans.*

NI: (A₁: [(F₁: DECL (F₁) (P₁)_S (P₂)_A (C₁: [{(T₁) (+idR₁)}_{FOC} (+ idR₂)] (C₁))] (Keizer, 2016, 1004)

No NI, *to spill the beans* é representado por um Subato Atributivo (T₁) e um Subato Referencial (R₁) que juntos atribuem a função pragmática de Foco, evitando a clivagem da unidade *beans*. Embora seja uma EI motivada e convencional, ainda é difícil derivar o significado idiomático resultante das partes componentes de modo isolado, não permitindo substituição.

No NR, *to spill the beans*, *beans* é argumento Undergoer do verbo *spill*, e é representado como um Conteúdo Proposicional (p₁), que reflete a natureza sua metafórica (*beans* – segredo – uma entidade de terceira ordem).

NR: (p₁: (past ep₁: (e₁: (f₁: [(f₂: **spill** (f₂)) (m p₁: **bean**_N (p₁))_U (x₂)_A] (f₁)) (e₁)) (ep₁)) (p₁)) (Keizer, 2016, p. 1004)

To spill the beans é representado no NM por três Palavras: *spill* como Palavra Verbal (V_{w1}) encabeçando um Sintagma Verbal (V_{p1}) flexionado no tempo passado, *the* uma Palavra Gramatical (G_{w1}) e *beans* como uma Palavra Nominal (N_{w2}).

NM: (Cl_1 : [(N_{p1} : (N_{w1} : he (N_{w1})) (N_{p1})) (V_{p1} : (V_{w1} : spill-past (V_{w1})) (V_{p1})) (N_{p2} : [(G_{w1} : the (G_{w1})) (N_{w2} : bean<pl> (N_{w2}))] (N_{p2}))] (Cl_1)) (Keizer, 2016, p. 1004)

To spill the beans se comporta no NF de modo previsível e, desde que não haja função pragmática, não há destaque prosódico, mantendo o padrão prosódico de uma Declarativa.

NF: (U_1 : (fIP_1 : (PP_1 : /hɪ/ (PP_1)) (PP_2 : /spɪldðə'bi:nz/ (PP_2)) (IP_1)) (U_1)) (Keizer, 2016, p. 1004)

Antes de seguirmos com um maior detalhamento sobre os conceitos de transparência e opacidade para GDF, trouxemos uma introdução sobre transparência linguística na próxima subseção.

2.5.4: Transparência linguística

O trabalho de Pezatti et al. (2017)¹⁶ sobre os conceitos de transparência linguística aborda desde a noção de arbitrariedade do signo linguístico nos estudos de Saussure até estudos mais atuais sobre os princípios de iconicidade e isomorfismo, o que significa que algo na estrutura da língua (forma) reflete a estrutura da realidade, de eventos, pensamentos, estados, etc. (significado).

A **iconicidade** na língua é vista por Givón (1985) como a correspondência entre forma e significado presente no sistema de cada língua. Essa concepção não é vista como um universal absoluto nas línguas, pois representa preferências do falante que afetam a determinação da expressão linguística em relação ao significado.

Leufkens (2015) aprofunda essa noção de iconicidade em duas leituras: a primeira refere-se à previsibilidade do significado pela forma; e a segunda, refere-se à iconicidade de sequência, na qual uma expressão linguística reflete a mesma ordem de suas categorias morfossintáticas e semânticas.

¹⁶ Os trabalhos de Pezatti et al. (2017) e de Leufkens (2015) dialogam entre si sobre transparência linguística. Aquele trouxemos como uma introdução ao tema, e este, como aporte principal sobre descontinuidade.

Para a autora, a transparência linguística está relacionada à **relação um-para-um entre forma e significado**, envolvendo as noções de homomorfismo e isomorfismo. Esta é a relação bidirecional entre dois conjuntos, resultando em biunivocidade, e aquela, a relação idêntica de unidades de dois conjuntos. Pezatti et al. (2017) discorrem que há uma relação de isomorfismo quando há uma correspondência um-para-um entre forma e significado de unidades de diferentes níveis, e.g., níveis semântico e morfossintático. Croft (2003) discorre sobre a diferença do isomorfismo paradigmático, “que são projeções de significado e forma no léxico”, e do isomorfismo sintagmático, “que são projeções de significado e forma na gramática” (Croft, 2003 apud Pezatti et al., 2017, p. 8).

O conceito de **transparência na GDF** passa por sua estruturação modular em níveis e camadas e sua organização descendente entre os processos de Formulação e Codificação, percurso que se inicia na intenção do falante e termina na Articulação. Tal conceito de transparência está apoiado na relação um-para-um entre forma e significado, i.e., Leufkens (2015) entende o termo “unidade de significado” como um primitivo fornecido pela teoria - operador, função ou camada do NI e NR, que são unidades de significado; e a “unidade de forma”, que é um primitivo do NM e do NF, e são unidades formais, tais como oração, sintagma, palavras, operadores e formas supletivas.

Com isso, a noção de transparência na GDF se dá no emparelhamento entre as interfaces de Formulação (NI e NR) e Codificação (NM e NF), i.e., ocorre “transparência quando uma unidade de um nível da organização linguística corresponde a uma unidade de qualquer outro nível da organização” (Leufkens, 2015 apud Pezatti et al., 2017, p. 15), sendo assim, uma relação de forma e significado entre interfaces, resultando em seis possíveis combinações de interfaces entre os quatro níveis: NI-NR, NI-NM, NI-NF, NR-NM, NR-NF, NM-NF.

A opacidade na GDF se dá então nos processos de violação a esse pareamento entre as interfaces, resultando em divergências nas unidades contidas em cada nível. Leufkens (2015) entende que transparência é uma noção escalar, e não binária, pois há graus de transparência nas línguas, algo possível de mensurar devido aos traços opacos manipulados pela ausência de determinados fenômenos na língua.

Para a autora, há cinco possíveis tipos de ruptura da transparência da relação um-para-um entre as unidades dos níveis da teoria: (i) relação um-para-nulo, que são categorias e argumentos vazios não visíveis ou audíveis em sua produção e com existência teoricamente motivada, como o morfema-zero; (ii) relação nulo-para-um, que refere-se a uma forma morfossintática ou fonológica não motivada ou relacionada aos níveis mais altos, como o sujeito vazio; (iii) relação um-para-muitos, em que há uma unidade de significado e mais de uma forma, como casos de redundância; (iv) relação muitos-para-um, que consiste da existência de múltiplos significados e apenas uma unidade de forma,

como por exemplo, os morfemas *portmanteau* – uma unidade morfológica que resulta da fusão de distintos morfemas (*motel, brunch, workaholic, etc.*); e (iv) violação de integridade de domínio, em que unidades de significado do NI ou NR ficam justapostas na representação no NM, resultando em **descontinuidade**:

A subcategoria de descontinuidade inclui violações da integridade do domínio, que resultam em unidades morfossintáticas ou fonológicas incompletas ou fragmentos que não podem ser usados de forma independente, mas requerem a presença de outra unidade. Isto dá origem a relações entre uma única unidade de significado e uma unidade morfossintática que é dividida em múltiplas partes, envolvendo uma relação não transparente entre uma unidade de significado e duas ou mais unidades morfossintaticamente incompletas. (Leufkens, 2015, p. 63. Tradução nossa)¹⁷

Esse quinto e último tipo de ruptura da transparência nas línguas é de fundamental importância em nossa fundamentação metodológica, pois será parte de um dos critérios de análise linguística: violação da transparência linguística na relação forma e significado, na qual vemos casos de descontinuidade nas situações de referência anafórica e catafórica com uma unidade da cena idiomática, nas quais as unidades de forma e significado da cena idiomática ocorrem de modo descontínuo, em um alinhamento não-paralelo das unidades, gerando casos de anáfora e catáfora invocadas por parte de uma unidade da EI.

Leufkens (2015) reforça o comportamento das EIs enquanto casos de **Lexemas compostos por mais de uma Palavra morfossintática**, ponto que nos auxiliará a distinguir os casos de descontinuidade.

Após a apresentação dos conceitos de transparência e opacidade para GDF, parte-se para sua aplicação, tendo como referência os tipos de EIs definidos por Keizer (2016).

2.5.5: Transparência e opacidade das EIs

Se há relação um-para-um entre unidades de dois níveis, isso contribui para a transparência da EI. As EIs do **Tipo 3** são mais transparentes entre NI e NR e NM, pois os Subatos Atributivo (T) e Referencial (R) correspondem a unidades únicas nos Níveis Representacional – um Estado-de-

¹⁷ Texto original: The subcategory of discontinuity includes violations of domain integrity, which result in incomplete morphosyntactic or phonological units or fragments that cannot be used independently but require the presence of another unit. This gives rise to relations between a single meaning unit and a morphosyntactic unit that is split-up into multiple parts, involving a non-transparent relation between one meaning unit and two or more morphosyntactically incomplete units. (Leufkens, 2015, p. 63)

Coisas (e) e um Conteúdo Proposicional (p) e no Nível Morfossintático (uma Palavra Verbal (Vw) e uma Palavra Nominal (Nw)).

As EIs do **Tipo 2** são menos transparentes entre NI e NR, pois um Subato Atributivo (T) corresponde a duas unidades no Nível Representacional (um Estado-de-Coisas (e) e um Indivíduo (x)), além de serem menos transparentes entre NI e NM. Já a relação entre NR e NM é transparente (uma Palavra Verbal (Vw) e uma Palavra Nominal (Nw)).

As EIs do **Tipo 1** são mais opacas (não-transparentes); embora haja transparência entre NI, com um Subato Atributivo (T), e NR, com um Estado-de-Coisas (e), suas representações no NM divergem em uma Palavra Verbal (Vw) e uma Palavra Nominal (Nw).

Há dois grupos de EIs que aparentemente não se comportam conforme a análise tipológica proposta para os três tipos de EIs: EIs com uma variável em forma de pronome possessivo, ligado ao sujeito; e as EIs que apresentam características de mais de um tipo. O primeiro grupo são EIs com uma variável em forma de pronome possessivo ligado ao sujeito: quando motivadas e decomponíveis, e.g. *to lose one's mind* (perder a cabeça/controlar), ou não decomponíveis, como *to grit one's teeth* (mostrar coragem/determinação). Essa característica pode ser resolvida no Nível Representacional (Tipo 3 – uma Propriedade Configuracional e um Conteúdo Proposicional e Tipo 2 – duas Propriedades Configuracionais), mas quando são EIs não-motivadas e não-decomponíveis, como *to pop one's clogs* (morrer), a expressão é tratada como uma única Propriedade no NR.

A solução seria tratar esse possessivo sintaticamente como predicados reflexivos, como *to enjoy oneself* (divertir-se), assim o pronome possessivo não faria parte do lexema complexo (*to pop clogs oneself*), mas seria representado no NM (2-34), concordando em número, pessoa e gênero com o sujeito:

(2-34) **To pop clogs oneself** (Tipo 1 não-motivada e não-decomponível)

NR: (p_i: (inf ep_i: (e_i: (f_i: [(f_j: **pop_clog_m** v (f_j)) (1x_i: **oneself** x_i)_U] (f_i)) (e_i)) (ep_i)) (p_i))

NM: (Cl_i: [(Np_i: (Nw_i: oneself (Nw_i))_{Subj} (Np_i)) (Vp_i: (Vw_i: **pop-inf** (Vw_i)) (Vp_i)) (Np_j: [(Nw_j: **clog<pl>**(Nw_j))] (Np_j))] (Cl_i))

Já o segundo grupo refere-se às EIs que apresentam características de mais de um tipo, sendo ambas as características motivadas pertencentes aos tipos 2 e 3. Por exemplo, a EI *to bite the dust* (morrer/ser derrotado/falhar) apresenta traços do tipo motivada e decomponível (aceita modificadores

incongruentes e nominalização – *to bite the financial dust/the digital dust/the biting of the dust for the U.S.A*) e ao mesmo tempo, parece não aceitar passivização (possível em EIs motivadas decomponíveis – *the beans were spilled*).

Essa ambiguidade entre elementos decomponíveis e não-decomponíveis pode ser explicada pelos traços semântico-cognitivos da EI, que levam os falantes a relacionarem a cena literal (engolir/morder poeira após cair no chão ao ser derrotado numa luta/ou morto) e seu significado idiomático (ser derrotado/estar morto), o que demonstra a familiaridade do falante com a EI e sua motivação ao reconhecer seu significado idiomático, mesmo com tais manipulações no comportamento morfossintático. A autora deixa esses casos atípicos de lado na classificação final, como *to bite the digital/financial dust, popped his footballing clogs*, ao entender sua incoerência com a classificação proposta.

2.5.6: Como representar EIs no Léxico pensando na distinção Léxico x Gramática

Os Lexemas são armazenados em quadros separados, e durante a Formulação se combinam em quadros Interpessoais ou Representacionais. A melhor maneira de representar as EIs no Léxico é utilizar os quatro níveis da teoria da GDF, com suas respectivas representações de quadros no NI e NR e os modelos do NM com os elementos invariáveis (parcialmente selados/fixos). Esses traços fixos devem ser armazenados no Léxico, enquanto o restante do quadro e as variações consistem em espaços abertos, que serão preenchidos com regras normais da Gramática.

Por exemplo, *to kick the bucket, to spill the beans, to hit the ceiling* serão as entradas lexicais parcialmente seladas, e funcionarão como uma combinação de quadros parcialmente instanciados (ComPIFs), representando os traços imprevisíveis e invariáveis das EIs, como uma entrada lexical ou como parte do Componente Conceitual. Os Lexemas fixos nas EIs também estão armazenados separadamente no Léxico, assim, o tipo 1 será listado como um lexema simples (com várias palavras, como em *to kick the bucket = morrer*).

(2-35) Lexicon:

Meaning: ‘die’

NI: $(\pi T_1)_\phi$

NR: $(f_1: \text{kick_the_bucket}_v(f_1))$

NM: $(Vp_1: (Vw_1: \text{kick} (Vw_1)) (Vp1)) (Np_2: [(Gw_1: \text{the} (Gw_1)) (Nw_2: \text{bucket} (Nw_2))] (Np_2))$
(Keizer, 2016, p. 1009)

Essa representação, conforme postulado pela autora, permite considerar distribuição e flexibilidade sintática, assim como composicionalidade, e ao considerar os Lexemas separados armazenados no Léxico, é possível depreender a leitura literal das partes componentes, o que permite avaliar a sua motivação (transparência e opacidade). Para tal possibilidade, basta acrescentar uma nova entrada lexical, composta por primitivos existentes: **ComPIF**. Essa nova entrada também auxiliaria na análise de outras construções semifixas, tais como *the thing is*, *Would you believe that?*, entre outras.

O trabalho de Keizer (2016) justificou a necessidade de se distinguir três tipos de EI com base em aspectos formais (comportamento morfossintático) e funcionais (semântico-cognitivos em termos de motivação e composicionalidade). Ao utilizar a GDF como ferramenta de análise, a autora explorou ao máximo os recursos da teoria em termos dos quatro níveis de análise e como resultados, argumentou que a proposta conseguiu representar a variação formal dessas expressões, a partir dos oito testes formais, e também sua relação às propriedades funcionais de motivação e composicionalidade, mesmo com a existência de casos atípicos à classificação proposta.

Ao final, a autora propõe a entrada lexical (**CombPIF – Combinação de quadros parcialmente instanciados**) para as EIs, associada à **combinação semifixa** de quadros interpessoais e representacionais e modelos morfossintáticos que especificam as restrições formais e funcionais das EIs, além dos diferentes graus de flexibilidade sintática, conforme explicitado anteriormente.

Já em seu trabalho de 2018, a autora discorre sobre novas **EIs AD-HOC**. Essas expressões são derivadas de EIs consagradas e sofrem alguma variação lexical em sua forma para ser utilizada em um contexto específico, com uma determinada finalidade que só é válida para tal contexto. Essa nova EI **AD-HOC** não tem funcionalidade em outros contextos e não se torna uma nova EI consagrada. Tratamos as EIs **AD-HOC** como quarto tipo, em consenso com a autora, alinhado durante estágio no exterior em 2024.

Para Keizer (2018), o falante reconhece a EI **AD-HOC** e seu grau mínimo de variação lexical em relação à forma consagrada, assim como consegue reconhecer essa variação representada como

um pronome anafórico, sem ambiguidade nem dependência de contexto. Não há ambiguidade, pois o falante, ao associar as formas consagrada e variada, entende que não há a criação de uma nova EI consagrada, mas que tal variação lexical é específica e tem determinada finalidade. Também não há dependência de contexto, pois o falante é capaz de apreender os significados literal e idiomático das duas formas - consagrada e *AD-HOC*, relacionando as partes componentes.

Nesse caso, há um **esforço extra-cognitivo** por parte do leitor para reanalisar a EI em um **processo de não-decomponibilidade para decomponibilidade**. Essa inferência ocorre pela mudança de foco na significação literal e idiomática: há uma substituição lexical, resultando em uma nova significação literal e conseqüentemente, uma nova significação idiomática. Como exemplo, a autora trouxe a expressão motivada não-decomponível *to bite the bullet* (*fazer um sacrifício ou passar por uma dificuldade*) no seguinte contexto de fala:

(2-36) The DLC did *bite a bullet* on trade, or at least ***gnawed it***. (Keizer, 2018, p. 187)

O DLC *mordeu uma bala* no comércio, ou pelo menos a *roeu*. (Tradução nossa)

Podemos ver que a alteração do verbo *bite* (morder) para *gnaw* (roer), que altera o significado da cena literal, trouxe também alteração significativa para a cena idiomática: houve uma suavização e cautela da ação de fazer um sacrifício, sendo uma situação desagradável, realizada pelo DLC, bem como se identifica um grau de abrandamento de intensidade entre morder e roer a bala. Em outra expressão motivada não-decomponível, temos o exemplo *to open the floodgates* (conceder exceções ou facilitar processos):

(2-37) Once *the floodgates open*, ***they*** cannot be closed. (Keizer, 2018, p. 187)

Uma vez que *as comportas se abrem*, ***elas*** não podem ser fechadas. (Tradução nossa)

Notamos que, nesse caso, há uma substituição de um elemento da EI (*floodgates*) pelo pronome ***they***, que passa também por um processo de referência anafórica, passando a referenciar com o item *floodgates* (comportas), fato que também se reflete na cena idiomática: pronominalização de exceções (elas - comportas) e alteração do verbo referente à cena literal – *abrir e fechar* as comportas e sua significação idiomática – conceder ou não exceções/facilidades.

As variações que ocorrem nos casos analisados em expressões não-decomponíveis nem sempre necessita ser altamente não convencional e não-sistemática. As funções discursivo-

pragmáticas - dispositivos coesivos e fortalecedores, e sua eficiência, i.e. confiança na inferência em vez da expressão explícita dos significados, superam o custo do esforço extra-cognitivo no processo de reanálise para compreensão das expressões.

Com base nos apontamentos de Keizer (2016, 2018), Bakker e Hengeveld (2012) e também nos estudos de Lakoff (1986, 1995), Lakoff e Johnson (1980) e Langlotz (2006), o quadro 3 apresenta a representação escalar dos quatro tipos de EIs vistas nos trabalhos de Keizer (2016, 2018):

Tipos de EIs	Motivação			Composicionalidade			Ativação de cena literal		Exemplos do inglês
	Global	Constitucional	Desmotivados	Literal	Parcial	Não-literal	Interpretação regular	Interpretação incomum	
Tipo 1- não motivadas e não-decomponíveis			x	x		x	x	x	<i>to kick the bucket, to pop one's clogs, to shoot the breeze</i>
Tipo 2- motivadas e não-decomponíveis	x			x			x		<i>to hit the ceiling, to go around in circles, to grit one's teeth</i>
Tipo 3- motivadas e decomponíveis	x	x		x			x		<i>to spill the beans, to grasp the nettle, to swallow a bitter pill</i>
Tipo 4- AD-HOC motivadas e não-decomponíveis	x			x			x	x	<i>to bite the bullet, to open the floodgates</i>

Quadro 3: Ativação semântica das EIs (Elaboração própria, 2025)

No capítulo de análise, apresentaremos um novo quadro (quadro 7) com EIs do português e com a inclusão do quinto tipo de EI. Nesse primeiro momento, o quadro 3 reúne propriedades

semântico-cognitivas encontradas nos autores citados anteriormente, com foco na proposta de Keizer (2016, 2018).

Observa-se que as EIs do tipo 1 podem apresentar composicionalidade literal e interpretação regular da cena literal (e.g. *to kick the bucket/morrer*) e casos de composicionalidade não-literal e interpretação incomum da cena literal (e.g. *to shoot the breeze/ter uma conversa casual*). O tipo 2 apresenta maior regularidade em termos de comportamento semântico-cognitivo, já que mostra motivação global, composicionalidade literal, o que facilita sua ativação semântica literal.

O tipo 3 apresenta casos com motivação global e constitucional, como *to grasp the nettle* (resolver um problema) e *to spill the beans* (revelar um segredo), respectivamente. A motivação em *to grasp the nettle* é global, pois sua significação idiomática é relacionada à expressão como um todo (agarrar como enfrentamento e urtiga como problema). Já em *to spill the beans* o termo de maior transparência para contribuição no significado idiomático é o verbo *spill* (divulgar), sendo um constituinte autônomo com significado lexicalizado figurativo. O tipo 4 possui casos com ativação da cena literal tanto regular (e.g. *to open the floodgates/abrir as comportas*) e incomum (*to bite the bullet/morder a bala*).

A figura 9, apresenta a representação gradual dos quatro tipos de EIs, com base na transparência e opacidade, para delimitar os tipos mais prototípicos:

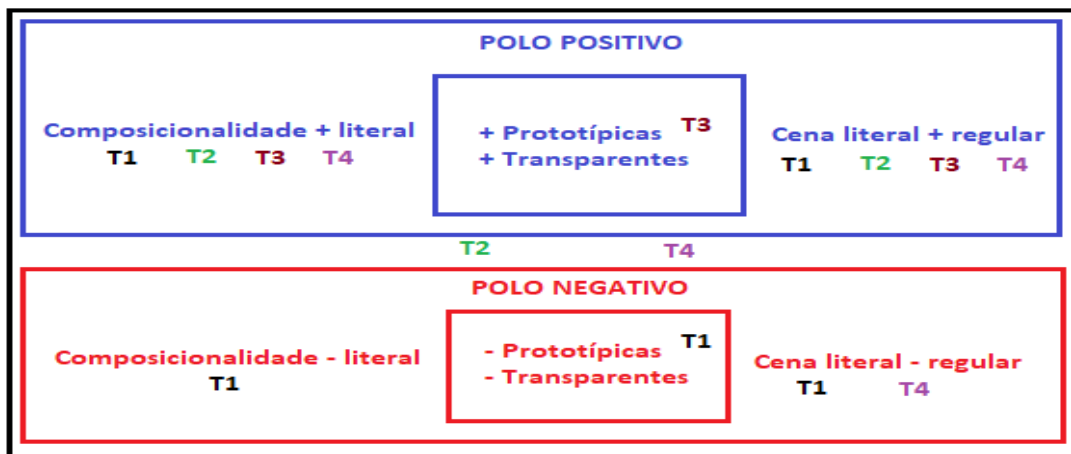


Figura 9: Graduação dos tipos de EIs (Elaboração própria, 2025)

Conforme dados do quadro 3 e da figura 9, todos os tipos de EIs podem apresentar casos de composicionalidade literal e interpretação regular da cena literal (e.g. *to spill the beans*). O tipo 3 é

prototipicamente mais transparente e motivado (e.g. *to grasp the nettle*), e o tipo 1, mais opaco e desmotivado (e.g. *to pop one's clogs/morrer*). Os tipos 2 e 4 ficaram numa posição intermediária de transparência, pois são definidos como motivados, mas não-decomponíveis (e.g. *to hit the ceiling/ficar furioso*). Alguns casos do tipo 4 e tipo 1 apresentam interpretação incomum da cena literal (e.g. *to bite the bullet*), e o tipo 1 apresenta dados de composicionalidade não-literal (e.g. *to shoot the breeze*).

Com a inclusão dos parâmetros de ativação semântica, vistos em Langlotz (2006), notamos a expansão de possibilidades de encaixe do mesmo tipo de EI em diferentes graus de análise entre os pólos positivo e negativo. Assim como vimos na proposta de Keizer (2016) e durante comunicação pessoal nas aulas de *English idioms*¹⁸, no primeiro semestre de 2024, na Universidade de Viena, o comportamento morfossintático das EIs pode variar em diferentes graus, em grande influência pela criatividade do falante em adaptar diferentes fenômenos da língua em diversos contextos de comunicação, assim como o trato do falante em familiarizar-se com tais expressões no momento de fala no processo de analisabilidade.

Por exemplo, a expressão *to bite de bullet* foi aqui entendida como incomum em sua interpretação literal, o que de fato não anula a condição de sua realização física ser possível. O mesmo não ocorre com *to rain cats and dogs* (chover muito). Para Langlotz (2006), a **criatividade linguística** é fruto do desejo comunicativo do falante em adaptar as unidades linguísticas por meio de expressões que estão enraizadas na Gramática Cognitiva, para codificar complexos e conceituações em constante mudança de uma forma contextualmente apropriada. Seu reflexo criativo na codificação idiomática é baseado na computação mental; que envolve **manipulação mental do conjunto de ativação idiomática** e sua integração com outros esquemas e construtos pré-estabelecidos em unidades simbólicas.

O autor entende que as EIs são frequentemente variadas no discurso e podem apresentar variantes institucionalizadas, pois seus dados de variação fornecem “fortes evidências empíricas que falam contra uma concepção uniforme de EIs como sendo semelhante a palavras longas” (Langlotz, 2006, p. 175, Tradução nossa¹⁹). Tais variações podem ser: (i) institucionalizadas (e.g. *to hit the ceiling/roof*); (ii) ocasional (e.g. *to spill the juicy beans*); (iii) gramaticalmente sistemática (e.g. *she didn't spill a single bean*); (iv) variante de trocadilho (e.g. *to spill the pasta* (máfia)); (v) variante

¹⁸ Realização do doutorado sanduíche com bolsa de fomento à pesquisa CAPES PRINT em 2024.

¹⁹ Texto original: “Idiom-variation data provide strong empirical evidence speaking against a uniform conception of idioms as being similar to long words. (Langlotz, 2006, p. 175)

errônea não intencional (e.g. *to blow one's lid* – *to blow one's top* + *to flip one's lid*) e (vi) pseudo-variante (e.g. *to spill the peas*).

O falante manipula as EIs, levando em conta sua complexa relação entre forma e significado, através de sua inserção em dado ambiente cognitivo específico, baseado em um tópico discursivo, um contexto que o influencia a selecionar determinada expressão que se encaixe em suas intenções comunicativas, levando em conta a interpretação lógica da relação dos constituintes e suas significações literal e idiomática.

Dentre os vários fenômenos de variação e manipulação das EIs, estão ainda os casos de referência anafórica. Seguimos com uma breve introdução de referências anafórica e catafórica para a GDF e na sequência, retomamos com o texto de Keizer (2018) sobre esse fenômeno.

2.5.7: Referências anafórica e catafórica para a GDF

Na GDF, o conceito de referência anafórica é usado para indicar uma relação de identidade entre duas unidades no mesmo nível, por exemplo, no Nível Representacional, a coindexação entre dois Indivíduos (Minha *blusa* é nova. Eu *a* comprei ontem) ou Propriedades (um *carro* azul e *um* vermelho). Assim, um pronome anafórico é usado para se referir a uma entidade anteriormente introduzida no discurso.

Os pronomes catafóricos diferem no processamento dos termos anafóricos em que a coindexação funciona dentro de seu domínio de operação. O domínio de operação da catáfora é geralmente o Ato do Discurso, como no exemplo em inglês visto em Hengeveld e Mackenzie (2008): *After he* (1x_i) *took a shower, Brian* (1x_i) *went to the movies*, onde a relação catafórica se estabelece entre descrições de dois Estados-de-Coisas em um único Ato Discursivo. No Nível Representacional, a relação entre Indivíduos (1x_i) se estabelece através de **coindexação**.

Sobre o fenômeno de referência anafórica, Keizer (2018) detalha três padrões para esse tipo de teste formal: o primeiro tipo é a **referência anafórica por parte da EI não-decomponível**, na qual há uma adaptação mínima da cena literal. A forma idiomática é preservada e reconhecida. Não há dependência do contexto, nem ambiguidade e nem marcação estilística, conforme (2-38) com EI motivada e (2-39), como exemplo de uma expressão desmotivada:

- (2-38) We are accused of centralising FE. If we are going *to be tarred with that brush* then let's do something with it. (Keizer, 2018, p. 184)

Somos acusados de centralizar a EF. Se vamos *ser manchados com esse pincel*, então vamos fazer algo com ele.) esse pincel – centralizar a EF. (Tradução nossa)

- (2-39) The win just might be the impetus the Jays need as they improved to 8-17 and just *avoided the worst March/April start in club history* (7- 15 in 1979 and 2004 takes *that cake*) (Keizer, 2018, p. 184)

A vitória pode ser o impulso que os Jays precisam, pois melhoraram para 8-17 e *evitaram o pior início de março/abril da história do clube* (7-15 em 1979 e 2004 leva *esse bolo*). Esse bolo – piores inícios de temporada em 1979 e 2004. (Tradução nossa)

O segundo tipo de referência anafórica é denominado **Referência anafórica a uma entidade do discurso previamente introduzida**, na qual o antecedente é a EI e a referência é acionada por parte da EI em um processo de inferência: o significado idiomático é relevante na compreensão da mensagem principal e o referente idiomático é previamente introduzido, conforme *this bacon* em (2-40) e o pronome *it* em (2-41):

- (2-40) Everybody wants *to bring home the bacon*. The difference is *this bacon* can blow up the world. (Keizer, 2018, p. 185)

Todo mundo quer *levar o bacon para casa*. A diferença é *que este bacon* pode explodir o mundo. Este bacon – arsenal nuclear. (Tradução nossa)

- (2-41) And, perhaps best of all, it's going to be a family tradition. It's a family with quite a few of them, but this one *takes the cake* – and then has *it* donated. (Keizer, 2018, p. 185)

E, talvez o melhor de tudo, será uma tradição familiar. É uma família com alguns deles, mas este *leva o bolo* – e depois o tem *doado*. *Esse bolo (doado) – generosidade da família e sua tradição familiar excessiva*. (Tradução nossa)

O terceiro tipo é denominado **Referência anafórica a uma entidade da cena literal**, na qual o significado literal da cena literal da EI é mais evidenciado e ativo. Como exemplo, temos em (2-42) e (2-43) as cenas literais das EIs motivadas *hanging on by a thread* e desmotivadas *to take the cake* em evidência:

- (2-42) The child was *hanging on by a thread*, and ***that thread*** was beginning to unravel. (Keizer, 2018, p. 186)

A criança estava *pendurada por um fio*, e ***esse fio*** estava começando a se desfazer. Esse fio – situação de risco. (Tradução nossa)

- (2-43) (from an article on alcohol consumption in the US in the 1790's) But it was the men in Independence Hall during the summer of 1787 who really *took the cake*, especially if ***it*** was drenched in bourbon. (Keizer, 2018, p. 186)

(de um artigo sobre o consumo de álcool nos EUA na década de 1790) Mas foram os homens do Independence Hall durante o verão de 1787 que realmente *levaram o bolo*, especialmente se ***ele*** estivesse encharcado de bourbon. Se ele (este bolo – pior aparição referente ao tópico do alcoolismo. Tradução nossa)

Finaliza-se aqui nosso aporte teórico, que perpassou pelo detalhamento do nosso objeto de estudo – as EIs, também pelos principais conceitos da GDF e pelos trabalhos norteadores de Keizer (2016, 2018), e de outros autores que nos motivaram a aplicar uma proposta de análise funcional semelhante à de Keizer (2016, 2018) em EIs de diferentes variedades lusófonas do português.

Seguimos agora para o capítulo de fundamentação teórico-metodológica, em que explicaremos como se deu a seleção de EIs a partir de um banco de dados e como se deu o levantamento de ocorrências nos *corpora* selecionados, além de explicitar as etapas e os critérios norteadores da análise.

CAPÍTULO 03: FUNDAMENTOS TEÓRICO-METODOLÓGICOS

3.1: Metodologia de pesquisa

Esta é uma pesquisa de cunho qualitativo e quantitativo (investigação *type* – tipo de ocorrência e *token* – número de ocorrências), com foco na análise das EIs encontradas no CRPC, em excertos da *web*, e em dados construídos, conforme necessário²⁰. Como visto nos capítulos anteriores, as EIs possuem valor expressivo, presente na oralidade, fato que contribui para sua complexidade e abstração, em termos de análise linguística, e até mesmo para sua dificuldade de dicionarização (fator que contribui também para a opacidade do fenômeno), conforme aponta Riva (2009).

Como **objetivo geral**, busca-se investigar o estatuto semântico-cognitivo das EIs encontradas em variedades lusófonas do português, utilizando como princípios norteadores, a motivação, para análise cognitiva, e a composicionalidade, para análise morfossintática, a fim de definir graus de transparência e opacidade das EIs (Keizer, 2016, 2018), para propor uma tipologia dessas expressões para o português.

Como **objetivos específicos**, busca-se: (i) verificar como se dá o processo de esforço extra-cognitivo do falante para reanálise das EIs *AD-HOC*; (ii) averiguar o alinhamento da representação formal das EIs, de acordo com os três níveis mais altos da GDF (Hengeveld; Mackenzie, 2008); e (iii) determinar como ocorre a violação da relação de um-para-um entre forma e significado dos elementos constituintes das EIs (Leufkens, 2015).

3.1.1: Problemáticas e hipóteses da tese

Os objetivos acima descritos foram elaborados a fim de responder às seguintes problemáticas de pesquisa: (i) É possível verificar uma relação sistemática entre funções pragmáticas e semântico-cognitivas que influenciam o comportamento morfossintático de EIs e sua representação na gramática das línguas? Em caso afirmativo, (ii) Como se dá a caracterização dessas expressões no português, sob a óptica da GDF, uma teoria funcional-estrutural (Butler, 2005), que leva em consideração a função, mas tem como foco primordial fenômenos que manifestam distinção na codificação, ou seja,

²⁰ Conforme explicitamos anteriormente, os dados construídos foram elaborados com o intuito de atestar a funcionalidade (ou não) de casos de variação morfossintática. Esses dados construídos serão utilizados quando não encontrarmos casos de variação morfossintática no CRPC e na *web*, sendo nossa terceira fonte de consulta.

na estrutura das unidades linguísticas? (iii) Como se avalia a transparência e a opacidade dessas expressões em situações de uso, considerando-se o significado que assumem e o papel dos Participantes na compreensão do sentido em diferentes contextos (Leufkens, 2015; Keizer, 2016, 2018)? (iv) Como as interpretações metafóricas de EIs podem ajudar no processo de ativação semântica (Langlotz, 2006) e de composicionalidade de suas partes?

Para responder a essas quatro problemáticas, lançamos quatro hipóteses a serem investigadas: (i) possibilidade de encontrar no português os quatro tipos de EIs que Keizer (2016, 2018) definiu para o inglês, tendo como norte de investigação os princípios de motivação e composicionalidade; (ii) as EIs do português podem apresentar resultados de variação morfossintática similares aos vistos no trabalho da autora, além de outros resultados específicos, visto que o português pode apresentar alguma especificidade com relação à estruturação de EIs; (iii) a possibilidade de ocorrências em referências anafórica e catafórica apresentarem algum tipo de violação da relação entre forma e significado; e (iv) a possibilidade de encontrar algum alinhamento padrão entre os níveis da GDF para os tipos de EI, situação vista no trabalho de Keizer (2016).

Para cumprimento do objetivo geral - investigar o estatuto semântico-cognitivo das EIs encontradas em variedades lusófonas do português, analisa-se o estatuto semântico-cognitivo das EIs através de **duas etapas de análise - análise cognitiva e análise morfossintática**, a fim de definir graus de transparência e de opacidade das EIs, para a proposta de uma tipologia para o português, de acordo com os trabalhos de Keizer (2016, 2018).

Para a **primeira etapa, da análise cognitiva**, busca-se compreender a definição dicionarizada da EI em sua forma consagrada e sua interpretação metafórica, além de analisar o contexto de inserção dessas expressões nos *corpora*, etapa importante para compreendermos sua interpretação como idiomática, tanto em sua forma consagrada, como também em sua forma variada. Incluem-se ainda, nessa proposta de análise, a composição metafórica como fator de motivação no processo de criação e codificação das EIs com base nos tratados de Lakoff e Johnson (1980), sobre interpretação metafórica, e de Langlotz (2006), sobre ativação semântica, que, conforme observado, engloba os fatores de motivação constitucional e composicionalidade literal, além das misturas metafóricas que são ativadas parcialmente como imagens metafóricas em partes constituintes das expressões, ponto-chave que nos interessa para delimitação do quinto tipo de EI.

Para a **segunda etapa, da análise morfossintática**, além de aplicar os oito critérios de variação morfossintática inspirados em Keizer (2016) - substituição, modificação, passivização,

nominalização, definitude, número, quantificação e anaforicidade – foram adicionados seis novos critérios - referência catafórica, inserção de advérbios, inserção de pronomes, inversão da ordem preferida, elisão de um termo fixo e possibilidade de conversão para modo imperativo - totalizando **quatorze critérios** de variação morfossintática para análise morfossintática das EIs. Observa-se, nessa etapa, se os resultados de variação morfossintática são similares aos vistos nos trabalhos de Keizer (2016,2018) e considera-se a possibilidade do surgimento de novas conjunturas semânticas nas EIs do português.

Busca-se, assim, aplicar ao português, a classificação dos quatro tipos de EIs definidos por Keizer (2016,2018): (i) EIs não-motivadas (opacas) e semanticamente não-decomponíveis (e.g. *to kick the bucket*); (ii) EIs motivadas (transparentes) e semanticamente não-decomponíveis (e.g. *to hit the ceiling*); (iii) EIs motivadas (+ transparentes) e decomponíveis (*to spill the beans*), (iv) EIs *AD-HOC* (e.g. *to gnaw the bullet*), além da possibilidade de outros tipos, conforme especificidades identificadas em EIs do português.

Para o cumprimento do **primeiro objetivo específico** - verificação dos processos de esforço extra-cognitivo do falante para reanálise das EIs *AD-HOC*, buscamos entender como ocorre a reanálise das novas expressões *AD-HOC*, levando em conta uma análise semântico-cognitiva, com base no processo de mudança de não-decomponibilidade para decomponibilidade (Keizer, 2018), através da interpretação das cenas literal e idiomática da EI *AD-HOC* e da sua relação com a EI consagrada.

O **segundo objetivo específico** - averiguação do alinhamento da representação formal das EIs de acordo com os três níveis mais altos da GDF, é alcançado com a aplicação da **terceira etapa de análise - análise das EIs de acordo com os níveis e camadas da GDF**. Nessa etapa toma-se como critérios: (i) a investigação dos aspectos pragmáticos e camadas do NI (se a análise da identificabilidade, especificidade, além da função pragmática são relevantes, se são Subatos); (ii) aspectos semânticos e camadas do NR (se funcionarão na camada do Conteúdo Proposicional, ou da Propriedade Configuracional e/ou do Indivíduo, se os constituintes estarão numa relação hierárquica ou não-hierárquica); e (iii) aspectos morfossintáticos e camadas do NM (ordenação e suas formas sintáticas). Ainda nos resultados dos níveis da GDF, entende-se que existe a possibilidade de um alinhamento padrão, fato encontrado nos trabalhos de Keizer (2016).

A exemplo disso, as expressões do tipo 1, trabalhadas pela autora (desmotivadas, semanticamente não-decomponíveis, e.g. *to kick the bucket*), funcionam como um Subato Atributivo

(T) no Nível Interpessoal, como um Estado-de-Coisas (e) no Nível Representacional e possui três unidades codificadas no Nível Morfossintático: Palavra Verbal (Vw), Palavra Nominal (Nw) e uma Palavra Gramatical (Gw). Assim, a autora entende que há transparência entre NI e NR, mas não entre NI e NM, nem NR e NM.

Já para cumprir o **terceiro objetivo específico** – determinação dos apontamentos de violação da relação de um-para-um entre forma e significado dos elementos constituintes das EIs, trabalha-se com os elementos que abrangem as três etapas de análise. Imaginamos haver a possibilidade de casos de referências anafórica e catafórica nas EIs *AD-HOC*, encontradas em camadas mais altas da teoria, o que contribui para o possível surgimento de casos de violação da transparência entre forma e significado. Leufkens (2015) trata essa violação como **descontinuidade** - divisão de uma unidade formal que resulta em um alinhamento não-paralelo entre os NI, NR e NM.

Com foco nos estudos de Leufkens (2015) e na GDF, analisa-se, assim, como ocorre essa violação em algumas ocorrências encontradas no CRPC. Entendemos que essa descontinuidade pode ser compreendida na camada do Discurso do NI (Stassi Sé, 2012), pois foram encontradas em uma extensão maior que o Movimento. Notamos esse fenômeno nas ocorrências em referência anafórica e catafórica entre os turnos de fala dos participantes.

3.1.2: Critérios e fatores de análise

Nossa metodologia de pesquisa foi organizada em três etapas de análise, detalhadas na introdução e na subseção anterior: (i) **análise cognitiva das EIs**, organizada com os seguintes critérios de análise: (i) a definição dicionarizada da EI em sua forma consagrada; (ii) interpretação do contexto de inserção da EI; e (iii) interpretação metafórica das EIs. Esclarecemos novamente que selecionamos excertos com as EIs em sua forma consagrada e também em sua forma com variação.

Com base nos bancos de dados das EIs, trabalha-se com a definição dicionarizada da EI em sua forma consagrada, e assim compreende-se os processos de interpretação metafórica das EIs, de acordo com a ativação semântica e composicionalidade das partes componentes dessas expressões. A leitura do contexto de inserção nos auxiliará a compreender esse processo de análise das leituras literal (veículo) e idiomática (alvo), conforme vimos em Langlotz (2006).

Já a segunda etapa, de **análise morfossintática das EIs**, é baseada na aplicação dos quatorze

critérios de variação morfossintática: substituição lexical, modificação adjetiva, quantificação, número, definitude, passivização, nominalização, referência anafórica, referência catafórica, inserção de advérbios, inserção de pronomes, inversão da ordem preferida, elisão de um termo fixo, possibilidade de conversão para modo imperativo.

Conforme análise cognitiva, tomaremos a forma consagrada como padrão dicionarizado da EI, e trabalharemos os quatorze critérios de variação morfossintática, a fim de analisar o comportamento morfossintático das EIs e seus graus de composicionalidade. Os casos de EIs *AD-HOC* e os casos de descontinuidade entre forma e significado das partes constituintes também serão investigados de acordo com os três critérios de análise cognitiva e com os quatorze critérios de análise morfossintática.

A terceira e última etapa é a **análise das EIS de acordo com os níveis e camadas da GDF**. Será feita a representação formal das EIs de acordo com os NI, NR e NM, tendo como critérios de análise a investigação dos aspectos (i) pragmáticos; (ii) semânticos; e (iii) morfossintáticos. Com os resultados obtidos, será investigado se existe algum padrão de alinhamento entre os resultados de cada nível. Nossa intenção é definir se há transparência ou opacidade nesses resultados de alinhamento.

Dessa forma, a análise de dados se organiza da seguinte maneira: (i) apresentação do excerto; (ii) análise cognitiva das EIs; (iii) análise morfossintática das EIs; (iv) análise das EIs de acordo com os níveis e camadas da GDF. Para melhor compreensão da análise e discussão dos dados, optou-se por ordenar em ordem crescente a apresentação dos cinco tipos de EIs encontrados no português, com seus respectivos excertos e fatores de análise, partindo do tipo um até o tipo cinco.

Ao final da análise de dados, espera-se que cada etapa de análise nos auxilie a cumprir os objetivos e possa confirmar as hipóteses de pesquisa. As análises cognitiva e morfossintática das EIs oferecem suporte para definir cinco tipos de EIs, de acordo com os graus de transparência e opacidade dessas expressões no português, além de nos ajudar a compreender o processo de reanálise das EIs *AD-HOC*, e das EIs em descontinuidade. A análise das EIs de acordo com os níveis e camadas da GDF apresenta-se como crucial para verificar se há um possível padrão de alinhamento entre os níveis, e pode apontar para alguma tendência de padrão de transparência em EIs no português.

Cumprir lembrar que nosso objetivo nesta investigação não passa por preocupações de cunho sociolinguístico, variação linguística, estudos etimológicos, lexicográficos ou traduções entre as

diferentes variedades de português, ou entre português e inglês.

Entendemos que, por ser um *corpus* limitado em diferentes variedades do português e sob determinada limitação temporal dos registros, poderá haver casos de diferentes variações morfossintáticas, motivadas pragmática ou semanticamente, que nossa classificação formal da variação morfossintática aqui proposta não seja suficientemente estruturada explicar, pois, conforme Simões (2018), as EIs são idiossincrasias que individualizam determinada comunidade, caracterizam e delimitam sua cultura, seus costumes e vivências que naturalmente acompanham e evoluem linguisticamente.

Resumimos nossa fundamentação metodológica no quadro abaixo:

FUNDAMENTAÇÃO METODOLÓGICA DA PESQUISA		
Problemáticas	1- É possível verificar uma relação sistemática entre funções pragmáticas e semântico-cognitivas que influenciam o comportamento morfossintático de EIs e sua representação na gramática das línguas?	
	2- Como se dá a caracterização das EIs em diferentes variedades lusófonas do português, sob a óptica da Gramática Discursivo-Funcional (GDF)?	
	3- Como avaliar possíveis graus de transparência/opacidade dessas expressões em situações de uso?	
	4- Como as interpretações metafóricas de EIs podem ajudar no processo de ativação semântica (Langlotz, 2006) e composicionalidade de suas partes?	
Objetivos	Geral 1- Investigar o estatuto semântico-cognitivo das EIs encontradas em variedades lusófonas do português, utilizando como princípios norteadores, a motivação e a composicionalidade a fim de definir graus de transparência e opacidade das EIs e propor uma tipologia para o português.	
	Específicos	1- Verificação dos processos de esforço extra-cognitivo do falante para reanálise das EIs AD-HOC.
		2- Averiguação do alinhamento da representação formal das EIs de acordo com os três níveis mais altos da GDF.
Hipóteses	3- Determinação dos apontamentos de violação da relação de um-para-um entre forma e significado dos elementos constituintes das EIs.	
	1- Possibilidade de encontrar no português os quatro tipos de EIs que Keizer (2016, 2018) classificou no inglês, tendo como norte de investigação os princípios de motivação e composicionalidade.	
	2- As EIs do português podem apresentar resultados de variação morfossintática similares aos vistos no trabalho da autora, além de outros resultados específicos, visto que o português pode apresentar alguma especificidade com relação à estruturação de EIs, o que nos levou a pensar na possibilidade de um novo tipo de EI.	

	3- Possibilidade de ocorrências em referências anafórica e catafórica apresentarem algum tipo de violação da relação entre forma e significado, pontos entrelaçados com a não-transparência dessas expressões, rompendo, possivelmente, a relação um-para-um entre níveis da GDF e também entre forma e significado, dado que contribui para a opacidade dessas expressões vistas nessas situações, além do surgimento de EIs <i>AD-HOC</i> , que entendemos poder haver uma dependência de contexto específico para atestação.	
	4- Possibilidade de encontrarmos algum alinhamento padrão entre os níveis da GDF para os tipos de EI, situação vista no trabalho de Keizer (2016).	
Metodologia	Etapas	Critérios
	1-Análise cognitiva das EIs	1-Definição dicionarizada da EI em sua forma consagrada; 2- Interpretação do contexto de inserção da EI; 3-Interpretação metafórica das EIs.
	2-Análise morfosintática das EIs	1-Substituição lexical; 2-Inserção de pronomes; 3-Inserção de advérbios; 4-Modificação adjetiva; 5-Número; 6-Quantificação; 7-Definitude; 8-Nominalização; 9-Passivização; 10-Possibilidade de conversão para modo imperativo; 11-Elisão de um termo fixo; 12-Inversão da ordem preferida; 13-Referência anafórica; 14-Referência catafórica.
	3-Análise das EIs de acordo com os níveis e camadas da GDF	1-Aspectos pragmáticos e camadas do NI; 2-Aspectos semânticos e camadas do NR; 3-Aspectos morfosintáticos e camadas do NM.

Quadro 4: Fundamentação metodológica da pesquisa (Elaboração própria, 2025)

Seguimos agora para a apresentação dos bancos de dados de EIs (dicionários) e dos *corpora* utilizados nesta pesquisa.

3.2: Banco de dados de EIs

3.2.1: Dicionário onomasiológico das EIs

O dicionário onomasiológico de EIs de Riva (2009) foi selecionado como nossa principal fonte de consulta. O dicionário é do tipo onomasiológico, apresenta uma formatação de busca linguística a partir de uma ideia determinada que examina diversos modos com os quais uma

expressão linguística é encontrada em dada língua, através de aspectos ligados à denominação dos signos, fato que auxilia na memorização e integração das expressões linguísticas. Tal aspecto o auxiliou na organização do dicionário por conceitos-chave, e não palavras-chave. Como por exemplo, o conceito de *embriaguez* e seu rol de EIs, *como encher a cara, entortar o caneco, estar mamado*, etc.

Dentre as diversas dificuldades apontadas pelo autor na elaboração de um dicionário fraseológico, estão a nomenclatura (se semasiológico ou onomasiológico, por exemplo); e a denominação do objeto a ser dicionarizado – como delimitar o que é uma EI, um provérbio, um adágio, fórmulas rotineiras, etc.; fonte de dados; gêneros textuais, etc. Ao consultar outros dicionários fraseológicos, o autor notou uma certa preferência por registros de linguagem formal extraídos de textos literários ou científicos, mas entende que o uso coloquial é um “terreno fértil” para EIs, principalmente na *internet*.

Conforme abordado no primeiro capítulo, Riva (2009) entende EI como lexias complexas, com valor idiomatizado, com certo grau de fixação dos elementos e usos, além de serem convencionalizadas na comunidade falante.

Já sobre a frequência de EI ser considerada minimamente suficiente para ser inserida num verbete dicionarizado, o autor elenca diversas dificuldades em se trabalhar com a dados da *web*, principalmente nos critérios para eliminar casos atípicos ou de emprego inusitado, além da análise de excertos que divergem entre expressões com ativação semântica literal apenas (e.g. *puxar o tapete*), dados que se perdem ao passar do tempo, etc., ou elementos formadores das expressões que caíram em desuso, como o disco de vinil, elemento presente na expressão *virar o disco*.

Há casos de expressões que expandiram seu nicho de aplicação, como *ter espírito de porco* (base bíblica referente ao evento em que Jesus Cristo expulsou demônios de um grupo de pessoas e os enviou a uma manada de porcos, e que hoje é utilizado para descrever pessoas desagradáveis ou que provocam constrangimento), e *mudar de time* (base no esporte e que hoje, além de remeter ao conceito de traição, ganhou um sentido pejorativo para mudança de orientação sexual – heterossexualidade para homossexualidade), etc.

Para trabalhar com dados da *web*, o autor utiliza a unidade de medida PMW (*per million words* – palavras por milhão), e toma como parâmetro, para aquele período de coleta de dados – 2004 a 2008, o valor de cinquenta e seis milhões de palavras para o português encontradas na *web*, com base em estudos de vários autores, como Evans et al. (2004). Assim, Riva (2009) aplicou o coeficiente de 1 PMW ao limiar da frequência de cinquenta e seis ocorrências para o português. Como gerenciador de busca, utilizou o *Google* e como *corpora* da *web*, o *SkecthEngine* e o Laboratório de Lexicografia

da UNESP de Araraquara (*corpus* fechado).

Como exemplo, a busca pela expressão *rato de biblioteca* no *Google* durante a elaboração do dicionário resultou no número de 31.700 entradas. Em 2024, a mesma busca resultou em número superior a quatro milhões de entradas em 0,27 segundos, apenas em páginas em português, o que significa um aumento gigantesco nesses dezesseis anos de avanços tecnológicos. O autor trabalhou com um total de 1.562 EIs, agrupadas em quatrocentos conceitos, sendo 32% entendidos como positivo (e.g. amor), 46% como negativo (e.g. falsidade), e 22%, neutro (e.g. tempo). Vejamos abaixo alguns exemplos encontrados em Riva (2009):

baixar o pau – falar mal de alguém ou de alguma coisa [hip.; orig. sup.: alusão ao ato de violência de bater com um pau para indicar difamação] ♦ Era ali que a esquerda se reunia na década de 70 para *baixar o pau* na ditadura militar. (www.terra.com.br/istoe/1765/comportamento/1765_cheio_graca.htm; acesso em 27/05/04).

baixar o sarrafo – falar mal de alguém ou de alguma coisa [hip.; orig. sup.: alusão à agressividade para indicar difamação] ♦ E aí começa a *baixar o sarrafo*, dizendo que um é viado, a outra é sapatão, o outro é bossal, o outro é ladrão, a outra é envolvida com tóxicos... (www.panoramaespirita.com.br/colunas/alamar_regis/artigos/o_critico_espirita.html; acesso em 27/05/04) (Riva, 2009, p. 135)

Como podemos ver, o autor traz uma descrição da significação idiomática da expressão e um exemplo real de uso (excertos de jornais, entrevistas, etc.), além de classificar as expressões por proximidade semântica e construção lexical, agrupadas no conceito de *crítica*. Há casos em que a mesma expressão se enquadra em diferentes conceitos, como *botar pra fora* (desabafar, reclamar ou despedir alguém). Seguimos agora para o dicionário de Silva (1997).

3.2.2: Dicionário sobre a origem das palavras e expressões da língua portuguesa

Já a obra de Silva (1997) é uma obra que traz frases e curiosidades da língua portuguesa, em forma de verbetes dicionarizados, ordenados em ordem alfabética para palavras, trazendo diversas origens históricas e socioculturais de EIs, como por exemplo “*Levar um puxão de orelhas*”: expressão de origem francesa, que significa repreender, ação praticada em tradições antigas no período das Ordenações Afonsinas (1469-1524), onde era costume cortar as orelhas de ladrões, que, após séculos,

as orelhas deixaram de ser cortadas, e eram puxadas em forma de castigo e repreensão.

Os verbetes dicionarizados também discorrem da origem etimológica das palavras e de suas explicações idiomáticas (quando tal palavra é vinculada a uma expressão), como *abacaxi*: palavra de origem tupi, que também é utilizada na expressão *descascar o abacaxi*, como sendo uma tarefa difícil, devido ao cheiro, aspecto espinhoso e a dificuldade na ação.

Apesar de o dicionário trazer uma mescla de diversos termos (e.g. *propina, quenga, quaresma, grife, etc.*), frases de pensadores famosos, como *O coração tem razões que a razão desconhece*, do matemático, filósofo e escritor francês Blaise Pascal (1623-1662), ou frases curiosas do português (e.g. *Ordem e Progresso*, lema inscrito na bandeira do Brasil, *Pagar o pato*, expressão presente em vários textos da literatura portuguesa e brasileira, como em Gregório de Matos (1636-1696)), a obra apresenta vários verbetes específicos que remetem e explicam sua relação com EIs, além da contextualização histórica de outras, como em (3-1) e (3-2), respectivamente um verbete e uma EI comentada:

- (3-1) Pitanga: do tupi, piranga, vale vermelho, veio a variação pitanga. Como em Portugal havia a expressão ‘chorar lágrimas de sangue’, entre nós transformou-se em chorar as pitangas, dado que este “fruto globular é de linda carnação rubra”, no dizer do folclorista brasileiro Luís da Câmara Cascudo (1899-1982). (Silva, 1997, p. 208)
- (3-2) Dourar a pílula: Esta frase tem o significado de apresentar algo difícil ou desagradável como coisa fácil de aceitar. Nasceu de conhecida prática das farmácias antigas, que consistia em embrulhar as pílulas em finos papéis, como o fim de preparar psicologicamente o cliente para engolir um remédio de gosto amargo. Do sentido literal, passou a metáfora e logo recebeu aplicação literária, estando em um de seus mais antigos registros na peça *Anfitrião*, de Jean-Baptiste Poquelin Molière (1622-1673), em que Sósia, na última cena do terceiro ato, diz: “o senhor Júpiter sabe dourar a pílula”. *Dourar a pílula* é ainda hoje tática sutil de persuadir renitentes, quando se procura destacar os aspectos positivos de algo desagradável. (Silva, 1997, p. 107)

Tais conexões entre palavras e expressões, embora que de modo aleatório, sem uma sequência direta entre verbetes e expressões separadas por categorias, como frases curiosas, termos, expressões idiomáticas, provérbios, etc., o dicionário de Silva (1997) mostrou-se de grande valia para o nosso trabalho, devido à variedade de verbetes que remetem ao contexto sociocultural entre o português de diferentes variedades e em muitos casos, sua conexão com línguas africanas e indígenas, como visto acima com a palavra *pitanga*, ou a palavra *maconha*, do quimbundo (língua falada na região de Luanda, em Angola) *ma’kaña*, droga derivada do cânhamo, também conhecida em latim como *cannabis sativa* (Silva, 1997). Seguimos agora para o dicionário de Almeida (2022).

3.2.3: Dicionário de locuções e expressões da língua portuguesa

O dicionário de Almeida (2022) é tratado como um dicionário em construção (aberto em linha), resultado inicial de um exercício de Linguística, de linguagens de programação de Perl.²¹ Com cerca de quatro mil entradas, o dicionário foi elaborado utilizando a linguagem de programação DPL (*Dictionary Programming Language*) do projeto Natura da Universidade do Minho, em Portugal.

Esse projeto, liderado pelo autor, refere-se a um grupo de investigação em Processamento de Linguagem Natural do Departamento de Informática, que contribui também para um grupo de Processamento e Especificação de Linguagens. A ideia do trabalho é desenvolver vários projetos e ferramentas para o campo da Linguística Computacional, tais como o *jSpell* (analisador morfológico derivado de um corretor ortográfico), *NATools* (processamento de *corpus*) e dicionários do português.

A estrutura das entradas dos verbetes contém o significado semântico, sinônimos, exemplos de uso, com sete níveis do calão – desde normal até calão “estupidamente carroceiro”/pejorativo. O dicionário define as entradas em termo, frase pitoresca, interjeição, provérbio, insulto, termo calão, EI e outras. Há também uma distinção geográfica entre diferentes regiões de Portugal (norte, sul e centro), Brasil, Angola, Açores, Madeira, e outros locais. Vejamos abaixo em (3-3) um verbe de da expressão idiomática *ficar nos cornos do touro*:

(3-3) **Ficar nos cornos do touro**

é um: idiomática

ficar mal visto, ser culpabilizado ou responsabilizado por uma situação ingrata

sin.: lixar-se; tramar-se; lerpar; já foste!

ex.: Sempre que há confusão, eu é que fico sempre nos cornos do touro.

local: pt centro (Almeida, 2022, p. 118)

Algumas expressões foram agrupadas de acordo com a mesma significação idiomática, como *viver à barba longa*, *viver à grande e a francesa* e *viver à larga*, significando ser rico, viver no luxo, além da mesma expressão com alguma alteração nos componentes de acordo com a variedade regional, como *vestir a camisa* (Brasil) e *vestir a camisola* (Portugal), algo similar ao que vimos nas EIs do inglês *to take the cake* e *to take the biscuit* (variedades americana e britânica, respectivamente).

Esses três dicionários foram trazidos a título de registro e atestação teórica da significação das

²¹ Linguagem de programação utilizada para manipulação de texto, administração de sistemas, desenvolvimento da *web* e programação de rede. É uma linguagem de código aberto, o que permite rápido desenvolvimento e popularidade entre os desenvolvedores de aplicativos da *web*.

EIs utilizadas nas análises, pois, mesmo que tais expressões sejam tratadas como de conhecimento trivial e partilhado ao longo dos anos, é importante para nossa pesquisa que tenhamos embasamento teórico sobre seus registros.

Os três dicionários se estendem em uma diferença temporal (entre publicações) de vinte e cinco anos - 1997-2022, e somam mais de dez mil entradas, que incluem, em sua maioria, EIs. Com essa quantidade relativamente elevada de registros, verificamos que os três dicionários contemplam registros das expressões selecionadas para análise funcional, com detalhamento da significação idiomática, exemplos de uso, e em alguns casos encontrados em Silva (1997), origem etimológica da expressão ou de termos-chave das expressões, e.g. termo *Diabo* e a expressão *como o Diabo fuge da cruz*, ou precisão geográfica, e.g. *dar gasosa – Angola, molhar a mão – Brasil*, em casos registrados em Almeida (2022), ou agrupamento de conceitos de expressões similares, e.g. conceito de *morte – vestir o paletó de madeira, bater as botas, estar com o pé na cova, comer capim pela raiz* em casos encontrados em Riva (2009). Vejamos outros bancos de dados de EIs, entendidos como materiais secundários de apoio.

3.2.4: Outros bancos de dados de EIs: consulta secundária

A ferramenta *Google* mostrou ser de grande valia para extração de casos e registros em diversas fontes, como *blogs* sobre a língua, textos literários, de entretenimento, etc., conforme podemos ver em (3-4) o excerto extraído do site <https://www.dicio.com.br> - Dicionário online do português, com a definição e exemplos da EI *quebrar a cara* (última atualização da página no momento da pesquisa: 2012):

- (3-4) Quebrar a cara: Ter um plano frustrado, sofrer decepção, se dar mal; desiludir-se, desenganar-se; falhar, errar: dizem que abrir um negócio e quebrar a cara é o que há de mais didático no empreendedorismo. Brigar, lutar com alguém; dar murros ou socos; esbofetear, esmurrar: eu vou quebrar a cara daquele bandido.
Sinônimos de quebrar a cara: falhar, errar.
Exemplos com Quebrar a cara:
Ele disse que os candidatos da oposição vão "**quebrar a cara**" se usarem o escândalo como arma política no segundo turno. *Folha de S. Paulo, 10/10/2012*
Mas a chance de **quebrar a cara** é alta por falta de formação em gestão. *Folha de S. Paulo, 13/12/2012*
O ministro avisou aos presentes que quem tentar "desestabilizar" a ministra vai "**quebrar a cara**". *Folha de S. Paulo, 12/05/2011* (Fonte da internet: Dicio, 2012)

Tal expressão também foi encontrada no dicionário de Riva (2009) com a mesma significação e exemplos extraídos da *web*. Durante o levantamento bibliográfico, foi considerado, para reforço ou

eventual consulta secundária, a inclusão de outros dicionários ou trabalhos acadêmicos sobre dicionários fraseológicos, como o artigo de Lauand (2022) sobre verbetes comentados para composição de dicionário fraseológico com EIs e gírias do português falado no Brasil (e.g. *pôr os pingos nos is* vista em (3-5)); o trabalho de Carvalho e Turé (2023) sobre EIs presentes nas línguas portuguesa brasileira e guineense (e.g. *segurar vela* em (3-6)); e a tese de doutorado de Pereira (2015) que conta com o registro de EIs do português europeu presentes na BADAF – Base de Dados Fraseológica do português europeu, constituída por colocações, locuções, enunciados fraseológicos e EIs, como em *apertar o cinto* (3-7):

- (3-5) **Pôr os pingos nos is:** Com o mesmo sentido que tem hoje (deixar as coisas claras, esclarecer uma situação dúbia), aparece na BN em 1861: (...) voltarei à questão, porei pingos nos iis – e declinarei nomes e pronomes (...). (“O Regenerador” RJ, 08-05-1861) (Lauand, 2022, p. 8)
- (3-6) **Expressões idiomáticas no guineense:** pegar vela/Expressões idiomática no Português Brasileiro: *segurar vela*/significado literal traduzido do guineense: *segurar vela*/significado conotativo: ficar na companhia de duas pessoas que namoram, bancar o cupido (Carvalho; Turé, 2023, p. 517-518)
- (3-7) **Apertar o cinto** (c. de 697 000 resultados, 17/01/14), fazer um grande sacrifício económico; passar por uma grande dificuldade económica; perder poder de compra: “Portugueses são quem vai apertar mais o cinto na Europa logo a seguir aos gregos” (Público, 02/10/10, 1.^a p., título) (Pereira, 2015, p. 496)

Novamente, atestamos o registro dessas três expressões vistas acima também no dicionário de Riva (2009), fonte principal de nossa consulta, o que reforça sua importância e dos demais dicionários e trabalhos citados que nos servem de base para consulta secundária dos registros dicionarizados das EIs, contemplando as variedades brasileira, europeia e africanas, fato que dialoga com o que vimos na expansão das metáforas culturais entre universais e quase universais (Kövecses, 2010) entre as línguas e culturas, e em especial a relação sociocultural e histórica entre Brasil, Portugal e África e sua materialização no português.

Para Bonvini (2015), os enlaces linguísticos e ideológicos entre as variedades brasileira, africana e indígenas contribuíram para a formação de uma ideologia nacionalista:

A problemática da relação entre as línguas africanas e o português do Brasil (PB), no século XX, colocou-se desde os anos 30. Desde o começo, buscou-se evidenciar a participação daquelas na constituição da variante brasileira do português. Esse debate desenvolveu-se em dois planos independentes, o lingüístico e o ideológico, prolongando assim as preocupações nascidas por ocasião da independência do Brasil em 1822. Tratava-se, na época, de marcar a diferença entre o Brasil e Portugal. A ideologia nacionalista orientou então os estudos lingüísticos na direção da pesquisa dos elementos autóctones diferenciadores, engendrados pela presença das línguas indígenas e africanas. (Bonvini, 2015, p. 16)

O autor destaca que os contatos múltiplos e constantes entre o português brasileiro e africano resultaram em maior flexibilidade na língua, fluidez estrutural, com dinamismo interno, mas disposto por lutas de resistência e inovação.

Ao procurar por materiais específicos das EIs do português encontradas nos excertos de diálogos dos países africanos, nos deparamos com seus registros oriundos de variedades africanas no dicionário de Almeida (2022), além de trabalhos acadêmicos e teses, como o trabalho de Simões (2018) sobre o ensino de EIs em países falantes de português como língua oficial, e também o trabalho de Teletin e Marques (2011), em que as autoras realizam estudos comparativos sobre EIs do tipo gentílico ou toponímico do português e também do francês, com o intuito de analisar os estereótipos linguísticos de cada cultura. Nesse estudo há muitos exemplos de EIs muito peculiares a determinados eventos socioculturais de cada país, como por exemplo, *chamar pelas cebolas do Egito*:

No caso do povo judeu, a Língua Portuguesa caracteriza-o de maneira negativa como podemos ver através destes poucos exemplos: *Fazer judiarias, Alma de judeu, Chamar pelas cebolas do Egito*. As cebolas do Egito tornaram-se famosas pelas lamentações do povo de Israel enquanto peregrinava pelo deserto fugindo da escravidão. Assim, podemos atestar que o povo judeu é conotado negativamente. (Marques; Teletin, 2011, p. 347)

Para este trabalho, nota-se que, em ambos os casos - excertos de Portugal, Brasil e de países africanos, as EIs selecionadas e identificadas nessas diferentes variedades tiveram o mesmo significado idiomático registrado nos dicionários citados, o que de fato não impede que as mesmas EIs tenham novos ou diferentes usos em determinadas regiões/países e épocas. Tal possibilidade não invalida nossa proposta, pois não pretendemos homogeneizar as significações das EIs como únicas para o português como um todo, e sim, entender seus respectivos usos nos casos e situações aqui selecionadas, com base na análise funcional proposta.

Para tal entendimento da expansão das EIs entre Brasil, Portugal e África, Holm (1987) realizou um estudo comparativo de diferentes línguas crioulas de base portuguesa e encontrou paralelismos entre expressões do português vernacular brasileiro, i.e. língua falada pelos brasileiros de extratos sociais mais baixos, e de línguas africanas, o que reforça a contextualização histórica evidenciada em seu estudo: o Brasil teve cerca de 40% dos escravos trazidos para a América e teve de modo tardio a abolição da escravidão. Seguimos para a apresentação dos *corpora* de análise.

3.3: Materiais de busca de EIs – *corpora* de análise

3.3.1: *Corpus* de Referência do Português Contemporâneo (CRPC), dados da *web* e dados construídos

O CRPC é um vasto *corpus* digital de variedades do português de Portugal, Brasil, Angola, Cabo Verde, Guiné-Bissau, Moçambique, São Tomé e Príncipe, Goa, Macau e Timor-Leste. O acervo conta com mais de 311,4 milhões de palavras, e abrange diferentes tipos de textos escritos (literário, jornalístico, técnico, etc.) e de registos orais (formal e informal), além de cerca de 1,6 milhões de palavras de gravações e transcrições faladas, conforme dados do quadro 5:

Pais	Escrito	Falada	Total
Portugal	289 840 619	1 470 593	291 311 212
Brasil	3 539 770	23 177	3 562 947
Angola	10 744 623	27 363	10 771 986
cabo Verde	1 449 269	25 413	1 474 682
Guiné-Bissau	364 421	25 016	389 437
Moçambique	1 126 299	26 166	1 152 465
São Tomé e Príncipe	537 600	25 287	562 887
Macau	2 086 763	6 775	2 093 538
Timor Leste	123 575	2 409	125 984
Goa	-	1 840	1840
Total	309 812 939	1 634 039	311 446 978

Quadro 5: Variedades abrangidas pelo CRPC e número de *tokens* por registro (CLUL, 2016)

Para acesso ao *corpus*, a plataforma CQPweb é dividida em: (i) CRPC (parte total escrita); (ii) Portugal (subparte do CRPC com textos da variedade do português europeu); (iii) Mini-CRPC (1% do CRPC disponível para usuários não cadastrados testarem a plataforma); e (iv) África (subparte do CRPC). Sobre os gêneros textuais presentes no *corpus*, há uma grande abrangência de tipos de textos escritos e transcrições faladas, com a inclusão de registros de programas de rádio, televisão, discursos e debates políticos, conferências, pregações e ensino.

Vejamos abaixo um exemplo de um excerto, encontrado na variedade de Angola, que contém as EIs “Dar com a língua nos dentes” (fazer fofoca, entregar alguém) e “ser amigo da onça” (ser falso amigo):

(3-8) Onde foi buscar esse dinheiro? Perguntem aos deslocados e refugiados Nzita Tiago **deu com a**

língua nos dentes Amigo de Samakuva ou da onça? Nzita Tiago, líder de uma das facções que se bate de armas na mão pela independência de Cabinda, revelou uma inconfidência que pode criar embaraços ao presidente da Unita. (Angola - Política 2005 J140417)

A filtragem por palavras e categorias gramaticais nos permitiu encontrar as EIs em diferentes variedades lusófonas do português, sem nos preocuparmos com estudos etimológicos, lexicográficos ou sociolinguísticos, nem mesmo tradução em diferentes variedades do português. Cada excerto filtrado possui um ficheiro de identificação com os dados relevantes, como código de identificação, fonte, localidade, título, nome do jornal, volume, ano de escolaridade, autor, data, gênero, número de palavras, entre outras informações.

Esse *corpus* foi selecionado como parte principal do nosso material de busca de EIs em situações de uso, base para extração de excertos para análise e em casos de insuficiência de dados relevantes, selecionamos **excertos da web**²², como no exemplo (3-9) com a expressão *chorar as pitangas* (reclamar):

(3-9) - Nem que ele **chore pitanga** lhe emprestarei dinheiro. (Fonte da *Internet*: Veja, 2014)

Conforme a proposta de Keizer (2016), em ocorrências em que as EIs já apresentam algum tipo de alteração em relação à forma consagrada, mantemos a forma variada com o(s) elemento(s) já introduzidos no excerto. Em casos de expressões que estão dispostas em sua forma consagrada, faremos a adição de **dados construídos**, e.g. *chore muita/essas pitangas grandes*. Essa necessidade de criação de dados construídos também surgiu na análise da autora, e.g. *to kick the empty bucket*, pois nem sempre os dados do CRPC e da *web* cobrem todas as possibilidades de variação morfossintática.

Com esses bancos de dados de EIs em mãos e com o levantamento dessas EIs nos *corpora* apresentados, tendo em mente esse pluralismo cultural refletido na língua, em especial, sobre as metáforas culturais que se fazem presentes nas EIs, partimos para a subseção de análise quantitativo-qualitativa das EIs.

²² Os dados quantitativos e qualitativos da *web*, bem como exemplos de usos das EIs encontradas nesse material de busca, estão descritos de modo completo na subseção 3.4: Análise quantitativo-qualitativa. Fizemos essa menção aos dados da *web* em 3.3.1 para reforçar sua inclusão como parte do material de busca de EIs em situações de uso.

3.4: Análise quantitativo-qualitativa

Com base nos bancos de dados de EIs e nos *corpora* de busca de EIs utilizados na pesquisa e descritos nas subseções anteriores, elencamos trinta e cinco EIs encontradas em diferentes variedades lusófonas do português. Chegamos a esse grupo de expressões após avaliar o registro em ao menos um dos dicionários utilizados como banco de dados de EIs, a frequência no CRPC e a frequência na *web*, dos casos de variação morfossintática, fator importante para investigação do comportamento semântico-cognitivo das expressões.

Sobre a frequência de dados encontradas no CRPC, agrupamos os resultados em três colunas no quadro 6 a seguir: a primeira coluna sobre os dados do CRPC apresenta os resultados quantitativos com a filtragem da expressão em sua forma consagrada (e.g. *tirar o cavalo da chuva*), a segunda coluna com resultados de filtragem por palavras-chave da expressão, pensando na variação dos verbos em diferentes tempos verbais, como por exemplo a expressão *perder a cabeça*, pesquisada como *perd* (raiz do verbo) + *a cabeça*, fator que notamos ser importante para captação de casos em variação morfossintática. Já a terceira coluna preenchemos com exemplos encontrados de variação morfossintática.

Como vimos anteriormente, o CRPC conta com variedades de diversos gêneros textuais e engloba variedades do português europeu, brasileiro e africano, além de dados de Goa e Macau. Não propomos aqui fazer distinções de resultados por país ou variedade, conforme já delimitamos anteriormente.

Durante a filtragem de resultados no CRPC, tivemos os resultados baseados na totalidade de palavras e textos que compõem o *corpus*: **309.812.943** palavras dispostas em 356.208 textos. Pensando no grande avanço tecnológico da *internet* durante o intervalo de 2009 para 2024, notamos que os parâmetros quantitativos vistos em Riva (2009) - a frequência de cinquenta e seis milhões de palavras do português da variedade brasileira encontradas na *web* e o valor de cinquenta e seis ocorrências de EIs em *PMW*, apresentam uma grande diferença com os dados atualizados no primeiro semestre de 2024.

Com relação aos dados de frequência da *web*, utilizamos o gerenciador de busca *Google* e trouxemos os dados quantitativos, que correspondem aos números de resultados e a velocidade de busca, encontrados na primeira semana de junho de 2024, em páginas de língua portuguesa, restrição que ativamos nas ferramentas de pesquisa.

Com a intenção de aprimorar a filtragem de dados e evitar ambiguidade ou casos de significações literais (e.g. *descascar o abacaxi*), utilizamos como palavra-chave o termo “expressão idiomática” + frase (e.g. expressão idiomática + *tirar o cavalo da chuva*), e percebemos maior refinamento de resultados em testes com e sem a palavra-chave. Como exemplo, tomamos a expressão *descascar o abacaxi*: na busca geral encontramos 105.000 resultados gerados em 0,36 segundos, e com o termo “expressão idiomática”, gerou 17.200 resultados encontrados em 0,38 segundos.

Segundo a plataforma de pesquisa sobre usos da tecnologia na *internet Web Technologies Surveys* (W3Techs, 2024), o português ocupa a sétima posição no ranking de conteúdo da *internet*, representando cerca de 3,6% (inglês representa a maior porcentagem – 49,9%). Já de acordo com dados da plataforma *Corpus* do Português.org, o corpus da *web NOW Corpus (News on the web)* conta com cerca de 1 bilhão de palavras encontradas em jornais e revistas em quatro países falantes de português – Brasil, Portugal, Angola e Moçambique, durante o período de 2012 a 2019.

Já o *Corpus Web/Dialects*, criado em 2016, também conta com cerca de 1 bilhão de palavras, extraídas de aproximadamente 1 milhão de páginas da *web* dos quatro países citados acima, e contou com o apoio do *National Endowment for the Humanities* dos Estados Unidos, o que permitiu atualizações do *Corpus* do português original (2006).

Esses dados se refletem em nossa pesquisa no *Google* para cada uma das trinta e cinco expressões selecionadas, pois notamos que dez expressões selecionadas ultrapassaram o número de 1 milhão de resultados na pesquisa em páginas de língua portuguesa, e a expressão que apresentou número mais baixo foi de 14.800 resultados (*pendurar a chuteira*), números distantes do padrão mínimo considerado por Riva em 2009 – cinquenta e seis ocorrências do português brasileiro por *PMW*, considerando, no ano de sua pesquisa, o valor de cinquenta e seis milhões de palavras do português brasileiro na *web*, dado que em 2024 ultrapassou a barreira dos trezentos milhões de palavras só na variedade brasileira, conforme dados do *Corpus* do Português.org (2024). Fizemos uma conversão atualizada e encontramos o valor de mil ocorrências por *PMW* para EIs de diferentes variedades lusófonas do português:

1. Riva em 2009: 56.000.000 de palavras do português brasileiro na *web/PMW* = 56 ocorrências de EIs em registros brasileiros.

2. Autor em 2024 (elaboração própria): 300.000.000 de palavras do português brasileiro na *web*/PMW = 300 ocorrências de EIs em registros brasileiros.
3. Autor em 2024 (elaboração própria): 1.000.000.000 de palavras do português na *web*/PMW = 1.000 ocorrências de EIs em registros de diferentes variedades do português.

Esses dados se refletem no CRPC também, que, conforme apontamos, conta com mais de 309 milhões de palavras. Por exemplo, a expressão *pôr minhocas na cabeça* não apresentou registro no CRPC, dado relevante a ser considerado, já que realizamos a busca através de construções do tipo V+ Np em um *corpus* fechado, o que dificulta encontrar um alto número de ocorrências, que, em sua natureza composicional, já apresentam um caráter restritivo.

Essa mesma expressão, pesquisada na *web*, gerou mais de 250.000 resultados em junho de 2024, com diversos exemplos de usos. Nesses casos, fizemos uma mescla de resultados entre CRPC (prioridade da pesquisa), dados da *web*, e dados construídos, quando necessário. Seguimos com o quadro 6, com as expressões selecionadas e a respectiva legenda dos dicionários selecionados, criada para facilitar a organização visual do quadro:

EI	Significado	Dicionário encontrado	Frequência na <i>web</i> (N ^o /segundo)	CRPC		
				Forma consagrada	Forma variada	Exemplos de variação
1- Chorar as pitangas	Reclamação	R/S	27.580 (0,24)	0	0	N/A ²³
2- Pôr minhocas na cabeça	Desconfiança, influência	R/S	266.000 (0,27)	0	0	N/A
3- Engolir sapo	Tolerar situação desagradável	R/A/P/ CT	1.620.000 (0,31)	1	62	Tempos verbais, digerir, vivo, algum, um, maior, indigesto, pequeno, engolido, que engolem, mais este, enorme, primeiro, segundo, grande,

²³ Não aplicado.

	sem reclamar					cada, esse, sapos.
4- Tirar o cavalo da chuva	Desilusão, desistência, abandono	R/A/P/CT/L	1.250.000 (0,31)	3	17	Tempos verbais, cavalinho, seu, tirem.
5- Bater as botas	Morte	R/A/P/CT/	1.460.000 (0,30)	0	28	Tempos verbais, a bota.
6- Fazer tempestade num copo de água	Exagero	R/S/A/P	425.000 (0,41)	1	19	Tempos verbais, tão evidentemente, em um copo, pequenina, mais uma, alguma, levantar uma, violenta, ser uma
7- Descascar o abacaxi	Resolver/enfrentar problemas	S	17.200 (0,38)	1	2	Carioca, um dos maiores abacaxis.
8- Dar com as línguas nos dentes	Revelar informações comprometedoras	R/A/P	2.580.000 (0,25)	14	16	Tempos verbais, alegremente, não dar.
9- Inês é morta	Situação tardia	A	679.000 (0,37)	7	1	Tempo verbal.
10- Como o diabo foge da cruz	Evitar alguém ou situação de qualquer custo	S/A/P	235.000 (0,28)	14	1	Tempos verbal.
11- Ser amigo da onça	Falsidade, traição	R/A/P/CT	310.000 (0,32)	0	20	Tempos verbais, verdadeiro.
12- Procurar sarna para se coçar	Arrumar problemas desnecessários	CT	19.800 (0,28)	2	5	Buscar, caçar, desejar, conveniente, marxista, do terrorismo, mais.
13- Ser um balde de água fria	Desilusão, desapontamento	A/P	144.000 (0,36)	0	92	Tempos verbais, enorme, relativo, autêntico, lançar, constituir, dar um, realmente, cair como, outro balde, merecer, deste, aquele, despejar, jogar, qualquer, verdadeiro, apanhar.
14- Ser pau para toda obra	Fazer um pouco de tudo,	A	2.460.000 (0,28)	2	40	Tempos verbais, considerar, confucionista, uma espécie de,

	ser prestativo					mesmo, esse autêntico.
15- Pendurar as chuteiras	Aposentar-se	A	14.800 (0,23)	11	7	Perdurar, de vez, dependurado.
16- Perder a cabeça	Perder o controle, enlouquecer	R/A/P	2.820.000 (0,26)	75	24	Tempos verbais.
17- Esquentar a cabeça	Preocupação	R	186.000 (0,26)	1	4	Tempos verbais.
18- Abrir as portas	Negócio, oportunidade	R/P	9.970.000 (0,26)	451	392	Tempos verbais, a porta, uma.
19- Ser a gota de água	Saturação, estopim para crise	R/A	862.000 (0,36)	2	58	Tempos verbais, última.
20- Acabar em pizza	Impunidade	R/A	517.000 (0,46)	0	1	Terminar.
21- Ter paciência de Jó	Paciência	R/A	237.000 (0,27)	0	1	Com uma.
22- Levantar um puxão de orelhas	Repreensão, punição	R/S/A/P	740.000 (0,32)	2	61	Tempos verbais, puxões de/na, orelha(s), apanhar, aceitar, adotar, prever, uns, bons, esses, alguns, autêntico.
23- Quebrar o galho	Ajuda, solução temporária	R/A	125.000 (0,23)	3	1	Tempos verbal, um.
24- Custar os olhos da cara	Alto custo financeiro	R/S/A/P	146.000 (0,33)	2	43	Tempos verbais, comer, empenhar, aos consumidores, roubar.
25- Começar com o pé direito	Sorte	R/S/A/P	8.100.000 (0,26)	18	8	Tempos verbais, entrar.
26- Para inglês ver	Aparência, fingimento, engano	R/A/S/P	82.500 (0,26)	115	0	N/A
27- Untar as mãos	Suborno, propina	A/P	85.300 (0,30)	3	1	Tempo verbal.

28- Acertar na mosca	Precisão	R/A	310.000 (0,26)	1	10	Tempos verbais, em cheio.
29- Dourar a pílula	Tentar melhorar a aparência de algo desagradável	R/S/A	1.120.000 (0,29)	28	9	Tempos verbais, um pouco, amarga.
30- Quebrar a cabeça	Persistência, preocupação	R	4.590.000 (0,26)	6	22	Tempos verbais.
31- Apertar o cinto	Dificuldade econômica	R/P	237.000 (0,27)	50	10	Tempos verbais.
32- Pôr os pingos nos is	Esclarecimento, explicação	R/L	133.000 (0,29)	1	0	N/A
33- Pagar o pato	Injustiça	R/S/A/ CP	665.000 (0,23)	1	3	Tempos verbais.
34- Abandonar o barco	Desistência	R/CT	178.000 (0,24)	37	21	Tempos verbais, não abandonar, governamental, de Dean, da solidariedade.
35- Chorar lágrimas de sangue	Sufrimento, desespero	A/S	645.000 (0,30)	3	11	Tempos verbais, nunca, grossas, jorrar.
Total: 35 EIs			43.235.380	856	983	
Legenda dos dicionários:	A- Almeida (2022); CT- Carvalho; Turé (2023); L- Lauand (2022); P- Pereira (2015); R- Riva (2009); S- Silva (1997)					

Quadro 6: EIs selecionadas (Elaboração própria, 2025)

Antes de explicar o quadro 6, cumpre esclarecer que essa é uma pesquisa de cunho quantitativo-qualitativo, assim, as expressões são organizadas sem a preocupação de se classificar quais são as expressões mais ou menos populares e tampouco questiona-se sua aplicabilidade por variedades geográficas ou recortes sociolinguísticos.

Nota-se, no quadro 6, que, dentre as milhares de expressões existentes no português, esse conjunto específico de trinta e cinco expressões nos pareceu quantitativamente interessante aos propósitos da pesquisa, já que as trinta e cinco EIs selecionadas resultaram de um total de 43.235.380

resultados na *web*, em pesquisa realizada na primeira semana de junho de 2024. Já no CRPC encontramos o total de 1.839 resultados, divididos entre 856 resultados na forma consagrada e 983 resultados em formas variadas. Chamamos aqui de forma consagrada a construção V + Np (Verbo + Sintagma Nominal), com o verbo na forma infinitiva. Realiza-se esse tipo de análise pois considera-se a definição de EI conforme Keizer (2016): expressões semifixas compostas por um verbo e um sintagma nominal ou preposicional, em que o significado literal das partes não se relaciona ou não tem significado (tão claro) no significado idiomático.

Entendemos também que há expressões centradas em predicção, como *bater as botas*, e outras centradas no Np, como *amigo da onça*. Ao pensar nessa diferença, nota-se que em todos os casos, a EI foi inserida por um verbo ou fazia parte de um conjunto de sequências verbais e, de acordo com o contexto, o verbo era selecionado para melhor adequação às necessidades do falante.

A exemplo disso, encontramos *ser ou fazer tempestade num copo de água* (uma situação ser um exagero ou alguém fazer de uma situação um exagero), *entrar ou começar com o pé direito* (alguém começar uma situação ou atividade com sorte ou entrar numa situação, atividade ou local com sorte), e outros. Assim, delimitamos um verbo principal, com base nos registros dicionarizados e sua ativação como parte do termo de busca na *web* e no CRPC, e os demais verbos consideram-se como forma variada.

A expressão que gerou mais resultados na *web* foi *abrir as portas* com 9.970.000 resultados, fato que notamos se refletir nos resultados do CRPC, com 451 ocorrências na forma consagrada e 392 resultados em formas variadas, num total de 843 resultados. Conforme vimos em Riva (2009), essa expressão aplica-se aos conceitos de oportunidade, negócio e esperança, ponto que auxilia a compreender seu número elevado de ocorrências.

Ainda que seja uma expressão que ative com certa regularidade a cena literal, realizamos um teste de pesquisa no *Google* sem a palavra-chave “expressão idiomática”, e vimos que o resultado foi muito maior – 132.000.000 ocorrências encontradas em um intervalo de tempo menor – 0,22 segundos. O refino da pesquisa direcionado para as EIs concentrou um resultado menor de ocorrências e demandou um tempo maior.

Já no CRPC, nota-se que, ao pesquisar por *abrir a porta* (no singular), os resultados possuem maior conexão com a cena literal, e na forma plural, os resultados são, em sua maioria, direcionados ao sentido idiomático, resultado que nos interessou aqui. A diferença entre singular e plural também foi identificada em *bater as botas* e *bater a bota*. A versão no plural gerou 1.460.000 resultados em 0,30 segundos, e no singular tivemos 1.310.000 resultados em 0,28 segundos, ambas as versões (singular e plural) levantadas com a palavra-chave “expressão idiomática” inclusa na pesquisa.

Essa diferença também teve reflexo no CRPC, já que a forma consagrada *bater as botas* não teve resultados (com a conjugação em outros tempos verbais, sim, com vinte resultados), e *bater a bota*, que aparenta ter maior conexão com a cena literal, tivemos seis resultados. Vimos no dicionário de Almeida (2022) as duas formas em verbetes diferentes, fato que explica os resultados aparecerem nos *corpora* também nas duas formas. Outra expressão com forte ativação literal tanto na *web* como no CRPC, foi *descascar o abacaxi*, com resultados ligados principalmente ao preparo de receitas e a forma verbal imperativa, com três resultados no CRPC e 229.000 resultados na *web*.

Outro teste que fizemos foi com a expressão *fazer tempestade num copo de água*, que entendemos possuir apenas a cena idiomática como uso comum. Com a inclusão do termo “expressão idiomática” na busca, encontramos 425.000 resultados encontrados em 0,41 segundos e sem o termo, gerou 1.890.000 resultados em 0,38 segundos. Entende-se que as EIs podem estar vinculadas a outros termos de pesquisa sobre sua categoria, como fraseologismos e ditados populares, e assim, fizemos o teste para essa mesma expressão com esses termos a título de teste, e encontramos apenas 170 resultados para *fraseologismo + fazer tempestade num copo de água* e 210.000 resultados para *ditado popular + fazer tempestade num copo de água*.

A EI com menor resultado na *web* foi *pendurar as chuteiras*, com 14.800 resultados encontrados em 0,23 segundos. Tal expressão no CRPC gerou dezoito resultados, sendo um caso de forma variada ligado ao verbo *perdurar* (e não *pendurar*), fato em que notamos que a similaridade sonora teve influência ao analisarmos que o contexto do excerto não tinha relação semântica com o verbo *perdurar* (ter longa duração, permanecer como recordação).

As EIs *chorar as pitangas* e *pôr minhocas na cabeça* não apresentaram resultados no CRPC, mas identificamos números consideráveis na *web*: 27.580 e 266.000 resultados, respectivamente. Conforme apontado em Silva (1997), *chorar as pitangas* é a versão brasileira para a expressão mais comum em Portugal *chorar lágrimas de sangue*, que gerou 645.000 resultados na *web* e 14 resultados no CRPC. A contextualização histórica da colonização portuguesa no Brasil ajuda a entender os resultados: a EI *chorar as pitangas*, pesquisada apenas em páginas brasileiras, gerou um total de 8.100 resultados em 0,20 segundos. Já apenas em páginas portuguesas, gerou 509 resultados em 0,27 segundos, com ambas as pesquisas filtrando resultados apenas em língua portuguesa.

Outras expressões em que identificamos resultados diferentes ao considerarmos as versões brasileira e portuguesa foram *untar as mãos* e *molhar a mão*. *Untar as mãos* gerou 85.300 resultados em 0,30 segundos e *molhar a mão* gerou 364.000 resultados em 0,28 segundos. *Pagar gasosa*, versão africana dessa expressão, encontrada comumente em Angola (Almeida, 2022) gerou cinco

ocorrências. Optamos pela versão mais comum na variedade portuguesa (Almeida, 2022), pois obtivemos quatro ocorrências no CRPC e com *molhar a mão*, nenhuma ocorrência.

No CRPC, encontramos cinco expressões sem registro em sua forma consagrada – *acabar em pizza* (uma ocorrência na forma variada), *ser um balde de água fria* (noventa e duas ocorrências na forma variada), *ser amigo da onça* (vinte ocorrências na forma variada), *bater as botas* (vinte e oito ocorrências em forma variada) e *ter paciência de Jó* (uma ocorrência na forma variada). Encontramos também duas EIs apenas na forma consagrada: *pôr os pingos nos is* (uma ocorrência) e *para inglês ver* (115 ocorrências). As expressões que apresentam maiores resultados de variação são *ser um balde de água fria* (noventa e duas ocorrências) e *engolir sapo* (sessenta e duas ocorrências).

Seguimos agora para o quarto capítulo, com a análise dos excertos, com base na fundamentação metodológica apresentada no capítulo anterior.

CAPÍTULO 4: ANÁLISE DE DADOS

4.1: Análise de dados

Antes de iniciar a análise dos excertos, retomam-se aqui os tipos de EI que são definidas no português: **Tipo 1 (menos produtivas)**: EIs não-motivadas (opacas) e semanticamente não-decomponíveis (e.g. *to kick the bucket, to pop one's clogs*); **Tipo 2**: EIs motivadas (transparentes) e semanticamente não-decomponíveis (e.g. *to hit the ceiling, to grit one's teeth*), **Tipo 3 (mais produtivas)**: EIs motivadas (+ transparentes) e decomponíveis (e.g. *to spill the beans, to grasp the nettle*); **Tipo 4**: EIs *AD-HOC* motivadas e/ou desmotivadas (finalidade específica) não-decomponíveis (e.g. *to bite the bullet/to gnaw the bullet, to open the floodgates/to close the floodgates*) e **Tipo 5**: EIs parcialmente motivadas por metáforas conceituais e decomponíveis (novo tipo aplicado aos casos do português). Reforçamos que os quatro primeiros tipos de EI foram definidos por Keizer (2016, 2018) e o quinto tipo é a proposta de nossa autoria.

A discussão dos dados se organiza conforme a identificação de cada tipo de EI e a subsequente análise cognitiva, morfossintática e dos níveis Interpessoal, Representacional e Morfossintático, seguida de uma proposta de alinhamento para cada tipo, conforme o modelo teórico da GDF.

Cumpramos lembrar que, em alguns casos, dados construídos foram adicionados, e aparecem sublinhados, enquanto os dados encontrados no CRPC e na *web* aparecem sinalizados dentro de parênteses. Para a discussão de dados, são analisados treze excertos, que representam treze análises detalhadas que exemplificam os tipos de EIs entre as quais se dividem as trinta e cinco expressões consideradas nesta investigação.

4.1.1: EIs do Tipo 1

As expressões que são aqui classificadas como parte do **Tipo 1** (menos produtivas) são as EIs não-motivadas (opacas) e semanticamente não-decomponíveis.

Veja em (4-1) o excerto com *tirar o cavalo da chuva*, extraído do CRPC, na variedade europeia. Note que, em todos os excertos analisados, apresenta-se a inserção do excerto com a EI destacada em negrito, seguida da frase destacada isolada e de sua respectiva descrição morfológica, de acordo com a norma padrão do português, com o intuito de facilitar a leitura e definição dos elementos das EIs. Após o excerto, começamos pela análise cognitiva da EI:

(4-1) - Sr. Deputado Nogueira de Brito, no seu tempo, faça favor de continuar, se assim o entender, a sua exposição.

- Eu “**tiro o cavalo da chuva**”. (Portugal NIL Política A137983)

Eu	tiro	o	cavalo	da	chuva.
1 SG	V tirar PRS	DEF ART M	cavalo	PREP. DEF F	chuva.

Conforme visto em seu registro dicionarizado (Almeida, 2022), *tirar o cavalo da chuva* significa desilusão, abandono, desistência a alguma situação. No excerto, podemos notar que o falante utiliza essa EI para manifestar sua desistência de seguir com sua exposição, em um debate político, sobre calamidades públicas que têm afetado o país e que era o tópico discursivo da sessão política.

A interpretação metafórica dessa EI aparenta não ter conexão entre as cenas literal e idiomática, pois não vemos contribuição das partes componentes para a compreensão do seu valor idiomático. Assim, nota-se o quão difícil é enxergar alguma motivação que facilite a leitura dos elementos *tirar o cavalo da chuva* com a leitura alvo – abandono. Os resultados na *web* (mais de 1 milhão de resultados) e sua frequência no CRPC (vinte ocorrências) mostram que sua convencionalidade na comunidade falante se dá por meio da fixação de seus elementos e da frequência.

Conforme Silva (1997), a origem etimológica da expressão conecta de modo superficial as duas cenas: em um tempo em que os cavalos e carroças eram o principal meio de transporte, era comum um visitante amarrar o cavalo na entrada da residência do anfitrião durante a visita e, ao começar a chover, o visitante era obrigado a tirar o cavalo da chuva e preservá-lo em um lugar coberto, significando uma estadia prolongada não-desejada.

Outra explicação etimológica é que, antigamente, o cavalo era utilizado como ferramenta de trabalho nos campos, devido à sua força para arar o solo e executar outras tarefas. Quando começava a chover, os trabalhadores eram forçados a tirar o cavalo da chuva, quando se sentiam vencidos pela intensidade da chuva. Ainda assim, notamos o quanto é difícil para o falante recuperar esse conhecimento e ativá-lo ao utilizar a EI. Na sequência, são apresentados os critérios de variação morfossintática, conforme a organização em tópicos que segue:

- Substituição lexical: Tirar o **pônei/égua** da chuva*/Tirar o cavalo da **tempestade/aguaceiro/vendaval***
- Inserção de pronomes: Tirar o **seu/meu** cavalo da chuva (resultado encontrado no CRPC e *web*)
- Inserção de advérbios: Tirar **agora** o cavalo da chuva (resultado encontrado na *web*)

- Modificação adjetiva: Tirar o cavalo **grande/pequeno/político** da chuva*
- Número: Tirar **dois/três** cavalos da chuva*
- Quantificação: Tirar **algum/nenhum/muitos** cavalo(s) da chuva*
- Definitude: Tirar **o** cavalo da chuva (resultado encontrado na *web* e CRPC)/Tirar **os/um/uns** cavalos da chuva*
- Nominalização: **Retirada** do cavalo da chuva*
- Passivização: O cavalo **foi tirado** da chuva*
- Possibilidade de conversão para modo imperativo: **Tire** o cavalo da chuva (resultado encontrado na *web* e no CRPC)
- Elisão de um termo fixo do predicado: Tirar o cavalo*/Tirar da chuva*
- Inversão da ordem preferida: Tirar da chuva o cavalo*
- Referência anafórica: Tire o cavalo da chuva. Isso vai ter ajudar/ Tire o cavalo da chuva. Ele está atrapalhando sua vida*
- Referência catafórica: Faça isso: tire o seu cavalo da chuva

Como se observa, o comportamento morfossintático é mais restrito, permitindo apenas a variação de pronomes possessivos, modificação adverbial e, em relação à substituição lexical, apenas o diminutivo *cavalinho* foi encontrado no CRPC e na *web*. Assim, opta-se por não o considerar como relevante para caracterizar a EI com possível substituição lexical, já que é um termo derivado da forma consagrada. Já a possibilidade de conversão para modo imperativo é possível, alterando o tipo de Ilocução (de declarativa para imperativa) no NI.

Casos de referências anafórica e catafórica por entidade do discurso previamente introduzida aparentam ser possíveis, acionados por pronomes. Já as referências anafórica e catafórica, por parte da EI, ou por ativação da cena literal, são muito improváveis. Veja em (4-2) o resultado da análise da EI, de acordo com os três níveis da GDF:

(4-2) Eu tiro o cavalo da chuva²⁴.

NI: (M_i: (A_i: [(F_i: DECL (F_i)) (P₁)_S (P₂)_A (C_i: [(T_i)_{FOC} (+id R_i)] (C_i))] (A_i))(M_i))

NR: (p_i: (PRS ep_i: (e_i: (f_i: [(f_j: **tirar o cavalo da chuva** (f_j)) (1x_i)_U] (f_i)) (e_i)) (ep_i)) (p_i))

²⁴ Esclarecemos que, durante a análise das EIs, trabalhamos com sua representação formal nos moldes da GDF tanto nos casos de forma consagrada, como também nos exemplos com variação morfossintática. No caso dos exemplos de forma variada, a representação alfabética dos itens de cada camada (e.g. propriedade: f_i, f_j, f_k, etc.) pode sofrer alterações de ordem em comparação à forma consagrada, tratada como resultado final. Um exemplo é a forma variada vista em (4-2), que contém uma Palavra Nominal (Nw_i: eu (Nw_i)_{Subj} e a forma consagrada, não. Assim, a primeira Palavra Nominal da forma consagrada passa a ser (Nw_i: cavalo (Nw_i)).

NM: (Le_i: (Cl_i: [(Np_i: (Nw_i: eu (Nw_i))_{Subj} (Np_i)) (Vp: (Vw_i: tirar.prs (Vw_i)) (Vp_i)) (Np_j: (Gw_i: o (Gw_i)) (Nw_j: cavalo (Nw_j)) (Np_j)) (Adpp: (Gw_j: de (Gw_j)) (Gw_k: a (Gw_k)) (Addp_i)) (Np_k: (Nw_k: chuva (Nw_k)) (Np_k))](Cl_i)) (Le_i))

No NI, vemos que a EI se comporta como um único Subato Atributivo (T_i: tirar o cavalo da chuva (T_i)) ²⁵com função Foco, pois entendemos que a EI só evoca a propriedade atributiva *desistir* de modo selado. *Cavalo* e *chuva* não são Subatos Referenciais independentes, portanto, não recebem nominalização, definitude (o artigo definido já faz parte da forma consagrada), referência anafórica ou catafórica por parte componente da EI. Em casos do uso de pronomes possessivos *seu/meu*, notamos a possibilidade de adição do operador de identificabilidade (+id).

No NR, a EI é analisada como um Estado-de-Coisas (e_i: tirar o cavalo da chuva (e_i)), com um único predicado verbal com um argumento *Undergoer* (U), i.e alguém desiste/abandona a situação. *Cavalo* e *chuva* não são unidades representacionais separadas, pois não contribuem de modo independente para a significação idiomática, fato que explica a impossibilidade de variação morfossintática de quantificação (*muitos cavalos/chuvas**) e de substituição lexical (*tirar o pônei da tempestade**), flexão de número (*dois cavalos**). Além disso, também não podem ser modificadas por adjetivos (*cavalos fortes*/chuva torrencial**). Já os pronomes possessivos *meu/seu*, no NR, exercem a função de Referência (R_{ef}).

Keizer (2016) também avalia casos de modificação com pronomes possessivos em expressões desmotivadas e não-decomponíveis (e.g. *he popped his ducal clogs*), mas não enxerga essa modificação como relevante para entender a EI como decomponível, e sim, um ajuste do contexto.

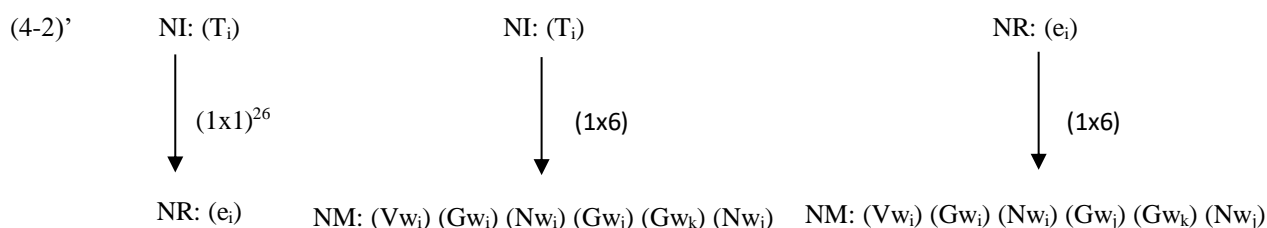
Já no NM, é possível fazer a distinção entre Palavra e Lexema da teoria, na qual a EI é um Lexema complexo codificado com seis Palavras: 1 Palavra Verbal (Vw_i: tirar (Vw_i)), duas Palavras Nominais (Nw_i: cavalo (Nw_i)) e (Nw_j: chuva (Nw_j)) e três Palavras Gramaticais (Gw_i:o (Gw_i)), (Gw_j: de (Gw_j)) e (Gw_k: a (Gw_k)), sendo a preposição *da* dividida em *de + a*). Assim, entendemos essa EI como Tipo 1 – não-motivadas e não-decomponíveis.

Como resultado do alinhamento entre os níveis da forma consagrada, vemos que há transparência apenas entre NI e NR, sendo 1 Subato Atributivo (T_i: tirar o cavalo da chuva (T_i)) e 1 Estado-de-Coisas (e_i: tirar o cavalo da chuva (e_i)). Entre NI e NM, há 1 Subato Atributivo (T_i) e seis

²⁵ Ao citar os elementos de cada nível no corpo do texto de análise das EIs, de acordo com os três níveis mais altos da GDF, optamos por representá-los de modo simplificado, como, por exemplo, o Subato Atributivo de (4-2) (T_i: tirar o cavalo da chuva (T_i)). Sua representação formal completa está registrada no exemplo (4-2) em seu respectivo nível (NI: (C_i: [(T_i)_{FOC} (+id R_i)] (C_i)). Adotamos essa prática (também vista no quadro 8) para facilitar a leitura, focando na descrição da camada mais relevante para cada caso analisado.

unidades no NM: 1 Palavra Verbal (V_{w_i} : tirar (V_{w_i})), duas Palavras Nominais (N_{w_i} : cavalo (N_{w_i})) e (N_{w_j} : chuva (N_{w_j})), e três Palavras Gramaticais (G_{w_i} : o (G_{w_i})), (G_{w_j} : de (G_{w_j})) e (G_{w_k} : a (G_{w_k})).

O mesmo vale para o NR, com 1 Estado-de-coisas (e_i : tirar o cavalo da chuva (e_i)) e seis unidades no NM, o que resulta em opacidade entre esses níveis, conforme esquema de representação (4-2)' abaixo. A segunda parte desse mesmo excerto será retomada como um caso do Tipo 4 –EI AD-HOC mais à frente.



Vejam agora a análise e a discussão da EI *para inglês ver*, encontrada na variedade de Angola. Na sequência segue sua análise cognitiva:

- (4-3) Dói -me como cidadão saber que o dinheiro do herário público fora mal empregue, tão somente porque não houve o devido acompanhamento.
 Se a intenção foi arranjar por arranjar, **para inglês ver**, então estamos conversados.
 Se é o contrário, então o governo provincial deve encarar com seriedade e preocupação a razão da minha insistência no assunto. (Angola NIL J16266 Esp_Leitor)

Para inglês ver.
 PREP. inglês V ver INF.

No excerto observa-se que o falante ativa a significação idiomática da EI– enganação, aparência, fingimento, para descrever seu descontentamento com o mau investimento de dinheiro público para construção e reformas de estradas em Angola. Se a intenção do governo era reformar as estradas apenas para enganar a população, por aparência, fingindo estar fazendo melhorias efetivas, tal objetivo foi alcançado, pois é de conhecimento do falante que o dinheiro investido para tal empreitada foi muito maior do que os resultados mostraram, reforçando que não houve acompanhamento e fiscalização de órgãos competentes de modo efetivo.

Não há uma motivação que facilite a conexão entre as significações literal e idiomática nessa EI, o que reforça seu caráter opaco e não-decomponível, pois, mesmo que tenhamos em mente a origem e o contexto sócio-histórico dessa EI, conforme vimos em Silva (1997), é difícil para o falante

²⁶ Entre parênteses está delimitada a relação numérica entre elementos de cada nível, por exemplo, (1x1): o primeiro número (1) é a representação do nível superior (NI) e o segundo número (1) é a representação do nível inferior (NR).

ver uma contribuição dos elementos da EI refletidos na significação idiomática de enganação, aparência ou fingimento. No levantamento de dados da *web* e do CRPC, vimos que essa EI só apresentou registros em sua forma consagrada (115 ocorrência no CRPC), fato que facilita sua rigidez, conforme critérios de variação morfossintática da EI:

- Substituição lexical: Para **britânico/inglesa** ver*/Para **inglês enxergar/visualizar***
- Inserção de pronomes: Para o **seu/meu/nosso** inglês ver*
- Inserção de advérbios: Para inglês **agora/aqui** ver*
- Modificação adjetiva: Para o inglês **novo/velho** ver*
- Número: Para **dois/três** ingleses verem*
- Quantificação: Para **algum/nenhum/muitos** inglês(s) ver*
- Definitude: Para **o/um/uns** inglês(s) ver*
- Nominalização: **Vista/visão/visualização** do inglês*
- Passivização: **Foi visto** pelo inglês*
- Possibilidade de conversão para o modo imperativo: N/A
- Elisão de um termo fixo do predicado: N/A
- Inversão da ordem preferida: **Para ver inglês/Inglês para ver***
- Referência anafórica: **A reforma foi feita apenas para inglês ver. Isso é comum aqui**
- Referência catafórica: **Fizeram a reforma apenas para isso: somente para inglês ver**

Como podemos ver, praticamente todos os critérios de variação morfossintática são improváveis para essa EI. Com algum esforço, é possível enxergar a possibilidade de referências anafórica e catafórica por entidade do discurso previamente introduzida, acionados por pronomes, mas praticamente impossível por parte componente da EI ou por entidade da cena literal.

Casos de possibilidade de conversão para modo imperativo, elisão de termos e nominalização são difíceis de formular como dado construído. Com isso, considera-se essa expressão como Tipo 1, com alto grau de opacidade também. Veja abaixo a representação da EI nos termos da GDF, com

foco na representação apenas de EI em destaque:

(4-4) Se a intenção foi arranjar por arranjar, **para inglês ver**, então estamos conversados.

NI: (M_i: [(A_i) (A_j: [(F_j: DECL (F_j)) (P₁)_S (P₂)_A (C_j: [(T_j)_{FOC}] (C_j))] (A_j)_{orient} (A_k)] (M_i))

NR: (p_i: (NON-FIN ep_i: (e_i: (f_i: [(f_j: **para inglês ver**_v (f_j)) (p_h) (x_i)_A (x_j)_U] (f_i)) (e_i)) (ep_i)) (p_i))

NM: (Le_i: (Cl_i: [(Addp_i: (Gw_i: **para**_{prep} (Gw_i)) (Addp_i)) (Np_i: (Nw_i: **inglês** (Nw_i)) (Np_i)) (Vp_i: (Vw_i: **ver**_{NON-FIN} (Vw_i)) (Vp_i))] (Cl_i)) (Le_i))

Na representação em (4-4), consideramos a EI como um Subato Atributivo (T_j: para inglês ver (T_i)), pois entendemos que a EI só evoca a propriedade atributiva *enganar* de modo selado. Vimos que esse Subato Atributivo (T_j) é parte do Ato Discursivo (A_j), este com função retórica de Orientação em relação aos outros dois Atos Discursivos (A_i: Se a intenção foi arranjar por arranjar (A_i)) e (A_k: então estamos conversados (A_k)). Esses Atos estão em uma relação de dependência, pois entendemos que a EI tem a função de orientar o que foi descrito em (A_i): *para inglês ver* é uma orientação do falante em relação à intenção do governo em arranjar as rodovias apenas por aparência, para enganar a população.

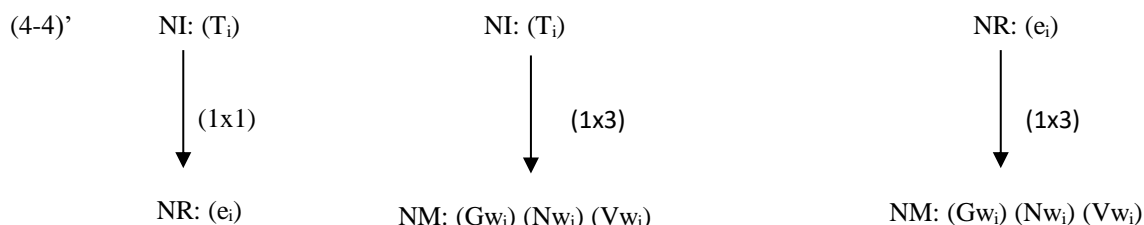
Novamente, temos a EI ativando a função de Foco (situação comum à todas as EIs) no discurso do falante, inserida com a intenção de descrever seu Conteúdo Comunicado de modo mais expressivo. *Inglês* não é Subato Referencial, por isso não recebe nominalização, definitude, referência anafórica ou catafórica por parte componente da EI nem por entidade da cena literal, pois as partes da EI não contribuem de modo independente para a significação idiomática, fato que explica essa restrição na variação morfossintática.

No NR, temos a EI como um único Estado-de-Coisas (e_i: para inglês ver (e_i)), representando o processo de enganar, realizado pelo governo (x_i)_A. Consideramos a população como (x_j) com função *Undergoer* (U), que é enganada pelo governo sobre o processo de uso do dinheiro público para a reforma das estradas. Também levamos em conta o Conteúdo Proposicional (p_h) referente à intenção, já que esse construto mental foi introduzido no Ato Discursivo anterior (A_i: Se a intenção foi arranjar por arranjar (A_i)). Com isso, temos a intenção (p_h) e o processo (e_i) pela qual essa intenção é manifestada entre os participantes (x_i) e (x_j).

Já no NM, a EI é codificada como três Palavras: 1 Palavra Verbal (Pw_i: *ver* (Pw_i)), a Palavra Nominal (Nw_i: *inglês* (Nw_i)) e a Palavra Gramatical (Gw_i: *para* (Gw_i)).

Como resultado do alinhamento entre os níveis, vemos que há transparência apenas entre NI e NR: 1 Subato Atributivo (T_i: para inglês ver (T_i)) e 1 Estado-de-coisas (e_i: para inglês ver (e_i)). Entre NI e NM, temos 1 Subato Atributivo (T_i) no NI e três unidades no NM (1 Palavra Verbal (Vw_i:

ver (Vw_i)), 1 Palavra Nominal (Nw_i : *inglês* (Nw_i)) e 1 Palavra Gramatical (Gw_i : para (Gw_i)), e a mesma relação não-transparente entre NR (1 Estado-de-Coisas (e_i)) e NM (1 Palavra Verbal (Vw_i), 1 Palavra Nominal (Nw_i) e 1 Palavra Gramatical (Gw_i)). Vejamos a representação em (4-4)':



4.1.2: EIs do Tipo 2

Seguimos agora com a análise de dois exemplos de EIs do Tipo 2: motivadas e semanticamente não-decomponíveis, a começar pela EI *abandonar o barco* encontrada na variedade de Moçambique:

- (4-5) Aquando da Assembleia Geral do Banco Austral, em Abril, os accionistas privados demonstraram não serem capazes de injectar fundos para a recapitalização do banco, ou seja, falta de capacidade ou vontade de aumentar o capital social e efectuar suprimentos, pelo que decidiram "**abandonar o barco**" e passar as suas acções ao Estado. (Moçambique NIL Revista R7410)

abandonar	o	barco.
V abandonar NON-FIN	DEF ART M	barco

Como podemos ver em (4-5), a discussão apontada pelo falante é sobre a desistência e a falta de interesse dos acionistas locais em aumentar o capital social do banco e gerar suprimentos para a movimentação da economia. Essa falta de capacidade ou vontade relaciona-se com a significação idiomática vista em *abandonar o barco* – conceito de abandono e desistência (Riva, 2009), com o total de 178.000 resultados na *web* e cinquenta e oito ocorrências no CRPC.

Entende-se essa EI como motivada, justamente pela ativação da significação literal do verbo *abandonar*. Já o termo *barco*, pode ser compreendido como a situação abandonada, mesmo que essa conexão não fique tão claramente relacionada, tanto na leitura literal – o abandono de um barco poderia gerar interpretações como suicídio (em alto mar), impossibilidade de movimentação (caso esteja quebrado), como na cena idiomática – o termo *barco* não carrega em si a ideia de um fardo, um problema, uma dificuldade ou um desafio a ser abandonado.

Por tal interpretação metafórica, interpretamos essa EI como **Tipo 2** – motivada pela ativação semântica literal do verbo *abandonar* e não-decomponível, pela dificuldade em relacionar o termo *barco* como parte contribuinte para a composição da cena idiomática. Veja a análise morfossintática da EI com a aplicação dos critérios de variação morfossintática:

- Substituição lexical: abandonar o **navio** (resultado encontrado no CRPC e na *web*)
- Inserção de pronomes: abandonar o **seu/meu** barco (resultado encontrado no CRPC e *web*)
- Inserção de advérbios: Abandone **agora/imediatamente** o barco
- Modificação adjetiva: Abandonar o barco **governamental** (resultado encontrado no CRPC)
- Número: Abandonar **dois/três** barcos*
- Quantificação: Abandonar **algum/nenhum/muitos** barcos*
- Definitude: Abandonar **o** barco (resultado encontrado na *web* e CRPC)/Abandonar **os/um/uns** barcos*
- Nominalização: **Abandono do seu barco** (resultado encontrado no CRPC)
- Passivização: O barco **foi abandonado**
- Possibilidade de conversão para o modo imperativo: **Abandone** o barco (resultado encontrado na *web* e no CRPC)
- Elisão de um termo fixo do predicado: **N/A**
- Inversão da ordem preferida: **O barco abandonar***
- Referência anafórica: Como Diógenes, não precisamos de muito para nos sentirmos felizes, pelo que não serão as manobras de Sybila as culpadas de abandonarmos **este barco** a curto prazo (resultado encontrado no CRPC com ampliação do excerto)
- Referência catafórica: A menos que, face ao avolumar das dificuldades, queira repetir, agora, **o que fez há 15 anos: abandonar o barco** (resultado encontrado no CRPC)

Como se observa, esse tipo de EI apresenta um carácter menos restrito em relação ao Tipo 1, permitindo substituição lexical, modificação adjetiva, adverbial, uso de pronomes possessivos,

nominalização, possibilidade de conversão para o modo imperativo, passivização e referências anafórica, tanto referência por entidade previamente introduzida no discurso, como por parte da EI e referência catafórica por entidade posteriormente introduzida no discurso. Observa-se também que o artigo definido "o" faz parte da forma consagrada. Vejamos a representação formal da expressão em sua forma consagrada nos moldes da GDF:

(4-6) **Abandonar o barco.**

NI: (M_i: (A_i: [(F_i: DECL (F_i) (P₁)_S (P₂)_A (C_i: [(T_i) (+id R_i: (T_j) (R_i))_{FOC}] (C_i))](A_i)) (M_i))

NR: (p_i: (NON-FIN ep_i: (e_i: (f_i: [(f_j: [(f_k: abandonar_V (f_k) (1x_i: barco_N (x_i))_U] (f_j)) (x_j)_A] (f_i))(e_i)) (ep_i))(p_i))

NM: (L_e: (C_l: [(Vp_i: (Vw_i: abandonar_{NON-FIN} (Vw_i)) (Vp_i)) (Np_i: [(Gw_i: o (Gw_i)) (Nw_i: barco (Nw_i))] (Np_i))] (C_l)) (L_e))

No Nível Interpessoal, *abandonar o barco* é representado com um Subato Atributivo (T_i: abandonar (T_i)), um Subato Referencial (R_i: barco (R_i)), e um operador de identificabilidade (+ id: o), já que o artigo definido (o) faz parte da forma consagrada. Consideramos a unidade *barco* como um Subato Referencial separado, pois notamos a marcação de definitude e a possibilidade de inserção de substituição lexical.

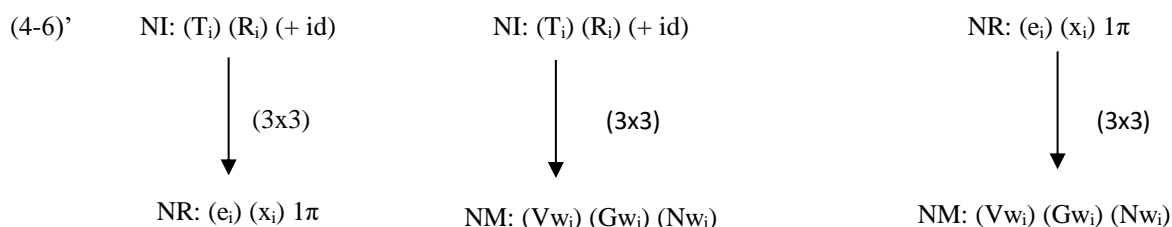
No NR, *barco* é uma unidade separada (x_i: barco (x_i)) funcionando como argumento *Undergoer* (U) do predicado verbal (f_k: abandonar (f_k)). Já o Estado-de-Coisas (e_i: *abandonar o barco* (e_i)) funciona como um predicado complexo de dois lugares (f_i), com um argumento *Actor* (x_j), que, no caso do excerto (4-5), poderíamos defini-lo como *os acionistas* - os accionistas privados que decidiram abandonar o barco - com a seguinte representação: (m x_j: acionista (x_j)), considerando o operador m para o plural.

O argumento (x_i: barco (x_i)) pode ser modificado por adjetivos que contribuam para o significado da expressão baseada no contexto extralinguístico, como *governamental*, e a posição do operador de (x_i) é selada, pois esse operador de número (1) é um recurso invariável e reforça a impossibilidade de quantificação e número.

No NM, temos a representação da EI com um Sintagma Verbal (Vp_i), composta de uma Palavra Verbal (Vw_i: abandonar (Vw_i)), uma Palavra Gramatical (Gw_i: o (Gw_i)) e uma Palavra Nominal (Nw_i: barco (Nw_i)).

Assim, vemos em (4-6)' que o alinhamento entre os níveis resultou em transparência entre todos os níveis: 1 Subato Atributivo (T_i: abandonar (T_i)), 1 Subato Referencial (R_i: barco (R_i)) e 1 operador de identificabilidade (+id: o) no NI (três unidades lexicais); 1 Estado-de-Coisas delimitado

pela propriedade f para o predicado verbal (f_k : abandonar_v (f_k)), 1 Indivíduo (x_i : barco (x_i)) e 1 operador de quantidade lexicalmente marcado (1 x_i) no NR (três unidades lexicais); e 1 Palavra Verbal (V_{w_i} : abandonar (V_{w_i})), 1 Palavra Nominal (N_{w_i} : barco (N_{w_i})) e 1 Palavra Gramatical (G_{w_i} : o (G_{w_i})) no NM (três unidades lexicais).



Seguimos para a EI *levar um puxão de orelhas*, na variedade de Cabo Verde:

- (4-7) Pelo contrário, entendeu que com os artigos o director mais não fez do que o exercício abusivo da liberdade de imprensa, o que não é tolerável num Estado de Direito Democrático. Após a leitura da sentença, a juíza aproveitou a oportunidade para **dar um autêntico puxão de orelha** ao R, por causa da sua já conhecida falta ao acto da leitura da sentença que estava marcada para o passado dia 10. (Cabo Verde NIL Jornal J10694)

Dar	um	autêntico	puxão	de	orelhas
V dar NON-FIN	INDEF ART	A M	puxão	PREP	orelha PL

Para análise cognitiva desse excerto, seleccionamos a EI em sua forma variada, pois enxergamos como uma oportunidade de tratarmos suas variações em um caso real. Como vimos no quadro 6, a EI *levar um puxão de orelhas* apresentou sessenta e um resultados em casos de variação no CRPC e 740.000 resultados na *web*, principalmente com variação verbal, conforme (4-7), com o verbo *dar*, adequada ao contexto de uso, onde notamos que a juíza realiza a ação de aplicar uma punição, repreensão ao indivíduo descrito como R, entendido no contexto como réu durante uma sessão de julgamento no Supremo Tribunal de Justiça em Cabo Verde.

Enxergamos que a motivação da EI baseia-se na relação entre a possível dor física de um puxão de orelha, algo que pode ser interpretado subjetivamente e não há graus de intensidade dessa ação, embora, conforme vimos em Silva (1997), tal prática iniciou-se com a intenção de ser uma punição torturante e dolorosa.

Porém, notamos que é difícil entender de modo claro a contribuição das partes componentes para a significação idiomática, pois, além de significar punição, castigo, repreensão, há vários casos em que a EI representa algo mais brando, como uma chamada de atenção, uma orientação mais rígida, assim como a ação literal de levar ou aplicar um puxão nas orelhas também não dá conta desses graus

de dor física e de seu resultado como punição, tampouco o porquê de sua aplicação nas orelhas e não em outra parte do corpo, o que também poderia resultar em dor. Observe-se a análise morfossintática da EI:

- Substituição lexical: **Dar/apanhar/aceitar/adotar/prever** um puxão de orelhas (resultados encontrados no CRPC)/ Levar um puxão no **nariz/dedo/cabelo***
- Inserção de pronomes: Levar **esse** puxão de orelhas (resultado encontrado no CRPC)
- Inserção de advérbios: Levar **agora/hoje** um puxão de orelhas
- Modificação adjetiva: Dar um **autêntico/bons** puxão(ões) de orelha (resultados encontrado no CRPC)
- Número: Levar **dois** puxões de orelhas
- Quantificação: Levar **alguns** puxões de orelhas (resultado encontrado no CRPC)
- Definitude: Levar **o/uns** puxão(ões) de orelhas (resultados encontrado na *web* e CRPC)
- Nominalização: **Levada de puxão de orelhas**
- Passivização: **O puxão de orelhas foi dado/levado**
- Possibilidade de conversão para o modo imperativo: **Dê um puxão de orelhas**
- Elisão de um termo fixo do predicado: **Levar um puxão/Levar nas orelhas***
- Inversão da ordem preferida: **Nas orelhas levar um puxão/Levar nas orelhas um puxão***
- Referência anafórica: O Orador: - Cada dia de mera contemplação passiva da realidade traduz -se, inevitavelmente, no futuro, em mais um “puxão de orelhas” de Bruxelas. Só que esses “**puxões de orelhas**” não são o mais grave. (resultado encontrado no CRPC)
- Referência catafórica: **Você está merecendo isso: levar um puxão de orelhas!**

Como podemos ver na análise morfossintática, essa EI permite um número maior de variações, embora entendamos que a substituição lexical esteja fortemente ligada ao tipo de construção sintática da EI – centrada em Np, com a abertura de substituição do verbo para adaptação ao contexto, principalmente por ser uma ação literal que pode ser realizada por um agente ou ser

recebida por um paciente, alguém pune, e alguém é punido pelo puxão de orelhas.

Essa ruptura se estende para a possibilidade de conversão para o modo imperativo, entendida como possível apenas para o agente (e.g. *dê um puxão de orelhas no João*). Tal substituição lexical não se aplica à orelha ou a puxão.

Outra variação que notamos está na construção singular e plural *de orelha*, fato que nos fez adicionar como dado construído *levar dois puxões de orelha(s)*, pois entendemos ser possível a interpretação de quantidade em um ou dois puxões, relacionada à quantidade de orelhas que temos. Outra possibilidade que pensamos sobre número e quantificação é seu uso relacionado à quantidade de sermões aplicados/sofridos.

Como resultado, entendemos a EI como Tipo 2 – motivada e não-decomponível, com uma menor restrição sintática, como em casos de inversão da ordem preferida e elisão de um termo fixo do predicado. Já os casos de referência anafórica e catafórica são possíveis pela entidade do discurso previamente/posteriormente introduzida, e a referência anafórica também é possível por parte da EI.

Observe-se o resultado da análise da EI de acordo com os níveis e camadas da GDF:

(4-8) **Dar um autêntico puxão de orelhas**

NI: (M_i: (A_i: [(F_i: DECL (F_i) (P₁)_S (P₂)_A (C_i: [(T_i) (- id + s R_i: (T_j) (T_k) (R_i) (R_j: (T_i) (R_j))_{FOC} (C_i)](A_i)) (M_i))

NR: (p_i: (NON-FIN ep_i: (e_i: (f_i: [(f_j: [(f_k: dar_V (f_k) (1 e_j: (f_i: puxão de_{ADP} (f_i) (m x_i: orelha_N (x_i))_{Ref} (e_j)): (f_m: autêntico_A (f_m))] (f_j) (x_j)_A] (f_i) (e_i) (ep_i) (p_i))

NM: (L_e: (C_i: [(Vp_i: (Vw_i: dar_{NON-FIN} (Vw_i) (Vp_i)) (Np_i: [(Gw_i: um (Gw_i) (Adp_i: (Adw_i: autêntico (Adw_i) (Adp_i)) (Nw_i: puxão (Nw_i) (Np_i)) (Adpp_i: (Gw_j: de (Gw_j) (Adpp_i)) (Np_j: (Nw_j: orelha_{pl} (Nw_j) (Np_j))] (C_i) (L_e))

No NI, temos a EI representada em um Subato Atributivo (T_i: dar (T_i)), e dois Subatos Referenciais: (R_i) composto pelo operador de identificabilidade *um* (- id), dois Subatos Atributivos (T_j: puxão (T_j)) e (T_k: autêntico (T_k)), e (R_j: *orelhas* (R_j)), com função de Foco na EI como um todo. Também adicionamos o operador de especificidade (+ s) referente ao adjetivo *autêntico*, o que significa que é um *puxão de orelhas* especificado pela propriedade *autêntico*.

Caso considerássemos apenas a forma consagrada, teríamos um Subato Atributivo (T_i: levar (T_i)), dois Subatos Referenciais ((R_i: (T_j: puxão (T_j) (R_i)) e ((R_j: (T_k: orelhas (T_k) (R_j)) e 1 operador de identificabilidade (-id), lexicalmente marcado – *um*, entendido como unidade fixa da forma consagrada.

No NR, o adjetivo *autêntico* modifica o Estado-de-Coisas (e_i): dar um puxão de orelhas (e_i)),

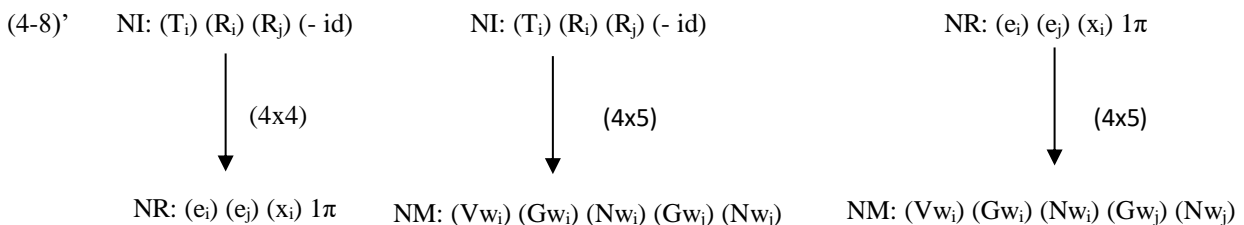
referente a EI como um todo, que pela significação do verbo (f_k : *dar* (f_k)), ativa a função semântica de Indivíduo *Actor* em (x_j), i.e alguém que aplica um sermão. Caso fosse a forma consagrada *Levar um puxão de orelhas*, o Indivíduo (x_j) seria caracterizado pela função *Undergoer* (v), i.e. alguém que recebe um sermão. Pensando na significação metafórica da EI como punição, entendemos *puxão* como um outro Estado-de-Coisas (e_j : *puxão* (e_j)) que ativa o operador de quantidade 1, pois vimos que há a possibilidade de quantificação e número, assim como em *orelhas* ($m x_i$: *orelha*_N (x_i))_{Ref}, esta com função de Referência, referente ao tipo de puxão em uma construção adposicional.

Tal percepção de um Estado-de-Coisas (e_j : *puxão* (e_j)) dentro de outro Estado-de-Coisas (e_i): *dar um puxão de orelhas* (e_i) para essa EI justifica-se pelo seu tipo - centrada em Np, e esse *slot* para diferentes verbos (*levar*, *dar*, *apanhar*, etc.) é representado no NR como um Estado-de-Coisas.

Na codificação morfossintática, temos uma Oração (Cl_i) composta por 1 Palavra Verbal (Vw_i : *dar* (Vw_i)), duas Palavras Nominais (Nw_i : *puxão* (Nw_i)) e (Nw_j : *orelha*_{pl} (Nw_j)), 1 Palavra Adjetival (Adw_i : *autêntico* (Adw_i)) e duas Palavras Gramaticais (Gw_i : *um* (Gw_i)) e (Gw_j : *de* (Gw_j)).

Considerando a forma consagrada *Levar um puxão de orelhas* para o alinhamento entre os níveis, temos 1 Subato Atributivo (T_i : *levar* (T_i)), dois Subatos Referenciais (R_i : *puxão de* (R_i)) e (R_j : *orelhas* (R_j)), e 1 operador de identificabilidade (-*id um*) no NI; dois Estados-de Coisas – sendo o primeiro (e_i) delimitado pela propriedade *f* para o predicado verbal (f_k : *levar*_v (f_k)) e o segundo (e_j : *puxão de*_{ADP} (e_j)), 1 Indivíduo ($m x_i$: *orelha* (x_i)), 1 operador de quantidade (1 e_j) no NR; e 1 Palavra Verbal (Vw_i : *levar* (Vw_i)), duas Palavras Nominais (Nw_i : *puxão* (Nw_i)) e (Nw_j : *orelha*_{pl} (Nw_j)), e duas Palavras Gramaticais (Gw_i : *um* (Gw_i)) e (Gw_j : *de* (Gw_j)) no NM.

Como resultado, há uma relação de transparência apenas entre NI e NR (quatro unidades). O NM apresentou cinco unidades, o que causa opacidade entre NM e os dois níveis da Formulação, conforme em (4-8)':



Apesar de operadores e funções ativarem estratégias gramaticais, considera-se sua representação lexical nos NI e NR como parte da contagem das unidades alinhadas em cada nível para fins de resultados de transparência e opacidade. A mesma contagem não foi considerada com a função *Foco* no NI, por não haver essa representação de um item lexical individualmente marcado no

NM. Em relação a EI anterior - *abandonar o barco*, notamos que no NM, há casos de diferenças na quantidade de itens lexicais, de acordo com a extensão da EI, o que resulta em diferentes relações de transparência entre Formulação e Codificação.

4.1.3: EIs do Tipo 3

A seguir, em (4-9), temos a EI *descascar o abacaxi*, encontrada na variedade brasileira e que defendemos pertencer ao Tipo 3 – motivadas e decomponíveis:

(4-9) E viu a salvação: " O que temos de fazer é exigir que o governo funcione, fazer um programa de apoio ao Estado para que a polícia funcione não só na defesa dos ricos mas também separando os pobres dos bandidos". MEDIDAS SIMPLES - Parece conversa de catecúmenos. Mas o prefeito eleito do Rio, César Maia, a quem caberá daqui a duas semanas **descascar o abacaxi carioca**, anda com idéias semelhantes. (Brasil NIL Revista COD1000069)

Descascar	o	abacaxi	carioca.
V descascar NON-FIN	DEF ART M	abacaxi	A

Como se observa no excerto, o redator da revista relata sobre problemas sociais que o novo candidato eleito, César Maia, terá pela frente em seu mandato como prefeito do Rio de Janeiro, e um desses problemas é a segurança pública. Encontramos três ocorrências no CRPC e 17.200 resultados na *web* da expressão como idiomática. Conforme vimos no dicionário de Silva (1997), *descascar o abacaxi* significa resolver um problema, enfrentar dificuldades, e em (4-9), vemos facilmente essa significação idiomática inserida no contexto político abordado na notícia.

Entendemos que há uma clara motivação que conecte as cenas literal e idiomática da EI, devido à dificuldade em manusear o fruto, devido ao seu cheiro forte, acidez e ao aspecto espinhoso, assim como é possível ver com clareza a contribuição das partes componentes para a ativação semântica da significação idiomática: *descascar* como resolver, enfrentar, e *abacaxi* como o problema ou situação adversa, desconfortável. Nota-se um alto grau de semelhança com a EI do inglês *to grasp the nettle* vista em Langlotz (2006) e Keizer (2016), também entendida como motivada e decomponível. Seguimos com a aplicação dos critérios de variação morfossintática:

- Substituição lexical: **Descascar o ananás/limão/abacate***
- Inserção de pronomes: Descascar o **seu** abacaxi
- Inserção de advérbios: Descascar **agora/depressa** esse abacaxi

- Modificação adjetiva: Descascar o abacaxi **carioca/um dos maiores abacaxis** (resultados encontrados no CRPC)
- Número: Descascar **dois** abacaxis
- Quantificação: Descascar **alguns/muitos** abacaxis (resultado encontrado na *web*)
- Definitude: Descascar **o/um/uns** abacaxi(s) (resultados encontrado na *web* e CRPC)
- Nominalização: **Descasque/Descascamento** de abacaxi*
- Passivização: **O abacaxi foi descascado**
- Possibilidade de conversão para o modo imperativo: **Descasque** o abacaxi (resultado encontrado na *web*)
- Elisão de um termo fixo do predicado: N/A
- Inversão da ordem preferida: **Um abacaxi descascar***
- Referência anafórica: É necessário **descascar esse abacaxi** da segurança pública. E **esse abacaxi** não é novidade para ninguém.
- Referência catafórica: Na sessão de hoje faremos **isso imediatamente: descascar o abacaxi!**

Como se observa na análise, o comportamento morfossintático é mais flexível, permitindo casos de referência anafórica tanto por parte da EI como por entidade do discurso previamente introduzida e referência catafórica apenas por entidade do discurso posteriormente introduzida, acionada por pronomes e modificação adjetiva relacionada ao contexto do discurso e à entidade da cena literal: quantificação, número, definitude, possibilidade de conversão para o modo imperativo, passivização, inserção de pronomes possessivos e advérbios.

Como resultado, definimos *descascar o abacaxi* como EI motivada e decomponível. Prosseguimos com a representação formal da GDF, considerando a modificação adjetiva do excerto:

(4-10) **Descascar o abacaxi carioca**

NI: (M_i: (A_i: [(F_i: DECL (F_i) (P₁)_S (P₂)_A (C_i: [(T_i) (+ id + s R_i: (T_j) (T_k) (R_i))_{FOC}] (C_i))](A_i)) (M_i))

NR: (p_i: (NON-FIN ep_i: (e_i: (f_i: [(f_j: [(f_k: **descascar**_V (f_k)) (1e_j: (f_i: **abacaxi**_N (f_i)) (e_j))_U: (f_m: **carioca**_A (f_m))] (f_j)) (x_i)_A] (f_i)) (e_i)) (ep_i)) (p_i))

NM: (Le_i: (Cl_i: [(Vp_i: (Vw_i: descascar_{NON-FIN} (Vw_i)) (Vp_i)) (Np_i: [(Gw_i: o (Gw_i)) (Nw_i: abacaxi (Nw_i)) (Adp_i: (Adw_i: carioca (Adw_i)) (Adp_i))] (Np_i))] (Cl_i)) (Le_i))

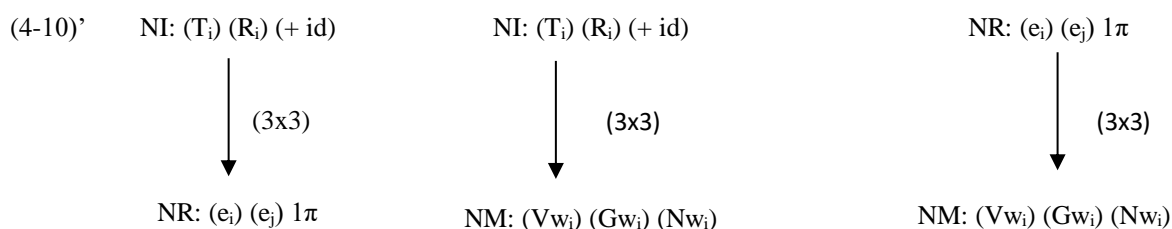
No Nível Interpessoal, a EI é representada por 1 Subato Atributivo (T_i: descascar (T_i)) e 1 Subato Referencial (R_i: abacaxi (R_i)) com os operadores (+ id) e (+ s) representando, respectivamente, o artigo definido *o* e o adjetivo *carioca*. Com a forma consagrada, teríamos apenas 1 Subato Atributivo (T_i: descascar (T_i)), 1 operador de identificabilidade (+ id: o) e 1 Subato Referencial (R_i: abacaxi (R_i)).

No NR, entende-se a expressão completa como 1 Estado-de-Coisas (e_i: descascar o abacaxi (e_i)) que possui um outro Estado-de-Coisas (e_j: abacaxi (e_j)) ativado pelo núcleo *abacaxi*, que, em sua base metafórica, representa *problema*.

Veja que (e_j: abacaxi (e_j)) possui o operador de quantidade 1 (1e_j: (f_i: abacaxi_N (f_i)) (e_j))_U, pois vimos nos critérios para classificação formal que a unidade *abacaxi* pode ser quantificada. Já o adjetivo *carioca* modifica (e_j: abacaxi (e_j)). Considerando a existência de um *Actor* (x_i), ativamos a camada (f_i) em (f_i: [(f_j: [(f_k: descascar_V (f_k)) (1e_j: (f_i: abacaxi_N (f_i)) (e_j))_U: (f_m: carioca_A (f_m))] (f_j)) (x_i)_A] (f_i)) para tal relação com a EI e seu verbo *descascar*, considerando assim que alguém (x_i) descasca o abacaxi na cena literal e alguém resolve o problema em sua significação idiomática.

No NM, a EI, em sua forma variada, é representada por 1 Palavra Verbal (Vw_i: descascar_{NON-FIN} (Vw_i)), 1 Palavra Gramatical (Gw_i: o (Gw_i)), 1 Palavra Nominal (Nw_i: abacaxi (Nw_i)) e 1 Palavra Adjetival (Adw_i: carioca (Adw_i)). Dentro da Oração (Cl_i) temos diferentes tipos de sintagmas: verbal (Vp_i), ativado pelo verbo *descascar*, nominal (Np_i), ativado pelo substantivo *abacaxi* e adjetival (Adp_i), ativado pelo adjetivo *carioca*.

Como resultados de alinhamento entre os três níveis da teoria para a forma consagrada *descascar o abacaxi*, temos relação de transparência entre todos os níveis, com três unidades: 1 Subato Atributivo (T_i: descascar (T_i)), 1 Subato Referencial (R_i: abacaxi (R_i)), e 1 operador de identificabilidade (+ id: o), lexicalmente marcado no NI. No NR temos 1 operador de quantidade lexicalmente marcado (1e_j: (f_i: abacaxi_N (f_i)) (e_j))_U e representado pelo número 1, devido à possibilidade de quantificação e número, e 2 Estados-de-Coisas, sendo o primeiro (e_i) delimitado pela propriedade f para o predicado verbal (f_k: descascar_V (f_k)) e o segundo (e_j: abacaxi (e_j)), codificados no NM como 1 Palavra Verbal (Vw_i: descascar_{NON-FIN} (Vw_i)), 1 Palavra Gramatical (Gw_i: o (Gw_i)) e 1 Palavra Nominal (Nw_i: abacaxi (Nw_i)), de acordo com a representação em (4-10)':



Seguimos agora para a análise cognitiva do segundo exemplo de EI do tipo 3 – motivada e decomponível, com a EI *procurar sarna para se coçar*, encontrada na variedade de Macau, em um excerto de revista literária:

- (4-11) Longe da vista, longe do coração
 Buscá sarna pa cuçá.
Buscar sarna para se coçar. (Macau 1999 Revista Provérbios R0457)

Buscar	sarna	para	se	coçar
V buscar NON-FIN	sarna	PREP	PRON REFL	V coçar NON-FIN

No excerto completo, observa-se que o texto se trata de um texto literário que reúne diversas EIs e provérbios, e tenta fazer uma conexão de forma poética. No trecho destacado, observa-se a EI, em sua forma variada, *buscar sarna para se coçar*, é repetida duas vezes: a primeira com uma escrita coloquial (*busca sarna pa cuçá*) e a segunda, escrita na forma culta (*buscar sarna para se coçar*). No entendimento da significação idiomática da EI, vimos que o autor fala de uma personagem que estava em um lugar calmo e tranquilo e buscou problemas para perder o sossego. Encontramos 19.800 resultados na *web* e cinco casos de forma variada no CRPC.

Nota-se que a motivação da EI é ativada semanticamente pelo termo *sarna*, conhecida também como escabiose, doença causada pelo parasita *Sarcoptes scabie*. A doença atinge animais silvestres e domésticos, como gatos, cachorros e aves. Os sintomas são coceira, irritação na pele, febre e inflamações cutâneas. Essa motivação semântica ativada pela doença, também dialoga culturalmente com contextos históricos em que homens sofriam com doenças transmitidas por animais, como o período da Peste Negra (1346-1353) e da Gripe suína (2009-2011).

Entendemos essa ativação semântica como motivada pelo fato de que é claramente possível enxergar a significação idiomática de procurar, buscar ou caçar problemas desnecessários para si, ou seja, o entendimento de *sarna* como problema é a primeira conexão visível, e a ideia de problema desnecessário, ou do sujeito procurar por esse problema é ativada pelos verbos *procurar*, *buscar* ou *caçar*, *desejar*, já que, considerando a EI centrada em Np, os verbos encontrados refletem a ideia do sujeito procurar por problemas, o que os tornam desnecessários.

Podemos decompor essa EI e, assim, compreender a ação de *se coçar* como sofrer as causas do problema – *sarna*, e as ações de *buscar* e *procurar* que tornam esse problema desnecessário. Seguimos para a análise morfossintática da EI:

- Substituição lexical: **Procurar/buscar/caçar/desejar** sarna para se coçar (resultados encontrados no CRPC e na *web*)
 - Inserção de pronomes: Procurar essa/sua sarna para se coçar
 - Inserção de advérbios: Procurar agora/(para) lá sarna para se coçar
 - Modificação adjetiva: Procurar sarna **conveniente/marxista/do terrorismo** para se coçar (resultados encontrados no CRPC)
 - Número: Procurar duas sarnas para se coçar*
 - Quantificação: Procurar **mais/muita** sarna para se coçar (resultados encontrado no CRPC e na *web*)
 - Definitude: Procurar a/uma sarna para se coçar
 - Nominalização: Procura/Busca de/por sarna para se coçar
 - Passivização: A sarna para se coçar foi procurada/A sarna (procurada) foi coçada*
 - Possibilidade de conversão para o modo imperativo: **Procure** sarna para se coçar (resultado encontrado na *web*)
 - Elisão de um termo fixo do predicado: sarna para se coçar/procurar sarna (resultado encontrado na *web*)
 - Inversão da ordem preferida: Para se coçar, procure/vá procurar sarna
 - Referência anafórica: Você procurou muita sarna para se coçar ao comprar esse carro velho. Agora, essa sarna vai te custar muito dinheiro e dor de cabeça.
 - Referência catafórica:
- (4-12) E ambas as forças, em vez de puxarem para o mesmo lado, passaram um mês a mandar recados pelos jornais. São estes tristes exemplos que desejamos para o nosso País?
Para se coçar, o Ministério da Justiça já tem os tribunais que ainda não conseguem julgar quantos processos lhes são remetidos, para esse efeito, pela DNIC.
 Ainda assim deseja **mais sarna** com a tutela desta polícia criminal e da polícia económica?

(Angola 2012 Ponto de Vista J162161)

Começamos com a análise da referência catafórica, vista em (4-12). No excerto extraído da variedade de Angola, vemos que a EI apresenta um caso de **descontinuidade** (além de seu acionamento em inversão da ordem preferida), já que há uma divisão da unidade formal dos componentes da EI, resultando em um novo alinhamento não-paralelo entre diferentes Movimentos (parágrafos). No primeiro Movimento (M_i), o falante utiliza as partes componentes referentes ao Sintagma Verbal da EI – *Para se coçar*, e no próximo Movimento (M_j), o falante completa a parte componente faltante da expressão de forma variada – *deseja mais sarna*. Após a introdução do segundo Movimento (M_j), entendemos que no parágrafo anterior – *para se coçar* refere-se à EI que apareceu de modo descontínuo, resultando na ativação da camada do Discurso (D_i), que será comentada mais adiante.

Com isso, analisamos que *os tribunais que ainda não conseguem julgar quantos processos lhes são remetidos, para esse efeito, pela DNIC é a construção literal que representa sarna* - os problemas que o Ministério da Justiça já tem, e *a tutela desta polícia criminal e da polícia económica representa mais problemas, i.e. mais sarna*.

A complementação da EI no segundo parágrafo é importante para entendimento completo da EI e, ao mesmo tempo, ajuda o leitor a diferenciar outras possíveis significações conotativas ativadas pelo verbo *coçar*, como inquietação (e.g. *coçar a cabeça* – 137.000 resultados na *web*), vontade de gastar dinheiro (e.g. *coçar a mão* - 158.000 resultados na *web*), vadiagem (e.g. *coçar o saco* – 122.000 resultados na *web*) ou desconfiança (e.g. *coçar o chifre* – 78.800 resultados na *web*) e ainda outras que podem ser específicas em Angola.

Na forma variada, observa-se a possibilidade de substituição lexical com o verbo *desejar*, que relaciona-se semanticamente com *procurar*, *caçar*, *buscar* por *sarna*. A EI apresentou maior flexibilidade quanto a variações, como modificação adjetiva, adverbial, uso de pronomes possessivos, quantificação, definitude, elisão de um termo fixo do predicado, nominalização, inversão da ordem preferida, possibilidade de conversão para modo imperativo, e referências anafórica por parte componente da EI e catafórica, de modo descontínuo, por parte componente da EI.

Apenas passivização e número não parecem possíveis, pois *sarna* não parece ser um termo contável (e.g. *Esse cachorro já pegou duas sarnas*/Esse cachorro já pegou sarna duas vezes/Esse cachorro já pegou dois tipos diferentes de sarna*). Como resultado, entendemos a EI *procurar sarna para se coçar* como motivada e decomponível. Seguimos para a representação formal da GDF com a forma variada vista em (4-11) e com comentários do excerto da variedade em referência catafórica

vista em (4-12):

(4-13) **Buscar sarna para se coçar.**

NI: (M_i: (A_i: [(F_i: DECL (F_i) (P₁)_S (P₂)_A (C_i: [(T_i) (T_j) (-s R_i: (T_k) (R_i)) (R_j: (T_i) (R_j))_{FOC} (C_i)](A_i)) (M_i))

NR: (p_i: (NON-FIN ep_i: (e_i: (f_i: [(f_j: [(f_k: buscar_V (f_k)) (e_j: (f_i:sarna_N (f_i)) (e_j))_U (e_k: (f_m: coçar_V (f_m)) (e_k))_{REC} (f_j)) (x_i)_A] (f_i)) (e_i)) (ep_i)) (p_i))

NM: (Le_i: (Cl_i: [(Vp_i: (Vw_i: buscar_{NON-FIN} (Vw_i)) (Vp_i)) (Np_i: (Nw_i: sarna (Nw_i))_{Obj} (Np_i)) (Adpp_i: (Gw_i: para (Gw_i)) (Adpp_i)) (Np_j: (Gw_j: se (Gw_j))_{Subj} (Np_j)) (Vp_j: (Vw_j: coçar_{NON-FIN} (Vw_j)) (Vp_j))] (Cl_i)) (Le_i))

No Nível Interpessoal, a EI é representada por dois Subatos Atributivos: (T_i: buscar (T_i)) e (T_j: coçar (T_j)), e dois Subatos Referenciais: (R_i) representando o pronome *se* – sujeito no qual a ação do verbo é reflexiva e marcado também com o operador de especificidade (– s), já que a forma variada não tem um sujeito especificado, e (R_j: sarna (R_j)), com função de Foco na EI como um todo.

No NR, entende-se a expressão completa como 1 Estado-de-Coisas (e_i) delimitado pela propriedade f para o predicado verbal (f_k: buscar_V (f_k)), além de outros dois Estado-de-Coisas: (e_j: sarna (e_j)), que, em sua base metafórica, representa *problema*, assim como vimos na EI do mesmo tipo 3 – *descascar o abacaxi*, e (e_k: coçar (e_k)). Considerando a existência de um Indivíduo *Actor* (x_i), ativamos a camada (f_i) para tal relação com a EI e o verbo *buscar*, i.e., alguém busca por problema em sua significação idiomática.

Nota-se também a relação de coindexação para o mesmo Indivíduo (x_i), que busca por problemas (*Actor*), e o pronome reflexivo *se*, marcando que o Indivíduo também é *Recipient* do problema, que, estando relacionado semanticamente com a preposição *para*, indicando finalidade (para se coçar), marcamos essa relação semântica com a Função *Recipient* (REC) (e_k: (f_m: coçar_V (f_m)) (e_k))_{REC}.

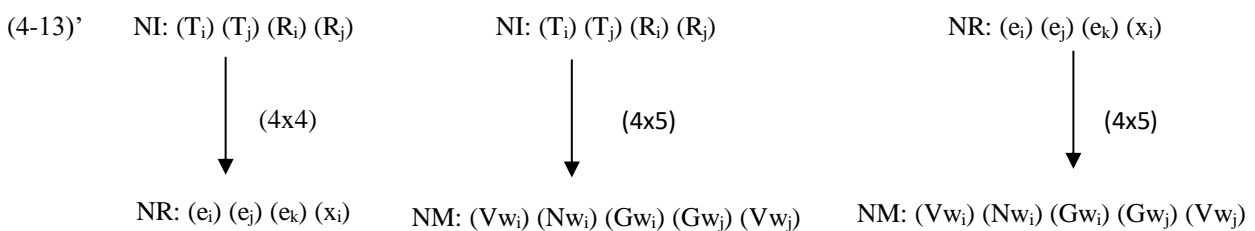
No NM, a EI em sua forma variada é representada por 2 Palavras Verbais (Vw_i: buscar (Vw_i)) e (Vw_j: coçar (Vw_j)), 2 Palavras Gramaticais (Gw_i: para (Gw_i)) e (Gw_j: se (Gw_j))_{Subj}, 1 Palavra Nominal (Nw_i: sarna (Nw_i)). O pronome reflexivo *se* ativa a função de sujeito (_{Subj}), pois consideramos a coindexação do Indivíduo (x_i) como *Actor* da EI como um todo e *Recipient* do verbo *coçar* no NR.

Ao analisar a EI em (4-12), entende-se que no NI o excerto ativa a camada do Discurso (D_i), pois, como vimos, a EI aparece de modo descontínuo em dois Movimentos: (D_i: [(M_i) (M_j)_{Resgate} (D_i)). Adicionamos a função Resgate entre os dois Movimentos, pois nota-se que o falante realiza um

resgate dos problemas existentes no Ministério, vistos em (M_i) - *os tribunais que ainda não conseguem julgar quantos processos lhes são remetidos*, e com a segunda parte componente da EI acionada em (M_j) – *deseja mais sarna*, o falante introduz os novos problemas (*mais sarna*) - *a tutela desta polícia criminal e da polícia económica*. Esse acionamento da camada do Discurso se faz necessário neste caso para que o falante possa entender a função da EI de modo completo, pois, além de a EI estar segmentada em dois Movimentos ((M_i) e (M_j)), também foi apresentada em inversão da ordem preferida – *para se coçar* no (M_i) e *deseja mais sarna* no (M_j).

No NR, adicionamos o advérbio *mais*, representado como um operador de quantidade "m" ao Estado-de-Coisas *sarna* (m e_j: (f_i: sarna (f_i)) (e_j))_U. Já na Codificação, a camada do Discurso no NI ativa a camada do Texto no NM: (Te_i: [(Le_i) (Le_j)] (Te_i)), e o advérbio *mais* é representado como 1 Palavra Adverbial em um Sintagma Adverbial: (Advp_i: (Adw_i: mais (Adw_i)) (Advp_i)).

Como resultados de alinhamento entre os três níveis da teoria para a forma consagrada *procurar sarna para se coçar*, temos relação de transparência apenas entre NI e NR (quatro unidades): dois Subatos Atributivos (T_i: procurar (T_i)) e (T_j: coçar (T_j)) e dois Subatos Referenciais (R_i: se (R_i)) e (R_j: sarna (R_j)) no NI; três Estados-de-Coisas: (e_i) delimitado pela propriedade (f_k) para o predicado verbal (f_k: procurar_v (f_k)), (e_j: sarna (e_j)) e (e_k: coçar (e_k)), e um Indivíduo (x_i: se (x_i)). Essas unidades são codificadas no NM em cinco unidades: 2 Palavras Verbais (Vw_i: procurar (Vw_i)) e (Vw_j: coçar (Vw_j)), 2 Palavras Gramaticais (Gw_i: para (Gw_i)) e (Gw_j: se (Gw_j))_{Subj}, e 1 Palavra Nominal (Nw_i: sarna (Nw_i)), o que resulta em opacidade entre os níveis NI e NM e NR e NM. Vejamos a representação do alinhamento em (4-13)’:



4.1.4: EIs do Tipo 4

Apresenta-se agora dois exemplos do tipo 4, EI *AD-HOC* (finalidade específica). Cumpre esclarecer que nosso objetivo em analisar as EIs desse tipo consiste na compreensão do processo de esforço extra-cognitivo do falante para reanalisar as novas expressões *AD-HOC*, com base no processo de mudança de não-decomponibilidade para decomponibilidade das expressões, pois o falante reanalisa o novo elemento lexical da EI baseado na contribuição das leituras e literal e

idiomática para a expressão previamente convencionalizada e introduzida no discurso. Assim, não aplicaremos os critérios de variação morfossintática ao novo elemento.

Começamos pela EI vista anteriormente em (4-1) *tirar o cavalo da chuva*, aqui repetida como (4-14), com foco na ampliação do excerto:

- (4-14) - Sr. Deputado Nogueira de Brito, no seu tempo, faça favor de continuar, se assim o entender, a sua exposição.
 - Eu “tiro o cavalo da chuva”.
 - Do tempo de que disponho, gostaria de registar que o termo que empreguei há pouco foi “calamidade” e não “**cavalidade**”, como parece ter sido entendido. (Portugal NIL Política A137983)

Eu	tiro	o	cavalo	da	chuva.
1 SG	V tirar PRS	DEF ART M	cavalo	PREP. DEF F	chuva

Como visto anteriormente, a EI *tirar o cavalo da chuva* pertence ao Tipo 1 – desmotivada e não-decomponível. Nosso foco está no elemento *AD-HOC cavalidade*. Como podemos ver com a extensão do excerto, o falante utiliza o termo *cavalidade* de modo irônico em resposta à fala do deputado Nogueira de Brito, que utilizou a EI *tirar o cavalo da chuva* para manifestar sua desistência de seguir com sua exposição sobre as calamidades públicas no país.

Com isso, o falante utilizou o termo *cavalidade* como junção do elemento da EI *cavalo* e o termo *calamidade*, tópico da sessão política. Nota-se que a intenção do falante que questiona a atitude do deputado Nogueira de Brito é de ironizar o uso da EI, pois notou que o deputado tentou fugir de sua responsabilidade sobre o tema das calamidades públicas, acionando as cenas literal e idiomática da EI: tirar o cavalo como sinal de desistência de sua explanação sobre as calamidades públicas.

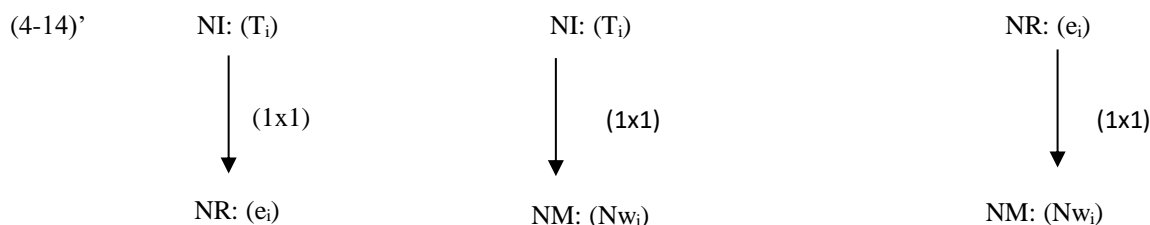
Com isso, entende-se que o uso do termo *cavalidade* é compreendido de acordo com sua conexão com a EI introduzida anteriormente pelo outro falante. Como há essa conexão entre turnos de fala, entendemos que o termo *AD-HOC* aparece em descontinuidade entre forma e significado e sua significação idiomática é entendida em um processo de referência anafórica, que norteia o processo de reanálise do novo termo, com base na leitura literal e idiomática da EI previamente introduzida.

O esforço extra-cognitivo do interlocutor consiste em decompor o termo *cavalo* da EI não-decomponível *tirar o cavalo da chuva*, relacioná-lo com o contexto do discurso, entendendo sua relação com o tópico *calamidade* para, assim, compreender a intenção irônica do falante ao criticar a desistência do deputado.

Tal reanálise perpassa também pela desambiguação do termo *cavalidade* – que significa, em sua leitura denotativa, brutalidade, estupidez, conforme 194 resultados encontrados na *web*, com filtragem em páginas de Portugal apenas, variedade da ocorrência. Notamos claramente na leitura do excerto completo que o termo *cavalidade* não significa brutalidade ou estupidez, e sim, é um termo com a finalidade específica – ironia, relacionado diretamente com a EI *tirar o cavalo da chuva*. Vemos, assim, que o termo *cavalidade* foge ao significado denotativo de sua significação dicionarizada – brutalidade ou estupidez - e incorpora o significado conotativo de ironia, combinando parte da EI *tirar o cavalo da chuva* e o tópico discursivo corrente – **calamidades** públicas, resultando em uma EI *AD-HOC* com uso específico.

Como resultados de sua representação na GDF, *cavalidade* representa um Subato Atributivo (T_i : cavalidade (T_i)) no NI. Considerando sua relação em referência anafórica com a EI, a camada do Discurso é ativada: (D_i : [(M_i) (M_j) $_{Trans}$] (D_i)), na qual (M_i) contém a EI – *Eu tiro o cavalo da chuva*, e (M_j) contém o Subato Atributivo (T_i : cavalidade (T_i)). A relação entre os dois Movimentos, que representam um turno de fala entre os dois falantes, é marcada pela função de Transição ($_{Trans}$), pois notamos que há uma preocupação de o falante explicar sua real intenção irônica ao empregar o termo *cavalidade* (Foco) em detrimento ao termo *calamidade* (Tópico).

No NR, *cavalidade* corresponde a um Estado-de-Coisas (neg e_i : cavalidade (e_i)), com o operador de negação *neg* (*não cavalidade*). No NM, o termo *AD-HOC* é representado por uma Palavra Nominal (Nw_i : cavalidade (Nw_i)), correspondendo a um substantivo, inserido em um Sintagma Nominal. Considerando a relação em referência anafórica com a EI, a camada do Texto é ativada na Codificação: (Te_i : [(Le_i) (Le_j)] (Te_i)), referente aos Movimentos do NI (M_i) e (M_j), respectivamente. Como resultado, temos uma relação transparente entre os três níveis, considerando apenas o termo *AD-HOC cavalidade*, conforme visto em (4-14)’:



A seguir temos em (4-15) o segundo exemplo de EI *AD-HOC*, com a expressão *ser a gota de água*, encontrada na variedade de Portugal:

(4-15) O Orador: - Simplesmente, estas não são situações de ontem, da semana passada ou de há 15 dias,

mas que se vêm arrastando e acentuando ao longo dos últimos anos. Entrou-se numa situação que tem a ver com a corrupção e com um conjunto de fenómenos que foram aqui denunciados, aos quais o Governo terá de fazer face. Este é **apenas um exemplo, uma gota de água**, apesar de tudo importante...

O Sr. Veiga de Oliveira (PCP): - **Uma gota de água?! E mas é um balde!**

O Orador: - em relação a toda esta conjuntura. (Portugal NIL Política A153167)

E	mas	é		um		balde!
CONJ	CONJ	COP PRES 3SG		INDEF ART M		balde

Conforme vimos, a EI *ser a gota de água* significa saturação, estopim para crise. Encontramos 862.000 resultados na *web* e sessenta ocorrências no CRPC. No excerto acima, novamente um debate político, o orador elenca uma série de problemas políticos relacionados à corrupção e compara esse conjunto de fenómenos denunciados e acumulados há anos, como mais uma gota de água, prestes a gerar uma nova crise no governo. Em uma resposta irônica, o deputado Veiga de Oliveira compreende que esse conjunto de problemas não representa apenas mais uma gota de água, e sim um balde.

Tal comparação acionada pela significação da cena literal da EI refere-se às dimensões físicas de uma gota de água e a proporção dessa gota como um balde de água, o que se reflete na significação idiomática, já que tais problemas no governo se acumularam em grandes proporções, de acordo com o contexto.

Ainda que haja uma imagem metafórica de algum recipiente transbordando ao receber uma (última) gota de água, é difícil depreender o significado idiomático de saturação pela interpretação literal da EI, assim como é muito difícil decompor suas partes componentes para compreender o valor idiomático. Vejamos a análise morfossintática da forma consagrada *ser a gota de água*:

- Substituição lexical: Ser a **porção/litro/mililitro** de **vinho/sangue***
- Inserção de pronomes: Foi **essa** a gota de água/Foi **aquela** gota de água
- Inserção de advérbios: Foi **agora/aqui** a gota de água
- Modificação adjetiva: Ser a **enorme/pequena** gota de água*
- Número: Ser as **duas** gotas de água*
- Quantificação: Ser **muita** gota de água*
- Definitude: Ser **a/uma** gota de água (resultados encontrados no CRPC e na *web*)
- Nominalização: **N/A**

- Passivização: N/A
- Possibilidade de conversão para o modo imperativo: **Seja a gota de água***
- Elisão de um termo fixo do predicado: **Ser a gota**
- Inversão da ordem preferida: **A gota de água foi/é**
- Referência anafórica: [...] - **uma última gota de água** faltava para fazer transbordar a taça. A administração dos Estaleiros de São Jacinto, entidade que há 35 anos vinha assegurando a ligação com Aveiro através de uma carreira de lanchas - inicialmente prevista para transporte dos seus trabalhadores, mas de que se vinha servindo desde sempre a população em geral -, fez cair **essa gota de água**, suspendendo a carreira a partir de 1 de Janeiro, invocando para tanto o estado de decadência das lanchas e o carácter não obrigatório do serviço prestado. (resultado encontrado no CRPC)
- Referência catafórica: A Oradora: - [...] Portanto, o Sr. Ministro colocou-se numa vulnerabilidade que pensamos ser também muito desagradável para si e, sobretudo, para o Governo.

O Sr. Jorge Ferreira (CDS-PP): - Muito bem!

A Oradora: - Finalmente, desejo referir um outro aspecto que penso ser muito importante. Julgo que **isto foi a gota de água**, porque o Sr. Ministro já foi obrigado a fazer três actos de governo que são perfeitamente contrários à sua maneira de ser e a tudo aquilo que disse: foi o totonegócio, o cinenegócio e agora é este. (resultado encontrado no CRPC)

Ao analisarmos os excertos, notam-se alguns casos do uso da expressão *ser a gota de água* com o sentido idiomático de significar insignificância, normalmente associado com outros termos, como *oceano, mar, imensidão, etc.*

Com foco nos excertos com o sentido idiomático de saturação, observa-se que a EI apresenta maior restrição sintática, permitindo inserção de pronomes demonstrativos, advérbios, inversão da ordem preferida (considerando o verbo *ser* como parte fixa da EI), definitude – considerando que o artigo definido faz parte da forma consagrada, referências anafórica e catafórica por entidade do discurso previamente/posteriormente introduzida.

A referência anafórica também pode ser acionada por parte da EI. Sobre modificação adjetiva, encontramos apenas o uso de *último*, adjetivo estritamente conectado com as cenas literal e idiomática da EI – última gota antes do transbordo e última situação tolerável, por isso não consideramos que essa EI permite casos de modificação adjetiva de modo tão claro e sem restrições. *Ser uma pequena gota de água* remeteria ao significado idiomático de insignificância, conforme dados da *web*.

Ao considerarmos o caso de elisão do termo *de água* como possível, identificamos que é necessário o direcionamento do contexto para a significação de saturação, devido à possibilidade de ambiguidade com a leitura idiomática de insignificância, ou outras expressões que referem-se a outros contextos, como *gota de suor e gota de sangue*, ambas remetendo a sacrifício, trabalho duro ou exploração. Como resultado, consideramos a EI *ser a gota de água* como Tipo 1 – desmotivada e não-decomponível.

Ao analisar o termo *AD-HOC balde* (*ser um balde de água*) visto em (4-15), observa-se que o termo foi acionado em referência anafórica por parte componente da EI introduzida anteriormente, dessa vez, pelos dois participantes do diálogo – orador e deputado, configurando mais um caso de **descontinuidade** entre forma e significado, pois só compreendemos a resposta irônica do falante ao utilizar o termo *balde* significando os problemas políticos acumulados há anos, através de sua relação direta com a EI *ser a gota de água*.

O processo de reanálise do termo em referência anafórica exige que o falante compare os termos *gota* e *balde* em suas significações literais, i.e. propriedades físicas, e assim, compreenda sua significação idiomática direcionada pelo contexto: os problemas de corrupção elencados anteriormente durante o debate político tomaram grandes proporções.

Tal contexto contribui para que não haja ambiguidade de significado idiomático da EI, tanto com a significação de insignificância, como também com outra EI – *ser um balde de água fria* (decepção) ou os possíveis casos citados de trabalho duro, sacrifício e exploração.

Assim como o termo *cavalidade*, o termo ***balde de água*** também foge ao significado denotativo de sua significação dicionarizada – recipiente de formato cilíndrico ou tronco de cone, utilizado para transportar ou armazenar líquidos ou sólidos, como resultado final de sua significação idiomática no contexto. O termo *balde* incorpora o significado conotativo da ironia evocada pelo falante, combinando parte da EI *ser a gota de água* e o tópico discursivo corrente – proporção dos problemas de corrupção política, resultando em uma EI *AD-HOC* com uso específico.

Similar à relação vista entre o termo *cavalidade* (EI *AD-HOC*) e o termo *cavalo*, que é parte da EI *tirar o cavalo da chuva*, ambos relacionados ao tópico discursivo de calamidade pública visto no caso anterior, aqui temos uma nova relação de ironia estabelecida entre *ser o balde de água* (EI

AD-HOC), *de água* (parte da EI *ser a gota de água*) e proporção dos problemas de corrupção (tópico discursivo).

Neste último caso, o falante causou o efeito de ironia ao comparar as dimensões físicas de uma gota de água e de um balde de água com as proporções dos problemas de corrupção. Reforçamos a importância do contexto para o processo de reanálise do termo *AD-HOC* nas EIs do tipo 4, assim como a importância do contexto para evitar ambiguidade com outras EIs (e.g. *ser um balde de água fria*, *ser/derramar gota de sangue/suor*), ou outras significações dicionarizadas, como visto em *cavalidade* como estupidez.

Realizamos abaixo a representação formal da EI em sua forma consagrada:

(4-16) Ser a gota de água.

NI: (M_i: (A_i: [(F_i: DECL (F_i)) (P₁)_S (P₂)_A (C_i: (+ id R_i: (T_i) (T_j) (R_i))_{FOC} (C_i))] (A_i))(M_i))

NR: (p_i: (NON-FIN ep_i: (1 e_i: (f_i: **gota de água** (f_i)) (e_i)) (ep_i)) (p_i))

NM: (L_{e_i}: (C_i: [(Vp_i: (Vw_i: ser_{NON-FIN} (Vw_i)) (Vp_i)) (Np_i: [(Gw_i: a (Gw_i)) (Nw_i: **gota** (Nw_i)) (Adpp_i: (Gw_j: de (Gw_j)) (Adpp_i)) (Np_j: (Nw_j: **água** (Nw_j)) (Np_j))] (Np_i))] (C_i)) (L_{e_i}))

No NI, vemos que a EI se comporta como um único Subato Referencial (R_i), composto por dois Subatos Atributivos (T_i: gota (T_i)) e (T_j: água (T_j)), com função Foco. Conforme a GDF, os verbos de cópula não possuem representação no NI, pois não evocam propriedade, sendo representados como elementos gramaticais no NR. Sabendo que o artigo definido faz parte da forma consagrada e encontramos outros casos de definitude no CRPC, adicionamos o operador de identificabilidade (+ id).

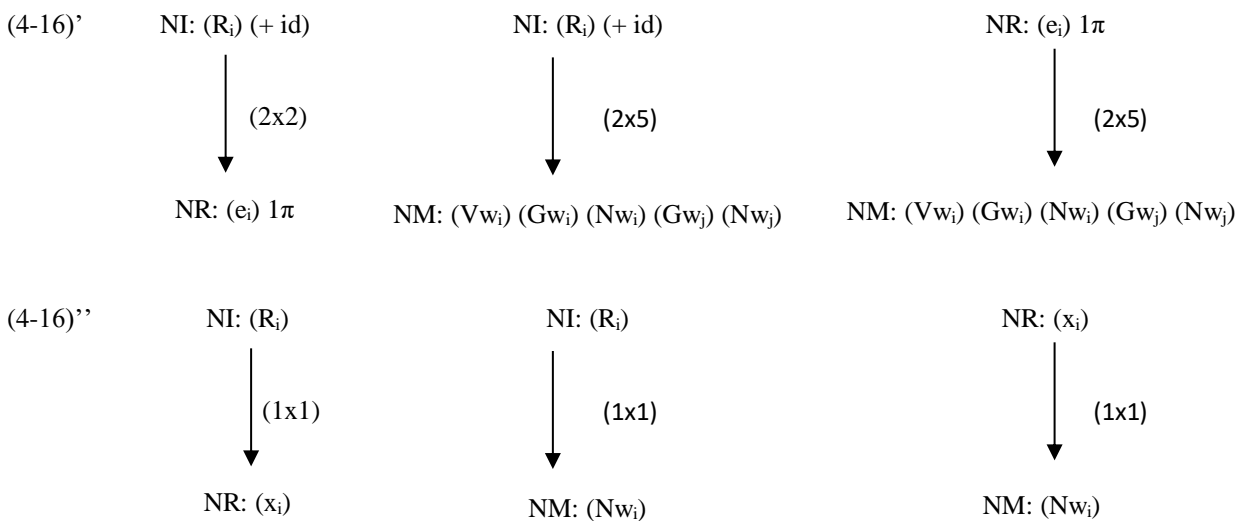
No NR, a EI é analisada como um Estado-de-Coisas (1 e_i: (f_i: gota de água (f_i)) (e_i)), e o verbo cópula *ser* não é representado lexicalmente, ativando apenas o operador de tempo NON-FIN na camada no Episódio (ep_i). Nesse caso, não consideramos a existência de um Indivíduo (x), como fizemos nos exemplos anteriores, e por isso, não há relação hierárquica identificada, e nem função.

Poderíamos adicionar um Estado-de-Coisas (e.g. reunião), um Indivíduo ou um Locativo (e.g. Londres), relacioná-los com a EI e assim ativar a relação hierárquica entre Estado-de-Coisas (e) e Indivíduo (x), como por exemplo, *Aquele carro velho foi a gota de água*, onde *carro* (x_i: (f_j: carro (f_j)) (x_i)) receberia a função de *Undergoer* (U). Adicionamos o operador de quantidade (1 e_i: (f_i: gota de água (f_i)) (e_i)) ao Estado-de-Coisas (e_i) referente à possibilidade de definitude e inserção de pronomes descrevendo *gota de água* como singular, pois não encontramos casos de quantificação e número.

Já no NM, a EI é representada por cinco unidades na Oração (Cl_i): 1 Palavra Verbal (Vw_i : ser (Vw_i)), ativando um Sintagma Verbal (Vp_i), 2 Palavras Nominais (Nw_i : gota (Nw_i)) e (Nw_j : água (Nw_j)) ativando dois Sintagmas Nominais (Np_i) e (Np_j), e duas Palavras Gramaticais (Gw_i : a (Gw_i)) e (Gw_j : de (Gw_j)), esta ativando um Sintagma Adposicional ($Adpp_i$).

Como o resultado do alinhamento entre os níveis, vemos que há transparência apenas entre os níveis de Formulação: 1 Subato Referencial (R_i : ser a gota de água (R_i)) e 1 operador de identificabilidade (+id: a) no NI, e 1 operador de quantidade (1 e_i) e 1 Estado-de-Coisas (1 e_i : (f_i : gota de água (f_i)) (e_i)) no NR, codificados em cinco unidades no NM: 1 Palavra Verbal (Vw_i : ser (Vw_i)), 2 Palavras Nominais (Nw_i : gota (Nw_i)) e (Nw_j : água (Nw_j)) e duas Palavras Gramaticais (Gw_i : a (Gw_i)) e (Gw_j : de (Gw_j)).

Considerando apenas a representação do termo *AD-HOC balde* visto em (15), temos uma relação de transparência entre os três níveis: 1 Subato Referencial (R_i : balde (R_i)) no NI, 1 Indivíduo (x_i : balde (x_i)) no NR e 1 Palavra Nominal (Nw_i : balde (Nw_i)) no NM. Em (4-16)' temos o esquema de alinhamento da EI em sua forma consagrada, e em (4-16)'' temos a relação de transparência do elemento *AD-HOC balde*:



4.1.5: EIs do Tipo 5

Prosseguindo com a análise, são discutidos nessa seção dois exemplos do tipo 5: EIs parcialmente motivadas por metáforas conceituais e decomponíveis, a começar com a EI *engolir sapo*, encontrada na variedade europeia:

- (4-17) “Vamos votar no outro”, tinham decidido os comunistas em Congresso. O "outro" acabou por ser Salgado Zenha, o ex-inimigo do "Verão quente". Foi a última tentativa "frentista", agora com o partido que recolhia alguns "capitães de Abril", apesar de ter como inspirador Ramalho Eanes, a principal figura do 25 de Novembro. Mas o PCP irá ainda **"engolir um sapo" maior**. Na segunda volta, reúne novo Congresso e acaba por votar Mário Soares. A contestação aumenta e transborda as paredes da "Soeiro" (Portugal 1991 Jornal J0037)

Engolir um sapo maior
 V engolir NON-FIN INDEF ART M sapo A

Na análise cognitiva da EI encontrada em um texto jornalístico, nota-se que o partido político PCP teve que lidar com a atuação de um deputado tido como desagradável ou de má reputação – Salgado Zenha, e novamente, após novas eleições no Congresso, aparentemente solicitadas com o aval e/ou participação do próprio PCP, terão que atuar outro político ainda mais desagradável – Mário Soares, entendido como um problema de maior proporção e que causou maiores contestações.

Conforme visto anteriormente, a EI *engolir sapo* significa tolerar uma situação desagradável sem reclamar, com 1.620.000 resultados encontrados na *web* e 63 ocorrências no CRPC. Constatamos essa significação idiomática no excerto acima, pois, o PCP terá que lidar com a presença desagradável do deputado eleito em votação, que teve aval e/ou participação do próprio partido, e não poderão reclamar de uma situação que tiveram algum tipo de participação.

Ao analisarmos as cenas literal e idiomática da EI, identificamos uma motivação parcial ativada pela metáfora conceitual do *sapo* como animal ou criatura desagradável, nojenta ou de aspecto físico que causa certa repulsa. Essa imagem metafórica esquemática associa um conceito – asco, e uma idiomatidade ativada na EI – situação desagradável.

Essa imagem metafórica e pejorativa do animal é um elemento também fortemente presente em diferentes culturas, faz parte de uma metáfora cultural quase-universal, como podemos ver em contos de fadas, onde a princesa beija o sapo (e.g. *Princesa Tiana* da *Disney*, Estados Unidos, baseada na obra do século XVIII dos irmãos Grimm “*The Frog Princess*”, Alemanha), canções infantis (e.g. *O sapo não lava o pé*, presente na coletânea da Galinha pintadinha e da Eliana, Brasil), sua imagem em animações do cinema que o colocam como criaturas traçoeiras e asquerosas (e.g. *A viagem de Chiriro*, do Studio Ghibli, Japão), entre outras espalhadas pelo mundo.

Tal conexão pode ser explicada como uma evocação mental constituída com base em subestruturas conceituais sem conexão com experiências baseadas na cena literal, na qual não é necessário literalmente engolir um sapo (ainda que seja humanamente possível) para apreender a sensação desagradável e de repulsa.

Essa imagem metafórica é semanticamente ativada apenas pelo termo *sapo*, o que explica sua composição parcial – o significado literal desse elemento contribui diretamente para o significado idiomático. Com isso, entendemos que *sapo* é semanticamente autônomo e possui significado figurativo lexicalizado motivado metaforicamente. Já a ação de *engolir* contribui para a compreensão da significação idiomática de tolerar tal situação desagradável, mas não é tão claro a significação de aceitar essa situação sem reclamar.

Como resultado, é proposto um novo tipo de EI que contempla esses fenômenos de natureza metafórica: **expressões idiomáticas parcialmente motivadas por metáforas conceituais e decomponíveis**. Para a elaboração de uma definição formal desse tipo, tomamos como base teórica os autores estudados nos capítulos anteriores: Langlotz (2006), Keizer (2016, 2018), Lakoff e Johnson (1980), Lakoff (1986), (Kövecses, 2010) e Vega Moreno (2007).

Segue a definição formal do Tipo 5: são expressões que apresentam **motivação parcial**, ativada por **metáforas conceituais**, que são imagens metafóricas esquemáticas associadas a um conceito e a uma idiomaticidade. Conforme vimos em Langlotz (2006), Lakoff e Johnson (1980) e Lakoff (1986), esse processo de criação de imagens metafóricas são evocações mentais constituídas com base em subestruturas conceituais sem conexão com experiências baseadas na cena literal.

Essas imagens metafóricas são ativadas semanticamente em partes das EIs, o que explica sua **composição parcial** – o significado literal de 1 elemento contribui diretamente para o significado idiomático, o que implica em seu caráter **decomponível**. Outro fator que contribui para sua decomponibilidade é a **motivação constitucional** vista em Langlotz (2006): seus constituintes são semanticamente autônomos e possuem significados figurativos lexicalizados motivados metaforicamente.

Por fim, conforme Vega Moreno (2007), o significado da EI, licenciado metaforicamente, atua como fator motivacional para sua convencionalização como parte integrante na cultura de comunidade falante, o que se fundamenta também no processo de criação e decodificação de **metáforas culturais universais** ou quase-universais (Kövecses, 2010).

Entende-se, assim, que a proposta de um novo tipo de EI contribui para a compreensão dos diferentes graus de motivação e de composicionalidade das EIs no português, que variam entre motivadas e desmotivadas, e decomponíveis e não-decomponíveis, através de uma análise semântico-cognitiva dos termos que ativam a interpretação metafórica, através das imagens metafóricas esquemáticas, e de todo o seu universo conceitual e convencional.

Somando esse novo tipo de EI aos demais tipos propostos por Keizer (2016), observa-se que o quinto tipo se diferencia dos demais tipos, pois contribui para um maior refinamento dos graus de

motivação e de composicionalidade das EIs, visto que tal proposta capta um olhar minucioso para os processos de interpretação metafórica das EIs, considerando as metáforas conceituais de cada língua, um maior detalhamento da composição dos elementos (se parcial ou literal) e sua autonomia no processo de interpretação metafórica, além de levar em conta os processos de criação e decodificação baseados nas metáforas culturais universais ou quase-universais.

Seguimos agora para os critérios de variação morfossintática da EI:

- Substituição lexical: **Digerir** sapo (resultado encontrado no CRPC)
- Inserção de pronomes: Engolir **esse/este** sapo (resultado encontrado no CRPC)
- Inserção de advérbios: Engolir **diariamente/já/aqui** esse sapo (resultados encontrado na *web*)
- Modificação adjetiva: Engolir sapo **vivo/maior/pequeno/enorme/indigesto** (resultados encontrados no CRPC)
- Número: Engolir **dois/primeiro/segundo** sapo(s) (resultados encontrados no CRPC e na *web*)
- Quantificação: Engolir **mais um/muito/cada** sapo (resultados encontrados no CRPC e na *web*)
- Definitude: Engolir **o/um/algum** sapo (resultados encontrados no CRPC e na *web*)
- Nominalização: **Engolimento de sapo***
- Passivização: **Muitos sapos foram engolidos** nessa reunião
- Possibilidade de conversão para o modo imperativo: **Engula** esse sapo (resultado encontrado na *web*)
- Elisão de um termo fixo do predicado: **N/A**
- Inversão da ordem preferida: **Sapo engolir***
- Referência anafórica: **João acabou de engolir um sapo enorme na reunião. Esse sapo atende pelo nome de chefe estressado.**
- Referência catafórica: Você deve parar de fazer **isso: engolir sapo** nesta empresa.

Podemos ver que a EI apresentou um comportamento morfossintático muito flexível, permitindo substituição lexical com verbo pertencente ao mesmo campo semântico (*engolir* e

digerir), inserção de advérbios, pronomes, adjetivos relacionados à cena literal (características físicas do *sapo*), número (incluindo numeração do tipo ordinal), quantificação, definitude, possibilidade de conversão para o modo imperativo, passivização, referência anafórica por parte da EI e pela entidade previamente introduzida no discurso.

A referência catafórica só é possível pela entidade posteriormente introduzida no discurso, acionada por pronome. Apenas os casos de nominalização e inversão da ordem preferida não foram possíveis, já que, pela limitação de elementos lexicais da EI, não se aplica a elisão de um termo fixo do predicado. Vejamos em (4-18) a representação formal nos níveis da GDF, utilizando a forma consagrada:

(4-18) Engolir sapo.

NI: (M_i: (A_i: [(F_i: DECL (F_i)) (P₁)_S (P₂)_A (C_i: [((T_i) (R_i: (T_j) (R_i))_{FOC}](C_i))] (A_i))(M_i))

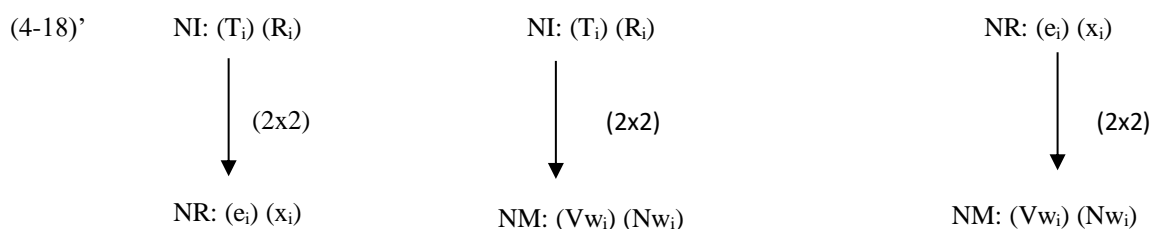
NR: (p_i: (NON-FIN ep_i: (e_i: [(f_i: (f_j: engolir_V (f_j)) (x_i: (f_k: sapo_N (f_k)) (x_i)_U (f_i))] (e_i)) (ep_i)) (p_i))

NM: (Le_i: (Cl_i: [(Vp_i: (Vw_i: engolir_{NON-FIN} (Vw_i)) (Vp_i)) (Np_i: (Nw_i: sapo (Nw_i)) (Np_i))] (Cl_i)) (Le_i))

No NI a expressão é representada por 1 Subato Atributivo (T_i: engolir (T_i)) e 1 Subato Referencial (R_i: sapo (R_i)), com função de Foco na EI como um todo. Por (R_i) ser um Subato Referencial independente, justifica-se assim a possibilidade de modificação adjetiva de modo produtivo. No NR temos 1 Estado-de-Coisas (e_i), caracterizado pela predicado verbal (f_j: engolir_V (f_j)), e 1 Indivíduo (x_i: sapo (x_i)), com função de Undergoer (U) ao verbo *engolir*.

Nesse caso, considera-se *sapo* como Indivíduo devido aos casos de modificação adjetiva, quantificação, número e definitude serem estritamente ligados às propriedades físicas do *sapo*, o que facilita sua representação como um Indivíduo (x_i), uma entidade de primeira ordem que pode ser localizada no espaço e ser avaliada em termos de sua existência (Hengeveld; Mackenzie, 2008, p. 131).

Já no NM a EI é expressa por 1 Palavra Verbal (Vw_i: engolir_{NON-FIN} (Vw_i)) e 1 Palavra Nominal (Nw_i: sapo (Nw_i)), como parte de uma Oração (Cl_i). Como resultado do alinhamento entre os níveis, representado em (4-18)', temos uma relação de transparência entre todos os níveis (duas unidades): 1 Subato Atributivo (T_i: engolir (T_i)) e 1 Subato Referencial (R_i: sapo (R_i)) no NI, 1 Estado-de-Coisas (e_i) caracterizado pela predicado verbal (f_j: engolir_V (f_j)), e 1 Indivíduo (x_i: sapo (x_i)) no NR, e 1 Palavra Verbal (Vw_i: engolir_{NON-FIN} (Vw_i)) e 1 Palavra Nominal (Nw_i: sapo (Nw_i)) no NM.



Abaixo seguimos com a análise cognitiva do segundo exemplo de EI do Tipo 5 – *Fazer tempestade no copo de água*, localizada na variedade de Portugal:

- (4-19) Orador: - Quanto ao conteúdo das suas palavras, como é “evidente”, o Sr. Deputado diz o que quer e muito bem. Mas não se pode é depreender das minhas palavras que estejamos a coagir o Sr. Deputado. Por outro lado e no aspecto de tempo concordo que aí fazemos parte do mesmo clube. Portanto, se o seu grupo parlamentar tem pouco tempo, nós não teremos mais. Parece-me assim que o Sr. Deputado **fez uma tempestade num copo de água**. É que a minha intenção era no sentido de lamentar - e espero que o Sr. Deputado me dê o direito de ter a minha opinião - que não tenha havido da parte da oposição aquilo que se esperaria, ou seja, a tomada de iniciativa. Ora, esta iniciativa pertenceu ao Sr. Presidente e não à oposição. Parece-me que isto é inquestionável.
- O Sr. Presidente: - Para pedir esclarecimentos ao Sr. Deputado Veiga de Oliveira tem a palavra o Sr. Deputado José Alberto Xerez. (Portugal NIL Política A163198)

o	Sr.	Deputado	fez	uma	tempestade
DEF ART M	PRON M	Deputado	V fazer PST	INDEF ART F	tempestade
num	copo	de	água.		
INDEF PREP	copo	PREP	água.		

Em mais um debate político, acompanhamos o diálogo entre os deputados Veiga de Oliveira e José Alberto Xerez. Com a leitura do excerto completo, nota-se que Xerez sentiu que foi coagido pela oposição e deixou sua insatisfação evidente em turno de fala anterior. Já Veiga de Oliveira, em resposta, acredita que Xerez cometeu um exagero ao enxergar a situação de tal modo, pois a intenção de Veiga na rodada de discussão anterior era apenas lamentar determinada tomada de iniciativa – que aparentemente partiu do presidente do Congresso, e não do lado partidário de Xerez, e assim, não quis incitar que Xerez foi coagido a fazê-la.

A EI *fazer tempestade num copo de água*, que significa cometer um exagero, teve vinte ocorrências encontradas no CRPC e 425.000 resultados na *web*. É possível ver claramente essa significação de exagero no excerto acima, na qual o deputado Veiga acredita que Xerez cometeu um enorme exagero ao entender a explanação anterior como uma acusação de coerção.

Assim como na EI *engolir sapo*, identificamos uma motivação parcial ativada pela metáfora conceitual de *tempestade* como um evento meteorológico que denota uma perturbação atmosférica

violenta, causando estragos, danos materiais, pânico, desordem natural do clima. Essa imagem metafórica esquemática associa um conceito – perturbação atmosférica, e uma idiomaticidade ativada na EI – sofrimento, desgraça, preocupação, medo.

Tal imagem metafórica pode ser compreendida como uma metáfora cultural universal, já que é possível ver sua representação em diferentes culturas, como nas EIs do inglês: *to make a storm in a teacup* (Fazer tempestade numa xícara de chá), *calm after the storm* (Calma depois da tempestade), e na EI em alemão *Wer Wind sät, wird Sturm ernten* (Quem semeia vento colherá tempestade), que refere-se ao texto bíblico do livro de Oséias, no Antigo Testamento. O termo também é comumente utilizado em nomes de franquias de equipes esportivas, significando que o time é poderoso, capaz de fazer estragos nos adversários, como a equipe de basquete *Seattle Storm*, que atua na liga nacional de basquete feminino dos Estados Unidos.

Verifica-se que a evocação mental do termo também é constituída sem conexão com experiências baseadas na cena literal, devido à impossibilidade de realização. A decomponibilidade da EI passa pela composição parcial do termo, já que seu significado literal contribui para o significado idiomático.

Já a contribuição do termo *num copo de água*, que expressa o local em que essa tempestade acontece na cena literal, parece limitar-se à dimensão física real do objeto, resultando em uma impossibilidade, mas a significação idiomática de exagero não fica tão clara, já que diversas situações de exagero são possíveis, tanto em construções concretas (e.g. *Você exagera na bebida todo fim de semana*), como em construções abstratas (e.g. *Você pensa nessa mulher de modo exagerado*).

Partimos agora para a análise morfossintática da expressão:

- Substituição lexical: **Ser/levantar** uma tempestade num copo de água (resultados encontrado no CRPC)
- Inserção de pronomes: Ser **essa** tempestade num copo de água (resultado encontrado no CRPC)
- Inserção de advérbios: Perante tempestade **tão evidentemente** em copo de água (resultado encontrado no CRPC)
- Modificação adjetiva: Fazer uma tempestade **verdadeira/pequenina/violenta** num copo de água (resultados encontrados no CRPC)
- Número: Fazer **duas** tempestades num copo de água*

- Quantificação: Fazer **mais uma/muita/cada** tempestade num copo de água (resultados encontrados no CRPC e na *web*)
- Definitude: Fazer **a/uma/alguma** tempestade num copo de água (resultados encontrados no CRPC e na *web*)
- Nominalização: **N/A**
- Passivização: Uma tempestade foi feita num copo de água
- Possibilidade de conversão para o modo imperativo: **Não faça** tempestade num copo de água (resultado encontrado no CRPC)
- Elisão de um termo fixo do predicado: Fazer tempestade num copo/Fazer tempestade de água*
- Inversão da ordem preferida: Num copo de água fazer tempestade*
- Referência anafórica:

(4-20) A ausência de dissolução tem outra face: a nomeação como chefe de governo do chefe do maior partido da coligação parlamentar maioritária - sem eleições. Ao nomeado faltaria legitimidade pessoal para o cargo.

Daí uma **tempestade num copo de água - tempestade mediterrânica**, forte e breve, que hoje, umas escassas três semanas mais tarde, nos parece já remota e quem sabe se merecedora de um modesto artigo de jornal. **Esta tempestade estival** deriva de uma outra e mais grave ilusão: a ilusão que temos um regime representativo. (Portugal 2009 Jornal noCOD_1069934)

- Referência catafórica:

(4-21) O Sr. Carlos Encarnação (PSD): - Mas isso não é facilitismo, é uma consideração especial!
 O Orador: - Francamente, não entendo como é que os Srs. Deputados do PSD entram tão facilmente em contradição na atitude que adoptam em relação ao Governo.
 O Sr. João Carlos da Silva (PS): - Quem os conhecer percebe!
 O Orador: - Penso, por outro lado, que, ponderadas as afirmações produzidas, até agora, neste debate, estamos talvez, de certo modo, mais perante “**uma tempestade num copo de água**”.
 O Sr. João Carlos da Silva (PS): - Muito bem!
 O Orador: - Isto não significa que a bancada do Partido Socialista não esteja também, mesmo tratando-se de **uma minitempestade**, a colaborar em transformá-la numa **bonança**.
 O Sr. Carlos Encarnação (PSD): - Ah! Assim está bem!
 O Orador: - Com certeza!
 O Sr. Lino de Carvalho (PCP): - É a **pequena tempestade**, porque **as grandes tempestades** fazem vocês no Governo! (Portugal 1997 Reunião Plenária A0150)

Observa-se que a EI apresenta um comportamento morfossintático menos restrito, permitindo

substituição lexical com o verbo, mas não com os demais elementos (*fazer vendaval num copo de água*/fazer tempestade numa garrafa de água*/fazer tempestade num copo de vinho**), inserção de advérbios, pronomes demonstrativos, adjetivos, quantificação, definitude, possibilidade de conversão para o modo imperativo (em caso de negação), passivização, referências anafórica e catafórica por parte da EI e pela entidade previamente/posteriormente introduzida no discurso.

No caso da referência anafórica em (4-20), nota-se que o falante retoma o termo *tempestade* e o modifica com adjetivos relacionados às propriedades físicas da tempestade – *mediterrânica* (evento meteorológico raro e de difícil detecção, similar a tempestades tropicais fortes e breves, e que ocorre na bacia do Mediterrâneo) no Ato Discursivo subsequente, e *estival* (tempo quente e seco que ocorre entre a primavera e o outono, com cessão das chuvas) presente no próximo Movimento.

Tais adjetivos operam na cena literal da EI, e, ao mesmo tempo, contribuem para o significado idiomático de exagero (nomeação do novo chefe de governo sem eleições como um evento impactante, mas ao tempo, breve e que cairá no esquecimento rapidamente), pois reduzem as propriedades destrutivas e imponentes de uma tempestade a eventos meteorológicos de diferentes proporções.

Em ambos os casos, definimos que esses termos não configuram o termo *tempestade* como uma nova EI *AD-HOC*, que, embora estejam empregados para uma finalidade específica – ironia do falante, tais termos são acessórios (modificadores) de *tempestade*, diferente dos termos *cavalidade* e *balde* vistos anteriormente, que substituíram o termo da forma consagrada da EI, resultando em uma nova significação literal e idiomática.

Definimos tais casos vistos em (4-20) como *idiomatic wordplay* (jogo de palavras idiomáticas). Langlotz (2006) entende que esse jogo de palavras idiomáticas é uma variação idiomática que explora diferentes recursos para a criação de efeitos estilísticos e poéticos. Tais recursos podem ser mensurados em três critérios: (i) **Marcação/conspicuidade estilística**: a EI é adaptada de uma forma marcante ou deslumbrante; (ii) **Ambiguidade**: tanto o significado literal quanto o significado idiomático são fortemente implicados; e (iii) **Dependência do contexto**: só é possível presumir o sentido da forma adaptada de uma EI no contexto de sua aplicação.

Com essa definição de Langlotz (2006), utilizada também no trabalho de Keizer (2018) para diferenciar *idiomatic wordplay* e *AD-HOC idioms*, entendemos que as construções *tempestade estival* e *tempestade mediterrânica* são casos que ativam recursos de ambiguidade e dependência do contexto ao mesmo tempo, pois tais adjetivos modificam propriedades semânticas das cenas literal e idiomática da EI, e necessitam do contexto para que o falante compreenda sua motivação de uso.

Já em (4-21), encontramos um caso complexo de descontinuidade entre forma e significado,

realizado em casos de referências anafórica e catafórica entre sete turnos de fala de quatro participantes: o orador da sessão e os deputados João Carlos da Silva, Carlos Encarnação e Lino de Carvalho.

Destaca-se que a **descontinuidade** referente ao significado da forma *tempestade* resulta em dois processos: o primeiro é a ativação semântica do significado idiomático da EI *fazer tempestade num copo de água*, i.e. exagero; e o segundo processo é a ativação semântica apenas da imagem metafórica de tempestade, i.e. perturbação, desordem, tribulação.

Nota-se que o jogo de significações do termo *tempestade* é acionado primeiramente em sua significação idiomática no turno de fala do orador: *Penso, por outro lado, que, ponderadas as afirmações produzidas, até agora, neste debate, estamos talvez, de certo modo, mais perante “uma tempestade num copo de água”*, demonstrando achar um exagero as problemáticas discutidas sobre facilitismo e uma série de contradições sobre o tema entre os deputados do PSD. Na sequência, o deputado do PS, João Carlos da Silva, endossa seu coro de apoio à crítica feita pelo orador ao partido rival: - *Muito bem!*

No próximo turno de fala, surge o segundo acionamento do termo *tempestade* – agora como *minitempestade*, novamente na fala do orador: *Isto não significa que a bancada do Partido Socialista não esteja também, mesmo tratando-se de uma minitempestade, a colaborar em transformá-la numa bonança*. Aqui, notamos uma crítica ao coro de apoio do deputado João Carlos da Silva (PS) vista no turno anterior, pois o orador entende que o PS também contribuiu para a formação da *minitempestade*, em que o termo é acionado anaforicamente, mantendo o significado idiomático da EI acionada no primeiro turno de fala, dessa vez adjetivado com o prefixo *mini*, mostrando que a contribuição do PS para a formação da *tempestade* (exagero) é menor que a contribuição do PSD.

Ainda no mesmo Ato Discursivo, o orador insere o termo **bonança**, que aciona assim, outra expressão popular: *depois da tempestade, vem a bonança*. Diante dessa nova significação, acionada cataforicamente em relação ao termo *minitempestade*, surge também o significado metafórico de *tempestade*: aflição, tribulação, dificuldades. Ao escutar a crítica feita pelo orador para os membros do PS, o deputado Carlos Encarnação (PSD) demonstra apoio ao orador: - *Ah! Assim está bem!* No próximo turno de fala, o orador demonstra concordar com o apoio recebido de Carlos Encarnação: *Com certeza!*

No último turno de fala, o deputado Lino de Carvalho (PCP) aciona o termo *tempestade* duas vezes em dois Atos Discursivos: - *É a pequena tempestade, porque as grandes tempestades fazem vocês no Governo!* O primeiro termo – *pequena tempestade* retoma anaforicamente o significado idiomático acionado pelo orador por *minitempestade*, i.e. exagero.

Sabendo que Lino de Carvalho pertence a outro partido (PCP), entendemos que o deputado coloca-se fora do contexto de *tempestade* enquanto tribulação, problemas, cabendo-lhe, como representante do seu partido, como parte da *pequena tempestade*, i.e. o pequeno exagero na qual a sua fala no início do debate sobre facilitismo gerou as contradições nos membros do PSD. Essa contradição foi a motivação do acionamento da EI pelo orador na primeira aparição do termo no debate, retomada no final por Lino de Carvalho numa tentativa de livrar seu partido (PCP) das responsabilidades dos problemas políticos que permeiam o Congresso.

Na última aparição do termo *tempestade*, nota-se que a adjetivação e que a pluralização *grandes tempestades* dialogam e criam um contraste com *pequena tempestade* (pequeno exagero). Esse contraste também é refletido na significação do termo *tempestade*, pois nota-se que *grandes tempestades* têm a significação metafórica de tribulação: *porque as grandes tempestades fazem vocês no Governo!* Lino de Carvalho agora transfere as responsabilidades dos problemas para os demais membros do debate: orador, João Carlos da Silva e Carlos Encarnação.

Nesse último jogo com o termo *tempestade*, que encerra o turno de fala de Lino de Carvalho, os significados idiomático e metafórico são compreendidos cataforicamente através da relação contrastiva entre *pequena tempestade e grandes tempestades*. Só compreendemos esse contraste entre significados da mesma forma após o acionamento do último termo. Para facilitar a compreensão, ilustramos esse percurso com base na representação simplificada no NI:

1º Acionamento – Orador: (M _i : [(A _i) (A _j : - uma tempestade num copo de água- (A _j))] (M _i))		
Forma: uma tempestade num copo de água		Significado de entrada: idiomático (exagero)
2º Acionamento - Orador: (M _i : [(A _i : - minitempestade - (A _j) (A _j : - bonança- (A _j))] (M _i))		
Forma: minitempestade		Significado de entrada: idiomático (exagero) Processo: Referência anafórica
		Significado de saída: metafórico (tribulação) Processo: Referência catafórica
3º Acionamento - Lino de Carvalho: (M _i : [(A _j : - pequena tempestade - (A _j))		
Forma: pequena tempestade		Significado de entrada: metafórico (tribulação) Processo: Referência anafórica
		Significado de saída: idiomático (exagero) Processo: Referência catafórica

4º Acionamento – Lino de Carvalho: (A _j : - grandes tempestades - (A _j))] (M _i)		
Forma: grandes tempestades		Significado de saída: metafórico (tribulação)

Figura 10: Descontinuidade na EI *Fazer tempestade num copo de água* (Elaboração própria, 2025)

Vejamos em (4-22) a representação formal nos níveis da GDF, utilizando a forma consagrada:

- (4-22) Fazer tempestade num copo de água
NI: (M_i: (A_i: [(F_i: DECL (F_i)) (P₁)_S (P₂)_A (C_i: [(T_i) (R_i: (T_j) (R_i)) (R_j: (T_k) (T_i) (R_j))_{Foc}](C_i))] (A_i))(M_i))
- NR:** (p_i: (NON-FIN ep_i: (e_i: [(f_i: (f_j: **fazer**_V (f_j)) (e_j: (f_k: **tempestade** (f_k)) (e_j))_U (x_i: (f_i: **copo** (f_i)) (x_i)): (f_i: **água**_N (f_i))_{Loc} (f_i))] (e_i)) (ep_i)) (p_i))
- NM:** (Le_i: (Cl_i: [(Vp_i: (Vw_i: **fazer**_{NON-FIN} (Vw_i)) (Vp_i)) (Np_i: (Nw_i: **tempestade** (Nw_i))_{Obj} (Np_i)) (Adpp_i: (Gw_i: **em** (Gw_i)) (Gw_j: **um** (Gw_j)) (Adpp_i)) (Np_j: (Nw_j: **copo** (Nw_j)) (Np_j)) (Adpp_j: (Gw_k: **de** (Gw_k)) (Adpp_j)) (Np_k: (Nw_k: **água** (Nw_k)) (Np_k))] (Cl_i)) (Le_i))

No NI, *fazer tempestade num copo de água* é representada por 1 Subato Atributivo (T_i: fazer (T_i)) e dois Subatos Referenciais (R_i: tempestade (R_i)) e (R_j: copo de água (R_j)), com função de Foco. Por *tempestade* ser um Subato Referencial independente, justifica-se, assim, a possibilidade de substituição lexical, modificação adjetiva, passivização e de inserção de operador de identificabilidade.

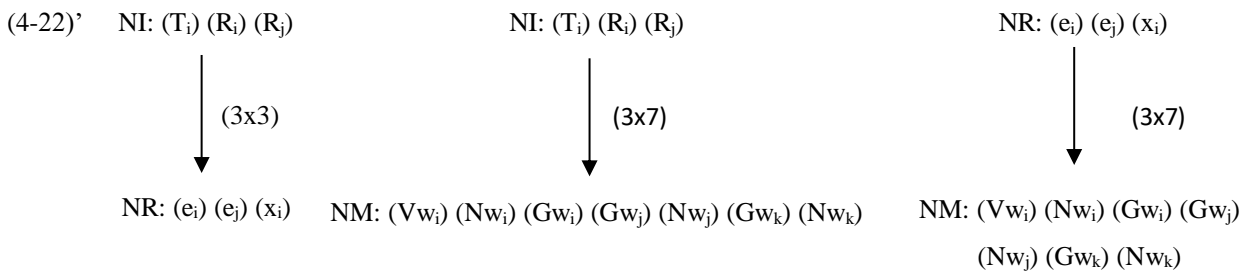
No NR há 1 Estado-de-Coisas (e_i) caracterizado pela propriedade (f_j: fazer (f_j)), e outro Estado-de-Coisas (e_j: (f_k: tempestade (f_k)) (e_j))_U representando *tempestade*, com função de Undergoer (U) em relação ao verbo *fazer*. Representa-se individualmente o termo *tempestade* como um Estado-de-Coisas (e_j) pensando em sua base metafórica de sofrimento, desgraça, preocupação e medo.

Diferente do que fizemos com o termo *sapo*, classificado como um Indivíduo (x_i) na EI do Tipo 5 *engolir sapo*, entendemos que a propriedade semântica de *tempestade* caracteriza-se mais fortemente como um Estado-de-Coisas (e_j: (f_k: tempestade (f_k)) (e_j))_U, por ser uma entidade que pode ser localizada num tempo relativo e espaço, e também pode ser avaliada em termos de realidade, conforme Hengeveld e Mackenzie (2008). Já *copo* representa um Indivíduo (x_i: (f_i: copo (f_i)) (x_i)): (f_i: **água**_N (f_i))_{Loc} modificado pela propriedade (f_i) *água* e que também ativa a função de Locativo (Loc).

Para finalizar, no NM a EI é codificada por sete unidades: 1 Palavra Verbal (Vw_i: fazer (Vw_i)), três Palavras Nominais (Nw_i: tempestade (Nw_i), (Nw_j: copo (Nw_j)) e (Nw_k: água (Nw_k)) e três

Palavras Gramaticais (Gw_i : em (Gw_i)) (Gw_j : um (Gw_j)), i.e. preposição *num* dividida em em + um, e (Gw_k : de (Gw_k)).

Como resultado do alinhamento entre os níveis, temos uma relação de transparência apenas entre os níveis de Formulação, com três unidades em cada nível: 1 Subato Atributivo (T_i : fazer (T_i)) e dois Subatos Referenciais (R_i : tempestade (R_i)) e (R_j : copo de água (R_j)) no NI; e 1 Indivíduo (x_i : (f_i: copo (f_i)) (x_i): (f_i: água_N (f_i))_{Loc} e dois Estados-de-Coisas: (e_i) caracterizado pela propriedade (f_j: fazer (f_j)), e (e_j : (f_k: tempestade (f_k)) (e_j))_U no NR. Em (4-22)' temos seu esquema de alinhamento:



Seguimos agora para a subseção de análise dos resultados finais encontrados na investigação detalhada dos excertos extraídos de seis variedades do português.

4.2: Resultados finais

Após investigação detalhada do estatuto semântico-cognitivo e do comportamento morfossintático das EIs selecionadas, representadas por treze excertos encontrados no CRPC, oriundos de seis variedades do português (sete ocorrências em Portugal, duas ocorrências em Angola, e 1 ocorrência no Brasil, Moçambique, Macau e Cabo Verde), organizamos os resultados qualitativos e quantitativos nesta subseção.

Conforme vimos nas subseções anteriores, analisamos nove EIs – *tirar o cavalo da chuva*, *para inglês ver*, *abandonar o barco*, *levar um puxão de orelhas*, *descascar o abacaxi*, *procurar sarna para se coçar*, *ser a gota de água*, *engolir sapo* e *fazer tempestade num copo de água*, que representam um grupo de trinta e cinco expressões selecionadas, e as classificamos em cinco diferentes tipos, de acordo com os princípios norteadores de **motivação** e de **composicionalidade**, seguindo a proposta dos trabalhos de Keizer (2016, 2018), autora que atuou na supervisão deste trabalho durante o estágio do autor na Universidade de Viena durante seis meses (abril a setembro de 2024).

Observa-se o quadro 7, com os resultados do comportamento morfossintático e com a identificação de descontinuidade, apresentados na análise morfossintática das EIs:

Tipos e exemplos de EI	Critérios de variação morfossintática e resultados de descontinuidade														
	Substituição lexical	Inserção de Pronomes	Inserção de Advérbios	Modificação adjetiva	Número	Quantificação	Definitude	Nominalização	Passivização	Possibilidade de conversão p/ imperativo	Elisão de um termo fixo	Inversão da ordem preferida	Referência anafórica	Referência catafórica	Descontinuidade
Tipo 1 – Desmotivadas e não-decomponíveis	-	±	±	-	-	-	±	-	-	±	±	±	±	+	+
Exemplos (13 EIs)	<i>Tirar o cavalo da chuva, Para inglês ver, Ser a gota de água, Chorar as pitangas, Bater as botas, Inês é morta, Ser pau para toda obra, Acabar em pizza, Quebrar o galho, Untar as mãos, Dourar a pílula, Apertar o cinto, Pagar o pato</i>														
Tipo 2 – Motivadas e não-decomponíveis	+	+	+	+	±	±	±	+	+	+	-	-	+	+	-
Exemplos (8 EIs)	<i>Abandonar o barco, Levam um puxão de orelhas, Dar com a língua nos dentes, Ser um balde de água fria, Pendurar as chuteiras, Acertar na mosca, Começar com o pé direito, Por os pingos nos is</i>														
Tipo 3 – Motivadas e decomponíveis	±	+	+	+	+	+	+	±	+	+	±	±	+	+	-
Exemplos (7 EIs)	<i>Descascar o abacaxi, Procurar sarna para se coçar, Ser amigo da onça, Abrir as portas, Ter a paciência de Jó, Custar os olhos da cara, Como o diabo foge da cruz</i>														
Tipo 5 – Parcialmente motivadas por metáforas conceituais e decomponíveis	+	+	+	+	±	+	+	-	+	+	-	-	+	+	+
Exemplos (7 EIs)	<i>Engolir sapo, Fazer tempestade num copo de água, Chorar lágrimas de sangue, Pôr minhocas na cabeça, Esquentar a cabeça, Quebrar a cabeça, Perder a cabeça</i>														

Quadro 7: Comportamento morfossintático das EIs em português (Elaboração própria, 2025)

Como é possível constatar no quadro 7, as EIs do Tipo 4 *AD-HOC* não entraram nos resultados dos critérios de variação morfossintática, pois analisamos a interpretação idiomática com base nos excertos e a representação nos termos na GDF. Conforme visto nos exemplos, além de derivarem de EIs de outros tipos – no caso, ambos os casos foram EIs do Tipo 1 (*ser a gota de água e tirar o cavalo da chuva*), essas expressões são de finalidade específica, i.e. não possuem aplicação fora dos seus respectivos contextos, o que as levam a não possuírem o estatuto de uma EI consagrada.

Tal característica reforça a impossibilidade de considerarmos tais casos como de substituição lexical, como por exemplo, *levar e dar um puxão de orelhas*, variação na qual nota-se que ambos são EIs consagradas na língua, com registros de uso mantendo a mesma significação idiomática em diferentes ações adaptadas ao contexto: dar e receber um sermão. O mesmo vale para os casos de *idiomatic wordplay* que vimos – *tempestade estival e tempestade mediterrânica*.

Em relação aos resultados dos testes formais de Keizer (2016, 2018), confirma-se que a aplicação dos critérios de variação morfossintática contemplou os oito testes vistos nos trabalhos da autora, e os outros seis elaborados com base na especificidade do português. Tais adições foram feitas com base na análise dos excertos encontrados no CRPC e na *web*, o que nos levou a incluí-los como critérios para análise morfossintática das EIs.

A autora realizou uma maior delimitação nos seus testes de modificação adjetiva, dividindo os resultados em congruentes, incongruentes e relacionados e não-relacionados à EI, resultados que também encontramos, mas não dividimos da mesma maneira. Pensamos que essa tarefa poderá ser desenvolvida em trabalho futuro, com maior aporte teórico focado nos adjetivos.

Realizamos uma breve comparação com os resultados dos oito critérios de variação morfossintática da autora com nossos resultados, sem adentrarmos em pormenores, pois sabemos que se tratam de resultados de línguas diferentes – inglês e português:

- Tipo 1 (EIs desmotivadas e não-decomponíveis): ambos os resultados apresentaram restrição em substituição, modificação adjetiva, passivização, número, quantificação e nominalização;
- Tipo 2 (EIs motivadas e não-decomponíveis): ambos os resultados apresentaram possibilidade de modificação adjetiva e substituição lexical;

- Tipo 3 (EIs motivadas e decomponíveis): ambos os resultados apresentaram possibilidade de anaforicidade, definitude, número, modificação adjetiva e quantificação.

Vejamos abaixo a interpretação dos resultados para cada critério:

(i) Substituição lexical: Encontramos quatro resultados, dos quais três foram aplicados a EIs centradas no Np – (*levar/dar um puxão de orelhas* (Tipo 2), *procurar/buscar sarna para se coçar* (Tipo 3) e *fazer/ser tempestade num copo de água* (Tipo 5)), que entendemos possuir o verbo como um *slot* que permite variação do verbo de acordo com a necessidade do contexto. Apenas *engolir sapo* (Tipo 5) não é uma EI centrada em Np, mas entendemos a proximidade semântica dos verbos (*engolir* e *digerir*), assim como entre *buscar, caçar, desejar e procurar* (*sarna para se coçar*), além de *abandonar o barco/navio*;

(ii) Inserção de pronomes: Todos os tipos permitiram pronomes demonstrativos (e.g. *esse puxão de orelhas*) ou possessivos (e.g. *meu cavalo da chuva*);

(iii) Inserção de advérbios: Todos os tipos permitiram advérbios de tempo (e.g. *abandone agora o barco*), modo (e.g. *descasque depressa esse abacaxi*) ou lugar (*procure para lá sarna para se coçar*);

(iv) Modificação adjetiva: Apenas o Tipo 1 não permitiu, já que consideramos *última* gota em *ser a gota de água* como um caso atípico, estritamente conectado com as cenas literal e idiomática da EI – última gota antes do transbordo e última situação tolerável. Identificamos resultados de adjetivação congruente (e.g. *sapo maior*), incongruente (e.g. *sarna marxista*) e relacionados ao significado idiomático (e.g. *levar uns bons puxões de orelhas*);

(v) Número: Apenas o Tipo 1 não permitiu. Notamos que a inserção de número dialogou também com a significação idiomática da EI, como na EI do Tipo 3 *descascar dois abacaxis* (resolver dois problemas);

(vi) Quantificação: Os Tipos 2, 3 e 5 permitiram casos de quantificação (e.g. *alguns puxões, muitos abacaxis, muito sapo*, respectivamente);

(vii) Definitude: Variação presente em todos os tipos, em alguns casos os artigos – definido ou indefinido já fazem parte da forma consagrada, como em *descascar o abacaxi* e *levar um puxão de orelhas*;

(viii) Nominalização: Encontramos três casos – *abandono do barco* e *levada de puxão de orelhas* (Tipo 2) e *busca/procura/caça por sarna para se coçar* (Tipo 3);

(ix) Passivização: Os tipos 2 (e.g. *o barco foi abandonado*), 3 (e.g. *o abacaxi foi descascado*) e 5 (e.g. *como a tempestade foi feita num copo de água*) permitiram passivização;

(x) Possibilidade de conversão para modo imperativo: Todos os tipos apresentaram possibilidade de conversão para modo imperativo. Houve casos de impossibilidade de conversão para modo imperativo, como por exemplo, *para inglês ver* (Tipo 1), e ao mesmo tempo, outra EI do mesmo tipo, houve tal possibilidade: *tire o cavalo da chuva*, situação que identificamos devido às diferentes ativações semânticas de cada EI;

(xi) Elisão de um termo fixo do predicado: As EIs *procurar sarna para se coçar* (Tipo 3) e *ser a gota de água* (Tipo 1) apresentaram resultados de elisão de um termo fixo: *procurar sarna* e *ser a gota*, respectivamente. Nesses casos, notamos a necessidade do contexto para evitar ambiguidade com outras EIs, conforme explicitado anteriormente. Consideramos a aplicação desse critério para os termos fixos da expressão, incluindo o verbo tratado como parte da forma consagrada, o que nos levou a considerar as EIs centradas em Np com seu respectivo verbo fixo, de acordo com nosso levantamento de dados das EIs;

(xii) Inversão da ordem preferida: Encontramos dois casos de inversão da ordem preferida. Em *procurar sarna para se coçar* (Tipo 3) vimos que a EI foi encontrada no CRPC em descontinuidade entre dois Movimentos e também com sua ordem preferida invertida – *Para se coçar* (M_i) e *deseja mais sarna* (M_j). Já a EI do Tipo 1 *ser a gota de água* foi testada com o dado construído *a gota de água foi/é*, e também vimos um caso extraído do CRPC – *uma última gota de água faltava para fazer transbordar a taça*, em que temos outro verbo – *faltar* ocupando a posição posterior ao Np da EI, em uma situação de adaptação ao contexto e à necessidade do falante;

(xiii) Referência anafórica: Todos os tipos apresentaram possibilidade de referência anafórica pela entidade previamente introduzida no discurso (e.g. *A reforma foi feita apenas para inglês ver. Isso é comum aqui*), acionadas por pronomes, e por parte da EI. Apenas as EIs *para inglês ver e tirar cavalo da chuva* (Tipo 1) apresentaram impossibilidade de referência anafórica por parte da EI, mas a EI do mesmo tipo *ser a gota de água* apresentou tal possibilidade (e.g. *aquela gota de água*);

(xiv) Referência catafórica: Todos os tipos apresentaram possibilidade de referência catafórica pela entidade posteriormente introduzida no discurso, acionadas por pronomes, fato previsível, assim como os casos de referência anafórica (e.g. *A menos que, face ao avolumar das dificuldades, queira repetir, agora, o que fez há 15 anos: abandonar o barco*). Também encontramos três resultados de referência catafórica por parte da EI: 1 caso do Tipo 4 *AD-HOC*: *tirar o cavalo da chuva - cavalidade*, um resultado no Tipo 3: *para se coçar – deseja mais sarna*, e um resultado do Tipo 5: *tempestade-minitempestade-bonança*;

(xv) Resultados de descontinuidade: Encontramos quatro resultados de descontinuidade entre forma e significado: dois EIs Tipo 4 *AD-HOC*: *ser a gota de água – balde* (acionado por referência anafórica), e *tirar o cavalo da chuva – cavalidade* (acionado por referência catafórica); 1 resultado do Tipo 3: *procurar sarna para se coçar – para se coçar – deseja mais sarna* (acionado por referência catafórica); e 1 resultado do Tipo 5: *uma tempestade num copo de água – minitempestade/bonança/pequena tempestade/grandes tempestades* (acionados por referências anafóricas e catafóricas entre sete turnos de fala de quatro participantes).

Todos os quatro resultados de descontinuidade ativaram a camada do Discurso no NI, e conseqüentemente, a camada do Texto no NM, pois ocorreram entre Movimentos acionados pelos participantes, o que gerou a necessidade de ampliar o contexto de inserção da EI para compreender todo o andamento do tópico discursivo, e assim apreender a significação idiomática da EI, além da compreensão da carga irônica nos casos das EIs *AD-HOC* evocadas pelos novos elementos – *balde* e *cavalidade*. Tal ampliação e dependência do contexto também nos auxiliou a evitar casos de ambigüidade para as EIs do Tipo 4, conforme subseção 4.1.4.

Conforme vimos, para a análise das nove EIs selecionadas, foram necessários um total de treze excertos extraídos do CRPC, pois repetimos as EIs *procurar sarna para se coçar* (duas vezes) e *fazer tempestade num copo de água* (três vezes). As ocorrências encontradas na variação

morfossintática foram extraídas do CRPC, na *web*, além de dados construídos, ambos devidamente identificados.

Dentre os quatorze critérios de variação morfossintática aplicados às nove EIs diferentes, nota-se que os critérios de referências anafórica e catafórica foram os mais produtivos em termos de compatibilidade, apresentando resultados possíveis em todas as nove EIs. Os menos produtivos foram os critérios de inversão da ordem preferida e elisão de um termo fixo do predicado, com dois resultados para cada critério.

Ao final, concluímos que EIs dos Tipos 2 (motivadas e não-decomponíveis) e 3 (motivadas e decomponíveis) foram as mais produtivas em termos de variações morfossintáticas possíveis: doze (12) (substituição lexical, inserção de pronomes, inserção de advérbios, modificação adjetiva, número, quantificação, definitude, nominalização, passivização, possibilidade de conversão para o modo imperativo, referência anafórica e referência catafórica) e todos os critérios permitidos, respectivamente.

As EIs do Tipo 5 também foram altamente produtivas – 11 critérios permitidos (substituição lexical, inserção de pronomes, inserção de advérbios, modificação adjetiva, número, quantificação, definitude, passivização, possibilidade de conversão para o modo imperativo, referência anafórica e referência catafórica). Já as EIs do Tipo 1 (desmotivadas e não-decomponíveis) foram, naturalmente, as menos produtivas: oito critérios permitidos (inserção de pronomes, inserção de advérbios, definitude, possibilidade de conversão para o modo imperativo, elisão de um termo fixo do predicado, inversão da ordem preferida, referência anafórica e referência catafórica).

Com isso, confirmamos o que vimos nos estudos de Keizer (2016, 2018): a importância dos fatores de motivação e composicionalidade das EIs como determinantes para entender as variações morfossintáticas das expressões como motivadas funcionalmente. Assim, comprovamos a aplicabilidade do conjunto de primitivos chamado **CombPIF** – Combinação de quadros parcialmente instanciados, agora com resultados de EIs no português.

Assim como os *idioms* da língua inglesa, as EIs de língua portuguesa oriundas de diferentes variedades se comportaram como entradas parcialmente seladas, combinadas em quadros interpessoais ou representacionais e armazenadas no Léxico do português.

Com os casos comprovados de variação morfossintática, vemos que os *slots* abertos de cada EI permitiram a inserção de elementos lexicais voltados ao contexto (e.g. *descascar o abacaxi carioca*), ao significado da cena literal da EI (e.g. *engolir um sapo maior*) e da cena idiomática (e.g. *autêntico puxão de orelhas*) e à cultura (e.g. *sarna marxista*). Os cinco tipos de EIs definiram o grau

de produtividade das EIs com base em seu funcionamento como expressões semifixas parcialmente imprevisíveis (lexicalmente) e parcialmente produtivas (gramaticalmente).

Agora seguimos com os resultados de alinhamento e transparência obtidos com as representações formais nos níveis da GDF, conforme quadro 8:

Exemplos analisados	Níveis da GDF			Transparência entre níveis	
	Interpessoal (NI)	Representacional (NR)	Morfossintático (NM)		
Tipo 1	Tirar o cavalo da chuva	1 Subato Atributivo (T: tirar o cavalo da chuva)	1 Estado-de-Coisas (e: tirar o cavalo da chuva)	1 Palavra Verbal (Vw: tirar), 2 Palavras Nominais (Nw: cavalo e chuva), 3 Palavras Gramaticais (Gw: o, de e a)	NI ↔ NR (1x1)
	Para inglês ver	1 Subato Atributivo (T: para inglês ver)	1 Estado-de-Coisas (e: para inglês ver)	1 Palavra Verbal (Vw: ver), 1 Palavra Nominal (Nw: inglês), 1 Palavra Gramatical (Gw: para)	NI ↔ NR (1x1)
	Ser a gota de água	1 Subato Referencial (R: gota de água), 1 operador de identificabilidade (id: a) π	1 Estado-de-Coisas (e: gota de água), 1 operador de quantidade (1: a) π	1 Palavra Verbal (Vw: ser), 2 Palavras Nominais (Nw: gota e água), 2 Palavras Gramaticais (Gw: a e de)	NI ↔ NR (2x2)
Tipo 2	Abandonar o barco	1 Subato Atributivo (T: abandonar), 1 Subato Referencial (R: barco), 1 operador de identificabilidade	1 Estado-de-Coisas (e: abandonar), 1 Indivíduo (x: barco), 1 operador de quantidade (1: o) π	1 Palavra Verbal (Vw: abandonar), 1 Palavra Nominal (Nw: barco), 1 Palavra Gramatical (Gw: o)	Entre todos os níveis

		ade (id: o) π			
	Levar um puxão de orelhas	1 Subato Atributivo (T: levar), 2 Subatos Referenciais (R: puxão e orelhas), 1 operador de identificabilidade (id: um) π	2 Estados-de-Coisas (e: levar e puxão), 1 Indivíduo (x: orelhas), 1 operador de quantidade (1: um) π	1 Palavra Verbal (Vw: levar), 2 Palavras Nominais (Nw: puxão e orelhas), 2 Palavras Gramaticais (Gw: um e de)	NI \leftrightarrow NR (4x4)
Tipo 3	Descascar o abacaxi	1 Subato Atributivo (T: descascar), 1 Subato Referencial (R: abacaxi), 1 operador de identificabilidade (id: o) π	2 Estados-de-Coisas (e: descascar e abacaxi), 1 operador de quantidade (1: o) π	1 Palavra Verbal (Vw: descascar), 1 Palavra Nominal (Nw: abacaxi), 1 Palavra Gramatical (Gw: o)	Entre todos os níveis
	Procurar sarna para se coçar	2 Subatos Atributivos (T: procurar e coçar), 2 Subatos Referenciais (R: sarna e se)	3 Estados-de-Coisas (e: procurar, sarna e coçar), 1 Indivíduo (x: se)	2 Palavras Verbais (Vw: procurar e coçar), 1 Palavra Nominal (Nw: sarna), 2 Palavras Gramaticais (Gw: para e se)	NI \leftrightarrow NR (4x4)
Tipo 4	<i>Cavalidade</i> (tirar o cavalo da chuva)	1 Subato Atributivo (T: cavalidade)	1 Estado-de-Coisas (e: cavalidade)	1 Palavra Nominal (Nw: cavalidade)	Entre todos os níveis
	<i>Balde</i> (ser a gota de água)	1 Subato Referencial (R: balde)	1 Indivíduo (x: balde)	1 Palavra Nominal (Nw: balde)	Entre todos os níveis

Tipo 5	Engolir sapo	1 Subato Atributivo (T: engolir), 1 Subato Referencial (R: sapo)	1 Estado-de-Coisas (e: engolir), 1 Indivíduo (x: sapo)	1 Palavra Verbal (Vw: engolir), 1 Palavra Nominal (Nw: sapo)	Entre todos os níveis
	Fazer tempestade num copo de água	1 Subato Atributivo (T: fazer), 2 Subatos Referenciais (R: tempestade e copo de água)	2 Estados-de-Coisas (e: fazer e tempestade), 1 Indivíduo (x: copo de água)	1 Palavra Verbal (Vw: fazer), 3 Palavras Nominais (Nw: tempestade, copo e água), 3 Palavras Gramaticais (Gw: em, um e de)	NI ↔ NR (3x3)

Quadro 8: Representação do alinhamento das EIs selecionadas (Elaboração própria, 2025)

Listamos abaixo os resultados do alinhamento entre os níveis da GDF para cada Tipo de EI e suas relações:

(i) Tipo 1 (desmotivadas e não-decomponíveis): encontramos apenas 1 tipo de alinhamento transparente entre níveis – NI-NR. Identificamos que a EI é representada por ao menos 1 Subato Atributivo (T) ou 1 Subato Referencial (R) no NI, 1 Estado-de-Coisas (e) no NR, 1 Palavra Verbal (Vw) e 1 Palavra Nominal (Nw) no NM;

(ii) Tipo 2 (motivadas e não-decomponíveis): encontramos dois tipos de alinhamento transparente entre níveis – NI-NR e entre todos os níveis. Vimos que a EI é representada por ao menos 1 Subato Atributivo (T) e 1 Subato Referencial (R) no NI, 1 Indivíduo (x) e 1 Estado-de-Coisas (e) no NR, 1 Palavra Verbal (Vw), 1 Palavra Nominal (Nw) e 1 Palavra Gramatical (Gw) no NM;

(iii) Tipo 3 (motivadas e decomponíveis): apresentou dois tipos de alinhamento transparente entre níveis – NI-NR e entre todos os níveis. A expressão é representada por ao menos 1 Subato Atributivo (T) e 1 Subato Referencial (R) no NI, 2 Estados-de-Coisas (e) no NR, 1 Palavra Verbal (Vw), 1 Palavra Nominal (Nw) e 1 Palavra Gramatical (Gw) no NM;

(iv) Tipo 4 (*AD-HOC* desmotivadas e não-decomponíveis): apresentou 1 tipo de alinhamento transparente entre níveis – todos os níveis. A EI é representada por ao menos 1 Subato Atributivo (T) ou 1 Subato Referencial (R) no NI, 1 Estado-de-Coisas (e) ou 1 Indivíduo (x) no NR, e 1 Palavra Nominal (Nw) no NM, a depender do tipo elemento *AD-HOC* introduzido na EI e qual o elemento retirado da forma consagrada;

(v) Tipo 5 (parcialmente motivadas por metáforas conceituais e decomponíveis): apresentou dois tipos de alinhamento transparente entre níveis – NI-NR e entre todos os níveis. Vimos que a EI é representada por ao menos 1 Subato Atributivo (T) e 1 Subato Referencial (R) no NI, 1 Indivíduo (x) e 1 Estado-de-Coisas (e) no NR, 1 Palavra Verbal (Vw) e 1 Palavra Nominal (Nw) no NM;

(vi) Relação entre os tipos: Notamos que todos os tipos apresentaram alinhamento transparente, ao menos entre os Níveis de Formulação (Interpessoal e Representacional). Outra situação comum é que, se um operador lexicalmente marcado e parte fixa da EI é representado no NI, o mesmo será representado no NR e codificado como uma Palavra Gramatical (Gw) no NM. Ressaltamos que o alinhamento entre os níveis de Formulação e o NM varia de acordo com a extensão lexical de cada EI, independente do seu tipo. Outro fato comum aos cinco tipos de EIs é que todas apresentaram função Foco no NI, situação que se justifica por sua natureza expressiva e uso intencional do falante para expressar suas intenções comunicativas, como por exemplo, os casos de ironia nos debates políticos.

Novamente comparamos os resultados do inglês encontrados nos trabalhos de Keizer (2016) com nossas ocorrências em variedades lusófonas do português. A EI do Tipo 1 *to kick the bucket* foi representada por 1 Subato Atributivo (T_i: *kick the bucket* (T_i)) no NI, 1 Estado-de-Coisas (e_i: *kick the bucket* (e_i)) no NR, 1 Palavra Verbal (Vw_i: *kick* (Vw_i)), 1 Palavra Nominal (Nw_i: *bucket* (Nw_i)) e 1 Palavra Gramatical (Gw_i: *the* (Gw_i)) no NM, resultando em alinhamento transparente apenas entre NI-NR.

Para a EI do Tipo 2 *to hit the ceiling*, motivada e não-decomponível, a autora definiu sua representação no NI como 1 Subato Atributivo (T_i: *hit the ceiling* (T_i)); 1 Estado-de-Coisas (e_i: *hit* (e_i)), 1 Indivíduo (x_i: *ceiling* (x_i)) e 1 operador de quantidade (1: *the*) no NR, 1 Palavra Verbal (Vw_i: *hit* (Vw_i)), 1 Palavra Nominal (Nw_i: *ceiling* (Nw_i)) e 1 Palavra Gramatical (Gw_i: *the* (Gw_i)) no NM. O alinhamento transparente ocorreu entre NR-NM.

O Tipo 3, exemplificado com a EI *to spill the beans* é marcado por 1 Subato Atributivo (T_i : *spill* (T_i)), 1 Subato Referencial (R_i : *beans* (R_i)) e 1 operador de identificabilidade (id: *the*) no NI; 1 Estado-de-Coisas (e_i : *spill* (e_i)), 1 Conteúdo proposicional (p_i : *beans* (p_i)) e 1 operador de quantidade (m : *bean*) no NR; 1 Palavra Verbal (Vw_i : *spill* (Vw_i)), 1 Palavra Nominal (Nw_i : *beans* (Nw_i)) e 1 Palavra Gramatical (Gw_i : *the* (Gw_i)) no NM. O alinhamento ocorre entre os três níveis.

Comparando os resultados temos: Tipo 1 apresentou o mesmo alinhamento transparente nas EIs das duas línguas (NI-NR); Tipo 2 apresentou diferentes tipos de alinhamento, e Tipo 3 apresentou casos com mesmo alinhamento – todos os níveis. No artigo de 2018, a autora trouxe dois termos *AD-HOC* para duas EIs do Tipo motivadas e não-decomponíveis: *to bite the bullet* – *to gnaw the bullet* e *to open the floodgates* – *to close the floodgates*.

Finalizando essa etapa de comparação entre resultados, nota-se que os termos *AD-HOC* são de classes gramaticais diferentes e portanto, acionam diferentes propriedades nos níveis da GDF: os verbos do inglês *to gnaw* e *to close* são Subatos Atributivos (T) no NI, Estado-de-Coisas (e) no NR e Palavras Verbais (Vw) no NM. Já os termos *AD-HOC cavilidade* e *balde* são, respectivamente, um Subato Atributivo (T) e um Subato Referencial (R) no NI; um Estado-de-Coisas (e) e um Indivíduo (x) no NR e ambos uma Palavra Nominal (Nw) no NM.

Já a proposta do novo tipo de EI (Tipo 5) contribuiu para compreensão mais detalhada dos diferentes graus de motivação e composicionalidade das EIs, pois exige uma análise semântico-cognitiva dos termos que ativam a interpretação metafórica, através das imagens metafóricas esquemáticas, e todo o seu universo conceitual e convencional.

Tal proposta tipológica de refinamento no grau de motivação e composicionalidade reforça a possibilidade de um maior desmembramento das estruturas esquemáticas de cada EI com base nas metáforas conceituais, a compreensão de seu caráter motivacional parcial, focando no elemento semântico autônomo de cada EI, que pode contribuir para o estudo das metáforas universais e quase-universais entre as línguas através da materialização linguística desse fenômeno cultural nas EIs.

Notamos que há uma troca cultural entre línguas por meio de metáforas materializadas nas EIs, como vimos nos exemplo da língua inglesa, *to take the cake* (variedade americana) e *to take the biscuit* (variedade britânica), entre variedades do português, e.g. *acertar na mosca*, *vestir a camisa*, *chorar as pitangas* (variedades brasileiras) e *acertar na muche*, *vestir a camisola*, *chorar lágrimas de sangue* (variedades portuguesas), EI de origem contextual no Brasil e utilizada em outras variedades, e.g. *para inglês ver*, imagens metafóricas partilhadas em diversas culturas, e.g. **tempestade** com o mesmo significado idiomático, e.g. *fazer tempestade em um copo de água* (português), *to make a storm in a teacup* (inglês), *Wer Wind sät, wird Sturm ernten* (alemão), *hacer*

una tempestad en un vaso de agua (espanhol).

Também vimos que o termo **sapo** pode apresentar diferentes significados idiomáticos, e.g. *engolir sapo* (português), *I no naka no kawazu tai kai o Shirazu* (japonês, que significa ignorância), *to have a frog in one's throat* (inglês, que significa rouquidão), *gaeguri olchaengijeok saenggak motanda* (coreano, que significa arrogância).

O termo **cabeça** também foi identificado em EIs do Tipo 5, conforme quadro 7, com variados significados na própria língua e em comparação à outras línguas, e.g. *pôr minhocas na cabeça*, *esquentar a cabeça*, *quebrar a cabeça*, *perder a cabeça* em português, *to lose one's head* (perder a paciência), *to keep one's head* (manter a calma), *to bang the head against the wall* (frustração) em inglês, *Féliül a feje* (ser orgulhoso), *Kérdez a fejemet* (estar confuso), *Fejét vágja* (estar nervoso), *Szépen, egy szál feje* (sem bom senso) em húngaro, etc.

Seguimos para as considerações finais da pesquisa e na sequência encerramos esta tese de doutorado com as referências bibliográficas utilizadas.

Considerações finais

Pode-se afirmar que duas perguntas de pesquisa perseguem esse pesquisador desde o início do doutorado, nutridas pelo que foi pesquisado sobre o universo da Fraseologia em trabalho de conclusão de curso da graduação: (i) Como se dá a caracterização das EIs encontradas em variedades lusófonas do português sob a óptica da GDF? (ii) Como avaliar os possíveis graus de transparência/opacidade dessas expressões?

Com o decorrer do doutorado e a oportunidade do estágio no exterior, sob supervisão da Profa. Dra. Evelien Keizer, o autor participou da disciplina intitulada *English idioms* e após novas leituras, foi definida uma nova pergunta de pesquisa, visando refinar a compreensão sobre o universo metafórico das EIs: (iii) Como as interpretações metafóricas das EIs podem ajudar no processo de ativação semântica e composicionalidade de suas partes?

Tais perguntas de pesquisa serviram de base para a proposta dos objetivos deste trabalho: o **objetivo geral** foi investigar o estatuto semântico-cognitivo das EIs encontradas em variedades lusófonas do português, utilizando como princípios norteadores, a motivação, para análise cognitiva, e a composicionalidade, para análise morfossintática, a fim de definir graus de transparência e opacidade das EIs e uma tipologia para o português, conforme (Keizer, 2016).

Já os **objetivos específicos** foram: (i) a verificação dos processos de esforço extra-cognitivo do falante para reanálise das EIs AD-HOC; (ii) a averiguação do alinhamento da representação formal das EIs de acordo com os três níveis mais altos da GDF (Hengeveld; Mackenzie, 2008), e (iii) determinação dos apontamentos de violação da relação de um-para-um entre forma e significado dos elementos constituintes das EIs (Leufkens, 2015).

Ao passo que as leituras foram desenvolvidas, tanto para o aporte teórico como para o levantamento no CRPC, os objetivos de pesquisa delimitaram as seguintes **hipóteses**: (i) possibilidade de encontrar no português os quatro tipos de EIs que Keizer (2016, 2018) classificou no inglês, tendo como norte de investigação os princípios de motivação e composicionalidade; (ii) as EIs do português podem apresentar resultados de variação morfossintática similares aos vistos no trabalho da autora, além de outros resultados específicos, visto que o português pode apresentar alguma especificidade com relação à estruturação de EIs, o que nos levou a pensar na possibilidade de um novo tipo de EI; (iii) a possibilidade de ocorrências em referências anafórica e catafórica apresentarem algum tipo de violação da relação entre forma e significado, pontos entrelaçados com a não-transparência dessas expressões, rompendo, possivelmente, a relação um-para-um entre níveis

da GDF e também entre forma e significado, dado que contribui para a opacidade dessas expressões vistas nessas situações, além do surgimento de EIs *AD-HOC*, que entendemos poder haver uma dependência de contexto específico para atestação, e (iv) a possibilidade de encontrarmos algum alinhamento padrão entre os níveis da GDF para os tipos de EI, situação vista no trabalho de Keizer (2016).

Após diversas modificações na fundamentação teórica, a **metodologia de pesquisa** foi estabelecida em três etapas e seus critérios. A primeira etapa - **análise cognitiva das EIs**, foi organizada com os seguintes critérios de análise: a definição dicionarizada da EI em sua forma consagrada; interpretação do contexto de inserção da EI, e interpretação metafórica das EIs.

A segunda etapa foi a **análise morfossintática das EIs**, baseada na aplicação dos quatorze critérios de variação morfossintática: substituição lexical, modificação adjetiva, quantificação, número, definitude, passivização, nominalização, referência anafórica, referência catafórica, inserção de advérbios, inserção de pronomes, alteração da ordem preferida, elisão de um termo fixo, possibilidade de conversão para modo imperativo. A terceira etapa estabelecida foi a **análise das EIS de acordo com os níveis e camadas da GDF**, tendo como critérios a investigação dos aspectos pragmáticos, semânticos e morfossintáticos.

Essa metodologia foi pensada para fornecer recursos e ferramentas para a compreensão do comportamento morfossintático e do estatuto semântico-cognitivo das EIs, com respaldo teórico à inclusão de diferentes variedades do português, além de dados estatísticos para a seleção das EIs para a análise.

Essa proposta de análise linguística tipológica só foi possível graças aos recursos e ferramentas que a GDF oferece, que nos permitiu aplicar essas etapas e critérios de análise, através dos quais pudemos compreender, de modo preciso, todo o percurso semântico-cognitivo pelo qual as EIs passaram durante os processos de variação morfossintática, por meio da investigação entre os níveis e camadas da teoria, das relações de escopo entre as partes componentes das EIs, e de como se deu sua representação formal no português. Além disso, foi possível compreender os graus de transparência e opacidade entre os níveis, de acordo com a relação sistemática entre funções pragmáticas e semântico-cognitivas que influenciaram o comportamento morfossintático das EIs e sua representação na gramática do português.

Com essa proposta metodológica, foi possível compreender como os princípios norteadores de motivação e composicionalidade atuaram no percurso de funcionamento da relação forma-

significado entre as formas consagrada e variada das EIs.

Ao final da investigação exaustiva do aporte teórico, coleta de dados nos bancos de dados de EIs e nos materiais de busca de EIs, e aplicação teórica da GDF, alcançamos os seguintes **resultados finais**, que seguem aqui estruturados:

- Ocorrências de EIs no português conforme os 4 tipos de EIs vistos em Keizer (2016, 2018);
- Surgimento de um quinto tipo de EI – parcialmente motivada por metáforas conceituais e decomponíveis, compreendida como uma contribuição que amplia a investigação de diferentes graus de motivação e composicionalidade das EIs, que variam entre motivadas e desmotivadas, e decomponíveis e não-decomponíveis, já que esse novo tipo aprofunda o olhar investigativo para os processos de interpretação metafórica das EIs, considera as metáforas conceituais de cada língua, aborda um maior detalhamento da composição dos elementos e sua autonomia no processo de interpretação metafórica, além de levar em conta os processos de criação e de decodificação baseados nas metáforas culturais universais ou quase-universais, o que pode ser útil para outras línguas que partilham as mesmas metáforas conceituais, já que mostram como sua materialização nas EIs pode ser produtiva e avaliada em graus de motivação e composicionalidade;
- Nova delimitação das expressões *AD-HOC*, com dependência do contexto, encontradas em referência anafórica e catafórica pela entidade introduzida previamente/posteriormente no discurso, acionando a camada do Discurso no NI e a camada do Texto no NM. Também realizamos a distinção entre os casos *AD-HOC* com casos de substituição lexical em EIs centradas no Np com *slot* do verbo aberto à adaptações de contexto e 1 caso de *Idiomatic wordplay*;
- Aplicabilidade dos critérios morfossintáticos: dentre os seis novos critérios de variação morfossintática propostos, todos apresentaram resultados, sendo eles: pronomes, advérbios, referência catafórica, possibilidade de conversão para modo imperativo, elisão de um termo fixo do predicado e inversão da ordem preferida. Considerando os oito testes vistos em Keizer (2016) e os respectivos resultados classificatórios das EIs entre (+) e (-) produtivas, encontramos resultados semelhantes para os tipos 1 (menos produtivos) e tipo 3 (mais produtivos) em ambos os trabalhos;

- Análise da transparência das EIs de acordo com a representação formal na GDF com padrão de alinhamento entre os cinco tipos: todos os tipos apresentaram ao menos, transparência entre NI-NR. Em comparação à análise de Keizer (2016), novamente encontramos similaridades nos resultados envolvendo os Tipos 1 e 3, além de todas as EIs serem focais no NI nas duas línguas;
- Foram encontrados quatro resultados de descontinuidade entre forma e significado, ativados por referências anafórica e catafórica, identificados na camada do Discurso (D) no NI;
- Atestação da aplicabilidade do conjunto de primitivos CombPIF – Combinação de quadros parcialmente instanciados (Keizer, 2016), agora com resultados de EIs de variedades lusófonas do português, o que pode contribuir para a ampliação de novos horizontes de investigação sobre a interpretação metafórica das EIs com base no CombPIF proposto por Keizer (2016), o que significa que o processo de armazenamento das EIs no Léxico como entradas parcialmente seladas passaria por um maior refinamento de análise do processo de criação e decodificação metafórica dos elementos;
- Atestação da GDF (Hengeveld; Mackenzie, 2008) como ferramenta eficaz para esse tipo de análise tipológica das EIs do português. Vimos que a abordagem tipológica da GDF nos ajudou a entender um fenômeno linguístico complexo – as EIs – de forma estruturada e ordenada, e nos permitiu associá-las à **combinação semifixa** de quadros interpessoais e representacionais e a modelos morfossintáticos que especificaram suas restrições formais e funcionais.

Esses resultados finais estão refletidos no quadro abaixo, anteriormente apresentado na metodologia, repetido aqui por conveniência, mas agora adaptado com as respostas às perguntas de pesquisa e à hipótese do estudo. Ademais, os objetivos gerais e específicos levantados na metodologia foram retomados e discutidos, conforme as contribuições alcançadas pela pesquisa. Além disso, o percurso da pesquisa é justificado, com base nos resultados encontrados:

FUNDAMENTAÇÃO METODOLÓGICA DA PESQUISA – PERGUNTAS E RESPOSTAS FINAIS

Problemáticas	<p>1- É possível verificar uma relação sistemática entre funções pragmáticas e semântico-cognitivas que influenciam o comportamento morfossintático de EIs e sua representação na gramática das línguas?</p> <p>– Sim, vimos que a relação entre motivação a composicionalidade das EIs foi fundamental para a compreensão do funcionamento e representação das EIs no português, desde a formulação até a codificação, passando por processos de natureza pragmática, semântica, cognitiva e morfossintática, resultando na classificação tipológica vista nos resultados.</p>	
	<p>2- Como se dá a caracterização das EIs em diferentes variedades lusófonas do português, sob a óptica da Gramática Discursivo-Funcional (GDF)?</p> <p>– Vimos que as EIs do português foram definidas em cinco tipos de EIs - Tipo 1 (desmotivadas e não-decomponíveis), Tipo 2 (motivadas e não-decomponíveis), Tipo 3 (motivadas e decomponíveis), Tipo 4 (<i>AD-HOC</i> desmotivadas e não-decomponíveis) e Tipo 5 (parcialmente motivadas por metáforas conceituais e decomponíveis), com base nos trabalhos de Keizer (2016, 2018), sob o aporte teórico da GDF, ferramenta que nos deu recursos para uma abordagem tipológica baseada nas relações de escopo vistas nos processos de Formulação e Codificação da teoria.</p>	
	<p>3- Como avaliar possíveis graus de transparência/opacidade dessas expressões em situações de uso?</p> <p>– Avaliamos o comportamento morfossintático das EIs com base nos princípios norteadores de motivação e composicionalidade, o que nos outorgou mensurar o estatuto semântico-cognitivo e como se deu sua representação morfossintática na língua, através dos critérios de análise que definimos para essa pesquisa- análises cognitiva, morfossintática e nos níveis e camadas da GDF.</p>	
	<p>4- Como as interpretações metafóricas de EIs podem ajudar no processo de ativação semântica (Langlotz, 2006) e composicionalidade de suas partes?</p> <p>– O processo de interpretação metafórica das EIs permitiu um maior detalhamento da composição dos elementos e sua autonomia no processo de interpretação metafórica, além de levar em conta os processos de criação e decodificação baseados nas metáforas culturais universais ou quase-universais, assim como a importância de sua materialização na língua, o que nos concedeu a possibilidade de uma avaliação em graus de motivação e composicionalidade.</p>	
Objetivos	Geral	<p>1- Investigar o estatuto semântico-cognitivo das EIs encontradas em variedades lusófonas do português, utilizando como princípios norteadores, a motivação e a composicionalidade a fim de definir graus de transparência e opacidade das EIs e propor uma tipologia para o português.</p>
		<p>– Essa investigação resultou na definição dos cinco tipos de EIs para o português com base nos graus de transparência e opacidade, que variaram entre EIs mais produtivas (mais transparentes) e menos produtivas (mais opacas).</p>

Específicos	<p>1-Verificação dos processos de esforço extra-cognitivo do falante para reanálise das EIs AD-HOC.</p> <p>- Verificamos que as EIs <i>AD-HOC</i> apresentaram dependência de contexto, e foram encontradas em referência anafórica e catafórica pela entidade introduzida previamente/posteriormente no discurso, acionando a camada do Discurso no NI e a camada do Texto no NM.</p> <hr/> <p>2-Averiguação do alinhamento da representação formal das EIs de acordo com os três níveis mais altos da GDF.</p> <p>- Averiguamos um padrão de alinhamento entre os cinco tipos: todos os tipos apresentaram ao menos, transparência entre NI-NR, além de todas as EIs serem focais no NI.</p> <hr/> <p>3-Determinação dos apontamentos de violação da relação de um-para-um entre forma e significado dos elementos constituintes das EIs.</p> <p>- Encontramos quatro resultados de descontinuidade entre forma e significado, ativados por referências anafórica e catafórica, identificados na camada do Discurso (D) no NI.</p>
Hipóteses	<p>1-Possibilidade de encontrar no português os quatro tipos de EIs que Keizer (2016, 2018) classificou no inglês, tendo como norte de investigação os princípios de motivação e composicionalidade.</p> <p>- Hipótese confirmada com os resultados das EIs do Tipo 1 (desmotivadas e não-decomponíveis), Tipo 2 (motivadas e não-decomponíveis), Tipo 3 (motivadas e decomponíveis) e Tipo 4 (<i>AD-HOC</i> desmotivadas e não-decomponíveis).</p> <hr/> <p>2- As EIs do português podem apresentar resultados de variação morfossintática similares aos vistos no trabalho da autora, além de outros resultados específicos, visto que o português pode apresentar alguma especificidade com relação à estruturação de EIs, o que nos levou a pensar na possibilidade de um novo tipo de EI.</p> <p>- Hipótese confirmada com a aplicação dos critérios morfossintáticos: dentre os seis novos critérios de variação morfossintática propostos, todos apresentaram resultados, sendo eles: pronomes, advérbios, referência catafórica, possibilidade de conversão para modo imperativo, elisão de um termo fixo do predicado e inversão da ordem preferida, além da definição do quinto tipo de EI - parcialmente motivadas por metáforas conceituais e decomponíveis.</p> <hr/> <p>3- Possibilidade de ocorrências em referências anafórica e catafórica apresentarem algum tipo de violação da relação entre forma e significado, pontos entrelaçados com a não-transparência dessas expressões, rompendo, possivelmente, a relação um-para-um entre níveis da GDF e também entre forma e significado, dado que contribui para a opacidade dessas expressões vistas nessas situações, além do surgimento de EIs <i>AD-HOC</i>, que entendemos poder haver uma dependência de contexto específico para atestação.</p> <p>- Hipótese confirmada através dos quatro resultados de descontinuidade entre forma e significado, ativados por referências anafórica e catafórica, identificados na camada do Discurso (D) no NI.</p>

	<p>4- Possibilidade de encontrarmos algum alinhamento padrão entre os níveis da GDF para os tipos de EI, situação vista no trabalho de Keizer (2016).</p> <p>- Hipótese confirmada após a verificação de um padrão de alinhamento entre os cinco tipos: todos os tipos apresentaram ao menos, transparência entre NI-NR. Em comparação à análise de Keizer (2016), encontramos similaridades nos resultados envolvendo os Tipos 1 e 3, além de todas as EIs serem focais no NI nas duas línguas.</p>	
Metodologia	Etapas	Aplicabilidade na pesquisa
	1-Análise cognitiva das EIs	Essa etapa nos permitiu compreender o universo formal da EI: sua dicionarização, sua aplicação nos contextos dos excertos e sua interpretação metafórica.
	2-Análise morfossintática das EIs	A segunda etapa de análise nos permitiu compreender o comportamento morfossintático das EIs através de catorze critérios de variação morfossintática: 1-Substituição lexical; 2-Inserção de pronomes; 3-Inserção de advérbios; 4-Modificação adjetiva; 5-Número; 6-Quantificação; 7-Definitude; 8-Nominalização; 9-Passivização; 10-Possibilidade de conversão para modo imperativo; 11-Elisão de um termo fixo; 12-Inversão da ordem preferida; 13-Referência anafórica; 14-Referência catafórica. Os resultados nos auxiliaram na abordagem tipológica que propomos, de acordo com os graus de transparência e opacidade das EIs.
	3-Análise das EIs de acordo com os níveis e camadas da GDF	O aporte teórico da GDF foi de fundamental importância para a análise dos aspectos pragmáticos, semânticos e morfossintáticos das EIs, etapa que nos permitiu entender padrões de alinhamento transparentes e opacos entre os três níveis mais altos da teoria.

Quadro 9: Fundamentação metodológica da pesquisa - perguntas e respostas finais (Elaboração própria, 2025)

A conclusão a que se chega, a essa altura, é que ainda existe um leque de opções para novos trabalhos em continuidade a esta tese, dentre eles estão: a exploração dos processos de gramaticalização das EIs, o agrupamento das EIs por gênero, como o de debate político e seu estudo por um viés sociofuncional. Também se abre possibilidade para a investigação dessas expressões no NF, utilizando material auditivo, ou para estudo da ironia expressa no uso das EIs (uso recorrente em nossos excertos). Ainda é possível pensar em um estudo comparativo entre EIs do inglês e do português. Ademais, há possibilidade de investigação das EIs centradas no Np, ou de propor um estudo detalhado da modificação adjetiva das EIs, ou ainda se pode pensar na investigação de casos de parentetização com EIs, ou de sua dicionarização, ou mesmo de investigação de processos de

aquisição dessas expressões, utilizando ferramentas da Inteligência Artificial para ensino e pesquisa, atualização de sistemas dicionarizados, etc.

Enfim, o campo das expressões idiomáticas é realmente fascinante, convidativo e desafiador. E dá muito pano pra manga...

REFERÊNCIAS BIBLIOGRÁFICAS

ALMEIDA, J. *Dicionário aberto de calão e expressões idiomáticas usuais na Língua Portuguesa do Brasil*. Portugal: Uminho, 2022.

BAKER, A. E.; HENGEVELD, K. *Linguistics*. Library of Congress Cataloging-in-Publication Data. New Delhi: Willey-Blackwell Publishing, 2012.

BALLY, C. *Traité de stylistique française*. Paris: Klincksieck. v. 1. 1909-1951.

BECHARA, E. *Moderna gramática da Língua Portuguesa*. Rio de Janeiro: Lucerna, 2009.

BIDERMAN, M. T. C. *Teoria linguística: teoria lexical e linguística computacional*. 2. ed. São Paulo: Martins Fontes, 2001.

BONVINI, E. *Línguas africanas e Português falado no Brasil*. Revista África no Brasil. 2015. Disponível em: <
http://www.educadores.diaadia.pr.gov.br/arquivos/File/2010/artigos_teses/2010/Lingua_Portugues/dissertacao/Aslinguas_escravos_brasil.pdf>.

BUTLER, C. S.; HENGEVELD, K. Functionalism. In: *Theories and methods in linguistics* (Wörterbücher der Sprach-und Kommunikationswissenschaft). KABATEK, J.; KORTMANN, B. (Ed.). Berlin: Mouton de Gruyter, 2009. Disponível em: <<http://www.home.hum.uva.nl>>.

_____. Functional approaches to language. In: BUTLER, C. S.; GÓMEZ-GONZÁLEZ, M. L. A.; DOVAL-SUÁREZ, S. M. (Ed.). In.: *The dynamics of language use: functional and contrastive perspectives*. Amsterdam: John Benjamins, 2005. p. 3-17. (Pragmatics and Beyond, 140).

_____. Functionalist Approaches to Language. In: *Structure and Function: A Guide to Three Major Structural-Functional Theories*. Part 1: Approaches to the Simplex Clause. Amsterdã/Filadélfia: John Benjamins Publishing Company, 2003.

BYBEE, J. *Language, Usage and Cognition*. Cambridge University Press, 2010.

CACCIARI, C.; GLUCKSBERG, S. Imagining Idiomatic Expressions: Literal or Figurative Meanings? In: *Idioms: structural and psychological perspectives*. Edited by Martin Everaert et al. New Jersey: Lawrence Erlbaum Associates Publishers, p.43-56, 1995.

_____.; TABOSSI, P. *The comprehension of idioms*. Journal of memory and Language, v.2, p. 668-683, 1988.

CARNEIRO, A. S. *Descrição e classificação das expressões cristalizadas com ser e estar do Português do Brasil*. Dissertação de Mestrado em Linguística. São Carlos: UFSCar, 2016.

CARVALHO, G. L.; TURÉ, A. S. *Expressão idiomática e cultura: um estudo sobre aspectos culturais na língua guineense e no Português Brasileiro*. Njinga; Sepé: Revista Internacional de Culturas, Línguas Africanas e Brasileiras. São Francisco do Conde, v.3, nº Especial II, p.508-526, 2023.

CENTRO DE LINGUÍSTICA DA UNIVERSIDADE DE LISBOA. *Corpus de Referência do Português Contemporâneo (CRPC)*. In: *Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC 2004)*. Paris. França, p. 507-510. <http://alfclul.clul.ul.pt/CQPweb/crpcf16/>, corpus-of contemporary-portuguese-crpc, 2016. Disponível em: < <https://gamma.clul.ul.pt/CQPweb/> >.

CIESLICKA, A. Formulaic language in L2: Storage, retrieval and production of idioms by second language learners. In: *Cognitive processing in second language acquisition: Inside the learner's mind*. M. Pütz; L. Scola (Eds.), p. 149–168. John Benjamins Publishing Company, 2010.

CORPAS PASTOR, G. *Manual de fraseologia española*. Madrid: Gredos, 1996.

CORPUS DO PORTUGUÊS. Disponível em: < <http://www.corpusdoportugues.org> >.

CROFT, W. Construction Grammar. In: *The Oxford Handbook of Cognitive Linguistics*. Dirk

Geeraerts, Hubert Cuyckens (Editors). Oxford University Press, 2007.

_____. *Typology and Universals*. 2nd edition. Cambridge: Cambridge University Press, 2003.

DA SILVA, D. O. *Expressões idiomáticas no Português: uma breve análise funcional*. Trabalho de conclusão de curso da graduação em Licenciatura em Letras Português/Inglês. Araraquara: FAVENI, 2023.

_____. *Fraseologia e Parafraseologia: uma análise semântico-formal*. Trabalho de conclusão de curso da graduação em Licenciatura em Letras português/espanhol. São Carlos: UFSCar, 2018.

DICIONÁRIO ONLINE DE PORTUGÊS. Disponível em: < <http://www.dicio.com.br>>.

DIK, S. *Gramática Funcional*. Tradução de Leocadio Martin Mingorance e Fernando Serrano Valverde. Madrid: Soc. Gen. Española de Librería, 1981.

_____. *The theory of functional grammar*. Ed. by Kees Hengeveld. Berlin: Mouton de Gruyter, 1997.

DREW, P.; HOLT, E. Idiomatic Expressions and Their Role in the Organization of Topic Transition in Conversation. In: *Idioms: structural and psychological perspectives*. Edited by Martin Everaert et al. New Jersey: Lawrence Erlbaum Associates Publishers, p.117-132, 1995.

FILLMORE, C. J.; O'CONNOR, P. K.; O'CONNOR, M. *Regularity and idiomatity in grammatical constructions*. *Language* 64: p. 501–38, 1988.

FONSECA, H. C. *Fraseologismos zoônimos: elaboração de base de dados Português-Francês*. UNESP, São José do Rio Preto: [s.n.], 2013.

GIBBS, R. *Spilling the beans on understanding and memory for idioms in conversation*. *Memory and cognition*, nº 8, p. 449-456, 1980.

GIVÓN, T. Iconicity, isomorphism and non-arbitrary coding in syntax. In: *Iconicity in Syntax*. Jon Haiman (ed.), p. 187-219. Amsterdam: John Benjamins, 1985.

GOLDBERG, A. E. *Constructions: a construction grammar approach to argument structure*. Chicago: University of Chicago Press, 1995.

GOOGLE. Disponível em: <<http://www.google.com.br>>.

HENGEVELD, K.; MACKENZIE, J. L. *Gramática Discursivo-Funcional*. Tradução de Marize Mattos Dall'Aglio Hattner para Língua Portuguesa, Oxford University Press, p. 1-47, 2010.

_____. *Functional Discourse Grammar: A typologically-based theory of language structure*. Oxford: Oxford University Press, 2008.

_____, KEIZER, E.; GIOMI, R. *Layering in Functional Discourse Grammar: The hierarchical structure of the language system*. Oxford: Oxford University Press. No prelo.

HOLM, J. Creole Influence on Popular Brazilian Portuguese. In: *Pidgin and Creole Languages*. GILBERT, G. (ed.). Honolulu: University of Hawaii Press, p. 406-429, 1987.

IRUJO, S. *Don't put your leg in your mouth: Transfer in the acquisition of idioms in a Second Language*. TESOL Quarterly, v. 20, nº 2, 1986.

KEIZER, E. *A Functional Discourse Grammar for English*. Oxford: Oxford University Press, 2015.

_____. *Idiomatic expressions in Functional Discourse Grammar*. De Gruyter Mouton. Linguistics, 54 (5): 981- 1016. DOI: 10.1515/ling-2016-0022, 2016.

_____. *Idiomatic expressions and anaphoric reference in Functional Discourse Grammar*. Quaderns de Filologia: Estudis Lingüístics XXIII: 177- 195. DOI: 10.7203/qf.23.13527, 2018.

KLARE, J. Lexicologia e fraseologia no Português moderno. In: *Revista de Filologia Românica*, IV. Madrid: Editorial de la Universidad Complutense, 1986.

KÖVECSES, Z. *Metaphor, language and culture*. Delta, nº 26: especial, p. 739-757, 2010.

LAKOFF, G. Embodied Minds and Meanings. In: *Speaking Minds — Interviews with Twenty Eminent Cognitive Scientists*. BAUMGARTNER, P. e PAYR, S. (Edit.). Princeton, Princeton University Press, 1995.

_____.; JOHNSON, M. *Metaphors We Live By*. Chicago: The University of Chicago Press, 1980.

_____. *Women, fire, and dangerous things: what categories reveal about the mind*. Chicago: The University of Chicago Press, 1986.

LANGLOTZ, A. *Idiomatic creativity: A cognitive-linguistic model of idiom representation and idiom-variation in English*. Amsterdam and Philadelphia: John Benjamins, 2006.

LAUAND, J. *Vinte e tantos verbetes para um Dicionário de Expressões*. International Studies on Law and Education, nº 41, 2022.

LEUFKENS, S. *The transparency in language: a typological study*. Tese de Doutorado. Universidade de Amsterdã. Utrecht: LOT Publications, 2015.

LEVELT, W. *Speaking*, Cambridge, Cambridge University Press, 1989.

LOPES, C. *Gramaticalização: definição, princípios e análise de casos*. Rio de Janeiro: Faculdade de Letras/UFRJ, 2007.

LYONS, J. *Introdução à linguística teórica*. São Paulo: Nacional/EDUSP, 1979.

MARQUES, I. S., TELETIN, A. *Quando os portugueses se veem gregos ou a questão dos estereótipos culturais em expressões idiomáticas portuguesas e francesas*. Textos seleccionados do XXVI Encontro da Associação Portuguesa de Linguística. Lisboa: APL, p.343-357, 2011.

MONTEIRO-PLANTIN, R. S. *Fraseologia: uma mão na roda na construção de sentido*. Synergies Tunisie, nº 3, p. 161-168. Paris: LDI - Paris 13, 2011.

NUNBERG, G., IVAN, A. S.; THOMAS, W. *Idioms*. Language, nº 70, v.3, p. 491–538, 1994.

OLIVEIRA, W. L.; XATARA, C. *Novo PIP - Dicionário de provérbios, idiomatismos e palavrões em uso Francês-Português /Português-Francês*. São Paulo: Cultura, 2008.

ORTÍZ-ALVAREZ, M. L. O. *Expressões idiomáticas do Português do Brasil e do Espanhol de Cuba: estudo contrastivo e implicações para o ensino do Português como língua estrangeira*. Tese de Doutorado em Linguística Aplicada. Campinas: UNICAMP, 2000.

PASTORE, P. C. F. *A simbologia dos animais em expressões idiomáticas Inglês-Português: uma proposta lexicográfica*. Tese de doutorado em Estudos Linguísticos. São José do Rio Preto: UNESP, 2009.

PEREIRA, A. C. *Competência linguística em Português Europeu Língua Materna e Língua Não Materna: Aquisição, Ensino e Aprendizagem de Expressões Idiomáticas*. Tese de doutorado em Ciências da Linguagem. Moinho: Universidade do Moinho, 2015.

PEZATTI, E. G. *A ordem das palavras no Português*. 1ª ed. São Paulo: Parábola Editorial, 2014.

_____. *Sintaxe descritiva da Língua Portuguesa: Gramática Discursivo-Funcional*. São José do Rio Preto: UNESP, 2010.

_____; et al. *Transparência linguística*. São José do Rio Preto: UNESP, 2017.

RIVA, H. C. *Dicionário onomasiológico de expressões idiomáticas usuais na língua portuguesa do Brasil*. UNESP São José do Rio Preto: [s.n.], 2009.

SABINO, M. A. *O campo árido dos fraseologismos*. Signótica, Goiânia, v. 23, nº 2, p. 385-401, 2011.

SCHAPIRA, C. *Les stéréotypes en français: proverbes et autres formules*. Paris: Ophrys, 1999.

SERPA, O. *Dicionário de expressões idiomáticas Inglês-Português/Português-Inglês*. 4.ed. Rio de Janeiro: FENAME, 1982.

SILVA, D. da. *De onde vem as palavras*. 3ª ed. São Paulo: Mandarim, 1997.

SIMÕES, D. *As expressões idiomáticas no ensino do Português*. Revista Educação e práticas interdisciplinares: Linguagens e diálogos, v. 3, São Carlos: Pedro e João Editores, p. 12-28. 2018.

STASSI SÉ, J. C. *Subordinação discursiva no Português à luz da gramática discursivo-funcional*. Tese de Doutorado em Estudos Linguísticos. São José dos Campos: UNESP, 2012.

TEIXEIRA, J. *Metáforas da vida co(t)vidiana*. Revista Estudos Linguísticos e Literários, nº 69, Salvador, p. 21-51, 2020.

TRISTÁ PÉREZ, M. A. *Fraseología y contexto*. Habana: Editorial de Ciencias Sociales, 1988.

VALE, O. A. *Expressões cristalizadas do Português do Brasil: uma proposta de tipologia*. Tese de Doutorado em Letras: Linguística e Língua Portuguesa. Araraquara: UNESP, 2001.

VEGA MORENO, R. E. *Creativity and Convention: The pragmatics of everyday Figurative Speech*. Pragmatics and Beyond new series. John Benjamins Publishing, 2007.

WEB TECHNOLOGIES SURVEYS. Disponível em < <https://www.w3techs.com/> >.

XATARA, C. *As expressões idiomáticas de matriz comparativa*. Dissertação de Mestrado em Estudos Linguísticos. Araraquara: UNESP, 1994.

_____. *Tipologia das expressões idiomáticas*. São Paulo: Alfa, v. 42, p. 169-176, 1998.

ZULUAGA, A. *Introducción al estudio de las expresiones fijas*. Frankfurt a. M. BernCirencester/U.K., Studia Románica et Lingüística, Verlag Peter D. Lang, 1980.

ANEXO 1

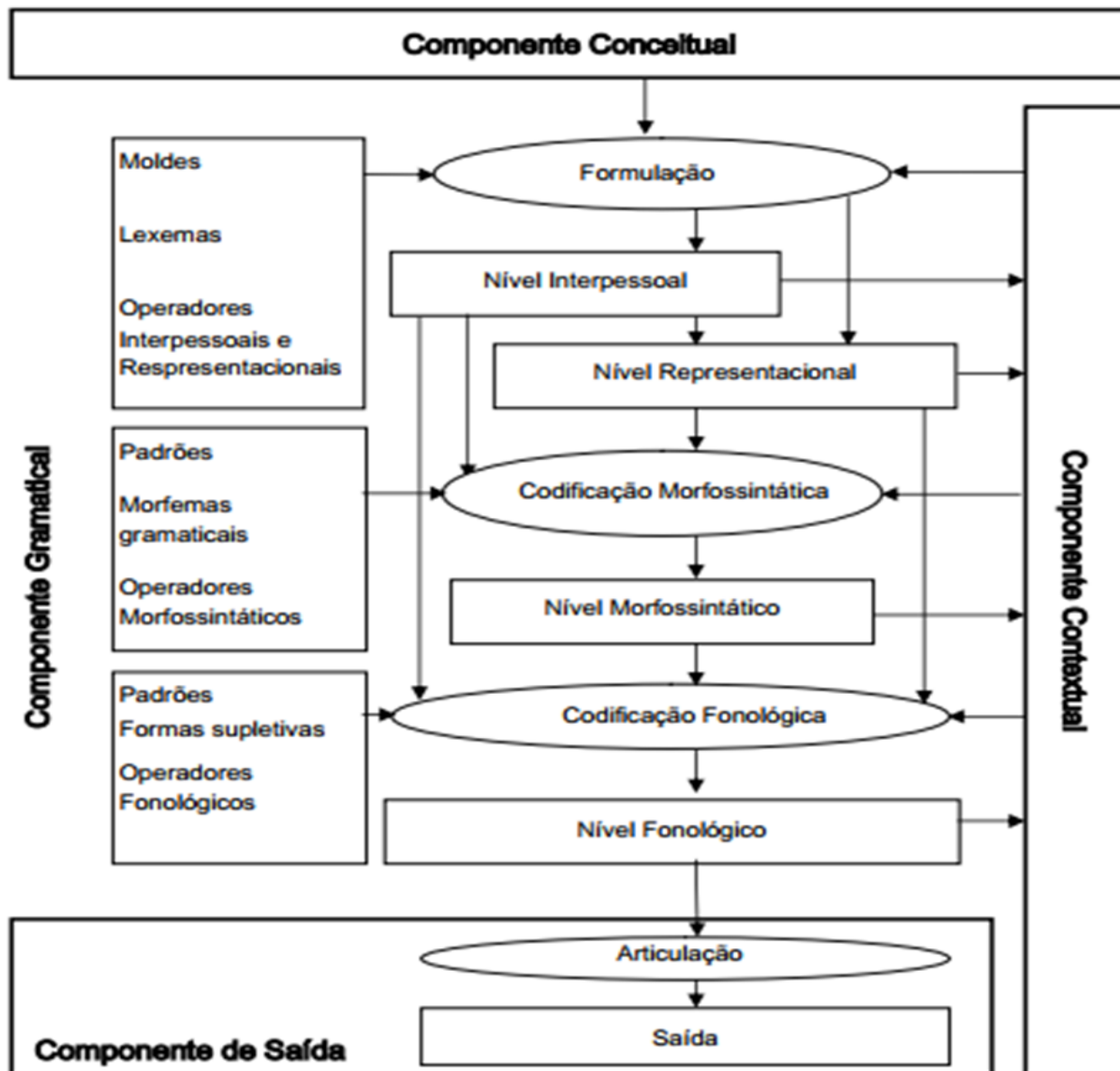


Figura: Arquitetura anterior da GDF (Hengeveld; Mackenzie, 2010, p. 13)