



Bianca Freitas de Jesus

**O *dizer* em português:
diálogos entre tradução, descrição
e linguística computacional**

Dissertação de Mestrado

Dissertação apresentada ao Programa de Pós-Graduação em Estudos da Linguagem da PUC-Rio como requisito parcial para obtenção do título de Mestre em Letras/Estudos da Linguagem.

Orientadora: Profa. Maria Cláudia de Freitas

Rio de Janeiro
Abril de 2016



Bianca Freitas de Jesus

**O *dizer* em português:
diálogos entre tradução, descrição
e linguística computacional**

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre pelo Programa de Pós-Graduação em Estudos da Linguagem da PUC-Rio. Aprovada pela Comissão Examinadora abaixo assinada.

Profa. Maria Cláudia de Freitas

Orientadora
Departamento de Letras – PUC-Rio

Profa. Maria Paula Frota

Departamento de Letras – PUC-Rio

Profa. Heliana Ribeiro de Mello

UFMG

Profa. Denise Berruezo Portinari

Coordenadora Setorial do Centro de Teologia
e Ciências Humanas – PUC-Rio

Rio de Janeiro, 08 de abril de 2016.

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização da autora, da orientadora e da universidade.

Bianca Freitas de Jesus

Graduou-se em Letras – Português/Inglês pela Universidade Federal do Rio de Janeiro (UFRJ) em 2011. Em 2008, começou a trabalhar como tradutora, especializando-se em tradução técnica. Além disso, é professora de idiomas desde 2009 e atua na formação de novos tradutores desde 2012. É também editora-geral da Traduzine, uma revista on-line voltada para tradutores.

Ficha Catalográfica

Jesus, Bianca Freitas de

O dizer em português : diálogos entre tradução, descrição e linguística computacional / Bianca Freitas de Jesus ; orientadora: Maria Cláudia de Freitas. – 2016.

133 f. : il. ; 30 cm

Dissertação (mestrado)—Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Letras, 2016.

Inclui bibliografia

1. Letras – Teses. 2. Verbos de elocução. 3. Descrição do português. 4. Linguística computacional. 5. Corpus. 6. Tradução. I. Freitas, Maria Cláudia de. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Letras. III. Título.

CDD: 400

Agradecimentos

Ao Renato, meu companheiro para toda a vida, que me apoiou incondicionalmente e ficou ao meu lado sempre. Ele fez de tudo para que eu ficasse bem e feliz durante esta jornada. Estou muito bem e muito feliz. Obrigada.

Aos meus pais e à minha irmã, que sempre foram a minha base sólida. Sem eles, o caminho até aqui teria sido muito mais árduo. Sem eles, eu não seria eu.

Aos meus queridos amigos, Luciane, Fábio e Luiza, que me amparam tanto ao longo desses dois anos. Eles fizeram tudo ficar mais leve.

Aos amigos que colaboraram diretamente com esta pesquisa: Gabriel, Laura, Elayne, Marcelo, Heloise, João Artur, Rony, Analice, Debora, Mariana F., Karinna, Mariana M., Aldo e Laís.

Às professoras Maria Paula Frota e Heliana Ribeiro de Mello, pela participação na comissão examinadora e pelos comentários construtivos.

À Diana Santos, cujos comentários e contribuições foram fundamentais para o andamento e o enriquecimento desta pesquisa.

Ao Alberto Simões, pela ajuda imensurável no desenvolvimento do nosso experimento no Rêve.

À Capes e à PUC-Rio, pelo suporte e pelos auxílios que viabilizaram esta pesquisa.

E especialmente à minha orientadora, Cláudia Freitas, que me motivou, me ensinou e me apoiou *muito* nos últimos dois anos. Ela foi a alma desta pesquisa maravilhosa. Obrigada por tudo.

Resumo

Jesus, Bianca Freitas de; Freitas, Maria Cláudia de. **O dizer em português: diálogos entre tradução, descrição e linguística computacional**. Rio de Janeiro, 2016, 133p. Dissertação de Mestrado – Departamento de Letras, Pontifícia Universidade Católica do Rio de Janeiro.

Os profissionais que traduzem textos do inglês para o português deparam-se com uma exigência dos clientes quanto à tradução de diálogos: diversificar os verbos de elocução, com o intuito de evitar a repetição do verbo “dizer”. Em resposta, esta pesquisa visa à elaboração de um glossário dos verbos que introduzem discurso relatado, chamados de verbos de elocução. Para tanto, conduzimos um estudo descritivo com base em corpus, com o intuito de compilar um léxico dos verbos de elocução do português. Assim, esta pesquisa promoveu um amplo levantamento dos verbos de elocução, estabeleceu os padrões de uso nos quais esses verbos costumam ser empregados e, ainda, propôs uma classificação desses verbos em grupos de sentido. Nosso estudo traz, portanto, dois objetivos principais, que se traduzem em contribuições concretas: (i) a elaboração de um glossário dos verbos introdutores de discurso relatado para tradutores, chamado DISSE, e (ii) a descrição dessa classe de verbos no português, tomando um viés semântico e com base em grandes corpora. Como contribuições secundárias, mas não menos importantes e já em fase final de implementação, destacam-se (iii) a criação e a disponibilização de um corpus anotado quanto aos verbos do grande grupo de dizer, também chamados de verbos de comunicação e (iv) a colaboração na elaboração de sistemas capazes de identificar automaticamente citações em textos da língua portuguesa.

Palavras-chave

Verbos de elocução; descrição do português; corpus; tradução; linguística computacional.

Abstract

Jesus, Bianca Freitas de; Freitas, Maria Cláudia de. **Say in Portuguese: a dialogue between translation, description and computational linguistics**. Rio de Janeiro, 2016, 133p. MSc Dissertation – Departamento de Letras, Pontifícia Universidade Católica do Rio de Janeiro.

Professionals that translate from English to Portuguese often face a certain demand from clients when it comes to translating dialogues: diversifying the verbs that introduce reported speech, in an attempt to avoid repeating the verb *dizer* (*say*) in Portuguese. In order to help solve such problem, this study aims at developing a glossary of verbs that introduce reported speech. To reach that aim, we conducted a corpus-based descriptive research, in order to compile a quotation verbs lexicon for the Portuguese language. Thus, this study promoted a wide collection of quotation verbs, as well as established patters of usage in which these verbs are commonly found and put forward a classification of these verbs into groups of meaning. This study proposes two main objectives, which have led to concrete contributions: (i) the elaboration of a reported speech verbs glossary for translators of Portuguese, which is called DISSE (*said*, in Portuguese), and (ii) the description of such verb class in Portuguese, with a semantic approach and based on large corpora. As secondary contributions, but far from unimportant and close to final implementation, it is possible to highlight (iii) the creation and the public availability of annotated corpora, including a semantic annotation for reported speech verbs and (iv) the collaboration with the preparation of systems capable of automatically identifying quotations in Portuguese-written texts.

Keywords

Reported speech verbs; Portuguese language description; corpus; translation; computational linguistics.

Sumário

Introdução	13
1. Perspectiva teórica	17
1.1 Perspectivas linguísticas: uma visão não essencialista para a descrição	17
1.1.1 O fracasso das teorias logocêntricas	17
1.1.2 Os caminhos para uma descrição não essencialista	20
1.2 Pesquisas com base em corpus: da descrição à anotação	22
1.3 Por que e como (continuar a) categorizar	25
2. Dos Estudos da Linguagem à Linguística Computacional: uma resenha de literatura sobre os verbos de elocução	28
2.1 Os verbos de elocução e os Estudos da Linguagem	28
2.1.1 Os verbos de elocução em outras línguas	28
2.1.2 Os verbos de elocução no português	39
2.2 Os verbos de elocução na Linguística Computacional	56
3. Levantamento dos verbos de elocução	60
3.1 Por onde começamos: as traduções de <i>said</i>	60
3.2 Depreensão de padrões de uso	64
3.2.1 Descrição dos padrões	69
3.2.2 Análise da produtividade dos padrões encontrados	82
3.3 Ampliação da lista de verbos de elocução: a construção de um léxico	84
3.4 Anotação como pesquisa: a implementação dos verbos de elocução e o refinamento na noção de discurso relatado	86
4. Criação do DISSE: organização dos verbos em grupos de sentido	90
4.1 A primeira classificação	90
4.2 Nova tentativa de classificação	94
5. Verbos de elocução e discurso relatado: uma descrição com base em corpus	105
5.1 Moura Neves: verbos de elocução e o tipo de discurso	106
5.2 Manual do Estadão: os verbos “mais que errados”	112
Considerações finais	116
Referências bibliográficas	120

Lista de quadros

Quadro 1:	Classes propostas por Levin (1993) para os verbos de comunicação da língua inglesa.	30
Quadro 2:	Esquema das classes criadas por Moura Neves (2000), com os respectivos exemplos.	47
Quadro 3:	Lista de traduções para a forma <i>said</i> em 958 ocorrências, compilada no COMPARA.	62
Quadro 4:	Distribuição dos seis verbos analisados, organizados por ordem de frequência no corpus CHAVE e distribuídos por variante.	66
Quadro 5:	Distribuição dos padrões por tipo de citação.	67
Quadro 6:	Complementos mais frequentes dos verbos de elocução no Padrão (5).	75
Quadro 7:	Verbos encontrados dentro do Padrão (6).	77
Quadro 8:	Número de verbos encontrados com as expressões de busca em cada padrão.	85
Quadro 9:	Classificação inicial dos 58 verbos de dizer obtidos no COMPARA, distribuídos em 21 grupos de sentido.	90
Quadro 10:	Lista de verbos e os grupos de sentido validados pelo primeiro teste.	91
Quadro 11:	Lista de classes atribuídas aos verbos que menos e mais receberam classificações diferentes.	94
Quadro 12:	Classes propostas por J. L. Austin (1962) para os verbos que tornam explícita a força ilocucionária.	95
Quadro 13:	Lista dos verbos da segunda versão do teste, com as respectivas classes atribuídas.	100
Quadro 14:	Linhas de concordância avaliadas na segunda versão do teste e que receberam o menor ou o maior número de classes.	101
Quadro 15:	Quadro comparativo dos verbos de elocução levantados em diversos estudos.	105
Quadro 16:	Verbos apresentados por Moura Neves (2000) como verbos de elocução e que não foram encontrados por esta pesquisa em um primeiro momento, mas foram confirmados pelos corpora do projeto AC/DC.	107

Quadro 17:	Organização dos verbos encontrados nesta pesquisa, de acordo com a classificação feita por Moura Neves (2000).	108
Quadro 18:	Análise do uso dos “verbos mais que errados”, listados pelo Manual de Redação e Estilo do Estado, nos corpora do AC/DC.	112

Lista de figuras

Figura 1:	Verbos mais comuns no domínio semântico da comunicação, organizados por ordem de frequência e distribuídos por registro.	36
Figura 2:	Gráficos com os 12 verbos mais frequentes em cada registro.	37
Figura 3:	Gráfico com os 12 verbos mais comuns no LSWE Corpus.	38
Figura 4:	Esquema da classificação dos verbos proposta por Moura Neves (2000), reduzida às classificações que interessam a este estudo.	45
Figura 5:	Tipos de complemento aceitos pelos verbos de simples dizer e verbos que qualificam o que é dito – parte 1.	46
Figura 6:	Tipos de complemento aceitos pelos verbos de simples dizer e verbos que qualificam o que é dito – parte 2.	49
Figura 7:	Tipos de complemento aceitos pelos verbos de simples dizer e verbos que qualificam o que é dito – parte 3.	49
Figura 8:	Tipos de complemento aceitos pelos verbos que instrumentalizam ou circunstanciam o que é dito – parte 1.	50
Figura 9:	Tipos de complemento aceitos pelos verbos que instrumentalizam ou circunstanciam o que é dito – parte 2.	50
Figura 10:	Número de ocorrências obtidas com uma busca por <i>said</i> em diálogos.	55
Figura 11:	Fluxograma com as etapas metodológicas desta pesquisa.	60
Figura 12:	Interface da busca no COMPARA.	62
Figura 13:	Interface do corpus Floresta em uma busca por distribuição de lema.	85
Figura 14:	Estrutura da primeira versão do teste.	91
Figura 15:	Gráfico da distribuição de verbos por classe.	97
Figura 16:	Interface da segunda versão do teste da classificação em grupos de sentido.	100

Lista de tabelas

Tabela 1:	Comparação dos tipos de complementos do verbo <i>continuar</i> no Padrão (5) entre as variantes do português.	76
Tabela 2:	Ocorrências dos seis verbos em cada padrão analisado.	83

Lista de anexos

Anexo 1:	Expressões de busca específicas para o projeto AC/DC, desenvolvidas para cinco dos padrões de uso.	123
Anexo 2:	Lista dos verbos encontrados, distribuídos por padrão.	125
Anexo 3:	Respostas dadas pelos dez participantes da primeira versão do experimento, que era composto por 58 verbos.	128
Anexo 4:	Glossário DISSE.	131

Introdução

Quando comecei a me dedicar profissionalmente à tradução, tive que me aventurar por diversas áreas para descobrir minha “identidade tradutória”. A tradução literária foi uma dessas áreas. No primeiro trabalho que consegui, fazendo tradução e copidesque de livros de ficção, a editora em questão me passou o manual de estilo que eu deveria seguir, desenvolvido pela própria editora. Ao ler as recomendações dadas, uma delas chamou a minha atenção: “Variar o uso dos verbos de elocução durante o texto. Não permitir que muitos ‘disse’ ou ‘falou’ fiquem acumulados num trecho só.”

A vivência profissional, aliada à minha formação continuada como tradutora, apenas confirmou que a recomendação dada pela editora é uma prática comum do mercado da tradução. A justificativa dada pelos profissionais que consultei (professores de cursos de formação e editores) é que, no inglês, é extremamente comum que a introdução de discurso seja feita pelo verbo *say* em sua forma pretérita, *said*, enquanto no português costumamos diversificar mais os verbos que introduzem fala.

Afinal, no português, temos tantos verbos que introduzem relato assim? Quantos verbos podem desempenhar essa função? Foram perguntas como essas que nortearam esta pesquisa, cuja principal motivação é a elaboração de um glossário de tais verbos, chamado DISSE, com o intuito de facilitar essa tarefa tradutória. Classificado quanto a classes de sentido, espero que o glossário sirva como auxílio ao tradutor, que, por conta da orientação de editoras e agências, deve lidar com uma variação lexical que não existe no texto fonte. No entanto, essa classe de verbos interessa também a outros campos, o que amplia também o escopo desta dissertação.

Os verbos que introduzem a fala são fundamentais para uma tarefa de Processamento de Linguagem Natural (PLN) chamada Identificação de Citação (*Quotation Extraction*), e atender a esse interesse constitui uma motivação secundária desta pesquisa. Na tarefa de *quotation*, o foco não está exatamente no verbo, mas em quem fala e no conteúdo dessa fala. Contudo, sabemos que a forma de introduzir o discurso relatado – isto é, o verbo escolhido – também é indicativa

de posicionamentos. Desse modo, a existência de uma lista de tais verbos, associada aos padrões léxico-sintáticos em que são usados, é um recurso de grande utilidade para o PLN em português.

A ampla frequência, na língua, dos verbos introdutórios de discurso relatado também é um grande motivador para estudos descritivos, sobretudo com base em corpora, além de indicar a sua relevância na língua. Em um dos corpora usados nesta pesquisa, chamado CHAVE¹, o verbo *dizer* é o sexto verbo mais empregado, atrás apenas dos verbos, *ser*, *ter*, *estar*, *fazer* e *poder* (verbos esses que são muito usados como auxiliares).

Esta dissertação também se enquadra no âmbito do projeto Gramateca² (Santos, 2014), que tem como um de seus objetivos a elaboração de uma gramática – em sentido amplo – da língua portuguesa, com base nos corpora públicos disponíveis por meio da interface AC/DC (Santos & Sarmento, 2003). Ao participar da Gramateca, a exploração e a descrição se beneficiam também dos serviços a ela associados, com o objetivo de auxiliar o pesquisador no estudo da língua (Simões e Santos, 2014). Especificamente, os serviços Distribuidor³ e Rêve⁴ contribuíram em diferentes momentos desta dissertação.

Um estudo descritivo amplo dos verbos de elocução também pode interessar a outras áreas, como o ensino de português – quando se tematiza discurso direto e indireto – e a pragmática, devido aos atos de fala (Austin, 1962). Apesar de existirem trabalhos sobre o tema, com uma perspectiva mais estilística ou pragmática, há uma carência de estudos descritivos com base em corpus. A exceção é a *Gramática de usos do português*, de Maria Helena de Moura Neves (2000), que, com base em corpus, aborda os verbos de elocução, listando-os conforme sua presença em discursos direto e indireto. Nossos objetivos são diferentes – e os resultados, complementares –, e o material (os corpora) que usamos está público, o que permite a outros confrontar ou continuar o trabalho apresentado aqui.

Este estudo tem, portanto, dois objetivos principais: (i) a elaboração de um glossário dos verbos de dizer para tradutores e (ii) a descrição dessa classe de verbos

¹

² No corpus CHAVE, encontrado no site <http://www.linguateca.pt/Gramateca>, usamos a expressão de busca [pos="V.*"] e pedimos a distribuição de lema. Esses conceitos serão explicados mais adiante.

³ <http://www.linguateca.pt/Distribuidor>

⁴ <http://www.linguateca.pt/Reve>

no português, tomando um viés semântico e com base em grandes corpora. É a descrição detalhada, realizada em (ii), que garantirá a qualidade da lista de verbos presente em (i). Como contribuições secundárias, mas não menos importantes, destacam-se (iii) a criação e a disponibilização de um corpus anotado quanto aos verbos do discurso relatado e (iv) a colaboração na elaboração de sistemas capazes de identificar automaticamente citações em textos da língua portuguesa. Para tanto, conduzimos um amplo levantamento dos verbos de elocução, propomos uma classificação desses verbos em grupos de sentido e, além disso, fazemos um levantamento dos padrões nos quais esses verbos costumam ser empregados.

Os verbos-alvo desta dissertação são comumente conhecidos como “verbos *dicendi*”, além de receber outros nomes, como “verbos de dizer” e “verbos de elocução”. O termo *dicendi* costuma ser usado em manuais de estilo, e essa classe de verbos seria mais restrita do que o escopo deste estudo, pois, para a maioria dos manuais e gramáticas, os verbos *dicendi* são os verbos de dizer propriamente ditos, isto é, verbos que indicam a elocução e cujo complemento direto é o conteúdo do que se diz. Fora desse grupo, encontram-se “verbos que introduzem discurso, mas não necessariamente indicam atos de fala”, como por exemplo, *ameaçar* e *espantar-se* (Moura Neves, 2000, p. 48-49), verbos esses que são contemplados neste estudo.

Em português, a maioria das pesquisas acadêmicas trata dos “verbos de elocução”, cuja definição, para Moura Neves, é “verbos introdutórios de discurso (discurso direto ou discurso indireto)” (Ibidem, p. 47). Como esse é o recorte que nos interessa, adotamos “verbos de elocução” como nossa nomenclatura e, como alternativa, empregamos também “verbos introdutórios de discurso relatado”.

Muitos dos estudos que descrevem verbos, principalmente os voltados para a língua inglesa, tratam dos chamados “verbos de comunicação”, que têm abrangência muito maior do que os verbos introdutórios de discurso relatado. Os verbos de comunicação correspondem a qualquer verbo relacionado à comunicação ou à fala, independentemente de haver algo sendo relatado. Como, mais uma vez, tínhamos grande interesse no relato, tampouco essa nomenclatura e essa abrangência nos pareceram proveitosas.

Como metodologia, utilizamos a exploração de grandes corpora e a ideia de anotação como forma de pesquisa (Sampson, 2001; Freitas, 2015; Santos et al., 2015). Como perspectiva teórica, adotamos uma visão não representacionista,

crítica do logocentrismo, inspirada sobretudo na leitura wittgensteiniana de Martins (1999).

A dissertação está estruturada da seguinte maneira: no capítulo 1, discorro sobre a perspectiva teórica que norteou esta pesquisa; no capítulo 2, desenvolvo a nossa visão sobre verbos de elocução e, ainda, apresento alguns estudos que já trataram dessa classe de verbos, tanto no português quanto em outras línguas; no capítulo 3, descrevo o processo de levantamento de verbos de elocução e, no capítulo 4, trato da elaboração do glossário DISSE, comparando-os, sempre que possível, com os demais trabalhos na área. Por fim, teço algumas considerações finais e apresento algumas possibilidades de trabalhos futuros.

1

Perspectiva teórica

Neste capítulo, apresento a perspectiva teórica que direciona esta pesquisa: (i) uma visão não logocêntrica da linguagem, tomando por base as visões de Rosemary Arrojo, que se fundamenta em Derrida, e de Helena Martins, que, por sua vez, se apoia no pensamento wittgensteiniano; (ii) a articulação entre essa perspectiva e trabalhos com base em corpora; e (iii) a proposta de categorização na língua como apresentada por John M. Ellis (1993).

1.1

Perspectivas linguísticas: uma visão não essencialista para a descrição

Para expor a perspectiva adotada para abordar a descrição de uma língua, escolhemos duas obras como norteadoras: *O signo desconstruído*, de Rosemary Arrojo, e a tese de doutorado de Helena Martins, chamada *Metáfora e polissemia no estudo das línguas do mundo: uma aproximação não representacionista*. Com a primeira, justificamos nossa posição de que teorias linguísticas com viés logocêntrico estão fadadas ao fracasso. Martins, por sua vez, nos aponta caminhos alternativos para a descrição de uma língua, uma vez que não nos embasamos em um modelo essencialista.

1.1.1

O fracasso das teorias logocêntricas

Em um dos artigos que compõem *O signo desconstruído*, Rosemary Arrojo explica que “toda teoria linguística parte necessariamente de um determinado pressuposto sobre a origem e a localização do significado”. E esse pressuposto, chamado “logocêntrico”, corresponde à visão de que a origem dos significados se encontra “*fora* do sujeito/leitor” (Arrojo, 2003a, p. 35, grifo da autora).

De acordo com Arrojo (e trazendo, aqui, a perspectiva da desconstrução), nós estaríamos inseridos em uma tradição cultural que estabelece dicotomias como sujeito e objeto, dado e interpretação, e que, ainda, atribui a origem do significado

ao “significante (no texto, na ‘mensagem’, na palavra), nas intenções (conscientes) do emissor/autor, ou numa combinação ou alternância dessas duas possibilidades” (Ibidem, p. 35). Assim, na análise de uma classe de verbos – nosso problema aqui –, perspectivas logocêntricas tentariam classificá-los de acordo com o seu significado “original” ou, ainda, com o que “o autor quis dizer” no momento da fala/escrita.

Quando se considera o significante como a fonte de significado, temos uma noção de literalidade que coloca o significado como “anterior a qualquer interpretação e independente de qualquer contexto”. Por outro lado, quando as intenções são consideradas como tal, o desejo do autor/emissor é que determina o significado, e, dessa forma, o leitor possui um papel passivo e se rende ao que “o autor quis dizer” (Ibidem, p. 35-36). O reconhecimento dessa dimensão do significado será fundamental, como afirmamos, para lidar com nossa tentativa de classificação (voltaremos a esse ponto no capítulo 4).

A tradição logocêntrica, segundo Arrojo, concebe um sujeito capaz de ser objetivo para com a realidade. Para essa tradição, que se insere em uma lógica dicotômica, o ato de interpretar se oporia ao de compreender, e essa oposição seria, de acordo com Arrojo, uma consequência teórica da demarcação entre sujeito e objeto. À oposição entre compreensão e interpretação se equiparariam outras, tais como literal e figurado, denotativo e conotativo, verdade e retórica.

Para a autora, a maior parte do conhecimento ocidental pressupõe a possibilidade de se resgatar significados estáveis, e nesse conhecimento se incluem as nossas concepções clássicas de linguagem. Na teoria da tradução, no entanto, são especialmente delicadas as questões envolvidas na aceitação dessa possibilidade de resgate. Eis o motivo:

Toda tradução, por mais simples e breve que seja, trai sua procedência, revela as opções, as circunstâncias, o tempo e a história de seu realizador. Toda tradução, por mais simples e breve que seja, revela ser produto de uma perspectiva, de um sujeito interpretante e, não, meramente, uma compreensão “neutra” e desinteressada ou um resgate comprovadamente “correto” ou “incorreto” dos significados supostamente estáveis do texto de partida. (Arrojo, 2003b, p. 68)

Trazemos aqui o campo dos Estudos da Tradução porque é a partir da tradução que demos os passos iniciais na nossa pesquisa, desde a motivação até o primeiro levantamento dos verbos de elocução. Arrojo explica que os modelos teóricos da linguística, que estão em mutação, são limitados demais para lidar com

as múltiplas faces da tradução. Com base nisso, é possível delinear um paralelo entre a limitação das teorias linguísticas para lidar com a tradução e para lidar com a linguagem de um modo geral.

Arrojo defende que o projeto logocêntrico estaria fadado ao fracasso, uma vez que ignora a temporalidade de tudo o que é produzido pelo homem e considera “categorias inevitavelmente marcadas pelo tempo e produzidas por sujeitos sempre situados em algum contexto sociocultural como instâncias ‘divinas’, acima de qualquer perspectiva ou interesse subjetivo”. De modo complementar, os modos de ver as coisas são ilimitados, o que coloca em xeque a possibilidade de haver um ponto de vista a partir do qual é possível visualizar tudo. Essa multiplicidade de modos de ver as coisas acaba por solapar as estruturas das pretensões de teorias universais.

Isso não quer dizer que qualquer interpretação seja possível, ou que qualquer teorização seja impossível. Sempre há perspectivas mais produtivas, que permitem ver mais coisas do que as demais. Neste trabalho, o nosso ponto de vista não parte de uma dada teoria, mas sim de uma prática, a Tradução, que tem como elemento central o sentido. Partimos do princípio, também, de que tradutores são bons jogadores no jogo da reconstrução dos sentidos de um texto em outra língua.

Esta dissertação trata, também, de categorização: categorização de certos verbos como verbos de elocução e, ainda, de uma categorização interna a essa. De uma perspectiva logocêntrica, essa categorização é uma tarefa independente de interpretação: para atribuir um elemento a uma classe (ou categoria), basta o *reconhecimento* das propriedades definidoras da classe – isto é, as propriedades e características estão lá, prontas para serem identificadas. E a classe dos verbos de elocução, ou de comunicação, é sempre dada como certa. De uma perspectiva não logocêntrica, a categorização é sempre um ato interpretativo, e aprofundaremos essa visão da categorização como ato decorrente de escolhas e intenções, sendo as categorias fruto dos interesses dos falantes, e não reflexos do mundo real, na seção 1.3. Em vez de criar uma teoria que ambicione dar conta de todos os verbos introdutórios de discurso relatado, o que podemos fazer é conduzir um estudo descritivo da língua em uso no momento em que nos encontramos, procurando chegar a um consenso entre os membros da nossa comunidade interpretativa.

Apesar das reflexões muito pertinentes para esta pesquisa, a saber, o fracasso do projeto logocêntrico e a impossibilidade de uma teoria, Arrojo apenas

volta seu olhar para a tradução, o ensino e a leitura. Considerando o tipo de reflexão que nos interessa, recorreremos à tese de Helena Martins (1999), que, inspirada pelo ângulo wittgensteiniano, nos aponta caminhos para compatibilizar as ideias não essencialistas, com a prática e a relevância da descrição linguística.

1.1.2

Os caminhos para uma descrição não essencialista

Logo no primeiro capítulo de sua tese, Helena Martins deixa claro que considera a visão wittgensteiniana da linguagem como um bom ponto de partida para descrever os significados regulares das palavras no estudo das línguas. A opção da autora deve-se ao fato de que Ludwig Wittgenstein, filósofo austríaco, permite a compreensão de metáfora e polissemia sem a exclusão da “noção de que a metáfora é um fato de resistência à estabilização conceitual” (Martins, 1999, p. 137). No entanto, nos interessa aqui a leitura positiva que Martins faz de Wittgenstein com relação aos Estudos da Linguagem, sobretudo com relação às possibilidades da descrição.

Segundo Martins, um dos intentos de Wittgenstein foi desvincular a reflexão filosófica sobre a linguagem de seu compromisso com o cientificismo. Por insistir com tanta veemência nessa desvinculação, há, de modo geral, menos controvérsia acerca da ideia de que, “quando se trata de pensar a linguagem e o significado, é útil abandonar a ‘obsessão pela generalidade’ característica da teorização científica” (Martins, 1999, p. 142).

A crítica wittgensteiniana, que corresponde em grande parte à crítica ao modelo logocêntrico, pode ser entendida como uma proposta, por parte de Wittgenstein, de abandonar toda a concepção da Linguística como ciência da linguagem. Martins ressalta, no entanto, que não é razoável acreditar que Wittgenstein indique a impossibilidade de aproximação científica de qualquer dimensão da linguagem. Em verdade, ao sugerir que a linguagem não está totalmente aberta à investigação científica, o filósofo parece tematizar, mais especificamente, o problema do sentido na linguagem. O que seria preciso abandonar, portanto, não é a ideia de uma ciência da linguagem, mas sim de uma ciência do significado.

Com isso, Wittgenstein enfatiza que, para pensar sobre a gramática, isto é, sobre as regras de uso de uma língua, é de maior valia *favorecer uma perspectiva*

descritiva em detrimento da atitude explicativa que caracteriza as ciências: “[e] não devemos construir nenhuma espécie de teoria. Não deve haver nada de hipotético em nossas considerações. Toda *explicação* deve desaparecer e ser substituída apenas por descrição” (Ibidem, p. 143). E é essa visão que justifica a nossa opção por um estudo descritivo, com base em corpus, dos verbos de elocução.

Martins aponta que, para muitos linguistas da contemporaneidade, abordar o uso dos signos a partir de uma perspectiva descritiva seria como abandonar o projeto da Linguística como um todo. De fato, a visão wittgensteiniana da linguagem é inconciliável com um projeto universalista e essencialista, além de não ser compatível com qualquer linguística representacionista. Assim, adotar a perspectiva de Wittgenstein corresponde a “abandonar a ambição explicativa no estudo da linguagem” (Ibidem, p. 144). No entanto, vale lembrar que, apesar de não encorajar a busca por uma explicação geral para o funcionamento das línguas, Wittgenstein não inviabiliza estudos sistemáticos e empíricos das línguas – e é nesse espaço que nos colocamos.

Uma vez reconhecido que não é possível conjugar a visão wittgensteiniana e “o representacionismo mentalista que domina o cenário teórico da Linguística pelo menos desde Chomsky”, Martins conclui que a Linguística

[n]ão será, pois, uma *linguística do sujeito*, voltada para a caracterização de um conhecimento mental tácito que ele supostamente deve possuir para produzir e interpretar enunciados linguísticos, mas antes um estudo que terá como propósito descrever regularidades nos jogos de linguagem que perfazem as línguas do mundo. (Ibidem, p. 146)

Trazendo tais reflexões para o âmbito desta dissertação, podemos dizer que o nosso ângulo de observação é o da tradução. Foi a partir da tradução que demos os passos iniciais na nossa pesquisa, desde a nossa motivação até o primeiro levantamento dos verbos introdutores de discurso relatado.

Também podemos afirmar que buscamos descrever regularidades observáveis no português, no que se refere a esse grupo de verbos, sem qualquer pretensão de dar conta da língua como um todo. Os padrões de uso dos verbos são o resultado da nossa busca por regularidades. Por fim, ilustramos, com anotação de corpus, a prática de descrição desvinculada da necessidade de explicação. Na próxima seção, abordaremos, com mais profundidade, a relação entre pesquisas com corpus e descrição.

1.2

Pesquisas com base em corpus: da descrição à anotação

A análise de grandes corpora é fundamental neste trabalho. Nesta seção, apresentamos um histórico dos trabalhos com corpus nos Estudos da Linguagem e introduzimos a ideia de anotação como forma de estudo de uma língua.

Com a evolução galopante das tecnologias, torna-se cada vez mais fácil ter acesso a textos de origens diversas. Essa facilidade, no entanto, entra em conflito com as nossas limitações enquanto seres humanos, uma vez que a quantidade de dados disponíveis é muito superior à nossa capacidade de analisá-los “de maneira consistente e confiável”. Por conta desse conflito, podemos encarar os computadores como “aliados valiosos, propiciando novas maneiras de leitura e auxiliando o pesquisador na percepção e na análise de padrões cuja identificação seria difícil (se é que possível) por meio de uma leitura convencional” (Freitas, 2015, p. 23).

Segundo Freitas (2015, p. 23), corpus é como chamamos as “grandes coleções de documentos de textos, compiladas segundo algum objetivo e, atualmente, processáveis por máquinas”. A autora diz *atualmente* porque os primeiros relatos de exploração de amostras de textos remetem ao século XIII, e, até o século XX, compilações de concordâncias eram feitas manualmente.

Hoje associados aos estudos linguísticos, os corpora teriam sido utilizados, originalmente, para conduzir estudos bíblicos e, mais tarde, literários. No que se refere ao estudo sistemático de línguas, as primeiras pesquisas com corpora extensos são comumente associadas ao projeto SEU (Survey of English Use), que teve início em 1959. Já o primeiro corpus eletrônico foi o Brown Corpus, desenvolvido nos Estados Unidos, em 1961, e composto por um milhão de palavras (Ibidem, p. 25). Na mesma época, também foram produzidos projetos de corpora para a língua portuguesa: em Portugal, um corpus vinculado ao projeto Português Fundamental (Bacelar do Nascimento et al., 1984, 1987 apud Santos et al., 2011); e, no Brasil, o projeto Nurc (Norma Urbana Culta), cujo objetivo era “documentar e estudar a norma culta falada em cinco grandes capitais[...], produzindo diferentes corpora orais” (Callou, 1999 apud Freitas, 2015).

A pouca influência das pesquisas com corpus no âmbito da Linguística se deve a duas razões: à crítica tecida pela escola gerativa chomskyana, que se coloca contra a língua em uso como fonte de evidências linguísticas; e à desconfiança da

análise manual de corpora de grande porte. No entanto, os computadores, desenvolvidos e popularizados a partir dos anos 1980, possibilitaram o armazenamento de grandes volumes de dados e, em seguida, “a análise de padrões da língua”, o que incitou a retomada da “utilização de corpus para estudos linguísticos” (Freitas, 2015, p. 26).

Sobre o começo do uso sistemático de grandes corpora eletrônicos nas pesquisas linguísticas, Freitas afirma:

[...] a partir dos anos 1990, [...] tem-se um contexto propício para pôr em xeque os modelos teóricos vigentes – uma vez que estes teriam sido construídos a partir de dados pouco confiáveis – e, considerando a linguagem como um sistema probabilístico, para estudar de maneira sistemática fenômenos que têm relação direta com a frequência, como a identificação de padrões lexicais e a preferência por certas combinações de palavras. (Ibidem, p. 28)

A partir dessa visão de que os modelos teóricos vigentes são questionáveis, surge o que se costuma chamar de abordagem *corpus-driven*. De acordo com essa perspectiva, “a exploração sistemática de dados reais da língua em uso, por si só, seria capaz de construir novos modelos teóricos, com novas categorias de análise” (Ibidem, p. 28). O pesquisador seria, então, capaz de se manter imparcial e apenas relataria dados e elaboraria hipóteses com base nos dados que aparecessem no corpus. A questão é que, ao indicar que a fonte dos *dados* e das *análises* feitas se encontra no corpus, em vez de ser um produto da *interpretação* do pesquisador, a abordagem *corpus-driven* presume que exista “uma entidade ‘significado’ autônoma, estável, independente de um sujeito, da história ou de circunstâncias de leitura, localizada no texto e pronta para ser revelada por meio das ferramentas adequadas” (Ibidem, p. 29).

A perspectiva *corpus-based*, por sua vez, entende o corpus como uma ferramenta que auxilia o pesquisador na hora de verificar hipóteses previamente elaboradas com base em modelos teóricos que já existiam. Assim, com base nos dados e/ou nos exemplos encontrados no corpus, uma teoria poderia ser validada, refutada ou especificada. De modo geral, as hipóteses direcionariam a observação, que, em seguida, poderia remodelar a hipótese, em um ciclo que se retroalimenta. Mesmo na abordagem *corpus-based*, o emprego da tecnologia é encarado como uma “garantia de neutralidade e objetividade” (Ibidem, p. 29-30). Com isso em mente, temos que as perspectivas *corpus-driven* e *corpus-based* não

necessariamente se opõem; as fronteiras entre as duas abordagens não são tão bem delimitadas.

Outra observação importante a respeito do uso de corpus para a investigação de fenômenos linguísticos diz respeito à representatividade. No âmbito dos estudos linguísticos, torna-se difícil determinar a representatividade de um dado corpus, uma vez que não temos como mensurar qual é o todo da língua – só podemos ter noção da representatividade de uma amostra se conhecemos a proporção dessa amostra em relação ao total. Todavia, Freitas (2015, p. 33) defende que não se deve abrir mão de buscar uma amostra representativa, ainda que esta pareça inalcançável, ou, no lugar da representatividade, deve-se buscar o contraste, a alteridade.

Os corpora usados nesta pesquisa, vinculados ao projeto AC/DC (Costa et al., 2009), reúnem mais de um bilhão de palavras. Ainda que esse número seja impressionante, os corpora correspondem a apenas uma amostra da língua portuguesa. Trazendo a reflexão sobre representatividade para esta dissertação, podemos pensar que o fato de nos termos voltado para corpora que contêm gêneros textuais distintos, principalmente textos literários e jornalísticos, provavelmente nos dá uma dimensão melhor de como o relato costuma ser introduzido no português do que se tivéssemos nos limitado a um tipo de texto apenas.

Freitas resume, então, o que devemos ter em mente ao pesquisar com base em corpus:

- (i) os dados obtidos referem-se sempre ao conteúdo do corpus, e o corpus é uma parcela de algo; (ii) para todo dado, há sempre uma interpretação humana; (iii) essa interpretação deve ser capaz de confrontar mesmo os dados do próprio corpus – o questionamento do corpus também deve fazer parte da pesquisa. (Ibidem, p. 33)

Neste trabalho, lidamos com corpora anotados. Chamamos de corpus cru o corpus que corresponde a apenas uma compilação de textos, em formato eletrônico, reunidos com uma dada finalidade. Nos corpora anotados, por sua vez, os textos coletados têm algum tipo de informação linguística a eles associada, informação essa que é chamada de anotação.

Como explica Freitas, “anotar é delimitar um segmento de texto e atribuir-lhe uma etiqueta (ou conjunto de etiquetas)” (Ibidem, p. 36). A anotação, assim, é uma atividade de *classificação*, feita conforme etiquetas (classes) pré-determinadas. Essas etiquetas podem conter informações de natureza semântica, morfossintática, entre outras. Os corpora do projeto AC/DC já estão anotados

morfossintaticamente, e alguns dos resultados desta pesquisa também foram incorporados a esse material, como uma anotação semântica do campo do dizer.

Além de ser o produto de uma pesquisa, quando a anotação é feita de forma manual ou semiautomática, ela se torna “um *processo* de investigação” (Ibidem, p. 48, grifo da autora), e esse é um dos pontos deste trabalho. Freitas explica que,

[a]o participar do desenvolvimento de um projeto de anotação, o pesquisador é obrigado a se apropriar do texto que será anotado e das classes que compõem o esquema de anotação. Assim, a anotação pode se configurar também como processo de pesquisa, como um teste de hipóteses relativo às categorias propostas para abordar uma questão. (Freitas, 2015)

Usarei as etapas de anotação pelas quais passamos como exemplo do que afirmou Freitas. Antes de dar início à anotação, estabelecemos os nossos critérios quanto ao que seriam os verbos de dizer, os verbos introdutores de discurso relatado, a citação direta, a citação indireta, a citação mista e, por fim, quanto ao que não seria *dizer*. No entanto, à medida que nos deparávamos com casos mais complexos, que não se enquadravam nos nossos critérios da maneira esperada, éramos forçadas a reformular os critérios, para que, enfim, pudéssemos dar conta desses casos. Assim, temos que a verificação no corpus pode levar à reformulação da hipótese (o que aconteceu durante o processo de anotação) ou, ainda, à confirmação do que foi estabelecido (o que buscamos atingir quando concluirmos a anotação).

A menção aos casos mais complexos nos leva à discussão de outro aspecto relevante da anotação. Inevitavelmente, o processo de anotação está longe de ser um trabalho mecânico; em vez disso, a anotação constitui uma interpretação. “Categorizar/classificar/anotar é estabilizar um certo fenômeno, e o processo de anotação nos confronta frequentemente com os limites de estabilizações, por mais precisos que estes nos pareçam” (Freitas, 2015, p. 48). Retomamos assim a ideia de classificação e categorização, fundamentais a essa pesquisa, apresentando as ideias de John Ellis (1993) na próxima seção.

1.3

Por que e como (continuar a) categorizar

Ao pensarmos sobre a organização do DISSE, decidimos que a melhor forma de o apresentar seria distribuindo os verbos de elocução em grupos de sentido, isto é, categorizando-os segundo classes previamente determinadas. A

classificação passou por dois testes, durante os quais encontramos diversos obstáculos. A seguir, traremos algumas das reflexões de Ellis (1993) sobre a atividade de classificação para o contexto desta dissertação e as discutiremos à luz do que foi feito nesta pesquisa.

Para Ellis (1993), categorias (e aqui se incluem categorias linguísticas) podem agrupar aquilo que *não é semelhante*. Segundo o autor, só é possível absorver a ideia de categorização se a entendemos como o agrupamento de coisas que não são semelhantes de modo que as possamos considerá-las como semelhantes.

Ellis (1993) aponta que a experiência de cada indivíduo é infinitamente variável, o que torna impossível rotular cada situação existente. Desse modo, nenhuma comunicação poderia ocorrer a não ser que cada falante de uma dada língua presenciasse cada situação, além de as ter memorizado e rotulado de forma semelhante. Ainda, o número de rótulos viria a ser tão extenso que nenhum dos indivíduos envolvidos em uma comunicação se lembraria de seus referentes. O mais importante, contudo, é que a língua não teria nenhuma aplicação real, pois se limitaria apenas à rememoração de situações passadas.

Nessa visão, a língua somente funciona como instrumento de conhecimento e comunicação porque *simplifica* a complexidade da experiência, ao reduzir uma variedade infinita a um conjunto finito de categorias. A consciência de que a categorização acaba por igualar o não igual é fundamental para lidar com casos difíceis de classificação.

Como lembra o autor, ficamos tentados a pensar que um número maior de categorias tornaria o sistema mais sensível às diferentes situações e mais útil. No entanto, à medida que o número de categorias se aproxima do número de situações categorizadas, toda a ideia de categorização se esvai, afinal, ter uma categoria para cada situação não se mostra nada útil.

Em suma, o argumento de Ellis consiste no seguinte: “as categorias linguísticas são principalmente um reflexo dos propósitos coletivos dos falantes de uma língua, em vez de serem reflexos diretos da estrutura do mundo” (1993, p. 34, tradução minha). A equivalência criada pelas categorias de uma língua é, portanto, funcional, uma vez que, mesmo não sendo idênticas, as coisas incluídas em uma categoria podem e serão tratadas como equivalentes para o propósito dessa categoria.

Na seção seguinte, esmiuçamos alguns dos estudos sobre os verbos introdutores de discurso relatado, conduzidos tanto aqui no Brasil quanto no exterior. Guiadas pela perspectiva teórica apresentada nesta seção, fazemos uma análise crítica de cada uma dessas obras, para, no capítulo 3, propor uma análise aprofundada dos verbos em questão.

2

Dos Estudos da Linguagem à Linguística Computacional: uma resenha de literatura sobre os verbos de elocução

No Brasil, há um número considerável de estudos que tratam dos verbos de elocução. No entanto, a maioria desses trabalhos adota um viés mais estilístico, pragmático ou normativo, o que não significa que não tenham contribuído para o estudo dessa classe de verbos da língua portuguesa. Alguns estudos voltados para outras línguas também foram consultados, e uma análise de cada um deles foi feita nas subseções a seguir, correspondentes a duas áreas: Estudos da Linguagem e Processamento de Linguagem Natural.

2.1

Os verbos de elocução e os Estudos da Linguagem

Nas próximas subseções, são analisados os estudos que se voltam para os verbos de elocução, direta ou indiretamente, de maneira breve ou aprofundada. Todos os trabalhos aqui analisados encontram-se no grande campo dos Estudos da Linguagem.

2.1.1

Os verbos de elocução em outras línguas

Nesta subseção, discuto dois trabalhos que constituem grandes referências dos estudos descritivos com base em corpus: Beth Levin, com *English verb classes and alternations*, e a *Longman grammar of spoken and written English*. Os principais aspectos de cada obra são analisados e, ainda, traço um paralelo entre tais trabalhos e a pesquisa desenvolvida por nós.

Levin (1993)

Em 1993, foi publicado o livro *English verb classes and alternations: a preliminary investigation*, de Beth Levin. Esse livro teve origem no estudo de organização lexical e representação lexical, conduzido no âmbito do Lexicon Project do MIT Center for Cognitive Science. *English verb classes and alternations*

corresponde, portanto, a uma lista ampliada e revisada das classes verbais compiladas. Os dicionários também serviram como embasamento para as listas apresentadas.

Levin explica que o que é apresentado no livro constitui os resultados iniciais de uma pesquisa sobre as propriedades sintáticas e semânticas dos verbos. Segundo a autora, o livro se baseia no princípio de que o comportamento sintático de um verbo é, em grande parte, determinado pelo seu significado, principalmente no que se refere à expressão e à interpretação dos argumentos dos verbos. Dessa forma, o comportamento sintático dos verbos poderia ser usado para investigar aspectos dos significados dessa classe de palavras (Levin, 1993, p. 1).

A autora também ressalta a dificuldade de se identificar os significados possíveis dos verbos com base apenas na intuição. Dicionários tampouco facilitam essa identificação. Para Levin, estudos do comportamento sintático de verbos ajudam a melhor indicar seus significados e a melhor organizar o léxico dos verbos da língua inglesa. As propriedades sintáticas, assim, facilitariam o enquadramento dos verbos em grupos semânticos e, do mesmo modo, seria possível prever o comportamento sintático de um verbo com base na sua classe semântica (Ibidem, p. 15).

Na Parte II do livro, que trata das classes de verbos, Levin apresenta classes de verbos que são “sintaticamente relevantes” e “semanticamente coerentes” (Ibidem, p. 22). Para cada classe, a autora fornece uma lista de verbos como exemplo, mas essas listas não têm a pretensão de ser exaustivas. Além disso, Levin seleciona um verbo representativo para exemplificar as propriedades características da classe analisada.

Entre as 49 classes de verbos propostas por Levin, a classe que nos diz respeito se chama “verbs of communication”. Segundo a explicação da autora, estão reunidos nessa classe os verbos que estão relacionados com a comunicação e a transferência de ideias, e esses verbos, por sua vez, são classificados em nove categorias semânticas. Vale ressaltar aqui que a menção à transferência de ideias remete diretamente à metáfora do conduto⁵, isto é, a uma visão da linguagem como

⁵ Arrojo cita Michael Reddy e a metáfora do conduto para aprofundar-se sobre a questão do transporte de significados. Reddy explica que, se considerarmos as palavras como recipientes de pensamento, sentimento ou emoção, é preciso levar em conta estas consequências:

“1) a linguagem funciona como um conduto, transferindo pensamentos de pessoa para pessoa;
2) ao escrever e falar, as pessoas inserem seus pensamentos ou sentimentos nas palavras;

algo que pode ser transportado de maneira objetiva, sem qualquer interferência externa à palavra – uma visão de língua de que não compartilhamos, como já explicado no capítulo 1.

No Quadro 1, encontram-se as nove classes propostas por Levin, bem como a lista de verbos para cada classe:

CLASSE	VERBOS
Verbos de transferência de mensagens (<i>verbs of transfer of a message</i>)	ask, cite, demonstrate, dictate, explain, explicate, narrate, pose, preach, quote, read, recite, relay, show, teach, tell, write
Contar (tell)	tell (apenas)
Verbos de modo de falar (<i>verbs of manner of speaking</i>)	babble, bark, bawl, bellow, bleat, boom, bray, burble, cackle, call, carol, chant, chatter, chirp, cluck, coo, croak, croon, crow, cry, drawl, drone, gabble, gibber, groan, growl, grumble, grunt, hiss, holler, hoot, howl, jabber, lilt, lisp, moan, mumble, murmur, mutter, purr, rage, rasp, roar, rumble, scream, screech, shout, shriek, sing, snap, snarl, snuffle, splutter, squall, squawk, squeak, squeal, stammer, stutter, thunder, tisk, trill, trumpet, twitter, wail, warble, wheeze, whimper, whine, whisper, whistle, whoop, yammer, yap, yell, yelp, yodel
Verbos de instrumento de comunicação (<i>verbs of instrument of communication</i>)	cable, e-mail, fax, modem, netmail, phone, radio, relay, satellite, semaphore, sign, signal, telephone, telecast, telegraph, telex, wire, wireless
Verbos de fala (<i>talk verbs</i>)	speak, talk
Verbos de bate-papo (<i>chitchat verbs</i>)	argue, chat, chatter, chitchat, confer, converse, gab, gossip, rap, schmooze, yak
Verbos de dizer (<i>say verbs</i>)	announce, articulate, blab, blurt, claim, confess, confide, convey, declare, mention, note, observe, proclaim, propose, recount, reiterate, relate,

3) as palavras contêm os pensamentos ou sentimentos até que sejam transferidos para outras pessoas;
4) ao ouvirem ou lerem, as pessoas extraem os pensamentos e sentimentos das palavras” (Orotny apud Arrojo, 2003c, 101).

Dessa forma, quando uma leitura é considerada inadequada ou inaceitável, atribui-se a responsabilidade ao leitor, que inseriu seus pensamentos nas palavras do texto, em vez de apenas extrair significados. Podemos concluir, então, que a tradição logocêntrica proíbe ao tradutor, ao leitor e ao ouvinte a interpretação, isto é, “a interferência de seu contexto histórico-social e de sua psicologia na construção de uma leitura ou tradução” (Arrojo, 2003c, p. 102).

	remark, repeat, report, reveal, say, state, suggest
Verbos de reclamar (<i>complain verbs</i>)	boast, brag, complain, crab, gripe, grouch, grouse, grumble, kvetch, object
Verbos de aconselhar (<i>advise verbs</i>)	admonish, advise, alert, caution, counsel, instruct, warn

Quadro 1: Classes propostas por Levin (1993) para os verbos de comunicação da língua inglesa.

A classe de verbos analisada por Levin corresponde aos verbos de comunicação que mencionei na Introdução. Trata-se de um grupo muito mais amplo de verbos, que têm relação com a fala, mas não necessariamente introduzem discurso. A princípio, a subcategoria dos “verbos de dizer” (*say verbs*) corresponderia melhor ao escopo desta dissertação, mas, ainda assim, os critérios que demarcam cada uma das subcategorias não são tão claros, e outras subcategorias parecem abrigar alguns dos verbos que são considerados, por nós, como verbos de elocução (como os “verbos de reclamar” e os “verbos de aconselhar”).

Levin deixa claro que seu tratamento dos verbos de comunicação é bastante breve, correspondendo basicamente aos verbos cujas propriedades não envolvem complementos sentenciais. Essa brevidade se deve ao fato de que o estudo sistemático dos complementos sentenciais de verbos não se encontra dentro do escopo do livro em questão (Ibidem, p. 202). Apesar do escopo limitado em termos sintáticos, Levin consegue reunir 163 verbos de comunicação, e o trabalho por ela desenvolvido é uma importante referência não apenas na descrição, mas também no PLN, uma vez que a sua classificação embasa um dos grandes recursos léxicos usados no processamento automático das línguas, a VerbNet (Kipper et al., 2006) e, por extensão, a sua versão brasileira, VerbNet-BR (Scarton et al., 2012). A seguir, apresento e discuto as nove categorias propostas por Levin.

(1) Verbos de transferência de mensagens (*verbs of transfer of a message*)

Esses verbos diferem dos demais no que se refere à natureza da mensagem e à maneira pela qual esta é transmitida. Aceitam sintagmas com *to*, que indicam o interlocutor e permitem a alternância dativa. Também podem aceitar complementos sentenciais, porém, como já foi dito, esse tipo de complemento não é abrangido pelo livro.

(2) Contar (*tell*)

O verbo *tell*, embora presente na categoria anterior, é melhor explorado nesta categoria. O significado desse verbo não envolve qualquer especificação. A autora provê, então, alguns exemplos que mostram que o verbo *tell* aceita outras estruturas sintáticas além das que foram explicitadas na categoria (1).

(3) Verbos de modo de falar (*verbs of manner of speaking*)

Esses verbos distinguem-se pela maneira com que o som é expresso. A maioria dos verbos aqui listados corresponde a verbos onomatopéicos ou de expressão não verbal. Verbos diferentes, ainda que na mesma classe, podem apresentar comportamento sintático diferente. Levin aponta que uma consideração acerca do uso comunicativo e não comunicativo dessa classe de verbos pode levar a uma melhor análise.

(4) Verbos de instrumento de comunicação (*verbs of instrument of communication*)

Todos os verbos dessa classe advêm de um substantivo que denomina um instrumento de comunicação e correspondem à comunicação via esses instrumentos. Apresentam alternância dativa, que, para a autora, sugere a interpretação da comunicação como uma troca de posse, pois a informação é passada de um indivíduo para outro.

(5) Verbos de fala (*talk verbs*)

Apesar de apresentarem propriedades diferentes, os verbos *speak* e *talk* foram agrupados porque o significado de ambos se relaciona à fala, porém sem qualquer especificação de meio ou modo. Não aceitam complementos sentenciais. Aceitam sintagmas com *to* para explicitar quem é o interlocutor, além dos sintagmas com *with*, que indicam outro participante na conversa.

(6) Verbos de bate-papo (*chitchat verbs*)

Verbos usados para indicar interações de fala entre dois ou mais participantes. Comumente aceitam sintagmas com *with*, mas nem sempre com *to*. Não admitem complementos sentenciais.

(7) Verbos de dizer (*say verbs*)

Levin indica que os verbos dessa classe são descritos como verbos de comunicação de proposições e de atitudes proposicionais (Gropen et al. apud Levin, 1993, p. 210). Esse grupo de verbos pode apresentar um sintagma com *to* que indica o interlocutor, mas não apresenta a alternância dativa. É limitada a variação de sintagmas nominais admitidos por essa classe de verbos. Complementos sentenciais finitos também são aceitos. Ainda na mesma subseção, Levin ressalta a necessidade de se estudar essa classe com mais cuidado e profundidade.

(8) Verbos de reclamar (*complain verbs*)

Esses verbos indicam a atitude ou os sentimentos do locutor em relação ao que é dito. Somente aceitam complementos sentenciais finitos. Dentro da classe, há algumas variações quanto à aceitação de sintagmas com *to* e *at*.

(9) Verbos de aconselhar (*advise verbs*)

São verbos relacionados a dar conselhos ou avisos. O verbo *alert* difere dos demais no que se refere à estrutura exigida, pois obriga o emprego de um objeto direto, enquanto para os demais verbos a explicitação do objeto direto é opcional.

Apesar da grande relevância e do impacto do trabalho de Levin, temos algumas críticas a tecer quanto à classificação proposta. Em primeiro lugar, os critérios semânticos não estão tão bem justificados. Não está claro, por exemplo, por que o verbo *tell* tem uma categoria só para ele. Tampouco é explicado por que há dois grupos de sentido bem específicos (*aconselhar* e *reclamar*), sendo que vários outros ficaram de fora. Por que essas duas classes foram separadas do *dizer*?

Ao contrário de *English verb classes and alternations*, o presente estudo não exclui os complementos oracionais, que são extremamente produtivos na língua, principalmente no que se refere ao relato de discurso. Por outro lado, tratamos de um grupo de verbos bem mais restrito dos ditos “verbos de comunicação” de Levin. Por conta das diferenças de escopo e, também, por conta das críticas que tecemos, decidimos que a nossa classificação partiria de pontos diferentes e tomaria rumos distintos dos que foram escolhidos pela autora.

Biber et al. (1999)

Longman grammar of spoken and written English é uma gramática do inglês baseada em corpus e, como tal, propõe-se a descrever o uso da língua inglesa em diferentes discursos, registros, etc. As descrições gramaticais dessa obra baseiam-se nos padrões de estrutura e uso encontrados em uma ampla coleção de textos falados e escritos, compilados no Longman Spoken and Written English Corpus.

Na introdução da gramática, os autores ressaltam que, para entender como a gramática do inglês realmente funciona, é necessário conjugar suas estruturas e seus usos. Assim, essa gramática apresenta, por um lado, descrições das estruturas gramaticais e, por outro lado, descrições dos padrões de uso de cada uma dessas estruturas.

Nessa gramática, é proposta uma divisão dos verbos em domínios semânticos, com base no sentido principal, que seria o sentido no qual os falantes costumam pensar primeiro. Os autores ressaltam, contudo, que muitos dos verbos apresentam sentidos variados, pertencentes a diversos domínios semânticos, e, em muitos casos, o verbo é mais comumente usado com seu sentido não principal (Biber et al., 1999, p. 361). Os exemplos (i) e (ii) ilustram o que os autores pretendem com sentido principal e não principal. O exemplo (i) faz referência à atividade física, enquanto o exemplo (ii) tem um sentido mais aspectual, voltado para o progresso de outra ação:

- (i) We **stopped** at the Market on the way back.
- (ii) And it was two o'clock when they **stopped** talking.

Os autores citam a existência de dois tipos de casos problemáticos. Para alguns verbos, parece não haver uma classificação correta e única, uma vez que os sentidos principais desses verbos podem enquadrar-se em mais de uma categoria. Esse primeiro caso corresponde, principalmente, aos verbos que denotam atividade, que possuem sentidos secundários em outros domínios. O verbo *confirm*, por exemplo, pode indicar tanto atos comunicativos quanto mentais.

No segundo tipo de casos problemáticos, alguns verbos seriam empregados com sentidos diferentes que pertencem a mais de um domínio semântico. Esses casos corresponderiam, principalmente, a verbos que denotam atividade, como o verbo *admit*, que pode indicar uma atividade física, comunicativa ou mental. A maioria dos verbos, no entanto, tem um sentido principal que pertence a apenas um domínio semântico.

Um dos domínios semânticos apresentados por Biber et al. corresponde aos verbos de comunicação, que interessam a esta pesquisa. Os autores definem os verbos de comunicação como aqueles que envolvem atividades comunicativas (fala e escrita). São citados os verbos a seguir como exemplos dessa categoria: *ask, announce, call, discuss, explain, say, shout, speak, state, suggest, talk, tell, write* (Biber et al., 1999, p. 362). Tal como explicamos na subseção anterior, esse grupo de verbos é muito mais amplo do que o objeto de estudo desta dissertação.

Para os autores, a comunicação, apesar de ser um tipo de atividade, é merecedora de um domínio semântico próprio. Os falantes de inglês costumam distinguir diferentes tipos de atividades comunicativas e relatar o que alguém disse ou escreveu, por meio de verbos como *ask, call, say* e *speak*. Além disso, os falantes com frequência relatam as próprias opiniões e os próprios sentimentos, bem como as opiniões e os sentimentos de outros. Para isso, os falantes usam verbos chamados de mentais, como *believe, consider, expect, need, want*, etc. (Biber et al., 1999, p. 365).

A Figura 1 foi retirada da gramática de Biber et al. e mostra os 12 verbos mais frequentes, dentro do domínio semântico da comunicação. Os dados mostram verbos que ocorrem mais de 300 vezes a cada um milhão de palavras, em pelo menos um dos quatro registros, a saber, conversa (CONV), ficção (FICT), notícias (NEWS) e redação acadêmica (ACAD). Já a Figura 2 corresponde a quatro gráficos que nos apresentam a frequência dos doze verbos mais usados em cada um dos registros analisados.

	CONV	FICT	NEWS	ACAD
communication				
<i>say</i>	██████████	██████████	██████████	████
<i>tell</i>	██████	██████	████	
<i>call</i>	██	██	██	██
<i>ask</i>	██	██████	██	
<i>write</i>	██	██	██	██
<i>talk</i>	████	████	██	
<i>speak</i>	██	████		
<i>thank</i>	██	██		
<i>describe</i>				██████
<i>claim</i>			██	
<i>offer</i>			██	
<i>suggest</i>				██

Figura 1: Verbos mais comuns no domínio semântico da comunicação, organizados por ordem de frequência e distribuídos por registro.
 Fonte: Biber et al. (1999).

Figure 5.9
Frequencies of the most common lexical verbs—conversation

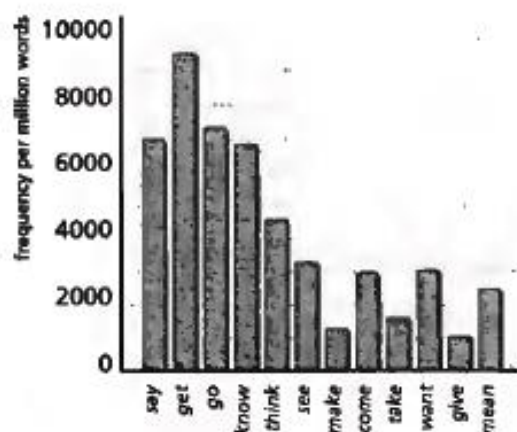


Figure 5.10
Frequencies of the most common lexical verbs—fiction

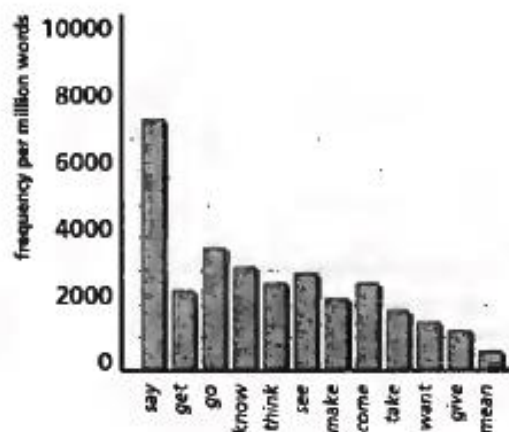


Figure 5.11
Frequencies of the most common lexical verbs—news

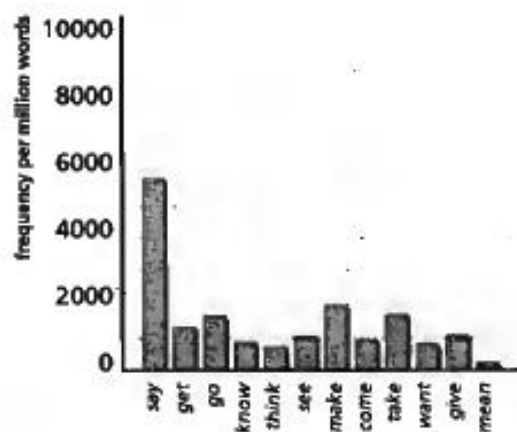


Figure 5.12
Frequencies of the most common lexical verbs—academic prose

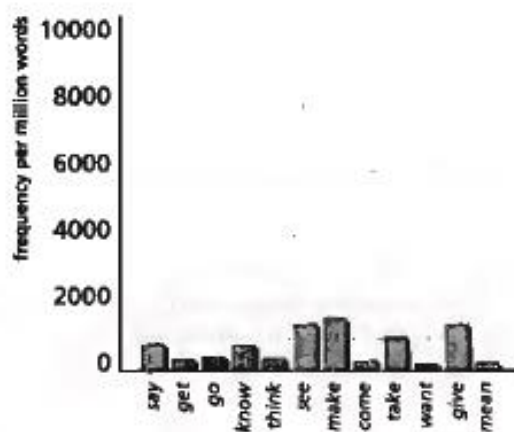


Figura 2: Gráficos com os 12 verbos mais frequentes em cada registro.
Fonte: Biber et al. (1999).

Biber et al. destacam o verbo *say* como o verbo mais comum no LSWE Corpus como um todo e, além disso, como o único verbo extremamente comum em mais de um registro. Nenhum dos demais verbos de comunicação se encontra entre os 12 verbos mais comuns. Conclui-se, assim, que o falante emprega muito esse verbo para relatar discurso. A Figura 3 mostra um gráfico com os verbos mais comuns dentro do LSWE Corpus:

Most common lexical verbs

Overall use of the most common lexical verbs

7 Frequency of the most common lexical verbs in the LSWE Corpus (over 1,000 per million words)

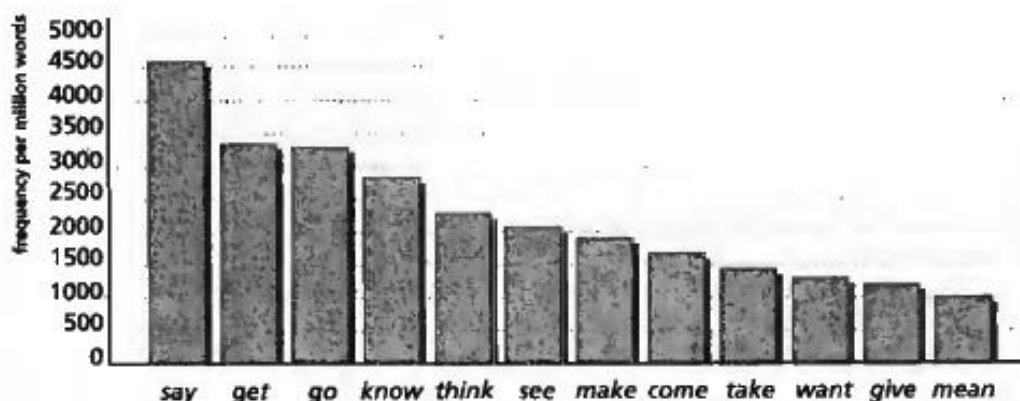


Figura 3: Gráfico com os 12 verbos mais comuns no LSWE Corpus.

Fonte: Biber et al. (1999).

Em todos os registros, o verbo *say* é mais comum no tempo pretérito, em referência a um proferimento passado, como nos exemplos abaixo, tirados de Biber et al. (1999, p. 374):

- (iii) You **said** you didn't have it. (CONV)
- (iv) No use sitting about, he **said**. (FICT)
- (v) He **said** this campaign raised "doubts about the authenticity of the eventual allegedly free choice". (NEWS)

No entanto, no registro de conversa, também é muito comum o uso do tempo presente (35% das vezes, segundo os autores). Biber et al. (1999, p. 375) descrevem brevemente em que situações se costuma usar o presente: (1) para marcar o discurso relatado passado como se fosse no presente, com um efeito de imediatismo e envolvimento pessoal; (2) para marcar um comportamento habitual; (3) para contar piadas em uma conversa. O uso mais frequente da forma *said* em relação a *say* justifica, mais adiante, o nosso ponto de partida para o levantamento dos verbos de elocução, relatado no capítulo 3.

Austin (1962) também se debruça sobre os verbos de elocução, no contexto dos "atos de fala". Uma vez que a tentativa de classificação dos verbos proposta

por ele – e o seu “fracasso” – são especialmente relevantes para o desenvolvimento do glossário, deixamos uma análise mais detalhada desse trabalho no capítulo 4.

2.1.2

Os verbos de elocução no português

Para apresentar os textos de língua portuguesa que consultamos durante a pesquisa, decidimos discutir, primeiro, dois manuais de redação que são importantes referências para tradutores e demais profissionais das palavras: *Comunicação em prosa moderna*, de Othon M. Garcia, e o *Manual de Redação e Estilo do Estado*, o manual que norteia a redação do jornal *O Estado de S. Paulo*. Em seguida, faremos uma resenha dos estudos acadêmicos que consultamos sobre os verbos de elocução.

Othon M. Garcia (2010)

Comunicação em prosa moderna, cujo subtítulo é *Aprenda a escrever, aprendendo a pensar*, é um manual de redação que se propõe a estudar e recomendar padrões “da língua dos nossos dias” (Garcia, 2010, p. 6), com o intuito de ensinar seus leitores a escrever. O manual de Othon M. Garcia é bastante relevante para este estudo, uma vez que muitos tradutores, no curso de suas formações, são orientados a estudá-lo e a consultá-lo. Apesar de Garcia garantir que não preza pela correção gramatical, *Comunicação em prosa moderna* é de fato um manual e, por isso, adota um tom normativo em todo o seu conteúdo.

Othon M. Garcia define os verbos *dicendi* como aqueles “cuja principal função é indicar o interlocutor que está com a palavra” (GARCIA, 2010, p. 149). Propõe, então, uma classificação semântica tentativa para esses verbos, que inclui nove categorias:

- 1) de *dizer*
- 2) de *perguntar*
- 3) de *responder*
- 4) de *contestar*
- 5) de *concordar*
- 6) de *exclamar*
- 7) de *pedir*
- 8) de *exortar*

9) de *ordenar*

O próprio autor, no entanto, reconhece que essas poucas categorias não dão conta dos muitos verbos que podem ser empregados como *dicendi*. Garcia explica que as categorias compreendem os verbos *dicendi* mais comuns, porém ressalta que há muitos outros, que são “mais específicos, mais caracterizadores da fala” (Garcia, 2010, p. 149). Em uma nota de rodapé, Garcia cita alguns exemplos de verbos que, mesmo não sendo *dicendi* por definição, são, para ele, comumente usados como tal.

No que chama de uma “lista caótica”, Garcia (2010) relaciona 50 verbos, e esses verbos corroboram o argumento de que a classificação por ele proposta não é capaz de dar conta da quantidade de verbos empregados como *dicendi*. Pode-se exemplificar essa questão com verbos como *sussurrar* e *mentir*, ambos presentes na lista de Garcia. É possível perceber que nenhum dos dois verbos se encaixa perfeitamente nas categorias sugeridas pelo autor, o que incita a necessidade de se criar novas categorias, de modo a abranger não só os verbos tradicionalmente usados como introdutores de discurso direto ou indireto, mas também aqueles que são ocasionalmente empregados dessa maneira e que descrevem a fala de forma mais específica.

Othon M. Garcia vai além das duas possibilidades mencionadas acima e comenta o uso de verbos que, a princípio, não estão relacionados com a elocução:

Chegam mesmo, os mais imaginativos, a empregar verbos que nenhuma relação têm com a ideia de elocução, o que, do ponto de vista da sintaxe, poderia ser considerado como inadmissível, pois os *dicendi* deveriam ser, teoricamente pelo menos, transitivos ou admitir transitividade. (GARCIA, 2010, p. 149)

Como exemplo, o autor cita apenas o verbo “fazer”, que normalmente não aparece associado à elocução.

Em seguida, Garcia introduz os verbos chamados *sentiendi*, “que não são propriamente ‘de dizer’ mas ‘de sentir’, e que, por analogia, podem ser chamados *sentiendi* [...], [verbos] que expressam estado de espírito, reação psicológica de personagem, emoções” (2010, p. 150), entre outras coisas. O autor explica que os verbos *sentiendi* são “vicários” dos *dicendi* e caracterizam as manifestações psíquicas dos personagens.

Garcia afirma, entretanto, que alguns verbos que não admitem a transitividade devem, no discurso direto, vir antepostos à fala:

- (a) Mas João de Deus, vendo que Vasco não lhe dá atenção, explode:
 - Você pensa, seu Vasco, que estou disposto a aturar suas malcriações? (Veríssimo apud Garcia, 2010, p. 155)
- (b) – Você pensa, seu Vasco, que estou disposto a aturar suas malcriações?
 - explodiu João de Deus.

O exemplo (a), fornecido por Garcia, seria a única possibilidade de uso do verbo “explodir”, a princípio intransitivo, no discurso direto. Para ele, o exemplo (b) seria inadmissível. Para que um verbo como “explodir” seja aceitável na posição posposta, seria necessário que um verbo “legitimamente” *dicendi* o acompanhasse, e, ainda assim, a ordem dos verbos teria de ser alterada, tal como no exemplo a seguir:

- (c) – O coitadinho tem andado aborrecido! – disse ela *lamentando-se*.

Apesar dessa restrição imposta aos verbos *sentiendi* por Othon M. Garcia, uma busca no corpus CHAVE (Santos & Rocha, 2005), corpus monolíngue que integra o projeto AC/DC (Santos & Sarmento, 2003), revela o emprego frequente do próprio verbo “lamentar” em posição posposta à fala descrita, como é possível observar nos exemplos a seguir:

- (d) «Pensei que conseguiria vencer e fiquei animado», *lamentou* Wilander.
- (e) «Qualquer dia eles me tiram», *lamenta*.
- (f) «Até hoje não obtivemos resposta», *lamenta* Pádua

A pesquisa em corpus permite a observação de que nem sempre normas prescritas por gramáticas e manuais de redação se aplicam a dados reais, compilados em corpora. Em apenas um corpus, procurando-se somente pela expressão de busca [word="»"] [word=","] [lema="lamentar"]⁶, foram encontradas 392 ocorrências do verbo posposto a uma fala. Fica evidente, assim, que os dados reais relativizam, quiçá invalidam a limitação do uso dos verbos *sentiendi* estipulada por Garcia, e com os dados é acentuado, também, o potencial do uso de corpus para este tipo de estudo.

⁶ Significado dos símbolos:

[] isolamos elementos pesquisados

word: pesquisa por palavra ou elemento exato

" ": a pesquisa refere-se a tudo o que está entre aspas

lema: pesquisa por todas as formas de uma mesma palavra, com base na forma encontrada nos dicionários.

Garcia separa, ainda, uma seção do capítulo dedicado aos discursos direto e indireto para tecer recomendações acerca da posição dos verbos *dicendi* nos períodos em que se inserem. Adotando novamente uma postura normativa, o autor elabora uma lista de casos em que o verbo pode ou deve anteceder, intercalar ou seguir a fala. Segundo Garcia, no discurso direto dito tradicional, o verbo *dicendi* costuma vir no meio ou no fim da fala. A posição anterior à fala é adotada, para o autor, apenas em casos excepcionais.

O verbo de dizer seria empregado no fim “quando a fala é muito breve e/ou constitui uma unidade com entoação íntegra que lhe torne desaconselhável a ruptura em dois fragmentos com intercalação do *dicendi*” (Ibidem, p. 158). A posição intercalada pode ocorrer nas seguintes situações: (i) “entre dois termos mutuamente dependentes (como sujeito e predicado, verbo e seu complemento, nome e seu adjunto), com propósito enfático”; (ii) após as duas ou três primeiras palavras às quais na fala “se segue uma pausa natural”, como um vocativo, por exemplo; e (iii) entre duas unidades independentes ou dois períodos (Ibidem, p. 159).

Entretanto, Garcia observa que o verbo *dicendi* raramente é colocado após a terceira linha da fala, mesmo estando intercalado ou posposto. O autor afirma que “o normal é vir na primeira [linha], como pudemos verificar em alguns milhares de amostras em algumas dezenas de autores, desde o romantismo até os nossos dias” (Ibidem, p. 159). Ocorreu, então, na literatura mais recente, uma mudança na preferência pela posição do verbo de elocução, que, segundo Garcia, passou a ser empregado anteposto à fala. Além disso, Garcia aponta um aumento no uso dos verbos que ele chama de “vicários do dizer”, isto é, os verbos que denotam reação psicológica. De acordo com Garcia, essa mudança teria ocorrido a partir do movimento realista.

Para corroborar seus argumentos, o autor apresenta alguns dados. Nas obras de José de Alencar, a percentagem de verbos de dizer antepostos é de 5%, enquanto em *Memórias póstumas de Brás Cubas*, por exemplo, o número chega a 25%. Foi a partir de 1930, todavia, que essa mudança se acentuou: em obras como *João Miguel* (1935), de Rachel de Queiroz, e *Música ao longe* (1935), de Érico Veríssimo, a percentagem de anteposição dos verbos de elocução é de aproximadamente 65%.

Por se tratar de um manual de consulta rápida – ao contrário do manual de Othon M. Garcia, que tem uma proposta mais didática –, o *Manual de Redação e Estilo do Estado* não se estende muito a respeito dos verbos que eles chamam de *dicendi*. Reproduzimos, na íntegra, tudo o que o manual diz a respeito dessa classe de verbos:

Verbos mais que errados. 1 - Há verbos que, por questão de significado, não podem ser acompanhados de *que*. Em geral, eles equivalem aos diversos sentidos de dizer (verbos *dicendi*). Veja um exemplo: alguém defende uma ideia, uma posição, mas nunca defende que alguma coisa se realize ou concretize. Assim, são erradas ou no mínimo impróprias as formas: acusar que, alertar que, antecipar que, apontar que, aprovar que, assumir que, citar que, comentar que, continuar que, defender que, definir que, denunciar que, desmentir que, difundir que, divulgar que, enfatizar que, indicar que, justificar que, mencionar que, narrar que, proferir que, prosseguir que, referir que, registrar que e relatar que. 2 - Podem, porém, ser normalmente usadas as formas: acrescentar que, adiantar que, admitir que, advertir que, afiançar que, afirmar que, aguardar que, assegurar que, asseverar que, atestar que, certificar que, comprovar que, concordar que, confirmar que, constatar que, declarar que, determinar que, dizer que, esperar que, garantir que, jurar que, negar que, ordenar que, prever que, prometer que, reiterar que, repetir que, ressaltar que, ressaltar que, revelar que e verificar que. (Estado de S. Paulo, 2016, grifo meu)

O caráter normativo já fica claro na entrada do verbete: “verbos mais que errados”. Segundo o manual, nenhum dos verbos listados acima, no item 1, poderia ser empregado na forma mais conhecida de discurso indireto, isto é, seguido pela conjunção integrante *que*. Contudo, no manual, não há nenhuma informação sobre a fonte dessa lista de verbos ou, ainda, dos critérios que os selecionaram.

A própria noção de *dicendi* está em conflito com o que afirmam os demais trabalhos analisados nesta dissertação. Pela explicação dada pelo manual, o verbo *defender*, em “alguém defende uma ideia”, seria também um verbo *dicendi*, quando, para muitos estudos (e para nós também), o complemento em forma de sintagma nominal não constitui um relato de discurso.

Após a consulta a dois manuais de redação, voltamos o nosso olhar para o que foi feito na academia no que se refere aos verbos de elocução. Mais adiante, no capítulo 5, problematizamos um pouco mais o uso do *Manual de Redação e Estilo do Estado* como referência na área de tradução, verificando suas afirmações com o uso de grandes corpora.

Maria Helena de Moura Neves (2000)

A *Gramática de usos do português*, um estudo descritivo com base em corpus, apresenta-se como “uma obra de referência que mostra como está sendo usada a língua portuguesa atualmente no Brasil” (Moura Neves, 2000, p. 13). A gramática foi desenvolvida a partir de uma base de dados de 70 milhões de ocorrências, armazenada no Centro de Estudos Lexicográficos da Unesp (Campus de Araraquara) e composta por textos de literaturas “romanesca, técnica, oratória, jornalística e dramática”, o que, segundo a autora, “garante a diversidade de gêneros e permite a abrangência de diferentes situações de enunciação, incluindo a interação, sendo notável a representatividade da língua falada, encontrada na simulação que dela fazem as peças teatrais” (Ibidem, p. 14).

Esse estudo aproxima-se do que desenvolvemos nesta dissertação, pois tem um caráter descritivo e resulta de consultas a bases de dados da língua (corpus). No entanto, o material consultado por Moura Neves não está disponível ao público, o que dificulta a reprodução das pesquisas feitas na gramática. Todo o material que consultamos durante a pesquisa, por outro lado, está disponível na internet.

Moura Neves ressalta que, embora uma gramática de usos não seja, em sua essência, normativa, a norma de uso é apresentada em sua gramática com intuitos comparativos, indicando as restrições tradicionalmente feitas a certos usos. Para estudar os verbos, Moura Neves propõe a divisão em dois grandes grupos que, por sua vez, são divididos em inúmeros subgrupos. Nós nos concentraremos, porém, nas categorias que interessam a este estudo.

Os dois grandes grupos são (1) os verbos que constituem predicados e (2) os verbos que não constituem predicados. Uma vez que o grupo (2) é constituído principalmente por verbos auxiliares e modalizadores, analisaremos aqui apenas o grupo (1).

A Figura 4 mostra um esquema parcial das subclassificações do grande grupo (1). Chamamos de “parcial” porque nos preocupamos em mostrar apenas as ramificações das subclassificações que se encontram dentro do escopo desta pesquisa.

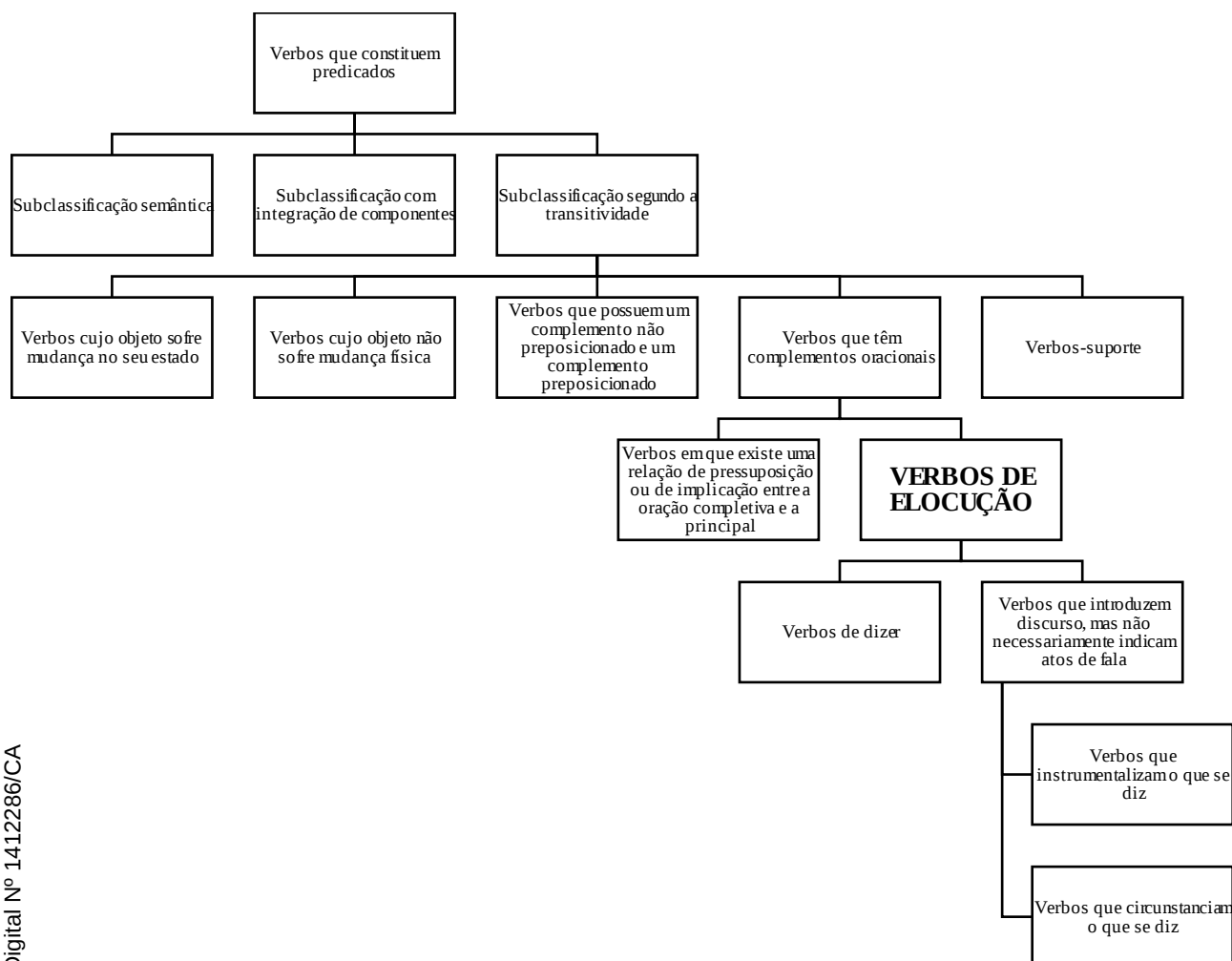


Figura 4: Esquema da classificação dos verbos proposta por Moura Neves (2000), reduzida às classificações que interessam a este estudo.

É possível notar que se trata de uma classificação bastante ampla. A autora parte de três critérios de classificação diferentes, sendo um semântico e dois sintáticos (integração de componentes e transitividade), para, então, analisar os verbos. Apesar de uma classificação semântica interessar a este estudo, a subclassificação semântica feita por Moura Neves é muito breve e apenas divide os verbos que constituem predicados em dinâmicos (isto é, que denotam ações ou processos) e não dinâmicos (que indicam estados). A subclassificação que nos interessa é a feita segundo a transitividade. A partir dessa divisão, foram criados diversos subgrupos, todos a partir de um critério sintático.

Dentro do subgrupo de “verbos que têm complementos oracionais”, Moura Neves apresenta os “verbos de elocução”, definidos como os “verbos introdutórios

de discurso (discurso direto ou discurso indireto)” (Ibidem, p. 47). Após defini-los, a autora estabelece uma importante distinção entre o discurso direto e o indireto:

No discurso direto, o falante tem uma responsabilidade muito menor sobre a oração completa, que é uma citação direta, ficando o controle das expressões correferenciais e dêiticas (de referência à situação) circunscrito à própria oração citada, e, portanto, independente de referência ao falante [...]. O discurso indireto não envolve citação literal do que o sujeito diz, mas constrói uma paráfrase pela qual o falante assume a responsabilidade do que é referido, além de controlar a correferência dos pronomes e dos advérbios dêiticos, já que a dêixis deixa de ficar centrada no sujeito do verbo da completa. (Ibidem, p-47-48).

Moura Neves chama de “verbos de dizer” os “verbos de elocução propriamente ditos”, isto é, os “verbos de ação cujo complemento direto é o conteúdo do que se diz” (Ibidem, p. 48). Ao lado dos verbos de dizer, a autora propõe os “verbos que introduzem discurso, mas não necessariamente indicam atos de fala”, grupo este que também integra os verbos de elocução.

Incluem-se no grupo dos verbos de dizer, também chamados pela autora de verbos *dicendi*, os verbos *falar* e *dizer*, que seriam “neutros”, juntamente com diversos outros verbos “cujo significado traz, somado ao dizer básico, informações sobre o modo de realização do enunciado”, como os verbos *gritar*, *berrar* e *sussurrar*. Também seriam verbos de dizer aqueles que acrescem “noções sobre a cronologia discursiva”, como é o caso dos verbos *retrucar* e *repetir* (Ibidem, p. 48).

Moura Neves destaca que, entre os verbos de dizer, existem também aqueles verbos nos quais se encontra “lexicalizado o modo que caracteriza esse dizer” (Ibidem, p. 48). A autora cita como exemplos os verbos *queixar-se*, *comentar* e *responder*, que podem ser parafraseados por *dizer uma queixa*, *dizer um comentário* e *dizer uma resposta*.

Para Moura Neves, os verbos de dizer implicam necessariamente um ato de fala, e, por isso, a autora optou por criar um segundo subgrupo para aqueles verbos que nem sempre introduzem discurso e indicam atos de fala. Esse segundo subgrupo, por sua vez, é dividido em (a) verbos que instrumentalizam o que se diz e (b) verbos que circunstanciam o que se diz.

Em (a) estão incluídos verbos como *ameaçar* e *acalmar*, “que indicam ações realizadas com o uso de um instrumento, que pode consistir, eventualmente, em um dizer. Pode-se, por exemplo, ameaçar alguém com uma faca, com um gesto ou com palavras”. Já em (b) a autora inseriu os verbos “que expressam uma ação ou um processo” realizado “ao mesmo tempo que o dizer. Indicam, então, as

circunstâncias que caracterizam o ato de fala” (Ibidem, p. 49). Nesse grupo incluem-se verbos como *rir*, *chorar* e *espantar-se*.

O Quadro 2 constitui uma esquematização das subclasses criadas por Moura Neves e que dizem respeito a esta pesquisa. As subclasses estão representadas por ordem de hierarquia, da mais ampla para a mais específica:

SUBCLASSE	EXEMPLO
Verbo de elocução	<i>E o pior é que ela sabia assinar. Aí, diz que o padre tirou o papel do bolso e disse: “Então assine aqui”.</i>
a) Verbos de dizer ou verbos <i>dicendi</i>	<i>Michelângelo, diante de um bloco de mármore de Carrara, exclamou que ali dentro estava Moisés.</i>
b) Verbos que introduzem discurso, mas não necessariamente indicam atos de fala	
b.1) Verbos que instrumentalizam o que se diz	<i>Raul ameaçou-o com os punhos: – Olhe, que eu lhe dou uns tabefes.</i>
b.2) Verbos que circunstanciam o que se diz	<i>–Como? – espantou-se. –Quer prestar exames no Ateneu e me vem com “um tiquinho” para Aritmética?</i>

Quadro 2: Esquema das classes criadas por Moura Neves (2000), com os respectivos exemplos.

Uma vez que os verbos de elocução se encontram inseridos no grupo dos verbos que têm complementos oracionais, Moura Neves oferece alguns quadros em que são analisados os tipos de complemento aceitos pelos verbos de elocução e nos quais o discurso introduzido se apresenta. Com base nas Figuras 5 a 9, contabilizamos 103 verbos compilados pela autora como verbos de elocução.

VERBOS	TIPOS DE DISCURSO		FORMA DO COMPLEMENTO			
	DISCURSO DIRETO	DISCURSO INDIRETO	ORAÇÃO INFINITIVA	ORAÇÃO CONJ. QUE	ORAÇÃO CONJ. SE	SINTAGMA NOMINAL (nominalização)
exclamar	x	x	x	x	---	---
explicar	x	x	x	x	---	x
expor	x	x	---	x	---	x
falar	x	x	x	x	---	de/sobre x
frisar	x	x	x	x	---	x
garantir	x	x	x	x	---	x
gritar	x	x	x	x	---	---
informar	x	x	x	x	---	x
insinuar	x	x	x	x	---	x
insistir (em)	x	x	x	x	---	x
jurar	x	x	x	x	---	---
lembrar	x	x	x	x	---	x
negar	x	x	x	x	---	x
observar	x	x	x	x	---	---
ordenar	x	x	---	x	---	---
participar	x	x	x	x	---	x
perguntar	x	x	---	---	x	de/sobre x
ponderar	x	x	x	x	---	---
pregar	x	x	x	x	---	x
prevenir	x	x	x	x	---	---
proclamar	x	x	x	x	---	x
prometer	x	x	x	x	---	---
protestar	x	x	x	x	---	---
queixar-se	x	x	---	x	---	---
questionar	x	x	x	x	x	x
reafirmar	x	x	x	x	---	x
reconhecer	x	x	x	x	---	x
reiterar	x	x	x	x	---	x
relatar	x	x	x	x	---	x
repetir	x	x	x	x	---	---
replicar	x	x	x	x	---	---
resmungar	x	x	---	x	---	---
responder	x	x	x	x	---	---
ressaltar	x	x	x	x	---	x
retrucar	x	x	x	x	---	---
revelar	x	x	x	x	---	x

Figura 5: Tipos de complemento aceitos pelos verbos de simples dizer e verbos que qualificam o que é dito – parte 1.

Fonte: Moura Neves (2000).

VERBOS	TIPOS DE DISCURSO		FORMA DO COMPLEMENTO			
	DISCURSO DIRETO	DISCURSO INDIRETO	ORAÇÃO INFINITIVA	ORAÇÃO CONJ. QUE	ORAÇÃO CONJ. SE	SINTAGMA NOMINAL (nominalização)
aconselhar	x	x	---	x	---	x
afirmar	x	x	x	x	---	x
alegar	x	x	x	x	---	x
antecipar (-se)	x	x	x	x	---	x
anunciar	x	x	x	x	---	x
argumentar	x	x	x	x	---	---
arrematar	x	---	---	---	---	---
assegurar	x	x	x	x	---	x
avisar	x	x	x	x	---	de x
berrar	x	x	x	x	---	---
boquejar	x	x	x	x	---	---
citar	x	---	---	---	---	---
cochichar	x	x	x	x	---	---
comentar	x	x	x	x	---	x
completar	x	---	---	---	---	---
comunicar	x	x	x	x	---	x
concluir	x	x	x	x	---	---
concordar	x	x	x	x	---	com x
confessar	x	x	x	x	---	x
confiar	x	x	x	x	---	x
confidenciar	x	x	x	x	---	x
confirmar	x	x	x	x	---	x
considerar	x	---	---	---	---	---
contar	x	x	x	x	---	---
continuar	x	---	---	---	---	---
criticar	x	---	---	---	---	---
declarar	x	x	x	x	---	x
determinar	x	x	x	x	---	x
destacar	x	x	x	x	---	x
diagnosticar	x	x	x	x	---	x
dizer	x	x	x	x	---	---
emendar	x	x	---	x	---	---
ênfatizar	x	x	x	x	---	x
esclarecer	x	x	x	x	---	---

Figura 6: Tipos de complemento aceitos pelos verbos de simples dizer e verbos que qualificam o que é dito – parte 2.

Fonte: Moura Neves (2000).

VERBOS	TIPOS DE DISCURSO		FORMA DO COMPLEMENTO			
	DISCURSO DIRETO	DISCURSO INDIRETO	ORAÇÃO INFINITIVA	ORAÇÃO CONJ. QUE	ORAÇÃO CONJ. SE	SINTAGMA NOMINAL (nominalização)
salientar	x	x	x	x	---	x
sugerir	x	x	x	x	---	x
suplicar	x	x	---	x	---	---
sussurrar	x	x	x	x	---	---
tornar	x	---	---	---	---	---

Figura 7: Tipos de complemento aceitos pelos verbos de simples dizer e verbos que qualificam o que é dito – parte 3.

Fonte: Moura Neves (2000).

VERBOS	TIPO DE DISCURSO		FORMA DO COMPLEMENTO			
	DISCURSO DIRETO	DISCURSO INDIRETO	ORAÇÃO INFINITIVA	ORAÇÃO CONJ. QUE	ORAÇÃO CONJ. SE	NOMINALIZAÇÃO
acalmar	x	---	---	---	---	---
agastar-se	x	---	---	---	---	---
aguilhoar	x	---	---	---	---	---
ameaçar	x	x	---	x	---	x
apelar	x	---	---	---	---	---
bronquear	x	---	---	---	---	---
bulir	x	---	---	---	---	---
caçoar	x	---	---	---	---	---
chamar	x	---	---	---	---	---
chorar	x	---	---	---	---	---
conchavar	x	---	---	---	---	---
consolar	x	---	---	---	---	---
cumprimentar	x	---	---	---	---	---
debicar	x	---	---	---	---	---
debochar	x	---	---	---	---	---
desafiar	x	---	---	---	---	---
desiludir	x	---	---	---	---	---
escarnecer	x	---	---	---	---	---
espantar-se	x	---	---	---	---	---
ferroar	x	---	---	---	---	---

Figura 8: Tipos de complemento aceitos pelos verbos que instrumentalizam ou circunstanciam o que é dito – parte 1.

Fonte: Moura Neves (2000).

VERBOS	TIPO DE DISCURSO		FORMA DO COMPLEMENTO			
	DISCURSO DIRETO	DISCURSO INDIRETO	ORAÇÃO INFINITIVA	ORAÇÃO CONJ. QUE	ORAÇÃO CONJ. SE	NOMINALIZAÇÃO
inclinar-se	x	---	---	---	---	---
interceptar	x	---	---	---	---	---
interromper	x	---	---	---	---	---
maldizer	x	---	---	---	---	---
remediar	x	---	---	---	---	---
rir	x	---	---	---	---	---
suspirar	x	---	---	---	---	---
zombar	x	---	---	---	---	---

Figura 9: Tipos de complemento aceitos pelos verbos que instrumentalizam ou circunstanciam o que é dito – parte 2.

Fonte: Moura Neves (2000).

Apesar de seu trabalho ser o mais exaustivo até então, Moura Neves não explica como foi o processo de levantamento dos verbos encontrados. Lembramos que essa falta de explicação se soma ao fato de que o corpus usado não pode ser acessado livremente, e isso tudo dificulta a expansão e o aprofundamento do que Moura Neves fez por parte de outros pesquisadores.

Dos trabalhos consultados no português, a *Gramática de usos do português* é o único que corresponde a um estudo descritivo com base em corpus. Apesar dos problemas que mencionamos, a gramática foi basilar para o nosso estudo, pois nos forneceu a nomenclatura usada aqui e, ainda, pudemos enriquecer a análise dos nossos resultados, tentando reproduzir o que Moura Neves fez, porém com os verbos que encontramos e com os corpora que consultamos. Dessa forma, os nossos contra-argumentos para alguns dos verbos citados por Moura Neves serão apresentados durante a análise dos nossos resultados, no capítulo 4.

Oliveira et al. (1985)

O artigo “Verbos introdutores de discurso direto” teve origem com as anotações das atividades de descrição de verbos, feitas para o desenvolvimento de um dicionário gramatical de verbos. As autoras indicam que o comportamento irregular dos verbos introdutores de discurso direto foi a motivação para tentar classificá-los. A partir de uma abordagem descritiva e que tem sempre em vista o ensino, Oliveira et al. (1985, p. 91) debruçam-se sobre as dificuldades na transformação do discurso direto em discurso indireto. É importante ressaltar que, tal como Othon M. Garcia, as autoras adotam uma postura um tanto normativa em seu artigo.

No artigo, Oliveira et al. (1985) propõem-se a estabelecer um critério para identificar e classificar os verbos que introduzem discurso direto. Ainda, as autoras procuram apontar “os fatores que condicionam a possibilidade [...] de transposição do discurso direto para o indireto” (Ibidem, p. 91). Para tal, fazem uso de exemplos retirados de gramáticas escolares, manuais de ensino de português e de coletâneas de crônicas.

Segundo Oliveira et al. (1985, p. 91), os manuais de ensino e as gramáticas escolares voltam suas atenções principalmente para a descrição da transformação do discurso direto em indireto, o que, na opinião delas, pode levar à conclusão errônea de que “qualquer enunciado em discurso direto pode ser transformado,

feitas as devidas adaptações, em discurso indireto”. Para pôr em questionamento essa conclusão, fornecem os exemplos (1) e (2), abaixo. De acordo com a classificação proposta pelas autoras, a conversão de discurso direto para discurso indireto nas construções com os verbos *tranquilizar* e *orientar* não seria possível em português.

- (1) *(a velhinha) tranquilizou também a vizinha de banco:*
 – *Ela vai trabalhar no ministério; eu vou para casa, moro no Rio Comprido.*
- (2) *a nave foi auxiliada em terra por dois crioulos que orientavam as manobras: “Dá ré, agora vira tudo pra direita, isso, chega um pouquinho à frente, tá bom aí.”*

A classificação de Oliveira et al. (1985) divide os verbos em dois grupos: verbos elocutórios e verbos não elocutórios. Estes grupos, por sua vez, definem-se e subdividem-se da seguinte maneira:

- 1) Verbos elocutórios: “contêm implícito em seu significado um comportamento de fala[,] ou seja, um *dictum*”.
 - a) Verbos de dizer: “verbos de ação cujo complemento direto é o *dictum*”. Exemplos: *dizer, falar, gritar*, etc.
 - b) Verbos que qualificam o *dictum*: “verbos de ação que apresentam lexicalizada a modalização que caracteriza o *dictum* neles implícito”. Exemplos: *queixar-se, comentar, protestar, avisar*, etc.
- 2) Verbos não elocutórios: “indicam ação, processo ou ação-processo e [...] se realizam independentemente de um comportamento de fala”.
 - a) Verbos que instrumentalizam o *dictum*: “verbos que indicam ações realizadas mediante o uso de um instrumento, que pode consistir [...] em um *dictum*”.
 Exemplos: “*ameaçar com uma faca/ameaçar com a fala*”.
 - b) Verbos circunstanciais: “expressam uma ação ou processo que pode realizar-se ao mesmo tempo que o *dictum*”.
 Exemplo: *rir, chorar, resmungar*, etc.

Com base nas classificações de Oliveira et al. (1985), o verbo *tranquilizar* poderia enquadrar-se na categoria de verbos circunstanciais, que se encontra dentro do grupo de verbos não elocutórios. Uma das principais conclusões do artigo é que os verbos não elocutórios não admitem a transformação em discurso indireto, uma

vez que o *dictum*, isto é, a coisa dita, não é o complemento do verbo. Assim, a conversão para o discurso direto só seria possível com a explicitação do verbo *dizer*. Convertemos abaixo o exemplo (1) para o discurso indireto:

- (3) A velhinha *tranquilizou* também a vizinha de banco, *dizendo* que ela trabalharia no ministério...

No entanto, em uma busca em todos os corpora disponíveis no projeto AC/DC⁷, foram encontradas as seguintes linhas de concordância:

- (4) *Então quero fazer uma declaração para país, para o mundo, no sentido de **tranquilizar** que a lei será aprovada, é uma lei revolucionária, estabelece novas bases para o mercado de telecomunicações do Brasil, gera um mercado competitivo, ao contrário do que alguns setores vêm dizendo, porque estabelece uma competição obrigatória para cada modalidade de serviço.*

Na concordância (4), “a lei será aprovada” corresponde ao *dictum* e, ao mesmo tempo, ocupa a posição de complemento do verbo *tranquilizar*. Portanto, esse emprego do verbo *tranquilizar*, convertido para o discurso indireto, põe em xeque as regras estabelecidas por Oliveira et al. (1985). Ainda que se argumente que a construção “tranquilizar que” causa estranhamento ao falante de português, a existência de uma ocorrência mostra que é possível, sim, transformar essa construção em discurso indireto, sem que a comunicação fique prejudicada de fato. Deve-se destacar, no entanto, que essa construção parece fazer-se presente mais na oralidade.

Enquanto o exemplo do verbo *tranquilizar* possa parecer mais radical, outros verbos citados pelos autores como não elocutórios são com mais frequência empregados no discurso indireto, como os verbos *orientar*, *resmungar* e *chorar*:

- (5) «Quando me divorciei, com duas filhas de 1 e 2 anos, uma psicóloga **orientou** que eu deixasse a elas a iniciativa de tocar na questão de minha homossexualidade.
- (6) Não surpreende que muitos homens sejam vistos **resmungando** que as mulheres estão roubando seus empregos.
- (7) Andaram para aí a **chorar** que o papá Dhlakama não vinha.

Oliveira et al. (1985) criticam o manual de ensino de Othon Garcia, por este apresentar como critério definidor de um verbo de elocução o fato de este indicar o interlocutor que detém a palavra. Para elas, esse critério permitiria a qualquer verbo

⁷ Usamos a expressão de busca [lema=“tranquilizar”] [word=“que”], que pede qualquer forma do verbo *tranquilizar*, seguida pela palavra *que*.

assumir a posição de introdutor de discurso. A categorização proposta pelas autoras, entretanto, exclui uma gama de verbos que, ao contrário do que elas afirmam, podem funcionar como verbos de elocução.

Loffredo et al. (2004)

As autoras do artigo “Verbos de elocução – as diferenças entre o inglês e o português” tiveram como objetivos principais (i) contrastar o uso dos verbos de elocução em diálogos extraídos de textos em inglês e português e (ii) analisar a forma como os tradutores lidam com esses verbos, tanto em traduções quanto em versões. Trata-se de um trabalho que se aproxima desta pesquisa, uma vez que foi movido pela mesma motivação inicial: “Suspeitávamos que no inglês não havia muita variação na escolha do léxico, enquanto no português a variação era quase obrigatória, para tornar o texto mais compreensível e natural” (Loffredo et al., 2004, p. 168).

As autoras conduziram um estudo descritivo com base em corpus, mais especificamente, a partir do COMPARA, mesmo corpus de onde começamos esta dissertação. Loffredo et al. (2004, p. 168) pretendiam determinar as diferenças entre inglês e português, a partir de textos autênticos, escritos originalmente nas duas línguas. Além disso, tinham como propósito sugerir o que usar na tradução de cada verbo encontrado. Por um lado, esta pesquisa não tem como foco a comparação do uso dos verbos de elocução entre inglês e português. A ideia de fornecer sugestões de verbos na hora de traduzir, por outro lado, é bem semelhante à nossa proposta de compilação de um glossário dessa classe de verbos.

Para iniciar as buscas, tal como fizemos, Loffredo et al. (2004, p. 168) optaram por pesquisar por *said*, a fim de poder filtrar melhor as ocorrências e ter mais chances de se deparar com discurso relatado. A justificativa foi que, ao buscar pelos verbos de elocução isoladamente, o número de resultados tinha sido excessivo, e, para melhor trabalhar com os dados, as autoras tiveram de desenvolver uma espécie de filtro na hora de avaliar quais casos estavam no escopo do artigo (buscando por *said* em vez de *say*, por exemplo).

originais para tradução - 174 (879)

tradução para originais - 162 (527)

Dizer (49%)	Dizer (91%)
Responder (14,4%)	Sem verbo (3%)
Sem verbo (9%)	Falar (2,5%)
Perguntar (8,6%)	Responder
Comentar	Perguntar
Exclamar	Insistir
Explicar	Refletir
Concluir	
Declarar	
Insistir	
Confirmar	

Figura 10: Número de ocorrências obtidas com uma busca por *said* em diálogos.
Fonte: Loffredo et al. (2004).

A análise feita pelas autoras sobre os resultados da primeira coluna é que “na maior parte das vezes em que *say* foi traduzido para o português optou-se pelo verbo dizer e raramente um verbo diferente de *dizer*, *perguntar* ou *responder* ocorreu nas traduções”. Já a segunda coluna, que mostra *say* como a tradução de algum verbo em português, mostra que esse verbo quase sempre corresponde a uma tradução de *dizer*. Em outras palavras, “quando o tradutor se depara com outro verbo diferente de dizer, praticamente não utiliza *say*, traduzindo-o por um equivalente mais próximo da língua de origem, mas que talvez não soe tão natural na língua de chegada” (Ibidem, p. 170).

Os resultados da segunda coluna não se encontram no escopo deste estudo, mas a primeira coluna reflete uma etapa da pesquisa bem semelhante à etapa inicial desta dissertação. A análise, no entanto, é um pouco diferente, pois as autoras chegaram à seguinte conclusão:

[F]oi comprovado com esse trabalho que, talvez por se preocupar muito com a fidelidade da tradução, por falta de tempo para analisar cada contexto, ou simplesmente por pouca observação da parte dos tradutores, os verbos de elocução são geralmente traduzidos de forma literal. (Loffredo et al., 2004, p. 183).

E, ainda, fazem a seguinte sugestão:

Seria bastante coerente, portanto, que em traduções do inglês para o português, o verbo *say* em expressões como *he said* se transformasse em um verbo diferente de *dizer*, dependendo da intenção comunicativa da personagem (e do autor) ou de seu tom de voz. (Ibidem, p. 183)

Curiosamente, partindo dos mesmos dados, nossa conclusão é radicalmente oposta: do nosso ponto de vista, é bastante interessante que, em metade das vezes, o tradutor tenha tido a iniciativa de empregar outros recursos, considerando que somos inevitavelmente influenciados pelo texto que traduzimos. Mais adiante, reproduzimos a mesma busca por *said*, porém com um número maior de ocorrências analisadas.

A sugestão dada pelas autoras, apesar de alinhada com as motivações desta pesquisa, é muito vaga e em pouco ajuda os tradutores que desejam se aprimorar como profissionais. Foi pensando nisso que decidimos ir além e compilar um léxico dos verbos que podem desempenhar essa função no português, de modo a não só chamar a atenção dos tradutores para essa diferença de estilo como também munir os tradutores de recursos variados.

2.2

Os verbos de elocução na Linguística Computacional

Na Linguística Computacional, o discurso relatado comparece de maneira direta na tarefa de Identificação de Citações (Quotation Extraction), cujo objetivo principal é identificar as citações em um texto, relacionando-as a seus autores. O foco está na identificação de quem fala e no conteúdo dessa fala, e em geral privilegia-se o discurso direto, cujas marcas formais facilitam a detecção automática. O resultado dessa tarefa interessa a organizações, analistas e atores políticos, motivados pela busca por feedback com relação a produtos, sejam eles objetos de mercado, pessoas ou iniciativas políticas. Relatamos abaixo o trabalho de Sagot et al. (2010), no âmbito das tarefas de identificação de citação.

Sagot et al. (2010)

O artigo de Sagot et al. (2010) apresenta o processo de criação de um léxico de *quotation verbs*, tendo em vista o desenvolvimento de uma ferramenta automática de extração de citação.

Segundo os autores, as técnicas de extração de informações para recuperar citações são úteis para sintetizar as informações sobre um determinado assunto. Essas informações compreendem inclusive citações de pessoas relevantes para o tema em questão. O principal requisito para o desempenho dessas técnicas é um

estudo que cubra as diversas formas de citação, bem como suas estruturas e suas propriedades lexicais (Sagot et al., p. 294).

Com base em um corpus da Agence France-Presse (AFP), Sagot et al. propõem-se a analisar as citações na língua francesa. Os autores indicam que as citações, em francês, podem apresentar as seguintes configurações:

(1) Citação indireta (IQ)

Exemplo: Le sénateur a déclaré (à la ministre) qu’il allait “défendre cette loi contre vents et marées.”

Tradução: O senador declarou (à ministra) que defenderia “essa lei contra tudo e todos”.

(2) Citação direta com o verbo que a introduz na posição inicial (DI)

Exemplo: Le sénateur a déclaré (à la ministre): “Je vais défendre cette loi contre vents et marées.”

Tradução: O senador declarou (à ministra): “Eu vou defender essa lei contra tudo e todos”.

(3) Citação direta com o verbo que a introduz em uma oração apositiva, em posição intermediária ou final (DP)

Exemplo: “Je vais défendre cette loi contre vents et marées”, a déclaré le sénateur (à la ministre).

Tradução: “Eu vou defender essa lei contra tudo e todos”, declarou o senador (à ministra).

Como o objetivo dos autores é a construção de uma ferramenta automática de extração de citação, eles estabeleceram algumas perguntas que norteariam a pesquisa:

[W]hich are the (French) verbs that introduce a quotation in an IQ, DI and/or DP configuration (such verbs are called “quotation verbs”)? Which lexical entry should be given to a quotation verb in a syntactic lexicon used by parsers, possibly within a quotation extraction system? (Sagot et al., 2010, p. 294)

A partir de uma lista prévia de 110 verbos, compilados em um estudo anterior, Sagot et al. analisaram um corpus de 5.000 notícias de jornais da AFP, procurando por estruturas que correspondessem às configurações IQ, DI e DP, explicadas acima. Após filtrar os resultados manualmente, chegaram a 836

configurações de citação associadas a verbos de citação. Dos 110 verbos iniciais, os autores encontraram 91 (Ibidem, p. 295).

Em seguida, os autores classificaram as ocorrências manualmente, o que os permitiu chegar às seguintes constatações:

- A maioria dos verbos ocorre em estrutura de DP, isto é, com o verbo posposto à citação direta: 70 de 91 verbos;
- Entre esses 70 verbos, 20 foram encontrados apenas nessa estrutura de DP, pois a maioria desses 20 verbos não é tipicamente associada a verbos introdutores de discurso relatado;
- A distribuição entre todos os verbos entre as três estruturas mostra que a configuração mais frequente é a de DP (65,8%), seguida por IQ (20,6%) e por DI (2,9%).

Uma vez que as estruturas de DP se revelaram mais produtivas, os autores decidiram concentrar-se nessa configuração e abandonar as demais. Depois de desenvolver expressões regulares para as estruturas de DP do corpus, Sagot et al. (2010, p. 295) obtiveram 232 verbos associados a essa configuração.

Em seguida, os autores reinseriram esses verbos em todas as expressões regulares que haviam desenvolvido e verificaram a distribuição dos verbos entre as três configurações, IQ, DI e DP. As conclusões foram que: (1) qualquer verbo em uma estrutura de IQ pode ocorrer também em DP; e (2) apenas metade dos 232 verbos são verbos de elocução que podem introduzir citação indireta (IQ).

Após uma análise cuidadosa dos 232 verbos, Sagot et al. (2010, p. 295) dividiram os verbos em três classes:

- Classe 1: 118 verbos de elocução transitivos, tais como *dire/dizer* e *répondre/responder*, que aparecem em estruturas de IQ e DP;
- Classe 2: 43 verbos intransitivos, como *ricaner/rir* e *fulminer/fulminar*, que aparecem em DP mas não em IQ;
- Classe 3: 71 verbos transitivos, como *commenter/comentar* e *continuer/continuar*, que também ocorrem em DP mas não em IQ.

A pesquisa de Sagot, Danlos e Stern foi uma importante referência para esta dissertação, pois inspirou algumas das nossas investigações e análises. O caminho

seguido por nós é bastante semelhante: com base em corpus, conduzimos um levantamento semiautomático dos verbos de elocução do português. Analisamos as estruturas em que esses verbos mais comumente apareciam e, a partir daí, levantamos centenas de verbos empregados nessas configurações.

3

Levantamento dos verbos de elocução

Para alcançar um dos objetivos principais desta pesquisa, a saber, a criação do glossário DISSE, precisávamos, antes de tudo, conduzir um levantamento dos verbos de elocução que comporiam o glossário. Neste capítulo, relatamos as etapas para a realização desse levantamento; também se encontra descrito aqui todo o processo subjacente à identificação dos verbos de elocução em grandes corpora, que culminou em uma lista de 293 verbos.

Antes de descrever essas etapas individualmente, apresentamos abaixo um fluxograma que ajuda a esclarecer o caminho que percorremos para compilar o glossário.

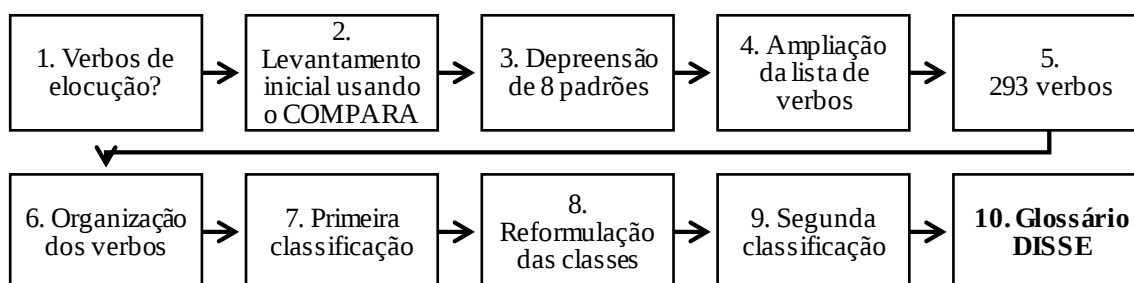


Figura 11: Fluxograma com as etapas metodológicas desta pesquisa.

3.1

Por onde começamos: as traduções de *said*

Para dar início à lista (ou léxico) de verbos, o primeiro passo foi realizar um levantamento dos verbos usados para introduzir a fala em português. Considerando o enquadramento desta pesquisa, optamos por obter os dados iniciais com o COMPARA (Frankenberg-Garcia & Santos, 2002), usando a versão 13.1.22. Trata-se um corpus paralelo com anotação morfossintática, disponível on-line.

A opção pelo COMPARA se deve ao fato de que o corpus é composto por obras literárias editadas, escritas originalmente em inglês ou em português, que são alinhadas com suas respectivas traduções. Nesse gênero textual, é grande a presença

de verbos que introduzem discurso, o que contribui para a relevância de tal corpus nesta pesquisa.

Vale ressaltar que o COMPARA constitui uma reunião de obras literárias que são célebres a ponto de serem traduzidas para outras línguas. Todas as traduções foram publicadas por editoras (em oposição a traduções “livres” publicadas na internet, por exemplo). Isso significa que as traduções encontradas no corpus são consideradas válidas por profissionais da área, o que nos dá mais segurança para aceitar os verbos escolhidos para traduzir *said* como verbos de elocução. Entre os autores traduzidos para o português, estão Henry James, Lewis Carroll e Oscar Wilde.

No COMPARA, buscamos traduções para a forma verbal *said* (no original). Usamos a expressão de busca [word="said"]⁸, com a restrição de resultados que correspondessem apenas a traduções do inglês para o português. A opção pela busca da forma *said* (em vez de *say*) se dá porque o verbo, na maioria das vezes, introduz discurso em sua forma pretérita, o que é confirmado por Biber et al. (1999, p. 374).

Nem sempre é óbvia a identificação de um dado verbo como introdutor de discurso relatado (por exemplo, *imaginar* ou *interromper*). A escolha pela exploração inicial no COMPARA com a forma *said* buscou, também, garantir segurança quanto à seleção dos verbos: quando o texto original usa a forma *said*, não há dúvidas de que o verbo escolhido, em português, se refere a um discurso relatado:

TEXTO FONTE: «Don't,» *I said*, in a muffled voice.

TEXTO META: – Não! – interrompi numa voz abafada.

A busca pelas traduções de *said* nos forneceu um material bastante rico para o nosso ponto de partida, sendo desnecessário recorrer a outras formas verbais de *say* ou a outros verbos equivalentes.

Como o COMPARA limita os resultados a mil linhas de concordância aleatórias, usamos exatamente essa amostra, apesar de o corpus apresentar 3.560 ocorrências de *said*. Foram analisadas, então, todas as linhas de concordância obtidas com a referida busca.

⁸ A expressão indica que queríamos apenas a palavra “said” nessa forma exata e mais nada, sem que se considerassem possíveis variações.

The screenshot shows the COMPARA search interface. On the left is a sidebar with navigation links like 'Linguatca', 'Estrutura', 'Equipe', 'Apresentação', 'Acesso a recursos', and 'Avaliação conjunta'. The main area is titled 'Resultados da pesquisa' and displays search results for the query 'said'. It includes a table with columns for the source text (ERDL1T1), the original text, and the translated text. The table shows several instances of 'said' being translated into Portuguese, such as 'respondei', 'disse-lhe', 'disse o Tom', 'disse-me', 'disse a enfermeira', 'disse-me num tom encorajador', 'exclamou', and 'respondeu-me'.

Figura 12: Interface da busca no COMPARA (Frankenberg-Garcia & Santos, 2002).

Com o objetivo de extrair os diferentes verbos escolhidos pelos tradutores como substitutos de *said*. Eliminamos 42 ocorrências, pois não correspondiam à introdução de um relato. Segue abaixo um exemplo dessas 42 ocorrências:

ORIGINAL: «What they do, what Polly *said*, oh please let's Ange!»

TRADUÇÃO: – O que eles fazem, o que *disse* a Polly, por favor, Ange!

O Quadro 3 lista as demais 958 ocorrências de *said*, organizada pelo número de ocorrências:

VERBO	OCORRENCIAS
dizer	561
responder	88
perguntar	60
omissão	53
comentar	20
explicar	17
acrescentar	9
afirmar	9
contrapor	9
exclamar	9
retorquir	9
continuar	8

VERBO	OCORRENCIAS
argumentar	2
contar	2
cumprimentar	2
esclarecer	2
propor	2
queixar-se	2
querer saber	2
admirar-se	1
admitir	1
advertir	1
anuir	1
anunciar	1

declarar	7
insistir	6
interromper	6
prosseguir	6
repetir	5
pedir	4
replicar	4
assegurar	3
comunicar	3
concluir	3
confessar	3
confirmar	3
falar	3
informar	3
observar	3
retrucar	3
ripostar	3
agradecer	2

balbuciar	1
berrar	1
brindar	1
censurar	1
concordar	1
desabafar	1
escrever	1
espantar-se	1
indagar	1
justificar-se	1
lembrar	1
prometer	1
saudar	1
sondar	1
sublinhar	1
sugerir	1
tornar	1
TOTAL	958

Quadro 3: Lista de traduções para a forma *said* em 958 ocorrências, compilada no COMPARA.

Fonte: COMPARA (Frankenberg-Garcia & Santos, 2002)

Apesar de o número de ocorrências do verbo *dizer* ser significativamente maior do que o número dos demais verbos, correspondendo a 58,56% das traduções, é interessante notar que os dados corroboram o que a prática tradutória já afirmava: em cerca de 40% dos casos, os tradutores optaram por não empregar o verbo *dizer*, fazendo uso de 58 verbos distintos, que incluem desde verbos frequentemente empregados como introdutores de discurso, como *perguntar* e *responder*, e outros verbos não tão comuns nessa posição, como *sondar*, *agradecer*, *brincar*, etc. Chama a atenção também o fato de que, em apenas 53 ocorrências, os tradutores preferiram omitir o verbo.

Como é possível verificar no Quadro 3, vários verbos escolhidos para substituir o *dizer* introduzem discurso com mais precisão, indicando, muitas vezes, intenções, sentimentos e ações específicas que não estão expressas, e são, portanto, inseridas pelo tradutor. A seguir, temos um exemplo dessa interferência:

ORIGINAL: So I *said* he couldn't do that, and he *said* yes he could, and then he *said* he wanted to anyway, he wants...'

TRADUÇÃO: Então eu *expliquei*-lhe que ele não podia fazer isso, ele *insistiu* que podia e depois *disse* que, de qualquer modo, era isso que ele queria, que ele havia de fazer...»

A etapa seguinte teve como objetivo ampliar a lista de 58 verbos, por meio de uma estratégia semiautomática. Para tanto, foi feita a exploração, em momentos diferentes, de grandes corpora monolíngues do projeto AC/DC (Costa et al., 2009): o CHAVE (Santos & Rocha, 2005), de textos jornalísticos (98 milhões de palavras); o OBras, de obras de literatura brasileira disponíveis em domínio público (1,2 milhão de palavras), e o Floresta (Freitas et al., 2008), majoritariamente jornalístico (6 milhões de palavras).

3.2

Depreensão de padrões de uso

Dos 58 verbos iniciais, 6 foram escolhidos para atuar como “sementes” na interface de busca em corpora AC/DC, usando inicialmente o corpus CHAVE, a fim de identificar *padrões* léxico-gramaticais tipicamente usados para introduzir um discurso relatado. Subjacente à identificação dos padrões de uso desses verbos, está a ideia de que os próprios padrões seriam capazes de facilitar a localização de demais verbos do dizer no corpus. Os verbos escolhidos foram *dizer*, *perguntar*, *responder*, *admitir*, *contar* e *continuar*. Os três primeiros verbos foram selecionados por serem frequentemente referenciados como verbos de elocução, e realmente foram os mais frequentes nas traduções (ver Quadro 3). Além disso, esses verbos são normalmente associados à classe de verbos *dicendi* em gramáticas e manuais de redação e, por isso, podemos considerá-los prototípicos como verbos de elocução.

Por outro lado, os verbos *admitir*, *contar* e *continuar* foram escolhidos por estarem associados a usos que nem sempre correspondem à elocução, como, por exemplo:

O evento contou com a presença do arcebispo de Fortaleza, d. Aloísio Loscheider.

Às 13h, a reportagem da Folha contou 83 pessoas na manifestação.

Apesar de ser o maior corpus paralelo revisto de português e inglês do mundo, o COMPARA dispõe de apenas 723.807 palavras de traduções para o português – além de estar restrito a textos ficcionais, em geral romances. Por isso, para a observação dos verbos-semente em contexto e consequente depreensão de seus padrões léxico-sintáticos, recorreremos a um corpus mais robusto: o CHAVE. Trata-se de um corpus monolíngue, composto por textos jornalísticos – portanto, repleto de indicações da fala de outros – retirados dos jornais Folha de S. Paulo (Brasil) e Público (Portugal). O CHAVE contém 97.884.763 palavras, um número mais de cem vezes maior do que a quantidade de palavras do COMPARA. Esse número mais elevado poderia nos dar acesso a uma variedade maior de estruturas usadas no português para introduzir discurso relatado.

Além disso, o fato de o CHAVE e o COMPARA serem de gêneros diferentes (jornalístico e literário, respectivamente), também contribuiu para essa escolha, uma vez que a variedade poderia fornecer uma visão mais ampla do uso dessa classe de verbos. Não deixamos de apontar, ao longo deste trabalho, as diferenças encontradas no emprego dos verbos introdutores de discurso relatado entre os textos literários e jornalísticos.

A ampla frequência dos verbos de elocução no CHAVE se, por um lado, corrobora o acerto na escolha do material, por outro, aponta para a inviabilidade da análise caso a caso de todas as ocorrências, nos obrigando a desenvolver estratégias para lidar com a grande quantidade de informações.

O Quadro 4, produzido com a ferramenta Distribuidor, apresenta a distribuição dos seis verbos no CHAVE, considerando a frequência e especificando a distribuição conforme a variante da língua portuguesa. No entanto, é importante mencionar que, na tabela, a frequência dos verbos não leva em conta o fato de estarem sendo usados como verbos de elocução ou não. Isso significa que ambas as ocorrências a seguir, por exemplo, foram consideradas nessa contabilização:

«A literatura é que seduz novos leitores», insiste Eliana Yunes.

Ao contrário: políticos europeus insistem numa ofensiva contra os estrangeiros.

VERBO	FREQUÊNCIA TOTAL	VARIANTE	FREQUÊNCIA PARCIAL	%
dizer	225.493	BR	121.618	53,93%
		PT	103.875	46,07%
continuar	52.614	BR	15.993	30,40%
		PT	36.621	69,60%
contar	30.309	BR	12.218	40,31%
		PT	18.091	59,69%
responder	17.480	BR	6.260	35,81%
		PT	11.220	64,19%
admitir	15.083	BR	4.321	28,65%
		PT	10.762	71,35%
perguntar	10.256	BR	5.072	49,45%
		PT	5.184	50,55%
TOTAL	351.235	-	-	-

Quadro 4: Distribuição dos seis verbos analisados, organizados por ordem de frequência no corpus CHAVE e distribuídos por variante.

Como estratégia para lidar com as mais de 350 mil ocorrências a fim de depreender regularidades, analisamos, primeiro, os três verbos que tipicamente indicam elocução: *dizer*, *perguntar* e *responder*. Lemos centenas de ocorrências desses verbos e avaliamos em que casos introduziam uma fala. A partir da análise do contexto desses verbos “gerais”, foi possível depreender oito padrões. Em seguida, verificamos se os demais verbos, *admitir*, *contar* e *continuar*, também eram usados com esses padrões.

Os oito padrões estão esquematizados a seguir. Nos padrões, chamamos de “citação completa” a citação marcada pela pontuação que costuma indicar fala (aspas curvas, aspas angulares e travessões) e que pode constituir um período completo:

- (1) Citação completa + verbo: «*Talvez tenha sido mal-interpretado*», disse.
- (2) Verbo + citação completa: *Até que uma amiga minha passou por ele e disse: «Oi, Fábio».*
- (3) Citação intercalada pelo verbo: «*Se não tivessem sido feitas*», disse, “*Portugal era hoje um país ao nível do Leste Europeu*”.
- (4) Verbo + citação introduzida por conjunção subordinativa: *Cauteloso, ele disse que não receberá empresários e empreiteiras.*

- (5) Verbo + citação em oração subordinada substantiva objetiva direta reduzida de infinitivo: *Em entrevista de dez minutos à TV russa, ele disse estar controlando o país.*
- (6) Verbo + pronome pessoal oblíquo + adjetivo ou particípio: *O governo de Israel se disse surpreso com críticas do enviado do Vaticano a Israel, Andrea Di Montezemolo.*
- (7) Verbo + oração reduzida de gerúndio + citação: *O Pp, por meio de Rocco Buttiglione, respondeu, aconselhando a leitura de «Mein Kampf» (Minha Luta), de Adolf Hitler, para entender porque o líder da Liga, senador Umberto Bossi, «age com um Führer (guia, como os alemães chamavam Hitler)».*
- (8) Verbo em oração subordinada adverbial conformativa + citação: *Afinal, como disse Boris Casoy, Hebe paga impostos e é assídua no trabalho.*

A maneira mais comum de se descrever o discurso relatado é de acordo com o tipo de citação envolvida: citação direta, indireta e mista (esta última presente apenas em alguns trabalhos). Esta classificação, no entanto, não captura o fato de que um mesmo padrão pode ser usado por mais de um tipo de citação.

A maior importância dada ao tipo de citação refere-se à relevância deste tipo de classificação para evidencialidade do que é relatado. No nosso caso, o interesse nos padrões (e não nos tipos de citação) é resultado de sua relevância para as buscas em corpora: os padrões – formalizações – correspondem a diferentes materializações das estruturas de elocução. Uma consequência da ênfase no tipo de citação é a ideia de que há uma relação entre um padrão e um tipo de citação. Por exemplo, o padrão Verbo + QUE invariavelmente se associaria ao discurso indireto. A observação dos contextos, no entanto, revelou que um mesmo padrão pode participar de diferentes tipos de citação. O Quadro 5 mostra a distribuição dos padrões por tipo de citação, e é interessante perceber como um mesmo padrão pode estar associado às diferentes formas de citar. No capítulo 5, retomamos os padrões, apresentando aspectos interessantes do seu uso em português.

PADRAO	DIRETA	INDIRETA	MISTA	EXEMPLOS
1	x			«O aumento não vale para as diárias, só mensalistas », afirmou Paulo Octávio.

2	x			Até que uma amiga minha passou por ele e disse : «Oi, Fábio».
3	x	x	x	D: «Na França », diz ela, « a história é uma atividade muito prestigiada, e portanto muito masculina». I: Antes de tentar um acordo, ele passou em outra concessionária (não informou o nome), que, disse , teria lhe oferecido arcar com parte do II. M: A contribuição de Robin Williams a «Aladdin », salienta ele, « nenhum diretor, escritor ou animador poderia suprir».
4		x	x	I: Cauteloso, ele disse que não receberá empresários e empreiteiras. M: Ambas acabam por reconhecer que « roubos, sempre os houve», mas acrescentam que «agora é muito pior».
5	x	x	x	D: Quando os jornalistas perguntaram se se tratava da reeleição dele ou de Fidel, FHC respondeu, rindo: «Para ele, não é reeleição, ele quer a eternidade». I: O banco respondeu, informando-os que o depósito em causa tinha sido liquidado em 30 de Agosto de 1985, e que havia pago um cheque ao portador nesse montante, assinado pelo João. M: Vital Moreira respondeu , afirmando manter «tudo o que disse» e que «a comunicação social escreveu, sem que tenha havido desmentidos».
PUC-Rio - Certificação Digital Nº 1412286/CA 6		x	x	I: Ontem, Farouk Kaddoumi, dirigente do departamento político da OLP, se disse favorável a realização de uma nova conferência multilateral de paz, conforme proposta da Rússia. M: Ontem, na reapresentação na Vila Belmiro, ele se disse abatido, «principalmente por ter sido contra o São Paulo, de quem eu queria ganhar».
7	x	x	x	D: Quando os jornalistas perguntaram se se tratava da reeleição dele ou de Fidel, FHC respondeu, rindo: «Para ele, não é reeleição, ele quer a eternidade». I: Charles respondeu dizendo que não podia se envolver nessa questão. M: A acusação partiu de Miltiadis Evert, o líder da Nova Democracia, e o porta-voz do governo socialista, Telmahos Hyritis, respondeu acusando-o de «críticas irresponsáveis».
8	x	x	x	D: Ou ainda, como escreveu o também poeta Joseph Brodsky: «A verdadeira biografia de um poeta é quase idêntica à dos pássaros». I: O presidente Itamar Franco deve vetar o projeto, segundo informou o ministro da Fazenda, Fernando Henrique Cardoso. M: Se há «graves impropriedades» no relatório, como afirma o Planalto, elas devem decerto ser identificadas e contestadas.

Quadro 5: Distribuição dos padrões por tipo de citação.

Abaixo, são apresentados os oito padrões, com a descrição do uso feita para cada um dos seis verbos analisados. Para facilitar a leitura, optamos por já indicar ao longo da descrição dos padrões, alguns resultados que já só fruto de análise. No entanto, como é a própria análise que contribui para o refinamento dos padrões, optamos por mantê-la nessa seção.

3.2.1

Descrição dos padrões

(1) Citação completa + verbo de dizer:

“ – « + citação completa + ” – » + (,) ⁹ + verbo + (agente)
“ – « + citação completa + ” – » + (,) + (agente) + verbo

- (a) «Ninguém da minha família participa de sequestro», **disse** Silva.
- (b) «Tudo que você escreve vende, não é mesmo?», **perguntou** Larry.
- (c) «É mais do que isso», **respondeu** Gabus Mendes.
- (d) «Continuamos sem grandes soluções atacantes», **admite** Toni, treinador adjunto de Manuel Cajuda.
- (e) «Perto de nós, 30 ou 40 pessoas comemoravam o aniversário de uma grã-fino», **conta** Luis Gustavo.
- (f) «O Sampaio, então, ficará a semana inteira sem fazer nada, pois nem jogou», **continuou** o técnico.

Este talvez seja um dos padrões mais comuns e mais reconhecíveis dos verbos introdutores de discurso relatado. Com citações diretas, fica fácil identificar os argumentos da citação. No caso da primeira linha de concordância, por exemplo, temos como identificar o autor da citação (“Silva”), o ato em si (“disse”) e a coisa dita (“Ninguém da minha família participa de sequestro”). Nesse tipo de estrutura, não é obrigatória a explicitação do autor da citação.

A pontuação característica deste padrão seria um dos facilitadores na identificação desses argumentos. No entanto, vale ressaltar que essa pontuação nem sempre está presente:

- (g) Os carros estavam saindo de traseira, **explicou** Keith Wiggins, o chefe da equipe.

⁹ Na representação dos padrões, os parênteses indicam que o elemento em questão é opcional.

A linha de concordância (g) constitui um caso em que o discurso é relatado de forma direta e sem o auxílio de pontuação típica de diálogos (aspas ou travessão). Esses casos não são raros e, portanto, não poderiam deixar de ser mencionados aqui.

Ainda sobre a pontuação, as aspas costumam figurar em textos jornalísticos, enquanto em outros gêneros textuais, como, por exemplo, no texto literário, as aspas são frequentemente substituídas por travessões. No caso das citações com travessão, a vírgula costuma ser omitida, como é possível observar no exemplo abaixo, que pertence à amostra de mil ocorrências retiradas do COMPARA:

(h) – *E como podemos fazer uma nova série sem Debbie? – eu **disse**.*

Este padrão foi identificado primeiro nas buscas pelas ocorrências do verbo *dizer*, mas foi encontrado, em seguida, entre as linhas de concordância de todos os outros verbos. O fato de encontrarmos os seis verbos neste padrão pode indicar que se trata de um padrão de uso bastante produtivo e, portanto, a partir dele, podemos ter mais chances de encontrar outros verbos introdutores de discurso relatado. Essa hipótese se confirma mais adiante, na seção 3.1.3.

(2) Verbo de dizer + citação completa:

verbo + : + “ | « + citação completa + ” | »

verbo + : + – + citação completa

- (a) *Zico, em conversa exclusiva com a Folha, **disse**: «Estou muito contente com essa homenagem».*
- (b) *Uma hora após Gaviria ter evitado responder de forma categórica à pergunta, a Folha **perguntou** a Fujimori: «Houve transparência nas eleições peruanas?»*
- (c) *E a mulher **respondeu**: «Alexandre não tem pão em seu reino?»*
- (d) *O ministro da Educação, Murílio Hingel, **admite**: «Apesar de todos os esforços, nosso ensino básico é vexaminoso.»*
- (e) *Com a autoridade de um dos primeiros diretores brasileiros a trabalhar para a MTV Brasil, ele **conta**: «Não existe um clipe meu que não tenha ficado entre os mais votados na MTV».*

- (f) *Ele **continua**: «Aqui eu acho que eu retomo a urgência, de um novo ângulo.»*

Pertencente ao discurso direto, tal como o Padrão (1), este também é um padrão extremamente corrente no português. A pontuação comumente empregada neste padrão (dois-pontos, travessões e aspas), tal como no Padrão (1), pode contribuir para a facilidade na identificação dos argumentos. Aqui, as aspas também podem ser substituídas por travessões, como mostra o exemplo (b):

- (g) *Ao ouvir isto, Sofia virou-se para sua mãe, e empunhando a escova de lavar pratos, **disse**:*
– Saiba que ele é perfeitamente bom da cabeça.

Tal como no anterior, o Padrão (2) foi inicialmente encontrado nas ocorrências de *dizer*, para, em seguida, ser corroborado pelos demais verbos. Por conta da peculiaridade dos Padrões (1) e (2), com marcas formais específicas e sempre associados a citações diretas, esses padrões podem ajudar bastante na extração de novos verbos introdutores de discurso relatado. Em outras palavras, essas duas estruturas raramente são empregadas para algo que não seja introdução de relato, o que nos dá uma garantia maior de que o verbo encontrado em tais padrões seja um verbo de elocução.

(3) Citação intercalada pelo verbo de dizer

“ | « + parte da citação + (” | ») + – | , + verbo + (agente) + – | , + (” | ») + parte da citação + “ | «
 “ | « + parte da citação + (” | ») + – | , + (agente) + verbo + – | , + (” | ») + parte da citação + “ | «
 (–) + parte da citação + – | , + (agente) + verbo + – | , + parte da citação
 (–) + parte da citação + – | , + verbo + (agente) + – | , + parte da citação

- (a) *«Na Califórnia – **diz** ela – é cada vez maior o número de pessoas que come pouca carne ou que se tornam vegetarianas».*
- (b) *Como, **perguntará** o leitor, o execrando Tratado de Maastricht prevê uma coisa dessas?*
- (c) *«Vossa Santidade», **respondi**, «não só eu estou contente, todos nós estamos muito contentes».*

- (d) «Sem dúvida alguma, ele é o melhor boxeador do mundo!», **admite**, «mas enquanto ele continuar dizendo que é o melhor lutador, estou pronto a desafiá-lo».
- (e) O helicóptero do líder do ANC foi apedrejado por manifestantes do Inkhata e a própria residência do rei foi atacada, **contaram** os seus conselheiros, obrigando o soberano a procurar refúgio numa quinta nos arredores.
- (f) Em 93, **continua**, a produção mundial foi de 18,7 milhões de toneladas e o consumo mundial, de 19,4 milhões.

Este tipo de padrão costuma ser empregado, principalmente, com citações mais longas. O Padrão (3) admite os três tipos de citação e pode ser encontrado em textos tanto jornalísticos quanto literários. O que difere o Padrão (3) nos dois gêneros textuais é a pontuação; enquanto o jornalismo costuma optar pelas aspas, a literatura tem predileção pelos travessões.

Nesse padrão, não é obrigatória a explicitação de todos os argumentos, mais especificamente do agente. É bastante comum que o sujeito da ação esteja em um período anterior.

(4) Verbo de dizer + citação introduzida por conjunção subordinativa

(agente) + verbo + (interlocutor) + conjunção integrante + citação
--

- (a) Sobre a lateral esquerda, Parreira **disse** que quem deve jogar amanhã é Leonardo.
- (b) Me **perguntaram** se esse era o momento oportuno.
- (c) Genro, em tom duro também, **respondeu** que aquela não era uma visão de um dirigente de expressão nacional como Dirceu.
- (d) O teólogo católico **admite**, no entanto, que o tema das mulheres é um problema real, ao qual a Igreja Católica também tem que dar resposta.
- (e) FHC **contou** aos parlamentares que coube ao próprio Arida a indicação de Loyola.

- (f) *Mas ante um persistente e interrogativo olhar do administrador, como que a intimá-lo a falar, ele **continuou** que não lhe tinha emprestado as pedras por uma questão de confiança, porque neste aspecto até poderia dizer que não confiava nem nos seus próprios dentes porque de quando em vez lhe mordiam a língua...*

Por se tratar de uma estrutura simples e comum na língua portuguesa, o Padrão (4) não impõe dificuldades que se refere à identificação dos argumentos. Essa fácil identificação pode ser atribuída, entre outros fatores, ao fato de que o padrão é rígido em termos de deslocamento sintático. O autor da citação ocupa o lugar convencional de sujeito, antecedendo a oração subordinada. Na linha de concordância (a), por exemplo, temos a citação, na sua forma indireta (“quem deve jogar amanhã é Leonardo”), a ação (“disse”) e o autor da citação (“Parreira”). Devemos ressaltar, entretanto, que a presença do autor da citação não é obrigatória; nesse caso, é possível encontrar o referente do sujeito em um período anterior.

Apesar de o Padrão (4) ser mais comumente associado às citações indireta e mista, encontramos nos corpora alguns casos de introdução de citação direta e completa por meio do verbo, seguido de conjunção integrante. Trata-se de uma variação do Padrão (4), com a adequação ao discurso direto (inclusão de dois-pontos):

- (g) ***Acrecentou** ainda que: «A prazo, o financiamento do Ensino Superior deverá crescer, a par do aumento da sua frequência e dos resultados alcançados.»*

O padrão (4) é bastante recorrente na língua portuguesa, por conta da ampla gama de verbos que podem ser empregados dessa forma. No entanto, é justamente essa variedade que dificulta o nosso trabalho de identificar verbos de elocução, pois essa estrutura não é necessariamente associada ao relato de discurso:

- (h) *Pessoalmente, **considero** que Mariano Gago foi, de longe, o melhor ministro da era «democrática».*

Na concordância (g), teríamos, a princípio, um indício de citação mista. Contudo, o verbo *considerar*, tal como os verbos *admitir*, *contar* e *continuar*, têm outras sentidos que não correspondem ao de verbos de elocução. No exemplo (h),

o fato de o verbo estar no presente e na primeira pessoa do singular já aparece como um impedimento para o relato, pois não julgamos ser possível relatar algo que ainda não foi dito. Justamente por conta do tempo verbal e da primeira pessoa, acreditamos que, em (g), o verbo expressa o sentido de “definir-se, depois de reflexão, sobre (determinada coisa); julgar” (Houaiss, 2016). Além disso, temos que uma das marcas formais típicas de diálogos, as aspas, não cumprem apenas essa função. Muitas vezes, usamos as aspas apenas como recurso de ênfase ou destaque, como acreditamos ser o caso acima.

A oração subordinada substantiva objetiva direta que compõe o padrão (4) também é possível em sua forma incompleta, com as respostas “sim” e “não”:

- (i) *Nem o governo, nos momentos de maior sinceridade, seria capaz de **responder** que sim.*

O verbo *perguntar* apresenta-se neste padrão de forma diferente dos demais verbos, assemelhando-se, por sua vez, a outros verbos que denotam pergunta. Como corresponde a uma pergunta convertida para o discurso indireto, encontramos comumente a conjunção integrante “se” no lugar de “que”, tal como na linha de concordância (b), além de outras conjunções integrantes e de pronomes interrogativos:

- (j) ***Perguntei** por que eles queriam que justo a Executive Outcomes fizesse isso.*
- (k) *Tivemos o desabafo do presidente da República quando um jornalista **perguntou** o que ele faria se dependesse do salário mínimo.*

(5) Verbo + citação em oração reduzida de infinitivo

(agente) + verbo + oração subordinada substantiva objetiva direta reduzida de infinitivo

- (a) *Um dia antes de disputar a eleição, Sanguinetti **disse** achar que seu país está chegando ao Mercosul em desvantagem com os demais.*
- (b) *Sem muita convicção, um deles **respondeu** saber de tal urgência.*
- (c) *O ministro russo de Cultura, Evguêni Sidorov, **admitiu** ter tocado no tesouro.*

- (d) *Contou ter feito a campanha de lançamento na Argentina da marca Cica.*

Em uma busca no corpus CHAVE, obtiveram-se 870 ocorrências de verbos de elocução com complemento no infinitivo. Encontramos 73 verbos diferentes na posição de complemento, e listamos os mais frequentes no Quadro 6.

LEMA	OCORRENCIAS
ter ¹⁰	458
ser	171
estar	38
saber	21
haver	14
fazer	12
passar	11
existir	9
ir	9
mostrar	7
melhorar	6
acabar	6
verificar	5
mudar	5
cair	5
parar	4
correr	4
voltar	4
criar	3
pertencer	3
deixar	3
aparecer	3
demonstrar	3

fechar	3
ficar	3
amadurecer	2
resistir	2
discordar	2
fumar	2
gastar	2
responder	2
rever	2
vir	2
funcionar	2
representar	2
repetir	2
chegar	2

¹⁰ O verbo *ter* possui o maior número de ocorrências porque, na maioria das linhas de concordância, funciona como verbo auxiliar.

Ouadro 6: Complementos mais frequentes dos verbos de elocução no Padrão (5).
Fonte: CHAVE (Santos & Rocha, 2005).

No que se refere aos argumentos, o Padrão (5) apresenta a mesma estrutura e o mesmo processo de identificação da citação, do autor e do verbo introdutor de discurso relatado, mostrados no padrão (4). A rigidez no deslocamento dos argumentos também se faz presente neste padrão, o que facilita o processo de identificação.

Foram encontradas apenas 21 ocorrências do verbo *continuar* seguido de infinitivo¹¹. No entanto, em todas elas, *continuar* tem o sentido de “persistir”. É bem incomum que o complemento do verbo *continuar*, quando se trata de um infinitivo, venha desacompanhado da preposição “a”, como na linha de concordância (e):

- (e) *Tirar um curso superior em Portugal **continua** depender das condições económicas.*
- (f) *O grupo dos times pequenos **continua** a existir.*

Na Tabela 1, são comparadas as ocorrências com e sem a preposição e, ainda, de acordo com as variantes do português:

VARIANTE	COM PREPOSIÇÃO	SEM PREPOSIÇÃO
PT	17253	48
BR	1664	3

Tabela 1: Comparação dos tipos de complementos do verbo *continuar* no Padrão (5) entre as variantes do português.

Fonte: CHAVE (Santos & Rocha, 2005)

(6) Verbo + pronome pessoal oblíquo + adjetivo ou particípio

pronome pessoal oblíquo + verbo + adjetivo particípio
verbo + pronome pessoal oblíquo + adjetivo particípio

- (a) *O governo do México se **disse** disposto a negociar uma trégua.*

¹¹ Para tanto, foi usada a expressão [pos="V.*" & lema="continuar"]@[pos="V" & temcagr="INF"]

Esse é um padrão relativamente frequente do verbo *dizer* – , foram encontradas 212 ocorrências¹². Contudo, é difícil determinar se o verbo, nesse padrão, está sendo usado como verbo de elocução, uma vez que tem como intuito atribuir característica ao autor da suposta citação. No caso da linha de concordância (a), o governo do México (autor) considerou que estava disposto a negociar uma trégua, ou seja, o governo do México atribuiu a si a característica “disposto”.

(b) *Ele se **disse** traído por Viljoen.*

Na linha de concordância (b), que difere da (a) por seu complemento ser um particípio e não um adjetivo, observa-se a mesma limitação quanto ao intuito do locutor. As questões são: “traído por Viljoen” é a coisa dita? Tem-se aqui, de fato, uma citação? O fato de que pouquíssimos verbos são empregados nessa estrutura já minimiza as chances de o Padrão (6) ser um bom indício de que temos um verbo de elocução. Entre os seis verbos analisados, apenas *dizer* apresentou ocorrências que se enquadram no padrão e, em uma busca posterior, obtivemos apenas cinco verbos empregados no Padrão (6):

LEMA	OCORRÊNCIAS
dizer	26
declarar	6
revelar	3
reconhecer	2
defender	1

Quadro 7: Verbos encontrados dentro do Padrão (6).

Fonte: CHAVE (Santos & Rocha, 2005)

(7) Verbo + oração reduzida de gerúndio + citação

verbo de dizer + (,) + gerúndio + conjunção integrante

¹² Foi usada a expressão [pos="PERS_obj"] [pos="V.*" & lema="dizer"] [pos="ADJ.*"], cujo sentido é: [pos="PERS_obj"]: expressão para buscar por um pronome pessoal oblíquo (objeto). [pos="V.*" & lema="dizer"]: pede-se que seja um verbo e que seja uma variante de “dizer”. [pos="ADJ.*"]: o verbo deve ser seguido por qualquer tipo de adjetivo.

O verbo *dizer* não foi encontrado introduzindo discurso com esse padrão. As linhas de concordância obtidas corroboram a teoria de que a oração reduzida de gerúndio acrescenta uma informação nova ao período, em vez de simplesmente reiterar o verbo de dizer que a antecede:

(a) «É, na época eu era aluno de pós-graduação», **diz** brincando.

Assim como o verbo *dizer*, o verbo *perguntar* não foi encontrado nesse padrão. As ocorrências de *perguntar* nessa estrutura correspondem a outros usos das orações reduzidas de gerúndio, como, por exemplo, para indicar ações simultâneas ou consecutivas:

(b) “Este aqui?”, **perguntou**, apontando para um homem moreno, cheio de brilhantina no cabelo e bigodinho à Errol Flyn.

(c) “Quem recebe é corrupto”, **perguntou**, provocando risos na plateia.

Os verbos *responder* e *admitir* são outros casos que deixam dúvidas. Por exemplo, seria a oração reduzida uma caracterização do *responder*? Seria uma ação adicional à ação de responder? As ocorrências encontradas para esses dois verbos parecem não corroborar o Padrão (7) como um forte candidato a padrão de uso dos verbos introdutores de discurso relatado.

(d) A hierarquia católica costuma **responder** dizendo que a igreja não é uma democracia.

(e) Mais modesto é, sem dúvida, o balanço em relação à evolução conseguida nas taxas de juro: «tem vindo a verificar-se uma evolução no sentido favorável, embora permaneçam no mercado de crédito taxas de juro reais elevadas», **admite**, sublinhando no texto entregue aos jornalistas que «a descida das taxas de juro não tem sido uniforme nos vários segmentos (dívida pública, crédito a grandes empresas, «PME Prestígio» e outras PME e particulares)».

Nas poucas ocorrências de oração reduzida de gerúndio após o verbo *contar*, o verbo da oração reduzida não auxilia na introdução de discurso, mas sim atribui

algum tipo de característica ou especificidade ao ato de contar, tal como se vê nos exemplos abaixo:

- (f) *Ele relata as histórias que lhe **contaram**, usando palavras que mostram que ele sabia aquilo a que se referia ao escrever.*

Não foram encontradas ocorrências do verbo *continuar* sendo empregado como introdutor de discurso relatado e seguindo o Padrão (7). Em todas as ocorrências em que é sucedido por um verbo no gerúndio, *continuar* apresenta o sentido de “prosseguir”:

- (g) *Disse que reconhecia nele a virtude de **continuar** dizendo o que fez.*

(8) Verbo em oração subordinada adverbial conformativa + citação

conjunção subordinativa conformativa + verbo + agente + , + (“ | «) + citação
 conjunção subordinativa conformativa + agente + verbo + , + (“ | «) + citação
 citação (“ | ») + , + conjunção subordinativa conformativa + verbo + agente
 citação (“ | ») + , + conjunção subordinativa conformativa + agente + verbo

- (a) *E como **diz** o locutor Fiori Giglioti em suas transmissões de futebol, «o tempo passa, torcida brasileira».*
- (b) *Então, como **perguntou** um professor presente no Curso de Verão, «não será necessário mudar o essencial da relação pedagógica, da filosofia do ensino e da organização da escola?»*
- (c) *«Boa pergunta!», como **responderia** ele.*
- (d) *Coca-Cola, eles provaram pela primeira vez no navio Merida, como **contou** à Folha o frentista Hugo Berto Isquierdo, 22.*
- (e) *«O reforço das acções necessários» para a prossecução desses objectivos de luta contra a evasão e a fraude é, como **continuou**, «uma das prioridades da DGCI» e para o qual tanto este departamento como a Inspecção-Geral de Finanças eram «organismos vitais».*

Esse padrão caracteriza um caso bem peculiar do uso do verbo de elocução no que se refere aos argumentos. No exemplo acima, pode-se ver que o que temos não é o simples relato de um discurso, mas sim a apropriação de uma fala proferida

anteriormente. Nesse caso, o autor da citação no momento da enunciação é, de fato, o locutor, mas, fora isso, não temos mais pistas que permitam a identificação desse indivíduo. Por outro lado, fica explícito o autor “original” da citação, isto é, o indivíduo que proferiu a citação pela primeira vez. Somente o contexto no qual a citação se insere poderia confirmar a identidade do autor que detém a fala no instante em questão.

Em contraste com os sete padrões anteriores, o Padrão (8) dispõe de uma estrutura sintática um pouco mais rígida. Mais especificamente, nesse padrão, não é possível omitir o sujeito do verbo introdutor de discurso relatado. A explicação provavelmente está fundada na peculiaridade da apropriação de uma fala anterior pelo locutor da frase; como o “dono original” da fala não é o sujeito do período que a introduz no relato, torna-se obrigatório explicitá-lo.

Abaixo, encontram-se algumas variações desse padrão, que é observado em todos os tipos de discurso:

- (a) A América, como *dizia* Miles Davis, é um país maravilhoso onde você pode reunir mil vozes para gravar um uníssono.
- (b) A resposta, meu amigo, é sussurrada pelos ventos, como *diria* a canção.
- (c) E como *disse* o pagodeiro na MTV: tudo na vida é passageiro, menos o cobrador e o motorista.
- (d) Até chegar a uma «produtora estruturada», como *diz* Sandra, foram «meses de angústia».

Os verbos *perguntar*, *responder* e *continuar*, segundo as nossas pesquisas, não costumam ser empregados no Padrão (8), apresentando três, duas e uma ocorrências, respectivamente. Apesar de os números não serem muito altos, a possibilidade de emprego dos seis verbos no padrão o torna um bom candidato para nos ajudar a levantar os verbos introdutores de discurso relatado.

Outra seção: Quando os verbos de elocução não *dizem*

Uma vez definidos os candidatos a padrões de uso para os verbos introdutores de discurso relatado, cabe também esclarecer em que circunstâncias os referidos verbos podem não desempenhar tal função ou de fato não o fazem. Foram

destacados os casos mais interessantes e mais frequentes, e listamos abaixo conforme o verbo envolvido.

Dizer

- 1) Presença em expressões multivocabulares: “dizer respeito”, “querer dizer”.

*Outro crime **diz** respeito a uma nota fria emitida em favor da Gonair Táxi Aéreo.*

*Isso quer **dizer** que os preços reais subiram.*

Perguntar

- 1) Apesar de ser o único desta lista, este é um caso complexo. O complemento preposicionado da linha de concordância abaixo, “sobre Josenildo da Silva”, não indica exatamente a coisa dita, mas sim o assunto da coisa dita. Por não se tratar do relato de uma fala, o *perguntar*, neste contexto, não foi considerado um verbo introdutor de discurso relatado. Tomamos o mesmo posicionamento em todos os casos em que o complemento do verbo em questão não era oracional.

*Os assassinos **perguntaram** sobre Josenildo da Silva, um dos sete filhos da mulher.*

Responder

- 1) Quando o complemento do verbo *responder* é um sintagma nominal ou preposicional, consideramos que não há introdução de citação. No exemplo abaixo, o conteúdo da resposta, isto é, a coisa dita, respondida, não está presente.

*Não serão **respondidas** cartas com pedidos de aconselhamento médico ou psicológico.*

*Como físico, sinto dificuldades para **responder** a essas questões.*

- 2) O verbo “responder”, acompanhado da preposição “por”, pode ter o sentido de “responsabilizar-se”.

*E para piorar ainda mais este já lamentável quadro, esses políticos não **respondem** por seus atos.*

- 3) Também seguido pela preposição “por”, “responder” pode significar “equivaler”.

*Os doze maiores **respondem** por 95 % do setor, segundo estudos da Fipe / USP.*

Admitir

- 1) Além do sentido de “reconhecer”, o verbo “admitir” é comumente empregado com a ideia de “permitir”.

*Pois lá a concorrência é muito grande e não são **admitidos** erros.*

Contar

- 1) Um dos sentidos considerados primários de “contar” é “calcular”.

*Toda vez que tem que entrar em seus novos trajes de Mauricinho, Marzo **conta** até dez.*

- 2) Outro sentido de “contar” que não introduz discurso é “dispor”.

*As próximas de Santarém **contam** com infraestrutura turística.*

- 3) “Contar” também é muito empregado com um complemento que é um sintagma nominal: “uma anedota”, “uma história”. Esses sintagmas nominais não relatam a coisa dita, mas sim categorizam o que foi dito: aquilo que foi dito é uma anedota, é uma história, etc.

*Hoje só vou **contar** piada.*

- 4) Na expressão “é o que conta”, o verbo “contar” não funciona como verbo introdutor de discurso relatado, pois expressa ideia de “valor”.

*Claro que se pode alegar que o critério do árbitro é o que **conta**.*

Continuar

- 1) Quando empregado como introdutor de discurso relatado, o verbo *continuar* de fato passa a ideia de “prosseguimento”. Essa mesma ideia se mantém, mesmo quando o verbo não se refere à fala.

*O setor empresarial não pode **continuar** nessa passividade.*

3.2.2

Análise da produtividade dos padrões encontrados

A análise da seção anterior está esquematizada na Tabela 2, a seguir, que mostra quais padrões foram aceitos por cada verbo analisado:

VERBO \ PADRÃO	1	2	3	4	5	6	7	8
DIZER	sim	sim	sim	sim	sim	sim	não	sim
PERGUNTAR	sim	sim	sim	sim	não	não	não	sim
RESPONDER	sim	sim	sim	sim	sim	não	sim	sim
ADMITIR	sim	sim	sim	sim	sim	não	sim	sim
CONTAR	sim	sim	sim	sim	sim	não	não	sim
CONTINUAR	sim	sim	sim	sim	não	não	não	sim

Tabela 2: Ocorrências dos seis verbos em cada padrão analisado.

Com base na Tabela 2, é possível esboçar duas conclusões. A primeira delas é que todos os verbos analisados, por serem aplicáveis na maioria dos padrões estabelecidos, podem ser classificados como introdutores de discurso relatado. Ainda que o verbo *continuar* não tenha sido encontrado em três dos oito padrões, isso não o exclui da classe de verbos, principalmente porque, na verdade, talvez devam ser considerados apenas cinco padrões – o que nos leva à segunda conclusão.

Como foi dito anteriormente, os padrões gramaticais utilizados com os verbos de elocução seriam propostos a partir da análise dos seis verbos “semente” e, simultaneamente, esses mesmos seis verbos colocariam os padrões encontrados à prova. Uma vez que somente o verbo *dizer* é aplicável ao Padrão (6) e só os verbos *responder* e *admitir* se encaixaram satisfatoriamente no Padrão (7), a generalidade desses padrões é questionável e estes não nos ajudam muito a identificar verbos de elocução. Por isso, daqui em diante, esses padrões serão deixados de lado.

O Padrão (5) foi aceito por quatro dos seis verbos, o que, para nós, é uma sugestão de que o padrão pode servir para introduzir discurso relatado. No entanto, a sua pouca precisão levaria, inevitavelmente, à recuperação de inúmeros outros verbos como candidatos a verbos de elocução, mesmo sem sê-lo. Por isso, decidimos desconsiderá-lo. De fato, tínhamos outros cinco que serviam melhor ao nosso propósito: construir um léxico dos verbos de elocução a partir dos padrões. Considerando os 5 principais padrões, passamos à etapa seguinte, de ampliação dos verbos.

3.3

Ampliação da lista de verbos de elocução: a construção de um léxico

Com base nos padrões, criamos as expressões que seriam utilizadas nos corpora do AC/DC para a ampliação do léxico dos verbos. As expressões estão listadas no Anexo 1. Depois de selecionar os cinco padrões relatados anteriormente, a partir deles, desenvolvemos expressões de busca na linguagem dos corpora do AC/DC. Com essas expressões, pudemos realizar um amplo levantamento dos verbos empregados em cada um dos padrões, e, assim, compilamos uma extensa lista dos verbos introdutores de discurso relatado da língua portuguesa.

É importante ressaltar que alguns padrões possuem variações quanto à forma como se apresentam. No entanto, selecionamos apenas uma ou duas variações de cada padrão para desenvolver as expressões, visto que as buscas, mesmo com um número mais limitado de estruturas, já foram bastante produtivas.

Logo percebemos que os padrões, sozinhos, não garantiriam a presença de um verbo de elocução, uma vez que recuperavam coisas como os exemplos abaixo, indicando a necessidade de outra análise caso a caso:

- (a) *Verificada a licitude da «pré-campanha », **regressemos** ao caso do leitor.*
- (b) *Ao longo de quase dois anos, os habitantes de Reveles ainda **acreditaram** **que** os trabalhos de estabilização dos solos pudessem vir a evitar o pior.*

Foi o que fizemos, desta vez com um corpus de tamanho inferior ao CHAVE, mas ainda assim de grandes dimensões: usamos o corpus Floresta (Freitas et al., 2008). O Floresta é composto por 6.046.541 palavras.

Para cada expressão de busca usada, pedimos que nos fosse fornecida a distribuição dos verbos por lema, como lista a Figura 13:

Resultados da procura

Wed Mar 23 04:21:48 WET 2016

Procura: [word="";*] » [—] [word=";"] @[pos="V.*" & temcagr!=";*.PCP.*|.*.GER.*|.*.PASSIVA.*|.*.FUT_SUBJ.*|.*.PR_SUBJ.*|.*.FUT_IND.*|0|.*.PR_PROG_IND.*|.*.IMPF_FUTHAVER_IND.*|.*.IMPF_SUBJ.*|.*.COND_IND.*"]

Distribuição de lema

Corpo: Floresta Sintá(c)tica v. 2.5

8040 casos.

Distribuição

Houve 581 valores diferentes de lema.

dizer	2318
afirmar	913
explicar	468
contar	365
ser	222
comentar	159
lembrar	124
revelar	108
acrescentar	91
declarar	90
concluir	83
completar	81
perguntar	77
garantir	71
brincar	64
responder	61
referir	55
observar	54

Figura 13: Interface do corpus Floresta em uma busca por distribuição de lema.

Fonte: Floresta (Freitas et al., 2008).

Para que a análise fosse viável, não consideramos verbos com apenas uma ocorrência, e dispensamos da leitura caso a caso os 58 verbos já validados pela primeira etapa.

O resultado desta etapa foi uma lista de 293 verbos. O Quadro 8 apresenta a distribuição dos verbos segundo os padrões de busca. Os dados incluem repetições, uma vez que um mesmo verbo pode participar de diferentes padrões. A lista completa dos verbos encontrados pode ser conferida no Anexo 2.

PADRÃO	VERBOS INTRODUTORES DE DISCURSO RELATADO
Padrão 1	219
Padrão 2	107
Padrão 3	137
Padrão 4	156
Padrão 8	52
TOTAL	293 (sem repetições)

Quadro 8: Número de verbos encontrados com as expressões de busca em cada padrão.

Com base no Quadro 8, é possível tirar algumas conclusões. Em primeiro lugar, vemos que o Padrão (1) é o mais produtivo de todos, uma vez que a porcentagem de verbos encontrados nesse padrão em relação ao total de 285 é a maior de todas: 74,38%. O Padrão (8), por sua vez, foi o menos produtivo, apresentando apenas 18,24% dos verbos compilados.

Partindo das expressões de busca desenvolvidas e da lista de verbos que conseguimos reunir, passamos a implementação do léxico e regras nos corpora do projeto AC/DC. Na subseção a seguir, relataremos de forma breve como se tem desenrolado esse processo, que permitirá a qualquer pessoa interessada buscar por verbos dentro do campo semântico do dizer.

3.4

Anotação como pesquisa: a implementação dos verbos de elocução e o refinamento na noção de discurso relatado

Como mencionado no capítulo 1, a anotação de corpus também pode ser vista como uma auxiliar no processo de pesquisa. Nesta seção, relatamos brevemente o processo de aplicação de regras (os padrões) e léxico aos corpora do AC/DC.

De fato, esta seção poderia figurar em um capítulo à parte, já como uma aplicação dos resultados encontrados segundo a metodologia proposta. Assim, a anotação dos verbos de elocução no AC/DC não é exatamente uma etapa desta dissertação, mas um trabalho complementa, feito à parte, com colaborações externas, cujo objetivo é, também, enriquecer a anotação semântica do AC/DC. As regras utilizadas na anotação não foram exatamente as mesmas descritas aqui e o escopo da anotação ultrapassa os objetivos dessa dissertação. Tudo isso está descrito em detalhes em Freitas et al.(2016), e relatamos aqui apenas os pontos diretamente vinculados à dissertação.

Como dissemos, o processo de anotação, longe de ser apenas a aplicação das regras, trouxe novas nuances à delimitação ao que consideramos verbos de elocução e discurso relatado, com impacto no que foi dito até aqui. Por isso, decidimos apresentar esta etapa.

Quando demos início ao levantamento dos verbos nos corpora, percebemos que, antes de determinar o que são os verbos introdutores de discurso relatado, era

fundamental ter bem claro o que é relato. Sendo assim, adotamos, inicialmente, o seguinte critério: não é relato se não foi dito anteriormente. Com isso em mente, o caso (i), abaixo, foi considerado como discurso relatado, enquanto (ii), não:

- (i) Silva **disse** que não teve como evitar a batida.
- (ii) Logo mais, **dirão** que também o povo está contra nós!

Como, no exemplo (ii), o candidato a verbo introdutor de discurso relatado está empregado no futuro, e como todo o período dá a entender que a citação ainda não foi proferida por ninguém, estabelecemos que o verbo em tempos futuros não poderia desempenhar tal função. O mesmo valeria para os verbos no modo subjuntivo.

Também estabelecemos como critério que, se não houver como assegurar que se trata de um relato, desconsideramos a ocorrência. Assim, os casos (iii) e (iv) introduzem, para nós, um relato de discurso, porém nos (v) e (vi) não temos como saber se algo foi dito de fato (ou se a “fala original” era, na verdade, uma tabela de dados, por exemplo):

- (iii) «No Sambódromo não dá», **conclui** Liz.
- (iv) «Depois, o negócio é sair», **indicam** as garotas.
- (v) A perícia **concluiu** que ele foi atingido com golpes na cabeça por blocos de concreto, encontrados próximo ao corpo.
- (vi) Pesquisa do Datafolha **indicou** que a maioria dos eleitores brasileiros é favorável à existência do horário eleitoral gratuito.

A partir desses critérios, desenvolvemos as expressões de busca descritas na seção anterior. Isso quer dizer que elaboramos as expressões de modo a excluir os verbos que se encontrassem nos tempos e modos verbais que não consideramos introdutores de discurso relatado.

Quando começamos o trabalho de implementação dessas regras no corpus, entretanto, acabamos nos deparando com ocorrências que puseram em xeque os nossos critérios.

Percebemos que mesmo a decisão sobre algo já “ter sido dito anteriormente”, crucial na definição inicial, pode não ser óbvia, como ilustram as frases (vii)-(x) a seguir.

Assim, a implementação e verificação da anotação, ao nos apresentar novos exemplos, nos obrigou a refinar os critérios, sem perder de vista um dos objetivos que esta pesquisa pretende atender: auxiliar na tarefa de identificação de citação, no âmbito do Processamento de Linguagem Natural (PLN). Vale destacar que, considerando a dificuldade, ou mesmo a impossibilidade, de uma solução definitiva, qualquer outra solução adotada, desde que adequada à pergunta/motivação inicialmente proposta, seria igualmente válida.

Na tarefa de extração de citação, busca-se identificar, grosso modo, *quem disse o quê*. Pensando nesse propósito, é relevante que os conteúdos abaixo, por exemplo, sejam identificados:

- (vii) Você **vai dizer** que eu ando devagar quase parando, uma vez que a Copa acabou em julho.
- (viii) Alvares não **disse** que era contra a divulgação do conteúdo da pasta.
- (ix) Mas cada um desses torcedores **diria** que o seu time fora o melhor.
- (x) Retiro apenas a pena que senti do Viana e **digo** que fechamos o ano com a esperança de que, afinal, a caixa fure.

Em outras palavras, a prática da anotação, antes de ser uma etapa puramente braçal, foi crucial para delimitar fronteiras do fenômeno estudado. Desconhecemos, nas pesquisas feitas ao longo desta dissertação, a problematização desses aspectos do discurso relatado, seja na literatura dos Estudos da Linguagem, seja na literatura do PLN e de extração de citações.

Com critérios mais bem delimitados, e considerando o interesse ampliado nos verbos *dicendi* (que não necessariamente são idênticos aos verbos de elocução, como explicitado em Freitas et al., 2016), criamos as seguintes classes de anotação, já implementadas nos corpora do AC/DC:

- dizer (são os verbos de comunicação listados no apêndice XXX mas que não estão em contexto de elocução, e portanto não interessam a essa dissertação);
- dizer_relatoDIRETO;
- dizer_relatoINDIRETO;
- dizer_relatoMISTO.

Em paralelo ao levantamento dos verbos, padrões e à anotação, foi feita a classificação dos verbos encontrados em grupos de sentido, tendo em vista a organização do glossário DISSE. A seção a seguir explica como foi conduzida essa organização e quais foram as dificuldades encontradas.

4

Criação do DISSE: organização dos verbos em grupos de sentido

4.1 A primeira classificação

Como já dissemos, todo o procedimento de levantamento dos verbos de elocução nos corpora, descritos nos capítulos anteriores, culminou com uma lista de 285 verbos. Tendo em vista o desenvolvimento do glossário DISSE, precisávamos, em primeiro lugar, decidir que tipo de organização faria mais sentido para o propósito do glossário. Na internet e até mesmo em alguns manuais de redação, como o manual do Othon M. Garcia, é possível encontrar diversas listas de verbos *dicendi*, porém sem qualquer organização a não ser em ordem alfabética. A ordem alfabética, no entanto, só é útil quando a pessoa sabe o que está procurando. No caso de um tradutor, ao consultar a lista, muito provavelmente ele tem em mente uma noção aproximada de sentido. Sendo assim, pensamos que a organização da lista em grupos de sentido seria a mais proveitosa para os tradutores.

Criamos, então, 21 grupos de sentido (21 classes) para enquadrar os 58 verbos obtidos na busca inicial no COMPARA. A seguir, temos o Quadro 9, com as classes criadas e os verbos neles classificados:

CLASSE	VERBOS						
acordo	anuir	concordar					
acréscimo	acrescentar	continuar	prosseguir				
afirmação	afirmar	comentar	contar	declarar	dizer	falar	
censura	censurar						
confirmação	assegurar	confirmar	repetir				
confissão	admitir	confessar	declarar	desabafar			
contestação	argumentar	contrapor	insistir	queixar-se	replicar	retorquir	retrucar
destaque	insistir	observar	sublinhar				
exclamação	berrar	exclamar					
explicação	argumentar	concluir	esclarecer	explicar	justificar-se	observar	
hesitação	balbuciar						
informação	anunciar	comunicar	contar	declarar	informar		
interrupção	interromper						
lembança	lembrar						
opinião	concluir	insistir					
pedido	pedir						

pergunta	indagar	perguntar	querer saber	sondar			
resposta	replicar	responder	retorquir	retrucar	ripostar		
sugestão	advertir	propor	sugerir				
surpresa	admirar-se	espantar-se					
outros	agradecer	brindar	cumprimentar	escrever	prometer	saudar	

Quadro 9: Classificação inicial dos 58 verbos de dizer obtidos no COMPARA, distribuídos em 21 grupos de sentido.

Fonte: Dados da pesquisa.

Uma vez classificados os verbos, desenvolvemos um teste com o intuito de validar as classes criadas. Isto é, a ideia era verificar se as pessoas concordariam com a classificação proposta. O teste consistia em 58 linhas de concordância, cada uma com um dos 58 verbos, e o voluntário deveria classificar o verbo em destaque conforme nossas classes, ou, ainda, sugerir novas classificações, caso acreditasse que nenhuma das já existentes desse conta do verbo em questão. A Figura 14 ilustra um trecho do teste.

TRADUÇÃO	CATEGORIA(S)	SUGESTÕES
– Obrigada, mas acho que não – respondeu com dificuldade.		
– E se eu não puder? – perguntei .		
– Eu não sei o que quero – confessou Alistair com ar de desânimo. – Só não quero que tudo isto continue.		

Figura 14: Estrutura da primeira versão do teste.

O teste, que foi distribuído entre dez pessoas, todas com formação de nível superior em Letras, apresentou resultados que questionaram a classificação proposta. O Anexo 3 apresenta as respostas dadas por todos os participantes.

Poucas classes foram validadas. Estabelecemos como critério para a validação da classe a concordância de 80% dos voluntários. Em outras palavras, se oito ou mais indivíduos atribuísem a mesma classe a um determinado verbo, essa classe ficaria, portanto, validada (mesmo que não tivesse sido sugerida por mim). Dos 58 verbos, apenas 18 tiveram suas classes confirmadas pelos participantes do teste, como é possível conferir a seguir:

VERBO	CLASSE	CONCORDÂNCIA
pedir	PEDIDO	90%

sublinhar	DESTAQUE	90%
acrescentar	ACRÉSCIMO	90%
anuir	ACORDO	80%
replicar	CONTESTAÇÃO	80%
censurar	CENSURA	90%
perguntar	PERGUNTA	80%
confessar	CONFISSAO	90%
propor	SUGESTÃO	90%
querer saber	PERGUNTA	90%
balbuciar	HESITAÇÃO	80%
confirmar	CONFIRMAÇÃO	80%
sugerir	SUGESTÃO	100%
admirar-se	SURPRESA	80%
exclamar	EXCLAMAÇÃO	90%
declarar	AFIRMAÇÃO	80%
indagar	PERGUNTA	80%
lembrar	LEMBRANÇA	80%

Quadro 10: Lista de verbos e os grupos de sentido validados pelo primeiro teste.

Fonte: Dados da pesquisa.

Apenas um verbo teve sua classe validada por todos os participantes, o verbo *sugerir*. E, mesmo nesse caso, as pessoas não deixaram de atribuir, simultaneamente, outras classes preestabelecidas, ou, ainda, de sugerir novas classificações. Assim, foram contabilizadas, no total, sem repetições, 304 classes (havia, originalmente, 51), o que corresponde a uma média de 5,24 classes distintas para o mesmo verbo.

Apesar de termos antecipado a possibilidade de os verbos se enquadrarem em mais de um grupo, ainda assim, consideramos o valor médio de grupos por verbo um tanto alto, o que poderia dificultar a consulta do tradutor ao futuro glossário. Se tivéssemos muitos verbos repetidos em grupos de sentido diferentes, a classificação poderia perder o seu propósito, já que teríamos várias listas longas com os mesmos verbos. Além do número relativamente baixo de validações e da média alta de grupos por verbo, ressaltamos também o fato de que nenhum verbo recebeu apenas um grupo de sentido.

Com os resultados desse primeiro teste, ficou evidente que uma classificação granular demais (21 classes para 58 verbos) suscita mais discordância e não facilita a tarefa do tradutor, pelo contrário: ele fica “obrigado” a decidir entre diferentes nuances de sentido, sendo levado a dúvidas que normalmente não existiriam. Quando se deparasse com o verbo *said*, o tradutor seria obrigado a fazer inferências e a enquadrar o verbo em questão em uma classificação muito refinada, sendo que ele não necessariamente precisa desse nível de refinamento para encontrar uma tradução que se encaixe no contexto em questão. Ficou claro para nós que uma reformulação da classificação deveria, necessariamente, envolver grupos mais abrangentes e amplos.

Retomamos aqui a posição de Ellis (1993), já apresentada no capítulo 1. A ideia de uma classificação é justamente simplificar o que é complexo, e, para isso, precisamos abrir mão de certas características distintivas, isto é, tornar igual o que não é igual. O processo de categorização não é natural (no sentido de que é uma criação humana para melhor entender o mundo) e, nesse processo, a comunidade interpretativa que a propõe encontra-se envolvida em uma sucessão de escolhas, feitas para que a classificação lhe sirva para uso próprio. Segundo Ellis, as categorias linguísticas são, em primeiro lugar, reflexo dos propósitos coletivos dos falantes de uma língua e não reflexos diretos da estrutura do mundo (Ellis, 1993, p. 34).

Desse modo, a classificação que buscamos nesta pesquisa é uma classificação *consensual*; uma vez que não há como escapar da minha interferência na classificação, procuramos um equilíbrio entre o que temos agora (todos os verbos agrupados em um grande grupo dos “verbos de elocução”) e os grupos de sentido

muito específicos, propostos no primeiro teste que realizamos. Propondo classes menos específicas, a chance de consenso deve ser maior.

Decidimos, então, reduzir o número de grupos de sentido. Vale ressaltar que o que almejamos aqui não é uma classificação *correta*, tampouco *definitiva*. Procuramos apenas alcançar um espaço em que haja concordância, em um nível que torne as categorias úteis para os profissionais da tradução. Ellis defende que as classes mais úteis não são as mais específicas, como muitos podem pensar, mas sim as mais amorfas, justamente porque é a amorfia que abre mais espaço para o consenso. No entanto, ainda que tenhamos reduzido o número de classes, estas permaneceram minimamente informativas, para que não se perca o propósito e a utilidade da classificação.

No teste, pedimos que os voluntários classificassem os verbos com base em apenas um exemplo. Mesmo assim, os resultados obtidos foram capazes de refletir a possibilidade de flutuação entre as classes, uma vez que nenhum verbo recebeu uma única classificação; *indagar* e *lembrar*, verbos aos quais foram atribuídos o menor número de classes, foram enquadrados em duas classes cada. Por outro lado, os verbos *agradecer*, *prosseguir* e *queixar-se* encaixaram-se, segundo os participantes, em nove classes (cada um). O Quadro 11 mostra o exemplo a ser classificado e os grupos atribuídos a cada um desses cinco verbos citados:

OCORRÊNCIA	CLASSES ATRIBUÍDAS
– Porquê? – indagou Mr. Walsh.	CONTESTAÇÃO, PERGUNTA
– Até os ovos – lembrou alguém.	ACRESCIMO, LEMBRANÇA
– Foi muito simpático da sua parte – agradeceu Bernard.	AFIRMAÇÃO, AGRADECIMENTO, CONFIRMAÇÃO, CORTEJO, DESTAQUE, EXCLAMAÇÃO, GRATIDÃO, OPINIÃO, RESPOSTA
– No entanto gosto do azul – prosseguiu , a cliente – e o verde é muito bonito.	ACORDO, ACRÉSCIMO, AFIRMAÇÃO, CONTRAPOSIÇÃO, INFORMAÇÃO, INTERRUPÇÃO, NEUTRALIDADE, OPINIÃO, SUGESTÃO
– De qualquer modo, é muito frio aquilo em Sea Cliff, em toda aquela ponta do Noroeste de S. Francisco – queixou-se .	CENSURA, CONCLUSÃO, CONFISSÃO, EXCLAMAÇÃO, LAMENTO, OPINIÃO, QUEIXA, RECLAMAÇÃO, RESPOSTA

Quadro 11: Lista de classes atribuídas aos verbos que menos e mais receberam classificações diferentes.

Mantivemos nossa opção pela distribuição em grupos de sentido, mesmo cientes das dificuldades (ou da impossibilidade) de classificação de tais verbos (Austin, 1962), porque este tipo de classificação faz sentido *no contexto do glossário pretendido* (em oposição a uma tentativa de classificação geral desses verbos, por exemplo). Além disso, a ênfase em uma classificação motivada pelo sentido deve-se à sua utilidade para os usuários do glossário – por isso não utilizamos uma classificação de viés sintático, por exemplo, como feita por Moura Neves (2000). No momento da tradução, informações sintáticas a respeito de um verbo, por exemplo, não são relevantes para o processo decisório do tradutor.

Antes de propor uma nova classificação, voltamos nosso olhar para o que já foi estudado a respeito dos atos de fala, no âmbito da Pragmática. Não é a teoria em si que nos interessa aqui, mas sim a tentativa de Austin de classificar os verbos que tornam explícita a força ilocucionária de uma declaração¹³.

Sobre esses verbos, Austin (1962) afirmou:

Using then the simple test (with caution) of the first person singular present indicative active form, and going through the dictionary (a concise one should do) in a liberal spirit, we get a list of verbs of the order of the third power of 10. I said I would attempt some general preliminary classification [...] I distinguish five more general classes: but I am far from equally happy about all of them. [...] I call them classes of utterance, classified according to their illocutionary force. (Austin, 1962, p. 149-150).

Vale mencionar aqui as classes propostas e os verbos que se incluem em cada classe, a fim de, em seguida, analisar de forma breve a conclusão à qual Austin chegou a respeito de sua classificação. O Quadro 12 apresenta as cinco classes de proferimentos de Austin, juntamente com alguns verbos nelas enquadrados.

CLASSES PROPOSTAS E DEFINIÇÕES	ALGUNS VERBOS
(1) Vereditivos: exercício de julgamento.	reckon, assess, characterize, describe, analyze
(2) Exercitivos: afirmação de influência ou de exercício de poder.	appoint, order, command, recommend, veto, pray

¹³ Segundo Austin (1962), o ato ilocucionário é aquele que é executado na fala. Em outras palavras, no momento do proferimento, um ato sendo realizado simultaneamente. Por exemplo, quando dizemos “prometo que não vou embora”, o ato de prometer é executado ao mesmo tempo em que a fala é proferida.

(3) Comissivos: compromisso em assumir uma obrigação ou de declarar uma intenção.	promise, plan, intend, guarantee, vow, consent
(4) Comportamentais: adoção de uma atitude.	apologize, thank, welcome, congratulate, dare, criticize, approve
(5) Expositivos: esclarecimento de razões, argumentos e opiniões.	affirm, deny, report, remark, accept, inform, tell, ask, agree

Quadro 12: Classes propostas por J. L. Austin (1962) para os verbos que tornam explícita a força ilocucionária.

Fonte: Austin (1962).

Antes mesmo de propor a classificação acima, Austin se mostra insatisfeito com o resultado obtido. Mais adiante, tecendo alguns comentários a respeito das classes por ele estabelecidas, Austin comenta:

The last two classes [behabitives and expositives] are those which I find most troublesome, and it could well be that they are not clear or are cross-classified, or even that some fresh classification altogether is needed. I am not putting any of this forward as in the very least definitive. Behabitives are troublesome because they seem too miscellaneous altogether: and expositives because they are enormously numerous and important, and seem both to be included in the other classes and at the same time to be unique in a way that I have not succeeded in making clear even to myself. It could well be said that all aspects are present in all my classes. (Ibidem, p. 151)

Assim, a tentativa, segundo Austin, fracassada de classificar esses verbos nos ajudou a compreender um pouco melhor as complicações envolvidas na classificação que propusemos. Entretanto, mantivemos nossa opção pela distribuição em grupos de sentido, mesmo cientes das dificuldades (ou da impossibilidade) de classificação de tais verbos (Austin, 1962), pois temos objetivos diferentes de Austin. Não temos a pretensão de uma sistematização absoluta da língua, não acreditamos ser essa uma tarefa possível. Temos, na verdade, um objetivo prático: organizar 285 verbos de modo a criar um glossário voltado para tradutores.

4.2 Nova tentativa de classificação

Passamos, então, à nova proposta de classificação dos verbos. Dessa vez, incluímos todos os verbos obtidos com as nossas pesquisas em corpus até então, que, somados aos que encontramos no COMPARA, chegaram ao impressionante

número de 293 verbos empregados como verbos de elocução. Com base nessa nova lista, criamos 11 classes, levando em consideração os resultados do experimento. Para chegar a essas onze classes, tivemos três critérios: (i) partimos dos resultados do primeiro teste, que acabaram por indicar alguns caminhos a respeito dos nomes dos grupos; (ii) mesclamos alguns dos grupos que eram mais próximos em termos de sentido, para que se reduzisse o número total de grupos; e (iii) levamos em conta a utilidade de cada grupo para o tradutor durante a sua tarefa laboral.

Analizamos as ocorrências dos verbos em contexto e os distribuímos nas novas classes propostas. A seguir, encontra-se um gráfico que explicita os 11 novos grupos de sentido e a distribuição dos 293 verbos entre esses grupos.

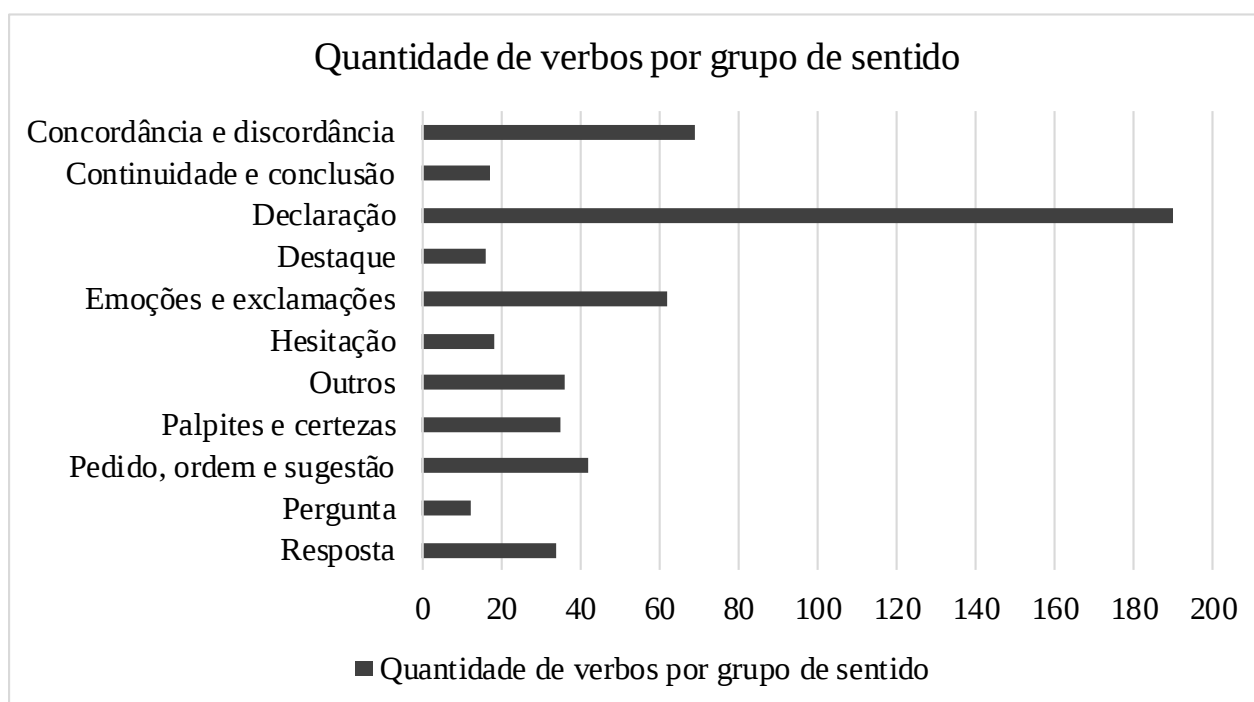


Figura 15: Gráfico da distribuição de verbos por classe.

Fonte: Dados da pesquisa.

O grupo *declaração* foi proposto por dois voluntários do teste para verbos que tinham sido enquadrados no grupo *afirmação*: os verbos *declarar* e *falar*. Consideramos que o grupo *declaração* era mais adequado do que *afirmação*, pois o primeiro poderia abrigar um número maior de verbos (incluindo-se aqui, por exemplo, os verbos de negação). Os grupos *informação* e *opinião* também foram incorporados a esse novo grupo. Dessa forma, *declaração* constitui o grupo mais amplo, com 189 verbos dos 293 (64,7%). Muitos desses verbos foram também colocados em mais de um grupo de sentido.

Os grupos *acordo* e *contestação* uniram-se para compor o grupo *concordância e discordância*. Entendemos que essa classe reúne verbos que indicam qualquer tipo de posicionamento em relação ao que o outro fala. A opção pela troca das palavras que denominam a classe se deveu à necessidade de maior clareza; com base nos resultados do experimento, chegamos à conclusão de que os nomes *acordo* e *contestação* nem sempre ficaram claros para os participantes. Além disso, foi incorporada a essa classe a *censura*, pois consideramos *contestação* e *censura* campos semânticos afins.

Pedido e *sugestão*, grupos já existentes, juntaram-se em um novo grupo, que conta também com os verbos com ideia de *ordem*. A classe *acréscimo* recebeu um novo nome e teve seu escopo ampliado: *continuidade e conclusão*. Também ampliamos a abrangência da classe *exclamação*, nela incluindo também os verbos que, de alguma forma, denotam emoção, sem necessariamente haver algum tipo de exclamação; temos, então, a classe *emoções e exclamações*. Por último, aumentamos o escopo da classe *hesitação*, que agora também abriga o que chamamos de *ponderação*.

A única classe realmente nova nessa segunda classificação é *palpites e certezas*. O único verbo que se enquadraria nesse grupo, *prometer*, estava antes classificado como *outros*. No entanto, considerando a nova lista (com muito mais verbos), essa classe foi ampliada, passando a incluir 35 verbos.

Apenas quatro grupos de sentido foram mantidos: *pergunta*, *resposta*, *destaque* e *outros*. Ainda que, comparativamente, os grupos *pergunta* e *resposta* englobem poucos verbos, decidimos por mantê-los pelo simples motivo de que o tradutor, quando busca uma tradução possível para *said*, costuma pensar, em primeiro lugar, se o contexto constitui uma declaração, uma pergunta ou uma resposta. Em geral, há pistas formais que facilitam esse pensamento, sendo a principal delas a pontuação. Ao se deparar com um ponto de interrogação, por exemplo, o tradutor possivelmente pensará em substituir o *said* por algum verbo que indique pergunta.

Já o grupo *destaque* foi mantido, sem alteração de nome, pois, na primeira classificação, figurava entre os grupos com o maior número de verbos. Agora, trata-se do segundo menor grupo, com 16 verbos, ficando atrás apenas de *pergunta*. Por fim, a manutenção do grupo *outros* se deve à necessidade de um grupo-coringa, que

abrigue aqueles verbos que não se enquadram bem nos demais grupos, mas que, ao mesmo tempo, não justificam a criação de um grupo só para si.

Implementamos a segunda versão do teste no Rêve (Santos et al, 2015), uma ferramenta on-line criada no âmbito da Gramateca justamente para auxiliar atividades de pesquisa linguística com base em grandes corpora, e que tomam a anotação como processo de investigação. De fato, o processo de classificação de elementos linguísticos em contexto (que corresponde ao teste criado para validar as classes propostas) pode ser visto como um processo de anotação, ainda que não haja implementação. O uso da ferramenta trouxe facilidades tanto para a realização do teste por parte dos voluntários quanto para a análise dos resultados, uma vez que o Rêve produz um relatório de estatísticas das respostas ao teste (tanto o teste quanto o relatório dos resultados estão disponíveis on-line, no site do Rêve¹⁴).

Para esse teste, selecionamos 30 verbos apenas, pois julgamos que a versão anterior tinha ficado muito longa. Os critérios de seleção desses 30 verbos foram os seguintes: (1) quinze verbos correspondem a verbos do teste anterior, mais especificamente aos verbos que apresentaram mais discordância na classificação (pois receberam um número maior de grupos) e (2) os outros quinze foram retirados aleatoriamente dos onze grupos de sentido, porém nos certificamos de que havia pelo menos um verbo de cada grupo. O teste foi feito por 17 participantes, todos com formação superior em Letras.

A Figura 16 mostra a interface do teste aplicado e apresenta, também, a definição que nós demos a cada um dos grupos de sentido:

¹⁴ <http://www.linguateca.pt/Reve/>

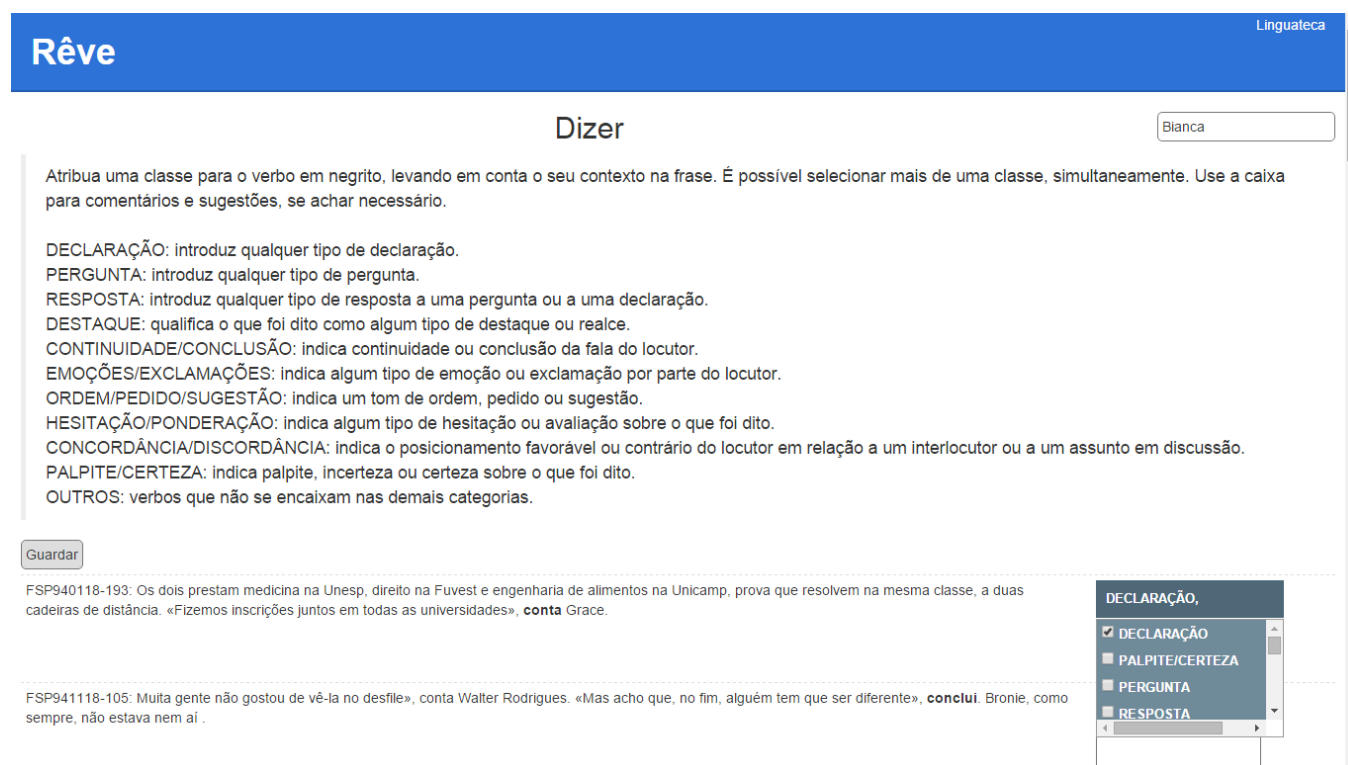


Figura 16: Interface da segunda versão do teste da classificação em grupos de sentido.

Apesar das reformulações feitas, os resultados do teste não foram muito diferentes da primeira versão, corroborando as impressões de Austin. Dos 30 verbos testados, em apenas 19 (cerca de 63%) conseguimos uma concordância igual ou superior a 70% por parte dos voluntários (decidimos reduzir o número de 80% para 70% por conta do aumento do número de voluntários). O Quadro 13 mostra os verbos cujas classes foram validados:

VERBO	CLASSES
contar	DECLARAÇÃO (94,1%)
concluir	CONTINUIDADE/CONCLUSÃO (82,4%)
queixar-se	EMOÇÕES/EXCLAMAÇÕES (70,6%)
prometer	DECLARAÇÃO (70,6%)
contrapor	CONCORDÂNCIA/DISCORDÂNCIA (82,4%)
prosseguir	CONTINUIDADE/CONCLUSÃO (100%)

espantar-se	EMOÇÕES/EXCLAMAÇÕES (100%)
anunciar	DECLARAÇÃO (100%)
esclarecer	DECLARAÇÃO (70,6%)
arrematar	CONTINUIDADE/CONCLUSÃO (76,5%)
continuar	CONTINUIDADE/CONCLUSÃO (100%)
convidar	ORDEM/PEDIDO/SUGESTÃO (70,6%)
censurar	CONCORDANCIA/DISCORDANCIA (82,4%)
frisar	DESTAQUE (94,1%)
emendar	CONTINUIDADE/CONCLUSÃO (100%)
festejar	EMOÇÕES/EXCLAMAÇÕES (94,1%)
apoiar	CONCORDANCIA/DISCORDANCIA (94,1%)
xingar	EMOÇÕES/EXCLAMAÇÕES (88,2%)
sondar	PERGUNTAR (82,4%)

Quadro 13: Lista dos verbos da segunda versão do teste, com as respectivas classes atribuídas.

Fonte: Dados da pesquisa.

Por outro lado, conseguimos reduzir a média de classes atribuídas por verbo, que ficou em 2,7.. Os verbos *emendar* e *anunciar* receberam apenas uma classe, enquanto o verbo *advertir* teve o maior número de classes atribuídas: sete. O Quadro 14 mostra algumas das frases testadas e as classes atribuídas pelos participantes:

LINHAS DE CONCORDANCIA	CLASSES ATRIBUÍDAS
Confessou a juíza ao condenar José à pena mais mínima que podia, 12 contos mais despesas de tribunal: «Às vezes a necessidade leva-nos a cometer coisas que são muito pouco aconselháveis», e depois emendou que «com certeza a sua	CONTINUIDADE/CONCLUSÃO

filha prefere ter um pai que seja um bom exemplo em vez de ter um pai preso por crimes» e remendou que «inclusivamente pode estar a criar sentimentos de culpa à sua filha -- 'o meu pai cometeu um crime por minha causa...'	
O anúncio foi feito ontem pelo futuro presidente do Santos, Miguel Kodja Neto, e pelo prefeito David Capistrano (PT). Pelé esteve presente e anunciou que até sexta-feira deve divulgar a empresa que irá patrocinar o clube.	DECLARAÇÃO
«O resultado de tanta incapacidade, hesitação e má vontade será a ONU desacreditada, a Otan arruinada e os europeus desmoralizados. " " As relações entre Ocidente e o mundo muçulmano nunca mais serão as mesmas», advertiu o presidente. Disse ainda que a Rússia «está do lado dos agressores».	DECLARAÇÃO EMOÇÕES/EXCLAMAÇÕES PALPITE/CERTEZA HESITAÇÃO/PONDERAÇÃO ORDEM/PEDIDO/SUGESTÃO DESTAQUE OUTROS

Quadro 14: Linhas de concordância avaliadas na segunda versão do teste e que receberam o menor ou o maior número de classes.

No entanto, é necessário ressaltar que decidimos desconsiderar classificações atribuídas por apenas 1 ou 2 pessoas – consideramos tais casos desvios¹⁵. Ainda que tenhamos escolhido apenas participantes formados em Letras, é importante cogitar a possibilidade de que esses participantes, embora tenham feito o teste voluntariamente, não tenham o mesmo “comprometimento com a tarefa” que nós temos e, por isso, as análises feitas por eles podem estar sujeitas a “erros” (Santos et al., 2015, p. 15).

Uma das dificuldades do nosso teste, percebida na primeira versão e confirmada na segunda, deve ser atribuída à própria natureza da tarefa. Ao darmos início à nossa pesquisa, talvez tenhamos subestimado a dificuldade de classificação dos verbos em grupos de sentido. No caso do teste, especificamente, percebemos

¹⁵ Poderíamos ter usado testes estatísticos para uma maior precisão das concordâncias, discordâncias e desvios, mas isso não foi possível.

que foi muito difícil para os participantes concentrar a classificação *apenas no verbo*. Em outras palavras, pensávamos, a princípio, que o verbo que introduzia o relato é que dava a carga semântica ao que era dito. O teste nos mostrou, no entanto, que as pessoas se deixam levar mais pelo que é dito do que pelo verbo empregado; a maioria das discordâncias do teste é fruto dessa influência do conteúdo relato, como, por exemplo, no item abaixo, classificado por um participante como “destaque”:

*O ala Dominique Wilkins, 34, deu uma entrevista coletiva em que se queixou dos dirigentes de sua ex-equipe, o Atlanta Hawks. «Eles não precisavam mostrar tanto desrespeito», **queixou-se** Wilkins, 34, trocado por Danny Manning, do Los Angeles Clippers.*

Ainda assim, consideramos os resultados satisfatórios, e organizamos o glossário DISSE, encontrado na íntegra no Anexo 4. O DISSE corresponde à união dos verbos encontrados com os padrões, dos verbos encontrados no COMPARA e dos verbos encontrados apenas em outros estudos (por exemplo, os verbos encontrados por Moura Neves, mas que não estavam em nossa lista inicial, como mostrarei no capítulo 5), mas que foram confirmados pelos corpora do projeto AC/DC.

Mantivemos os onze grupos de sentido que propusemos, sem acréscimos das sugestões dos participantes do teste. Embora o verbo *cumprimentar* tenha recebido quatro sugestões de “saudação” como grupo de sentido, optamos por desconsiderá-la, visto que testamos esse grupo no primeiro teste e os resultados não foram muito produtivos. Esse grupo de sentido é muito restrito, o que foge à nossa segunda proposta de grupos mais amplos e que deem conta de mais verbos.

Além disso, dentro dos grupos, listamos os verbos por ordem de frequência e não em ordem alfabética. Mais uma vez, levamos em conta que o fato de que a ordem alfabética é mais útil para quem sabe o que procura. O tradutor, ao buscar um verbo para usar em sua tradução, terá, além do campo semântico, um critério relacionado à frequência de uso: ou estará em busca de um verbo mais comumente usado, ou então de um verbo mais heterodoxo. No primeiro caso, o tradutor poderá direcionar o olhar para o início da lista, enquanto para o segundo caso, o fim da lista lhe dará opções mais interessantes.

Apesar das dificuldades encontradas, seguimos com o nosso propósito de reunir um glossário de verbos introdutores de discurso relatado. Ainda enxergamos a classificação em grupos de sentido como a melhor forma de organizar o léxico que reunimos. Acreditamos que fizemos uma primeira tentativa satisfatória e deixamos o caminho aberto para quem quiser aprimorá-lo de alguma forma, seja na forma de organização do léxico, seja na forma de apresentação do glossário.

5

Verbos de elocução e discurso relatado: uma descrição com base em corpus

Apresentamos aqui os resultados que obtivemos com a identificação dos padrões, comparando-os, sempre que possível, com os trabalhos já mencionados sobre o *dizer*. O Quadro 15 mostra a quantidade de verbos de elocução encontrados ou listados por trabalhos citados nesta pesquisa:

PUBLICAÇÃO	IDIOMA	CLASSE DE VERBOS ANALISADA	QUANTIDADE DE VERBOS
Sagot et al (2010)	francês	verbos de elocução	232
Biber et al (1999)	inglês	verbos de comunicação	66
Levin (1993)	inglês	verbos de comunicação	163
<i>Manual de Redação e Estilo do Estado</i> (2016)	português	verbos de elocução	56
Garcia (2010)	português	verbos de elocução	77
Moura Neves (2000)	português	verbos de elocução	103
Este trabalho	português	verbos de elocução	293

Quadro 15: Quadro comparativo dos verbos de elocução levantados em diversos estudos.

O Quadro 15 evidencia o sucesso da abordagem que adotamos nesta pesquisa. No entanto, os números, por si só, podem dar a entender que todos os autores tratam dos mesmos fenômenos, com os mesmos interesses, o que não é o caso. A terceira coluna do quadro mostra qual foi a classe de verbos abordada por cada um dos trabalhos, de acordo com a nossa classificação.

Sagot et al (2010) tem interesse na extração automática de citação, e seu interesse nos verbos refere-se exclusivamente a resolver a tarefa pretendida; Biber et al (1999) e Levin (1993) são trabalhos descritivos, que têm o inglês como língua de interesse e consideram a classe mais ampla dos *verbos de comunicação*; Garcia (2010) e o *Manual de Redação e Estilo do Estado* (2016) são manuais de redação; e Moura Neves (2000) refere-se a uma gramática de usos do português.

Nas seções a seguir, detalhamos as comparações e diálogos com os referidos trabalhos.

5.1

Moura Neves: verbos de elocução e o tipo de discurso

Moura Neves, tal como já afirmamos anteriormente, compilou 103 verbos de elocução, em contraste com os 293 verbos desta pesquisa. Entre os 103 verbos da *Gramática de usos do português* (Moura Neves, 2000), apenas 25 não estavam presentes na lista de verbos que reunimos. Os 25 verbos são: acalmar, agastar-se, aguilhoar, antecipar, boquejar, bronquear, bulir, caçoar, cochichar, conchavar, consolar, debicar, debochar, desiludir, diagnosticar, escarnecer, ferroar, inclinarse, interceptar, maldizer, participar, remediar, suspirar, sussurrar, zombar. Os verbos que estão sublinhados pertencem todos a uma mesma categoria de Moura Neves, *verbos que instrumentalizam ou circunstanciam o que é dito*, que são os verbos que podem introduzir discurso, mas que não necessariamente indicam atos de fala. Diante desse número, e também porque, apesar do grande número de verbos encontrados, sabemos que a nossa lista não é exaustiva, resolvemos, então, testar essa lista de Moura Neves (2000) nos três corpora monolíngues usados neste estudo. Para fazer essa pesquisa, usamos todas as expressões de busca desenvolvidas para os seis padrões.

Primeiramente, queríamos entender por que esses 25 verbos não apareceram entre os 293 verbos que compilamos. Consultamos então todo o material do AC/DC e identificamos três motivos diferentes para a ausência: ou (i) tinham apenas uma ocorrência (e, portanto, foram descartados na nossa seleção inicial), ou (ii) os verbos não foram encontrados entre as ocorrências, ou (iii) as ocorrências não correspondiam a verbos de elocução (exemplo: “Com o vídeo “Circular”, *participou* este ano do prestigiado “Curta Cinema”, Festival Internacional de Curtas do Rio.”).

Em seguida, precisávamos confirmar se tais verbos poderiam funcionar como verbos de elocução (e, em caso afirmativo, seriam incluídos em nossa lista).

Dos 25 verbos, 15 foram empregados como verbos de elocução em todos¹⁶ os corpora:

VERBO	LINHA DE CONCORDÂNCIA
acalmar	«E normal, cada criança tem um ritmo próprio de desenvolvimento», acalmavam os vários pediatras -- no Porto e em Lisboa, frisa a mãe -- que o observaram.
antecipar	«Pode haver muita contestação, inclusive judicial, o que só prejudicará o partido », antecipa o ex-governador.
boquejar	Ele foi a ordem do dia de muitos dias; apontaram-no a dedo, boquejaram , por portas travessas, que ia sair um jornalzinho, intitulado O Bode ” só para botar os podres do ordinário na rua!
bronquear	Ele vai pagar por isso», bronqueou o dirigente.
caçoar	Vamos ver se jogando tanto a turma aprende», caçoou .
cochichar	«Não consigo falar », cochichei .
consolar	O psicólogo a consola : « Não se atormente, minha senhora. »
debochar	Enquanto Os Incríveis cantavam «as praias do Brasil ensolaradas, lá, lá, lá», uma paródia debochava : « maconha no Brasil foi liberada... »
desiludir	«Quanto a doenças, acabou-se!", desiludiu Ieltsin a oposição.
diagnosticar	«Faltam metas, objetivos, quantificação ao programa», diagnosticou .
escarnecer	«Actualmente a nossa tertúlia é o aparelho de televisão, o prato do bife e nós próprios », escarneceu Silva Pereira, um dos responsáveis pela Tertúlia Momentos de Reflexão e Fraternidade, assembleia que se junta uma vez por mês na Padaria para, após um «jantar frugal», debater os temas mais diversos e mesmo controversos.
participar	Com alguma satisfação, no final, o Ministério do Interior participava que a «infraestrutura da FAP foi eliminada através da apreensão de telefones móveis, faxes,

¹⁶ Lembramos que a utilização de todos os corpora corresponde a uma busca em mais de 1 bilhão de palavras, distribuídas em diferentes gêneros textuais.

	listas de membros e de dinheiro, correio, disquetes de computador e correspondência».
suspirar	«Graças a Deus», suspirou .
sussurrar	«Levantei-a nos meus braços», sussurrou Carol Ann Timmel aos jornalistas.
zombar	«Caymmi entra com a música e o Carlinhos, com o uísque», zombavam .

Quadro 16: Verbos apresentados por Moura Neves (2000) como verbos de elocução e que não foram encontrados por esta pesquisa em um primeiro momento, mas foram confirmados pelos corpora do projeto AC/DC.

O fato de que não encontramos 10 dos verbos propostos em Moura Neves (2000) não significa necessariamente que tais verbos não podem ser empregados como introdutores de discurso relatado. Devemos lembrar que, apesar da dimensão dos corpora consultados, um corpus jamais será exaustivo e tampouco dará conta de todos os fenômenos da língua. Como o corpus usado por Moura Neves não está disponível para o público, não temos como verificar se as ocorrências analisadas na elaboração da gramática correspondem, para nós, a verbos de elocução. Ainda assim, vale ressaltar que a ausência desses 10 verbos em corpora que, juntos, somam mais de um bilhão de palavras, pode indicar que o emprego dos verbos como introdutores de discurso relatado seja incomum.

Além disso, traçamos também um paralelo entre o estudo de Moura Neves e esta pesquisa no que se refere ao tipo de discurso. No Quadro 2, que se encontra no capítulo 2, é possível ver que Moura Neves determinou se os verbos de elocução são aceitos no discurso direto (DD) e/ou no indireto (di). No Quadro 17, reproduzimos a classificação de Moura Neves, considerando (a) os verbos em que houve divergência quanto ao tipo de discurso associado ao verbo, ou seja, quando nossos resultados e os de Moura Neves diferem; e (b) os verbos encontrados apenas por nós – contribuindo, assim, com o levantamento já feito pela autora:

VERBO	DD	DI
acentuar	x	x
acreditar	x	x
acrescentar	x	x

acrescer		x
acusar	x	x
adiantar	x	x
admirar-se	x	

admitir	x	x
advertir	x	x
advogar	x	x
aferir	x	
afiançar	x	x
agradecer	x	x
ajudar	x	
alardear	x	x
alegrar-se	x	
alertar	x	x
alfinetar	x	
aludir	x	x
amenizar	x	
analisar	x	x
animar-se	x	
anotar	x	x
antever	x	x
anuir	x	x
apelar	x	x
apojar	x	
apontar	x	x
apostar	x	x
apregoar	x	x
apresentar-se	x	
apurar	x	x
arriscar	x	x
asseverar	x	x
assinalar	x	x
assumir	x	x
atacar	x	
atestar	x	x
atirar	x	
avaliar	x	x
bradar	x	x
brincar	x	x
brindar	x	
calcular	x	x
cantar	x	x
censurar	x	

chamar	x	x
chutar	x	x
citar	x	x
clamar	x	x
clarificar	x	x
colocar	x	x
combinar	x	x
começar	x	
comemorar	x	x
comparar	x	
complementar	x	
completar	x	x
conceder	x	x
considerar	x	x
constatar	x	x
contabilizar	x	
contestar	x	x
continuar	x	x
contrapor	x	x
convencer	x	x
convidar	x	x
convocar	x	x
cortar	x	
crer	x	x
criticar	x	x
cumprimentar	x	x
debater		x
decidir	x	x
decretar	x	x
defender	x	x
definir	x	x
deixar escapar	x	x
deliberar	x	x
demonstrar	x	
denunciar	x	x
desabafar	x	x
desafiar	x	x
desconfiar	x	x
descrever	x	x

desculpar	x	x
desmentir	x	x
despistar	x	
detalhar	x	x
detonar	x	
devolver	x	
diferenciar	x	
discordar	x	x
discorrer	x	x
discursar	x	
discutir		x
disparar	x	
ditar	x	x
divertir\+se	x	
divulgar	x	x
elogiar	x	x
empolgar\+se	x	
ensinar	x	x
entregar	x	x
entusiasmar\+se	x	
enumerar	x	x
enunciar	x	x
esmiuçar	x	
especificar	x	x
especular	x	x
esperar	x	x
estabelecer	x	x
estimar	x	x
estipular	x	x
estranhar	x	
evocar	x	x
exemplificar	x	x
exigir	x	x
explicitar	x	x
expressar	x	x
exprimir	x	x
exultar	x	
festejar	x	
filosofar	x	

finalizar	x	
gabar-se	x	
gargalhar	x	
gracejar	x	
hesitar	x	
homenagear	x	x
idealizar	x	x
ilustrar	x	
imaginar	x	x
impor	x	x
incentivar	x	x
indagar	x	x
indicar	x	x
indignar\+se	x	
interrogar	x	
intimar	x	
inventar	x	x
ironizar	x	x
julgar	x	x
justificar	x	x
lamentar	x	x
mandar	x	x
manter	x	x
mencionar	x	x
mentir	x	x
minimizar	x	
mostrar(-se)	x	
murmurar	x	x
narrar	x	x
notar	x	x
noticiar	x	x
opinar	x	x
orgulhar\+se	x	
pedir	x	x
pensar	x	x
planejar	x	
pontuar	x	
pormenorizar	x	x
precisar	x	x

preconizar	x	x
preferir	x	x
presumir		x
prever	x	x
profetizar	x	x
proibir		x
pronunciar	x	x
propor	x	x
prosear	x	
prosseguir	x	x
provocar	x	x
querer saber	x	x
reagir	x	
realçar	x	x
rebater	x	x
recear	x	
recitar	x	x
reclamar	x	x
recomendar	x	x
recordar	x	x
recusar	x	x
referir	x	x
refletir	x	x
reforçar	x	x
registrar	x	x
regozijar\+se	x	
rejeitar	x	x
relembrar	x	x
rematar	x	x
rememorar	x	
repreender	x	
resignar\+se	x	
ressalvar	x	x
resumir	x	x
retomar	x	
retorquir	x	x
retratar	x	
rezar	x	x
ripostar	x	x

rir	x	x
saudar	x	
sentenciar	x	x
simplificar	x	
sinalizar	x	x
sintetizar	x	
situar	x	
solicitar	x	x
soltar	x	
sorrir	x	x
sublinhar	x	x
supor	x	x
suspeitar	x	x
sustentar	x	x
teimar	x	x
temer	x	x
teorizar	x	x
testemunhar	x	x
torcer	x	
vaticinar	x	x
xingar	x	x

Quadro 17: Organização dos verbos encontrados nesta pesquisa, de acordo com a classificação feita por Moura Neves (2000).

Diante da análise feita acima, merecem destaque os cinco verbos que não introduziram citação direta em nenhuma ocorrência: *acrescer*, *debater*, *discutir*, *presumir* e *proibir*. Novamente, ressaltamos que o fato de não encontrarmos ocorrências desses verbos no discurso direto não significa, necessariamente, que o emprego nesse tipo de contexto não seja possível. Como já afirmamos, apesar de os corpora analisados apenas darem conta de uma parcela ínfima do que pode ser produzido na língua portuguesa, a não ocorrência pode, talvez, indicar que o emprego desses verbos como introdutores de discurso direto seja pouco comum.

5.2

Manual do Estadão: os verbos “mais que errados”

A observação dos verbos em um contexto nos permitiu, também, mais segurança para confrontar outros trabalhos que tematizam o discurso relatado. Damos destaque ao que é afirmado pelo Manual de Redação do Estadão. Segundo o manual, que tem caráter estilístico e, por isso, adota tom normativo, existem verbos que equivalem ao sentido de dizer e que, por isso, as ocorrências desses verbos no discurso indireto ou misto são “erradas ou no mínimo impróprias”. Esta pesquisa permitiu apresentar uma realidade distante daquela desenhada no manual. Tomando por base textos dos jornais *Folha de S. Paulo* e *Público* (Portugal), apresentamos, no Quadro 18¹⁷, os verbos que, para o manual, “não podem ser acompanhados de *que*”, isto é, não admitem discurso indireto na forma mais usual, analisados nos corpora do projeto AC/DC:

VERBO	ACEITA "QUE"?	LINHA DE CONCORDÂNCIA
acusar	sim	Muitos dos radicais acusavam que a política econômica e a proposta de reforma agrária eram socialistas», registrava.

¹⁷ Nem todos os verbos listados no quadro constam de nosso léxico, uma vez que nem todos correspondem ao nosso entendimento de discurso relatado. Tais verbos estão indicados com um *.

alertar	sim	O papa João Paulo 2º alertou que o mundo precisa cuidar para que não haja outros campos de Auschwitz.
antecipar*	sim	No mesmo dia, o presidente do Sindicato dos Motoristas, Edivaldo Santiago, antecipou que a decisão judicial seria desrespeitada.
apontar	sim	Ele apontou que tirou lições do episódio.
aprovar*	sim	O cardeal aprovou que a arquidiocese entre na luta contra o projeto que dá calote na Educação.
assumir	sim	Sabrina se define como «meio preguiçosa» e assume que não tem paciência para ajudar em casa.
citar	sim	Para embasar o ceticismo sobre o fim da inflação, Maluf citou que algumas lojas paulistanas estariam calculando juros de 100% ao ano nas compras pelo crediário.
comentar	sim	Altman comentou que ele havia agido assim para ganhar mais publicidade, uma vez que os outros estilistas colaboraram com o filme.
continuar	sim	Mas ante um persistente e interrogativo olhar do administrador, como que a intimá-lo a falar, ele continuou que não lhe tinha emprestado as pedras por uma questão de confiança, porque neste aspecto até poderia dizer que não confiava nem nos seus próprios dentes porque de quando em vez lhe mordiam a língua...
defender	sim	Ele defende que o banco mantenha seu sistema atual, que limita os empréstimos às empresas que se comprometam a usar o dinheiro em investimentos.
definir	sim	O próprio presidente eleito definiu que a pasta será uma das mais importantes de seu governo.
denunciar	sim	Leônidas Menezes da Silveira denunciou que seu filho Ulisses de Souza da Silveira, 14, ficou parálítico após tomar vacina anti-rábica, em Manaus.
desmentir	sim	Ele desmente que não goste desta canção, mas não consegue explicar sua popularidade por aqui.
difundir*	sim	Henequim, porém, insistia em difundir que o trono de Deus flutuava sobre o centro do Brasil.

divulgar	sim	Em 94, o candidato derrotado, Andrés Pastrana, divulgou que um dos Orejuela insinuava ter dado dinheiro para a campanha.
ênfatizar	sim	O economista e ex-ministro Mário Henrique Simonsen não chega a dizer que o programa está lento, mas ênfatiza que o governo não pode retardar as privatizações.
indicar	sim	Mas Clinton indicou que não pretende mexer no auxílio à reforma econômica na Rússia.
justificar	sim	Osiris justificou que esse critério facilita a operacionalização das informações.
mencionar	sim	Antes mesmo que Benito Gama denunciasse a existência do cheque, Righi mencionou que tinha feito «uma única transação comercial» com João Alves.
narrar	sim	Um ministro do STJ narra que é comum nas cidades do interior as prefeituras darem todo o material de apoio aos juízes, função que é do governo estadual.
preferir*	sim	No início do século XX, Harvey Cushing, além de introduzir a incisão sublabial na via transesfenoidal para o tratamento cirúrgico dos tumores hipofisários, profeticamente preferiu que as drogas seriam a terapia do futuro para estes adenomas.
prosseguir	sim	«Acredito pouco que as coisas se modifiquem», afirma, antes de prosseguir que «quando a lógica das instituições não coincide com a minha, elas continuam, eu saio».
referir	sim	Pessoa e Costa referiu que seria conveniente alterar, nesta ação, o termo 'beneficiária' para o termo 'gestora'.
registrar	sim	Ele pára e registra que aquela renda, oito anos atrás, mostra bem como «o cinema acabou».
relatar	sim	Ela relata que a fissura do crack é maior que a da maconha, droga que havia experimentado antes.

Quadro 18: Análise do uso dos “verbos mais que errados”, listados pelo Manual de Redação e Estilo do Estado, nos corpora do AC/DC.

Vale ressaltar que todas as ocorrências relatadas acima advêm de fontes jornalísticas, gênero de que se ocupa o manual. Uma vez que o Manual de Redação

e Estilo do Estado é uma das principais referências dos tradutores e revisores brasileiros, temos que apontar que há um certo descolamento entre o que é recomendado pela obra e o que está sendo produzido pelos falantes – o que costuma ser frequente em manuais de caráter normativo.

Considerações finais

Esta dissertação se propôs a criar um glossário dos verbos de elocução em português, a fim de auxiliar profissionais de tradução. Para tanto, fizemos um amplo estudo descritivo dessa classe de verbos. O trabalho tomou por base grandes corpora da língua portuguesa e dialogou ainda com a Linguística Computacional.

Desenvolvemos a nossa pesquisa em duas vertentes. Com base em corpus, estabelecemos oito padrões gerais de uso dos verbos de elocução, a partir dos quais foi possível um levantamento desses verbos e, assim, construímos um grande léxico dos verbos de elocução do português, constituído por 308 verbos (os 293 verbos encontrados com a nossa abordagem somados aos os 15 verbos encontrados apenas por Moura Neves, porém confirmados por nós). Paralelamente, nos propusemos a classificar os verbos encontrados em grupos de sentido. Após duas tentativas e dois testes, alocamos o total de verbos em 11 grupos de sentido, usados para organizar o glossário DISSE.

Sabemos que a classificação dos verbos de elocução em grupos de sentido é tarefa complexa – talvez impossível, como sugere Austin. No entanto, tínhamos um problema concreto para resolver: organizar quase 300 verbos, tendo em vista as necessidades de um tradutor. Esperamos que a solução proposta seja satisfatória.

Esta pesquisa oferece, de maneira geral, três contribuições concretas. A primeira consiste no referido glossário, disponível na página da Gramateca¹⁸ e no Anexo 4. Com o léxico distribuído entre 11 grupos de sentido, organizamos, dentro desses grupos, os verbos por ordem decrescente de frequência, para facilitar a busca do tradutor. Assim, se o consultante estiver procurando por um verbo mais usual, pode voltar-se para o início da lista; se estiver buscando um verbo mais heterodoxo, o consultante pode dirigir-se ao final da relação.

Ressaltamos, aqui, a importância da compilação desse glossário para os tradutores e os demais profissionais que atuam na área de produção textual. As listas encontradas na maioria dos manuais de redação e na internet (que são, provavelmente, as duas principais fontes de consulta dos tradutores) parecem ser arbitrárias e, de acordo com as análises feitas nesta pesquisa, estão descoladas da língua efetivamente em uso. O fato de que não há nenhum material com base em

¹⁸ <http://linguateca.pt/Gramateca/>

corpus, voltado especificamente para tradutores e disponível on-line também torna o glossário DISSE ainda mais relevante.

A segunda contribuição consiste em uma um léxico de verbos específicos do campo do dizer, associado a padrões de uso (ou regras), elementos de grande relevância para a tarefa de Identificação de Citação, como ilustrado por Sagot et al. De fato, este léxico e verbos já estão sendo utilizados na criação de um sistema extrator de citações, desenvolvido no âmbito da dissertação de mestrado de Rafael Reis, do Departamento de Informática da PUC-Rio (Reis, 2015).

Também resultado do léxico e dos padrões, está em andamento a anotação completa dos corpora do projeto AC/DC, permitindo, assim, a busca pelo campo semântico *dizer*. Como explicitamos ao longo da dissertação, o material é público e está disponível para consultas on-line. Com isso, tornamos possível a busca por verbos de elocução (além da especificação do tipo de relato introduzido: direto, indireto e/ou misto). Hoje, o material já está implementado e revisto em apenas um dos corpora, mas o processo está em andamento e, em poucas semanas, todo o material estará anotado.

Vale mencionar também que, do ponto de vista da Linguística Computacional, os 308 verbos compilados ao longo do trabalho estão sendo integrados à OpenWordNet-PT¹⁹, uma base de dados lexical de amplo uso. A OpenWordNet-PT é usada, por exemplo, pelo Google Tradutor.

Com relação à descrição do português, as contribuições são bastante significativas. Em termos quantitativos, o fato de termos compilado 293 verbos, isto é, mais verbos do que qualquer estudo consultado por nós, seja em português, seja em outras línguas, aponta o sucesso do caminho que escolhemos – o caminho da consulta em grandes corpora. A apreensão de padrões envolvidos no relato, isto é, a observação de que um mesmo tipo de citação (direta, indireta ou mista) pode ser expressa por meio de diferentes padrões também são contribuições que julgamos valiosas, e que explicitam as potencialidades dos estudos descritivos com base em grandes corpora da nossa língua, ainda incipientes.

Gostaríamos de poder comparar os resultados de Biber et al. (1999), para o inglês, e de Sagot et al. (2010), para o francês, com a língua portuguesa. No entanto, sabemos que, considerando a imensa lista de verbos de elocução e o

¹⁹ <http://wnpt.brcloud.com/wn/>

reconhecimento de que a presença de tais verbos, por si só, não é conclusiva quanto à presença de um discurso relatado, a comparação com Biber et al. só pode ser viabilizada diante de um corpus com anotação semântica relacionada ao *dizer*. Pretendemos apresentar esses resultados assim que a anotação dos corpora do AC/DC estiver concluída.

Apesar de os resultados desta pesquisa serem bastante amplos, ainda há muito a ser feito no que diz respeito aos verbos de elocução. Uma das linhas que podem ser seguidas se refere à diferença entre os usos no inglês e português, evidenciada pela nossa primeira busca no COMPARA. Não buscamos entender o motivo da diferença, pois julgávamos que se tratava apenas de uma questão de estilo. No entanto, é possível que essa diferença tenha origem na tipologia das línguas. Nesse caso, o maior emprego de *say*, em comparação com a variação dos verbos no português, estaria ligado ao fato de o inglês ser uma língua com frame no verbo (*verb-framed*), ao passo que o português seria uma língua com o frame no satélite (*satellite-framed*), ou, ainda, que o português estaria em uma posição intermediária entre as línguas com frame no verbo e as línguas com frame no satélite.

Seria necessário, entretanto, investigar a pertinência dessa hipótese, uma vez que essa classificação tipológica dos verbos se refere apenas aos verbos de movimento. Infelizmente, por ser esta uma questão bastante complexa e, portanto, merecedora de um aprofundado estudo, tivemos que abrir mão de investigá-la, devido à limitação de tempo. Fornecemos aqui, entretanto, um bom mapeamento do emprego desses verbos no português, o que pode ser fundamental para entender o porquê dessa distinção entre as duas línguas.

Também acreditamos ser de enorme relevância um estudo relacionado à prosódia semântica dos verbos aqui estudados. Sabemos que a escolha por um determinado verbo para relatar um evento nunca é neutra, e a escolha lexical nas frases abaixo evidencia esse ponto:

Prefeito **alertou** o COI sobre possibilidade de atraso.

Prefeito **admitiu** ao COI possibilidade de atraso.

Não foi possível aqui a inclusão dessa dimensão dos verbos de elocução, mas lembramos que o material produzido é público e está disponível para a comunidade.

Boa parte dos estudos dos verbos na língua inglesa trata da classe de verbos de comunicação como um todo. Seria de grande valia uma investigação mais aprofundada dessa classe mais abrangente dos verbos de comunicação na língua portuguesa, considerando-se o papel proeminente do *dizer* dentro da língua (lembramos aqui que o verbo *dizer* é o quarto mais frequente dentro dos corpora que consultamos). A anotação semântica dos verbos dos corpora do AC/DC, que ganhou o campo semântico *dizer* (expresso por meio do sema “dizer”), poderá ser de enorme utilidade para aqueles que desejam explorar uma das classes de verbos mais amplas do português.

Além disso, na comparação com outras línguas, nossos resultados sugerem um caminho de investigação relevante: se nossa classificação semântica dos verbos (o glossário) for uma classificação produtiva na língua (e isso ainda está em aberto), os resultados põem em xeque toda a proposta de Levin – não pelo fato de a classificação ser diferente, mas porque, em uma mesma classe, temos verbos que participam de diferentes padrões sintáticos. Ou seja, a correspondência entre uma classificação semântica e um padrão sintática parece não se verificar.

Por fim, esperamos ter demonstrado, ao longo deste trabalho, como pode ser produtivo o diálogo entre a descrição de uma língua, a tradução e a linguística computacional. A descrição, de qualquer tipo, não pode ser uma atividade abstrata, descolada de intenções e objetivos, o que leva a nossa pesquisa a apontar para a adequação de parcerias com a Tradução e com a Linguística Computacional, áreas aplicadas.

Da Tradução, surgiu não apenas a motivação inicial do trabalho, mas também a garantia de reconhecimento dos verbos como introdutores do discurso, através da utilização do corpus COMPARA. A parceria com a Linguística Computacional possibilitou não apenas a “oferta” de “produtos” como léxico e regras, mas a própria necessidade de “entregar” tais produtos obrigou a uma precisão ainda maior na descrição. Como pano de fundo e costurando as interfaces, encontram-se o corpus e o processo de anotação.

Referências bibliográficas

ARROJO, Rosemary. A desconstrução do logocentrismo e a origem do significado. In: **O signo desconstruído: implicações para a tradução, a leitura e o ensino**. Rosemary Arrojo (Org.). 2ª ed. Campinas, SP: Pontes, 2003a.

_____. Compreender x interpretar e a questão da tradução. In: **O signo desconstruído: implicações para a tradução, a leitura e o ensino**. Rosemary Arrojo (Org.). 2ª ed. Campinas, SP: Pontes, 2003b.

_____. O ensino de tradução e seus limites: por uma abordagem menos ilusória. In: **O signo desconstruído: implicações para a tradução, a leitura e o ensino**. Rosemary Arrojo (Org.). 2ª ed. Campinas, SP: Pontes, 2003c.

AUSTIN, J. L. **How to do things with words**. 2ª ed. Cambridge, MA: Harvard University Press, 1962.

BEAUGRANDE, Robert-Alain de. Descriptive linguistics at the millennium: corpus data as authentic language. **Journal of Language and Linguistics**. 1 (2): 91–131, 2002.

BIBER, Douglas, JOHANSSON, Stig, LEECH, Geoffrey, CONRAD, Susan e FINEGAN, Edward. **Longman Grammar of Spoken and Written English**. Harlow: Pearson Education ESL, 1999.

ELLIS, John Martin. **Language, thought and logic**. Evanston, Illinois: Northwestern University Press, 1993.

FRANKENBERG-GARCIA, Ana e SANTOS, Diana. COMPARA, um corpus paralelo de português e de inglês na Web. **Cadernos de Tradução** IX.1 (2002), pp. 61-79. Universidade de Santa Catarina. ISSN: 1676-7047.

FREITAS, Cláudia, FREITAS, Bianca & SANTOS, Diana. "QUEMDISSE?: Reported speech in Portuguese", **Proceedings of LREC 2016**. (no prelo)

FREITAS, Cláudia. Corpus, Linguística Computacional e as Humanidades Digitais. In: LEITE, Miriam S., GABRIEL, Carmen Teresa. (org.) **Linguagem, discurso, pesquisa e educação**. 1ª ed. Petrópolis, RJ: De Petrus; Rio de Janeiro: FAPERJ, 2015.

GARCIA, Othon M. **Comunicação em prosa moderna: aprenda a escrever, aprendendo a pensar**. 27 ed. Rio de Janeiro: Editora FGV, 2010.

KIPPER K., KORHONEN A., RYANT N. & PALMER, M. Extending VerbNet with Novel Verb Classes. In: **Proceedings of the 5th International Conference on Language Resources and Evaluation** (LREC 2006), Genoa, Itália, 2006.

LOFFREDO, Laís, GROSSMAN, Deborah, BITAR, Gladys & GONÇALVES, Janice. Verbos de Elocução - As Diferenças entre o Inglês e o Português. **CROP - Revista da Área de Língua e Literatura Inglesa e Norte-Americana** 10 (2004), pp. 167-184. Brasil : Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo.

LEVIN, Beth. **English verb classes and alternations**. Chicago: The University of Chicago Press, 1993.

Manual de redação e Estilo do Estado. Disponível em: <http://www.estadao.com.br/manualredacao/>. Verbete “Verbos mais que errados”. Acesso em 10 jan 2016.

MARTINS, Helena Franco. **Metáfora e polissemia no estudo das línguas do mundo: uma aproximação não representacionista**. Tese de doutorado. Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1999.

NEVES, Maria Helena de Moura. **Gramática de usos do português**. São Paulo: Editora Unesp, 2000.

OLIVEIRA, Ana Maria P., de, LONGO, Beatriz N. O., DEZOTTI e Maria Celeste C. **Verbos introdutórios de discurso direto**. Alfa, São Paulo, 29:91-96, 1985.

REIS, Rafael. **Indirect Quotation Extraction for Portuguese**. Proposta de Qualificação de Mestrado. Orientador: Ruy Milidiú. Departamento de Informática, PUC-Rio. Novembro, 2015.

SAGOT, Benoît, DANLOS, Laurence e STERN, Rosa. A lexicon of French quotation verbs for automatic quotation extraction. In: **Proceedings of LRED 2010**, La Valette, Malta (2010).

SAMPSON, Geoffrey. *Empirical Linguistics*. London: Continuum, 2001.

SANTOS, Diana. Gramateca: corpus-based grammar of Portuguese. In Jorge Baptista, Nuno Mamede, Sara Candeias, Ivandré Paraboni, Thiago A.S. Pardo & Maria das Graças Volpe Nunes (eds.), **International Conference on Computational Processing of Portuguese (PROPOR'2014)** (São Carlos, Outubro de 2014), Springer, pp. 214—219.

SANTOS, Diana. Linateca's infrastructure for Portuguese and how it allows the detailed study of language varieties. OSLa 3 (2), 2011, pp. 113-128. J.B. Johannessen (ed.), Language variation Infrastructure.

SANTOS, Diana, MARQUES, Rui, FREITAS, Cláudia, MOTA, Cristina & SIMÕES, Alberto. Comparando anotações na Gramateca: filosofia, ferramentas e exemplos. Domínios da Lingu@agem, 2015.

SANTOS, Diana, SILVA, Rosário & FREITAS, Cláudia. Pluralidades na cor: contrastando a língua do Brasil e de Portugal. In Augusto Soares da Silva, Amadeu Torres & Miguel Gonçalves (orgs.), **Línguas Pluricêntricas: Variação Linguística e Dimensões Sociocognitivas. Pluricentric Languages: Linguistic Variation and Sociocognitive Dimensions**. Braga: Aletheia, Publicações da Faculdade de Filosofia da Universidade Católica Portuguesa, 2011, pp. 555-572.

SANTOS, Diana & ROCHA, Paulo. The key to the first CLEF in Portuguese: Topics, questions and answers in CHAVE. In: Carol Peters, Paul Clough, Julio Gonzalo, Gareth J. F. Jones, Michael Kluck & Bernardo Magnini (eds.), **Multilingual Information Access for Text, Speech and Images**, 5th Workshop of the Cross-Language Evaluation Forum, CLEF 2004, Bath, UK, September 15-17, 2004, Revised Selected Papers. Berlin/Heidelberg: Springer, Lecture Notes in Computer Science, 2005, pp. 821-832. Revised version of Santos & Rocha (2004).

SANTOS, Diana & SARMENTO, Luís.. O projecto AC/DC: acesso a corpora/disponibilização de corpora. In Amália Mendes & Tiago Freitas (eds.), **Actas do XVIII Encontro Nacional da Associação Portuguesa de Linguística (APL 2002)** (Porto, Portugal, 2-4 de Outubro de 2002), Lisboa : APL, pp. 705-717.

SCARTON, C. e ALUISIO, S. Towards a cross-linguistic VerbNet-style lexicon to Brazilian Portuguese. In: **Proceedings of the LREC 2012 Workshop on Creating Cross-language Resources for Disconnected Languages and Styles (CREDISLAS 2012)**, Istambul, Turquia, 2012, p. 11-18.

Anexo 1

Expressões de busca específicas para o projeto AC/DC, desenvolvidas para cinco dos padrões de uso.

PADRÃO	EXPRESSÃO DE BUSCA ²⁰
Padrão (1) ²¹	[word=""".* > -- —"] [word=","] @[pos="V.*" & temcagr!=".*PCP.* .GER.* .PASSIVA.* .FUT_SUBJ.* .PR_SUBJ.* .FUT_IND.* 0 .PR_PROG_IND.* .IMPF_FUTHAVER_IND.* .IMPF_SUBJ.* .COND_IND.*"]
Padrão (2) ²²	[pos="V.*" & temcagr!=".*PCP.* .GER.* .PASSIVA.* .FUT_SUBJ.* .PR_SUBJ.* .FUT_IND.* 0 .PR_PROG_IND.* .IMPF_FUTHAVER_IND.* .IMPF_SUBJ.* .COND_IND.*"] ":" [word=""".* > -- —"]
Padrão (3) ²³²⁴	[word=""".* > -- —"] [word=","] [pos="V.*" & temcagr!=".*PCP.* .GER.* .PASSIVA.* .FUT_SUBJ.* .PR_SUBJ.* .FUT_IND.* 0 .PR_PROG_IND.* .IMPF_FUTHAVER_IND.* .IMPF_SUBJ.* .COND_IND.*"] [pos="DET.*"]* [pos="N PROP PERS"]* [word=","] [word=""".* < -- —"] [word!=""".* > -- —"] [word="-- ,"] @[pos="V V_fmc V-REGENTE_fmc" & temcagr!=".*PCP.* .GER.* .PASSIVA.* .FUT_SUBJ.* .PR_SUBJ.* .FUT_IND.* 0 .PR_PROG_IND.* .IMPF_FUTHAVER_IND.* .IMPF_SUBJ.* .COND_IND.*"] [pos="DET.*"]* [pos="N PROP PERS"]* [word="-- ,"] [word!=""".* < -- —"]

²⁰ Para facilitar a leitura dos resultados, filtramos alguns casos em que o verbo não pode ser ou dificilmente é introdutor de discurso relatado, eliminando os seguintes tempos ou formas verbais (na ordem em que aparecem nas expressões de busca): particípio, gerúndio, voz passiva, futuro do subjuntivo, pretérito do subjuntivo, futuro do indicativo, presente progressivo do indicativo e futuro perifrástico com haver.

²¹ Com essa expressão, pede-se ocorrências que contenham o que está sublinhado: «Acabou a democracia? », ironizou a sindicalista.

²² Eles disseram: « Bom, então você apresente suas críticas».

²³ «Benzinho », dizele, « me pega mais um whiskey, mas, ué, você não está aqui? »

²⁴ Isto porque, adianta, se os terrenos em causa forem integrados no Parque Periférico, terão de ir à posse da Câmara.

Padrão (4) ²⁵	[pos="V.*" & temcagr!=".*PCP.* . *GER.* . *PASSIVA.* . *FUT_SUBJ.* . *PR_SUBJ. * . *FUT_IND.* 0 . *PR_PROG_IND.* . *IMPF_FUTHAVER_IND.* . *I MPF_SUBJ.* . *COND_IND.*"] [pos="KS" & word="que"]
Padrão (8) ²⁶	[word=","] [word="como conforme segundo"] @[pos="V.*" & temcagr!=".*PCP.* . *GER.* . *PASSIVA.*"] [pos="DET.*"]* [pos="N PROP PERS"]

²⁵ Archângelo disse que não houve nenhum tipo de incidente durante as provas.

²⁶ Serão realizadas até até, segundo afirma ele, visitas a empresas.

Anexo2

Lista dos verbos encontrados, distribuídos por padrão.

Padrão 1

acentuar, aconselhar, acreditar, acrescentar, acusar, adiantar, admirar\.*, admitir, advertir, aferir, afirmar\.*, alegar, alegrar\.*, alertar, alfinetar, ameaçar, amenizar, analisar, animar\.*, anotar, anuir, anunciar, apelar, apoiar, apontar, apostar, argumentar, arrematar, arriscar, assegurar, assinalar, assumir, atacar, atestar, atirar, avaliar, avisar, balbuciar, berrar, bradar, brincar, brindar, cantar, censurar, chamar, citar, clarificar, colocar, comemorar, comentar\.*, começar, comparar, complementar, completar, concluir, concordar, confessar, confirmar, considerar, constatar, contabilizar, contar.*, contestar, continuar, contrapor, convidar, cortar, criticar, decidir, declarar, decretar, defender.*, definir, deixar escapar, demonstrar, denunciar, desabafar, desafiar, despistar, destacar, detalhar, detonar, devolver, diferenciar, discorrer, discursar, disparar, divertir\.*, dizer, dizer\+eu, dizer\+nós, elogiar, emendar, empolgar\.*, enfatizar, ensinar, entregar, entusiasmar\.*, enumerar, esclarecer, espantar\.*, esperar, estranhar, exclamar, exemplificar, explicar, explicar\.*, expor, exultar, falar, festejar, filosofar, finalizar, frisar, gabar\.*, garantir, gargalhar, gracejar, gritar, idealizar, ilustrar, imaginar, indagar, indicar, indignar\.*, informar, insistir, interrogar, interromper, inventar, ironizar, jurar, justificar\.*, lamentar\.*, lembrar, minimizar, mostrar\.*, murmurar, narrar, notar, observar, opinar, orgulhar\.*, pedir, pensar, perguntar, planejar, ponderar, pontuar, precisar, preconizar, prever, proclamar, profetizar, prometer, propor, prosear, prosseguir, protestar, provocar, queixar\.*, querer saber, questionar\.*, reafirmar, reagir, rebater, recear, reclamar, reconhecer, recordar, referir, referir\.*, refletir|reflectir, reforçar, registrar, regozijar\.*, reiterar, relatar, relembrar, rematar, rememorar, repetir, resignar\.*, responder, ressaltar, resumir, retomar, retorquir, retratar, revelar, rir.*, salientar, saudar, sentenciar, simplificar, sintetizar, situar, sorrir, sublinhar\.*, sugerir, sustentar, temer, teorizar, testemunhar, torcer, vaticinar

Padrão 2

aconselhar, acrescentar, acusar, adiantar, advertir, afirmar, ajudar, alertar, alfinetar, anunciar, apontar, argumentar, arrematar, arriscar, assegurar, assinalar, atacar, atestar, avisar, brincar, cantar, chamar, chorar, clamar, comentar, complementar, completar, concluir, concordar, confessar, confidenciar, confirmar, contar, continuar, convidar, declarar, defender, definir, deixar escapar, desabafar, descrever, destacar, discordar, disparar, dizer.*, emendar, enfatizar, ensinar, esclarecer, exclamar, exemplificar, explicar, falar, filosofar, finalizar, garantir, gritar, hesitar, indagar, insistir, interrogar.*, interromper, intimar, ironizar, justificar, lamentar, lembrar, narrar, notar, observar, pedir, pensar, perguntar, precisar, proclamar, profetizar, prometer, propor, prosseguir, provocar, questionar, rebater, reclamar, recomendar, recordar, refletir, registrar, relatar, relembrar,

rematar, repetir, replicar, resmungar, responder, ressaltar, resumir, retrucar, revelar, ripostar, sentenciar, soltar, sorrir, sugerir, teorizar

Padrão 3

acentuar, acreditar, acrescentar, acusar, adiantar, admitir, afirmar, agradecer, ameaçar, analisar, anotar, anunciar, apontar, apresentar\.*, argumentar, assegurar, assumir, avaliar, brincar, cantar, chamar, citar, comemorar, comentar, complementar, conceder, concluir, concordar, confessar, confirmar, considerar, contar, contestar, continuar, convidar, convocar, crer, criticar, cumprimentar, debater, decidir, declarar, defender, desabafar, desafiar, desculpar, desmentir, destacar, discutir, divulgar, dizer, elogiar, enfatizar, ensinar, enunciar, esclarecer, esmiuçar, especificar, estipular, evocar, exigir, explicar, expor, expressar, exprimir, falar, frisar, gracejar, gritar, homenagear, idealizar, imaginar, impor, indicar, informar, insinuar, insistir, interrogar.*, inventar, ironizar, julgar, jurar, justificar.*, lamentar.*, lembrar.*, mentir, narrar, negar, notar, observar, opinar, orgulhar\.*, pedir, pensar, perguntar.*, pormenorizar.*, precisar, prometer, propor, prosseguir, protestar, provocar, querer saber, questionar, realçar, recitar, reclamar, reconhecer, recordar, referir, refletir, registrar, rejeitar, relatar, relatar.*, relemburar, rematar, repetir, repreender.*, responder.*, ressaltar, ressalvar, resumir, retorquir, revelar, rezer, rir, salientar, sentenciar, sintetizar, solicitar, sorrir, sublinhar, sugerir, supor, sustentar, xingar,

Padrão 4

acentuar, aconselhar, acrescentar, crescer, adiantar, admitir, advertir, advogar, afiançar, afirmar, alardear, alegar, alertar, aludir, analisar, antever, anuir, anunciar, apontar, apostar, apregoar, apurar, argumentar, assegurar, asseverar, assinalar, assumir, atestar, avaliar, avisar, berrar, brincar, calcular, cantar, chamar, chutar, citar, colocar, combinar, comentar, completar, comunicar, conceder, concluir, concordar, confessar, confiar, confidenciar, confirmar, constatar, contar, contestar, convencer, criticar, declarar, defender, definir, deixar escapar, deliberar, denunciar, desconfiar, desmentir, destacar, detalhar, determinar, discordar, discutir, ditar, divulgar, dizer, enfatizar, esclarecer, especificar, especular, esperar, estabelecer, estimar, estipular, exclamar, exemplificar, exigir, explicar, explicitar, falar, frisar, garantir, gritar, impor, incentivar, indagar, indicar, informar, insinuar, insistir, inventar, jurar, justificar, lamentar, lembrar, manter, mencionar, negar, notar, noticiar, observar, opinar, ordenar, pedir, perguntar, ponderar, precisar, preferir, pregar, presumir, prevenir, prever, proclamar, proibir, prometer, pronunciar, propor|propôr, queixar, questionar, reafirmar, realçar, reclamar, recomendar, reconhecer, recordar, recusar, referir, reforçar, registrar, rejeitar, relatar, relemburar, repetir, replicar, resmungar, responder, ressaltar, revelar, rezer, salientar, sentenciar, sinalizar, solicitar, sublinhar, sugerir, suplicar, supor, suspeitar, sustentar, teimar, temer, testemunhar, vaticinar

Padrão 8

adiantar, afirmar, alertar, anunciar, apontar, assegurar, assinalar, cantar, citar, colocar, comentar, constatar, declarar, defender, definir, denunciar, destacar, divulgar, dizer, ensinar, esclarecer, explicar, explicitar, expressar, garantir, informar, insistir, lembrar, ler, mandar, mostrar, notar, noticiar, observar, preconizar, prever, propor, publicar, querer saber, recomendar, recordar, referir, registrar, relatar, ressaltar, revelar, rezer, salientar, sintetizar, sublinhar, sugerir, sustentar,

Anexo3

Respostas dadas pelos dez participantes da primeira versão do experimento, que era composto por 58 verbos.

VERBO ▾	PARTICIPANTES									
	1 ▾	2 ▾	3 ▾	4 ▾	5 ▾	6 ▾	7 ▾	8 ▾	9 ▾	10 ▾
dizer	afirmação	neutralidade	acordo confirmação	confirmação opinião resposta	resposta	afirmação resposta	afirmação	opinião	afirmação	afirmação
concordar	confirmação	acordo confirmação	confirmação	confirmação resposta	confirmação	acordo afirmação	acordo	confirmação	acordo resposta	confirmação
ripostar	contestação	censura	censura contestação	contestação destaque	resposta	afirmação contestação opinião resposta		contestação	afirmação contestação	
observar	afirmação	opinião	constatação	acréscimo opinião detalhamento	afirmação	opinião	afirmação informação	opinião	afirmação opinião	
anunciar	informação	afirmação exclamação	informação	afirmação anúncio	explicação	informação	afirmação destaque exclamação	informação	afirmação informação	afirmação
pedir	pergunta	pedido	pedido	pedido	pedido	pedido	pedido sugestão	pedido	pedido comando	pedido
argumentar	contestação	contestação opinião	explicação opinião	contestação explicação sugestão	contestação	contestação opinião	contestação opinião	opinião	contestação explicação	explicação
sublinhar	destaque	destaque	destaque	destaque	destaque	acréscimo contestação destaque	destaque	confirmação	destaque	destaque
admitir	confissão	confissão hesitação	confissão	acordo confissão	acordo	acordo confirmação	afirmação	confissão	confissão	confirmação
acrescentar	acréscimo	acréscimo sugestão	acréscimo	acréscimo	acréscimo	acréscimo afirmação	acréscimo	destaque	acréscimo	acréscimo
anuir	acordo	acordo confirmação	acordo	acordo confirmação	acordo	acordo	acordo	confirmação opinião	acordo confirmação	
berrar	exclamação	censura pedido	exclamação	censura destaque exclamação	afirmação	exclamação	destaque exclamação	censura	exclamação	exclamação
replicar	resposta	censura contestação	resposta	censura contestação opinião resposta	contestação	afirmação contestação resposta	contestação resposta	censura contestação	contestação resposta	contestação
tornar	resposta	neutralidade	resposta	amenização	afirmação			censura	resposta	opinião
retrucar	resposta	confissão	resposta	contestação resposta	resposta	afirmação resposta	contestação resposta	confissão	resposta	contestação
esclarecer	explicação	contestação	explicação	informação resposta	censura	explicação	afirmação explicação resposta	contestação	afirmação informação	informação
prometer	resposta	confirmação	promessa	destaque resposta	contraposição	afirmação	promessa	acordo		acordo
contrapor	sugestão	acréscimo	contestação	acréscimo destaque opinião amenização	opinião	opinião resposta	contestação opinião	contestação opinião	acréscimo contestação resposta	exclamação

censurar	sugestão	censura	censura	censura interrupção pedido	censura	afirmação censura exclamação	censura	censura	censura	censura
agrade cer	resposta	agradecimento	exclamação	resposta cortejo	opinião	afirmação destaque	gratidão	confirmação		afirmação
advertir	sugestão	opinião advertência	censura sugestão	censura contestação opinião sugestão advertência	censura	censura pedido sugestão	contestação sugestão	conselho	censura	explicação
sondar	hesitação	hesitação pergunta	pergunta especulação	pergunta	pergunta	contestação indagação	pergunta	afirmação confirmação	pergunta	opinião
saudar	exclamação	neutralidade	exclamação	cortejo	exclamação	exclamação saudação cumprimento	destaque exclamação	exclamação	saudação	exclamação
escrever	afirmação	neutralidade	destaque	afirmação destaque cortejo	afirmação	lembança	afirmação	informação		lembança
responder	resposta	hesitação	resposta	censura contestação hesitação opinião resposta	hesitação	hesitação	resposta	hesitação	resposta	contestação
perguntar	pergunta	contestação hesitação pergunta	pergunta	contestação pergunta	pergunta	contestação pergunta	pergunta	acordo	pergunta	hesitação
confessar	confissão	confissão	confissão	confissão contestação resposta	confissão	afirmação confissão	afirmação confissão	hesitação	confissão	confissão
propor	sugestão	sugestão	sugestão	afirmação exclamação sugestão	sugestão	exclamação sugestão	pedido sugestão	exclamação	sugestão	sugestão
comunicar	informação	afirmação informação	afirmação explicação	afirmação informação	afirmação	informação	afirmação informação	informação	afirmação informação	afirmação
pros e guir	acrécimo	acrécimo neutralidade	interrupção	acrécimo opinião	contraposição	acrécimo	acrécimo afirmação	acordo sugestão		informação
re pet ir	interrupção	interrupção neutralidade	destaque lembança	destaque interrupção	destaque	acrécimo	destaque	destaque interrupção		interrupção
queixar-se	opinião	queixa	confissão exclamação	censura opinião resposta	opinião	reclamação	censura opinião	conclusão	opinião reclamação	lamento
re torquir	interrupção	censura contestação	resposta	censura contestação exclamação	censura	resposta	contestação resposta	censura	opinião resposta	contestação
que re r saber	pergunta	pergunta	pedido pergunta	pergunta surpresa	pergunta	contestação pergunta	pergunta	pergunta	pergunta	interrupção pedido
asse gurar	confirmação	afirmação	confirmação destaque	afirmação destaque	afirmação	afirmação	afirmação confirmação	confirmação	afirmação	acordo
balbuciar	hesitação	hesitação	hesitação	exclamação hesitação pedido		hesitação		hesitação surpresa	hesitação	hesitação
confirmar	resposta	acordo confirmação	confirmação	confirmação resposta	confirmação	acordo confirmação resposta	acordo afirmação	confirmação	afirmação confirmação	confirmação
suge rir	sugestão	sugestão	sugestão	sugestão	sugestão	pergunta sugestão	sugestão	acordo sugestão	pergunta sugestão	sugestão

espantar-se	surpresa	surpresa	surpresa	censura surpresa cortejo	surpresa	pergunta	destaque exclamação pergunta	hesitação surpresa		exclamação surpresa
cumprimentar	exclamação	neutralidade	exclamação	cortejo	exclamação	exclamação saudação cumprimento	exclamação	exclamação	saudação	afirmação
contar	afirmação	relato	afirmação	informação	afirmação	contestação explicação informação lembança	afirmação	informação	informação relato	confissão
afirmar	afirmação	explicação	afirmação informação	afirmação contestação informação	afirmação	afirmação destaque	afirmação	censura contestação	afirmação	contestação
admirar-se	surpresa	exclamação surpresa	surpresa	surpresa cortejo	surpresa	surpresa	destaque pergunta	surpresa	surpresa deslumbramento	surpresa
continuar	acrécimo	afirmação	interrupção	acrécimo afirmação informação	confirmação	acrécimo	acrécimo	destaque		afirmação
exclamar	exclamação	exclamação	exclamação	exclamação opinião	exclamação	exclamação	explicação	exclamação opinião	exclamação	exclamação surpresa
declarar	afirmação	afirmação	afirmação	afirmação informação	afirmação	afirmação	afirmação	adversatividade	afirmação declaração	confissão
brindar	exclamação	destaque	exclamação	exclamação	exclamação	exclamação desejo	destaque	exclamação		exclamação
comentar	afirmação	acrécimo	afirmação opinião	afirmação exclamação opinião surpresa	exclamação	exclamação	afirmação	exclamação opinião	opinião	lamento
concluir	acrécimo	explicação	conclusão	acrécimo afirmação explicação informação resposta	opinião	afirmação opinião	afirmação	explicação	conclusão	confirmação
desabafar	confissão		confissão exclamação	censura confissão exclamação	pergunta	exclamação pergunta	destaque			
explicar	explicação	neutralidade	explicação	confirmação explicação resposta	confirmação	explicação	explicação		explicação	
falar	afirmação	acrécimo	declaração	afirmação lembança cortejo		afirmação	afirmação			
indagar	pergunta	pergunta	pergunta	contestação pergunta	pergunta	pergunta	pergunta		pergunta	
informar	informação	hesitação informação	informação	afirmação informação amenização	informação	afirmação	afirmação informação		informação	
insistir	opinião	destaque	destaque	afirmação destaque	afirmação	afirmação opinião	afirmação destaque opinião			
interromper	interrupção	exasperação	interrupção	contestação exclamação interrupção	exclamação	exclamação interrupção	exclamação interrupção		interrupção	
justificar	resposta	explicação	explicação	afirmação confissão explicação	contestação	afirmação	afirmação explicação resposta		explicação	
lembrar	lembança	acrécimo lembança	lembança	acrécimo lembança	lembança	acrécimo lembança	acrécimo lembança		lembança	

Anexo 4

Glossário DISSE.

DECLARAÇÃO

dizer - afirmar - apresentar - falar - considerar - mostrar - manter - prever - pedir - contar - esperar - colocar - decidir - anunciar - defender - precisar - garantir - explicar - chamar - acreditar - revelar - acusar - lembrar - assumir - exigir - informar - definir - referir - confirmar - determinar - divulgar - apontar - discutir - reconhecer - ajudar - estabelecer - indicar - concluir - entregar - registrar - preferir - apoiar - declarar - prometer - citar - negar - impor - justificar - criticar - mandar - apurar - analisar - assegurar - demonstrar - conceder - julgar - sugerir - comentar - alegar - atacar - proibir - imaginar - observar - adiantar - sustentar - convidar - comparar - convencer - esclarecer - supor - expor - temer - denunciar - convocar - situar - recordar - estimar - rejeitar - notar - descrever - avisar - constatar - recomendar - comunicar - comemorar - contestar - confessar - argumentar - ensinar - refletir - inventar - disparar - alertar - debater - protestar - lamentar - atirar - salientar - assinalar - resumir - brincar - mencionar - combinar - decretar - expressar - reafirmar - ameaçar - noticiar - confiar - relatar - ilustrar - pronunciar - elogiar - desmentir - ordenar - incentivar - homenagear - ditar - discordar - agradecer - apresentar-se - diferenciar - retratar - estranhar - torcer - frisar - evocar - reforçar - advertir - pregar - detalhar - proclamar - exprimir - estipular - anotar - narrar - advogar - minimizar - saudar - especificar - chutar - sinalizar - mentir - ponderar - jurar - contabilizar - atestar - contrapor - testemunhar - deliberar - simplificar - ironizar - cumprimentar - explicitar - insinuar - relemburar - discursar - preconizar - sintetizar - desculpar - idealizar - exemplificar - ressaltar - desabafar - aludir - especular - enunciar - clarificar - opinar - pontuar - censurar - antever - enumerar - detonar - apregoar - intimar - sentenciar - arrematar - discorrer - aferir - alardear - recitar - pormenorizar - despistar - confidenciar - resignar - teorizar - deixar escapar - asseverar - profetizar - filosofar - realçar - repreender - rememorar - suplicar - resignar-se - esmiuçar - afiançar

PERGUNTA

colocar - perguntar - convidar - solicitar - convocar - questionar - querer saber - interrogar - indagar - sondar - suplicar

RESPOSTA

defender - garantir - explicar - assumir - responder - confirmar - entregar - negar - justificar - assegurar - recusar - alegar - adiantar - esclarecer - reagir - constatar - contestar - confessar - argumentar - inventar - devolver - soltar - agradecer - especificar - contrapor - rebater - sintetizar - especular - clarificar - replicar - retrucar - deixar escapar - ripostar - retorquir - alfinetar - anuir

CONCORDÂNCIA & DISCORDÂNCIA

considerar - manter - defender - provocar - acusar - assumir - confirmar - reconhecer - admitir - preferir - apoiar - negar - justificar - criticar - avaliar - analisar - conceder - recusar - alegar - atacar - observar - concordar - reclamar - sustentar - convencer - cortar - denunciar - situar - reagir - rejeitar - contestar - confessar - argumentar - disparar - debater - protestar - lamentar - atirar - devolver - combinar - ameaçar - elogiar - desmentir - discordar - desafiar - queixar+se - desconfiar - estranhar - advertir - queixar - advogar - minimizar - contrapor - ironizar - rebater - insinuar - preconizar - ressalvar - desabafar - enunciar - censurar - detonar - teimar - apregoar - xingar - aferir - resignar-se - retrucar - retorquir - repreender - alfinetar - anuir

EMOÇÕES & EXCLAMAÇÕES

continuar - manter - tornar - começar - concluir - acrescentar - retomar - completar - cortar - prosseguir - complementar - crescer - rematar - reiterar - finalizar - emendar - arrematar - provocar - lembrar - assumir - discutir - entregar - cantar - atacar - temer - reagir - comemorar - contestar - interromper - gritar - disparar - animar - protestar - atirar - rir - ameaçar - confiar - divertir(-se) - chorar - desmentir - incentivar - homenagear - admirar(-se) - desafiar - queixar(-se) - desconfiar - rezar - sorrir - saudar - festejar - espantar-se - cumprimentar - entusiasmar-se - rebater - relembrar - preconizar - desculpar - indignar-se - desabafar - censurar - orgulhar(-se) - detonar - clamar - empolgar(-se) - brindar - exclamar - berrar - xingar - alardear - confidenciar - gabar-se - alegrar(-se) - bradar - murmurar - resmungar - exultar - repreender - rememorar - gargalhar - regozijar - gracejar - gabaritar

CONTINUIDADE & CONCLUSÃO

continuar - manter - tornar - começar - concluir - acrescentar - retomar - completar - cortar - prosseguir - complementar - crescer - rematar - reiterar - finalizar - emendar - arrematar

DESTAQUE

mostrar - apontar - repetir - observar - sublinhar - insistir - destacar - notar - salientar - assinalar - acentuar - ressaltar - frisar - enfatizar - ressalvar - realçar

HESITAÇÃO & PONDERAÇÃO

considerar - apurar - avaliar - analisar - calcular - refletir - planejar - desconfiar - hesitar - recear - ponderar - deliberar - amenizar - sondar - aferir - despistar - teorizar - filosofar - balbuciar

PALPITES E CERTEZAS

prever - garantir - acreditar - confirmar - prometer - assegurar - demonstrar - julgar
 - imaginar - comparar - supor - apostar - estimar - crer - calcular - arriscar - confiar
 - desmentir - suspeitar - desconfiar - torcer - prevenir - chutar - sinalizar - jurar -
 atestar - insinuar - especular - antever - presumir - sondar - teorizar - asseverar -
 profetizar - vaticinar - afiançar

ORDEM, PEDIDO & SUGESTÃO

considerar - pedir - decidir - propor - exigir - determinar - ajudar - estabelecer -
 indicar - impor - mandar - sugerir - proibir - observar - convidar - convencer -
 solicitar - convocar - comandar - recomendar - alertar - apelar - combinar - decretar
 - ameaçar - ordenar - incentivar - aconselhar - ditar - reforçar - advertir - pregar -
 estipular - prevenir - contabilizar - deliberar - idealizar - clamar - apregoar - intimar
 - sentenciar - suplicar

OUTROS

chamar - lembrar - convidar - comparar - ensinar - inventar - resumir - brincar -
 divertir (-se) - soltar - pronunciar - desafiar - rezer - narrar - chutar - mentir -
 contabilizar - simplificar - relembrar - sintetizar - idealizar - ressalvar - recitar -
 murmurar - rememorar - gracejar