UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, LETRAS E CIÊNCIAS HUMANAS
DEPARTAMENTO DE LETRAS MODERNAS
PROGRAMA DE PÓS-GRADUAÇÃO EM LÍNGUA, LITERATURA
E CULTURA ITALIANAS

RENATO RAILO RIBEIRO

Atos de fala em dicionários híbridos italiano>português-brasileiro:
sugestão para dicionarização de ilocuções via *corpora*

Versão Corrigida

São Paulo
2015

RENATO RAILO RIBEIRO

**Atos de fala em dicionários híbridos italiano>português-brasileiro:
sugestão para dicionarização de ilocuções via *corpora***

Versão Corrigida

Dissertação apresentada ao Programa de Pós-Graduação em Língua, Literatura e Cultura Italianas do Departamento de Letras Modernas da Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, para obtenção do título de Mestre em Letras.

Linha de pesquisa: o italiano falado e escrito sob a perspectiva brasileira.

Orientadora: Prof^a. Dr^a. Angela Maria Tenório Zucchi.

De acordo: _____

Prof^a. Dr^a. Angela Maria Tenório Zucchi

São Paulo

2015

Nome: RIBEIRO, Renato Railo

Título: Atos de fala em dicionários híbridos italiano>português-brasileiro: sugestão para dicionarização de ilocuções via *corpora*

Dissertação apresentada ao Programa de Pós-Graduação em Língua, Literatura e Cultura Italianas do Departamento de Letras Modernas da Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, para obtenção do título de Mestre em Letras.

Aprovado em:

Banca Examinadora

Prof. Dr.: _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Prof. Dr.: _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Prof. Dr.: _____ Instituição: _____

Julgamento: _____ Assinatura: _____

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio impresso ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

renatorailo@yahoo.it

<http://minhateca.com.br/Railo-Ribeiro>

Catálogo na Publicação
Serviço de Biblioteca e Documentação
Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo

R484a Ribeiro, Renato Railo
 Atos de fala em dicionários híbridos
 italiano>português-brasileiro: sugestão para
 dicionarização de ilocuções via corpora / Renato Railo
 Ribeiro ; orientadora Angela Maria Tenório Zucchi. -
 São Paulo, 2015.
 144 f.

Dissertação (Mestrado)- Programa de Pós-Graduação
em Língua, Literatura e Cultura Italianas. Área
de concentração: O italiano falado e escrito no Brasil.

1. Lexicografia. 2. Pragmática Linguística. 3.
Linguística de Corpus. 4. Dicionários híbridos. 5.
Atos de fala. I. Zucchi, Angela Maria Tenório,
orient. II. Título.

Sapere aude!¹

Horácio, *Epistulae*, I, 2, 40 (aprox. 20 a. C.),
citado por Immanuel Kant em
Beantwortung der Frage: Was ist Aufklärung? (1784).

¹ Ousa saber!

AGRADECIMENTOS

À **Profª. Drª. Angela Maria Tenório Zucchi**, pela oportunidade que me concedeu ao me aceitar como seu orientando, bem como por tudo o que esta ação lhe implicou ao longo dos últimos três anos: leituras, sugestões, críticas, revisões, encontros, empréstimo de livros, troca de emails, os trâmites para obtenção da bolsa de estudos etc. E eu lhe agradeço também pela autonomia que me concedeu, pois mesmo possivelmente discordando de alguns pontos do trabalho em momento algum direcionou minha pesquisa ao encontro de suas próprias concepções;

À **Profª. Drª. Elisabetta Santoro**, por ter participado da banca examinadora do relatório de qualificação e da comissão examinadora da dissertação, bem como pelas sugestões, críticas, correções, apontamentos e livros emprestados ao longo dos dois últimos anos;

À **Profª. Drª Vivian Orsi**, por ter participado da comissão examinadora da dissertação, bem como pelas críticas, sugestões e demais observações de grande valia para o aperfeiçoamento do projeto como um todo;

À **Profª. Drª. Stella Tagnin**, por ter participado da banca examinadora do relatório de qualificação, bem como pelas sugestões e críticas que indiscutivelmente contribuíram para o aprimoramento do trabalho;

À **Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)**, fundação do Ministério da Educação (MEC) cuja função é possibilitar a expansão e consolidação da pós-graduação *stricto sensu* (mestrado e doutorado) em todos os estados da Federação. Como pesquisador, agradeço pela concessão da bolsa de pesquisa que tornou possível este trabalho, e, como cidadão, agradeço pelas bolsas e demais incentivos concedidos aos inúmeros pesquisadores espalhados pelo país, que através de suas pesquisas buscam, de uma forma ou de outra, contribuir para a educação e a ciência brasileiras;

Ao **Programa de Pós-Graduação em Língua, Literatura e Cultura Italianas** do Departamento de Letras Modernas da Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, pela confiança depositada no trabalho;

Aos **professores doutores dos departamentos de Letras Modernas, de Linguística e de Filosofia** da Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo com os quais tive a oportunidade de estudar – seja como aluno regularmente matriculado, seja como aluno-ouvinte;

À **Profª. Drª. Maria Aparecida Barbosa**, pelas sugestões e obras gentilmente fornecidas;

Aos **funcionários do Serviço de Pós-Graduação** da Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, pelo imprescindível trabalho;

À **Profª. Ms. Pamela Andrade**, pelas valiosas conversas acerca da teoria dos atos de fala de Austin;

À **tradutora e pesquisadora Ana Mhereb** (ana.mhereb@usp.br), por ter elaborado o meu *abstract* mesmo tendo sido solicitada em cima da hora, o que apenas reforça sua gentileza e competência costumeiras;

À **Prof^a. Ms. Stela Danna**, pelas leituras e sugestões gentilmente fornecidas ao longo da primeira metade do trabalho, bem como pelo apoio que me deu no momento da qualificação – motivos pelos quais lhe serei eternamente grato;

À **professora e pesquisadora Kamunjin Tanguale**, colega de Pragmática e de Filosofia da Linguagem que muito me ajudou ao me sugerir alguns caminhos antes ignorados por mim;

À **bibliotecária Maria Paula Galdino Miyashiro**, pelos quase dez anos de amizade, resistente a qualquer uma das possíveis incompatibilidades conjugais que possamos ter enfrentado no passado. *“De todos os meios pelos quais a sabedoria busca a felicidade, o mais importante é a amizade. A noção que nos encoraja a acreditar que nada de temível é eterno, tampouco dura muito tempo, é a mesma que nos garante que a segurança, nestas mesmas condições limitadas, é obtida, sobretudo, por meio da amizade”* (Epicuro, *Máximas Capitais*, 148, XXVII-XXVIII).

À **minha mãe, Solange**, e ao **meu irmão, Cézar**, pelos valiosos incentivos, sem os quais nada seria possível.

RESUMO

RIBEIRO, Renato Railo. Atos de fala em dicionários híbridos italiano>português-brasileiro: sugestão para dicionarização de ilocuções via *corpora*. Dissertação (Mestrado). Orient. Prof^ª. Dr^ª. Angela Maria Tenório Zucchi. Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo (FFLCH/ USP), São Paulo, 2015.

O objetivo do estudo foi sugerir a inserção, em dicionários híbridos italiano>português-brasileiro tais como o *Parola Chiave* (2012), de informações acerca da dimensão pragmático-ilocucionária de ambas as línguas, a partir do uso de *corpora* eletrônicos. A metodologia empregada foi a seguinte: (a) fundamentação teórica, baseada: no conceito de Lexicografia de Krieger (2008) e no de dicionário híbrido de Höfling, Silva e Tosqui (2004); no de Pragmática Linguística e no de atos de fala de Bianchi (2008) e Sbisà (2009); no de Linguística de *Corpus* de Berber-Sardinha (2000), (2003) e (2004) e Tagnin (2004) e (2013); (b) estabelecimento de critérios para a investigação de ilocuções em *corpora* italiano e português-brasileiro, a saber: adoção da teoria dos atos de fala proposta por Austin (1990[1975]) (em função de seu critério para reconhecimento de ilocuções, a forma verbal na primeira pessoa do singular do presente do indicativo na voz ativa, e de sua associação ilocuções-verbos); adoção de dois *corpora online*, *Corpus Paisà* e *Corpus do Português* (o primeiro do italiano, o segundo do português-brasileiro, escolhidos em função da extensão de cada um, o que os torna representativos de suas respectivas línguas); adoção de cinco verbos ilocucionários comissivos do italiano, *promettere*, *giurare*, *assicurare*, *impegnarsi*, *garantire* (em função da mútua relação sinonímica que possuem entre si, segundo o dicionário *Sinonimi e contrari minore* (2009)); adoção de cinco verbos ilocucionários comissivos do português-brasileiro, *prometer*, *jurar*, *assegurar*, *comprometer-se*, *garantir* (em função de serem semanticamente equivalentes aos respectivos verbos italianos, segundo o *Parola Chiave*); adoção da forma verbal citada acima como sintaxe de busca para pesquisa em *corpora*; (c) análise qualitativa e quantitativa: obtiveram-se os números totais de ocorrências de cada ilocução; realizou-se uma análise qualitativa de modo a excluir casos anômalos, cujos critérios de exclusão foram: casos de homonímia, negação, repetição e ininteligibilidade; após obtenção dos números reais, excluíram-se da pesquisa as ilocuções *mi impegno* e *comprometo-me/ me comprometo* (em função da baixa frequência); (d) discussão acerca das possibilidades de dicionarização de ilocuções a partir dos resultados de *corpora*. Como resultado final, a sugestão foi a de inserir: na microestrutura, marcas de uso referentes às classes de ilocução; na micro-medioestrutura, remissivas de modo a conduzir o leitor a um texto externo à nomenclatura; na macroestrutura, um texto externo à nomenclatura que contenha: (i) explicações referentes às classes de ilocuções e lista de respectivas espécies convencionalmente recorrentes do italiano dispostas segundo frequência; (ii) espécies de ilocuções equivalentes do português-brasileiro dispostas segundo frequência; (iii) exemplos de uso, retirados dos *corpora*, de tais verbos ilocucionários desempenhando sua função ilocucionária convencional.

Palavras-chave: Lexicografia; Dicionário Híbrido; Atos de Fala; Pragmática; Linguística de *Corpus*.

RIASSUNTO

RIBEIRO, Renato Railo. Atti linguistici in dizionari ibridi italiano>portoghese-brasiliano: proposta di registro lessicografico di verbi illocutori tramite l'analisi dei corpora. Tesi (Master). Orient. Prof.ssa Dott.ssa Angela Maria Tenório Zucchi. Facoltà di Filosofia, Lettere e Scienze Umane dell'Università di Sao Paulo (FFLCH/ USP). Sao Paulo, 2015, 130 pp.

Lo scopo di questo lavoro è fornire una proposta di indicazione delle illocuzioni all'interno di dizionari ibridi italiano > portoghese-brasiliano, tramite l'analisi di *corpora* testuali. La metodologia utilizzata per questo lavoro integra: (a) per la base teorica: il concetto di Lessicografia di Krieger (2008) e di dizionario ibrido di Höfling, Silva e Tosqui (2004); i concetti di Pragmatica e atti linguistici di Bianchi (2008) e Sbisa (2009); i concetti di Linguistica dei *corpora* di Berber-Sardinha (2000), (2003) e (2004) e Tagnin (2004) e (2013); (b) i criteri di ricerca di illocuzioni nei *corpora*: l'adozione della teoria degli atti linguistici sviluppata da Austin (1990 [1975]) (in ragione del suo criterio di riconoscimento delle illocuzioni, identificato con la forma verbale alla prima persona singolare del presente indicativo attivo e della corrispondenza tra i verbi e le illocuzioni); l'adozione di due *corpora online* di ampia dimensione (e perciò rappresentativi) delle rispettive lingue (il corpus di italiano *Paisà* e il corpus di portoghese-brasiliano *Corpus do Português*); la scelta di cinque verbi commissivi italiani (*promettere, giurare, assicurare, impegnarsi, garantire*, in ragione della relazione sinonimica tra loro (secondo il dizionario *Sinonimi e contrari minore* (2009)); e di cinque verbi illocutivi del portoghese-brasiliano (*prometer, jurar, assegurar, comprometer-se, garantir*, semanticamente equivalenti ai verbi italiani sopracitati (secondo il dizionario *Parola Chiave* (2009)); l'adozione della forma verbale sopracitata per la ricerca dei verbi nei corpora; (c) l'analisi qualitativa/ quantitativa: è stata realizzata la ricerca dei verbi nei *corpora* e se ne è ottenuto il numero totale per ogni verbo selezionato; è stata poi effettuata un'analisi qualitativa al fine di escludere i casi anomali (casi di omonimia, negazione, ripetizione e incomprensibilità); si è così ottenuto il numero reale di occorrenze per ogni verbo e sono state poi escluse le illocuzioni *mi impegno* e *comprometo-me/ me comprometo* (a causa della loro bassa frequenza); (d) la proposta di registro lessicografico delle illocuzioni all'interno di dizionari ibridi può essere così descritta: inserire all'interno della microstruttura marche d'uso sulle classi di illocuzione; inserire all'interno della microstruttura l'indicazione di riferimenti incrociati che permettano al lettore di accedere ad un'altra parte del dizionario (vale a dire, a una sezione esterna al lemmario); inserire all'interno della macrostruttura una sezione che: (i) spieghi che cos'è l'atto linguistico e esponga quali sono le classi ed i tipi più frequenti in italiano (ii) contenga le stesse informazioni sugli atti linguistici corrispondenti in portoghese-brasiliano; (iii) faccia riferimento a esempi d'uso estratti dai *corpora*.

Parola-chiave: Lessicografia; Dizionario Ibrido; Atti Linguistici; Pragmatica; Linguistica dei Corpora.

ABSTRACT

RIBEIRO, Renato Railo. Speech acts in Italian>Brazilian-Portuguese Hybrid Dictionaries: a suggestion for lexicographical records of illocutionary acts through Corpus Linguistics. Master's Thesis. Orient. Prof^a. Dr^a. Angela Maria Tenório Zucchi. Faculty of Philosophy, Languages and Human Sciences at the University of Sao Paulo (FFLCH/ USP). Sao Paulo, 2015.

This study aims to suggest a method of inserting, in hybrid Italian>Brazilian-Portuguese dictionaries such as *Parola Chiave* (2012), information about the pragmatic-illocutionary dimension of both languages, through electronic *corpora*. The methodology used was as follows: (a) theoretical foundation based on: Krieger's concept of Lexicography (2008) and Höfling, Silva & Tosqui's hybrid dictionary (2004); Bianchi (2008) and Sbisà's (2009) Pragmalinguistics and speech acts; Berber-Sardinha (2000, 2003, 2004) and Tagnin's (2004, 2013) Corpus Linguistics; (b) establishing criteria for illocution research in corpora, namely: adopting the theory of speech acts proposed by Austin (1990[1995]) as research paradigm (due to its criteria for recognizing illocutions, the verb form in the first-person singular of the present indicative in active voice and its illocution-verbs association); adopting two online corpora, *Corpus Paisà* and *Portuguese Corpus* (due to the length of each one, which makes them representative of their languages); adopting five Italian illocutionary verbs of commissive class, *promettere*, *giurare*, *assicurare*, *impegnarsi*, *garantire* (due to the mutual synonymy relation they have one to each other, according to *Sinonimi e contrari minore* (2009)); adopting five illocutionary verbs of Brazilian-Portuguese, as known: *prometer*, *jurar*, *assegurar*, *comprometer-se*, *garantir* (for being semantically equivalent to their corresponding verb in Italian, according to *Parola Chiave* (2009)); adopting the verb form mentioned above as syntax search to *corpora* research; (c) quantitative and qualitative analysis: the total number of occurrences of each illocution was obtained; a qualitative analysis was conducted in order to exclude anomalous cases, of which exclusion criteria were: cases of homonyms, denial, occurrences repetition and unintelligibility; a quantitative analysis was conducted in order to exclude the illocutions *mi impegno* and *comprometo-me / me comprometo* (due to their low frequency); (d) discussion around the possibilities of lexicographical records of illocutions from the results of *corpora*. As a final result, the suggestion was to insert: in the dictionary microstructure, signs of usage referring to illocution classes; in its micro-mediumstructure: cross references in order to conduct the reader to a section out of nomenclature; in its macrostructure: a text, external to the nomenclature, containing: (i) explanations related to illocution classes and a list of conventionally recurring species of Italian arranged by frequency; (ii) equivalent illocutions species of Portuguese-Brazilian arranged by frequency; (iii) examples of usage, taken from the *corpora*, of illocutionary verbs performing their conventional illocutionary function.

Key-words: Lexicography; Hybrid Dictionary; Speech Acts; Pragmatics; Corpus Linguistics.

Lista de abreviações

HTD – *How to Do Things with Words* (obra de J. L. Austin)

LC – Linguística de *Corpus*

PC – *Parola Chiave* (dicionário)

TAF – Teoria dos atos de fala

Lista de figuras

Figura 1: lista de marcas de uso pragmáticas fornecidas pelo *Salamanca* (1996: xi) em sua Introdução (Introdução, I, p. 6).

Figura 2: marcas de uso pragmático-ilocucionárias contidas no *Salamanca* (1996: xi) e listados por Guerra Salas e Gómez Sánchez (2005: 354) (Introdução, I, p. 6).

Figura 3: sugestão de notas de uso feita por Duran (2004: 110-111) (1.5.3., p. 30).

Figura 4: lista de símbolos gráficos do PC (2009) (1.8.1., p. 37).

Figura 5: página inicial do *Corpus Paisà*, (3.5., p. 76).

Figura 6: página inicial do *Corpus do Português*, (3.5., p. 77).

Figura 7: integração das três estruturas, (5.4., p. 106).

Figura 8: nº de ocorrências totais da ilocução “assicuro” no *Corpus Paisà* (Anexo 1, p. 120).

Figura 9: nº de ocorrências totais da ilocução “giuro” no *Corpus Paisà* (Anexo 1, p. 120).

Figura 10: nº de ocorrências totais da ilocução “garantisco” no *Corpus Paisà* (Anexo 1, p. 121).

Figura 11: nº de ocorrências totais da ilocução “prometto” no *Corpus Paisà* (Anexo 1, p. 121).

Figura 12: nº de ocorrências totais da ilocução “mi impegno” no *Corpus Paisà* (Anexo 1, p. 122).

Figura 13: nº de ocorrências totais da ilocução “juro” no *Corpus do Português* (Anexo 1, p. 122).

Figura 14: nº de ocorrências totais da ilocução “garanto” no *Corpus do Português* (Anexo 1, p. 123).

Figura 15: nº de ocorrências totais da ilocução “prometo” no *Corpus do Português* (Anexo 1, p. 123).

Figura 16: nº de ocorrências totais da ilocução “asseguro” no *Corpus do Português* (Anexo 1, p. 124).

Figura 17: nº de ocorrências totais da ilocução “comprometo” no *Corpus do Português* (Anexo 1, p. 124).

Figura 18: página principal da mídia Folha Online. Disponível em: <http://www.folha.uol.com.br/>. Acesso: 02.09.2013 (Anexo 2, p. 127).

Figura 19: página principal da mídia Folha Online contendo categorias de assuntos, dentre as quais ‘Política’. Disponível em : <http://www.folha.uol.com.br/>. Acesso: 02.09.2013 (Anexo 2, p. 127).

Figura 20: página principal da categoria ‘Política’ da mídia Folha Online, com matérias passíveis de serem coletadas. Disponível em: <http://www1.folha.uol.com.br/poder/>. Acesso: 02.09.2013 (Anexo 2, p. 128).

Figura 21: matéria contida na categoria ‘Política’ da mídia Folha Online. Disponível em: <http://www1.folha.uol.com.br/poder/2013/09/1335461-dilma-aumenta-a-seguranca-no-7-de-setembro-por-temer-violencia.shtml/>. Acesso: 02.09.2013 (Anexo 2, p. 128).

Figura 22: cópia e colagem de uma matéria contida na categoria ‘Política’ da mídia Folha Online para o software ‘Bloco de Notas’ (Anexo 2, p. 129).

Figura 23: comentários de uma matéria contida na categoria ‘Política’ da mídia Folha Online. Disponível em: <http://www1.folha.uol.com.br/poder/2013/09/1335461-dilma-aumenta-a-seguranca-no-7-de-setembro-por-temer-violencia.shtml/>. Acesso: 02.09.2013 (Anexo 2, p. 130).

Figura 24: cópia e colagem de comentários de uma matéria contida na categoria ‘Política’ da mídia Folha Online para o software ‘Bloco de Notas’ (Anexo 2, p. 131).

Figura 25: dados que não interessam à pesquisa e que foram apagados (Anexo 2, p. 131).

Figura 26: dados que interessam à pesquisa e que serão mantidos (Anexo 2, p. 132).

Figura 27: arquivos do ‘Bloco de Notas’, um contendo as matérias e outro os comentários das matérias (Anexo 2, p. 132).

Lista de tabelas

Tabela 1: classes de força ilocucionária e algumas de suas espécies, tais como propostas por Austin (1990[1975]: 123-130) (2.3.2., p. 59-60).

Tabela 2: algumas das ilocuições comissivas propostas por Austin (1990[1975]: 127) (4.1., p. 79-80).

Tabela 3: verbos ilocucionários comissivos propostos por Sbisà (2009[1989]: 71) (4.1., p. 80).

Tabela 4: relação sinonímica entre verbos ilocucionários comissivos italianos pesquisados (4.1., p. 81).

Tabela 5: verbos ilocucionários comissivos italianos e respectivos equivalentes semânticos do português-brasileiro dados pelo PC (2012) (4.1., p. 82).

Tabela 6: nº de ocorrências totais de ilocuições comissivas pesquisadas no *Corpus Paisà* (4.2.1., p. 82-83).

Tabela 7: exemplos de casos excluídos referentes à ilocução comissiva “assicuro” (4.2.1., p. 83).

Tabela 8: exemplo de caso excluído referente à ilocução comissiva “garantisco” (4.2.1., p. 84).

Tabela 9: exemplo de caso excluído referente à ilocução comissiva “prometto” (4.2.1., p. 84).

Tabela 10: exemplo de caso excluído referente à ilocução comissiva “garantisco” (4.2.1., p. 84).

Tabela 11: exemplo de caso excluído referente à ilocução comissiva “giuro” (4.2.1., p. 84).

Tabela 12: exemplo de caso excluído referente à ilocução comissiva “mi impegno” (4.2.1., p. 84).

Tabela 13: nº de ocorrências reais de ilocuições comissivas pesquisadas no *Corpus Paisà* (4.2.1., p. 85).

Tabela 14: comparação entre o nº de ocorrências totais e reais de ilocuições comissivas pesquisadas no *Corpus Paisà* (4.2.1., p. 85).

Tabela 15: nº de ocorrências totais de ilocuições comissivas pesquisadas no *Corpus do Português* (4.2.2., p. 86).

Tabela 16: nº de ocorrências reais de ilocuições comissivas pesquisadas no *Corpus do Português* (4.2.2., p. 86-87).

Tabela 17: exemplos de casos excluídos referentes à ilocução comissiva “juro” (4.2.2., p. 87).

Tabela 18: exemplo de caso excluído referente à ilocução comissiva “prometo” (4.2.2., p. 87).

Tabela 19: exemplo de caso excluído referente à ilocução comissiva “comprometo” (4.2.2., p. 87).

Tabela 20: comparação entre o nº de ocorrências totais e reais de ilocuições comissivas pesquisadas no *Corpus do Português* (4.2.2., p. 88).

Tabela 21: comparação entre o nº de ocorrências reais de ilocuições comissivas pesquisadas no *Corpus Paisà* e no *Corpus do Português* (4.2.2., p. 88).

Tabela 22: os seis sites que fornecem notícias jornalísticas e opção para comentários mais acessados do Brasil. Fonte: <http://www.alexa.com/topsites/countries/BR>. Acesso em: 07.01.13 (Anexo 2, p. 126).

Tabela 23: os seis sites que fornecem notícias jornalísticas e opção para comentários mais acessados da Itália. Fonte: <http://www.alexa.com/topsites/countries/IT>. Acesso em: 07.01.13 (Anexo 2, p. 126).

Tabela 24: categorização dos temas segundo os doze sites pesquisados (Anexo 2, p. 136).

Tabela 25: nº de matérias por assunto/ idioma (Anexo 2, p. 138).

Tabela 26: nº de matérias por assunto/ idioma do italiano (Anexo 2, p. 138-139).

Tabela 27: nº de comentários por assunto/ idioma (Anexo 2, p. 139).

Tabela 28: nº de *tokens/ types* por assunto, segundo *corpus* do português-brasileiro (Anexo 2, p. 140).

Tabela 29: nº de *tokens/ types* por assunto, segundo *corpus* do italiano (Anexo 2, p. 140).

Tabela 30: nº de *tokens/ types* dos *corpora* do português-brasileiro e do italiano (Anexo 2, p. 140).

Tabela 31: nº de *tokens/ types* por mídia do *corpus* português-brasileiro (Anexo 2, p. 141).

Tabela 32: nº de matérias por assunto do *corpus* português-brasileiro (Anexo 2, p. 141).

Tabela 33: nº de comentários por assunto do *corpus* português-brasileiro (Anexo 2, p. 142).

Tabela 34: nº de *tokens/ types* por assunto, segundo *corpus* do português-brasileiro (Anexo 2, p. 142).

Tabela 35: nº de *tokens/ types* dos *corpora* do português-brasileiro e do italiano (Anexo 2, p. 142).

SUMÁRIO

INTRODUÇÃO	4
I. Contextualização da pesquisa.....	4
II. Objetivos	8
III. Justificativas.....	9
IV. Metodologia.....	9
V. Estrutura da dissertação	14
1. LEXICOGRAFIA.....	17
1.1. A ciência dos dicionários	17
1.2. Lexicografia prática e Metalexicografia	18
1.3. Léxico e dicionário geral de língua.....	20
1.4. Macroestrutura e nomenclatura	23
1.4.1. Textos externos.....	24
1.5. Microestrutura e verbetes.....	25
1.5.1. Marcas de uso	26
1.5.2. Exemplos de uso.....	29
1.5.3. Notas de uso	29
1.6. Medioestrutura ou Remissões.....	30
1.7. Dicionários gerais de língua: monolíngues, bilíngues e híbridos	31
1.8. O Parola Chiave (PC)	36
1.8.1. Macroestrutura do PC.....	36
1.8.2. Microestrutura do PC	38
1.8.3. Medioestrutura do PC.....	38
2. TEORIA DOS ATOS DE FALA (TAF)	40
2.1. O agir pela linguagem.....	40
2.2. A TAF como objeto de estudo da Pragmática Linguística	42
2.2.1 A TAF e o conceito de contexto.....	43
2.2.2. A TAF e os conceitos de comunicação e interação	44
2.2.3 A TAF e o conceito de competência comunicativa.....	49
2.3. A TAF de Austin.....	51
2.3.1. Performativos e constativos.....	52
2.3.2. Dizer é fazer	55
2.4. A TAF de Austin como paradigma de pesquisa	60

2.4.1. A classe dos comissivos	64
3. LINGÜÍSTICA DE CORPUS (LC).....	66
3.1. A LC como metodologia.....	66
3.2. Corpus e abordagens de investigação	68
3.3. Principais ferramentas oferecidas pela LC	70
3.4. Padrão, extensão e representatividade	71
3.5. Delimitação dos corpora: Corpus Paisà e Corpus do Português.....	74
4. ANÁLISE DE CORPORA.....	79
4.1. Critérios para busca de ilocuções via LC.....	79
4.2. Procedimentos de busca.....	82
4.2.1. Verbos ilocucionários comissivos do italiano	82
4.2.2. Verbos ilocucionários comissivos do português-brasileiro	85
5. SUGESTÃO DE INCLUSÃO DE ILOCUÇÕES EM DICIONÁRIOS	
HÍBRIDOS	90
5.1. Microestrutura.....	90
5.1.1. Marcas de uso	91
5.1.2. Notas de uso	94
5.1.3. Exemplos de uso.....	95
5.2. Medioestrutura	96
5.2.1. Remissivas entre verbetes.....	97
5.2.2. Remissivas entre verbetes e textos externos.....	98
5.3. Macroestrutura	99
5.3.1. Definições de ilocuções e lista exaustiva de suas espécies	100
5.3.2. Definições de ilocuções e algumas espécies seguidas de exemplos.....	101
5.3.3. Definições de ilocuções e algumas espécies do italiano e do português-brasileiro, seguidas de exemplos	101
5.3.4. Definições de ilocuções e algumas espécies do italiano e do português-brasileiro, seguidas de exemplos e ordenadas por frequência.....	102
5.4. Integração entre as três estruturas	103
CONSIDERAÇÕES FINAIS.....	107
REFERÊNCIAS	112
ANEXOS	120
Anexo 1. Figuras correspondentes às análises de corpora.....	120

Anexo 2. Proposta de construção de corpora apresentada no relatório de qualificação	
.....	125

INTRODUÇÃO

I. Contextualização da pesquisa

Desde os protótipos tabulares¹ até as mais recentes e recursivas versões disponibilizadas na internet², os dicionários vêm se apresentando como obras indispensáveis de preservação e difusão de um aspecto particular da cultura: a língua. Longe de se constituírem como meras listas de palavras, os produtos lexicográficos são concebidos de modo a fornecer diversas informações linguísticas (e até mesmo extralinguísticas), organizados de várias maneiras e destinados aos mais diversos fins e públicos. Parte dessa variedade pode ser observada na classificação de obras lexicográficas propostas por Bugueño Miranda (2014), composta por uma infinidade de obras – estas que, seguindo as palavras de Barros (2004: 133), “procura[m] atender às necessidades de uma sociedade heterogênea, marcada pela multiplicidade de discursos”.

Um dos elementos componentes da multiplicidade discursiva é a variação lexical, evidenciada tanto no interior de uma mesma língua como – e, sobretudo – na comparação entre línguas diversas. Segundo Silva (2006: 16), “o sistema lexical de uma língua não corresponde exatamente ao sistema de outra, pois a cultura de uma comunidade linguística influi diretamente na constituição de seu léxico”. Assim, torna-se imprescindível que, para facilitar a compreensão mútua entre indivíduos pertencentes a comunidades linguísticas diversas, e, conseqüentemente, favorecer a comunicação humana e a integração entre culturas e sociedades, ações e instrumentos de natureza linguístico-pedagógica sejam concebidos como pontos de contato entre as variadas práticas discursivo-comunicativas. Van Hoof (1998: 243) inclui os dicionários entre os instrumentos de tal natureza ao dizer que “o advento da educação compulsória gerou demanda por material pedagógico de todos os tipos, e em particular os dicionários”. O mesmo faz Seide (2006), para quem a função de todo dicionário é educativa e seu papel é contribuir para “o aperfeiçoamento das habilidades linguísticas dos consulentes”. Por

¹ Barros (2004: 29) aponta que a existência de dicionários temáticos monolíngues – em forma de tijolos de argila – é atestada desde 2600 a.C., entre os sumérios; Welker (2004: 56 e 61) também faz menção às listas lexicais produzidas em tábuas pelos sumérios, mas sugere a data aproximada de 3300 a.C. Segundo observações de Biderman (1984: 1) e (2001a: 17), a Lexicografia tal como se conhece hoje só se inicia na Idade Moderna, razão pela qual Welker, ao citar Boisson et al (2004: 56) nomeou aquelas primeiras experiências de “paleolexicografia”. BOISSON, C.; KIRTCHUK, P.; BÉJOINT, H. Aux origines de la lexicographie: les premiers dictionnaires monolingues et bilingues. *International Journal of Lexicography* 4.4 : 261-315.

² Como exemplo, veja-se o *Thinkmap Visual Thesaurus* (<https://www.visualthesaurus.com/>).

sua vez, Krieger (2007: 301, grifo nosso) ressalta o papel dos dicionários no processo de aprendizagem e comunicação:

De forma indiscutível, os dicionários de língua são instrumentos potenciais para o aprendizado e desenvolvimento da leitura, da redação e da comunicação em geral. Os que seguem o padrão prototípico podem também contribuir para o conhecimento descritivo da língua, em razão do conjunto de informações que os verbetes costumam oferecer. Junto a isso, eles permitem saber da existência de alguma palavra, de como é escrita ou pronunciada, de seus sentidos primeiros e os conotados. Além dos conhecimentos que favorecem o desenvolvimento da *competência comunicativa do sujeito e do saber sobre o funcionamento da língua*, a consulta lexicográfica pode contribuir para o processo de alfabetização.

Um tipo de obra lexicográfica dessa natureza é o dicionário híbrido, que Welker (2004: 202), citando Marelló³, entende como aquele que está entre os bilíngues e os monolíngues. Por outras palavras, os híbridos são dicionários de língua que: (a) possuem por base um dicionário monolíngue de uma dada língua – que apresenta, como resume Borba (2003: 16), “(...) a estrutura e funcionamento da língua, se possível num sistema bem nítido de notação”, a partir de informações de natureza fonológica, sintática, semântica etc; (b) fornecem equivalências (em geral, de natureza semântica) em outra língua – tal como faz um dicionário bilíngue. Exemplo concreto de dicionário híbrido é o *Parola Chiave* (2012) (doravante, PC), composto a partir de um dicionário monolíngue da língua italiana *mais* a equivalência semântica em português-brasileiro, esta última fornecida por uma equipe de lexicógrafos/ tradutores brasileiros.

Uma característica que torna os dicionários de língua geral (como os híbridos) úteis para o processo de aquisição de competência comunicativa é a inclusão, em seu interior, de informações acerca da *dimensão pragmática* da língua – i. e., informações léxico-linguísticas acerca da língua em uso interacional e/ ou comunicativo. Segundo Guerra Salas e Gómez Sánchez (2005: 353), se em determinado momento fora paradoxal imaginar informações de cunho pragmático em dicionários, uma vez que, em tese, “(...) o dicionário conserva o sistema (a língua, o significado), enquanto que a pragmática estuda a linguagem em uso (a fala, o sentido)”⁴, contemporaneamente tal paradoxo inexistente. Como exemplo, os mencionados autores examinaram o *Diccionario Salamanca de la Lengua Española* (1996) em razão de este fornecer explicitamente informações de natureza pragmática – entre as quais aquelas relativas aos *atos de fala*.

³ MARELLÓ, C. Les différents types de dictionnaires bilingües. In : Béjoint, H. ; Thoiron, P. *Les dictionnaires bilingües*. Louvain-la-Neuve: Duculot, 1996, p. 31-52.

⁴ “El diccionario recoge el sistema (la lengua, el significado) mientras que la pragmática estudia el lenguaje en uso (el habla, el sentido)”.

Segundo os pesquisadores citados (2005: 355), o que caracteriza a teoria dos atos de fala (doravante TAF) é a possibilidade de representar, por meio de categorias, aquilo que um falante tenciona “(...) ao proferir determinado enunciado”⁵ – melhor dizendo, aquilo que um falante *faz* ao dizer. Esse *fazer ao dizer*, como posteriormente se verá em detalhes (cap. 2), é conhecido como *dimensão ilocucionária*⁶ da língua.

Tal categorização, no *Salamanca* (1996, xi), é dada do seguinte modo, de acordo com as figuras abaixo:

La pragmática		
Las marcas pragmáticas, aunque no por completo, intentan señalar las posibilidades de uso en ciertas circunstancias de algunas expresiones y términos. Son las siguientes, que indican bastante en sí mismas:		
<i>amenaza</i>	<i>afectivo</i>	<i>afirmación</i>
<i>ánimo</i>	<i>anticipador narrativo</i>	<i>contestación a presentación</i>
<i>despedida</i>	<i>disgusto y enfado</i>	<i>eufemismo</i>
<i>humorístico</i>	<i>infantil</i>	<i>insulto</i>
<i>intensificador</i>	<i>ironía</i>	<i>llamada de atención</i>
<i>negación</i>	<i>petición</i>	<i>peyorativo</i>
<i>presentación</i>	<i>resumidor final</i>	<i>saludo</i>
<i>sorpreza</i>	<i>tratamiento</i>	
Los refranes, algunas frases y locuciones señalan, en principio, su uso pragmático general; por ejemplo: <i>¿Qué tal?: Se usa para saludar.</i>		

Figura 1: Lista de marcas de uso pragmáticas fornecidas pelo *Salamanca* (1996: xi) em sua Introdução.

TABLA 1. Conceptos pragmáticos y marcas pragmáticas del Diccionario Salamanca

Conceptos Pragmáticos	Marcas pragmáticas del Diccionario Salamanca
Teoría de los actos de habla	<i>Afirmación, negación, amenaza, petición, saludo, despedida, disgusto y enfado, insulto, peyorativo, ánimo, sorpresa humorístico, intensificador, afectivo</i>

Figura 2: Marcas de uso pragmático-ilocucionárias contidas no *Salamanca* (1996: xi) e listados por Guerra Salas e Gómez Sánchez (2005: 354).

A figura 1 mostra as marcas pragmáticas inseridas nos verbetes de algumas entradas do *Salamanca*, em conjunto com algumas informações a respeito do que tais marcas (enquanto categorias referentes à dimensão pragmática da língua) significam. Tais informações estão presentes na seção introdutória do referido dicionário – que, junto com a nomenclatura e demais seções, como se verá no capítulo 1, compõem a macroestrutura de um dicionário. Já a figura 2 mostra aquelas marcas pragmáticas contidas no *Salamanca* que, segundo Guerra Salas e Gómez Sánchez, são categorias de

⁵ “(...) al proferir un determinado enunciado”.

⁶ Posteriormente se verá (2.3.2) que a TAF comporta três dimensões: locucionária, ilocucionária e perlocucionária. Neste trabalho, contudo, tal como na maioria dos estudos sobre a TAF, “ato de fala” é usado como sinônimo de “ato ilocucionário”, “ilocução”, “força ilocucionária”.

atos de fala, ou seja, são *también* marcas ilocucionárias – marcas que, como se verá no capítulo 1, são inseridas nos verbetes de algumas entradas, constituindo o que se conhece como microestrutura de um dicionário.

Eis aqui, como exemplo introdutório, um trecho da entrada *siempre* tal como fornecida pelo *Salamanca* (1996: 1460):

siempre

adv. temp. **1** En todo momento o todo el tiempo. *El escritor siempre fumaba puros. Te echaré de menos siempre.* OBSERVACIONES: Normalmente se refiere al periodo de tiempo del que se habla o a uno que se da por consabido. (...) **4** AFIRMACIÓN. Resalta una cosa que se dice o refuerza una afirmación que no admite duda: *Siempre vivirás mejor solo que mal acompañado.*

Percebe-se que no verbete da entrada *siempre* do *Salamanca* está incluída a categoria ilocucionária *afirmación*. Isso significa que, de acordo com tal dicionário, proferir *siempre* equivale – se não em todos, pelo menos na maioria dos casos – a realizar a ação de afirmar. Por outras palavras, aquilo que *convencionalmente* se espera ao proferi-lo é a realização de uma afirmação.

Entretanto, os mencionados autores questionam os critérios utilizados pelos elaboradores do *Salamanca* para legitimar/ justificar tais inserções. De acordo com eles (2005: 361-362), podem-se apontar problemas tais como confusão e ambiguidade entre marcas de uso e definições, de modo que seria desejável que “na ‘Introdução’ o *Diccionario* explicasse o significado das marcas”⁷ – além de explicar certas ausências: “salta à vista que expressões como *por supuesto*, *claro* ou *seguro* (...) não sejam marcadas pragmaticamente como afirmações”⁸ (GUERRA SALAS; GÓMEZ SÁNCHEZ: 2005, 357).

Tendo isso em mente, pretendeu-se com o presente trabalho sugerir maneiras de inserção, em dicionários híbridos italiano>português-brasileiro tais como o *PC* (2012), de informações de natureza ilocucionária das duas línguas em questão, levando-se em consideração a composição tradicional dos dicionários⁹ de língua (macro, médio e microestrutura) e apontando as vantagens e desvantagens de cada uma das sugestões,

⁷ “(...) sería deseable que en la ‘Introducción’ del Diccionario se explicara el significado de las marcas”.

⁸ “Nos llama la atención que expresiones como *por supuesto*, *claro* o *seguro* (...) no aparezcan marcadas pragmáticamente como afirmaciones (...)”.

⁹ Vale dizer que o presente trabalho destina-se a sugerir inserções em *dicionários impressos*, o que não impede, contudo, que tais sugestões possam também ser adotadas (e aprimoradas) a casos específicos de dicionários eletrônicos, já que os recursos que estes últimos possuem são infinitamente superiores aos dos primeiros.

estabelecendo, ao fim, aquela aqui considerada, *segundo o caminho aqui percorrido*, como *ideal*¹⁰.

Antes de chegar a tal proposta final, porém, pretendeu-se verificar se determinada metodologia poderia auxiliar o lexicógrafo a obter tais informações com vistas a uma posterior dicionarização. Tal metodologia é aquela composta pelas ferramentas oferecidas pela Linguística de *Corpus* (doravante LC), que segundo Oliveira (2009: 49) é o campo de estudo da linguagem cujas análises linguísticas são realizadas com auxílio de programas de computador capazes de, por um lado, compilar e disponibilizar *corpora* de pesquisa a partir de textos autênticos em formato eletrônico, e, por outro, de permitir a extração de padrões linguísticos neles observados. O uso de *corpora* eletrônicos facilita, de acordo com Tagnin (2005: 21), “a identificação das unidades convencionais da língua”, e, segundo Berber-Sardinha (2004: 235), a realização de análises comparativas entre padrões linguísticos de duas línguas. Isso conduz à conclusão de que, em função de sua natureza informatizada, os *corpora* permitem, na perspectiva do dicionarista, agilidade na investigação, precisão e legitimidade nas escolhas, o que, por consequência, na perspectiva do usuário, garante produtos lexicográficos cada vez mais satisfatórios.

Tal metodologia será mais bem investigada no cap. 3. Por sua vez, os objetivos da presente investigação são mais bem explicitados abaixo.

II. Objetivos

A pergunta de pesquisa que direcionou o presente trabalho é a seguinte: considerando-se a possibilidade de um dicionário híbrido italiano>português-brasileiro conter informações acerca dos atos de fala, de que maneira esse tipo de informação poderia ser disponibilizada por tal dicionário, a partir do uso da LC como metodologia auxiliar para extração de padrões linguísticos aptos a serem dicionarizados?

Ao se tentar respondê-la, vislumbrou-se a possibilidade de alcançar os seguintes objetivos específicos: (i) apresentar um procedimento de pesquisa de espécies de ilocuções em *corpora* eletrônicos; (ii) sugerir, a partir das informações obtidas segundo o uso de *corpora* eletrônicos como fonte, maneiras de incluí-las em dicionários híbridos italiano>português-brasileiro, indicando-se ao fim uma em especial.

¹⁰ Demonstrativamente falando.

III. Justificativas

Justifica-se o presente trabalho na medida em que: são poucos os estudos voltados à investigação acerca da presença e/ ou inclusão de informações ilocucionárias em dicionários, sobretudo com relação aos híbridos; são poucos os estudos voltados à investigação de ilocuições em *corpora* com o objetivo de dicionarizá-los. Afora isso, considera-se que os dicionários híbridos, bem como outros dicionários de língua, são instrumentos linguístico-pedagógicos úteis a todos que estudam línguas estrangeiras, aos quais semelhanças e diferenças acerca de padrões de uso interacional e/ ou comunicativo entre duas línguas podem favorecer o aprendizado (o reconhecimento e o uso adequado) de tais padrões pragmáticos.

IV. Metodologia

Em primeiro lugar, a partir da leitura de obras da literatura corrente das áreas de Lexicografia, LC, Pragmática e Filosofia da Linguagem, estabeleceu-se a fundamentação teórica do projeto, a saber:

- a Lexicografia como ciência que cobre diversos aspectos de registros lexicais, problematiza a constituição e tratamento de unidades simples quando registrados em dicionários de língua e discute aspectos metodológicos da produção dicionarística (KRIEGER, 2008), podendo ser dividida em dois ramos: aquele que se debruça sobre a construção de dicionários (Lexicografia prática) e o se debruça sobre aspectos metodológicos dessa construção (Metalexicografia) (WELKER, 2004);
- a presente investigação como de natureza metalexicográfica teórica, uma vez que o objetivo é abordar um aspecto metodológico determinado, a saber, a de inserção de ilocuições em dicionários híbridos;
- os dicionários gerais de língua como aqueles tendem a apresentar, até certo ponto, a totalidade dos itens lexicais de uma língua (WELKER, 2004) (ou mesmo de mais de uma língua), em conjunto com informações acerca de sua estrutura e funcionamento;
- os dicionários gerais de língua como divididos em três partes: macroestrutura (conjunto de todas as seções: nomenclatura, textos

externos etc), microestrutura (conjunto de entradas, i. e., a nomenclatura em si) e medioestrutura (sistema de remissivas);

- o dicionário híbrido como tipo de dicionário geral de língua cuja função é colocar duas línguas em contato a partir de uma base monolíngue (a da língua-fonte) *mais* o fornecimento de equivalências (em geral semânticas) na língua-alvo, e com a possibilidade de colocar duas línguas em contato segundo aspectos pragmáticos;
- adoção do dicionário híbrido *PC* como paradigma de investigação;
- a Pragmática Linguística como disciplina que se ocupa da língua em uso, e que investiga tanto a influência do contexto interacional e/ ou comunicativo sobre a linguagem (de modo a apontar/ estabelecer quais regras/ manifestações linguísticas são as mais adequadas em determinadas situações), como a influência da linguagem sobre o contexto interacional e/ ou comunicativo (de modo a apontar/ estabelecer quais são as regras/ manifestações linguísticas que constituem determinadas situações) (BIANCHI, 2008);
- a TAF como teoria que compreende a linguagem como forma de ação (KOCH, 1992);
- os atos de fala como aqueles que são realizados ao se dizer algo (AUSTIN, 1990 [1975]), ou seja, disposições que a partir de dado proferimento linguístico constituem o contexto interacional e/ ou comunicativo em questão (BIANCHI, 2008; SBISÀ, 2009 [1989]);
- a competência pragmática como aquela que é necessária para que o falante possa produzir e compreender proferimentos em determinados contextos, abarcando, assim, a competência ilocucionária, i. e., o saber reconhecer e usar expedientes linguísticos que realizam determinados atos de fala (QECR, 2001);
- a LC como metodologia de investigação da língua por meio de análise de *corpora* representativos de linguagem natural compostos de textos disponíveis eletronicamente, permitindo a visualização e extração de padrões linguísticos segundo ferramentas tais como lista de frequência e lista de concordância, ou seja, baseadas no número de ocorrências (TAGNIN & TEIXEIRA, 2004; BERBER-SARDINHA, 2003);

- a LC como útil aos lexicógrafos, já que padrões linguísticos são justamente aqueles que compõem dicionários gerais de língua;
- adoção da abordagem baseada em *corpus* (*corpus-based*), segundo a qual o *corpus* é utilizado para testar e exemplificar teorias e descrições linguísticas pré-existentes (no caso, a TAF).

Em seguida, foram estabelecidos critérios para a investigação de ilocuções em *corpora* do italiano e do português-brasileiro. Isso pode ser dividido em dois momentos, estando o primeiro deles (e que se refere à pesquisa anterior à qualificação) abaixo:

- sendo os atos de fala aspectos da língua em uso, pensou-se primeiramente em constituir dois *corpora*, um do italiano e um do português-brasileiro, com textos compostos de *comentários* feitos por internautas a matérias jornalísticas publicadas por veículos de comunicação virtuais. Isso pelo motivo de que: (a) os comentários, apesar de representarem a língua escrita, possuem características similares à oralidade cotidiana (informalidade, espontaneidade, interatividade e flexibilidade), o que em tese facilitaria a busca por atos de fala; (b) em função de serem representantes da língua escrita, possuem a vantagem de proporcionarem a construção de *corpora* bem mais extensos do que *corpora* orais (logo, mais representativos em termos estatísticos da língua em geral); (c) em função de sua extensão, possuem a vantagem de melhor representarem a língua *stantard*, conceito imprescindível ao se pensar dicionários gerais de língua;
- foram escolhidas seis mídias brasileiras (UOL, Globo.com, Yahoo.br, Terra, Ig, Folha Online) e seis italianas (Libero, Yahoo.it, La Repubblica, Virgilio, Corriere della Sera, Tiscali) das quais foram extraídos comentários de internautas feitos a matérias relativas a cinco assuntos (Política, Cotidiano, Internacional, Esporte, Celebidades), procurando-se manter certa equivalência quantitativa entre os *corpora*;
- com os *corpora* prontos, foram investigados cinco procedimentos para observação e extração de atos de fala: (1) frequência de palavras, (2) frequência de palavras com a mesma função, (3) fórmulas usadas em contextos determinados, (4) verbos ilocucionários, (5) frequência de colocações/ coligações;

- em seguida, foram apresentadas vantagens e desvantagens de cada procedimento.

Esse primeiro momento foi debatido no exame de qualificação, quando então a banca apontou algumas falhas na proposta. Eis aqui:

- *corpora* orais podem ser tão extensos quanto (ou até mais do que) *corpora* escritos. Concordou-se com tal observação, mas manteve-se a opção por trabalhar com *corpora* escritos porque para a constituição de um dicionário geral de língua: (a) a economia de tempo/ trabalho é fundamental ao lexicógrafo, já que ele trabalha com uma grande quantidade de dados – economia esta que os *corpora* escritos lhe proporcionam, uma vez que a facilidade em se obter e/ ou compor *corpora* escritos são maiores, pois não precisam ser gravados/ filmados (o que demanda participação, nem sempre disponível, de um amplo número de pessoas), bem como prescinde da tarefa de transcrição (quando então passariam a se constituir como *corpora* escritos, tanto quanto os *corpora* que já são naturalmente escritos, o que logicamente significaria o mesmo); (b) a obtenção de dados da língua *standard* é importante para a constituição de um dicionário geral de língua e, considerando que em *corpora* orais a presença de regionalismos é em tese maior do que em *corpora* escritos, optou-se pelos últimos;
- os *corpora* confeccionados a partir dos comentários de internautas eram pouco extensos, violando assim o critério da representatividade da língua geral. Concordou-se com a observação, optando-se por abandoná-los e decidiu-se então pela adoção de *corpora* escritos disponibilizados na internet, a saber, *Corpus Paisà* e *Corpus do Português*;
- não se mostrou claramente o motivo pela escolha dos cinco procedimentos de investigação de atos de fala em *corpora*. Concordou-se com a observação;
- não se mostrou claramente como tais procedimentos poderiam auxiliar na inserção de atos de fala em dicionários. Concordou-se com a observação, já que os cinco procedimentos pertenciam a níveis categoriais diversos. A frequência de palavras (1) não é um procedimento de busca, bem como não o é a frequência de coligações/ colocações (lista de concordância) (5),

que são sim *ferramentas* de busca; os verbos ilocucionários (4) não são procedimentos de busca, bem como não os são as fórmulas usadas em contextos determinados (3), pois são sim *objetos* a serem buscados; a frequência de palavras com a mesma função (2) não é um procedimento de busca, pois é uma *análise qualitativa* a ser realizada após a obtenção de resultados quantitativos com base na frequência. Neste sentido, tais apontamentos foram úteis para a posterior adoção da frequência de palavras (1) como ferramenta de busca de verbos ilocucionários (4), de modo a, em seguida, analisá-los qualitativamente (2) para separá-los de casos anômalos. Concordou-se com isso, motivo pelo qual se decidiu reelaborar a proposta.

Uma parte desse primeiro momento da pesquisa foi disponibilizada no Anexo 2 desta dissertação – a saber, a proposta de construção de *corpora* compostos de *comentários* feitos por internautas à matérias jornalísticas publicadas por veículos de comunicação virtuais.

Diante de tais modificações outro caminho foi trilhado, constituindo-se assim o segundo momento do projeto. Eis aqui os novos passos:

- adoção de dois *corpora* escritos representativos (considerando-se suas extensões) do italiano e do português-brasileiro previamente elaborados e disponibilizados na internet, a saber, o *Corpus Paisà* e o *Corpus do Português*;
- adoção de uma classificação de atos ilocucionários específica, no caso, aquela proposta por Austin (1990[1975]), em função de ele ter associado ilocuições a verbos ilocucionários – o que para o presente trabalho facilita o reconhecimento das ilocuições em *corpora*;
- adoção de cinco espécies de verbos ilocucionários italianos (*promettere, giurare, impegnarsi, garantire, assicurare*) da classe dos comissivos (aqueles cujo proferimento compromete o falante a determinada linha de ação), com base na equivalência semântica (sinonímia) que tais verbos possuem entre si (com base no dicionário *Sinonimi e contrari* (2009)); em seguida, foram adotados os seus equivalentes semânticos em português-brasileiro (*prometer, jurar, comprometer-se, garantir, assegurar*);

- adoção de uma sintaxe de busca capaz de permitir o reconhecimento de ilocuções em *corpora*, no caso, a forma verbal na primeira pessoa do singular do presente do indicativo na voz ativa, a mesma usada por Austin para reconhecer ilocuções.

Em terceiro lugar, realizou-se a busca das cinco ilocuções em italiano e em português-brasileiro, obtendo-se assim o número total de ocorrências de cada um deles nos *corpora* mencionados. Entretanto, de modo a possuir números reais de ocorrências de cada um deles, ou seja, ter certeza (ou, pelo menos *mais* certeza) de que em todas as ocorrências os verbos em questão desempenharam a função ilocucionária esperada, realizou-se uma análise qualitativa de modo a excluir ocorrências anômalas. Os critérios para estabelecimento e exclusão de tais ocorrências foram os seguintes: (a) homonímia; (b) negação; (c) repetição e (d) inteligibilidade das ocorrências.

Em quarto lugar, tendo-se obtido os números reais de ocorrência, optou-se em primeiro lugar pela exclusão das ilocuções *mi impegno* e *comprometo-me/ me comprometo* devido ao baixo número de ocorrências. Em seguida, investigaram-se possibilidades de inserção de informações relativas a tais ilocuções em dicionários híbridos, segundo consideração da divisão tradicional dos dicionários em três partes e apontando os prós e contras de cada possibilidade, sugerindo ao fim uma possibilidade de inserção em especial, a saber: (1) a inserção de marcas de uso referentes às classes de ilocução no interior das quais se encontram as espécies de ilocução realizadas segundo proferimento de seus verbos ilocucionários correspondentes; (2) a inserção de uma remissiva (*vedi r*) após as marcas de uso ilocucionárias, de modo a conduzir o leitor a um material externo à nomenclatura do dicionário (aqui chamada de “seção *r*”); (3) a inserção de tal seção no dicionário, sendo esta composta de: (a) explicações referente às classes de ilocuções e lista de respectivas espécies convencionalmente recorrentes do italiano dispostas segundo frequência; (b) espécies de ilocuções equivalentes do português-brasileiro dispostas segundo frequência; (d) exemplos de uso de tais verbos ilocucionários desempenhando sua função ilocucionária convencional retirados dos *corpora*.

V. Estrutura da dissertação

Os **capítulos 1, 2 e 3** compreendem a **fundamentação teórica** da pesquisa.

O **capítulo 1** foi destinado a apresentar aspectos relacionados à **Lexicografia** que são relevantes para o presente trabalho, a saber: a sua definição (ciência dos dicionários); suas ramificações (Lexicografia Teórica e Prática); os seus objetos (dicionários gerais de língua, entre os quais os híbridos), considerando-se suas divisões internas (macro, micro e medioestrutura) e respectivas subdivisões (nomenclatura e materiais externos; marcas, notas e exemplos de uso; remissivas); o dicionário híbrido *PC* tomado como paradigma de investigação.

O **capítulo 2** foi destinado a apresentar aspectos relacionados à **TAF** que são relevantes para o presente trabalho, a saber: a sua definição (teoria segundo a qual dizer é fazer); a definição da disciplina que a possui como objeto de estudo (Pragmática Linguística: disciplina que examina a linguagem em contextos de proferimento); os principais conceitos com os quais se relaciona (contexto, comunicação, interação, competência); a TAF de Austin, com suas respectivas classe e espécies de atos de fala, tomada como paradigma, bem como o seu critério para estabelecimento de ilocuções tomado como sintaxe de busca em *corpora* (uso da forma verbal na primeira pessoa do singular do presente do indicativo na voz ativa como sintaxe de busca).

O **capítulo 3** foi destinado a apresentar aspectos relacionados à **LC** que são relevantes para o presente trabalho, a saber: a sua definição (metodologia de exploração da linguagem por uso de *corpus*); a definição de *corpus* (coletânea de textos eletrônicos); apresentação de suas principais ferramentas (lista de frequência de palavras, sobretudo); a abordagem de pesquisa adotada (*corpus-based*); o critério para escolha dos *corpora* (o de representatividade com base na extensão); os *corpora* online escolhidos como fonte de pesquisa (o *Corpus Paisà* e o *Corpus do Português*).

O **capítulo 4** foi destinado a apresentar os **procedimentos e resultados quantitativos da análise de corpora** a partir dos quais dez ilocuções foram investigadas (*prometto, giuro, garantisco, mi impegno, assicuro*, do italiano; *prometo, juro, garanto, comprometo-me/ me comprometo, asseguro*, do português-brasileiro); obtiveram-se também **resultados qualitativos** a partir do momento em que ocorrências anômalas foram excluídas (casos de: homonímia, negação, repetição e ininteligibilidade); excluíram-se também, por fim, duas das ilocuções investigadas, uma do italiano e outra do português-brasileiro, em função da baixa frequência.

O **capítulo 5** foi destinado a apresentar as **sugestões de dicionarização** de tais ilocuções, bem como a sugestão considerada ideal em razão de sua decorrência da fundamentação teórica e da discussão conceitual acerca de algumas sugestões

inicialmente vislumbradas. Tal sugestão ideal é a seguinte: inserção de marcas de uso referentes às classes de ilocução na microestrutura, inserção de remissivas após aquelas de modo a conduzir o leitor a uma seção externa à nomenclatura do dicionário, e inserção de tal seção, composta de informações acerca das ilocuições, lista de ilocuições mais frequentes do italiano e do português-brasileiro e exemplos de concretos retirados dos *corpora*.

Por fim, foram apresentadas as **considerações finais**, a saber: breve resumo do percurso e dos resultados; apontamento e reflexões de tópicos negligenciados e/ ou destinados a pesquisas futuras; comentários acerca de questões levantadas na banca de defesa.

Além disso, foram oferecidos também: **referências bibliográficas** e **dois anexos**, a saber: imagens relativas à pesquisa em *corpus*; os procedimentos de compilação de dois *corpora* com base em comentários de internautas a matérias jornalísticas publicadas *online*, desenvolvidos na primeira metade da pesquisa.

1. LEXICOGRAFIA

Neste capítulo, define-se a Lexicografia como área científica que se debruça sobre a teorização e a elaboração de dicionários gerais de língua. Entre eles, o dicionário híbrido é apresentado como aquele cujas características principais são: possuir uma base monolíngue na língua-fonte; fornecer equivalentes semânticos na língua-alvo. Além disso, apontam-se as três partes que compõem um dicionário geral de língua (macroestrutura, microestrutura e medioestrutura). Por fim, o PC, tomado como paradigma desta investigação de natureza metalexicográfica, é examinado e decomposto em todas as suas partes.

1.1. A ciência dos dicionários

A Lexicografia é a área que, segundo Biderman (2001a: 15-17), pertence às ciências do léxico¹¹, e se constitui, mais especificamente, como ciência dos dicionários. Ao fazer um apanhado das ciências do léxico, Barros (2004: 63-64) aponta como objetivo da Lexicografia elaborar dicionários de língua ou especiais¹², e ressalta que tais dicionários podem ser monolíngues, bilíngues ou multilíngues. Ampliando-se tais definições, Krieger (2008: 170) diz que a Lexicografia é a área científica que cobre diversos aspectos de registros lexicais, problematiza a constituição e tratamento de unidades simples e complexas (além de outras faces do léxico geral, quando registrados em dicionários de língua) e discute aspectos metodológicos da produção dicionarística. Definição semelhante, embora proposta do ponto de vista não da ciência e sim de seu

¹¹ As outras ciências são: a Lexicologia, que na visão de Lorente (2007: 19) é a disciplina que se ocupa do léxico das línguas de forma completa e integrada, e a Terminologia, que na visão de Barros (2004: 40) analisa o termo, “(...) unidade lexical com um conteúdo específico dentro de um domínio específico”. Biderman (2001a: 16-19) define a Lexicologia como ciência cujo objeto de estudo é a palavra, a categorização lexical e a estruturação do léxico, enquanto que a Terminologia como aquela cujo domínio é um subconjunto do léxico de uma língua, a saber, “cada área específica do conhecimento humano”. Para Barros (2004: 64), Lexicologia, Lexicografia e Terminologia mantêm relações de alimentação e realimentação científicas, “primando, no entanto, pela autonomia metodológica e assumindo estatutos epistemológicos próprios”. O mesmo diz Marelló (1996 *apud* ZUCCHI, 2010: 119, grifo nosso): “a relação entre lexicologia e lexicografia não é mais, supondo-se que tenha sido, uma relação de filiação, com a teoria que vem diminuía na prática: a lexicologia é *uma das muitas* fontes de informação do lexicógrafo (...)”. MARELLO, C. *Le Parole dell’italiano*. Bologna: Zanichelli. 1996.

¹² Os dicionários especiais são entendidos como aqueles que registram apenas um tipo de unidade lexical ou fraseológica (p. ex.: dicionários de expressões idiomáticas, de provérbios, de ditados, de sinônimos, de uso etc).

objeto, é a que apresenta Strehler (2001: 171), para quem os produtos lexicográficos são, em essência, descrições da língua.

De todas essas definições adota-se aqui a de Krieger, pelo fato de esta, de certo modo, abarcar todos os aspectos mencionados pelas outras, além de conter outro aspecto de grande relevância para o presente trabalho. Ao salientar que a constituição e tratamento de unidades simples e complexas abarcam outras “faces do léxico geral”, isso abre caminho para a presença de informações pragmáticas em dicionários gerais de língua. Tal inferência encontra respaldo na própria Krieger (2008: 171), pois a pesquisadora salienta que os atuais desenvolvimentos da Lexicografia se alicerçam na “noção linguística de uso da língua”, e que, portanto, este merece ser considerado pela Lexicografia. E além de fazer menção aos “aspectos metodológicos da produção dicionarística”, Krieger (2008: 171) ainda ressalta a contribuição dos *corpora* para tal, o que também favorece o estudo que aqui se pretende realizar.

Por dicionário de língua, assim, compreende-se uma determinada compilação dos itens lexicais de uma ou mais línguas, organizada de determinado modo e a partir da qual são fornecidas informações diversas acerca desta(s) língua(s), em termos de estrutura e uso, de acordo com os objetivos que se pretende alcançar com tal obra e o seu público-alvo.

1.2. Lexicografia prática e Metalexiconografia

Borba (2003: 15) apresenta a Lexicografia sob duplo aspecto: o primeiro, como técnica de montagem de dicionários – ocupando-se, assim, de critérios para seleção de nomenclaturas/ conjunto de entradas, de sistemas definitórios, de estruturas de verbetes, critérios para remissões e registro de variantes etc; o segundo, como teoria – dedicando-se, desta forma, ao estabelecimento de princípios que permitam descrever o léxico de uma língua. Tal proposta assemelha-se aquela que Welker (2004: 11), ao elencar diversos autores, mostra como distinção frequente da Lexicografia, a saber: prática (ciência, técnica, prática, arte de elaborar dicionários) e teórica (que o autor nomeia de metalexiconografia, que abrange o estudo de problemas relacionados à elaboração, críticas, pesquisas históricas e pesquisas de uso de dicionários).

Neste sentido, pode-se dizer que, grosso modo, cabe ao lexicógrafo produzir um dicionário, ao passo que analisar e discorrer teoricamente acerca de tal produção é tarefa

do metalexicógrafo¹³, noção que corresponde, por sua vez, com a distinção proposta pelo Grupo de Trabalho de Lexicologia, Lexicografia e Terminologia da Associação Nacional de Pós-Graduação e Pesquisa em Letras e Linguística (GTLEX-ANPOLL) segundo o qual a Lexicografia prática é a que tem por finalidade elaborar obras de referência, principalmente dicionários (impressos ou *online*), o que a distingue, portanto, da Metalexigrafia, cujo objetivo é estudar “todas as questões ligadas aos dicionários (história, problemas de elaboração, análise, uso)” (GTLEX, online).

De acordo com o exposto, deve-se dizer que este trabalho é de natureza *metalexigráfica*, uma vez que o objetivo aqui não é produzir um dicionário, mas focar sobre um aspecto particular de sua constituição (a inserção de ilocuções em dicionários híbridos), tendo-se em vista a função da Metalexigrafia tal como explicitada por Duran (2004: 18): “os estudos metalexigráficos têm por objetivo fornecer subsídios que propiciem melhoria na qualidade dos trabalhos lexicográficos”. Vale dizer ainda que segundo a divisão que Schneider (2001 *apud* DURAN, 2004: 17)¹⁴ propõe entre Metalexigrafia Teórica (que produz estudos a partir da análise de dicionários) e Metalexigrafia Empírica (que produz estudos baseados em pesquisas de campo acerca da interação dicionário-usuário), este trabalho insere-se no âmbito da *Metalexigrafia Teórica*.

Em termos práticos, isso significa dizer que este trabalho busca tão somente sugerir modos possíveis de inserção de ilocuções em dicionários híbridos, e não, como se poderia pensar, estabelecer classes/ espécies de ilocuções para em seguida dicionarizá-los. Assim, vale-se neste trabalho da característica *interdisciplinar* da Lexicografia tal como apontada por Marengo (1.1, nota de rodapé nº 11), ao se adotar uma classificação de ilocuções previamente constituída como paradigma de investigação, a saber, aquela proposta por Austin (1975[1990]) – cujos detalhes serão examinados posteriormente (2.3.2 e 2.4). Por consequência, esta é a razão pela qual se prefere utilizar *corpora* escritos, e não orais, pois considerando que o que se busca são padrões representativos da língua em geral passíveis de dicionarização, os escritos apresentam-se não somente como mais extensos, entre os que são disponibilizados na internet, e, logo, mais representativos em termos quantitativos da língua geral, como

¹³ Importante ressaltar, porém, que tal distinção não é absoluta, pois existem pesquisadores que são ao mesmo tempo lexicógrafos e metalexicógrafos. Welker (2004: 11) cita os nomes de Maria Tereza C. Biderman e Francisco S. Borba como casos concretos disso.

¹⁴ SCHNEIDER, K. Review of “Using Dictionaries: Studies of Dictionary Use by Language Learners and Translators. ATKINS, B. T. S. (ed). Tübingen: Niemeyer, 1988 (Lexicographica: Series maior 88)”. *System* 29, p. 153-159, 2001.

também são mais representativos (em termos qualitativos) da língua *standard*, por serem menos afeitos a oferecer dados de variações diatópicas, diastráticas etc. Este ponto será posteriormente examinado em detalhes (3.4).

1.3. Léxico e dicionário geral de língua

Se a Lexicografia se debruça sobre o léxico de uma ou mais línguas de modo a dicionarizá-lo, deve-se compreender o significado de *léxico* para esta ciência, uma vez que este termo, utilizado correntemente, é semanticamente ambíguo.

Segundo Welker (2004: 15), léxico, etimologicamente falando, vem do grego *lexis*, literalmente *palavra*. Porém, ressalta que o significado de *palavra* não é consensual. Neste caso, léxico poderia significar conjunto de morfemas, sejam estes lexicais ou gramaticais¹⁵ (REY, 1977 *apud* WELKER, 2004: 15)¹⁶, morfemas presos (unidades abaixo do nível de palavra), fraseologias (unidades acima do nível de palavra) (SCHINDLER, 2002 *apud* WELKER, 2004: 16)¹⁷ ou sinônimo de dicionário/vocabulário (WELKER, 2004: 16). Analogamente, De Mauro (2005: 12-19) apresenta várias definições possíveis, como “palavra gráfica” (sequência de caracteres separados de outros por um espaço em branco ou pontuação), “palavra fonológica” (palavras gráficas caracterizadas segundo por unidades fônicas), “palavra como bloco significante” (morfemas livres) etc¹⁸.

Tecnicamente falando, porém, no interior da ciência lexicográfica, Borba (2003: 16) propõe a compreensão de léxico como “o conjunto dos itens vocabulares da língua, ou seja, (...) a soma das formas livres que circulam nos discursos da comunidade”. Estes itens vocabulares livres são apontados pelo autor (2003: 19) como *lexias* ou *lexemas*, termos sinônimos entre si, que possuem características básicas: forma livre mínima, configuração fônica estável¹⁹, expressão de conteúdo significativo único e preenchimento de funções gramaticais específicas. Para Borba (2003: 22), existem

¹⁵ Os morfemas lexicais abarcariam os morfemas presos, como *sapat-*, de *sapatos*, e os livres, como *mar*, enquanto que os gramaticais abarcariam adjetivos, pronomes etc.

¹⁶ REY, A. *Le lexique: images et modèles. Du dictionnaire à la lexicologie*. Paris : Colin, 1977.

¹⁷ SCHINDLER, W. *Lexik, Lexikon, Wortschatz: Probleme der Abgrenzung*. In: CRUSE, A. et al. (ed.). *Lexicology. An International Handbook on the Nature and Structure of Words and Vocabularies*. Berlin-New York: de Gruyter, 1986, 34-44.

¹⁸ *Parola grafica, parola fonologica e parola come blocco significante* são os termos em italiano utilizados pelo autor.

¹⁹ Explica Borba (2003: 19): “bloco fônico que não permite a inserção de outros elementos mórficos”.

lexias simples e complexas, as primeiras formadas por uma única forma livre e as complexas formadas pela combinação de formas livres.

Barbosa (2001: 35), por sua vez, prefere o termo *lexema*, ao dizer que “o dicionário de língua tende a reunir o universo dos *lexemas*, apresentando, para cada um deles, os vocábulos que representam suas diferentes acepções” – entendendo aqui *lexema* como unidade da língua e *vocabulo* como unidade da fala/ discurso. Ou seja, seguindo a autora (2001: 35), enquanto *voar* seria um *lexema*, por conter um *semema* (unidade semântica) polissêmico, o *voar* dos pássaros ou o “*voar*” dos pensamentos seriam vistos como duas acepções possíveis de *voar*, i. e., *vocábulos*, que variam de acordo com seu emprego segundo contexto de proferimento. Outro autor oferece o exemplo do verbo *cortar*, este podendo ser visto como *lexema*, enquanto que *cortei*, *cortaremos* poderiam ser vistos como formas desse *lexema* (LYONS, 1979: 206 *apud* WELKER 2004: 21)²⁰ – aquilo que Barbosa chamou de *vocabulo*. Borba (2003: 19) adota postura semelhante ao dizer que o *vocabulo* se aplica à fala, discurso, enquanto que a *lexia/lexema* se aplica ao sistema da língua, e Biderman (2001b: 140) dá a essa noção de *lexia/lexema* o significado de *paradigma*, ou seja, elemento permanente do sistema linguístico que assume várias formas (os *vocábulos*, tal como mencionado acima).

No presente trabalho, o termo *item lexical* é adotado como sinônimo de *lexia/lexema* – forma livre mínima com significado, fonética estável e função gramatical específica, atuando como *paradigma* do sistema linguístico que abrange diversos *vocábulos* (formas diversas, variações do *item lexical* tomado por *paradigma linguístico estável*)²¹. Por sua vez, *vocabulo* é adotado com referência a acepções e variantes de um *item lexical*.

Para Barbosa (2001: 40), o *item lexical* é a unidade-padrão do dicionário de língua, resultante da recuperação de normas semânticas diversas, suas variações e que compreende vários *vocábulos* correspondentes a distintas acepções, que, no entanto, mantêm uma intersecção. Importante dizer que os *itens lexicais*, como unidades-padrão do dicionário geral de língua, ao serem selecionados para compor um dicionário são tecnicamente chamados de *lema* ou *entrada*. Segundo Welker (2004: 91), “(...) toma-se

²⁰ LYONS, J. *Introdução à linguística teórica*. Trad. R. V. Mattos e Silva e H. Pimentel. São Paulo: Companhia Editora Nacional-Edusp, 1979.

²¹ Essa definição é suficiente para os propósitos deste trabalho e não exclui, portanto, a possibilidade de se entender *item lexical* como abrangendo também *morfemas presos* e *fraseologias* – menos ainda a possibilidade desses serem inseridos em dicionários de língua.

como lema a forma ‘básica’ ou ‘canônica’ do lexema: o infinitivo dos verbos, o singular masculino dos substantivos e dos adjetivos”²². Neste sentido, ainda com o autor na mesma página, “(...) lematizar significa encontrar uma forma canônica dos lexemas”. Além do lema, há também o sublema (ou subentrada), “item lexical – geralmente lexema composto ou complexo”²³ – que é tratado dentro do mesmo verbete do lema” (WELKER, 2004: 93).

Vale explicitar que por dicionário geral de língua²⁴ entende-se, segundo Welker (2004: 77), aquele que “tende a apresentar a totalidade dos lexemas de uma língua”. Entretanto, aponta-se a seguinte ressalva (não encontrada explicitamente em Welker, mas que certamente é pressuposta por ele): a totalidade de que aqui se fala é entendida em termos ideais, devido à impossibilidade de se compor um dicionário a partir da totalidade real-concreta de todas as informações léxico-linguísticas possíveis. Isso pelo simples fato de que nenhuma proposta de registro pode ser idêntica ao próprio objeto registrado (caso contrário, não faria sentido distinguir o registro do objeto registrado, já que seriam idênticos e tratar-se-ia da mesma coisa). Logo, aceitando-se a noção de que nenhum dicionário é capaz de conter o registro de todo o léxico corrente de uma língua segundo todas as suas dimensões e todas as suas possibilidades de uso – razão pela qual existem *n* tipos de dicionários, voltados a *n* públicos-alvo, organizados de *n* maneiras, de acordo com *n* interesses envolvidos em sua publicação – dois equívocos poderiam ser evitados: (i) o de se pensar *nos* dicionários, i. e., classe abstrata que conteria *todos* os dicionários existentes, que já existiram e que irão existir (o que, segundo Van Sterkenburg (2003: 3), não passa de um dicionário prototípico) ao invés de *o dicionário tal*, exemplar historicamente específico; (ii) o de que mesmo no caso de um dicionário específico este não deve ser concebido como se fosse capaz de conter todos os registros/ descrições da língua em uso possíveis – pois, caso contrário, configurar-se-ia um paradoxo, uma vez que a língua em uso está em constante devir, sendo portanto incapaz

²² Que o autor reconhece ser uma convenção, mas que, por outro lado, discutindo o assunto a partir da posição de outros pesquisadores, diz não haver razão para mudança.

²³ Segundo Borba (2003: 22), as lexias complexas (as lexias que combinam mais de uma forma livre – João-de-barro, porta-luvas etc) têm mostrado importância tanto para a organização de gramáticas pedagógicas como para montagem de dicionários de uso. Biderman (2001b: 140-141) reconhece a dificuldade de se reconhecer as lexias complexas cujo valor esteja bem constituído – e decidir sobre qual dessas comporá a macroestrutura de um dicionário é outra dificuldade.

²⁴ Um dicionário de regência verbal ou um dicionário de uso são dicionários de língua, mas não são dicionários gerais de língua, já que se debruçam sobre um aspecto particular da língua. Daí a necessidade de definir dicionário geral de língua, uma vez que o objetivo aqui é abordar uma variante deste, o dicionário híbrido. Assim, doravante, toda referência a dicionários será feita pensando-se em dicionários gerais de língua, até o momento em que se começar a falar de dicionários híbridos.

de ser recortada em sua totalidade. Porém, isso não implica a impossibilidade de registro do uso, pelo contrário; contudo, deve-se ter em mente essa distinção significativa entre *um* tipo de registro *tal* e *o* registro.

Em resumo, de acordo com o que foi exposto até aqui e levando em conta o caso dos atos de fala, o padrão a ser utilizado neste trabalho será o seguinte: tendo-se em vista um dicionário geral de língua, o infinitivo dos verbos (suponha-se, *prometer*) será entendido como *item lexical*, e ao ser selecionado para ser dicionarizado será entendido como *entrada*. Já a forma do verbo *prometer* na primeira pessoa do singular do presente do indicativo na voz ativa, *prometo*, que será utilizada como sintaxe de busca em *corpora* eletrônicos, será, se dicionarizada, vista como *vocábulo*.

1.4. Macroestrutura e nomenclatura

Em termos de estrutura, o dicionário é organizado em dois eixos principais (SANTIAGO, 2012): a macro e a microestrutura. A macroestrutura corresponde à organização do dicionário como um todo – aquilo que Hartmann e James (1998: 92) chamam de megaestrutura²⁵. De acordo com Biderman (2001b: 133-135), alguns critérios macroestruturais devem ser levados em conta pelo lexicógrafo no momento da elaboração de uma obra lexicográfica: público a que se destina; uso de um “enorme corpus de dados linguísticos, de discursos realmente realizados – língua escrita e oral – para daí extrair a nomenclatura desejada com base em critérios léxico-estatísticos” (sic); extração de itens lexicais correntes na língua comum; entre outros. Tais critérios apontados por Biderman encontram eco em Miranda (2007), de quem se pode depreender que fazem parte do âmbito macroestrutural desde o universo léxico que o dicionário pretende conter até a consideração sobre a qual tipo de público a obra se destina, bem como, também, a de princípios funcionais entre maiúsculas e minúsculas²⁶.

O constituinte principal da macroestrutura de um dicionário é a nomenclatura²⁷. Viu-se anteriormente (1.3) que um item lexical selecionado em sua forma canônica ao ser dicionarizado é chamado de entrada. Na constituição de um dicionário geral de língua a entrada faz parte da nomenclatura, o conjunto das entradas ou a “(...) forma

²⁵ *Megastructure*.

²⁶ Diz Miranda (2007: 271) que “em algumas línguas existe um princípio de oposição funcional entre o uso de maiúsculas e minúsculas em relação à primeira letra do signo-lema”.

²⁷ Que Hartmann e James (1998: 91) chamam de *macrostructure*, esta como parte da megaestrutura. Neste trabalho, optou-se por nomenclatura como parte da macroestrutura.

como o corpo do dicionário é organizado” (WELKER, 2004: 81). Normalmente, as entradas são organizadas em ordem alfabética de significantes – tanto em dicionários semasiológicos, “(...) vai da forma, do lexema, ao significado” como em dicionários onomasiológicos, cujo movimento é oposto: “parte-se de conceitos para encontrar signos” (BALDINGER, 1960: 523 *apud* WELKER, 2004: 47)²⁸. Além dessa disposição, outra questão que compete à nomenclatura é sua extensão. Isso se refere, de acordo com Silva (2007: 285), “ao formato e número de entradas, em que se devem verificar as dimensões, o número de volumes, de páginas, de entrada em cada volume”.

1.4.1. Textos externos

A nomenclatura é o constituinte principal da macroestrutura de um dicionário, mas não o único. Como aponta Welker (2004: 78) um dicionário geral de língua não se constitui como mera lista de itens lexicais e respectivas informações, pois a grande maioria das obras possuem outros elementos, que juntos compõem a macroestrutura, a saber:

prefácio, introdução, lista de abreviaturas usadas no dicionário, informações sobre a pronúncia, resumo da gramática, lista de siglas e/ ou abreviaturas, lista de verbos irregulares, lista de nomes próprios, lista de provérbios, bibliografia, fontes, às vezes, certas curiosidades.

Welker (2004: 79) chama tais elementos de “textos externos”. De acordo com o autor (2004: 80), os textos externos não têm posição fixa, com exceção, por razões óbvias, do prefácio e da introdução, que precedem à nomenclatura.

Selistre (2009: 197), por sua vez, inclui esses elementos sob a expressão “material externo”²⁹ de um dicionário, sendo que o material externo pode ser dividido em três partes: anteposto, interposto e posposto. O material anteposto, segundo autora, é aquele que é apresentado antes da nomenclatura e em geral “apresenta uma introdução ou prefácio com informações sobre a elaboração do dicionário e um guia para o usuário, que traz o tipo de informações incluídas no dicionário e a forma de serem acessadas” (2009: 197). Já o material interposto corresponde a elementos inseridos entre os

²⁸ BALDINGER, K. Alphabetisches oder begrifflich gegliedertes Wörterbuch? *Zeitschrift für romanische Philologie* 76: 521-536, 1960.

²⁹ Selistre (2009) considera o material externo no mesmo plano da macroestrutura e não como parte desta, tal como se considera neste trabalho. Isso se deve, provavelmente, ao fato da autora entender o termo “macroestrutura” como sinônimo de “nomenclatura”, sendo que aqui a nomenclatura é tomada como parte da macroestrutura – esta composta por aquela *mais* o material externo.

verbetes, tais como “quadros mostrando as relações entre várias palavras sinônimas, as diferenças entre pares de palavras que costumam gerar confusão e questões gramaticais que envolvem o uso do item em questão” (2009: 198). Por sua vez, o material posposto é o que se segue à nomenclatura, podendo constar de “lista dos verbos irregulares, listas com nomes de países e nacionalidades, tabelas de pesos e medidas e apêndices gramaticais” (2009: 198). Santiago (2012) propõe caracterização semelhante, do mesmo modo que Hartmann e James (1998: 92)³⁰.

Neste trabalho, *texto externo* e *material externo* (e até mesmo *seção externa*) são concebidos como termos equivalentes, referindo-se a todas as partes macroestruturais externas à nomenclatura.

1.5. Microestrutura e verbetes

A microestrutura “(...) responde à pergunta sobre as diversas acepções da palavra” (BALDINGER, 1960: 524 *apud* WELKER, 2004: 107)³¹, ou ainda, “(...) conjunto das informações ordenadas de cada verbete após a entrada” (REY-DEBOVE, 1971: 21 *apud* WELKER, 2004: 107)³². Hartmann e James (1998: 94), por sua vez, dizem que a microestrutura corresponde às informações internas de uma unidade lexical. Novamente com Welker (2004: 83), “o número de verbetes é influenciado por decisões do lexicógrafo a respeito da inclusão ou não de diversos tipos de lexemas” – entendendo por verbete o conjunto das acepções, exemplos e demais informações que seguem cada entrada.

Cabe ao lexicógrafo, então, planejar como as diversas informações (verbetes) de uma entrada são dispostas. Entre as informações mais usuais estão: as que identificam o lema na sincronia (grafia, pronúncia, acentuação, classe gramatical, flexão) e na diacronia (etimologia); marcas de uso; informações explicativas (definição); informações sintagmáticas (construções, colocações, exemplos de uso); informações paradigmáticas (sinônimos, antônimos etc); outras informações de natureza semântica

³⁰ Embora, novamente, estes últimos chamem de *megaestrutura* aquilo que aqui se chama de *macroestrutura*.

³¹ REY-DEBOVE, J. *Étude linguistique et sémiotique des dictionnaires français contemporains*. Paris : Hachete, 1971.

³² BALDINGER, K. Alphabetisches oder begrifflich gegliedertes Wörterbuch? *Zeitschrift für romanische Philologie* 76: 521-536, 1960.

(metáforas etc); elementos de ordenamento (p. ex., uso de símbolos); remissões; entre outros (HAUSMANN; WIEGAND, 1989: 341 *apud* WELKER, 2004: 108)³³.

1.5.1. Marcas de uso

Por marcas de uso, como o próprio termo indica, entendem-se as categorias presentes em verbetes de dicionários gerais de língua cujo objetivo é fornecer alguma informação relativa a algum aspecto do uso concreto que *convencionalmente* se faz de determinado item lexical/ vocábulo. Neste sentido, as marcas de uso possuem função pedagógica relevante, uma vez que auxiliam o usuário de um dicionário (p. ex., no caso de um dicionário bilíngue/ híbrido) a selecionar e empregar itens lexicais adequados da língua estudada, segundo consideração de dado contexto interacional e/ ou comunicativo. Como se verá no capítulo 2, o proferimento de dado item lexical (no interior de uma sentença) tanto pode desempenhar um ato de fala *convencionalmente* adequado, como pode desempenhar um ato de fala possível e *convencionalmente* inadequado (ou menos adequado). Por exemplo, se é verdade que *promettere* e *garantire* são verbos italianos semanticamente equivalentes, é falso que ambos possam ser usados indiscriminadamente em qualquer contexto comunicativo – se um sujeito pretende realizar uma promessa, “*prometto che...*” é mais adequado do que “*garantisco che...*”. Logo, as marcas de uso podem contribuir indicando quais itens lexicais são mais ou menos adequados para a realização de atos de fala em dado contexto.

Considera-se relevante aqui indicar algumas espécies de marcas de uso, já que o objetivo deste trabalho é abordar um aspecto particular de uso da língua que, julga-se, merece ser dicionarizado – e que, ao sê-lo, segundo a discussão aqui feita, poderia sê-lo também sob marcas de uso.

Ao se pensar em marcas de uso de natureza sociolinguística, Borba (2003: 315), que se refere a tais marcas por “rótulos” (aqui, ambos tomados como sinônimos), diz que estas se relacionam com a variação “tanto espacial, de uma região para outra, como social, no mesmo espaço considerada quanto aos diferentes registros utilizados pelas pessoas nas diferentes situações da vida social”. O mesmo autor (2003: 315-7) diz que

³³ HAUSMANN, F. J.; WIEGAND, H. E. Component Parts and Structures of General Monolingual Dictionaries: a Survey. In: HAUSMANN, F. J. et al. (ed.), v. 1, 328-360. *Wörterbücher – Dictionaries – Dictionnaires. Ein Internationales Handbuch zur Lexicographie*. Berlin/ New York: de Gruyter, 1989, v. 1, 328-360.

“os dicionários costumam dar este tipo de informação por um conjunto de rótulos (...)”, tais como: rótulos de natureza diatópica (informa diferenças entre subcomunidades geograficamente situadas em áreas diversas), diastrática (informa diferenças entre subgrupos de uma mesma comunidade segundo gênero, faixa etária etc) e diafásica (informa diversidade de estilos segundo circunstâncias discursivas).

Além dessas três mencionadas por Borba, Welker (2004: 131), com referência a Hausmann³⁴, acrescenta outras: diacrônicas (p. ex., *antiquado*, *neologismo* etc), diaintegrativas (para assinalar estrangeirismo), diamediais (assinalam diferenças entre linguagem oral e escrita), diatextuais (assinalam que o item lexical é restrito a um gênero textual: jornalístico, poético etc), diatécnicas (informam que a acepção pertence a uma linguagem técnica), diafrequentes (raro, muito raro etc), diaevaluativas (assinalam atitude por parte do falante: pejorativo, eufemismo etc), dianormativas (assinalam que o item lexical é errado em termos de norma da língua padrão). Welker (2004: 134-135) defende que, apesar das dificuldades³⁵ de delimitação e inclusão de marcas de uso seria desejável que houvesse mais delas do que o que se verifica na maioria dos dicionários – e ressalta (2004: 136) que as marcas de uso são “(...) particularmente importante[s] no dicionário bilíngue quando consultado na produção de textos na L2”, além de reforçar que tais marcas, quando adotadas, devem ser explicada ao consulente do dicionário.

Se Welker tem razão ao dizer que a presença de marcas de uso deve ser explicada, isso significa que a macroestrutura de um dicionário é relevante nesse processo, uma vez que a explicação será dada em textos externos. Esse é o caso concreto do *Diccionario Salamanca de la Lengua Española* (1996), monolíngue dirigido a estudantes, *hispanohablantes* ou não, que tenham por objetivo aprimorar o domínio da língua espanhola, bem como a professores que se dedicam a ensinar espanhol (GUTIÉRREZ CUADRADO, 1996: v).

Como se viu anteriormente (Introdução, I), o *Salamanca* oferece marcas de uso especificamente pragmáticas em sua microestrutura, bem como explicações adicionais a

³⁴ HAUSMANN, F. J. *Einführung in die Benutzung neufranzösischer Wörterbücher*. Tübingen: Niemeyer, 1977, p. 112-113.

³⁵ O próprio Welker (2004: 135) cita K. Steyer (Usuelle Wortverbindungen des Deutschen. Linguistisches Konzept und lexikografische Möglichkeiten. *Deutsche Sprache* 28: 101-125, 2000), para quem “as marcas de uso na forma de uma única palavra são cada vez mais criticadas na metalexicografia” – talvez pelo fato de haver pouca explicação quanto aos critérios de apresentação de tais marcas, ou mesmo, o que não é de se menosprezar, pela própria falta de espaço no dicionário (isso, sobretudo, em relação aos dicionários impressos, confeccionados com base em diretrizes editoriais que em grande parte transcendem a atuação do lexicógrafo).

respeito destas em sua “Introdução”. De acordo com Guerra Salas e Gómez Sánchez (2005: 354), tal dicionário “(...) incorpora explicitamente uma lista de marcas pragmáticas (...) [buscando] especificar as possibilidades de uso de algumas expressões e termos em certas circunstâncias”³⁶ (2005: 354). Neste ponto, vale a pena rever as figuras 1 e 2 que mostram, respectivamente, as marcas de uso pragmáticas do *Salamanca* e aquelas que são especificamente ilocucionárias. Além disso, vale a pena, também, retomar o exemplo concreto de marca de uso pragmático-ilocucionária fornecido na mesma parte da Introdução:

siempre

adv. temp. **1** En todo momento o todo el tiempo. *El escritor siempre fumaba puros. Te echaré de menos siempre.* OBSERVACIONES: Normalmente se refiere al periodo de tiempo del que se habla o a uno que se da por consabido. (...) **4** AFIRMACIÓN. Resalta una cosa que se dice o refuerza una afirmación que no admite duda: *Siempre vivirás mejor solo que mal acompañado.*

Ao analisarem o *Diccionario Salamanca*, Guerra Salas e Gómez Sánchez notaram que o rótulo *Afirmación* foi atribuído a verbetes tais como *siempre*, *evidente*, *dos*, *tres* e *luego* (2005: 356). No caso do advérbio *siempre*, p. ex., o *Salamanca* (1996: 1460) informa na microestrutura dessa entrada que tal item lexical “ressalta algo que se diz ou reforça uma afirmação que não admite dúvidas”³⁷, ao lado de um exemplo (“*Siempre vivirás mejor solo que mal acompañado*”) e do rótulo *Afirmación*. Dito de outro modo: para tal dicionário, a ilocução de *siempre* é a de afirmação, ou seja, o ato de fala convencional/ padrão ligado ao proferimento do item lexical próprio da língua espanhola *siempre* é o *ato de afirmar* algo. O mesmo ocorre com o adjetivo *evidente*, cuja microestrutura sugere que tal item lexical é “usado para responder afirmativamente a uma pergunta”³⁸, seguido do seguinte exemplo: “– ¿Crees que va a ganar? – *Evidente*”³⁹ (1996: 687) (e seguido também, naturalmente, da mesma marca de uso *Afirmación*).

Outros dois casos notados no citado dicionário, embora não comentados por Guerra Salas e Gómez Sánchez, são os dos substantivos masculinos *eunuco* e *charloteo*. No caso de *eunuco*, sua ilocução, tal como apontada pelo *Diccionario Salamanca*, é a de *Insulto*⁴⁰ (1996: 686), enquanto que a de *charloteo* é apontada como sendo

³⁶ “(...) incorpora explicitamente un catálogo de marcas pragmáticas (...) [intentando] señalar las posibilidades de uso en ciertas circunstancias de algunas expresiones y términos”.

³⁷ “Resalta una cosa que se dice o refuerza una afirmación que no admite duda”.

³⁸ “Se usa para responder afirmativamente a una pregunta”.

³⁹ Provavelmente “- Você acha que vai ganhar? – *Evidente*”.

⁴⁰ Cujas explicações são as seguintes: “hombre cuyo comportamiento no coincide con la imagen masculina que predomina en la sociedad”.

*Pejorativo*⁴¹ – esta última que, segundo o *Salamanca*, é usada para se referir à “conversa entre duas ou mais pessoas que incomodam outras pessoas”, sendo que o exemplo dado foi o seguinte: “*El profesor dijo a los alumnos que dejaran el charloteo para después de la clase*”⁴² (1996: 307, grifos nossos). Neste último caso, impossível deixar de reconhecer certa confusão entre o significado (semântica) do item lexical em questão e sua ilocução (pragmática)⁴³, algo que os próprios Guerra Salas e Gómez Sánchez (2005: 361-362)⁴⁴ concluíram ao final de seu artigo.

1.5.2. Exemplos de uso

Viu-se com Welker (1.5) que exemplos de uso são informações que podem ser dispostas nos verbetes das entradas. Veja-se, mais uma vez, a entrada *siempre* do *Salamanca*:

siempre

adv. temp. **1** En todo momento o todo el tiempo. *El escritor siempre fumaba puros. Te echaré de menos siempre.* OBSERVACIONES: Normalmente se refiere al periodo de tiempo del que se habla o a uno que se da por consabido. (...) **4** AFIRMACIÓN. Resalta una cosa que se dice o refuerza una afirmación que no admite duda: *Siempre vivirás mejor solo que mal acompañado.*

“*El escritor siempre fumaba puros*”, “*Te echaré de menos siempre*”, “*Siempre vivirás mejor solo que mal acompañado*” são casos de exemplos de uso. Ao final do trabalho (5.1.3 e 5.3.2) são apresentadas possibilidades de inserção de exemplos de uso ilocucionário, bem como a razão pela qual tal inserção não figura entre as sugestões ideais apresentadas (ao menos não no interior da microestrutura).

1.5.3. Notas de uso

Outra possibilidade de inserir informações de natureza pragmático-ilocucionária na microestrutura se dá através de notas de uso dispostas no interior dos verbetes. Tais

⁴¹ “*Pejorativo*”.

⁴² Provavelmente “*O professor pediu aos alunos que deixassem de conversa paralela*”.

⁴³ A confusão reside no fato de que tanto a explicação quanto o exemplo dado pelo *Diccionario Salamanca* referem-se ao conteúdo semântico da palavra, e não à ação realizada por aquele que a proferiu. Isso ficará mais evidente quando for apresentada a TAF (2.1 e 2.3.2). Afora isso, tal foi apontado pela banca examinadora, *insulto* e *peyorativo* não são atos de fala, ainda que o *Salamanca* assim os compreenda.

⁴⁴ “*Por ultimo, se aprecia que en ocasiones hay cierta ambigüedad en cuanto a las marcas otorgadas a las acepciones (es el caso de pejorativo e insulto, por ejemplo)*”.

recursos foram propostos por Duran (2004: 110-111), por considerá-los capazes de “trazer informações valiosas (...), tais como aspectos pragmáticos (...)”. Vale dizer que sua proposta foi dada a partir da consideração do dicionário *Password*, dicionário pertencente à *família* de dicionários híbridos a qual o *PC* pertence.

Reproduz-se, aqui, na íntegra a sugestão feita pela autora:

Necessarily – *adv* **Modalizador [tipo 6]** indica que o falante espera de seu interlocutor um comportamento determinado; ou que o que é dito seja realizado obrigatoriamente; □ **necessariamente**: *Necessarily, all the students will take the test tomorrow. We will necessarily meet in New York by the end of the week.* Obs: pode ser colocado antes ou depois da partícula negativa, com mudança de sentido: *They necessarily won't leave before paying; They won't necessarily leave before paying.* (TOSQUI, 2002. p. 132, grifo meu)

sauf [sof] *prép* salvo, fora, com exceção de, exceto.

A expressão **sauf si** é seguida de indicativo e significa “salvo”, “a menos que”: *ils iront sauf s'il pleut* / eles irão, a menos que chova.

A expressão **sauf que** é seguida de indicativo e significa “mas”, “salvo”: *c'est un bon film, sauf qu'il est trop long* / é um bom filme, mas é muito longo.

Figura 3: sugestão de notas de uso feita por Duran (2004: 110-111).

Os acréscimos feitos pela pesquisadora aparecem sobre o pano de fundo azul e a princípio representam um procedimento viável capaz de tornar possível a inserção de informações mais completas acerca dos atos de fala.

1.6. Medioestrutura ou Remissões

Ainda que as principais divisões de um dicionário sejam, como se viu anteriormente (1.4 e 1.5), a macro e a microestrutura, pode-se apontar uma terceira, talvez não tão imprescindível quanto as outras duas, no entanto capaz de desempenhar uma função importante. Segundo Welker (2004: 177-178), entre a microestrutura e a macroestrutura há aquilo que aponta como medioestrutura, i. e., um sistema de remissões que conduz o usuário da obra a outra parte da mesma e cuja função principal é evitar repetições. Diz Selistre (2009: 196) que a medioestrutura “é a rede de referências cruzadas que facilita a pesquisa de informações relacionadas em diferentes partes do dicionário”. Por sua vez, diz Welker (2004: 178-179) que “os meios usados no sistema remissivo variam de dicionário para dicionário. Muito frequente é a utilização

do verbo *ver*, geralmente abreviado (*v.*), ou setas” – sendo que importante aqui é que seus usos “sejam explicados na introdução”. Citando H. Wiegand⁴⁵, o mesmo autor (2004: 178) diz que as remissões podem estar nas seguintes posições: no verbete, no índice, numa inserção dentro da nomenclatura (remissão interna), num dos textos externos (fora, portanto, da nomenclatura, sendo esta uma remissão externa), em ilustrações.

1.7. Dicionários gerais de língua: monolíngues, bilíngues e híbridos

Segundo Krieger (2008: 170), “embora tenha havido um predomínio de problemáticas próprias aos dicionários monolíngues, as obras bilíngues também compõem o quadro de nossas pesquisas” – a saber, o quadro da pesquisa (meta)lexicográfica. Isso salienta um tipo de critério de diferenciação dos dicionários a partir do número de línguas. De acordo com Duran (2004: 20), tal critério permite caracterizar um dicionário como monolíngue (aquele que apresenta uma só língua), bilíngue (o que coloca duas línguas em contato), trlíngue (o que coloca três línguas em cotato) ou multilíngue (o que coloca mais de três línguas em contato). Definições semelhantes apresentam Hartmann e James (1998: 15 e 95-6).

Como se viu antes (Introdução, I), Borba (2003: 16) concebe o dicionário monolíngue como um tipo de dicionário geral de língua que informa “(...) a estrutura e funcionamento da língua, se possível num sistema bem nítido de notação”. Para Welker (2004: 193), “há vários aspectos em que os dicionários bilíngues diferem muito pouco dos monolíngues (...) sendo a maior [das diferenças] o fato de que, no lugar da definição, são fornecidos equivalentes⁴⁶”. Ou seja, o dicionário bilíngue é aquele que “coloca duas línguas em contato” (DURAN; XATARA, 2007: 312), ou ainda, “(...) obra de referência que trata da equivalência das unidades lexicais de duas línguas” (HÖFLING; SILVA; TOSQUI, 2004: 2).

Schmitz (2001) e Zucchi (2010) ressaltam a importância dos dicionários bilíngues no ensino de línguas e salientam a existência de certa resistência por parte de professores e demais profissionais em relação ao uso desses instrumentos. Na visão de Schmitz (2001: 163), p. ex., uma das possíveis causas para tal é a limitação de tais

⁴⁵ Über die Mediostrukturen bei gedruckten Wörterbüchern. In: Zettersten, A.; Pedersen, V. H. (ed.). *Symposium on Lexicography VII*. Tübingen: Niemeyer, 1996, 11-43.

⁴⁶ Welker (2004: 194-199) discute amplamente, a partir de outros autores, a noção de equivalência em dicionários bilíngues.

dicionários “(...) no que diz respeito ao número de vocábulos arrolados e à má qualidade das definições apresentadas”, mas segundo o próprio tal crítica não serve para limitar o potencial de instrumentos, uma vez que, em primeiro lugar, “este tipo de dicionário apresenta uma variedade de tamanhos”, e, em segundo lugar, inúmeros aperfeiçoamentos podem ser realizados, como p. ex., os “dicionários bilíngues podem ser aperfeiçoados com a consulta e a utilização das definições apresentadas no dicionário padrão” – ou seja, um dicionário bilíngue pode ser constituído a partir de um dicionário monolíngue tomado por base. Outro caminho apresenta Welker (2004: 194), ao sugerir “o aproveitamento de *corpora* paralelos. Trata-se de *corpora* de certos textos e de suas traduções. Se esses *corpora* forem grandes, os lexemas (...) ocorrem frequentemente, e (...) são encontrados adequados em contextos”.

São vários os tipos de dicionários bilíngues. Segundo Welker (2004: 199-200), apoiando-se em Hausmann⁴⁷, esses podem ser:

- de duas direções ou bi-escopal (p. ex., italiano>português e português>italiano); de apenas uma direção ou monoescopal (p. ex., italiano>português ou português-italiano);
- monodirecional (voltado a falantes de uma das duas línguas); bidirecional (voltado a falantes de ambas).

Segundo Duran (2004: 34), os dicionários bilíngues podem ser compreendidos segundo critérios de:

- direcionalidade (monodirecional, se o dicionário apresenta apenas uma direção, p. ex., italiano>português ou português>italiano; bidirecional, se apresenta duas direções, p. ex., italiano>português e português>italiano);
- reciprocidade (recíproco, se atender a falantes de duas línguas; não-recíproco, se atender a falantes de apenas uma língua);
- funcionalidade (apoio à codificação, L1>L2; apoio à decodificação, L2>L1).

Vale dizer que a terminologia adotada neste trabalho é essa tal como apresentada por Duran.

Como visto na Introdução (seção I), Welker (2004: 202) entende por dicionários híbridos os que estão entre os bilíngues e os monolíngues, e cita algumas de suas peculiaridades: “a) a separação das acepções é feita com base num dicionário

⁴⁷ HAUSMANN, F. J. Typologie du dictionnaire général bilingue. In : Figge, U. (ed.). *Portugiesische und portugiesisch-deutsche Lexicographie*. Tübingen : Niemeyer, 1994, 11-33.

monolíngue; b) antes de se fornecer o equivalente, é dada a definição do lema, ou melhor, de suas várias acepções, como no monolíngue”. Por sua vez, Höfling, Silva e Tosqui (2004: 3) dizem que o dicionário híbrido possui como objetivo “(...) anular a oposição entre o uso do monolíngue de LE [língua estrangeira] (útil e recomendado pelos professores) e o uso do bilíngue (preferido pelos estudantes devido à sua praticidade)”. Por sua vez, Duran (2004: 27) diz que

o surgimento dos dicionários bilíngües para aprendizes, também chamados de semibilíngües, híbridos, password ou bilingualizados, ocorreu simultaneamente às publicações de um segundo momento do movimento comunicativo, quando os teóricos passaram a reavaliar várias das premissas da abordagem comunicativa e acabaram por atribuir relevância, dentre outras coisas, ao papel do vocabulário e da língua materna na aquisição de línguas estrangeiras.

Segundo entendimento de Schmitz (2001: 165), que dos sinônimos mencionados acima adota o termo “semibilingue”, tal dicionário é “realmente um avanço no campo de lexicografia e possivelmente este tipo de dicionário venha a substituir o dicionário bilíngue tradicional no futuro”. Uma de suas vantagens, diz, é que esse tipo de dicionário “(...) não apresenta uma série de alternativas tradutórias fora de contexto e sem orações modelos [tal] como se observa no caso dos dicionários bilíngues tradicionais (...)” (SCHMITZ, 2001: 166), razão pela qual diz que o formato do dicionário híbrido “com a inclusão de orações-modelo e definições apresenta melhores condições para um aprendizado mais adequado do que o dicionário bilíngue tradicional (...)”. Entretanto, tal entendimento não é unânime. Duran e Xatara (2005: 48), p. ex., discordam da possibilidade de se considerar os dicionários híbridos como avanços, uma vez que

se o termo “semibilíngüe” já havia sido cunhado anteriormente por Quémada [em 1971], como aponta Marelló (1996) e o “conceito lexicográfico” da bilingualização é, conforme James (2000), um recurso antigo, nem o termo nem o conceito são novos⁴⁸.

Importante dizer que as autoras (2005: 51) consideram esse apontamento relevante pelo fato de ser comum encontrar na bibliografia corrente de Lexicografia autores que usam os termos “semibilíngue” e “dicionário bilingualizado” como sinônimos, enquanto que, de acordo com elas, há uma diferença conceitual entre os dois, pois o segundo refere-se a um dicionário bilíngue composto com base em um monolíngue e o primeiro é um dicionário concebido “originalmente como tal e não baseado em outra obra”. Assim,

⁴⁸ MARELLO, C. Les différents types de dictionnaires bilingues. In: BÉJOINT, H.; THOIRON, P. *Les dictionnaires bilingues*. Louvain-la-Neuve: Duculot, 1996, cap. 2.; JAMES, G. Bilingualisation as a genre. In: JAMES, G. *A history of Tamil dictionaries*. Chennai: Cre-A, 2000, p. 450-458.

apontam que tais termos, tomados ou não como sinônimos, são antigos a ponto de não justificar a noção de que o dicionário híbrido é um avanço (no sentido de novidade).

O objetivo das autoras ao fazerem tal apontamento é o de propor os dicionários híbridos como “dicionários bilíngues pedagógicos”. Primeiramente, Duran e Xatara (2005: 51-52) apresentam as características-chave dos dicionários híbridos, a saber, dicionários que: trazem simultaneamente definições e equivalências; são dirigidos a aprendizes de línguas estrangeiras; na maioria das vezes são bilingualizados, i. e., derivados de obras monolíngues; podem ser utilizados como monolíngues ao se desprezar os equivalentes e quando a metalinguagem é feita na mesma língua das entradas. As autoras (2005: 46) apontam como exemplo prototípico desse dicionário o *Password – English Dictionary for Speakers of Portuguese*, publicado pela Editora Martins Fontes, em 1991. Segundo as autoras (2005: 46), o *Password*

seria um dicionário bilíngüe, monodirecional, monofuncional para decodificação, pedagógico, dirigido a aprendizes de inglês de nível intermediário e avançado, falantes nativos do português, e derivado de dicionário monolíngüe para aprendizes.

Posteriormente, publicações similares a essa destinadas a estudantes de outras línguas foram lançadas pela mesma editora, tais como o *Señas*, destinado a estudantes de espanhol, o *Wahrig*, destinado a estudantes de alemão e o *Parola Chiave*, destinado a estudantes do italiano.

Em seguida, Duran e Xatara (2005: 52) dizem o seguinte:

Acreditamos, contudo, que todas essas características não são incompatíveis com uma tipologia de dicionários bilíngües. Julgamos que a própria definição de bilíngüe deve ter como cerne o fato desse tipo de dicionário apresentar duas línguas em contato, deixando as demais características como passíveis ou não de ocorrerem. Assim, não “engessamos” a definição, pois as características que normalmente ocorrem não são vistas como características que obrigatoriamente têm que ocorrer.

Não é objetivo do presente trabalho discutir sobre o melhor termo a ser utilizado para o que em geral se entende por dicionários híbridos – e que Duran e Xatara optam por chamar de dicionários bilíngues pedagógicos. Pode-se dizer, contudo, que se por um lado concorda-se com Duran e Xatara por apontarem inconsistências de natureza histórica no apontamento dos híbridos (ou semibilingues) como novidades lexicográficas, por outro lado opta-se por manter o termo “híbrido”. Em primeiro lugar, porque se está de acordo com Van Hoof (Introdução, I), para quem todos os dicionários, em algum grau, desempenham função pedagógica – caso se entenda o adjetivo “pedagógico” em sentido amplo, a saber, como conjunto de métodos capazes de garantir

a adaptação de dado conteúdo informativo aos indivíduos para os quais tal conteúdo é destinado. Pois, neste caso, assim como um dicionário híbrido pode desempenhar adequadamente a função pedagógica caso esteja de acordo com as expectativas de um estudante de língua estrangeira, um dicionário de usos da língua pode desempenhar adequadamente a função pedagógica caso esteja de acordo com um estudante de pragmática linguística. Não haveria, portanto, a necessidade do adjetivo “pedagógico”, já que, considerando o presente parágrafo, nota-se que todos os dicionários seriam, em maior ou menor grau, instrumentos pedagógicos⁴⁹.

Em segundo lugar, porque se extraindo o adjetivo “pedagógico” da denominação restaria o nome “dicionário bilíngue”. Entretanto, nem todo dicionário bilíngue apresenta informações variadas a respeito da língua-fonte – isso é algo que se encontra propriamente em monolíngues, sendo que o diferencial dos híbridos é justamente aliar tal característica monolíngue com aquela bilíngue de fornecer equivalências na língua-alvo. Assim, se é verdade que embora todo híbrido seja bilíngue, nem todo bilíngue é híbrido, é também verdade que, embora todo híbrido seja monolíngue, nem todo monolíngue é híbrido. Ou seja, nomear o dicionário híbrido de “bilíngue pedagógico” faz desaparecer sua característica monolíngue – por sinal, ao menos para este trabalho, de grande relevância.

Em terceiro lugar, porque não se considera adequado dizer, como Duran (2004: 90), que “o fato de os bilíngues tradicionalmente não apresentarem definições não significa que não possam trazê-las”, pois, se assim fosse, não haveria distinção entre os dicionários – um dicionário de uso não traz, mas *poderia* trazer informações fonológicas de outras línguas, um bilíngue não traz, mas *poderia* trazer informações acerca de regência verbal, um dicionário de regência verbal não traz, mas *poderia* trazer informações pragmáticas etc. Em outras palavras, as definições de tais dicionários estariam sendo subvertidas.

Neste sentido, o trabalho em questão pretende se debruçar sobre os híbridos pelo fato de esses, por sua própria definição, colocarem duas línguas em contato, sendo que tal contato, embora em geral a partir de equivalências semânticas, não necessariamente se limita a essas, já que tomam por base dicionários monolíngues que, por sua vez,

⁴⁹ Importante dizer que se tem consciência do ramo da Lexicografia conhecido como “Lexicografia Pedagógica” que se debruça, sobretudo, sobre questões relativas aos dicionários voltados a estudantes de LE, cuja legitimidade de existência é inegável e irrefutável. O que se questiona aqui, contudo, é, tão somente, o *argumento* utilizado para justificar o uso do termo *pedagógico* a um determinado tipo de dicionário, desde que se entenda que todo dicionário apresenta, em maior ou menor grau, relevância pedagógica.

trazem informações sobre as várias dimensões linguísticas – o que abre espaço para a inserção de informações de natureza pragmática (especificamente falando, ilocucionária).

1.8. O Parola Chiave (PC)

O híbrido aqui tomado como paradigma de investigação é o *PC* (2012). Este possui as seguintes características: monodirecional, não-recíproco, voltado à decodificação, destinado aos falantes do português-brasileiro (língua-alvo) e derivado de dicionário monolíngue italiano (língua-fonte).

1.8.1. Macroestrutura do PC

O *PC* (2012: índice) é macroestruturalmente composto da seguinte maneira:

1. Apresentação da edição brasileira;
2. Guia gráfico para consulta;
3. Abreviaturas e símbolos gráficos;
4. Dicionário;
5. Gramática de uso da língua italiana;
6. Verbos irregulares;
7. Como se fala, como se escreve;
8. Glossário português-italiano.

De acordo com o que se viu no subcapítulo 1.4.1, os itens 1-3 compõem o que se pode chamar de material externo anteposto. Já os itens 5-8 compõem o que se pode chamar de material externo posposto. Por fim, o item 4 nada mais é do que a nomenclatura, que abrange o que se poderia chamar de material interposto

Na Apresentação (1), o editor do dicionário explicita que tal dicionário foi composto de uma base monolíngue constituída pelo *Dizionario italiano per stranieri*, da Editora Giunti. Em seguida, foi sublinhado que “o *Dizionario italiano* foi enriquecido com o acréscimo, ao final de cada definição, de um termo equivalente, em português, à palavra ou expressão definida em italiano”.

Em seguida é fornecido um guia gráfico para consulta (2), indicando todas as partes que compõem os verbetes. Já na seção (3), “abreviaturas e símbolos gráficos”, são fornecidas definições de cada um dos símbolos especiais, a saber:

Figura 4: lista de símbolos gráficos do PC (2012)⁵⁰.

Em relação aos materiais pospostos, a seção (5) apresenta informações gramaticais variadas da língua italiana, tais como frase e oração, verbos, substantivos, preposições etc. Vale citar, desta seção, a subseção intitulada “Usos e regras pragmáticas”, onde o dicionário aborda fórmulas de cortesia, formas de cumprimento e apresentação, marcadores de discurso, interjeições etc. Nesta subseção, os atos de fala não são indicados diretamente, mas pode-se notar sua presença implícita em alguns momentos. Exemplos:

- na definição de regras pragmáticas (2012: 1010, grifos nossos), o dicionário informa que estas são as regras “regem os *usos* da língua: estão fortemente ligadas ao *contexto de comunicação* e aos *sentidos das ações*, presentes nas escolhas linguísticas”;
- ao abordar formas de tratamento (2012: 1011), o dicionário informa modos de realizar ações tais como pedidos, convites, ordem (“- *Mi daresti, per cortesia, un bicchiere d’acqua?*”);
- ao abordar “turnos de fala” (2012: 1012), como instrumentos linguísticos utilizados para evitar sobreposição de fala (“- *Senti una cosa...*”/ “- *Dimmi*”);
- ao abordar formas de cumprimento e de apresentação (2012: 1013): *ciao!*, *buongiorno* etc;
- ao abordar interjeições (2012: 1016), como palavras invariáveis que servem para “expressar ordens, incentivos etc”.

⁵⁰ Ao mencionar tais símbolos em texto corrido, o símbolo usado para uso figurado foi “♣” e para sinônimo foi “§”. Isso porque os símbolos usados pelo PC não existem no *software* de texto Word.

Já na seção (6) é dada uma lista de verbos irregulares, na (7) informações de natureza fonética e morfológica, entre outras, e na (8) um miniglossário português-brasileiro>italiano.

1.8.2. Microestrutura do PC

Viu-se anteriormente (1.5) que a microestrutura é o conjunto de verbetes. Tome-se como exemplo o verbete da entrada *promettere*, tal como fornecido pelo PC (2012: 587)

prométtere

v.tr.[conjugato come *mettere*]. **1** Impegnarsi a fare o a dare qualcosa o a comportarsi in un certo modo: *mi ha promesso di venire, di aiutarmi, di non disturbarmi più; promettere un regalo, un lavoro*. **2** ♣ Far sperare, far prevedere: *un cielo che promette bel tempo; il suo atteggiamento non promette niente di buono; un ragazzo che promette bene*. □ **Prometer**.

Segundo informações do gráfico para consulta (2012), tal verbete pode ser assim decomposto:

- **prométtere**: “em negrito a palavra com indicação da sílaba tônica”;
- *v.tr.*: “categoria gramatical”;
- [conjugato come *mettere*]: “conjugação do verbo”;
- **1** Impegnarsi a fare o a dare qualcosa o a comportarsi in un certo modo: *mi ha promesso di venire, di aiutarmi, di non disturbarmi più; promettere un regalo, un lavoro*: primeira acepção seguida de exemplos de uso;
- **2** ♣ Far sperare, far prevedere: *un cielo che promette bel tempo; il suo atteggiamento non promette niente di buono; un ragazzo che promette bene*: segunda acepção seguida de exemplos de uso;
- ♣: “O símbolo do trevo indica o uso figurado”;
- □ **Prometer**: “*parola chiave*: tradução indicativa”.

1.8.3. Medioestrutura do PC

A medioestrutura, como se viu (1.6), corresponde ao sistema de remissivas. O PC possui esse recurso, como se nota ao visualizar o verbete *sonare* (2012: 734):

sonàre vedi suonàre.

Nota-se que o *vedi* é o sinal indicador de remissiva que pede ao leitor interessado no item lexical *sonare* que pesquise o verbete *suonare*.

Ao que parece, todas as remissivas do *PC* estão na microestrutura e conduzem o leitor à macroestrutura (ou seja, para outra entrada), não existindo assim alguma⁵¹ que faça remissão aos demais materiais externos do dicionário.

⁵¹ Alguma que se tenha verificado.

2. TEORIA DOS ATOS DE FALA (TAF)

Neste capítulo, define-se a TAF como teoria segundo a qual as palavras são ações, o que significa que, segundo tal perspectiva, a investigação e compreensão da linguagem se dá em meio ao contexto interacional e/ ou comunicativo concreto em que aparece, de modo a se tornar possível reconhecer qual a *força ilocucionária* de dado proferimento – ou seja, *aquilo que se fez ao se ter dito algo*. Assim, a TAF é apresentada como objeto de estudo da Pragmática Linguística, disciplina que examina a língua em contextos de uso, e também como elemento da *competência pragmática*, aquela cuja posse permite ao falante/ ouvinte reconhecer (e usar) regras linguísticas convencionalmente estabelecidas para o desempenho de determinadas ações. Por fim, apresenta-se a TAF de Austin (1975[1990]) como paradigma de pesquisa, em função, sobretudo, da associação que o filósofo fez entre atos ilocucionários e verbos ilocucionários e do seu critério para reconhecimento de atos ilocucionários – o verbo na primeira pessoa do singular do presente do indicativo na voz ativa.

2.1. O agir pela linguagem

A TAF é aquela que permite a compreensão da linguagem como forma de ação (KOCH, 1992: 19). De acordo com tal abordagem, a linguagem, além de ser um meio utilizado para expressar pensamentos, descrever o mundo e transmitir mensagens/ informações, é também usada para realizar ações diversas, sendo que estas se constituem pelo próprio dizer (SBISÀ, 2009[1989]:45).

De acordo com Bianchi (2008: 57), postular a linguagem como ação significa dizer que quando expressões linguísticas são proferidas novos fatos são criados, modificando a realidade. Deste modo, as ações devem ser entendidas, tal como foram por Sbisà (2009[1989]: 40), como disposições que “operam transformações em seu contexto”⁵², em função da emissão de determinado proferimento, e, logo, pensar nos atos de fala como transformação contextual conduz à noção de contexto como instante “construído pelo ato linguístico” (SBISÀ, 2009[1989]: 45)⁵³. Em outras palavras, a

⁵² “(...) che operano trasformazioni nel loro contesto”.

⁵³ “(...) trattandosi piuttosto di effetti di senso, che di conseguenze materiali, essi riguarderanno piuttosto il contesto in quanto costruito dall’atto linguistico, che il contesto in quanto dato precedentemente ad esso”.

realização de um determinado ato de fala por intermédio da emissão de algum expediente linguístico constitui o contexto interacional e/ ou comunicativo em questão.

De acordo com Caffi (2011: 35), todo proferimento é um ato, e, logo, seja qual for o expediente linguístico utilizado para sua emissão, todo proferimento é passível de ser analisado em termos da TAF – o que significa *reconhecer qual é a ação constituída pelo proferimento* em questão. Afirmações, promessas, pedidos, cumprimentos, condenações, apostas são exemplos de ações “que o falante cumpre com o objetivo de produzir no interlocutor, através do uso da linguagem, uma reação, uma mudança de estado mental ou um comportamento” (ANDORNO: 2009: 62)⁵⁴. “Afirmo que é isso mesmo”, “Prometo!”, “Poderia fechar a porta?”, “Olá!”, “Condeno o réu à pena máxima”, “Aposto imediatamente!” são exemplos de proferimentos que ao serem emitidos realizam, respectivamente, as ações acima elencadas. Em outras palavras, tais proferimentos, a partir dessa abordagem, indicam *atos de fala, ilocuções*.

Uma importante noção que se deve ter em mente logo de início é que um ato de fala/ ilocução/ ato ilocucionário não equivale ao significado de uma palavra/ sentido de uma sentença. Como se viu no parágrafo anterior, dizer “Prometo!” significa, *convencionalmente* falando, realizar uma promessa – mas *não necessariamente*. Nada impede que um sujeito x, p. ex., em determinado contexto, diga “Prometo!” ao sujeito y não de modo a constituir uma promessa, mas sim de afirmar/ reafirmar uma promessa realizada previamente: “– Prometo”. “– Promete mesmo?” “– *Prometo!*”. Ou então, que x diga “Poderia fechar a porta?” não para constituir um pedido, mas para realizar uma ordem – ainda que fórmulas de cortesia tenham sido usadas (no caso, o uso do verbo no futuro do pretérito): “– Feche a porta, por favor?” “– Não.” “– *Poderia fechar a porta?*” Em outras palavras: o proferimento de determinada palavra/ sentença não necessariamente se dá de modo a realizar um determinado ato de fala, i. e., aquele ato que *convencionalmente* se espera a partir do proferimento de determinada palavra/ sentença, sendo que só o contexto de proferimento é capaz de fornecer elementos capazes de permitir o reconhecimento do ato de fala em questão e a constatação de se o ato realizado foi aquele *convencionalmente* esperado ou não. Não à toa a TAF é objeto de estudo da Pragmática Linguística.

⁵⁴ “(...) *che il parlante compie con l'intento di produrre nell'interlocutore, attraverso l'uso del linguaggio, una reazione, un cambiamento di stato mentale o un comportamento*”.

2.2. A TAF como objeto de estudo da Pragmática Linguística

De acordo com Levinson (2007: 288), a TAF é “um dos fenômenos centrais que qualquer teoria pragmática geral deve explicar”. Para este autor (2007: 64), Pragmática⁵⁵ é a disciplina linguística cujos estudos, partindo de sequências de proferimentos considerados em conjunto com algumas suposições de fundo a respeito do uso linguístico em dado contexto, admitem extrair “(...) inferências altamente detalhadas a respeito da natureza das suposições que os participantes estão fazendo e dos fins para os quais as enunciações estão sendo usadas”. Caffi (2011: 19), por sua vez, diz que a Pragmática se ocupa da “língua em situação”, i. e., “ocupa-se dos usos da língua em contextos; é o estudo dos usos linguísticos contextualizados”⁵⁶, o que permite a compreensão das práticas comunicativas cotidianas (2011: 15). Para Gutierrez (2008: 195), a Pragmática é a disciplina linguística que se ocupa do estudo da língua em relação às necessidades de comunicação.

Por abarcar todas as anteriores, a definição tomada aqui é a de Bianchi (2008: 10-11), para quem a Pragmática Linguística é a disciplina que se ocupa: do uso da linguagem; daquilo que um falante comunica; do contexto; do significado em contexto; do significado nas interações sociais; da distância entre os interlocutores.

A autora (2008: 11) ainda sugere que a Pragmática Linguística pode tomar duas direções: (1) aquela que considera a influência do contexto sobre a linguagem (o exame da linguagem deve considerar o contexto discursivo em questão); (2) aquela que considera a influência da linguagem sobre o contexto (o exame da linguagem deve considerá-la como modificadora do contexto discursivo em questão). Ainda que a presente investigação seja de natureza metalexigráfica, e não pragmática, como se viu anteriormente (1.2), pode-se dizer que este trabalho se alinha à segunda possibilidade, já que entende os atos de fala como modificadores de seu contexto de proferimento. Em (2), como esclarece Bianchi (2008: 55), é acentuada

a *dimensão social* da linguagem, e em particular a variedade dos usos discursivos de frases da linguagem natural: afirmações, ordens, perguntas, ameaças etc. Nesta perspectiva, falar significa agir: todo enunciado serve para cumprir um ato, regulado por normas, convenções ou costumes; a

⁵⁵ A obra de Levinson (2007) expõe e examina, além da TAF, outras teorias/ perspectivas de análise pragmática da língua, como a *dèixis* e a *implicatura conversacional*.

⁵⁶ “(...) *la pragmatica si occupa della lingua in situazione. Si occupa degli usi della lingua in contesto; è lo studio degli usi linguistici contestualizzati*”.

linguagem como tal vem concebida no mesmo nível de uma instituição social⁵⁷.

2.2.1 A TAF e o conceito de contexto

Cabe compreender melhor, neste momento, o que se entende por contexto⁵⁸. Viu-se anteriormente (2.1) que para Sbisà o contexto é construído pelo proferimento, construção equivalente ao ato de fala realizado. Segundo Bianchi (2008: 24), existem dois modos de avaliação de um contexto: (i) o semântico ou situação objetiva de proferimento, que se limita a “determinar o conteúdo de um pequeno número de variáveis e, particularmente, fixar a identidade do falante/ interlocutor, o tempo e o lugar de proferimento etc”⁵⁹; (ii) o pragmático, que corresponde à “rede de crenças, intenções, atividade dos interlocutores, e contribui para a determinação de suas intenções comunicativas”⁶⁰.

O contexto pelo viés semântico engloba o exame das condições de verdade de uma proposição proferida. P. ex., para saber se a sentença “*Ho delle credenze*” (em português literal, possui duplo significado: “Possuo alguns armários” ou “Tenho algumas crenças”) é verdadeira, é preciso saber: quem a proferiu (foi de fato uma pessoa que as possui?), quando a proferiu (tal pessoa ainda as tinha quando proferiu?) etc. Entretanto, diz Bianchi (2012[2003]: 163), o significado convencional de uma sentença “não é suficiente para determinar completamente suas condições de verdade”, razão pela qual se torna necessário analisá-la pragmaticamente – e isso em dois momentos possíveis: o momento pré e o momento pós-semântico.

O momento pré-semântico é aquele em que se opera a desambiguação do proferimento linguístico em questão. P. ex., a palavra *credenze* do caso acima faz referência às crenças/ opiniões/ visões/ concepções de dada pessoa ou ao seu mobiliário? Tal desambiguação, segundo Bianchi (2008: 53), é “um procedimento

⁵⁷ “Ad essere sottolineata è ora la dimensione sociale del linguaggio, e in particolare la varietà degli usi discorsivi delle frasi del linguaggio naturale: affermazioni, ordini, domande, minacce ecc. In questa prospettiva parlare significa agire: ogni enunciato serve a compiere un atto, regolato da norme, convenzioni o consuetudini (...)”.

⁵⁸ Van Dijk (1980) é um autor que procurou analisar a fundo o conceito de contexto. Para o presente trabalho, contudo, a proposta de Bianchi é suficiente.

⁵⁹ “(...) determinare il contenuto di un piccolo numero di variabili e, in particolare, a fissare l'identità del parlante e interlocutori, il tempo e il luogo del proferimento, e così via”.

⁶⁰ “(...) corrisponde alla rete di credenze, intenzioni, attività degli interlocutori, e contribuisce alla determinazione delle loro intenzioni comunicative”.

pragmático”, pois a desambiguação “ocorre antes que a semântica entre em campo”⁶¹, sendo que assim acontece porque os problemas de ambiguidade linguística não podem ser resolvidos a partir de um sistema fixo de regras semânticas preestabelecido. Assim, somente considerando o contexto pragmático (interacional e/ ou comunicativo) de proferimento se poderá fixar *credenze* como se referindo especificamente, p. ex., ao mobiliário – contexto hipotético em que o sujeito x telefonou ao sujeito y após ter visto o seu anúncio segundo o qual y manifestara interesse em vender alguns de seus objetos domésticos, quando então x lhe perguntou se haviam mobiliários à venda, obtendo uma resposta positiva: “*Ho delle credenze*”.

Já o momento pós-semântico é aquele em que se pode reconhecer o que está implícito a dado proferimento. Não existem regras semânticas previamente constituídas capazes de avaliar *in abstracto* se um proferimento como *Ho delle credenze* é um ato de fala de afirmação ou de ironização, razão pela qual a autora sublinha que tal informação só é possível de ser obtida examinando o contexto pragmaticamente. Por outras palavras, a autora (2008: 53-54) diz que o contexto pós-semântico “nos introduz aos casos de influência da palavra sobre o contexto – em outras palavras, dos modos em que os falantes modificam a situação discursiva”⁶². Em suma, abrange o exame dos atos de fala enquanto modificadores/ construtores do contexto de proferimento. Na visão de Bianchi (2008: 53), portanto, tal concepção de contexto pode ser expressa pelo seguinte esquema: *pragmática* > *semântica* > *pragmática*, entendendo a pragmática como situada antes e depois da semântica, fornecendo o *input* da semântica e dela recebendo, posteriormente, o *input* que lhe permitirá avaliar pragmaticamente o implícito. Este é o momento em que se pode verificar, tal como expresso no subcapítulo 2.1., se “Prometo!” constitui uma promessa ou uma afirmação.

2.2.2. A TAF e os conceitos de comunicação e interação

Note-se que o conceito de contexto é irremediavelmente ligado aos conceitos de comunicação e/ ou interação. Isso porque ao se pensar em alguém proferindo determinado expediente linguístico é natural pensar que assim o fez dirigindo-se a

⁶¹ “*La disambiguazione è una procedura pragmatica (...), è una procedura pre-semantica, che avviene prima che entri in gioco la semantica*”.

⁶² “*L’esame degli usi post-semantici del contesto (...) ci introduce al tema dell’influenza della parola sul contesto – dei modi, in altri termini, in cui i parlanti modificano la situazione di discorso*”.

alguém que *coparticipava* desse contexto de proferimento. Logo, se x diz a y “Prometo que lhe trarei o livro” é de se supor que x desempenhou, diante de y, a ação de prometer – x comunicou/ expressou a y o seu comprometimento com ele, ao constituir o contexto de *promessa*. Segundo tal interpretação, considera-se contraintuitivo negar que tenha havido certa interação entre x e y, ainda que não se saiba qual foi a reação de y em relação à promessa feita por x. Logo, contexto comunicativo e contexto interativo apresentam-se como noções interrelacionadas neste trabalho.

Tal visão, contudo, não é unânime. De acordo com Richards e Rodgers (2001), existem ao menos três pontos de vista teóricos sobre a linguagem: a estrutural, que vê a língua como uma estrutura, sistema internamente coeso e passível de compreensão fora de qualquer contexto usual; a funcional, que vê a língua como veículo para expressão de significados; a interacional, que vê a língua como veículo para a realização de relações interpessoais.

Assim, a princípio o presente trabalho poderia ser alinhado exclusivamente à perspectiva funcionalista, pois segundo Andrade (2014: 33-34) para o funcionalismo o conhecimento de regras linguísticas “vem do uso, da prática e, portanto, está relacionado com a ideia de que é possível ver e descrever o mundo como ele é, isto é, a ideia de realismo está por trás dessa visão”. A abordagem comunicativa, assim, segundo a qual a língua é usada, sobretudo, para a comunicação, ancora-se no funcionalismo, uma vez que para haver comunicação é necessário que duas pessoas entendam o significado dos sons articulados que emitem – significado este que, de certo modo, seria invariável. Em casos como esse, proferir “Prometo que...” necessariamente implica realizar uma promessa, “Ordeno que...” necessariamente implica realizar uma ordem, “Olá” necessariamente implica em um cumprimento etc. Assim sendo, e considerando-se a pesquisa de natureza metalexigráfica que aqui se pretende, a perspectiva funcional se mostraria mais afeita à presente pesquisa, uma vez que não se pretende investigar ilocuções segundo contextos particulares de ocorrência de modo a estabelecê-las conforme particularidades de cada um dos casos de interação, mas sim investigar ilocuções *convencionalmente* estabelecidas de modo a dicionarizá-las enquanto padrões de uso da língua geral voltadas para a comunicação.

O problema é que isso parece contradizer o que se disse anteriormente (2.1), a saber: que determinado ato de fala não equivale a determinado expediente linguístico, pois as ilocuções não se reduzem a manifestações linguísticas, ainda que estas últimas sejam necessárias para sua realização. Entretanto, tal contradição é aparente, pois a

noção de que “Prometo!” pode realizar tanto o ato de fala de promessa, como o de afirmação, como o de qualquer outro, não implica a desconsideração de que, segundo usos *convencionalmente* estabelecidos do português-brasileiro, “Prometo!” é associado ao ato de fala de promessa. Em resumo: nada impede que alguém diga “Prometo!” para afirmar, mas é *convencionalmente* esperado que “Prometo!” seja usado para constituir uma promessa – e tal convenção é o que se pretende capturar aqui, de modo que as ilocuções possam ser dispostas em dicionários híbridos. E levando-se em conta que, como se disse parágrafos acima, é contraintuitivo negar que tenha havido certa interação entre x e y, quando x disse “Prometo!” a y (mesmo que não se saiba qual a reação de y diante de tal promessa), uma vez que, se uma promessa foi feita, foi feita *por alguém para alguém*, razão pela qual comunicação e interação estão intimamente relacionadas⁶³.

Kerbrat-Orecchioni, contudo, é uma autora que prega a distinção rígida entre o que é funcional e o que é interacional – aliando-se a esta última perspectiva. Segundo ela (2005: 69), tal divisão se deve ao fato que os atos de linguagem, tal como considerados por Austin e Searle, segundo sua leitura, “aparecem como entidades abstratas e isoladas, i. e., destacadas (...) de seu contexto de atualização e dos outros atos que podem precedê-los e segui-los no encadeamento discursivo”. Ainda com a autora (2005: 70-71), a teoria-padrão (segundo ela, a de ambos os “Johns”) prega que “um locutor transmite a um ouvinte sua intenção ilocutória por meio de um enunciado dotado de uma ‘força’ apropriada” – algo que para ela é equivocado em razão de isso desconsiderar o “encadeamento discursivo”, já que “toda conversação (...) se apresenta como uma arquitetura complexa e hierarquizada” (2005: 76), propondo assim a distinção entre o elemento ilocutório de um proferimento (valor que possui fora do contexto) e o elemento interativo de um proferimento (valor que possui dentro do contexto). O ilocutório, assim, seria característica do paradigma funcional-comunicacional, enquanto o interativo, como o termo já indica, do paradigma interacional. A autora, então, nomeia a sua perspectiva de *interacionista*, pois para ela falar, mais do que agir, é “*interagir*, no sentido de que, ao longo do desenrolar de uma troca comunicativa qualquer, os diferentes participantes exercem entre eles *influências* de natureza diversa”.

⁶³ Razão pela qual mantém-se a fórmula “interação e/ ou comunicação”.

Cabem aqui algumas observações. Em primeiro lugar, as teorias de Austin e Searle não são idênticas – melhor dizendo, não é unânime a visão de que a teoria de Searle automaticamente segue a de Austin. Para Searle (1981[1969]: 35), quando um sujeito x diz algo ele realiza três tipos de atos: atos de enunciação (ato de enunciar palavras), atos proposicionais (ato de se referir a algo e/ ou predicar algo de alguma coisa) e atos ilocucionários (ato de prometer, ordenar, perguntar etc). De acordo com tal tipologia, não seria de todo equivocado interpretá-la como que fazendo pouca ou nenhuma referência ao interlocutor do falante – provavelmente, a interpretação feita por Kerbrat-Orecchioni. Entretanto, o próprio Searle (1981[1969]: 41 e 44) diz que “os atos de fala são realizados não por palavras, mas pelos falantes que as pronunciam” e que “é o contexto que permitirá fixar a força ilocucional da enunciação”. Logo, parece claro que Searle não desconsidera o contexto de proferimento.

Há, entretanto, um ponto pelo qual se dá razão à Kerbrat-Orecchioni: Searle (1981[1969]: 11) no início de seu livro diz que tal não é um ensaio “(...) sobre as línguas, como o francês, o inglês ou o swahili, mas sobre a linguagem”. Em outra passagem, Searle (1981[1969]: 22) diz:

considerando os elementos linguísticos, podemos propor caracterizações linguísticas que não se relacionem com emissões particulares mas que tenham um caráter geral, derivado do fato de que os elementos são regidos por regras.

A interpretação que Rajagopalan (2010d: 48-49) faz de Searle é a seguinte: este desenvolve suas teses a partir de critérios de classificação que, partindo obviamente de sua língua materna, o inglês, estão “acima das línguas particulares, isto é, [são] universais na acepção plena do termo”. Por outras palavras, na ótica de Rajagopalan, Searle “reconhece que existem importantes diferenças entre as línguas (...), mas prefere ignorá-las, como se tais diferenças estivessem relacionadas apenas a restrições de idiomatidade” (2010b: 108) – cuja desconsideração resulta nessa “teorização de cima para baixo”. Ainda com o mesmo autor (2010d: 48-49), tais posturas diferem das de Austin, pois este último “tinha uma posição fechada em relação à tese de que os atos ilocucionários se definem em relação a línguas específicas”. E mais, como se nota na seguinte passagem (2010b: 106):

Talvez em nenhum outro ponto a diferença de Austin e Searle seja mais evidente do que na maneira como eles fizeram generalizações. Contra o hábito típico de Austin de interrogar cada caso inúmeras vezes, de tentar testar diferentes matizes e nuances de significado de cada item lexical, de se voltar inúmeras vezes ao que uma pessoa comum normalmente diria, a abordagem de Searle consiste numa teorização de cima para baixo.

Supondo, p. ex., que Kerbrat-Orecchioni tenha realizado uma análise exaustiva da obra de Searle (e dada sua autoridade, é possível que a tenha feito) e desse modo tenha razões capazes de justificar sua visão acerca de Searle, ainda assim valeria se perguntar se as mesmas justificariam sua visão acerca de Austin. Isso porque, a princípio, a crítica que a autora faz ao filósofo inglês parece problemática pelo fato de que ele, além de deixar claro ao longo de sua obra que os proferimentos devem ser analisados em contexto, oferece uma tipologia que claramente evidencia tal visão. Para Austin (1990[1975]) (mais detalhes em 2.3.2), quando um sujeito x diz algo, ele realiza três tipos de atos: o ato locucionário (proferir palavras de determinada língua segundo determinadas regras gramaticais de modo a se referir a determinado objeto/ evento do mundo), o ato ilocucionário (realizar ações tais como prometer, ordenar, afirmar etc) e o ato perlocucionário (influenciar/ convencer/ causar determinado efeito *em alguém* em função do proferimento). Neste caso, parece indubitável que só se poderia causar efeitos em alguém caso houvesse *alguém capaz de sofrer os efeitos* e, se há tal pessoa, e se tal pessoa é afetada pelo proferimento de outra, seria contraintuitivo pensar não ter havido neste momento um contexto interativo (tanto quanto comunicativo) – mesmo que se tenha optado por não avaliar qual foi o efeito causado.

Ao que parece, Kerbrat-Orecchioni exclui a teoria de Austin do paradigma interacional apenas pelo fato de ele ter focado seu interesse sobre as ilocuições (os atos que são realizados ao serem proferidos pelo falante), deixando de lado as perlocuições (atos que são recebidos pelo ouvinte). Porém, “abrir mão” de investigar as segundas, tal como faz Austin, não significa que sua investigação seja abstrata (pois, novamente, um sujeito x só pode prometer, pedir, ordenar, afirmar, negar algo caso haja alguém para o qual ele se dirige). A abordagem austiniana só poderia ser considerada abstrata no sentido de ele não ter se valido de *corpora* autênticos de sua língua natural (afinal, era filósofo), mas não no sentido de que ele desconhecesse ou menosprezasse a relação falante-ouvinte. Tanto é que Sbisà (2009[1989]: 40), p. ex., evidencia que Austin considerou a recepção (perlocução) de um proferimento como importante, entendendo-a como parte da interação verbal.

Logo, concorda-se com Kerbrat-Orecchioni por ela propor que os atos de fala se desenvolvem em uma cadeia em que x, ao dizer algo a y, realiza determinada ação, esta recebida por y de determinado modo, em função da qual diz algo a x de modo a realizar outra ação, recebida por x de determinado modo, e assim sucessivamente; entretanto, ao

que parece a diferença entre ela e Austin reside tão somente no recorte contextual, ou, melhor dizendo, na *extensão do recorte* contextual que cada um faz (se Austin analisa a realização de *um* ato de fala em um discurso concentrando-se na *ilocução*, Kerbrat-Orecchioni analisa a realização de *vários* atos em um mesmo discurso, levando em conta a *ilocução* e a *perlocução*). Em suma: interação e comunicação coocorrem em ambos os casos.

Em conclusão, considera-se que o que é funcional-comunicativo e o que é interacional diferem apenas em termos da extensão do recorte contextual a ser examinado. A princípio, Bazzanella (2009: 105) mostra-se de acordo com a presente reflexão, pois ao falar da perspectiva funcional parece não distingui-la da perspectiva interacional:

a linguagem é, antes de tudo, um instrumento de interação social entre indivíduos (...) [em que] as expressões linguísticas não são objetos formais arbitrários, mas suas propriedades são não apenas sensíveis, mas codeterminadas pelas características pragmáticas da interação verbal. A estrutura linguística é portanto estreitamente correlata às necessidades comunicativas⁶⁴.

Neste caso, conclui-se que a presente investigação, embora se debruce tão somente sobre as ilocuições tais como aparecem em *corpora*, prescindindo-se assim das perlocuções (já que o contexto de proferimento fornecido por *corpora* escritos são limitados, sendo difícil – e em muitos casos, impossível – reconhecer a recepção de dada ilocução) e mostrando-se desse modo mais próximo ao funcionalismo, nem por isso deixa de reconhecer o contexto como, além de comunicativo, também interativo.

2.2.3 A TAF e o conceito de competência comunicativa

Por sua vez, tanto interação quanto comunicação conduzem ao conceito de competência comunicativa. Uma vez compreendendo competência como “o conjunto dos conhecimentos, capacidades e características que permitem a realização de acções (sic) (QECR, 2001: 29)”, por competência comunicativa entende-se aquela que permite a um indivíduo “agir utilizando especificamente meios linguísticos” (QECR, 2001: 29).

Esta competência difere da mera competência linguística, ou seja,

⁶⁴ “(...) il linguaggio è prima di tutto uno strumento di interazione sociale tra individui (...). Le espressioni linguistiche non sono oggetti formali arbitrari, ma le loro proprietà sono non solo sensibili, ma co-determinate dalle caratterizzanti pragmatiche dell’interazione verbale. La struttura linguistica è quindi strettamente correlata ai bisogni comunicativi (...)”.

conhecimento tácito da estrutura da língua, i. e., conhecimento que geralmente não é consciente ou disponível para um relato espontâneo, mas necessariamente implícito ao que o falante-ouvinte (ideal) quer dizer (HYMES, 1979 [1971]: 271)⁶⁵.

Por outras palavras, segundo Levinson (2007: 39) a competência linguística é o “conhecimento de uma língua idealizada, distanciada (especialmente) da irregularidade ou do erro e da variação” – ao que se pode associar, sobretudo, o conhecimento sintático-gramatical e fonológico da língua.

Segundo Hymes (1979[1971]: 281)⁶⁶, “existem muitos setores da competência comunicativa, dos quais a gramática é um”, o que permite a consideração de que a competência linguística está incluída no conceito de competência comunicativa – para alguns em conjunto com a competência sociolinguística, esta referente “às condições socioculturais do uso da língua” (QECR, 2001: 35), e com a competência pragmática, referente “ao uso funcional dos recursos linguísticos (produção de funções linguísticas, actos de fala)” (QECR, 2001: 35).

Entretanto, tal posição não é unânime – provavelmente, tal como diz Levinson (2007: 35), por ser difícil delimitar uma fronteira clara entre fenômenos sociolinguísticos e pragmáticos. Note-se, p. ex, a seguinte definição de competência pragmática como

conjunto de regras internalizadas de como usar a linguagem de modo sócio-culturalmente apropriado, tendo em conta os participantes em uma interação comunicativa e características do contexto em que a interação ocorre (CELCE-MURCIA; OLSHTAIN, 2000: 20 *apud* GUTIERRES, 2008: 196)⁶⁷,

segundo a qual é reduzida à competência sociolinguística. Outra definição de competência pragmática é a de Bachman (1991: 42), para quem é

o conhecimento necessário, acrescido à competência organizacional, para produzir e compreender discursos apropriadamente. Especificamente, isso inclui a competência ilocucionária, ou conhecimento de como usar atos de fala, e competência sociolinguística, ou conhecimento das convenções sociolinguísticas que governam o uso da linguagem⁶⁸.

⁶⁵ “*Linguistic competence is understood as concerned with the tacit knowledge of language structure, that is, knowledge that is commonly not conscious or available for spontaneous report, but necessarily implicit in what the (ideal) speaker-listener can say*”.

⁶⁶ “*There are several sectors of communicative competence, of which the grammatical is one*”.

⁶⁷ CELCE-MURCIA, M. & OLSHTAIN, E. *Discourse and context in language teaching. A Guide for Language Teachers*. CUP, 2000.

⁶⁸ “*The knowledge necessary, in addition to organizational competence, for appropriately producing or comprehending discourse. Specifically, it includes illocutionary competence, or the knowledge of how to perform speech acts, and sociolinguistic competence, or the knowledge of the sociolinguistic conventions which govern language use*”.

Ou seja, com ele se tem o inverso: a competência pragmática engloba a competência sociolinguística.

Como não é objetivo deste trabalho adentrar em tais pormenores, as posições de Hymes e do QECR, segundo os quais a competência pragmática é um subconjunto da competência comunicativa⁶⁹, estando no mesmo plano das competências linguística e sociolinguística, serão adotadas. Desse modo, entende-se que tais competências estão intimamente relacionadas, sendo que ambas (em conjunto com a competência linguística) são necessárias para a compreensão e realização da “comunicação” humana.

2.3. A TAF de Austin

A TAF foi explicitamente desenvolvida⁷⁰ por Austin⁷¹ e exposta em *How to Do Things with Words* (doravante HTD), obra póstuma de 1962 organizada a partir do material que Austin utilizou para ministrar suas conferências em Harvard, em 1955, como parte das *William James Lectures* – e que se assemelham ao material usado em suas conferências proferidas entre 1952 e 1954 em Oxford, chamadas *Words and Deeds* (Palavras e Ações)⁷². A compilação foi organizada pelo filósofo inglês James O. Urmson com a colaboração de outro filósofo inglês, Geoffrey J. Warnock. Vale dizer que a segunda edição de HTD, publicada em 1975, contou com a colaboração e revisão da filósofa italiana Marina Sbisa e, no Brasil, esta última foi traduzida pelo filósofo Danilo Marcondes Filho por *Quando dizer é fazer: palavras e ações* (1990[1975]).

Segundo Rajagopalan (2010a: 242), os atos de fala passaram a interessar à Linguística “(...) a partir da leitura de [John] Searle, [quando então] foi possível acoplar toda a investigação austiniana no terreno dos atos de fala ao modelo de gramática em franca ascensão (...), a gramática gerativo-transformacional (...)” (2010a: 246-7).

⁶⁹ Gutierrez (2008: 196) diz que a comunicação eficaz ocorre “quando os falantes compartilham conhecimento, crenças e pressuposições, além de regras específicas de sua conversação e de seu contexto”.

⁷⁰ Embora a “paternidade” de tal teoria seja algo em disputa, devido à obra de Adolf Reinach (1883-1917), intitulada *Die apriorischen Grundlagen des bürgerlichen Rechtes* (1913) (*Fundamentos a priori da lei civil*), onde o autor fala de *atos sociais* – solicitar, advertir, questionar etc (STANFORD, 2014) –, e à obra de Alan H. Gardner (1879-1963), intitulada *The theory of speech and language* (1932), onde o autor fala de atos linguísticos (BIANCHI, 2008: 148).

⁷¹ Para biografia de Austin, ver ‘John Langshaw Austin, a biographical sketch’, escrito por G. J. Warnock (FANN, 1969: 3-21).

⁷² Segundo Warnock (1989: 105), contudo, o conteúdo de HTD passou a ser formulado já em 1939.

Rajagopalan (2010b: 93) também cita Falkenberg⁷³, segundo o qual “o livro de Searle de 1969 (...) foi uma quebra de paradigma para a linguística e ajudou a estabelecer uma nova disciplina: a pragmática linguística”. Vale dizer que o livro de Searle é o já citado *Speech Acts*, que para Levinson (2007: 302) é uma “(...) influente sistematização da obra de Austin (...), a cujos escritos a teoria dos atos de fala talvez deva a maior parte do seu impacto sobre os linguistas (...)”.

Não há, entretanto, acordo sobre se Searle de fato sistematizou a obra de Austin ou se, na verdade, o primeiro elaborou uma teoria própria e distinta da do segundo (ainda que dele tenha partido para formulá-la). Se, por um lado, Levinson opta pela primeira alternativa, Rajagopalan (2010c: 17), por seu turno, opta pela segunda, dizendo sentir-se surpreso com o fato de “(...) muitos estudiosos sequer reconhecerem que tais mudanças ocorreram [entre Austin e Searle], continuando a insistir em uma continuidade automática entre os dois” (algumas diferenças foram expostas no subcapítulo 2.2.2). Como, porém, não é objetivo deste trabalho investigar a fundo tal questão, basta expor a proposta⁷⁴ de Austin aqui adotada e o motivo de tal adoção.

2.3.1. Performativos e constativos

Cabe, antes, verificar como Austin chegou à noção de linguagem como ação. Para Alcoforado (1980: 15-18), a teoria de Austin pode ser dividida em “restrita” e “generalizadora”. A teoria restrita seria marcada pela distinção inicial proposta por Austin entre *proferimentos constativos* (aqueles usados para meramente descrever/relatar/ predicar algo de algo), que são passíveis de avaliação segundo princípio lógico da bivalência (ou seja, a cada constativo pode ser atribuído um e apenas um dos dois seguintes valores de verdade: verdadeiro ou falso), e *proferimentos performativos*⁷⁵ (aqueles que seriam usados para desempenhar uma ação), aos quais o princípio da

⁷³ FALKENBERG, G. Searle on Sincerity. In: BURKHARDT, A. (org.). *Speech Acts, Meaning and Intentions*. Berlin: Walter de Gruyter, 1990, p. 129-145.

⁷⁴ O Relatório de Qualificação ofereceu um esboço do contexto histórico-filosófico no interior do qual a TAF de Austin foi concebida. Entretanto, por sugestão da banca optou-se por excluí-lo. Ainda assim, considera-se fundamental indicar, além das obras citadas no corpo do texto, algumas outras que são relevantes para a compreensão do assunto em sua totalidade: Auroux (1995), Baldwin (2005), Biletzki (1998), Dummett (1996), Hacker (1997), Glock (2011), Leclercq (2008), Marcondes (2004) e (2005), Martinich e Sosa (2006), Milkov (2003), Miller (2010), Oliveira (2006), Passmore (1994[1957]), Rorty (1992[1967]).

⁷⁵ A noção de performatividade da linguagem aparece explícita em Austin, segundo Ottoni (1998: 32), já no artigo *Other Minds*, de 1946.

bivalência não se aplica (ou seja, não podem ser avaliados como verdadeiros ou falsos, e sim como *felizes* ou *infelizes*).

No artigo *Performative Utterances* (1979 [1956]), Austin critica aquilo que considera como *falácia descritiva*, i. e., argumento segundo o qual todo e qualquer proferimento só possui significado quando tomado como *proposição* – sentença declarativa usada para descrever/ relatar etc fatos do mundo, suscetível a ser classificada como verdadeira ou falsa. Afinal, ele se pergunta, os proferimentos não poderiam servir também para, “ao invés de [só] registrar fatos, [também] influenciar pessoas desta ou daquela maneira, ou desabafar deste ou daquele modo?” (1979 [1956]: 234)⁷⁶. Neste sentido, para o filósofo inglês um proferimento performativo seria aquele

que se parece com uma proposição e que gramaticalmente certamente seria classificada como tal, que não carece de sentido e, no entanto, não é nem verdadeiro nem falso. (...) São proferimentos simples, com verbos correntes na primeira pessoa do singular do presente do indicativo na voz ativa (1979 [1956]: 235)⁷⁷.

Um exemplo de performativo é o “Eu *aceito* esta mulher como minha legítima esposa”, segundo o qual nada está sendo descrito, mas a partir do qual o falante concretiza o seu casamento. Isto não significa que bastaria dizer “Eu *aceito*...” para consumir um casamento, pois outros requisitos devem ser satisfeitos: circunstâncias apropriadas; fatos contextualmente verdadeiros; convenção aceita; procedimento desenvolvido correta e completamente; sentimentos/ intenções/ pensamentos adequados etc. Assim, ao se proferir “Eu *aceito*...” em circunstâncias cujos critérios são respeitados, casa-se (situação em que o proferimento pode ser considerado como *feliz*). Em caso contrário, o casamento não é efetuado (situação em que o proferimento pode ser considerado como *infeliz*). Nota-se, assim, que ao contrário dos proferimentos constativos, avaliados em termos de verdadeiro/ falso, os performativos são avaliados em termos de *felicidade/ infelicidade*.

O filósofo inglês (1979[1956]: 241), entretanto, se pergunta, em seguida, sobre como se poderia ter certeza de que um proferimento é (ou não é) um performativo. Isso porque passa a suspeitar que dificilmente haja um critério gramatical que sustente tal peculiaridade em relação aos demais proferimentos. A princípio, seria possível supor que a forma verbal na primeira pessoa do singular do presente do indicativo na voz ativa

⁷⁶ “Mightn’t they perhaps be intended not to reports facts but to influence people in this way or that, or to let off steam in this way or that?”

⁷⁷ “(...) a kind of utterance which looks like a statement and grammatically, I suppose, would be classed as a statement, which is not nonsensical, and yet is not true or false. (...) They will be perfectly straightforward utterances, with ordinary verbs in the first person singular present indicative active”.

fosse um critério seguro (p. ex., “*Autorizo o senhor a...*”); entretanto, em um performativo como “O senhor *está autorizado a...*” encontra-se a mesma função ilocucionária de “*Autorizo...*”, sem que se faça uso aqui do verbo na primeira pessoa. Afora isso, outros expedientes linguísticos também poderiam ser utilizados para o mesmo fim (tal como dizer “*Sim*” ou “*Claro, senhor*”, com o sentido de autorizar). Diante disso, Levinson (2007: 293) sugere a distinção entre performativos explícitos e implícitos⁷⁸, sendo que os explícitos (tal como “*Autorizo o senhor a...*”) seriam apenas “maneiras relativamente especializadas de alguém ser inequívoco e específico a respeito do ato que está executando ao falar” (2007: 296), uma vez ser possível que maneiras menos inequívocas (implícitas) desempenhem a mesma função (tal como “*Claro, senhor*”).

No entanto, se um performativo é como tal considerado pelo fato de poder ser caracterizado como feliz/ infeliz, o que dizer de um proferimento do tipo “Todas as carnes do açougue são de segunda, mas não há carnes lá”, que aparenta ser uma proposição (proferimento constativo), mas que não possui um valor de verdade (já que não possui referência, i. e., não se refere a nada, já que não há carnes, não sendo, portanto, nem verdadeira nem falsa)? Ou ainda, e no caso de um proferimento do tipo “Olha o carro!” (proferimento performativo implícito), que poderia ser expresso por “Eu te alerto que um carro vem em sua direção” (proferimento performativo explícito), mas cuja avaliação em termos de felicidade não procederia caso carro algum estivesse se aproximando? Sobre isso, diz Levinson (2007: 298) que

se oferecemos conselho ou fazemos uma advertência, somos obrigados a ter bons fundamentos para fazê-lo, da mesma maneira que devemos ser capazes de fundamentar uma asserção ou uma constatação.

Assim, a conclusão é a de que tanto os performativos como os não-performativos poderiam ser julgados tanto em termos de verdadeiro/ falso como em termos de felicidade/ infelicidade. Logo, não há critério para distinguir um performativo de um constativo, o que significa que qualquer expediente linguístico pode ser analisado, de acordo com a abordagem, como performativo ou como constativo.

⁷⁸ Ou performativos explícitos e simples, como diz Levinson (2007: 297).

2.3.2. Dizer é fazer

Foi ao longo de suas sete primeiras conferências (sete primeiros capítulos de HTD) que Austin examinou essas questões e apresentou tais dificuldades. A partir disso, reavaliou seu percurso e elaborou a proposta segundo a qual todo proferimento possui três dimensões de atos – adentrando assim naquilo que Alcoforado (1980: 19) chamou de teoria generalizadora, segundo a qual “(...) seria um erro pensar que um proferimento é a realização de um único ato”. Não por acaso Alcoforado (1980: 19) diz que a TAF é aquela que considera que “(...) todo proferimento é um ato complexo ou um complexo de atos”, momento em que “a dicotomia entre performativas e constativas é rejeitada em favor de uma teoria completa dos atos de fala” (LEVINSON, 2007: 299).

Austin explicita os três atos subjacentes a qualquer proferimento (*utterance*):

Em primeiro lugar, distinguimos um conjunto de coisas que fazemos ao dizer algo, que sintetizamos dizendo que realizamos um *ato locucionário*, o que equivale, a grosso modo, a proferir determinada sentença com determinado sentido e referência, o que, por sua vez, equivale, a grosso modo, a ‘significado’ no sentido tradicional do termo. Em, segundo lugar dissemos que também realizamos *atos ilocucionários* tais como informar, ordenar, prevenir, avisar, comprometer-se, etc, isto é, proferimentos que têm uma certa força (convencional). Em terceiro lugar também podemos realizar *atos perlocucionários*, os quais produzimos *porque* dizemos algo (1990[1975]: 95)⁷⁹.

Por outras palavras, em sua visão proferir (*utter*) significa realizar:

(a) um *ato locucionário*, i. e., “(...) ato de ‘dizer algo’ (...)” (1990[1975]: 85)⁸⁰, composto por três subatos: (a.1) ato fonético, que consiste na emissão de ruídos; (a.2) ato fático, que consiste na emissão de certas palavras, ou seja, ruídos de tipo específico pertencentes a um determinado vocabulário e em conformidade com uma determinada gramática; (a.3) ato rético, que consiste na emissão de tais ruídos constitutivos de palavras de determinada língua com um determinado significado;

(b) um *ato ilocucionário*, ato que se realiza ao dizer algo, diferindo assim do ato de dizer algo e que está intimamente ligado com a noção de que “(...) a ocasião de um proferimento tem enorme importância, e que as palavras utilizadas têm de ser até certo ponto ‘explicadas’ pelo ‘contexto’ em que devem estar ou em que foram realmente

⁷⁹ “We first distinguished a group of things we do in saying something, which together we summed up by saying we perform a locutionary act, which is roughly equivalent to uttering a certain sentence with a certain sense and reference, which again is roughly equivalent to ‘meaning’ in the traditional sense. Second, we said that we also perform illocutionary acts such as informing, ordering, warning, undertaking, etc, i. e. utterances which have a certain (conventional) force. Thirdly, we may also perform perlocutionary acts: what we bring about or achieve by saying something (...)” (1975[1962]: 109).

⁸⁰ “The act of ‘saying something’ (...)” (1975[1962]: 94).

faladas numa troca linguística” (1990[1975]: 89)⁸¹ – uma vez que o que caracteriza tal ato é a ação realizada pelo falante emissor, expressa por determinada *força* de natureza convencional;

(c) um *ato perlocucionário*, a saber, quando dizer algo “(...) produzirá certos efeitos ou consequências sobre os sentimentos, pensamentos, ou ações dos ouvintes, ou de quem está falando, ou de outras pessoas” (1990[1975]: 89)⁸².

Em linhas gerais, pode-se dizer que se o ato locucionário é o ato *de* dizer algo e o ilocucionário é o ato que se realiza *ao* dizer algo, o perlocucionário é o ato que se realiza *por* se ter dito algo. Neste sentido, suponha-se que um indivíduo, x, diz o seguinte a outro indivíduo, y: “Prometo que farei isso”. O proferimento desse conjunto de sons articulados (ato fonético), que na língua portuguesa do Brasil foram categorizados como sendo dois verbos, uma conjunção e um pronome demonstrativo (ato fático), e que juntos formam a sentença da língua portuguesa “Prometo que farei isso” com certo sentido (ato rético), a saber, x prometeu a y fazer algo no futuro, constituem o que Austin chamou de *ato locucionário*. Por sua vez, como o “Prometo que farei isso”, ao ser dito por x, consistiu na realização de uma promessa, pode-se dizer que o *ato ilocucionário* em questão foi o de prometer. Finalmente, suponha-se que y, por conhecer o caráter espirituoso de x, compreendeu, aquilo que deveria ter sido compreendido como uma promessa, como uma brincadeira; logo o *ato perlocucionário* foi o de brincar, divertir, fazer sorrir. Vale dizer que o “Prometo que farei isso” pode ser descrito como proferimento performativo explícito, enquanto que “Eu farei” ou “Claro!” podem ser descritos como performativos implícitos.

Considerando que o primeiro ato (o locucionário) se refere ao “(...) estudo de locuções, ou de unidades completas do discurso” (AUSTIN, 1990[1975]: 85)⁸³, estudo este que sempre esteve na mira de estudiosos os mais diversos, e considerando que o terceiro ato (o perlocucionário) trata das consequências de uma ação – e que, portanto, muitas vezes é de difícil compreensão (entre outros motivos, pelo fato de nem sempre determinada ação é recebida como tal) – Austin concentra sua análise no segundo ato, o ilocucionário, já que “(...) por demasiado tempo os filósofos negligenciaram este estudo,

⁸¹ “(...) *the occasion of an utterance matters seriously, and that the words used are to some extent to be ‘explained’ by the ‘context’ in which they are designed to be or have actually been spoken in a linguistic interchange*” (1975[1962]: 100).

⁸² “*Saying something (...) produce certain consequential effects upon the feelings, thoughts, or actions of the audience, or of the speaker, or other persons (...)*” (1975[1962]: 101).

⁸³ “(...) *the study of locutions, or of the full units of speech*” (1975[1962]: 94).

tratando todos os problemas como problemas de ‘uso locucionário’ (...)”⁸⁴ (1990[1975]: 89). Esse é o motivo pelo qual frisa seu interesse em “(...) distinguir *força* de significado” (1990[1975]: 89)⁸⁵. Assim, uma vez que o ato locucionário diz respeito à dimensão linguística e o perlocucionário ao efeito extralinguístico produzido pelo interlocutor, o ato ilocucionário é “a realização da ação por meio do enunciado, seu aspecto fundamental é a força ilocucionária, que determina o ato de fala, ou sua função” (OTTONI, 2002, *apud* ANDRADE, 2014: 49)⁸⁶.

Logo, por *força* compreende-se o fenômeno de natureza convencional associado a todo e qualquer proferimento da linguagem ordinária, a partir do qual se torna possível inferir qual a ação realizada por sua emissão. Por outras palavras, “o ato ilocucionário é o que estabelece a função ou ação social de determinado enunciado” (ANDRADE, 2014: 49). Segundo Levinson (2007: 300-301), para Austin o estudo do significado (locução) pode prosseguir independentemente do estudo da força (ilocução) e, se o ato ilocucionário se realiza “diretamente pela força convencional associada à emissão de certo tipo de enunciado em conformidade com um procedimento convencional”, por consequência esse ato pode ser determinado.

Essa associação que Austin faz entre força ilocucionária e proferimentos da linguagem ordinária é marcada pela associação entre ato ilocucionário e verbo ilocucionário. Como se viu parágrafos acima, o ato ilocucionário é o ato que se realiza ao se dizer algo; já o verbo ilocucionário é aquele responsável por realizar determinado ato de fala, de acordo com as convenções linguísticas. O verbo *prometer*, *convencionalmente* falando, proferido em primeira pessoa do singular do presente do indicativo na voz ativa, realiza o ato de prometer (promessa); já o verbo *afirmar* realiza o ato de afirmar (afirmação) etc.

Austin foi criticado por isso⁸⁷, pois, como se vem observando desde o subcapítulo 2.1, os atos de fala associam-se aos expedientes linguísticos utilizados para realizá-los, mas não se confundem com os mesmos, uma vez que a força ilocucionária não equivale ao significado semântico – logo, não faz sentido associar a ação de prometer ao verbo “prometer”, uma vez que “*Prometo que...*” pode desempenhar a ação de prometer, sim,

⁸⁴ “It may be said that for too long philosophers have neglected this study, treating all problems as problems of ‘locutionary usage’ (...)” (1975[1962]: 100).

⁸⁵ “(...) to distinguish force and meaning (...)” (1975[1962]: 100).

⁸⁶ OTTONI, P. John Langshaw Austin e a visão performativa da linguagem. *D.E.L.T.A.*, v. 18, n. 1, p. 117-143, 2002.

⁸⁷ Ver Searle (1968). Rajagopalan (2010d: 51), contudo, aponta em Searle a mesma associação que Searle apontou em Austin.

mas também a de afirmar, fazer sorrir etc (dependendo do contexto de proferimento). Contudo, vale a pena fazer uma observação. Ainda que as críticas a Austin sejam (e de fato são) válidas, deve-se notar que a preferência que teve pelos verbos ilocucionários se dá em função de estes serem explícitos – ou seja, mais facilmente reconhecíveis. Isso não significa que Austin tenha descartado a possibilidade de que “Prometo que...” possa, ao invés de desempenhar a ação de prometer, desempenhar a ação de afirmar (seria contraditório, pois foi o próprio que pela primeira vez apontou essa possibilidade); simplesmente, ele optou por privilegiar os verbos ilocucionários.

Vale dizer que também neste trabalho os verbos ilocucionários serão privilegiados, uma vez que se pretende buscar ilocuições em *corpora* escritos, valendo-se assim das ferramentas oferecidas pela LC (de acordo com as quais os contextos de proferimento são ou negligenciados, ou considerados de modo restrito). Assim, a opção por verbos ilocucionários (ilocuições explícitas) se deu em função de ser possível prescindir dos contextos de proferimento em sentido amplo sem com isso prescindir das forças ilocucionárias convencionalmente estabelecidas pelo uso de tais verbos.

Austin (1990[1975]: 123-4) tenta determiná-los ao propor cinco classes de força ilocucionária: (i) *vereditivos*; (ii) *exercitivos*; (iii) *comissivos*; (iv) *comportamentais*; (v) *expositivos*. Os vereditivos

caracterizam-se por dar um veredito, como o nome sugere, por um corpo de jurados, por um árbitro, ou por um desempatador (...). Constituem essencialmente o estabelecimento de algo – fato ou valor (...). Os vereditivos consistem em emitir um juízo, oficial ou extra-oficial, sobre evidências ou razões quanto ao valor ou ao fato, na medida em que estes são passíveis de distinção. (...) Os vereditivos têm conexões óbvias com verdade e falsidade com relação ao fundamento ou falta de fundamento, à justiça ou injustiça (1990[1975]: 123-125)⁸⁸.

Por sua vez, os exercitivos

consistem no exercício de poderes, direitos ou influências. (...) Um exercitivo consiste em tomar uma decisão a favor ou contra um determinado curso da ação, ou advoga-la. É decidir que algo tem de ser de determinada maneira, em oposição a julgar que tal coisa é assim. É advogar que seja assim (...). Suas consequências podem ser que outros sejam ‘compelidos’ ou ‘autorizados’ ou ‘não autorizados’ a fazer certos atos (1990[1975]: 123 e 126)⁸⁹.

⁸⁸ “(...) are typified by the giving of a verdict, as the name implies, by a jury, arbitrator, or umpire. (...) It is essentially giving a finding as to something – fact, or value (...). Verdictives consist in the delivering of a finding, official or unofficial, upon evidence or reasons as to value or fact, so far as these are distinguishable. (...) Verdictives have obvious connexions with truth and falsity, soundness and unsoundness and fairness and unfairness” (1975[1962]: 151-153).

⁸⁹ “(...) are the exercising of powers, rights, or influence. (...) An exercitive is the giving of a decision in favour of or against a certain course of action, or advocacy of it. It is a decision that something is to be

Já os comissivos

caracterizam-se por prometer ou de alguma forma assumir algo; *comprometem* a pessoa a fazer algo, mas incluem também declarações ou anúncios de intenção (...) e incluem também coisas um tanto vagas que podemos chamar de adesões, como, por exemplo, tomar partido. (...) O importante de um comissivo é comprometer quem o usa a uma determinada linha de ação (1990[1975]: 123 e 127)⁹⁰.

Em relação aos comportamentais, diz Austin que estes

constituem um grupo muito heterogêneo, e têm a ver com atitudes e *comportamento social*. (...) Os comportamentais incluem a ideia de reação diante da conduta e da sorte dos demais, e de atitudes e expressões de atitudes diante da conduta passada ou iminente do próximo (1990[1975]: 123 e 129)⁹¹.

Finalmente, Austin esclarece o que entende por expositivos:

esclarecem o modo como nossos proferimentos se encaixam no curso de uma argumentação ou de uma conversa, como estamos usando as palavras, ou seja, são, em geral, expositivos. (...) Os expositivos se usam nos atos de exposição que consistem em expressar opiniões, conduzir debates e esclarecer usos e referências (1990[1975]: 124 e 130)⁹².

A tabela 1 indica algumas das espécies de força ilocucionária incluídas por Austin nas cinco classes descritas (1990[1975]: 123-130):

Classes/ Espécies	Vereditivos	Exercitivos	Comissivos	Comportamentais	Expositivos
Em inglês ⁹³	<i>acquit,</i> <i>convict, find,</i> <i>rule, reckon,</i> <i>rate, describe,</i> <i>assess, grade,</i>	<i>order,</i> <i>command,</i> <i>grant, appoint,</i> <i>nominate,</i> <i>choose,</i>	<i>promise,</i> <i>covenant,</i> <i>undertake,</i> <i>contract, mean</i> <i>to, guarantee,</i>	<i>apologize, thank,</i> <i>deplore,</i> <i>commiserate,</i> <i>felicitate, don't</i> <i>mind, criticize,</i>	<i>affirm, deny,</i> <i>remak, accept,</i> <i>postulate,</i> <i>interpret,</i> <i>object to,</i>

so, as distinct from a judgement that it is so: it is advocacy that it should be so (...). Its consequences may be that others are 'compelled' or 'allowed' or 'not allowed' to do certain acts" (1975[1962]: 151 e 155).

⁹⁰ "(...) *are typified by promising or otherwise undertaking; they commit you to doing something, but include also declarations or announcements of intention (...) and also rather vague things which we may call espousals, as for example, siding with. (...) The whole point of a commissive is to commit the speaker to a certain course of action*" (1975[1962]: 151-152 e 157).

⁹¹ "(...) *are a very miscellaneous group, and have to do with attitudes and social behaviour. (...) Behabitives include the notion of reaction to other people's behaviour and fortunes and of attitudes and expressions of attitudes to someone else's past conduct or imminent conduct*" (1975[1962]: 152 e 160).

⁹² "*They make plain how our utterances fit into the course of an argument or conversation, how we are using words, or, in general, are expository. (...) Expositives are used in acts of exposition involving the expounding of views, the conducting of arguments, and the clarifying of usages and of references*" (1975[1962]: 152 e 161).

⁹³ Tais como propostas por Austin (1975[1962]: 153-163).

	<i>characterize, analyse, value, locate</i>	<i>resign, vote for, press, advise, warning, warn</i>	<i>oppose, contemplate, favour, am determined to</i>	<i>welcome, bid you farewell, wish, defy, protest, challenge</i>	<i>agree, rejoin, deduce, argue, distinguish</i>
Em português ⁹⁴	absolvo, condeno, constato, determino, coloco, qualifico, descrevo, valorizo, classifico, caracterizo, analiso, avalio, situo	ordeno, mando, concedo, designo, indico, escolho, renuncio, voto em, pressiono, aconselho, aviso, advirto,	prometo, compactuo, comprometo-me, contrato, pretendo, garanto, oponho-me, aposto, juro, sou a favor de, estou decidido a	peço desculpas, agradeço, deploro, compadeço-me, felicito, não me importo, critico, seja bem-vindo, desejo-te, duvido que, protesto, desafio-o	afirmo, nego, observo, aceito, postulo, interpreto, faço objeção, repudio, concordo, replico, deduzo, argumento, distingo

Tabela 1: classes de força ilocucionária e algumas de suas espécies propostas por Austin (1990[1975]: 123-130).

Resumindo, diz Austin que

[e]m suma, podemos dizer que o vereditivo é um exercício de julgamento, o exercitativo é uma afirmação de influência ou exercício de poder, o comissivo é assumir uma obrigação ou declarar uma intenção, o comportamental é a adoção de uma atitude e o expositivo é o esclarecimento de razões, argumentos e comunicações (1990[1975]: 131)⁹⁵.

2.4. A TAF de Austin como paradigma de pesquisa

Uma vez que as ilocuições possuem uma natureza convencional, este trabalho visa a sugerir modos de dicionarizá-los, permitindo assim o conhecimento/ reconhecimento de tais convenções por aqueles que tomam contato com línguas estrangeiras. Reforça-se, assim, o que se disse anteriormente (1.2): a presente investigação metalexiconográfica não pretende analisar a TAF de Austin em si mesma, nem compará-la com teorias

⁹⁴ Segundo tradução de Danilo Marcondes Filho (1990[1975]: 124-131).

⁹⁵ “To sum up, we may say that the verdictive is an exercise of judgment, the exercitive is an assertion of influence or exercising of power, the commissive is an assuming of an obligation or declaring of an intention, the behabitive is the adopting of an attitude, and the expositive is the clarifying of reasons, arguments, and communications” (1975[1962]: 163).

classificatórias posteriores⁹⁶, tampouco estabelecer classes/ espécies de ilocuções a partir da análise de casos concretos. Aqui, algumas ilocuções presentes na classificação austiniana serão tomada como objeto de pesquisa, de modo a, primeiramente, investigá-las em *corpora* escritos, de acordo com determinada sugestão de pesquisa, e, em seguida, sugerir meios de inserção em dicionários híbridos.

Em termos prático-operacionais, a razão pela escolha da proposta austiniana baseia-se no critério do verbo na primeira pessoa do singular do presente do indicativo na voz ativa tal como apontado por ele. Nas palavras do próprio Austin (1990[1975]: 62):

todo proferimento que seja de fato um performativo deveria ser capaz de ser reduzido, expandido ou analisado de modo tal que se obtivesse uma forma na primeira pessoa do singular do presente do indicativo da voz ativa (gramatical)⁹⁷.

Viu-se (2.3.1 e 2.3.2) que o próprio Austin considerou este critério como restritivo, pois limita as ilocuções a verbos ilocucionários explícitos. Em termos práticos, isso significa que proferimentos do tipo “*Farei*” (com a força de prometer), “O senhor *está autorizado* a...” (com a força de autorizar), “*Olá!*” (com a força de cumprimento) sejam negligenciados da pesquisa. Entretanto, as vantagens decorrentes da opção por focar nos verbos ilocucionários são duas: por serem ilocuções explícitas, o critério do verbo na primeira pessoa permite, primeiramente, a criação de uma sintaxe de busca padronizada, e, em segundo lugar, evita-se a necessidade de se analisar todo o *corpus* na busca por casos indubitáveis de ilocuções (já que as ilocuções não dependem de expedientes linguísticos determinados para se manifestar).

Se as ilocuções não dependem de expedientes linguísticos determinados, não é imediatamente evidente como obter ilocuções via LC – já que essa metodologia baseia-se em análise de *corpora* cujos dados são *eminentemente* linguísticos. Tal obtenção, então, torna-se possível com a TAF de Austin, seja porque ele associou as ilocuções a verbos ilocucionários (categoria gramatical facilmente reconhecível), seja porque propôs o critério do verbo na primeira pessoa do singular (que pode ser utilizada como sintaxe de busca). Reconhece-se, assim, que tal critério restringe a pesquisa; contudo, o

⁹⁶ Rajagopalan (2010d: 46) apresenta uma série de estudiosos que procuraram classificações semelhantes, tal como faz Levinson (2007: 305). Ver também: HANCHER, M. The classification of co-operative illocutionary acts. *Language in Society*, 8.1, i-14, 1979.

⁹⁷ “Thus what we should feel tempted to say is that any utterance which is in fact a performative should be reducible, or expandible, or analysable into a form, or reproducible in a form, with a verb in the first person singular present indicative active (grammatical)” (1975[1962]: 61-62).

que se pretende aqui é sugerir a dicionarização e não dicionarizar (ou seja, construir um dicionário), quando então seria relevante considerar todos os (ou grande parte dos) casos de ilocuções⁹⁸.

Outras razões, agora em termos teóricos, podem ser dadas para justificar a escolha pela teoria de Austin. Este último, segundo Rajagopalan (2010d: 48-49), “(...) tinha uma posição fechada em relação à tese de que os atos ilocucionários se definem em relação a línguas específicas (ou melhor, comunidades de fala que as sustentam)”, noção relevante para o presente trabalho, já que o objetivo aqui, ao se ter por norte um dicionário híbrido italiano>português-brasileiro, é justamente propor contato entre duas línguas a partir do realce de particularidades e semelhanças entre elas.

Além disso, Austin havia, novamente com Rajagopalan (2010b: 106), o hábito de “interrogar cada caso inúmeras vezes, de tentar testar diferentes matizes e nuances de significado de cada item lexical”, o que permite a ideia de que, se por um lado, sua classificação não é fechada nem exaustiva, por outro, sua classificação se deu a partir de uma ampla investigação⁹⁹. O próprio Austin reconheceu a existência de “casos marginais ou embaraçosos, ou casos de sobreposição entre essas classes” (1990[1975]: 124)¹⁰⁰, ou ainda, como no caso das duas últimas classes citadas (comportamentais e expositivos), demonstrou consciência de que “(...) bem pode ser que não estejam nitidamente classificadas ou que estejam um tanto embaralhadas, ou mesmo que necessitem de uma classificação inteiramente distinta” (1990[1975]: 124)¹⁰¹ – além da hipótese de que “bem se poderia dizer que todos os aspectos estão presentes em todas as classes” (1990[1975]: 124)¹⁰². Vê-se com vantagem essa classificação aberta porque

⁹⁸ A escolha por quais ilocuções buscar poderia ter sido outra. Tendo-se, entretanto, que fazer uma escolha, e, sobretudo, devido ao fato de esta ser uma dissertação de mestrado (texto de natureza acadêmico-científica), cujos objetivos e o tempo destinado a alcançá-los são finitos, optou-se por um procedimento de pesquisa que decorresse logicamente dos pressupostos estabelecidos. Outras explicações são dadas no subcapítulo 2.4.1.

⁹⁹ Como apontou Andrade (2014: 53-54) ao citar Jucker (Speech act research between armchair, field and laboratory: The case of compliments. *Journal of Pragmatics*, 41(8):1611-1635, 2009), pode-se caracterizar três tipos de investigação de atos de fala: as “pesquisas de poltrona”, conduzidas a partir de exemplos imaginados pelos próprios autores; as “pesquisas de campo”, conduzidas em situações elaboradas de modo a parecerem “reais”; as “pesquisas de laboratório”, as mais comuns e em geral conduzidas por meio de questionários aplicados a um público determinado, cujas respostas são analisadas e transformadas em dados estatísticos. Diante disto, pode-se dizer que, embora “de poltrona”, a investigação de Austin parece ter sido ampla. Afora isso, pode-se incluir o presente trabalho na categoria “de laboratório”, pois, ainda que não utilize questionários, trabalha com dados concretos.

¹⁰⁰ “We should be clear from the start that there are still wide possibilities of marginal or awkward cases, or of overlaps” (1975[1962]: 152).

¹⁰¹ “(...) it would well be that they are not clear or are cross-classified, or even that some fresh classification altogether is needed” (1975[1962]: 152).

¹⁰² “It could well be said that all aspects are present in all my classes” (1975[1962]: 152). ”.

permitiria aos lexicógrafos (e demais linguistas) inúmeras possibilidades de investigação dos atos de fala segundo suas necessidades e possibilidades, i. e., seu(s) dicionário(s) e seu(s) público(s), o que não seria possível se tal classificação fosse rígida¹⁰³. Como se viu antes (2.2) com Sbisà (2009[1989]: 38), se é possível descobrir fatos novos com relação à ação praticada por um proferimento, já que o efeito produzido nem sempre corresponde a esta, pode-se ampliar essa noção dizendo que novos fatos podem ser descobertos ao se possuir uma classificação aberta – o que, inclusive, está de acordo com a própria natureza das línguas naturais, estas em constante devir.

A teoria de Austin foi escolhida também por reconhecer a utilidade dos dicionários. Austin os menciona esparsamente em vários momentos de sua obra (1979[1956]), (1990[1975]), (1979[1956-57]), mas sugere explicitamente o uso de dicionários por sua utilidade ao pesquisador na execução desta ampla tarefa de análise da linguagem, podendo ser consultados “(...) de cima a baixo (...) [ou] a partir de uma seleção prévia de termos obviamente relevantes (...)” (1979[1956-57]: 186)¹⁰⁴, de forma a se conhecer as diversas significações de cada um, segundo contexto de uso. Portanto, para Austin tais instrumentos são auxiliares à análise, e, logo, não dão a palavra final, já que se deve sempre privilegiar o exame da linguagem ordinária; esta, entretanto, tampouco é a palavra final¹⁰⁵. Portanto, se é verdade que Austin avalia importante, por um lado, para se depreender o sentido de um proferimento, considerar a sentença, pois “(...) o que sozinho tem significado é a *sentença*” (1979[1940]: 56)¹⁰⁶, é verdade também que o filósofo não desconsidera a investigação dos itens lexicais que constituem tal sentença (daí sua consideração pelas obras lexicográficas). E, ainda que, hipoteticamente, Austin desconsiderasse a análise de palavras fora de sentenças proferidas, o seu critério do verbo na primeira pessoa permite por si só a busca de ilocuções associadas a itens lexicais (“Prometo”, “Asseguro” etc).

¹⁰³ Além do fato de que, sendo rígida, caminharia em sentido contrário ao que ocorre no uso linguístico cotidiano.

¹⁰⁴ “(...) *to read the book through (...) [or] to start with a widsish selection of obviously relevant terms (...)*”.

¹⁰⁵ “Certamente a linguagem ordinária *não* é a última palavra (...). Mas, não nos esqueçamos, *é* a primeira palavra” (1979[1956-57]: 185) [*“Certainly, then, ordinary language is not the last word (...). Only remember, it is the first word”*].

¹⁰⁶ “(...) *what alone has meaning is a sentence*”.

2.4.1. A classe dos comissivos

Como se viu anteriormente (2.3), a classe de ilocuções comissivas são aquelas que, uma vez proferidas, “empenham o falante a certa linha de ação” (SBISÀ, 2009[1989]: 71)¹⁰⁷. As espécies de ilocuções que serão investigadas neste trabalho pertencem a tal classe, a saber: os italianos *Promettere*, *Assicurare*, *Garantire*, *Giurare*, *Impegnarsi*, e seus equivalentes semânticos em português-brasileiro: *Prometer*, *Assegurar*, *Garantir*, *Jurar*, *Comprometer-se*.

As razões para a escolha de tais ilocuções são as seguintes. Em primeiro lugar, por estarem associadas a verbos ilocucionários, mostram-se como ilocuções explícitas, o que permite a observação de suas ocorrências sem necessidade de se percorrer todo o *corpus*. Esta primeira razão, contudo, não é suficiente, uma vez que outras ilocuções associadas a verbos ilocucionários poderiam ter sido escolhidas (afirmar, condenar etc). Entretanto, a peculiaridade destas ilocuções é a seguinte: como se apontou no parágrafo anterior, os verbos ilocucionários que correspondem a essas ilocuções são semanticamente equivalentes e, logo, a partir desta escolha vislumbra-se a possibilidade de apontar, após a pesquisa em *corpora*, se também há equivalência pragmática. Ao se analisar a Tabela 1 (2.3.2), nota-se que não existem outros verbos ilocucionários explícitos que sejam sinônimos entre si e, portanto, como aqui se pretende, ao sugerir a dicionarização de ilocuções, apontar diferenças e semelhanças entre ilocuções realizadas na língua italiana e no português do Brasil, considera-se que possa ser útil para um brasileiro interessado no estudo/ contato com a língua italiana saber em quais casos pragmáticos se usa *prometto* e em quais se usa *assicuro*, p. ex. – já que, segundo o dicionário *Sinonimi e contrari* (2009), são semanticamente sinônimos.

Esta última razão se impõe ainda mais quando se coloca a seguinte questão: por quê, ao invés de verbos ilocucionários comissivos, não se investigou interjeições e demais expressões de saudação (“olá!”, “como vai?”, “fala aí!”, “oi!”, “tudo bom?”, “opa!”, “satisfação!”, “meu brother!” etc) ou advérbios (tal como fez o *Salamanca*) etc – já que, p. ex., “olá” é *convencionalmente* utilizado para saudar, mesmo que em uma versão implícita (uma versão explícita seria algo bizarro do tipo: “Eu lhe saúdo”, “Eu lhe cumprimento”)? A resposta é a seguinte: em primeiro lugar, é verdade que “olá!” e “oi!” são convencionalmente usados para saudar; porém, o mesmo não ocorre com “como vai?”, “tudo bem?”, “opa!”, que *convencionalmente* podem desempenhar

¹⁰⁷ “*impegnano il parlante a una certa linea d’azione*”.

ilocuções diversas – quando então seria necessário investigar todo o *corpus*, de modo a verificar se em todas as ocorrências elas desempenham a ação de cumprimentar ou não¹⁰⁸. A consequência é que a pesquisa de apenas duas ilocuções de cumprimento seria restrita, o que não ocorre no caso dos verbos ilocucionários comissivos, já que foi possível escolher cinco deles.

Assim, a escolha de cinco verbos – e não um, ou dois – baseia-se na consideração de que o presente número é suficiente, mesmo a um trabalho metalexicográfico de sugestão, para garantir o apontamento de algumas diferenças e semelhanças entre atos de fala do italiano e do italiano em comparação com o português-brasileiro. Importante neste ponto é alertar para o seguinte: não se pretende aqui sugerir que “*Prometto...*”, “*Assicuro...*” etc sejam sinônimos perfeitos¹⁰⁹, mas apenas que possuem certa sinonímia. Sabe-se, portanto, que *prometto* não pode ser substituído por *assicuro* em todos os casos e vice-versa, pois *prometto* corresponde ao ato de fala de prometer, *assicuro* ao de assegurar, e assim por diante.

¹⁰⁸ Não é que não se possa investigar apenas “oi!” e “olá!” em um *corpus*, ou advérbios ou outros expedientes linguísticos. Entretanto, neste trabalho, considerou-se a hipótese de que ao serem escolhidos cinco verbos sinônimos entre si, os resultados pragmático-ilocucionários pudessem ser interessantes.

¹⁰⁹ Como diz Ilari (2011: 169 *apud* COELHO; REZENDE, 2012: 331), “os sinônimos são palavras de sentido próximo, que se prestam, ocasionalmente, para descrever as mesmas coisas e as mesmas situações” – e ressalta a inexistência de sinônimos perfeitos. ILARI, R. *Introdução ao estudo do léxico* – brincando com as palavras. 5. ed. São Paulo: Contexto, 2011. Zucchi (2010b) também faz tal apontamento (ver 4.1). Reforça-se, então, que os objetos de investigação aqui *não são verbos semanticamente equivalentes*, pois a equivalência de que se vale aqui é tão somente usada como critério de escolha dos verbos. Em outras palavras, por ser necessário um critério para escolha dos verbos ilocucionários, adotou-se a sugestão de sinonímia dada pelo *Sinonimi e contrari* (2009), uma vez que não seria possível abordar a fundo tal questão neste trabalho.

3. LINGÜÍSTICA DE CORPUS (LC)

Neste capítulo são oferecidas informações mais completas acerca da LC, definida como metodologia de pesquisa capaz de permitir a exploração e extração de unidades convencionais de línguas naturais por meio da observação simultânea de grande quantidade de dados linguísticos compilados e disponibilizados eletronicamente, reunidos em um ou mais *corpora*. O *corpus*, assim, é definido como coletânea de textos orais e/ ou escritos a partir do qual é possível extrair padrões linguísticos, segundo a utilização de ferramentas adequadas – tais como lista de frequência e lista de concordância. Apresentam-se também as duas abordagens tradicionais de pesquisa, a *corpus-based* e a *corpus-driven*, salientando que este trabalho adota a primeira, em função de já se saber previamente aquilo que será investigado. Por fim, segundo a noção de representatividade, segundo a qual quanto mais extenso um *corpus*, mais representativo da língua ele é, opta-se por dois *corpora* amplamente extensos: o *Corpus Paisà* e o *Corpus do Português*.

3.1. A LC como metodologia

Muitos autores reconhecem a utilidade dos *corpora* para a Lexicografia. Welker (2004: 89) menciona que desde a década de 1970 os lexicógrafos têm aproveitado *corpora* eletrônicos. Krieger (2008: 171) também ressalta que os atuais desenvolvimentos da Lexicografia têm se alicerçado na constituição de *corpora*. Biderman¹¹⁰ (2001b: 134), ao comentar o processo de confecção do *Collins Cobuild Dictionary of English Language*, de 1987, frisa que seus autores “(...) basearam-se em resultados léxico-estatísticos para selecionar sua nomenclatura” e reforça a importância dos *corpora* de dados linguísticos como fonte a partir da qual é possível extrair dados de modo a compor dicionários. Oliveira (2009: 57) diz que a possibilidade de análise de grandes quantidades de dados da língua natural, segundo observação de uso da mesma em contextos diversos, “tem aberto novas perspectivas para estudos aplicados de diferentes naturezas como estudos lexicográficos, léxico-gramaticais, tradutórios e de gêneros discursivos” (2009: 57).

¹¹⁰ Vale dizer que foi a primeira linguista a trabalhar com LC – já nos anos de 1970.

A LC, segundo definição proposta por Tagnin e Teixeira (2004: 321) e aceita neste trabalho, é

uma metodologia de pesquisa científica em que a exploração da linguagem é feita por meio de evidências, extraídas por meio do uso de ferramentas computacionais, de um *corpus* de linguagem natural e autêntica criteriosamente reunido e disponível eletronicamente.

Segundo Gutierrez Gonzalez (2007: 8), a principal característica da LC é a possibilidade de

observação de dados empíricos de uma ou mais línguas – ou variedades da língua – armazenados em bancos de dados que compõem um *corpus* (sic), com a utilização de ferramentas eletrônicas especialmente desenvolvidas para auxiliar o pesquisador na análise dos dados, facilitando assim o seu trabalho quanto à verificação dos fenômenos da língua em uso.

O valor dessa exploração da linguagem por meio da observação de dados de uma ou mais línguas com base em *corpus* eletrônico é ressaltado por Berber-Sardinha (2003: 104) ao dizer que um *corpus* computadorizado é um ótimo objeto para a investigação da linguagem porque por ele pode ser estimada a ocorrência de uma ampla gama de traços linguísticos – morfológicos, morfossintáticos, sintáticos, semânticos, discursivos, etc. Ainda com o mesmo (2003: 104),

o emprego de um *corpus* (sic) na pesquisa linguística traz vários benefícios, entre eles a possibilidade da explicação de diferenças de uso de palavras, expressões, formas gramaticais e outros traços por meio da probabilidade de ocorrência em contextos específicos (Biber et al., 1998), a possibilidade de descoberta de fatos novos não disponíveis pela intuição ou eliciação (Sinclair, 1991) e a descrição objetiva da linguagem enquanto um sistema probabilístico (Halliday, 1992)¹¹¹.

Por sua vez, Oliveira (2009: 63) lista as seguintes características da LC:

constituem-se em investigações da língua em uso; baseiam-se em coleções de textos selecionados de acordo com certos critérios; usam computadores para a análise automática ou interativa; incluem análises quantitativas e/ou interpretações qualitativas para descreverem padrões; possibilitam a análise de textos longos e variados; possibilitam o uso de um mesmo *corpus* (sic) para verificar ou procurar novos resultados; podem trazer subsídios para linguistas teóricos e aplicados; proporcionam maior precisão e credibilidade às análises quantitativas.

¹¹¹ BIBER, D. et al. *Corpus linguistics - Investigating language structure and use*. Cambridge University Press, 1998; HALLIDAY, M.A.K. Language as system and language as instance: The corpus as a theoretical construct. IN: J. SVARTVIK (org.). Directions in corpus linguistics. *Proceedings of Nobel Symposium 82*, Stockholm, 4-8 August 1991. De Gruyter, 1992; SINCLAIR, J. *Corpus, concordance, collocation*. Oxford University Press, 1991.

Vale ressaltar que essa utilidade é melhor compreendida quando comparada com os métodos de investigação da linguagem antes da era da informática¹¹². De acordo com Tagnin (2013: 29), se no passado a identificação das unidades convencionais da língua dependia da análise de sua recorrência no dia a dia (ou, como se viu, através da constituição de *corpora* impressos), hoje essa identificação “(...) é facilitada pela observação simultânea de uma grande quantidade de dados a partir de um *corpus* eletrônico”. Afora isso, segundo Oliveira (2009: 49) “análises feitas com auxílio de programas de computador podem também levar a novas descobertas sobre aspectos linguísticos até então não considerados como relevantes pelos pesquisadores”.

A LC, então, a partir de tais considerações, é tida aqui como metodologia¹¹³ de pesquisa que permite a exploração de unidades convencionais de uma ou mais línguas naturais por meio da observação simultânea de grande quantidade de dados linguísticos empíricos, compilados e disponibilizados em formato eletrônico, graças às possibilidades oferecidas por ferramentas computacionais como computadores, *softwares* específicos e internet para tal compilação.

3.2. Corpus e abordagens de investigação

Na visão de Berber-Sardinha (2004: 18), *corpus* é um conjunto de dados linguísticos¹¹⁴ (do uso oral ou escrito da língua) sistematizados segundo critérios, suficientemente extensos¹¹⁵ em amplitude e profundidade e dispostos de tal modo que possam ser processados por computador. Por sua vez, Tagnin (2013: 29) entende *corpus* como “(...) coletânea de textos, necessariamente em formato eletrônico, compilados e organizados segundo critérios ditados pelo objetivo de pesquisa a que se destina”. Shepherd (2009: 151) o define como “uma coletânea de textos em linguagem natural,

¹¹² Gutierrez Gonzalez (2007: 10) aponta que *corpora* existem desde a Grécia Antiga, compostos por “(...) documentos ‘manuscritos’, adotando o sentido primário da palavra *corpus*, ou seja, uma grande quantidade de documentos”.

¹¹³ Isso não significa que se considera a LC como *mera* metodologia – bem como nenhum dos autores aqui mencionados a considera. Por exemplo, Berber-Sardinha (2000: 357) contesta a ideia de que a LC seja só metodologia porque esta não se resume a um conjunto de ferramentas, sendo, portanto, uma abordagem/ perspectiva de investigação linguística. Já Oliveira (2009: 51) defende a ideia da LC como campo de estudo com uma perspectiva particular da linguagem. Concorde-se com tais autores, mas adota-se neste trabalho a LC como metodologia pelo fato de que apenas seus instrumentais serão utilizados.

¹¹⁴ Pode-se neste ponto reforçar a questão da autenticidade: como diz Berber-Sardinha (2000: 336), quando se fala de compilação de *corpus* significa que os textos coletados são autênticos, i. e., “existem na linguagem (...), não foram criados com o propósito de figurarem no corpus”.

¹¹⁵ O *corpus* deve ser suficientemente extenso de modo a ser representativo da totalidade do uso linguístico. Isso será mais bem explicado no subcapítulo 3.4.

escritos ou falados, geralmente armazenados de forma organizada e informada, além de serem digitalizados a fim de que possam ser lidos por computador”. Em resumo, considerando-se tais definições e segundo a perspectiva da LC enquanto metodologia, por *corpus* entende-se fonte de pesquisa linguística, organizada segundo critérios elaborados/ adotados pelo pesquisador que o compõe de textos da língua corrente em formato eletrônico, de modo a nele observar e dele extrair padrões linguísticos desejados pela pesquisa.

Shepherd (2009: 153) indica a existência de duas abordagens consagradas de pesquisa com uso de *corpora* eletrônicos: a abordagem baseada em *corpus* (*corpus-based*) e a abordagem direcionada pelo *corpus* (*corpus-driven*). Segundo Shepherd (2009: 153), a abordagem baseada em *corpus* é “(...) uma metodologia que se aproveita do *corpus* essencialmente para testar e exemplificar teorias e descrições linguísticas pré-existentes”, sendo *corpus*, aqui, usado primordialmente como fonte de exemplos, quantificáveis em sua frequência e extensão – ou ainda, segundo Biber (2009: 276), a abordagem baseada em *corpus* “assume a validade das estruturas linguísticas derivadas da teoria linguística”, sendo que o objetivo primário com essa abordagem é “analisar padrões sistemáticos de uso por essas características linguísticas pré-definidas”¹¹⁶.

Já no caso da *corpus-driven*, esta em geral começou a ser praticada em função da “(...) ausência de uma teoria que explicasse as relações lexicais na gênese dos trabalhos com corpora eletrônicos” (SINCLAIR, 2004: xviii *apud* SHEPHERD, 2009: 153)¹¹⁷ – ou seja, neste caso não se utiliza o *corpus* como fonte de exemplo e/ ou como modo de testar e justificar empiricamente uma dada teoria, mas, de acordo com Shepherd (2009: 153), o que se busca é, sobretudo, a “observação e análise de padrões e frequências lexicais” – padrões e frequências de aspectos negligenciados por teorias linguísticas pré-estabelecidas. Sobre tal abordagem, diz Biber (2009: 276) que “(...) é mais indutiva, de modo que os constructos linguísticos emergem da análise do *corpus*” e difere, portanto, do *corpus-based* porque “faz afirmações *a priori* em número reduzido” e “assume somente a existência de palavras”¹¹⁸.

¹¹⁶ “‘Corpus-based’ research assumes the validity of linguistic structures derived from linguistic theory; the primary research goal is to analyze the systematic patterns of use for those pre-defined linguistic features”.

¹¹⁷ SINCLAIR, J. Interview. In.: KRISHNAMURTHY, R. (ed.) *English collocation studies: the OSTI report*. London: Continuum, 2004, 208 p.

¹¹⁸ “(...) ‘corpus-driven’ research is more inductive, so that the linguistic constructs themselves emerge from analysis of a corpus. The corpus-driven approach differs from the standard practice of linguistics in that it makes minimal *a priori* assumptions regarding the linguistic constructs that should be employed for the analysis. In its most basic form, corpus-driven analysis assumes only the existence of words (...)”.

Neste trabalho, vale dizer que a abordagem adotada será a *corpus-based* (abordagem baseada em *corpus*), uma vez que não se pretende levantar dados linguísticos de modo a estabelecer e/ ou elaborar teorias em si, mas sim investigar padrões linguísticos a partir de determinados pressupostos (as ilocuições comissivas). Em outras palavras, sabe-se previamente aquilo que será investigado, razão pela qual considera-se a *corpus-based* como abordagem mais adequada.

Outra distinção em relação ao uso de *corpus* como fonte de pesquisa é a de *corpus* comparável e paralelo. No primeiro caso, segundo Ribeiro (2004: 163), pretende-se comparar textos originais escritos numa dada língua com textos originais similares (similaridade em termos de assunto, tipologia textual etc) em outra língua. O segundo caso, por sua vez, ocorre quando se alinha texto original e sua tradução – i. e., pelo menos dois sub-*corpora* alinhados entre si, tendo-se de um lado um *corpus* composto de textos originais numa dada língua e de outro um *corpus* com os mesmos textos traduzidos para outra língua.

O presente estudo optou por *corpora* comparáveis, já que terá por base um *corpus* do italiano e um do português-brasileiro, independentes entre si e cuja equivalência se dá em termos de extensão.

3.3. Principais ferramentas oferecidas pela LC

A ferramenta de frequência é uma das principais oferecidas pela LC para verificação e extração de padrões da língua em textos pré-selecionados – qualquer que seja o tipo de padrão que se busca: o de itens lexicais, colocações, coligações¹¹⁹ etc. Segundo Gutierrez Gonzalez (2007: 15), “a lista de frequência de palavras é valiosa para a LC, pois revela a ocorrência de cada palavra naquele *corpus* específico, além de listar todas as palavras que o compõem”. *Softwares* de LC como *Wordsmith Tools* e *AntConc* são alguns dos que oferecem a possibilidade de se verificar, segundo listas de frequência, quais são os itens linguísticos mais ou menos frequentes em dado *corpus*. Tal ferramenta, contudo, pode também ser encontrada em *corpora* disponíveis *online*,

¹¹⁹ Tagnin (2005: 30) define coligação como “combinação consagrada de elementos linguísticos em que o colocado é uma palavra gramatical, p. ex., (...) *obedecer a, cumpridor de*”. Já por colocação entende “combinação lexical consagrada de duas ou mais palavras de conteúdo: p.ex., (...) *repolho roxo, (...), cadeira de balanço*”.

aqueles compilados por pesquisadores e grupos de pesquisadores e disponibilizados gratuitamente na internet¹²⁰.

Outra ferramenta de natureza estatística é lista de concordância, que segundo Berber-Sardinha (2004: 187), é a “(...) listagem de ocorrências de um item específico, dispostas de tal modo que a palavra de busca (aquela que se tem interesse em investigar) aparece centralizada na página (ou na tela do computador)” (sic). Vale ressaltar que, neste caso, o item lexical investigado aparece em seu contexto linguístico de uso (ou seja, como componente de uma sentença), em geral acompanhado do número preciso de suas ocorrências em todo o *corpus*. Essa ferramenta também é fornecida tanto pelos *softwares* mencionados no parágrafo anterior como por *corpora online*, tais como o *Corpus Italiano Paisà* e os *corpora* disponibilizados pelo site *Sketch Engine* (entre os quais, o *Corpus do Português*, utilizado neste trabalho).

3.4. Padrão, extensão e representatividade

Tanto em termos de lista de frequência quanto em termos de lista de concordância a noção de padrão é essencial. De acordo com Gutierrez Gonzalez (2007: 17), a padronização “apoia-se teoricamente no princípio idiomático (*idiom principle*)”, segundo o qual o usuário de determinada língua tem à disposição uma quantidade de sentenças pré ou semiconstruídas que se constituem em escolhas únicas. A LC, neste caso, auxilia no reconhecimento das escolhas linguísticas mais significativas da língua. Como diz Berber-Sardinha (2000: 351), a padronização se evidencia pela recorrência, isto é, um item lexical que se repete significativamente, podendo isto significar ser um caso de *padrão* lexical. O mesmo diz Tagnin (2002) ao reforçar que um *corpus* oferece “a forma mais usual na língua sob investigação”.

Novamente, para Berber-Sardinha (2000: 350) a visão que a LC tem da linguagem enquanto sistema probabilístico pressupõe a noção de que muitos traços linguísticos não ocorrem com a mesma frequência. P. ex., ainda que ambos os verbos “jurar” e “prometer” possam ser usados de modo a comprometer o falante a determinada linha de ação, é possível que não possuam a mesma frequência quando investigados em determinado *corpus*. Neste sentido, a LC permite ao pesquisador da linguagem a

¹²⁰ Gutierrez Gonzalez (2007: 15) menciona, como exemplo, o *Corpus Brasileiro* da Universidade Estadual Paulista (UNESP), que “serviu de subsídio para a preparação do Dicionário Didático de Português”.

extração de padrões linguísticos segundo métodos quantitativos, que “nos revelarão quais palavras são mais ou menos usadas numa dada língua, ou que tipos de padrão são mais frequentemente empregados” (GUTIERREZ GONZALEZ, 2007: 21). Em termos de lista de frequência, a LC indica quantas vezes um item lexical ocorre em um determinado *corpus*, considerando sua posição entre aqueles que ocorrem em um maior ou menor número de vezes. Já em termos de lista de concordância, a LC indica quantas vezes um item lexical ocorre em um determinado *corpus*, considerando os itens lexicais que coocorrem junto com aquele que foi pesquisado.

Vale dizer que, neste trabalho, a ferramenta fundamental a ser utilizada é a lista de frequência, segundo os critérios adotados para pesquisa de ilocuções (ver 2.4).

Neste sentido, como aponta Oliveira (2009: 50), pelo fato da principal particularidade da LC ser o conjunto de ferramentas de natureza estatística de que dispõe seria possível pensar que essa se resume a isso – ou seja, a de simplesmente permitir a extração de resultados numéricos dos *corpora* (o que, note-se bem, já é valiosíssimo). Em tal situação hipotética seria possível por em xeque a representatividade de tais resultados de pesquisas linguísticas, dado que o fato de um item lexical possuir ampla ocorrência não significa que em todas elas desempenhou a mesma função. Porém, como aponta a própria autora (2009: 50), é equivocado pensar que “bastam os números ou as estatísticas para descrever fatos linguísticos”, uma vez que, para interpretar os dados obtidos por intermédio do *corpus*, em muitos casos é necessário também examiná-lo qualitativamente.

Assim, se por um lado, como disse Gonçalves (2009: 81), com o uso da LC “palavras podem ser selecionadas dentro de milhares, sem a necessidade de ler todos os textos com que se está trabalhando”, por outro essa não-obrigatoriedade de leitura *não é* obrigatória. E a autora reforça (2009: 51):

Os resultados quantitativos produzidos com base no corpus (sic) são assim indicadores numéricos que devem ser discutidos à luz de diferentes posicionamentos teórico-metodológicos, para serem compreendidos. Da mesma forma que o corpus (sic) oferece apenas evidências linguísticas, e não informações, os números extraídos dos dados linguísticos não são ainda informações em si mesmos, precisando ser interpretados pelo pesquisador para que possam servir de apoio para novas descrições linguísticas ou para a proposta de novas perspectivas teóricas.

Porém, a questão da representatividade ainda está em aberto, já que, como diz Berber-Sardinha (2000: 336) o *corpus* “é um artefato produzido para a pesquisa. Assim, se por um lado os textos devam ser naturais (autênticos e independentes do *corpus*), o

corpus (sic) em si é artificial, um objeto criado com fins específicos de pesquisa” – momento em que se poderia duvidar se um objeto artificial conseguiria representar fielmente um objeto real, tal como as línguas naturais. O próprio autor tenta responder a essa questão (2000: 342) dizendo que “a característica mais facilmente associada à representatividade é justamente a extensão do corpus (sic), o que significa em termos simples que para ter representatividade o corpus deve ser o maior possível” – e isso por duas razões: em primeiro lugar, porque quanto mais extenso um *corpus* maior a probabilidade de que palavras de baixa frequência apareçam (sendo que o mesmo vale para a função desempenhada pelas palavras); em segundo lugar, como um *corpus* é uma amostra de uma população cuja dimensão não se conhece (a linguagem como um todo), não se pode estabelecer qual seria o tamanho ideal da amostra capaz de representá-la (2000: 343). Como conclusão, ao se dizer que um *corpus* deve ser representativo, tal representatividade deve ser entendida em termos de extensão – a maior possível. “Um corpus (sic) maior é em geral mais representativo do que um menor devido ao fato de conter mais instâncias de traços linguísticos raros” (2000: 343). Logo,

a representatividade está ligada à questão da probabilidade. A linguagem é de caráter probabilístico (...) e, sendo assim, há a possibilidade de estabelecer uma relação entre traços que são mais comuns e menos comuns em determinado contexto (2000: 343).

O autor diz também que a extensão de um *corpus* comporta três dimensões: número de itens lexicais (“medida da representatividade do corpus (sic) no sentido de que quanto maior o número de palavras maior será a chance do corpus (sic) conter palavras de baixa frequência, as quais formam a maioria das palavras de uma língua”); número de textos (“um número de textos maior garante que este tipo textual, gênero, ou registro, esteja mais adequadamente representado”); número de gêneros, registros ou tipos textuais (“um número maior de textos de vários tipos permite uma maior abrangência do espectro genérico da língua”) (2000: 344).

Ou seja, embora representativo o *corpus* possui limitações – ou, de acordo com Oliveira (2009: 58), deve-se considerar um *corpus* como “fragmento de língua”, mas que, por outro lado, “representa o seu sistema global (ou parte dele) e que, mesmo incompleto e fragmentado, pode refletir as possibilidades de ocorrência de usos linguísticos potenciais (OLIVEIRA; DIAS, 2006)”¹²¹. O *corpus*, assim, é capaz de

¹²¹ OLIVEIRA, L. P.; DIAS, M. C. P. *Representatividade na compilação de corpus: o projeto CORPOBRAS PUC-Rio*. Trabalho apresentado na Jornada de metodologia para recolha e sistematização de corpora para fins dicionarísticos. Rio de Janeiro: União Latina, 2006.

responder apenas a algumas perguntas e não todas, dependendo da pesquisa para a qual ele foi concebido.

Logo, ao se utilizar a LC como metodologia de pesquisa para a busca de ilocuções, o que se pretende é valer-se dos critérios quantitativos de modo a, primeiramente, comprovar se as espécies de ilocuções extraídas da TAF de Austin são de fato frequentes¹²² no italiano e no português-brasileiro, e, em segundo lugar, verificar quais são tais frequências.

Vale mencionar que Tagnin (2013: 109-110) inseriu em sua obra exemplos de informações que dizem respeito ao “(...) nível *pragmático* da convencionalidade (...)” a partir de investigação da língua por intermédio de *corpora*. No seu caso foram oferecidos exemplos de marcadores conversacionais que indicam:

- opinião: ‘Acho que talvez...’; ‘Sinceramente, acho...’; ‘A meu ver...’;
- restrição: ‘Sim, mas veja...’; ‘Tudo bem, mas...’; ‘Mas o problema é...’;
- digressão: ‘Isso me lembra...’; ‘Por falar em...’;
- sugestão: ‘O que você pode fazer é...’.

Assim, o que se pretende aqui é algo semelhante: pesquisar em *corpora* informações que dizem respeito ao nível pragmático da convencionalidade, mas, dessa vez, referentes a ilocuções *convencionalmente* associadas a itens lexicais específicos (verbos ilocucionários na primeira pessoa...).

3.5. Delimitação dos corpora: Corpus Paisà e Corpus do Português

O *corpus* de italiano *Paisà* é uma das fontes utilizadas na presente investigação. Segundo informações dadas por seus desenvolvedores (PAISÀ, 2014), o projeto *Paisà* foi criado com “o objetivo de contribuir para superar barreiras tecnológicas” que eventualmente possam impedir o acesso, por parte de um público amplo, à grande quantidade de textos escritos em italiano contemporâneo. O projeto volta-se primordialmente a imigrantes italianos que possuam o italiano como língua materna,

¹²² Considera-se oportuna a pergunta que faz Oliveira (2009: 62): “até que ponto podemos confiar em nossas intuições linguísticas para explicar algumas questões relativas ao uso da língua, e em que ocasiões as evidências linguísticas são essenciais?” Considera-se tal questionamento fundamental para que se compreenda a utilidade de investigação da língua via *corpora*.

mas que a utilizam pouco, e imigrantes de outras nacionalidades que vivem na Itália e que possuem o italiano como segunda língua.

O *corpus* do projeto *Paisà* é composto, segundo informações dispostas em seu site (2014), de

uma ampla coleção de textos autênticos da língua italiana extraídos da internet. A coleta de textos contemporâneos foi criada no âmbito do projeto *Paisà* (Plataforma para Aprendizagem do Italiano por *corpora* anotado), com o objetivo de fornecer material autêntico e disponível gratuitamente para aprendizagem do italiano¹²³.

Ainda segundo o site (2014), o projeto durou de 2009 a 2012 e os textos selecionados a compor o *corpus* foram extraídos da internet entre setembro e outubro de 2010, este último composto por cerca de 250 milhões de *tokens*¹²⁴ – número também apontado por Lyding *et al* (2014: 36-37). Ainda com os últimos autores (2014: 36-37), o *corpus* contém cerca de 388 mil documentos a partir de 1067 sites diferentes – em geral os da Fundação Wikimedia (Wikipedia, Wikibooks, Wikinews etc).

O *corpus Paisà* é disponibilizado para download e pode também ser consultado através de sua interface *online* (LYDING ET AL, 2014: 40). Vale dizer com os mesmos autores que a interface foi projetada para fornecer dois tipos básicos de pesquisa: um através de busca simples, onde “termos de pesquisa individuais, bem como combinações de várias palavras ou sequências, podem ser pesquisados através de sua inserção em uma caixa de texto simples” (2014: 40), e outro através de busca avançada, que permite a investigação segundo critérios de anotação linguística, entre outras opções de refinamento¹²⁵. Os resultados, por sua vez, são exibidos como KWIC (palavra-chave em contexto), com o item lexical centrado, em meio aos demais componentes da sentença a qual pertence.

¹²³“(…) ampia collezione di testi autentici in lingua italiana tratti da Internet. La raccolta di testi contemporanei è stata creata nell’ambito del progetto PAISÀ (Piattaforma per l’Apprendimento dell’Italiano Su corpora Annotati) allo scopo di fornire materiale autentico e disponibile gratuitamente per l’apprendimento dell’italiano”.

¹²⁴ Segundo Ribeiro (2004: 170), os *tokens* correspondem ao número de itens lexicais de um *corpus*, contando as repetições, enquanto que *types* correspondem ao número de itens desconsiderando as repetições. P. ex., uma sentença tal como “minha casa é amarela, sua casa é azul” possui oito *tokens* (“minha”, “casa”, “é”, “amarela”, “sua”, “casa”, “é”, “azul”) e seis *types* (“minha”, “casa”, “é”, “amarela”, “sua”, “azul”).

¹²⁵ Embora não haja possibilidade de refinamento de busca segundo tipo textual, tal como há no *Corpus do Português*.

Figura 5: página inicial do *Corpus Paisà*.

Por sua vez, o *Corpus do Português*¹²⁶, segundo informações do site (2014, online) foi criado por Michael Ferreira, da *Georgetown University*, com apoio de Mark Davies, da *Brigham Young University*. Davies criou o projeto *CORDE* no final de 1990, o primeiro grande *corpus* histórico do espanhol (Davies 2009: 140), e colaborou com a criação do *Corpus do Português*, concluído em 2006 e revisto em 2008.

Dos 45 milhões de itens lexicais do *Corpus do Português*, cerca de 15 milhões correspondem a documentos que compreendem os séculos 1200 a 1400, 10 milhões os séculos 1500 a 1700, e 15 milhões os séculos 1800 a 1900. Para os anos 1700 em diante, os textos são divididos entre Portugal e Brasil, e para 1900 estão igualmente divididos entre falado, fictício, jornalístico e acadêmico. Pode-se optar por quais textos comporão a pesquisa, bem como refinar a busca a partir de várias anotações, como século a ser pesquisado, variante nacional a ser pesquisada etc. Detalhes completos sobre cada um dos mais de 60 mil textos podem ser encontrados e os usuários podem baixar uma lista de arquivos Excel com todos eles (DAVIES, 2009: 140-141).

O *Corpus do Português* oferece, tal como o *Paisà*, como resultado, palavra-chave em contexto (KWIC), fornecendo informações acerca do número de ocorrências etc. Na presente pesquisa, restringiu-se a busca ao século XX, com textos compostos da variante brasileira do português presente em todos os tipos textuais (falado, ficção, jornalístico e acadêmico).

¹²⁶ Vale dizer que os textos brasileiros do século XX foram obtidos em parceria com o LacioWeb, da USP.

Figura 6: página inicial do *Corpus do Português*.

Importante dizer que o critério para escolha de tais *corpora* (*Paisà e Corpus do Português*) foi o de representatividade¹²⁷. Tal como anteriormente exposto por Berber-Sardinha (2000) no subcapítulo 3.4, quanto mais extenso for um *corpus* maior a probabilidade de que contenha uma ampla variedade de expedientes linguísticos, bem representando assim a língua geral. Ambos os *corpora* aqui escolhidos, de acordo com a escala¹²⁸ de tamanho de *corpus* proposta por Berber-Sardinha (2003: 119), podem ser considerados grandes, uma vez que possuem mais de 10 milhões de palavras. Ainda segundo ele, “se se quiser construir um corpus (sic) geral de uma língua, deve-se fazer uma escolha a mais variada possível: ela deve incluir o maior número possível de registros encontrados na língua-alvo” (2000: 337). Cabe dizer que essa variedade existe nos *corpora* escolhidos¹²⁹.

Uma questão poderia ser colocada: por que não trabalhar com *corpora* orais, já que em geral esses são mais adequados para a investigação de ilocuições e demais fenômenos pragmáticos? Um exemplo seria a utilização de técnicas de gravação e/ou

¹²⁷ Também Coelho e Rezende (2012: 323) adotaram em suas pesquisas o *Corpus do Português*, em função de sua representatividade.

¹²⁸ Tal escala, baseada no levantamento do tamanho dos corpora elaborados por pesquisadores da comunidade linguística é a seguinte: menos de 80 mil: pequeno; de 80 a 250 mil: pequeno-médio; de 250 mil a 1 milhão: médio; de 1 milhão a 10 milhões: médio-grande; 10 milhões ou mais: grande.

¹²⁹ Posteriormente Berber-Sardinha (2004) questiona essa escala de tamanho. Entretanto, para o presente trabalho manteve-se a adoção da antiga escala proposta pelo autor (2000), pois, como se disse, ambos os *corpora* são representativos no sentido de serem compostos por grande quantidade de palavras/ textos. O próprio Berber-Sardinha (2000) ressalta que qualquer escala de tamanho utilizada é útil como indicação, mas que não deve ser tomada ao pé da letra.

filmagem de diálogos orais autênticos (seguidas de sua transcrição e exame), ou, ainda, transcrição e exame de eventos televisivos¹³⁰, obras cinematográficas e outros do gênero audiovisual, pois seria assim possível observar gestos e expressões faciais dos interlocutores, pausas e hesitações etc – o que certamente auxiliaria no reconhecimento de atos de fala. Em suma, tal metodologia é reconhecidamente adequada para o *estabelecimento* de espécies de ilocuções.

Em primeiro lugar, como já salientado (1.2), a investigação que se realiza aqui é de natureza metalexicográfica, tendo por objetivo *sugerir a dicionarização* de espécies de ilocuções – e não *estabelecer* espécie de ilocuções. A pesquisa de ilocuções que se pretende aqui não é um fim em si mesmo, pois o fim é sua dicionarização (ou melhor, sugerir um modo para dicionarizá-las). Logo, vislumbra-se a possibilidade de que a Lexicografia, tal como apontado anteriormente (1.1), possa se valer de estudos e resultados de outras disciplinas, de modo a atingir propósitos específicos. Assim, tendo-se em vista que a dicionarização pressupõe padrões linguísticos, optou-se por *corpora* escritos, já que, se não mais extensos¹³¹, ao menos são mais representativos da língua geral, uma vez que *corpora* orais podem apresentar muitos traços de regionalismo – cujo valor para um dicionário de língua geral é restrito (a não se que se esteja pensando em um dicionário de regionalismos, o que não é o caso).

Outra necessidade da dicionarização é a economia de tempo e de custos, pois a compilação de *corpora* orais é mais difícil e custoso do que *corpora* escritos, motivo pelo qual, novamente, optou-se por *corpora* escritos. Por fim, ainda que o processo de compilação de ambos fosse igual: não se encontra disponíveis online *corpora* orais tanto quanto é fácil de se obter *corpora* escritos – como no caso do C-Oral (<http://www.c-oral-brasil.org/>), *corpus* de textos orais que, infelizmente, não está livremente disponível *online*.

¹³⁰ Tal como realizado pela Prof^a. Dr^a. Silvia Kaul de Marlageon, da Universidad de Rio Cuarto, Argentina, segundo comunicação proferida em dezembro de 2014 na conferência “Descortesía de fustigación de personajes mediáticos: entre la extimidad del hablante y la extimidad impuesta o atribuída al oyente”, organizada pelo PPG em Filologia e Língua Portuguesa do Depto. de Letras Clássicas e Vernáculas da FFLCH/ USP.

¹³¹ A maioria dos *corpora* escritos existentes são mais extensos do que os orais, provavelmente pela maior facilidade de serem compostos; contudo, não é impossível que existam ou venham a existir *corpora* orais mais extensos do que os escritos. Entretanto, como é mais frequente encontrar traços sociolinguísticos na oralidade do que na escrita, optou-se pela escrita, devido aos interesses deste projeto.

4. ANÁLISE DE CORPORA

Neste capítulo são apresentados os procedimentos para adoção dos dez verbos ilocucionários (cinco do italiano, cinco do português-brasileiro). Em seguida esses são investigados nos *corpora* adotados, de modo a se poder verificar se são de fato frequentes em cada uma das línguas e se, em caso afirmativo, obter qual o número total de ocorrências de cada um. Em seguida, uma análise qualitativa é realizada para excluir casos anômalos, obtendo-se assim o número real de ocorrências de cada caso e as diferenças quantitativas entre as línguas, a partir do qual fez-se a exclusão de duas ilocuções (em função da baixa frequência).

4.1. Critérios para busca de ilocuções via LC

Os critérios para a investigação de atos de fala em *corpora* linguísticos foram os seguintes – alguns dos quais já mencionados:

(a) adoção de cinco verbos ilocucionários italianos (*Promettere, Assicurare, Garantire, Giurare, Impegnarsi*) da classe austiniana das ilocuções comissivas, em função de serem entendidos como sinônimos, segundo o *Sinonimi e contrari* (2009) (2.4.1);

(b) adoção dos cinco verbos correspondentes em português-brasileiro (*Prometer, Assegurar, Garantir, Jurar, Comprometer-se*), em função da equivalência semântica que respectivamente possuem em relação aos verbos italianos (2.4.1);

(c) adoção da forma verbal na primeira pessoa do singular do presente do indicativo da voz ativa como sintaxe de busca para investigação em *corpora* (2.4).

Como neste trabalho se tem por base o dicionário híbrido *Parola Chiave* (2012), destinado a estudantes brasileiros do italiano, considerou-se que a pesquisa devesse iniciar com verbos da língua italiana pautando os passos seguintes – pois não se consideraria sensato partir da língua-alvo ao invés da língua-fonte numa obra monodirecional para decodificação.

Viu-se anteriormente (2.3.2) uma lista com alguns dos verbos comissivos propostos por Austin, a saber:

Classes/ Espécies	Comissivos
-------------------	------------

Em inglês ¹³²	<i>promise, covenant, undertake, contract, mean to, guarantee, oppose, contemplate, favour, am determined to</i>
--------------------------	--

Tabela 2: ilocuições comissivas propostas por Austin (1990[1975]: 127).

A primeira tarefa, então, foi obter verbos italianos correspondentes a estes. Para tal, adotaram-se os verbos que Sbisà (2009[1989]: 71) aponta em sua obra como os comissivos mais representativos de sua língua, a saber:

Classes/ Espécies	Comissivos
Em italiano	<i>Promettere, impegnarsi, proporsi di, proporre, giurare, scommettere, acconsentire, sostenere, opporsi</i>

Tabela 3: verbos ilocucionários comissivos propostos por Sbisà (2009[1989]: 71).

Note-se que alguns desses verbos possuem certa sinonímia, tal como se pode confirmar ao examiná-los no dicionário *Sinonimi e contrari minore Zanichelli* (2009). Neste ponto, vale a pena reforçar a ressalva já feita (ver 2.4.1): *sinonímia* é aqui entendida em sentido comum, i. e., como relação entre palavras de significado próximo – e não em sentido técnico, pois como apontou Zucchi (2010b) inexitem sinônimos absolutos, motivo pelo qual sugere ser mais adequado utilizar, em contextos técnicos, o termo proposto por Greimas (*parassinonímia*) para conotar a identidade parcial entre itens lexicais (identidade reconhecida segundo possibilidade de mútua substituição em determinados contextos).

O primeiro dos verbos ilocucionários presentes na lista de Sbisà é *Promettere* (na de Austin, *Promise*), e ao consultá-lo no referido dicionário obtiveram-se os seguintes sinônimos:

(i) *promettere* (2009: 665): *assicurare, garantire, giurare, impegnarsi*.

Veja-se que entre os sinônimos de *Promettere* aparecem dois verbos presentes na lista de Sbisà, a saber, *Impegnarsi* e *Giurare*. Após consultá-los no *Sinonimi e contrari*, notou-se os seguintes sinônimos de cada um deles:

(ii) *giurare* (2009: 387): *promettere solennemente, affermare solennemente, attestare, asserire, asseverare, assicurare, dichiarare; prestare giuramento*.

(iii) *impegnarsi* (2009: 417-418): *obbligarsi, incaricarsi, promettere, imbarcarsi, impegnarsi, vincolarsi, ingolfarsi; esporsi, comprometersi*.

¹³² Tais como propostas por Austin (1975[1962]: 153-163).

Neste ponto, nota-se que como sinônimo de *Giurare* aparece o já mencionado *Promettere* e também *Assicurare* (ambos mutuamente sinônimos). No caso de *Impegnarsi*, aparece como seu sinônimo o já citado *Promettere*. Consultando-se, então, a entrada *Assicurare*, obteve-se como resultado:

(iv) assicurare (2009: 81): *rendere sicuro, mettere al sicuro, dar sicurezza; difendere, proteggere, preservare, cautelare, tutelare, salvaguardare, riparare; ancorare, attaccare, agganciare, fermare, legare; salvare; dire, affermare, dichiarare; attestare, dimostrare, certificare; testificare, testimoniare; giurare, asseverare; promettere; garantire.*

Aqui aparecem como sinônimos *Promettere* e *Giurare*, já mencionados, e *Garantire*, sinônimo também de *Promettere*. Tal verbo, então, é o próximo a ser consultado no referido dicionário:

(v) garantire (2009: 373): *assicurare, mallevare, dar cauzione, cauzionare, avallare, coprire; cautelare; sponsorizzare; salvaguardare, tutelare, proteggere, preservare; confermare, affermare, dare per certo, asserire; promettere; certificare, attestare.*

Este verbo, como se percebe, possui como sinônimos dois verbos já mencionados, *Promettere* e *Assicurare*. Tem-se assim, então, a seguinte tabela¹³³ com os verbos investigados e seus equivalentes semânticos:

	Promettere	Assicurare	Garantire	Giurare	Impegnarsi
Promettere	-	x	x	x	x
Assicurare	x	-		x	
Garantire	x	x	-		
Giurare	x	x		-	
Impegnarsi	x				-

Tabela 4: relação sinonímica entre verbos ilocucionários comissivos italianos pesquisados.

Explica-se, assim, o critério (a). Em seguida, de modo a obter os verbos ilocucionários do português-brasileiro, verificou-se quais são aqueles fornecidos pelo *Parola Chiave* (2012) como equivalentes semânticos dos verbos ilocucionários italianos apontados acima. O resultado, explicando o critério (b), é o seguinte:

¹³³ *Asserire, attestare* e *affermare* são também alguns sinônimos recorrentes de alguns desses verbos, segundo o *Sinonimi e contrari*; porém, não são de *promettere*. Logo, ficaram de fora da presente investigação. Além disso, é de se supor que, segundo a classificação de Austin, tais verbos não pertencem à classe dos comissivos; *affermo*, p. ex., pertence à classe dos expositivos.

Verbos ilocucionários do italiano	Equivalentes semânticos do português-brasileiro dados pelo <i>Parola Chiave</i> (2012)
Assicurare	Assegurar (2012: 49)
Giurare	Jurar (2012: 306)
Garantire	Garantir (2012: 293)
Promettere	Prometer (2012: 587)
Impegnarsi	Comprometer-se (2012: 333)

Tabela 5: verbos ilocucionários comissivos italianos e respectivos equivalentes semânticos do português-brasileiro dados pelo *Parola Chiave* (2012).

Por fim, quanto ao critério (c), a razão para adoção da forma na primeira pessoa do singular do presente do indicativo na voz ativa se deve ao fato de que, deste modo, obtém-se uma sintaxe de busca padronizada capaz de ser utilizada em todos os casos, além do fato de facilitar o reconhecimento de ilocuções em *corpora*.

4.2. Procedimentos de busca

Cabe então realizar a pesquisa propriamente dita nos *corpora* adotados – em primeiro lugar, no *corpus* italiano e, em segundo, no do português-brasileiro.

4.2.1. Verbos ilocucionários comissivos do italiano

Nesta seção será descrita a pesquisa de tais verbos na forma da primeira pessoa do singular do presente do indicativo na voz ativa no *Corpus Paisà*, de modo a verificar se suas ilocuções correspondentes, semanticamente próximas, são também pragmaticamente próximas. Além da adoção da sintaxe de busca mencionada, a pesquisa foi refinada segundo as possibilidades oferecidas pelo *corpus*, a saber: adoção da anotação gramatical *verbo*. Assim, obteve-se o seguinte resultado¹³⁴:

Verbo ilocucionário/ ilocução	Nº de ocorrências totais
Assicuro	785

¹³⁴ Para ver imagens das primeiras páginas da pesquisa feita com o *corpus*, ver Anexo 1. Vale dizer que as cinco pesquisas feitas no *Corpus Paisà* foram realizadas no dia 30 de junho de 2014.

Giuro	318
Garantisco	158
Prometto	114
Mi impegno	31

Tabela 6: nº de ocorrências totais de ilocuições comissivas pesquisadas no *Corpus Paisà*.

Percebe-se que ao contrário do que a intuição indicava, “Prometto” não é a ilocução comissiva com maior número de ocorrências no *corpus* da língua italiana. Tal intuição, além de ser reforçada pelo fato de Austin e Sbisà terem apontado, respectivamente, “Promisse” e “Prometto” como a primeira das ilocuições comissivas, foi reforçada também pelo fato de que dos cinco verbos investigados no dicionário *Sinonimi e contrari* (2009), *promettere* é o único apontado como sinônimo de todos os outros. A ilocução com maior número de ocorrências é, na verdade, *assicuro*, seguida por *giuro* e *garantisco*.

Entretanto, surge aqui a seguinte questão: será que tais verbos ilocucionários desempenham de fato sua função ilocucionária *convencional*? Diante de tal questionamento, o passo seguinte foi salvar em arquivo Word todas as ocorrências dos cinco verbos pesquisados, de modo a verificar qualitativamente se de fato desempenham a força ilocucionária esperada.

O resultado foi o seguinte: tais desempenhos não ocorreram em todos os casos. Assim, casos anômalos foram excluídos da contagem real, a saber, casos de:

- repetição de ocorrências;
- negação (“*non prometto...*” etc);
- homonímia (p. ex., *giuro* como substantivo e não como verbo na primeira pessoa...);
- proferimentos/ ocorrências ininteligíveis.

Abaixo, alguns exemplos de exclusão.

9	Vi consiglio di provare se non altro per la soddisfazione di preparare deliziosi finger food , che visti dal di fuori sembrano difficilissimi da realizzare con 4 ore di corso , ma vi assicuro il risultato è garantito .
10	Vi consiglio di provare se non altro per la soddisfazione di preparare deliziosi finger food , che visti dal di fuori sembrano difficilissimi da realizzare con 4 ore di corso , ma vi assicuro il risultato è garantito .

Tabela 7: exemplos de casos excluídos referentes à ilocução comissiva “assicuro”.

A ocorrência de número 10 da ilocução “Assicuro” é repetição da ocorrência de número 9.

36	Pertanto , non garantisco circa i risultati ...
----	--

Tabela 8: exemplo de caso excluído referente à ilocução comissiva “garantisco”.

112	Non uccidete- mi , domani avrete il commento sul film ! (o magari stasera stesso , ma non prometto nulla) .
-----	--

Tabela 9: exemplo de caso excluído referente à ilocução comissiva “prometto”.

A ocorrência de número 36 de “Garantisco” constitui uma negação, ou seja, não constitui a ação de garantir. O mesmo se dá com a ocorrência de número 112 de “Prometto”, razão pela qual a ação de prometer não se desenvolve.

105	3) Ciò che prima di ogni altra cosa qualsiasi malato (presente o futuro) vorrebbe sentir- si dire non è : « garantisco la tua autonomia » , bensì : « garantisco che non sarai mai abbandonato » .
-----	--

Tabela 10: exemplo de caso excluído referente à ilocução comissiva “garantisco”.

275	Nel novembre del 2005 esce il libro " Eros : Lo giuro " , biografia ufficiale di Eros Ramazzotti , che racconta la sua vita e i suoi successi .
-----	--

Tabela 11: exemplo de caso excluído referente à ilocução comissiva “giuro”.

30	I regali per me sono sempre un incubo : ci penso da giugno , ma allora o sono io che mi impegno poco o il tempo passa veramente in fretta : perchè mi ritrovo sempre che il 23 dicembre devo correre su e giù per tutti i negozi-mercatini-secchioni della spazzatura che trovo sul mio cammino .
----	--

Tabela 12: exemplo de caso excluído referente à ilocução comissiva “mi impegno”.

As ocorrências de número 105 de “Garantisco” e 275 de “Giuro” não constituem ações correspondentes. O primeiro caso é uma menção a um enunciado; no segundo é uma menção ao título de uma obra. Já no caso da ocorrência de número 30 de “Mi impegno”, tem-se uma disjunção: “*o sono io che mi impegno poco o il tempo passa veramente in fretta*”.

Feita a “limpeza”¹³⁵, obteve-se os números reais:

¹³⁵ Como a “limpeza” foi feita manualmente, tem-se consciência de que um ou outro caso anômalo pode ter sido incluído indevidamente na contagem das ocorrências reais. Entretanto, dado o caráter crítico-metodológico do trabalho, tal equívoco, se de fato aconteceu, não invalida/ inviabiliza a proposta final.

Verbo ilocucionário/ ilocução	Nº de ocorrências reais
Assicuro	704
Giuro	272
Garantisco	143
Prometto	89
Mi impegno	23

Tabela 13: nº de ocorrências reais de ilocuições comissivas pesquisadas no *Corpus Paisà*.

Comparando os números totais com os números reais, percebe-se o seguinte:

Verbo ilocucionário/ ilocução	Nº de ocorrências totais	Nº de ocorrências reais
Assicuro	758	704
Giuro	318	272
Garantisco	158	143
Prometto	114	89
Mi impegno	31	23

Tabela 14: comparação entre o nº de ocorrências totais e reais de ilocuições comissivas pesquisadas no *Corpus Paisà*.

Alterações consideráveis ocorreram na passagem da contagem de ocorrências totais para reais. Entretanto, a posição ocupada pelas ilocuições permaneceu a mesma.

4.2.2. Verbos ilocucionários comissivos do português-brasileiro

O próximo passo, então, foi pesquisar as ilocuições do português-brasileiro a partir da mesma sintaxe de busca utilizada para a busca de ilocuições italianas: verbos ilocucionários na forma da primeira pessoa do singular do presente do indicativo da voz ativa. A investigação se deu, como se sabe, no *Corpus do Português*. Tal como no caso da pesquisa feita no *Corpus Paisà*, utilizou-se como filtro a anotação gramatical *verbo*.

Outro filtro utilizado foi aquele capaz de restringir a pesquisa a textos do português-brasileiro, excluindo-se assim textos do português-europeu¹³⁶. Eis o resultado¹³⁷:

Verbo ilocucionário/ ilocução	Nº de ocorrências totais
Juro	5055
Garanto	505
Prometo	404
Asseguro	157
Comprometo ¹³⁸	71

Tabela 15: nº de ocorrências totais de ilocuições comissivas pesquisadas no *Corpus do Português*.

Percebe-se aqui uma diferença considerável entre “Juro” e os demais verbos. Contudo, isso se deve ao fato de que, em português, “Juro” pode tanto atuar como verbo na primeira pessoa do singular do presente do indicativo na voz ativa como termo da economia (fazendo referência, grosso modo, ao rendimento de capital investido e/ ou percentagem imposta sobre determinada quantia emprestada). Logo, faz-se necessário, aqui, uma “limpeza”, de modo a serem obtidos números reais, de acordo com os mesmos critérios de exclusão que foram tomados no caso das ilocuições em italiano (ver 4.2.1).

Obteve-se, assim, com a análise qualitativa, o seguinte:

Verbo ilocucionário/ ilocução PT	Nº de ocorrências reais
Garanto	421
Juro	388

¹³⁶ Como se viu anteriormente (3.5), o *Corpus Paisà* é eminentemente composto por textos escritos, enquanto que o *Corpus do Português*, mesmo restrito à variante do português-brasileiro, é composto por textos escritos e orais. Pensou-se, inicialmente, em excluir textos orais, mas por fim optou-se por mantê-los, pelo motivo de que o critério para escolha do *corpus* se deu em função representativa decorrente de sua ampla extensão.

¹³⁷ Para ver imagens das primeiras páginas da pesquisa feita com o *corpus*, ver Anexo 1. Vale dizer que as cinco pesquisas feitas no *Corpus do Português* foram realizadas no dia 22 de agosto de 2014.

¹³⁸ Inclui tanto casos de *comprometo-me* como de *me comprometo*.

Prometo	292
Asseguro	145
Comprometo	60

Tabela 16: nº de ocorrências reais de ilocuções comissivas pesquisadas no *Corpus do Português*.

Houve aqui uma troca de posições: “Juro” passou para o segundo lugar dos mais recorrentes, enquanto que “Garanto” passou para o primeiro.

Exemplos de casos de exclusão:

3	São Paulo , Brasiliense , 1947 , p. 35 . 2 Como nasceram as estrelas . Rio de Janeiro , Rocco , 1999 , pp . 40-41 . M2U8T4 1 Só não juro que o Saci existe porque não se deve ficar jurando à-toa , por aí . Você é provavelmente de cidade e não me acredita . Mas que nas matas tem Saci , lá isso tem . E eu garanto essa
5	, estar contíguo 1) . 2) Em poesia : jurar : adjuro Stygii (Cés . B. Gal . 6 , 33 , 2) . caput fontis (Verg . En . 12 , 816) « juro adjacuí , perf. de adjacè. pela fonte do Estige » . adjectiõ , -õnis , subs. m. 1) Adição , aumento , 2. adjúrõ = adjavero . anexaço : Romana res adjectio-ne populi adjútõ

Tabela 17: exemplos de casos excluídos referentes à ilocução comissiva “juro”.

Aqui, a ocorrência de número 3 de “Juro” foi desconsiderada por ser uma negação. Já a ocorrência de número 5 foi excluída por ser ininteligível.

186	esteja presente em uma mensagem constitui-se como condição necessária para sua validade ou eficácia comunicativa . 6Denomina-se performativo o proferimento lingüístico cuja execução consiste na realização de determinado ato . Assim , o proferimento da frase \ " eu prometo voltar amanhã \ " é a efetuação do ato de prometer . É também a denominação do verbo que marca a performatividade do proferimento . 7Enunciação é o ato individual de utilização da língua . Opõe-se a enunciado
-----	--

Tabela 18: exemplo de caso excluído referente à ilocução comissiva “prometo”.

A ocorrência nº 186 de “Prometo” foi excluída ser menção a um enunciado.

43	trabajo remunerado . 4. Sigo a la espera de tus prometidos panoramas de las letras brasileñas. ¿ No has pensado reunir tus ensayos sobre moderna literatura brasileña en un volumen ? 361 Serían utilísimos y desde ya me comprometo a agenciar una editorial que los tradujera al español . 5. No sé qué nuevos libros interesantes brasileños han aparecido ni en qué estás actualmente trabajando . Sugierele a tu amigo el de la librería que me envíe catálogos
----	---

Tabela 19: exemplo de caso excluído referente à ilocução comissiva “comprometo”.

Por sua vez, a ocorrência de nº 43 de “Comprometo” foi excluída por estar em espanhol – incluída nos casos de ininteligibilidade.

Veja-se agora a comparação entre os números totais e reais busca no *Corpus do Português*:

Verbo ilocucionário/ ilocução	Nº de ocorrências totais	Nº de ocorrências reais
Juro	5055	388
Garanto	505	421
Prometo	404	292
Asseguro	157	145
Comprometo	71	60

Tabela 20: comparação entre o nº de ocorrências totais e reais de ilocuições comissivas pesquisadas no *Corpus do Português*.

Por fim, a comparação entre os verbos ilocucionários em italiano e em português-brasileiro:

Verbo ilocucionário/ ilocução PT	Nº de ocorrências reais	Verbo ilocucionário/ ilocução IT	Nº de ocorrências reais
Garanto	421	Assicuro	704
Juro	388	Giuro	272
Prometo	292	Garantisco	143
Asseguro	145	Prometto	89
Comprometo	60	Mi impegno	23

Tabela 21: comparação entre o nº de ocorrências reais de ilocuições comissivas pesquisadas no *Corpus Paisà* e no *Corpus do Português*.

Percebe-se uma diferença considerável entre as línguas: enquanto que, no caso da língua italiana, a ilocução comissiva com maior número de ocorrências é “Assicuro”, no caso do português-brasileiro é “Garanto”.

Importante dizer aqui que em ambas as línguas as ilocuições pesquisadas são frequentes – embora, naturalmente, umas mais frequentes que outras. Logo, os resultados das análises de *corpora* justificam a inserção de tais ilocuições em

dicionários. É possível, entretanto, que *mi impegno* e *comprometo-me/ me comprometo* fossem excluídas devido ao baixo número de ocorrências – p. ex., se um lexicógrafo tivesse realizado uma pesquisa com 50 ilocução em cada uma das línguas, provavelmente aquelas com menos ocorrência fossem excluídas. A título de exemplo, tais ilocução serão excluídas da próxima etapa do trabalho.

5. SUGESTÃO DE INCLUSÃO DE ILOCUÇÕES EM DICIONÁRIOS HÍBRIDOS

Chega-se, finalmente, ao capítulo onde é feita a sugestão de inserção de ilocuições em dicionários híbridos italiano>português-brasileiro. Vale dizer que tal sugestão é pautada pela composição tripartite dos dicionários (microestrutura, medioestrutura e macroestrutura), a saber:

- quanto à microestrutura, inserir marcas de uso referentes às classes de ilocução no interior dos verbetes das espécies de ilocução realizadas segundo proferimento de seus verbos ilocucionários correspondentes;
- quanto à microestrutura-medioestrutura: inserir remissivas após as marcas de uso, de modo a conduzir o leitor a uma seção externa à nomenclatura do dicionário;
- quanto à macroestrutura: inserir um texto externo que contenha: (i) explicações referente às classes de ilocuições e lista de respectivas espécies convencionalmente recorrentes do italiano dispostas segundo frequência; (ii) espécies de ilocuições equivalentes do português-brasileiro dispostas segundo frequência; (iii) exemplos de uso de tais verbos ilocucionários desempenhando sua função ilocucionária convencional retirados dos *corpora*.

Vale agora explicitar como se chegou a tal proposta.

5.1. Microestrutura

Como se viu anteriormente (1.5), por microestrutura entende-se a parte do dicionário que diz respeito ao conjunto das informações ordenadas de cada verbete após sua entrada. Em primeiro lugar, então, vale apresentar como os verbetes dos quatro verbos ilocucionários italianos¹³⁹ mais recorrentes investigados anteriormente são dispostos pelo PC (2012):

assicuràre (2012: 49)

v.tr. **1** Rendere sicuro § garantire: *risparmia per assicurare l'avvenire alla famiglia*. **2** Dare per certo; affermare con sicurezza: *lo assicurarono che non sarebbe successo nulla; assicura di non saperne niente*. □ **assegurar** **3** Fermare, fissare saldamente: *assicura le imposte prima di uscire* □ **fixar** **4** Proteggere un

¹³⁹ *Impegnarsi* foi excluída devido à sua baixa frequência. O mesmo ocorreu *comprometer-se*.

bene da eventuali rischi e danni facendo un'assicurazione: *assicurare l'auto contro il furto* □ **pôr no seguro** ♦ **assicurarsi** *v.pr.* **1** Accertarsi, controllare: *assicurati che le porte siano tutte chiuse* □ **assegurar-se** **2** Garantirsi da eventuali rischi e danni facendo un'assicurazione: *assicurarsi contro il furto* □ **fazer seguro**.

giuràre (2012: 306)

v.tr. **1** Dichiarare, promettere qualcosa con un giuramento: *giuro di dire la verità; giurami che non mi tradirai; giurare a Dio, sulla Bibbia, sul proprio onore*. □ **jurar** | *Giurare il falso* = mentire sotto (spec. in un processo) □ **prestar falso-testemunho** ♦ Affermare con certezza **é** assicurare: *ti giuro che non lo sapevo* □ **jurar** | *Non ci giurerei* = non ne sono sicuro □ **não apostaria nisso** | *Giurarla a qualcuno* = fare propositi di vendetta nei suoi confronti □ **jurar vingança contra alguém**.

garantìre (2012: 293)

v.tr.[*garantisco, garantisci*]. **1** Assicurare sotto la propria responsabilità il rispetto di un impegno proprio o altrui: *garantire la restituzione di un prestito* ♦ Assicurare la qualità e il perfetto funzionamento di una merce, impegnandosi con l'acquirente a ripararla o a sostituirla gratuitamente qualora si guasti entro un certo periodo: *la fabbrica garantisce questo televisore per un anno* **2** Dare per certo, assicurare: *ti garantisco che ha detto la verità; la buona pensione gli garantisce una vecchiaia tranquilla* □ **garantir** ♦ **garantirsi** *v.pr.* Assicurarsi, prendere delle precauzioni contro eventuali danni o rischi □ **garantir-se**.

prométtere (2012: 587)

v.tr.[*coniugato come mettere*]. **1** Impegnarsi a fare o a dare qualcosa o a comportarsi in un certo modo: *mi ha promesso di venire, di aiutarmi, di non disturbarmi più; promettere un regalo, un lavoro*. **2** ♣ Far sperare, far prevedere: *un cielo che promette bel tempo; il suo atteggiamento non promette niente di buono; un ragazzo che promette bene*. □ **Prometer**.

Feito isso, apontam-se três possibilidades de inserção de ilocuições na microestrutura de um dicionário tal como o *PC*: (1) marcas de uso; (2) notas de uso; (3) exemplos de uso.

5.1.1. Marcas de uso

Como visto anteriormente (Introdução, I e 1.5.1), o *Diccionario Salamanca* (1996) oferece informações ilocucionárias a partir de marcas de uso. Veja-se aqui um caso concreto desse dicionário:

siempre (1996: 1460)

adv. temp. **1** En todo momento o todo el tiempo. *El escritor siempre fumaba puros. Te echaré de menos siempre*. OBSERVACIONES: Normalmente se refiere al periodo de tiempo del que se habla o a uno que se da por consabido. (...) **4 AFIRMACIÓN**. Resalta una cosa que se dice o refuerza una afirmación que no admite duda: *Siempre vivirás mejor solo que mal acompañado*.

Nota-se que na aceção de número 4 o *Salamanca* insere a marca de uso *Afirmación*, seguida de um exemplo – sendo possível inferir ser esta uma marca de uso pragmática em função da explicação dada na introdução do dicionário. Importante notar também que a informação pragmática é disposta após as informações sintáticas e

semânticas, de acordo com o que disse Borba (2003: 309) a respeito da disposição-padrão de informações microestruturais:

O primeiro passo para a organização de um dicionário de língua é descrever o uso, quer dizer, mostrar como funcionam efetivamente os diversos setores da língua num determinado estágio de seu desenvolvimento. (...) Tal seleção está relacionada com o objetivo específico do dicionário. Se o objetivo é registrar todo o uso, então caberiam informações relacionadas com todos os níveis da estrutura linguística. Quanto à fonética, caberiam transcrições, informações sobre a prosódia, além da ortografia naturalmente; quanto à morfologia, dar-se-ia atenção principalmente ao sistema flexional com informações plurais e femininos irregulares, conjugação de verbos irregulares e defectivos etc. Em seguida, viriam as informações sintáticas, semânticas e pragmáticas. Se se pensa em uso como combinatória léxica para efeito de comunicação, então serão estas três últimas as privilegiadas.

Neste caso, o mesmo poderia ser feito no *PC* (aqui, as sugestões serão feitas em sublinhado vermelho, de modo a facilitar a visualização). Entretanto, considera-se que tais não são boas soluções. Tomem-se como exemplos os verbetes de *assicurare* e *giurare* e suas marcas de uso indicativas da espécie de ato de fala correspondentes:

assicuràre

v.tr. **1** Rendere sicuro § garantire: *risparmia per assicurare l'avvenire alla famiglia*. **2** Dare per certo; affermare con sicurezza: *lo assicurarono che non sarebbe successo nulla; assicura di non saperne niente*.
□ **assegurar** Assegurar **3** Fermare, fissare saldamente: *assicura le imposte prima di uscire* □ **fixar** **4** Proteggere un bene da eventuali rischi e danni facendo un'assicurazione: *assicurare l'auto contro il furto*
□ **pôr no seguro** ♦ **assicurarsi** *v.pr.* **1** Accertarsi, controllare: *assicurati che le porte siano tutte chiuse* □ **assegurar-se** **2** Garantirsi da eventuali rischi e danni fazendo un'assicurazione: *assicurarsi contro il furto*
□ **fazer seguro**.

giuràre

v.tr. **1** Dichiarare, promettere qualcosa con un giuramento: *giuro di dire la verità; giurami che non mi tradirai; giurare a Dio, sulla Bibbia, sul proprio onore*. □ **jurar** Jurar | *Giurare il falso* = mentire sotto (spec. in un processo) □ **prestar falso-testemunho** ♦ Affermare con certezza § assicurare: *ti giuro che non lo sapevo* □ **jurar** | *Non ci giurerei* = non ne sono sicuro □ **não apostaria nisso** | *Giurlarla a qualcuno* = fare propositi di vendetta nei suoi confronti □ **jurar vingança contra alguém**.

Os problemas desta sugestão são dois: a redundância e a confusão que pode causar ao consulente entre a força ilocucionária e o significado semântico. A redundância se dá porque a marca de uso que designa a espécie de ato de fala é idêntico ao equivalente semântico fornecido em português-brasileiro. Como decorrência, tem-se a confusão entre ambas as dimensões da língua (pragmática e semântica).

E o mesmo aconteceria caso se inserisse como a marca de uso não o nome da ação a se praticar, mas o da ação praticada, tal como se poderia notar no exemplo fornecido abaixo:

promettere

v.tr.[conjugato come *mettere*]. **1** Impegnarsi a fare o a dare qualcosa o a comportarsi in un certo modo: *mi ha promesso di venire, di aiutarmi, di non disturbarmi più; promettere un regalo, un lavoro.* **Promessa 2** ♣ Far sperare, far prevedere: *un cielo che promette bel tempo; il suo atteggiamento non promette niente di buono; un ragazzo che promette bene.* □ **Prometer.**

Importante explicitar que tal redundância não ocorre no caso do *Salamanca* porque neste a espécie de ato de fala é atribuída a uma entrada cujo ato de fala realizado é implícito – pois não é imediatamente evidente que o advérbio *sempre* possa ser usado para constituir a ação de afirmar da mesma forma que é evidente usar o verbo *promettere* para constituir uma promessa.

Dessa forma, conclui-se que a inserção de marcas de uso referentes a espécies de ilocuições só seria informativa caso se desse com ilocuições implícitas. P. ex., caso se verificasse em *corpus* que a expressão “*Fuori!*” é correntemente utilizado como ação de ordenar, então em no verbete da entrada correspondente a tal expressão poderia ser inserida uma marca de uso do tipo “ordem”, indicando assim um tipo de uso de tal item lexical que um estudante brasileiro do italiano poderia vir a conhecer. Contudo, no caso de verbos ilocucionários, cuja força é explícita, possivelmente esta não seria uma alternativa enriquecedora.

Por essa razão, deve-se buscar uma possibilidade que englobe o caso de verbos ilocucionários, a saber: inserir marcas de usos referentes às *classes* de atos de fala em questão. Logo, os verbetes considerados acima ficariam assim:

assicuràre

v.tr. **1** Rendere sicuro § garantire: *risparmia per assicurare l'avvenire alla famiglia.* **2** Dare per certo; affermare con sicurezza: *lo assicurarono che non sarebbe successo nulla; assicura di non saperne niente.* □ **assegurar** **Comissivo** **3** Fermare, fissare saldamente: *assicura le imposte prima di uscire* □ **fixar** **4** Proteggere un bene da eventuali rischi e danni facendo un'assicurazione: *assicurare l'auto contro il furto* □ **pôr no seguro** ♦ **assicurarsi** v.pr. **1** Accertarsi, controllare: *assicurati che le porte siano tutte chiuse* □ **assegurar-se** **2** Garantirsi da eventuali rischi e danni facendo un'assicurazione: *assicurarsi contro il furto* □ **fazer seguro.**¹⁴⁰

giuràre

v.tr. **1** Dichiarare, promettere qualcosa con un giuramento: *giuro di dire la verità; giurami che non mi tradirai; giurare a Dio, sulla Bibbia, sul proprio onore.* □ **jurar** **Comissivo** | *Giurare il falso* = mentire sotto (spec. in un processo) □ **prestar falso-testemunho** ♦ Affermare con certezza § assicurare: *ti giuro che non lo sapevo* □ **jurar** | *Non ci giurerei* = non ne sono sicuro □ **não apostaria nisso** | *Giurarla a qualcuno* = fare propositi di vendetta nei suoi confronti □ **jurar vingança contra alguém.**

¹⁴⁰ Vale dizer que não haveria distinção entre a força de *assicurare* e o significado das duas acepções. Pois se x dissesse a y “*risparmia per assicurare l'avvenire alla famiglia*”, y desempenharia a ilocução de assegurar ao responder, p. ex., “*Te lo assicuro*”. O mesmo com a acepção de *dare per certo*: “*Ti assicuro che non sarebbe successo nulla*”.

Desta forma, evita-se a ambiguidade e a redundância. Entretanto, tal sugestão só faria sentido caso fosse acompanhada de um material externo explicativo informando ao consultante o que significa essa marca (tal como ocorre com o *Salamanca*). Tal explicação pode ser dada tanto na introdução, quanto em outras seções. Isso, contudo, será mais bem investigado posteriormente (5.3).

Afora isso, caso se opte pela inserção de ilocuções por marcas de uso, seria necessário que também em casos implícitos (tal como no caso de “*Fuori!*”) fossem inseridas marcas referentes às classes de ilocuções, e não mais às espécies, de modo a manter uma padronização. Ou, caso o lexicógrafo opte pelos dois tipos de marcas, que informe os critérios para tal – o que não seria fácil. Logo, prefere-se o emprego de marcas de uso apenas relativas à classe de ilocuções.

5.1.2. Notas de uso

Afora esse caso concreto do *Salamanca*, com base na proposta de Duran mencionada no subcapítulo 1.5.3 oferece-se aqui outra sugestão: a de notas de uso.

Logo, tomando-se como exemplo o caso do verbete *assicurare* e *giurare*, faz-se a seguinte sugestão:

assicuràre

v.tr. **1** Rendere sicuro § garantire: *risparmia per assicurare l'avvenire alla famiglia*. **2** Dare per certo; affermare con sicurezza: *lo assicurarono che non sarebbe successo nulla; assicura di non saperne niente*. □ **assegurar** Obs: ilocução da classe dos comissivos, que, ao ser enunciada, compromete o falante a determinada linha de ação. **3** Fermare, fissare saldamente: *assicura le imposte prima di uscire* □ **fixar** **4** Proteggere un bene da eventuali rischi e danni facendo un'assicurazione: *assicurare l'auto contro il furto* □ **pôr no seguro** ♦ **assicurarsi** v.pr. **1** Accertarsi, controllare: *assicurati che le porte siano tutte chiuse* □ **assegurar-se** **2** Garantirsi da eventuali rischi e danni facendo un'assicurazione: *assicurarsi contro il furto* □ **fazer seguro**.

giuràre

v.tr. **1** Dichiarare, promettere qualcosa con un giuramento: *giuro di dire la verità; giurami che non mi tradirai; giurare a Dio, sulla Bibbia, sul proprio onore*. □ **jurar** Obs: ilocução da classe dos comissivos, que, ao ser enunciada, compromete o falante a determinada linha de ação. | *Giurare il falso* = mentire sotto (spec. in un processo) □ **prestar falso-testemunho** ♦ Affermare con certezza § assicurare: *ti giuro che non lo sapevo* □ **jurar** | *Non ci giurerei* = non ne sono sicuro □ **não apostaria nisso** | *Giurarla a qualcuno* = fare propositi di vendetta nei suoi confronti □ **jurar vingança contra alguém**.

Essa possibilidade é visivelmente mais esclarecedora do que o mero rótulo “Comissivo”. Afora isso, pode ser levada a cabo tanto para casos explícitos quanto implícitos de força ilocucionária, o que representa outra vantagem em relação ao uso de

marcas de uso. Entretanto, a desvantagem desta alternativa é sua extensão, que certamente aumentaria o tamanho do dicionário.

Em acréscimo, sabe-se que termos tais como “ato de fala”, “verbo ilocucionário”, “ilocuções”, “força ilocucionária” e correlatos não são publicamente conhecidos tais como os termos “verbo”, “substantivo”, “significado” etc. Logo, informações auxiliares para este caso devem ser dadas, ou na introdução, ou em outro texto auxiliar, indicando o que se pretende dizer com tais termos. Isso é o que se verá em seguida (5.3).

5.1.3. Exemplos de uso

Outra possibilidade é a inserção de exemplos correspondentes ao ato de fala que determinado vocábulo convencionalmente produz. Viu-se no subcapítulo 1.5 que Welker apresenta tal possibilidade, sendo que uma vantagem ao adotá-la é que os exemplos poderiam ser retirados do próprio *corpus* utilizado. Tome-se, como exemplo, o caso de *assicurare*:

assicuràre (2012: 49)

v.tr. **1** Rendere sicuro § garantire: *risparmia per assicurare l'avvenire alla famiglia.* **2** Dare per certo; affermare con sicurezza: *lo assicuraron che non sarebbe successo nulla; assicura di non saperne niente.* □ **assegurar** **3** Fermare, fissare saldamente: *assicura le imposte prima di uscire* □ **fixar** **4** Proteggere un bene da eventuali rischi e danni facendo un'assicurazione: *assicurare l'auto contro il furto* □ **pôr no seguro** ♦ **assicurarsi** *v.pr.* **1** Accertarsi, controllare: *assicurati che le porte siano tutte chiuse* □ **assegurar-se** **2** Garantirsi da eventuali rischi e danni facendo un'assicurazione: *assicurarsi contro il furto* □ **fazer seguro.**

Note-se que os exemplos dados pelo dicionário não se referem às ilocuções correspondentes ao verbo em questão, pois “*Lo assicuraron che non sarebbe successo nulla*” não explicita ação de assegurar, já que é um relato (‘x’ diz a ‘y’ que ‘b’ e ‘c’ asseguraram a ‘d’ que não aconteceu nada). Já no caso de “*Assicura di non saperne niente*”, tal proferimento não realiza a ação de assegurar, mas se constitui enquanto ordem/ pedido (‘x’ pede/ ordena a ‘y’ que lhe assegure de não saber nada a respeito de algo). Logo, exemplos correspondentes aos atos de fala poderiam ser os seguintes:

assicuràre

v.tr. **1** Rendere sicuro § garantire: *risparmia per assicurare l'avvenire alla famiglia.* **2** Dare per certo; affermare con sicurezza: *ti assicuro che non è facile tirare conclusioni; te lo assicuro tu non sai cogliere il punto.* □ **assegurar** **3** Fermare, fissare saldamente: *assicura le imposte prima di uscire* □ **fixar** **4** Proteggere un bene da eventuali rischi e danni facendo un'assicurazione: *assicurare l'auto contro il furto* □ **pôr no seguro** ♦ **assicurarsi** *v.pr.* **1** Accertarsi, controllare: *assicurati che le porte siano tutte*

chiuse □ **assegurar-se** 2 Garantir-se da eventuais riscos e danos fazendo uma seguradora: *assicurarsi contro il furto* □ **fazer seguro**.

Tais exemplos foram retirados do *corpus* (4.2.1) e constituem atos de fala (no primeiro caso, ‘x’ assegura a ‘y’ que não é fácil tirar conclusões; no segundo, ‘x’ assegura a ‘y’ que este último tem dificuldade de captar a questão em discussão).

O problema dessa sugestão é que a simples inserção de exemplos de uso não garante que o consulente infira o ato de fala em questão (talvez pelo fato de que, em geral, exemplos são utilizados como informação sintagmática, não ilocucionária).

Neste caso, uma possibilidade seria combinar exemplos de uso com marcas de uso ou exemplos de uso com notas de uso, explicitando assim a dimensão ilocucionária:

assicuràre

v.tr. 1 Rendere sicuro § garantire: *risparmia per assicurare l'avvenire alla famiglia*. 2 Dare per certo; affermare con sicurezza: ti assicuro che non è facile tirare conclusioni; te lo assicuro tu non sai cogliere il punto. □ **assegurar** **Comissivo** 3 Fermare, fissare saldamente: *assicura le imposte prima di uscire* □ **fixar** 4 Proteggere un bene da eventuali rischi e danni facendo una seguradora: *assicurare l'auto contro il furto* □ **pôr no seguro** ♦ **assicurarsi** v.pr. 1 Accertarsi, controllare: *assicurati che le porte siano tutte chiuse* □ **assegurar-se** 2 Garantir-se da eventuais riscos e danos fazendo uma seguradora: *assicurarsi contro il furto* □ **fazer seguro**.

assicuràre

v.tr. 1 Rendere sicuro § garantire: *risparmia per assicurare l'avvenire alla famiglia*. 2 Dare per certo; affermare con segurança: ti assicuro que não é fácil tirar conclusões; te lo assicuro tu não sai cogliere il punto. □ **assegurar** **Obs: ilocução da classe dos comissivos, que, ao ser enunciado, compromete o falante a determinada linha de ação**. 3 Fermare, fissare saldamente: *assicura le imposte prima di uscire* □ **fixar** 4 Proteggere um bem da eventuais riscos e danos fazendo uma seguradora: *assicurare l'auto contro il furto* □ **pôr no seguro** ♦ **assicurarsi** v.pr. 1 Accertarsi, controllare: *assicurati che le porte siano tutte chiuse* □ **assegurar-se** 2 Garantir-se da eventuais riscos e danos fazendo uma seguradora: *assicurarsi contro il furto* □ **fazer seguro**.

A desvantagem é que, mesmo com tais combinações, os exemplos de uso ainda estão “soltos”, ou seja, sem ligação direta com aquilo que se deseja que eles explicitem. É necessário, assim, pensar em alternativas.

5.2. Medioestrutura

Tal como se viu antes (1.6), medioestrutura nada mais é do que outro termo para designar as remissivas, sinais inseridos em verbetes (logo, dentro da microestrutura) para remeter a materiais externos e/ ou a outras entradas. Neste sentido, duas possibilidades se apresentam: (1) remissivas entre verbetes; (2) remissivas entre verbetes e materiais externos.

5.2.1. Remissivas entre verbetes

O PC possui sistema de remissivas, como se nota no caso do verbeito *sonare* (2012: 734):

sonàre vedi suonàre.

Assim, no caso de espécies de atos de fala pertencentes a uma mesma classe de atos de fala poderia haver remissão de todas as espécies a outra espécie tomada como paradigma de sua classe. Por exemplo:

giuràre

v.tr. **1** Dichiarare, promettere qualcosa con un giuramento: *giuro di dire la verità; giurami che non mi tradirai; giurare a Dio, sulla Bibbia, sul proprio onore.* □ **vedi assicurare jurar** | *Giurare il falso* = mentire sotto (spec. in un processo) □ **prestar falso-testemunho** ◇ *Affermare con certezza* § *assicurare*: ti giuro che non lo sapevo □ **jurar** | *Non ci giurerei* = non ne sono sicuro □ **não apostaria nisso** | *Giurarla a qualcuno* = fare propositi di vendetta nei suoi confronti □ **jurar vingança contra alguém.**

garantìre

v.tr.[*garantisco, garantisci*]. **1** Assicurare sotto la propria responsabilità il rispetto di un impegno proprio o altrui: *garantire la restituzione di un prestito* ◇ *Assicurare la qualità e il perfetto funzionamento di una merce, impegnandosi con l'acquirente a ripararla o a sostituirla gratuitamente qualora si guasti entro un certo periodo: la fabbrica garantisce questo televisore per un anno* **2** Dare per certo, assicurare: *ti garantisco che ha detto la verità; la buona pensione gli garantisce una vecchiaia tranquilla* □ **vedi assicurare garantir** ♦ **garantirsi** v.pr. Assicurarsi, prendere delle precauzioni contro eventuali danni o rischi □ **garantir-se.**

assicuràre

v.tr. **1** Rendere sicuro § *garantire: risparmia per assicurare l'avvenire alla famiglia.* **2** Dare per certo; affermare con sicurezza: *lo assicuraron che non sarebbe successo nulla; assicura di non saperne niente.* □ **assegurar** **Obs: ilocução da classe dos comissivos, que, ao ser enunciado, compromete o falante a determinada linha de ação.** **3** Fermare, fissare saldamente: *assicura le imposte prima di uscire* □ **fixar** **4** Proteggere un bene da eventuali rischi e danni facendo un'assicurazione: *assicurare l'auto contro il furto* □ **pôr no seguro** ♦ **assicurarsi** v.pr. **1** Accertarsi, controllare: *assicurati che le porte siano tutte chiuse* □ **assegurar-se** **2** Garantirsi da eventuali rischi e danni facendo un'assicurazione: *assicurarsi contro il furto* □ **fazer seguro.**

Como, segundo o *corpus*, *assicurare* possuiu um maior número de ocorrências do que os demais, então esse poderia ser tomado como paradigma dos atos de fala comissivos. Neste caso, seria desnecessário que todos os verbetes correspondentes a comissivos possuam notas de uso, bastando que apenas aquele tomado como paradigma a possua – o que impediria o crescimento desordenado do dicionário.

Entretanto, apenas “*vedi verbo x*” não garante o objetivo almejado, uma vez que o consulente poderia se perguntar: “Devo ver *assicurare* para quê?” Logo, uma possibilidade seria a seguinte: manter a nota de uso no verbo-paradigma e inserir nos demais casos o “*vedi verbo x*” mais a marca de uso referente à classe dos atos de fala em questão. Desse modo, ficaria assim:

giurare

v.tr. **1** Dichiarare, promettere qualcosa con un giuramento: *giuro di dire la verità; giurami che non mi tradirai; giurare a Dio, sulla Bibbia, sul proprio onore.* □ **giurar** **Comissivo - vedi assicurare** | *Giurare il falso* = mentire sotto (spec. in un processo) □ **prestar falso-testemunho** ◇ *Affermare con certezza* **§ assicurare**: *ti giuro che non lo sapevo* □ **giurar** | *Non ci giurerei* = non ne sono sicuro □ **não apostaria nisso** | *Giurarla a qualcuno* = fare propositi di vendetta nei suoi confronti □ **giurar vingança contra alguém**.

garantire

v.tr.[*garantisco, garantisci*]. **1** Assicurare sotto la propria responsabilità il rispetto di un impegno proprio o altrui: *garantire la restituzione di un prestito* ◇ *Assicurare la qualità e il perfetto funzionamento di una merce, impegnandosi con l’acquirente a ripararla o a sostituirla gratuitamente qualora si guasti entro un certo periodo: la fabbrica garantisce questo televisore per un anno* **2** Dare per certo, assicurare: *ti garantisco che ha detto la verità; la buona pensione gli garantisce una vecchiaia tranquilla* □ **garantir** **Comissivo - vedi assicurare** ◇ **garantirsi** v.pr. Assicurarsi, prendere delle precauzioni contro eventuali danni o rischi □ **garantir-se**.

assicurare

v.tr. **1** Rendere sicuro **§ garantire**: *risparmia per assicurare l’avvenire alla famiglia.* **2** Dare per certo; affermare con sicurezza: *lo assicurarono che non sarebbe successo nulla; assicura di non saperne niente.* □ **assegurar** **Obs: ilocução da classe dos comissivos, que, ao ser enunciado, compromete o falante a determinada linha de ação.** **3** Fermare, fissare saldamente: *assicura le imposte prima di uscire* □ **fixar** **4** Proteggere un bene da eventuali rischi e danni facendo un’assicurazione: *assicurare l’auto contro il furto* □ **pôr no seguro** ◇ **assicurarsi** v.pr. **1** Accertarsi, controllare: *assicurati che le porte siano tutte chiuse* □ **assegurar-se** **2** Garantirsi da eventuali rischi e danni facendo un’assicurazione: *assicurarsi contro il furto* □ **fazer seguro**.

O problema dessa sugestão é a seguinte: *giurare, promettere* etc são remetidos a *assicurare*; mas e no caso do consulente pesquisar diretamente *assicurare*? Ele seria remetido ao quê? Caso não seja remetido a nada, tem-se uma deficiência, uma vez que, do mesmo modo que alguém poderia gostar de saber que *promettere* tem força ilocucionária semelhante a *assicurare*, poderia também gostar de saber o inverso. Logo, seria interessante que *assicurare* remeta a algo. De modo a sanar essa questão, veja-se o subcapítulo seguinte.

5.2.2. Remissivas entre verbetes e textos externos

Vislumbra-se aqui a possibilidade de manter as remissivas dentro dos verbetes, indicando a leitura de um material externo. Tomando a questão deixada sem resposta no

subcapítulo anterior, viu-se que *garantire* e *giurare* são remetidos a *assicurare*, sendo que este último, por um lado, esclarece parcialmente o que é um verbo ilocucionário comissivo, mas, por outro, não remete a nada. Logo, uma possibilidade é a de fazer com que remeta a um texto externo capaz de explicar o assunto em questão. Tal poderia se dar da seguinte forma:

garantire

v.tr.[*garantisco, garantisci*]. **1** Assicurare sotto la propria responsabilità il rispetto di un impegno proprio o altrui: *garantire la restituzione di un prestito* ◇ Assicurare la qualità e il perfetto funzionamento di una merce, impegnandosi con l'acquirente a ripararla o a sostituirla gratuitamente qualora si guasti entro un certo periodo: *la fabbrica garantisce questo televisore per un anno* **2** Dare per certo, assicurare: *ti garantisco che ha detto la verità; la buona pensione gli garantisce una vecchiaia tranquilla* □ **Comissivo - vedi assicurare** **garantir** ♦ **garantirsi** v.pr. Assicurarsi, prendere delle precauzioni contro eventuali danni o rischi □ **garantir-se**.

assicuràre

v.tr. **1** Rendere sicuro § *garantire: risparmia per assicurare l'avvenire alla famiglia.* **2** Dare per certo; affermare con sicurezza: *lo assicurarono che non sarebbe successo nulla; assicura di non saperne niente.* □ **assegurar** **Obs: ilocução da classe dos comissivos, que, ao ser enunciado, compromete o falante a determinada linha de ação; vedi r** **3** Fermare, fissare saldamente: *assicura le imposte prima di uscire* □ **fixar** **4** Proteggere un bene da eventuali rischi e danni facendo un'assicurazione: *assicurare l'auto contro il furto* □ **pôr no seguro** ♦ **assicurarsi** v.pr. **1** Accertarsi, controllare: *assicurati che le porte siano tutte chiuse* □ **assegurar-se** **2** Garantirsi da eventuali rischi e danni facendo un'assicurazione: *assicurarsi contro il furto* □ **fazer seguro**.

Aqui a seção auxiliar foi nomeada “r”, mas poderia se chamar *atti linguistici* etc. O importante é remeter o verbo-paradigma ao texto externo do dicionário capaz de esclarecê-lo.

5.3. Macroestrutura

Aqui é onde se verifica a maior liberdade para o lexicógrafo, uma vez que a inclusão de textos externos não altera a microestrutura padrão dos dicionários. Neste sentido, as possibilidades são infinitas, sendo que aqui são apontadas as mais evidentes:

(1) seção contendo a definição de cada uma das classes de ilocuições e a lista exaustiva de suas respectivas espécies;

(2) seção contendo a definição de cada uma das classes de ilocuições e algumas de suas respectivas espécies em conjunto com exemplos;

(3) seção contendo a definição de cada uma das classes de ilocuições e algumas de suas respectivas espécies em conjunto com exemplos, e com espécies e exemplos correspondentes em português-brasileiro;

(4) seção contendo a definição de cada uma das classes de ilocuções e algumas de suas respectivas espécies em conjunto com exemplos, e com espécies e exemplos correspondentes em português-brasileiro de modo a esclarecer nuances de uso existentes entre as duas línguas.

5.3.1. Definições de ilocuções e lista exaustiva de suas espécies

Um exemplo de organização de tal seção pode ser expresso tal como se segue, que para tal se vale de alguns trechos retirados de algumas fontes. Tal seção será aqui chamada de “*seção r*”:

Usos e regras pragmáticas¹⁴¹

Nas situações reais de comunicação, além das regras de fonética, morfologia e sintaxe, são aplicadas as **regras pragmáticas** da língua. As regras pragmáticas regem os **usos da língua**: estão fortemente ligadas ao contexto de comunicação e aos sentidos das ações, presentes nas escolhas linguísticas. As regras pragmáticas refletem a condição social dos interlocutores, os objetivos da comunicação, os efeitos que se procura atingir sobre os interlocutores e sobre a situação de comunicação. Apresentamos alguns usos linguísticos e as relativas regras pragmáticas da língua italiana.

Atos de Fala

A teoria dos atos de fala é aquela que permite a compreensão da linguagem como forma de ação¹⁴². De acordo com essa abordagem, a linguagem, além de ser um meio utilizado para expressar pensamentos, descrever o mundo e transmitir mensagens/ informações, é também usada para realizar ações¹⁴³ - sendo que estas se constituem pelo próprio dizer. A ação levada a cabo por um enunciado é também conhecido pelo nome de ilocução.

Existem cinco classes de ilocuções.

(i) **Comissiva** – abrange ilocuções que caracterizam promessas ou a assunção de algo; por outras palavras, ao serem enunciadas, comprometem o falante a determinada linha de ação¹⁴⁴. Veja-se a lista ilocuções comissivas: *assicurare, giurare, promettere, garantire, impegnarsi, ...n*

Primeiramente, informou-se o que são as regras pragmáticas e o que são atos de fala; em segundo lugar, definiu-se as cinco classes de atos de fala; em seguida, forneceu-se uma lista exaustiva de espécies de ilocuções relativas a cada uma das classes.

A desvantagem dessa sugestão é que ela não coloca o italiano (a língua-fonte) e o português-brasileiro (língua-alvo) em contato. Afora isso, para tal sugestão de pouca utilidade teria o *corpus*, já que bastaria tomar uma TAF qualquer e inserir a lista de espécies que tal teoria fornece, sem necessidade de pesquisá-los em *corpus*. Outra

¹⁴¹ Extraído da seção “Usos e regras pragmáticas” do PC (2012: 1010).

¹⁴² Extraído de Koch (1992: 19).

¹⁴³ Extraído de Sbisà (2009[1989]:45).

¹⁴⁴ Extraído de Austin (1990[1975]: 123 e 12).

questão: seria verdadeiramente possível apresentar uma lista completa de todas as ilocuções possíveis? Considera-se que não, da mesma forma que é impossível registrar todo o léxico de uma língua e fornecer informações acerca de todas as suas dimensões.

Neste sentido, apresenta-se outra possibilidade.

5.3.2. Definições de ilocuções e algumas espécies seguidas de exemplos

Neste caso, a introdução acerca das regras pragmáticas e a explicação feita dos atos de fala e suas cinco classes seriam idênticas à sugestão (1). O que mudaria é que, ao invés de uma lista exaustiva de espécies de ilocuções, seriam informadas apenas algumas espécies mais recorrentes, segundo resultado obtido pelo *corpus*. Uma possibilidade de apresentação seria a seguinte:

(...)

(i) **Comissiva** – abrange ilocuções que caracterizam promessas ou a assunção de algo; por outras palavras, ao serem enunciadas, comprometem o falante a determinada linha de ação. Vejam-se as ilocuções comissivas mais comuns:

Assicurare: Se si parla di economia, ti **assicuro** che non è facile tirare conclusioni.

Giurare: Io **giuro** che se mai sarò libero, ti sposerò.

Promettere: Gli **prometto**, appena tornata a Milano, di uscire anche a cena con lui.

Garantire: Io ti **garantisco** che ero a scuola.

Porém, ainda falta aqui o contato entre o italiano e o português-brasileiro. Isso tentará ser resolvido no próximo tópico.

5.3.3. Definições de ilocuções e algumas espécies do italiano e do português-brasileiro, seguidas de exemplos

Neste caso, a introdução acerca das regras pragmáticas, a explicação dos atos de fala e das cinco classes de ilocuções seriam idênticas à sugestão (1) e (2). Além disso, a inserção das espécies mais recorrentes e respectivos exemplos seriam idênticos à sugestão (2). A diferença estaria na inserção de espécies e exemplos correspondentes em português-brasileiro, também retirados do *corpus*. Uma vantagem, também, é que o lexicógrafo poderia indicar que não há equivalência pragmática entre os verbos italianos (com respectivas ilocuções) e seus equivalentes semânticos do português-brasileiro:

(...)

Vejam-se as ilocuições comissivas mais comuns do italiano e do português-brasileiro:

<i>Assicurare</i>	Se si parla di economia, ti assicuro che non è facile tirare conclusioni.
<i>Assegurar</i>	E eu lhe asseguro que o Exército não está dormindo.
<i>Giurare</i>	Io giuro che se mai sarò libero, ti sposerò.
<i>Jurar</i>	E te juro pela alma de minha mãe que eu caso com Maria.
<i>Garantire</i>	Io ti garantisco che ero a scuola.
<i>Garantir</i>	No entanto, garanto ao povo brasileiro que quem está no Governo tem compromisso.
<i>Promettere</i>	Gli prometto , appena tornata a Milano , di uscire anche a cena con lui .
<i>Prometer</i>	Deixe-me ir e prometo não incomodá-lo mais

Entretanto, essa sugestão não leva em conta as diferenças, p. ex., no número de ocorrências entre *assicurare* e *assegurar*, dando a impressão de que ambos são utilizados com a mesma frequência em suas respectivas línguas.

5.3.4. Definições de ilocuições e algumas espécies do italiano e do português-brasileiro, seguidas de exemplos e ordenadas por frequência

Neste caso, a introdução acerca das regras pragmáticas, a explicação feita dos atos de fala e das cinco classes de ilocuições seriam idênticas às sugestões (1), (2) e (3). Além disso, a inserção de algumas espécies e respectivos exemplos seria idêntica às sugestões (2) e (3), e a inserção de espécies e exemplos correspondentes em português-brasileiro seria idêntica à sugestão (3). A diferença residiria na explicitação de nuances, como por exemplo, a informação de que, apesar de *assicurare* e *assegurar* serem semanticamente equivalentes, pragmaticamente falando *assicurare* é um comissivo mais utilizado em italiano do que o *assegurar* em português-brasileiro – segundo resultado apresentado pelos *corpora*. Veja-se uma possibilidade de apresentação:

(...)

Vejam-se as ilocuições comissivas mais comuns do italiano e do português-brasileiro, segundo frequência de uso:

1	<i>Assicurare</i>	Se si parla di economia, ti assicuro che non è facile tirare conclusioni.
2	<i>Giurare</i>	Io giuro che se mai sarò libero, ti sposerò.
3	<i>Garantire</i>	Io ti garantisco che ero a scuola.
4	<i>Promettere</i>	Gli prometto , appena tornata a Milano , di uscire anche a cena con lui.
1	<i>Garantir</i>	No entanto, garanto ao povo brasileiro que quem está no Governo tem compromisso.
2	<i>Jurar</i>	E te juro pela alma de minha mãe que eu caso com Maria.

3	<i>Prometer</i>	Deixe-me ir e prometo não incomodá-lo mais
4	<i>Assegurar</i>	E eu lhe asseguro que o Exército não está dormindo.

De acordo com isso, a apresentação geral do texto externo presente na macroestrutura seria a seguinte:

Usos e regras pragmáticas

Nas situações reais de comunicação, além das regras de fonética, morfologia e sintaxe, são aplicadas as **regras pragmáticas** da língua. As regras pragmáticas regem os **usos da língua**: estão fortemente ligadas ao contexto de comunicação e aos sentidos das ações, presentes nas escolhas linguísticas. As regras pragmáticas refletem a condição social dos interlocutores, os objetivos da comunicação, os efeitos que se procura atingir sobre os interlocutores e sobre a situação de comunicação. Apresentamos alguns usos linguísticos e as relativas regras pragmáticas da língua italiana.

Atos de Fala

A teoria dos atos de fala é aquela que permite a compreensão da linguagem como forma de ação. De acordo com essa abordagem, a linguagem, além de ser um meio utilizado para expressar pensamentos, descrever o mundo e transmitir mensagens/ informações, é também usada para realizar ações - sendo que estas se constituem pelo próprio dizer. A ação levada a cabo por um enunciado é também conhecido pelo nome de ilocução.

Existem cinco classes de ilocuições.

(i) **Comissiva** – abrange ilocuições que caracterizam promessas ou a assunção de algo; por outras palavras, ao serem enunciadas, comprometem o falante a determinada linha de ação. Vejam-se as ilocuições comissivas mais comuns do italiano e do português-brasileiro, segundo frequência de uso:

1	<i>Assicurare</i>	Se si parla di economia, ti assicuro che non è facile tirare conclusioni.
2	<i>Giurare</i>	Io giuro che se mai sarò libero, ti sposerò.
3	<i>Garantire</i>	Io ti garantisco che ero a scuola.
4	<i>Promettere</i>	Gli prometto , appena tornata a Milano , di uscire anche a cena con lui.

1	<i>Garantir</i>	No entanto, garanto ao povo brasileiro que quem está no Governo tem compromisso.
2	<i>Jurar</i>	E te juro pela alma de minha mãe que eu caso com Maria.
3	<i>Prometer</i>	Deixe-me ir e prometo não incomodá-lo mais
4	<i>Assegurar</i>	E eu lhe asseguro que o Exército não está dormindo.

5.4. Integração entre as três estruturas

No subcapítulo 5.1.1 indicou-se preferência por inserir na microestrutura de um verbete marcas de uso relativas à classe de atos de fala. No tópico 5.1.2., mostrou-se a utilidade das notas de uso, de modo que ambas podem ser combinadas, considerando-se o que se propôs no subcapítulo 5.2, segundo o qual as marcas são úteis para remeter diversos verbos ilocucionários a um verbo ilocucionário tomado como paradigma, ao passo que esse conteria uma nota de uso e remeteria o consulente a explicações mais amplas contidas em textos externos.

Entretanto, nota-se certa ambiguidade entre aquilo que a seção *r* (texto externo, macroestrutura) informa e o que informa a microestrutura de *assicurare*, com notas de uso. Veja-se:

assicuràre

v.tr. **1** Rendere sicuro § garantire: *risparmia per assicurare l'avvenire alla famiglia*. **2** Dare per certo; affermare con sicurezza: *lo assicuraron che non sarebbe successo nulla; assicura di non saperne niente*.

□ **assegurar** **Obs:** ilocução da classe dos comissivos, que, ao ser enunciado, compromete o falante a determinada linha de ação; vedi r **3** Fermare, fissare saldamente: *assicura le imposte prima di uscire* □ **fixar** **4** Proteggere un bene da eventuali rischi e danni facendo un'assicurazione: *assicurare l'auto contro il furto* □ **pôr no seguro** ♦ **assicurarsi** v.pr. **1** Accertarsi, controllare: *assicurati che le porte siano tutte chiuse* □ **assegurar-se** **2** Garantirsi da eventuali rischi e danni facendo un'assicurazione: *assicurarsi contro il furto* □ **fazer seguro**.

(...)

Atos de Fala

A teoria dos atos de fala é aquela que permite a compreensão da linguagem como forma de ação. De acordo com essa abordagem, a linguagem, além de ser um meio utilizado para expressar pensamentos, descrever o mundo e transmitir mensagens/ informações, é também usada para realizar ações - sendo que estas se constituem pelo próprio dizer. A ação levada a cabo por um enunciado é também conhecido pelo nome de ilocução.

Existem cinco classes de ilocuições.

(i) Comissiva – abrange ilocuições que caracterizam promessas ou a assunção de algo; por outras palavras, ao serem enunciadas, comprometem o falante a determinada linha de ação. (...)

Ou seja, tanto o verbete quanto o texto externo explicam o que é uma ilocução comissiva. Afora isso, o texto externo amplia a explicação ao dizer o que é uma ilocução e como esta se insere dentro do que se entende por regras pragmáticas. Como se não bastasse, apontou-se anteriormente (5.1.2) que as notas de uso poderiam aumentar o tamanho do dicionário. E mais: se tal dicionário contém um texto externo tal como o apresentado, qual a razão de vários verbos serem remetidos a um verbo-paradigma ao invés de todos serem remetidos diretamente ao texto externo? E tal texto não poderia conter, também, os exemplos de uso, que foram rejeitados da microestrutura por não se apresentarem explicitamente como relacionados aos atos de fala?

Logo, a sugestão *ideal* deste trabalho é uma organização do seguinte tipo – primeiramente, os quatro verbetes e, em seguida, a seção externa *r*:

(1) Verbetes

giuràre

v.tr. **1** Dichiarare, promettere qualcosa con un giuramento: *giuro di dire la verità; giurami che non mi tradirai; giurare a Dio, sulla Bibbia, sul proprio onore*. □ **jurar** Comissivo - vedi r | *Giurare il falso* = mentire sotto (spec. in un processo) □ **prestar falso-testemunho** ♦ Affermare con certezza § assicurare: *ti giuro che non lo sapevo* □ **jurar** | *Non ci giurerei* = non ne sono sicuro □ **não apostaria nisso** | *Giurarla a qualcuno* = fare propositi di vendetta nei suoi confronti □ **jurar vingança contra alguém**.

garantìre

v.tr.[*garantisco, garantisci*]. **1** Assicurare sotto la propria responsabilità il rispetto di un impegno proprio o altrui: *garantire la restituzione di un prestito* ◇ Assicurare la qualità e il perfetto funzionamento di una merce, impegnandosi con l'acquirente a ripararla o a sostituirla gratuitamente qualora si guasti entro un certo periodo: *la fabbrica garantisce questo televisore per un anno* **2** Dare per certo, assicurare: *ti garantisco che ha detto la verità; la buona pensione gli garantisce una vecchiaia tranquilla* □ **garantir** **Comissivo - vedi r** ◇ **garantirsi** v.pr. Assicurarsi, prendere delle precauzioni contro eventuali danni o rischi □ **garantir-se**.

assicuràre

v.tr. **1** Rendere sicuro § garantire: *risparmia per assicurare l'avvenire alla famiglia*. **2** Dare per certo; affermare con sicurezza: *lo assicurano che non sarebbe successo nulla; assicura di non saperne niente*. □ **assegurar** **Comissivo - vedi r** **3** Fermare, fissare saldamente: *assicura le imposte prima di uscire* □ **fixar** **4** Proteggere un bene da eventuali rischi e danni facendo un'assicurazione: *assicurare l'auto contro il furto* □ **pôr no seguro** ◇ **assicurarsi** v.pr. **1** Accertarsi, controllare: *assicurati che le porte siano tutte chiuse* □ **assegurar-se** **2** Garantirsi da eventuali rischi e danni facendo un'assicurazione: *assicurarsi contro il furto* □ **fazer seguro**.

prométtere

v.tr.[conjugato come *mettere*]. **1** Impegnarsi a fare o a dare qualcosa o a comportarsi in un certo modo: *mi ha promesso di venire, di aiutarmi, di non disturbarmi più; promettere un regalo, un lavoro*. **2** ♣ Far sperare, far prevedere: *un cielo che promette bel tempo; il suo atteggiamento non promette niente di buono; un ragazzo che promette bene*. □ **Prometer** **Comissivo - vedi r**

(2) Texto externo¹⁴⁵

Seção r - Usos e regras pragmáticas

Nas situações reais de comunicação, além das regras de fonética, morfologia e sintaxe, são aplicadas as **regras pragmáticas** da língua. As regras pragmáticas regem os **usos da língua**: estão fortemente ligadas ao contexto de comunicação e aos sentidos das ações, presentes nas escolhas linguísticas. As regras pragmáticas refletem a condição social dos interlocutores, os objetivos da comunicação, os efeitos que se procura atingir sobre os interlocutores e sobre a situação de comunicação. Apresentamos alguns usos linguísticos e as relativas regras pragmáticas da língua italiana.

Atos de Fala

A teoria dos atos de fala é aquela que permite a compreensão da linguagem como forma de ação. De acordo com essa abordagem, a linguagem, além de ser um meio utilizado para expressar pensamentos, descrever o mundo e transmitir mensagens/ informações, é também usada para realizar ações - sendo que estas se constituem pelo próprio dizer. A ação levada a cabo por um enunciado é também conhecido pelo nome de ilocução.

Existem cinco classes de ilocuições.

(i) Comissiva – abrange ilocuições que caracterizam promessas ou a assunção de algo; por outras palavras, ao serem enunciadas, comprometem o falante a determinada linha de ação. Vejam-se as **ilocuções comissivas mais comuns** do italiano e do português-brasileiro, segundo frequência de uso:

1	<i>Assicurare</i>	Se si parla di economia, ti assicuro che non è facile tirare conclusioni.
2	<i>Giurare</i>	Io giuro che se mai sarò libero, ti sposerò.
3	<i>Garantire</i>	Io ti garantisco che ero a scuola.
4	<i>Promettere</i>	Gli prometto , appena tornata a Milano, di uscire anche a cena con lui.

1	<i>Garantir</i>	No entanto, garanto ao povo brasileiro que quem está no Governo tem compromisso.
2	<i>Jurar</i>	E te juro pela alma de minha mãe que eu caso com Maria.
3	<i>Prometer</i>	Deixe-me ir e prometo não incomodá-lo mais

¹⁴⁵ Ver subcapítulo 5.3.2, notas de rodapé 142-145 para as referências das quais retiraram-se os trechos que compõem esta seção sugerida.

4	Assicurare	E eu lhe asseguro que o Exército não está dormindo.
---	------------	--

Aos verbetes foi inserido o nome da classe de ilocuções à qual pertence a ilocução correspondente ao verbo em questão, seguido de uma remissiva que remete o leitor ao texto externo. Neste material externo, foram oferecidas informações a respeito da pragmática e dos atos de fala, bem como uma lista com as classes de ilocuções existentes e algumas de suas espécies, a saber, as mais recorrentes nos *corpora* do italiano e do português-brasileiro pesquisado. Além disso, exemplos de uso concreto de tais ilocuções de ambas as línguas foram fornecidas. Vale dizer que a disposição de tais ilocuções se deu também em função dos dados obtidos em *corpora*.

De modo a facilitar a visualização das três estruturas interligadas, veja-se a imagem abaixo:

Figura 7: integração das três estruturas.

Ou seja, supondo que um dado consultante do PC tenha procurado informações acerca do item lexical *assicurare*, ele então se depara, ao visualizar o verbete, com as informações “Comissivo – veja seção *r*”. O usuário pode então, em seguida, ir até a seção *r* (material externo) e lá obter informações acerca do que é a pragmática, o que são atos de fala, o que são verbos comissivos, quando utilizá-los, qual a frequência de uso na língua italiana em comparação com o português-brasileiro etc – além de lá encontrar diversos verbos pertencentes à mesma categoria ilocucionária.

CONSIDERAÇÕES FINAIS

A motivação para o presente trabalho foi a de sugerir um modo de inserção de ilocuções em dicionários híbridos italiano>português-brasileiro. Tal objetivo foi pautado pela consideração de que o (re)conhecimento de atos de fala de dada língua estudada por um falante de outra língua pode, em primeiro lugar, favorecer o seu aprendizado, e, em segundo lugar, favorecer a comunicação/ interação satisfatória entre tal indivíduo e um indivíduo pertencente à outra comunidade linguística. Como, porém, são poucos os estudos que se debruçaram sobre a dicionarização de ilocuções, acredita-se (e espera-se) que o presente estudo possa contribuir na medida em que incentive outros pesquisadores a seguir os caminhos aqui trilhados – seja para torná-los mais transitáveis do que aqui se conseguiu, seja para abrir um atalho/ bifurcação que leve os interessados a terrenos aqui não explorados.

Desde o início, pretendeu-se que as conclusões oferecidas pelo trabalho fossem logicamente decorrentes das premissas (fundamentação teórica) das quais se partiu. Esse é o motivo pela opção de restringir a investigação aos verbos ilocucionários e excluir uma série de outros expedientes linguísticos comumente usados para a realização de atos de fala.

Advérbios, conjunções etc poderiam ter sido investigados nos *corpora* selecionados. Mas os advérbios não são oferecidos pela TAF de Austin (é verdade que ele faz menção a advérbios no artigo *A plea for excuses*; contudo, não fazem parte de sua TAF, tal como expressa em *How To Do Things with Words*), pois sua teoria associa ilocuções a verbos. Sem dúvida que outra classificação de atos de fala poderia ter sido adotada – p. ex., uma classificação que não fizesse tal associação. O problema, contudo, é que a análise de *corpora* se tornaria difícil, uma vez que, não havendo um critério-padrão (sintaxe de busca) *convencionalmente* entendido como realizando determinado ato de fala, *todo o corpus* teria que ser analisado – só assim para saber se “*Sim!*” foi dito de modo a prometer (e não afirmar), se “*De acordo!*” foi dito para prometer (e não concordar) etc. Isso porque não se pode fazer com as ilocuções o que se pode fazer com as classes gramaticais, a saber, uma anotação prévia. Esse é o caso em que o pesquisador insere uma etiqueta tal como [v.] ao lado de verbos, [prep.] ao lado de preposições etc, tornando mais fácil obter o número de ocorrências de verbos, de preposições etc (muitos *softwares* oferecem tal possibilidade). Acontece que as

ilocuções, como se viu ao longo do trabalho, nem sempre se confundem com o expediente linguístico utilizado para realiza-las, e, logo, não podem ser categorizadas *a priori* do mesmo modo que as classes gramaticais podem. Logo, considerou-se relevante que, para uma pesquisa de ilocuções em *corpora*, sobretudo escritos, onde o contexto de proferimento é restrito (ou inexistente), que fosse estabelecido um critério capaz de permitir não apenas a padronização na busca, como também o fácil reconhecimento de ilocuções. O critério encontrado foi o teste de Austin do verbo na primeira pessoa do singular do presente do indicativo na voz ativa – o que permitiu, inclusive, o uso da anotação *verbo* nos *corpora* utilizados.

Por quê, então, trabalhar com *corpus* escrito? Novamente, pelo motivo de que o lexicógrafo trabalha com um volume de dados imenso, sendo necessário, assim, encontrar meios que facilitem o seu trabalho. Considerou-se que a LC realiza essa tarefa, ao permitir a observação simultânea de grande quantidade de dados linguísticos.

Evidentemente seria possível abrir mão da LC, momento em que o lexicógrafo poderia se valer, p. ex., de pesquisas pragmáticas específicas – uma que tenha estudado apenas proferimentos de dêiticos, outros de fórmulas de cortesia, outro de atos de fala expressos por meio de advérbios etc. O problema é que a legitimidade de dicionarizar informações com base em informações restritas poderia ser questionada. Além do mais, optou-se por adotar uma TAF já estabelecida porque uma investigação que tivesse dois objetivos complexos tais como estabelecer espécies de atos de fala, por um lado, e dicionarizá-las, por outro, demandaria mais tempo do que o tempo disponível para a conclusão de um mestrado.

Uma possibilidade de pesquisa que envolvesse atos de fala e que não se realizou aqui seria a de pesquisar ilocuções que, segundo Austin, figuram entre classes de ilocuções diversas – p. ex., uma que pertence tanto à classe dos exercitivos quanto a dos vereditivos ao mesmo tempo. Uma pesquisa de tal natureza poderia apresentar resultados interessantes, uma vez que, neste caso, o dicionário poderia conter modos de informar ao consulente tal ambiguidade pragmática. Aqui, contudo, tal pesquisa não foi levada a cabo pelo motivo de que, de todas as classes de ilocuções propostas por Austin, além das razões já expressas para a sua adoção, a classe dos comissivos é a única que permite certa desconsideração contextual, uma vez que prescinde de informações detalhadas acerca do falante – que, como se viu em várias passagens do texto, não podem ser informadas pelo *corpus*. P. ex., caso se investigasse em *corpora* o ato de fala de *condenar*, o único meio de saber se um proferimento do tipo “*Condeno*” realiza de

fato a ação de condenar seria saber se tal falante é um juiz. Como, porém, obter essa informação de um *corpus* escrito extenso? O mesmo no caso da ação de *ordenar*: um proferimento do tipo “*Ordeno que faça isso*” é de fato uma ordem? Depende de quem o proferiu – certamente seria uma ordem, caso o falante fosse uma mãe e o ouvinte o seu filho. Mas, de novo, como obter essa informação ao se trabalhar com *corpora* escritos? Tal informação poderia ser facilmente obtida em casos de *corpora* orais pouco extensos, mas em um caso assim pergunta-se: caso um lexicógrafo se valesse de *corpora* pequenos para escolher o que lematizar em seu dicionário, haveria legitimidade em tal escolha? Em outras palavras, um *corpus* pequeno é capaz de representar a língua geral? Acredita-se que não, motivo pelo qual se manteve a investigação de ilocuções em *corpora* extensos.

Além disso, o problema que ocorre com ilocuções exercitativas/ vereditivas não ocorre com as ilocuções comissivas. Isso porque não é necessário que uma promessa seja realizada apenas quando proferida por “falantes especiais” – juízes, mães etc. Qualquer um pode prometer dizendo “*Prometo que...*”, e, logo, a convenção que permeia os comissivos é mais geral do que a convenção que permeia outras classes de ilocuções, motivo pelo qual é possível prescindir do contexto nos primeiros casos (embora, claro, sua presença ajudaria na análise).

Acredita-se que a análise *corpora* foi útil, em primeiro lugar, para legitimar as escolhas daquilo a ser dicionarizado. As ilocuções *mi impegno* e *comprometo-me/ me comprometo* foram excluídas, pelo baixo número de ocorrências em comparação com as demais. Assim, se por um lado a exclusão de tais casos pode ser justificada, por outro a dicionarização dos demais pode ser legitimada.

No último capítulo, o da sugestão de dicionarização de ilocuções, tal sugestão se deu em função da divisão tradicional dos dicionários gerais de língua em macro, micro e medioestrutura. Isso não significa que essa sugestão seja a única possível, mas, considerando-se tal divisão tripartite, bem como os modos tais como em geral informações pragmáticas são dispostas em dicionários, tal sugestão se mostrou viável e plausível. Isso no caso de um dicionário híbrido; em um dicionário de italiano escolar, p. ex., a sugestão de notas de uso pode ser o suficiente. E no caso de dicionários eletrônicos, a sugestão feita pode ser ampliada, já que as funcionalidades são bem maiores do que as de um dicionário impresso. De certo modo, vale dizer, a função que em um dicionário impresso as remissivas desempenham é a mesma que a função desempenhada pelos *links* em dicionários eletrônicos. P. ex., após a marca de uso

“Comissivo” posto no verbete de *promettere*, um *hipertexto* poderia ser posto sobre tal marca, de modo a conduzir o usuário a um texto explicativo semelhante (ou mais bem elaborado) do que aquele que aqui se apresentou. Tal discussão, contudo, transcende os objetivos deste trabalho.

Uma questão pertinente a se fazer diante de todas tais considerações seria a seguinte: é possível generalizar a proposta aqui apresentada? Acredita-se que sim – o que não significa, entretanto, que isso pudesse ser feito sem ajustes. Deve-se ter em mente o que se disse no cap. 1: é equivocado pensar *no* dicionário, sendo adequado pensar *em um dado* dicionário. Caso se queira realizar concretamente a proposta aqui apresentada, nenhuma modificação é necessária – ainda que os comissivos investigados fossem outros. Caso se pretenda, porém, investigar outras classes de atos de fala (p. ex., exercitivos), então se deve pensar em meios de ampliar o contexto oferecido pelo *corpus* de modo a reconhecer quem é o falante. A proposta apresentada no anexo 2 deste trabalho pode ser um caminho. O mesmo se pode dizer de casos em que se pretenda investigar outros expedientes linguísticos realizadores de atos de fala, tais como advérbios, interjeições etc – com exceção do fato de que, agora, talvez a teoria de Austin seja útil apenas como arcabouço teórico, e não mais prático-operacionalmente (pois ela não forneceria uma sintaxe de busca padronizada). A abordagem *corpus-driven* poderia ser útil nesse caso, por meio da qual o pesquisador não iniciaria a busca por ilocuções já estabelecidas, mas verificaria, de acordo com o que diz o *corpus*, quais são os atos de fala mais frequentes. A desvantagem, como já apontada, seria a necessidade de investigar todo ou grande parte do *corpus*.

Por fim, vale ressaltar que a sugestão de dicionarização de ilocuções feita neste trabalho mostra-se de utilidade ao consulente de um dicionário híbrido italiano>português-brasileiro, uma vez que tais informações não são tão bem conhecidas pelo público em geral tais como verbos, conjunções, interjeições etc. É verdade que muitos dicionários atuais trazem informações ilocucionárias mediante emprego da expressão “usado para”, entretanto pouca explicação é dada acerca do significado de tal recurso. Afora isso, no contexto de verbos ilocucionários o “usado para” configuraria redundância, motivo pelo qual foi desconsiderado nesse trabalho (p. ex.: *Promettere* – traduzido por *prometer*, *usado para* prometer/ constituir uma promessa). Considera-se, então, que a mera ampliação do conhecimento público acerca da pragmática e dos aspectos (extra) linguísticos sobre os quais se debruça já justifica inserção de atos de fala em dicionários. Porém, a utilidade não para neste ponto. Ao serem inseridos atos de

fala, leva-se a público regras implícitas a todo ato comunicativo e/ ou interativo, favorecendo a comunicação/ interação entre os indivíduos, sociedades e culturas.

REFERÊNCIAS

- ALCOFORADO, Paulo. A teoria da linguagem ordinária. *Acta Semiotica et Lingvistica* – Revista Internacional de Semiótica e Linguística, v. 4, 1980, p. 9-30.
- ANDORNO, C. M. *Che cos'è la pragmatica linguística*. Roma: Carocci, 2009.
- ANDRADE, P. *Proposta de ensino intercultural crítico de PFOL: os atos de fala e a visão performativa da linguagem*. Dissertação (Mestrado). Orient.: Prof^a. Dr^a. Maria Helena da Nóbrega. Dep. de Letras Clássicas e Vernáculas da FFLCH/ USP. São Paulo, 2014.
- AUROUX, S. *A filosofia da linguagem*. Trad. J. H. Nunes. Campinas: Ed. Unicamp, 1995.
- AUSTIN, J. L. *How to Do Things with Words* [1962]. 2^a ed. Ed. J. O. Urmson e M. Sbisà. Cambridge: Harvard University Press, 1975.
- _____. *Quando dizer é fazer: palavras e ação* [How to Do Things with Words, 2^a ed., 1975]. Trad. D. Marcondes. Porto Alegre: Artes Médicas, 1990.
- _____. Performative Utterances. Unscripted Talk, BBC, 1956. In: _____. *Philosophical Papers*. 3^a ed. Ed. J. O. Urmson e G. J. Warnock. Oxford: Oxford University Press, 1979, p. 233-252.
- _____. A Plea for Excuses. Proceeding of the Aristotelian Society, v. 57, 1956-57. In: _____. *Philosophical Papers*. 3^a ed. Ed. J. O. Urmson e G. J. Warnock. Oxford: Oxford University Press, 1979, p. 175-204.
- BACHMAN, L. F. *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press, 1991.
- BALDWIN, T. *Contemporary Philosophy: Philosophy in English since 1945*. Oxford: Oxford University Press, 2005.
- BARBOSA, M. A. Dicionário, vocabulário, glossário: concepções. *Caderno de Terminologia*, n. 1., 2001, p. 23-45.
- BARROS, L. A. *Curso básico de terminologia*. São Paulo: Edusp, 2004.
- BAZZANELLA, C. *Linguistica e pragmatica del linguaggio: un'introduzione*. Roma-Bari: Laterza, 2009.
- BERBER-SARDINHA, T. Linguística de Corpus: histórico e problemática. *DELTA*, v. 16, n. 2, 2000, p. 323-367.
- _____. Tamanho de corpus. *the ESP*, São Paulo, v. 23, n. 2, 2003, p. 103-122.

_____. Linguística de Corpus e tradução: prosódia semântica. In: *Linguística de Corpus*. Barueri: Manole, 2004, p. 233-250.

BIANCHI, C. Significado como uso: uma interpretação [Significato come uso: um'interpretazione, 2003]. Trad. Renato Railo. *Fundamento – Revista de Pesquisa em Filosofia*, v. 1, n. 4, jan.-jun/2012, p. 163-184. Disponível em: <http://www.revistafundamento.ufop.br/index.php/fundamento/article/view/58/47>. Acesso em: 15.06.2015.

_____. *Pragmatica del linguaggio*. Roma-Bari: Laterza, 2008.

BIBER, D. A corpus-driven approach to formulaic language in English: Multi-words patterns in speech and writing. *International Journal of Corpus Linguistics*, 14:3 (2009), p. 275-311.

BIDERMAN, M. T. C. A ciência da lexicografia. *Alfa*, São Paulo, 28 (supl.): 1-26, 1984.

_____. As ciências do léxico. In: OLIVEIRA, A. M. P. P.; ISQUERDO, A. N. (orgs.). *As ciências do léxico: lexicologia, lexicografia, terminologia*. 2ª ed. Campo Grande: Ed. UFMS, 2001a, p. 13-22.

_____. Os dicionários na contemporaneidade: arquitetura, métodos e técnicas. In: OLIVEIRA, A. M. P. P.; ISQUERDO, A. N. (orgs.). *As ciências do léxico: lexicologia, lexicografia, terminologia*. 2ª ed. Campo Grande: Ed. UFMS, 2001b, p. 131-144.

BILETZKI, A. (ed.). *The Story of Analytic Philosophy: Plot and Heroes*. New York: Routledge, 1998.

BORBA, F. S. *Organização de dicionários: uma introdução à lexicografia*. São Paulo: Ed. Unesp, 2003.

BUGUEÑO MIRANDA, F. Da classificação de obras lexicográficas e seus problemas: proposta de uma taxonomia. *Alfa*, 58 (1): 215-231, 2004.

CAFFI, C. *Pragmatica: sei lezioni*. Roma: Carocci, 2011.

COELHO, C. M.; REZENDE, F. A. A utilização de corpus na prática lexicográfica. *Revista Eletrônica de Linguística*, v. 6, n. 1, 2012. Disponível em: <http://www.seer.ufu.br/index.php/dominiosdelinguagem>. Acesso em: 07.10.14.

CORPUS do Português. Disponível em: <http://www.corpusdoportugues.org/>. Acesso em: 01.10. 2014.

CORRIERE della Sera. Disponível em: <http://www.corriere.it/>. Último acesso: 14.10.2013.

DAVIES, M. Creating useful historical corpora: a comparison of CORDE, The Corpus del Español, and The Corpus do Português. *Arias-01*, set./ 2009, p. 139-168.

DE MAURO, T. *La fabbrica delle parole: il lessico e problemi di lessicologia*. Torino: UTET, 2005.

DUMMETT, M. *Origins of Analytical Philosophy*. Cambridge: Harvard University Press, 1996.

DURAN, M. S. *Dicionários bilíngues pedagógicos: análise, reflexões e propostas*. Dissertação (Mestrado). Orient. Profa. Dra. Claudia Maria Xatara. UNESP, 2004.

DURAN, M. S.; XATARA, C. M. Dicionários semibílingues: uma inovação? *Revista de Estudos da Linguagem*, Belo Horizonte, v. 13, n. 1, p. 45-57, jan./jun. 2005.

_____. Critérios para categorização de dicionários bilíngues. In: ISQUERDO, A. N.; ALVES, I. M. (orgs.). *As ciências do léxico: lexicologia, lexicografia, terminologia*, v. 3. Campo Grande: Ed. UFMS, 2007, p. 311-320.

FANN, K. T. (org.). *Symposium on J. L. Austin* [1969]. Oxon: Routledge & Kegan Paul, 2011.

FOLHA Online. Disponível em: <http://www.folha.uol.com.br/>. Último acesso: 14.10.2013.

GLOBO.com. Disponível em: <http://www.globo.com/>. Último acesso: 14.10.2013.

GLOCK, H.-J. *O que é filosofia analítica?* Trad. R. H. Pich. Porto Alegre: Penso, 2011.

GONÇALVES, L. B. Linguística de Corpus como instrumento de avaliação de tradução literária. *TradTerm*, n. 15, 2009, p. 79-100.

GTLEX – Grupo de Trabalho de Lexicologia, Lexicografia e Terminologia da Associação Nacional de Pós-Graduação e Pesquisa em Letras e Linguística. Disponível em: <http://www.letras.ufmg.br/gtlexNovo/>. Acesso em: 14.07.14.

GUERRA SALAS, L.; GÓMEZ SÁNCHEZ, E. Pragmática y lexicografía: análisis de las marcas pragmáticas en el Diccionario Salamanca de la lengua Española. *Actas del XVI Congreso Internacional de Ásele*, 2005, p. 353-362. Disponível em: http://cvc.cervantes.es/ensenanza/biblioteca_ele/asele/pdf/16/16_0351.pdf. Acesso em: 29.04.12.

GUTIERRES, A. Relevância da pragmática e da teoria dos atos de fala para o ensino de inglês como língua estrangeira: tratamento dado a greetings em materiais didáticos. *The ESPECIALIST*, b. 29, n. especial (187-209), 2008, p. 187-209.

GUTIÉRREZ CUADRADO, J. Introducción. In: *Diccionario Salamanca de la lengua española*. Ed. J. Gutiérrez. Madrid: Santillana-Univ. de Salamanca, 1996, p. v-xii).

GUTIERREZ GONZALEZ, Z. M. *Linguística de Corpus na análise do internetês*. Dissertação (Mestrado). Orient. Prof. Dr. Antonio Paulo Berber Sardinha. São Paulo: PUC-SP, 2007.

HACKER, P. M. S. *Wittgenstein's Place in Twentieth-Century Analytic Philosophy*. Oxford: Blackwell Publishers, 1997.

HARTMANN, R. R. K.; JAMES, G. *Dictionary of Lexicography*. London/ New York: Routledge, 1998.

HERRERO CECÍLIA, J. *Teorías de pragmática, de lingüística textual y de análisis del discurso*. Cuenca: Ediciones de la Univ. de Castilla-La Mancha, 2006.

HÖFLING, C.; SILVA, M. C. P. da; TOSQUI, P. O dicionário como material didático na aula de língua estrangeira. *Intercâmbio*, v. 13, 2004, p. 1-7.

HYMES, D. H. On Communicative Competence [1971]. In: PRIDE, J. B.; HOLMES, J. *Sociolinguistics*. New York: Penguin Books, 1979, p. 269-293.

IG. Disponível em: <http://www.ig.com.br/>. Último acesso: 14.10.2013.

INGLEZ, K. G. O fórum eletrônico no Orkut: uma análise discursiva do hipertexto. In: GIL, B. D. et al. (orgs.). *Modelos de análise linguística*. São Paulo: Contexto, 2009, p. 183-199.

KERBRAT-ORECCHIONI, C. *Os atos de linguagem no discurso: teoria e funcionamento*. Trad. F. A. de Almeida e I. E. Dias. Niterói: Ed. da UFF, 2005.

KOCH, I. V. *A interação pela linguagem*. São Paulo: Contexto, 1992.

KRIEGER, M. G. O dicionário de língua como potencial instrumento didático. In: ISQUERDO, A. N.; ALVES, I. M. (orgs.). *As ciências do léxico: lexicologia, lexicografia e terminologia*, v. 3. Campo Grande: Ed. UFMS, 2007, p. 295-310.

_____. Lexicologia, lexicografia e terminologia: impactos necessários. In: ISQUERDO, A. N.; FINATTO, M. J. B. (orgs.). *As ciências do léxico: lexicologia, lexicografia, terminologia*, v. 4. Campo Grande: Ed. UFMS, 2008, p. 161-176.

LA REPUBBLICA. Disponível em: <http://www.repubblica.it/>. Último acesso: 14.10.2013.

LECLERCQ, B. *Introduction à la philosophie analytique: La logique comme méthode*. Paris: De Boeck, 2008.

LEVINSON, S. C. *Pragmática*. Trad. L. C. Borges e A. Mari. São Paulo: Martins Fontes, 2007.

LIBERO. Disponível em: <http://www.libero.it/>. Último acesso: 14.10.2013.

LORENTE, M. A lexicologia como ponto de encontro entre a gramática e a semântica. Trad. C. D. Schreiber, J. S. Padilha e C. R. Bevilacqua. In: ISQUERDO, A. N.; KRIEGER, M. G. (orgs.). *As ciências do léxico: lexicologia, lexicografia e terminologia*, v. 2. Campo Grande: Ed. UFMS, 2007, p. 19-30.

LYDING, V. et al. The Paisà Corpus of Italian Web Texts. In: BILDHAUER, F; SCHÄFER, R. (eds.). *Proceedings of the 9th Web as Corpus Workshop (WaC-9)*. Gothenburg: ACL, 2014, p. 36–43.

MARCONDES FILHO, D. *Filosofia analítica*. Rio de Janeiro: Jorge Zahar, 2004.

_____. *A pragmática na filosofia contemporânea*. Rio de Janeiro: Zahar, 2005.

MARCUSCHI, A. C. Gêneros textuais emergentes no contexto da tecnologia digital. In: _____; XAVIER, A. C. (orgs.). 2ª ed. *Hipertexto e gêneros digitais: novas formas de construção de sentido*. São Paulo: Lucerna, 2005, 13-68.

MARTINICH, A. P.; SOSA, D. (eds). *A Companion to Analytic Philosophy*. Oxford: Blackwell Publishing, 2006.

MILKOV, N. *A Hundred Years of English Philosophy*. Dordrecht: Kluwer Academic Publishers, 2003.

MILLER, A. *Filosofia da linguagem*. Trad. E. L. Gomes. 2ª ed. São Paulo: Paulus, 2010.

MIRANDA, F. B. O que é macroestrutura no dicionário de língua? In: ISQUERDO, A. N.; ALVES, I. M. (orgs.). *As ciências do léxico: lexicologia, lexicografia e terminologia*, v. 3. Campo Grande: Ed. UFMS, 2007, p. 261-272.

OLIVEIRA, L. P. Linguística de Corpus: teoria, interface e aplicações. *Matraga*, v. 16, n. 24, jan./ jun. 2009, p. 48-76.

OLIVEIRA, M. A. *Reviravolta linguístico-pragmática na filosofia contemporânea*. 3ª ed. São Paulo: Loyola, 2006.

OTTONI, Paulo. *Visão performativa da linguagem*. Campinas: Ed. da Unicamp, 1998.

PAISÀ. Corpus Italiano. Disponível em: <http://www.corpusitaliano.it/>. Acesso em: 30.08.2014.

PAROLA CHIAVE. *Dizionario di italiano per brasiliani*. Trad. C. A. Dastoli et al. São Paulo: Martins Fontes, 2012.

PASSMORE, J. *A Hundred Years of Philosophy* [1957]. London: Penguin Books, 1994.

QECR. Quadro Europeu Comum de Referência para as Línguas: aprendizagem, ensino, avaliação. Porto: Edições ASA, 2001.

RAJAGOPALAN, K. O Austin do qual a linguística não tomou conhecimento e a linguística com a qual Austin sonhou. *Cadernos de Estudos Linguísticos*, v. 30, 1996, p. 105-116. In: _____. *Nova pragmática: fases e feições de um fazer*. São Paulo: Parábola Editorial, 2010a, p. 241-254.

_____. Searle falando de [Austin falando de] linguagem [Searle [on Austin] on Language]. *Language & Communication*, v. 20, n. 4, 2000, p. 347-391. Trad. R. Ferreira. In: _____. *Nova pragmática: fases e feições de um fazer*. São Paulo: Parábola Editorial, 2010b, p. 67-118.

_____. A nova pragmática: fases e feições de um fazer; relato de uma trajetória pessoal. In: _____. *Nova pragmática: fases e feições de um fazer*. São Paulo: Parábola Editorial, 2010c, p. 7-22.

_____. A irredutibilidade do ato ilocucionário como fator inibidor do êxito das tentativas taxonômicas. *DELTA*, v. 8, n. 1, 1992, p. 91-133. In: _____. *Nova pragmática: fases e feições de um fazer*. São Paulo: Parábola Editorial, 2010d, p. 45-66.

RICHARDS, J. D.; RODGERS, T. S. *Approaches and methods in language teaching*. New York: Cambridge University Press, 2001.

RIBEIRO, G. C. B. *Tradução técnica, terminologia e Linguística de Corpus: a ferramenta Wordsmith Tools*. UFSC, 2004, p. 159-174.

RORTY, R. (ed.). *The Linguistic Turn: essays in philosophical method* [1967]. 2ª ed. Chicago: The University of Chicago Press, 1992.

SALAMANCA. *Diccionario de la lengua española*. Ed. J. Gutiérrez. Madrid: Santillana-Univ. de Salamanca, 1996.

SANTIAGO, M. S. Análises contrastivas de microestruturas em dicionários escolares. *Pesquisas em Discurso Pedagógico*, 1, 2012. Disponível em: <http://www.maxwell.vrac.puc-rio.br/20000/20000.PDFXXvmi=9LgHg323nsmqvWbSKk89DQiiULLrzM0IzggPksSJWgatelmZGv9kmsDiepLNdLwM3xDnrw7oOiS4Ef3IMdG9GksUKeJZ0gM4fpsHxB6MT0QxgUU39BdPWGLIUuHVMt02UtJS5QGZRxx8ttcD5pt37fL2LsZxWZAStkMzcsz428dvCpaHQELxTw1G00ex0PQtbD4n7DsaLHHn8R60wRZro08p1mn2dfwoTBN2egc1ZUo7Nv7ko6MN7Ve8iDbvSIVN>. Acesso em: 10.01.2015.

SBISÀ, M. *Linguaggio, ragione, interazione*. Per una pragmatica degli atti linguistici [1989]. Trieste: EUT Edizioni Università di Trieste, 2009.

SCHMITZ, J. R. A problemática dos dicionários bilíngues. In: OLIVEIRA, A. M. P. P.; ISQUERDO, A. N. (orgs.). *As ciências do léxico: lexicologia, lexicografia, terminologia*. 2ª ed. Campo Grande: Ed. UFMS, 2001, p. 161-170.

SEARLE, J. R. Austin on Locutionary and Illocutionary Acts. *The Philosophical Review*, v. 77, n. 4, Out./ 1968, p. 405-424.

_____. *Os actos de fala: um ensaio de filosofia da linguagem* [Speech acts: an essay in the philosophy of language, 1969]. Coimbra: Almedina, 1981.

SEIDE, M. S. Lexicografia pedagógica no Brasil: avanços e desafios. *Resumo expandido de comunicação proferida no VIII ENGTLEX – Encontro Intermediário do*

Grupo de Trabalho de Lexicologia, Lexicografia e Terminologia da ANPOLL – Associação Nacional de Pós-Graduação e Pesquisa em Letras e Linguística. 2006. Disponível em: http://www.letras.ufmg.br/gtlex/viiiengtlex/pdf/resumos_expandidos/Marcia%20Sipavicius%20Seide.pdf. Acesso em: 01.11.2014.

SELISTRE, I. C. T. O uso de dicionário bilíngue escolar no ensino médio para a compreensão de textos em língua inglesa. *Signo*. Santa Cruz do Sul, v. 34 n. 57, p. 191-204, jul./dez., 2009.

SERRA, L. H. Uma visita linguística à casa de farinha e ao engenho: a produção de verbetes socioterminográficos. *Littera Online*, n. 4, 2011, p. 156-173. Disponível em: <http://www.periodicoeletronicos.ufma.br/index.php/littera/article/viewFile/765/480>. Acesso em: 02.09.2013.

SHEPHERD, T. M. G. O estatuto da Linguística de Corpus: metodologia ou área da Linguística? *Matraga*, v. 16, n. 24, jan./ jun. 2009, p. 150-172.

SILVA, M. C. Lexicografia bilíngue: uma verificação dos substantivos mais frequentes em dicionários bilíngues francês-português e português-francês. In: LONGO, B. N. O; SILVA, B. C. D. *A construção de dicionários e de bases de conhecimento lexical*. Araraquara: Laboratório Editorial FCL/ UNESP; São Paulo: Cultura Acadêmica, 2006, p. 13-44. (Série Trilhas Linguísticas, 9).

_____. Para uma tipologia geral de obras lexicográficas. In: ISQUERDO, A. N.; ALVES, I. M. (orgs.). *As ciências do léxico: lexicologia, lexicografia e terminologia*, v. 3. Campo Grande: Ed. UFMS, 2007, p. 283-292.

SINONIMI e contrari minore. Dizionario fraseologico delle parole equivalente analoghe e contrarie. 3ª ed. Bologna: Zanichelli, 2009.

STANFORD Encyclopaedia of Philosophy. Adolf Reinach. Disponível em: <http://plato.stanford.edu/entries/reinach/#ReiTheSocAct>. Acesso em: 01.08.2014.

STREHLER, R. G. Marcas de uso nos dicionários. In: OLIVEIRA, A. M. P. P.; ISQUERDO, A. N. (orgs.). *As ciências do léxico: lexicologia, lexicografia, terminologia*. 2ª ed. Campo Grande: Ed. UFMS, 2001, p. 171-180.

TAGNIN, S. E. O. Os Corpora: instrumentos de auto-ajuda para o tradutor. In: *Cadernos de Tradução IX*. Florianópolis, UFSC, 2002. Disponível em: <http://www.cadernos.ufsc.br/online/9/stella.htm>.

_____. *O jeito que a gente diz*. Barueri: DISAL, 2005.

_____. *O jeito que a gente diz: combinações consagradas em inglês e português*. Ed. rev. e ampl. Barueri: DISAL, 2013.

_____; TEIXEIRA, E. D. Linguística de Corpus e tradução técnica: relato da montagem de um corpus multivarietal de culinária. *TradTerm*, 10, 2004, p. 313-358.

TERRA. Disponível em: <http://www.terra.com.br/portal/>. Último acesso: 14.10.2013.

TISCALI. Disponível em: <http://www.tiscali.it/>. Último acesso: 14.10.2013.

UOL. Disponível em: <http://www.uol.com.br/>. Último acesso: 14.10.2013.

VAN DIJK, T. A. *Texto y contexto: semántica y pragmática del discurso*. Madrid: Cátedra, 1980.

VAN HOOFF, H. Os tradutores e os dicionários. In: DELISLE, J.; WOODSWORTH, J. (orgs.). *Os tradutores na história*. Trad. S. Bath. São Paulo: Ática, 1998, p. 241-253.

VAN STERKENBURG, P. 'The' dictionary: Definition and history. In: _____. (Org.). *A Practical Guide to Lexicography*. Amsterdam-Philadelphia: John Benjamins Publishing Company, 2003, p. 3-17. (Terminology and Lexicography Research and Practice, v. VI).

VIRGILIO. Disponível em: <http://www.virgilio.it/>. Último acesso: 14.10.2013.

WARNOCK, G. J. J. L. *Austin*. New York: Routledge, 1989. (The Arguments of the Philosophers).

WELKER, H. A. *Dicionários: uma pequena introdução à lexicografia*. 2ª ed. rev. e ampl. Brasília: Thesaurus, 2004.

XAVIER, A. C.; SANTOS, C. F. E-forum na internet: um gênero digital. In: ARAÚJO, J. C.; BIASI-RODRIGUES, B. (orgs.). *Interação na internet: novas formas de usar a linguagem*. Rio de Janeiro: Lucerna, 2005, p. 30-38.

ZUCCHI, A. M. T. *O dicionário nos estudos de línguas estrangeiras: os efeitos de seu uso na compreensão escrita em italiano*. Tese (Doutorado). Orient.: Profª. Drª. Maria Aparecida Barbosa. Dep. de Linguística da FFLCH/ USP. São Paulo, 2010.

_____. Dicionário monolíngue no ensino de língua estrangeira: uma experiência de uso. In: ISQUERDO, A.N.; FINATTO, M.J.B. (org.). *As ciências do léxico: lexicologia, lexicografia e terminologia*, v. 4. Campo Grande: Ed. UFMS, 2010b, p. 255-268.

ANEXOS

Anexo 1. Figuras correspondentes às análises de *corpora*

1-15/785

Commercio a Cavarzere ... Le	assicuro	che ci sono grani di
colpite dalla crisi e vi	assicuro	che per un supporto di
e i cresimati 2008 Vi	assicuro	che non câ Ã tesoro
mauro era terrorizzato . Vi	assicuro	che una esperienza così è
ed in ogni caso vi	assicuro	che di teloni di senegalesi
non essere offensiva paris ti	assicuro	che la gente di fianco
bandiere populiste . Te lo	assicuro	tu non sai cogliere il
nell' essere accertata . Ti	assicuro	che le imprese che investono
di corso , ma vi	assicuro	il risultato è garantito .
di corso , ma vi	assicuro	il risultato è garantito .
54 anni . Ma vi	assicuro	che Gigi Sabani non soffriva
conoscere la moto . Vi	assicuro	che la differenza tra la
quanto può costare ? Vi	assicuro	che supera quella cifra abbondantemente
la nuotata , ma vi	assicuro	che se cadete in acqua
è simile , ma vi	assicuro	che il suo allarme è

esporta tutti i 785 risultati in

Figura 8: nº de ocorrências totais da ilocução “assicuro” no *Corpus Paisà*.

1-15/318

Va

passo , abbozza ridendo ,	giuro	che mi ammazzo . Quattro
limiterò ad un terzo (	giuro	poi basta) . è
me è successo e vi	giuro	da allora non ho più
stati in Tunisia , lo	giuro	, e Ahmed Hafiene è
. Sarò sincera , lo	giuro	. " Be ! si
, poi , te lo	giuro	, quando alla fine lei
spaventa il loro (vi	giuro	! e magari poi hanno
Non conoscevo Mills , lo	giuro	sui miei cinque figli .
ma di sabbia . Vi	giuro	che non ho capito come
un film di Pieraccioni e	giuro	che non ne ho mai
Comina nel libro " Non	giuro	a Hitler . La testimonianza
grissini . I grissini ,	giuro	. Ho avuto anche il
faccio un uso proprio (	giuro	!) e le torte
e la maglietta , vi	giuro	, le ho cercate ma
ho parole ... io ti	giuro	rispetto il movimento cattolico ,

esporta tutti i 318 risultati in formato **K**

Figura 9: nº de ocorrências totais da ilocução “giuro” no *Corpus Paisà*.

1-15/158	Vai
come reagirebbe , e le garantisco che non è piacevole !	
vivo in montagna e ti garantisco che è talvolta difficile consegnare	
che a parità di trattamento garantisco la validità delle fonti .	
gruppo che , ve lo garantisco , porterà aiuti direttamente alla	
arabi . E , ti garantisco , militarmente sarebbe un osso	
? Saluti . Pietro Vi garantisco che anche in questo ultimo	
impressionante , che , vi garantisco irraggiungibile per le cantanti famose	
in giro nessuno , lo garantisco ; anzi sempre , pur	
è accaduto mi scuso e garantisco che faremo tutto il possibile	
sex de ferr ti garantisco che l 'n95 il fix	
ritenevo giusto sostenere Federico e garantisco di non essere una adolescente	
era abbandonato , io gli garantisco un' esistenza di amore e	
sento di abbandonare . Ti garantisco che , comunque , tengo	
al muccassassina c' ero e garantisco che le parole , come	
deturpati dalle loro buche (garantisco ...) sono stanca di	
esporta tutti i 158 risultati in	formato KWIC

Figura 10: n° de ocorrências totais da ilocução “garantisco” no *Corpus Paisà*.

1-15/114	Vai al
compito . Da parte mia prometto di impiegare in tale lavoro	
in faccia , ma mi prometto di restituir- glie- lo io	
meglio di me . E prometto niente di meno » .	
sono le stesse . Non prometto niente se non l' impegno	
jonsi , grazie . e prometto che acquisterò il libro al	
prometter- lo . Ma vi prometto che le cercheremo – attivamente	
alle staminali embrionali . Vi prometto anche che non intraprenderemo questo	
quasi fratelli e ai quali prometto , anche se l' ho	
promis ad abbuffar- me , prometto che seguirò le regole del	
dei contemporanei , io vi prometto di sacrificare alla vostra sete	
! D' ora in avanti prometto che starò più attento ai	
lo so , ma vi prometto che se vi interessa ci	
Farò qualcosa , te lo prometto !! ” . Mi	
Diciamo subito che non vi prometto " 7 chili in 7	
inizio campionato : " vi prometto un anno di B "	
esporta tutti i 114 risultati in	formato KWIC

Figura 11: n° de ocorrências totais da ilocução “prometto” no *Corpus Paisà*.

1-15/31 V

buona , e prego e	mi impegno	a riscattare quella cattiva ,
no e scriverò loro che	mi impegno	personalmente a non votar- li
genere . Con questo spirito	mi impegno	affinché la Rai svolga sempre
di lettura corretta . Personalmente	mi impegno	perché dalla Regione Lombardia venga
" . Sì , e	mi impegno	ogni giorno per riuscir- ci
metto in testa una cosa	mi impegno	fino in fondo a costo
vedere " questi filemtti "	mi impegno	in tutto ciò che ci
amministrazione , in tal modo	mi impegno	a garantire tutta la copertura
. Da parte mia ,	mi impegno	a non sbagliare mira ed
là delle mie valutazioni politiche	mi impegno	perché nel territorio si viva
me alla tua cellula e	mi impegno	ad inserire la proposta dell'
dal ministro Di Pietro ,	mi impegno	da un lato a proporre
dell' esame (" non	mi impegno	così conosco già l' esito
paragone venga rimosso , e	mi impegno	a segnalar- lo a chi
Segretaria generale Filtea-Cgil Aderisco e	mi impegno	a promuovere la più ampia

esporta tutti i 31 risultati in [formato K1](#)

Figura 12: nº de ocorrências totais da ilocução “mi impegno” no *Corpus Paisà*.

Query **juro**, V 5,055 (4.5 per million)

Page 1 of 253 [Go](#) [Next](#) | [Last](#)

doc#3	e , se o BC não interviesse , a taxa de juro mergulharia . Todo dia , o chefe do open
doc#3	acham a escravidão ao dinheiro e à taxa de juro uma coisa que o futuro mostrará ser \"
doc#3	juro . No mundo do dinheiro , a taxa de juro é a mãe de todas as taxas . No Brasil ,
doc#3	mesma de sempre. ___ Mas agora é sério , juro . Não tome nenhuma decisão antes da minha
doc#3	correção monetária mais meio por cento de juro ao mês . A diferença está na medição desta
doc#4	agora não tem pagar para norde só pagar só juro (...) porque eles espreitou precisar
doc#4	ajudarei . Diante do Mais Terrível Nome , eu juro ! A ânfora sacudiu-se toda quando o gênio
doc#4	Rocco , 1999 , pp . 40-41 . M2U8T4 1 Só não juro que o Saci existe porque não se deve ficar
doc#4	Bastante me protegeu ; Eu não pude estudá . Eu juro que Jesus deu No verdô de minha idade Sarvação
doc#4	caput fontis (Verg . En . 12 , 816) « juro adjacui , perf. de adjacèò . pela fonte
doc#4	. ad Br . 1 , 18 , 5) . 2) No plural : juro a um oraco fôr negativa , com ne ou ut
doc#4	jridicilis Jno juridicti , JURO JUVENLITER juste juro justitia justus jussi , Juvenla justa juvenlter
doc#4	para a vida civil. estrutura das taxas de juro e a for- Propõe-se , em consequência ,
doc#4	deposita todo dia 15 , se quiser . - O que é \ " juro \ " ? - Juro é um aumento que o seu dinheiro
doc#4	viu . \ ' . . . Manuela . . . Eu não per juro ! Juro Pela luz dos olhos teus . . . Morrer
doc#4	ajudarei . Diante do Mais Terrível Nome , eu juro ! - A ânfora sacudiu-se toda quando o gênio
doc#4	ter confessado . E que vergonha . So não juro que o Saci existe porque não se deve ficar
doc#4	agora não tem pagar para norde só pagar só juro (...) porque eles espreitou precisar
doc#4	poder da poesia para que aconteça , como juro que acontece , que esta pergunta simples
doc#4	Azeitona ! -- replicou Andréia chorando . -- Eu juro ! -- Calma , Andréia ... Podem-nos notar

Page 1 of 253 [Go](#) [Next](#) | [Last](#)

Figura 13: nº de ocorrências totais da ilocução “juro” no *Corpus do Português*.

Query garanto , V 505 (0.4 per million)	
Page 1 of 26	Go Next Last
doc#1	?) Se nao der resultado , tente outra - garanto que no final do dia irá encontrar uma excelente
doc#2	quero sofrer ... sofrer ... Uma coisa eu garanto ... o Garotinho não ganha... o Chefe aqui
doc#3	ano todo . \ " Neste mês faço 35 shows e garanto 500 000 reais \ " , contabiliza , feliz
doc#3	que não posso saber o nome de todas . Só garanto que são todas legais \ " . Não é ilegal
doc#3	em você . Cincoenta e três votos eu lhe garanto . Pouco a pouco , impiedosamente , meu
doc#4	professores da educação infantil , nas creches , garanto que elas não sabem fazer essa estimulação
doc#4	conta dessas diferenças . Nessa rotina , garanto sempre uma atividade de leitura para as
doc#4	privilegiar o aspecto discursivo do texto , mas garanto em todas elas momentos para discutir também
doc#4	de história , gibis . Na minha rotina , garanto algumas atividades comuns a todos , como
doc#4	aquelas que envolveram mais os alunos e garanto o trabalho com textos , que são os que
doc#4	seu moço ! Onde já se viu isso ? Pois garanto que dá uma sopa pra lá de boa . Demora
doc#4	nas matas tem Saci , lá isso tem . E eu garanto essa verdade que até parece mentira , garanto
doc#4	garanto essa verdade que até parece mentira , garanto , porque já vi esse meio-gente e meio-bicho
doc#4	que és tão cruel ? Tem piedade , pai ; eu garanto por ele . Este é o segundo homem que vejo
doc#4	as empresas fizeram opção pelo Sistema de garanto educação às crianças de 0 a 6 anos em creche
doc#4	as empresas fizeram opção pelo Sistema de garanto educação às crianças de 0 a 6 anos em creche
doc#4	senhor não conversa com o seu filho ? Eu garanto ao senhor que o que ele estudou ele sabe
doc#4	situação . E se eu ganhar para vereador eu garanto a vocês que ajudarei os índios . Pode ter
doc#4	floresta meu amigo tivesse pé pra andar eu garanto meu amigo , com o perigo não tinha ficado
doc#4	seu moço ! Onde já se viu isso ? - Pois garanto que dá uma sopa pra lá de boa . - Demora
Page 1 of 26	Go Next Last

Figura 14: nº de ocorrências totais da ilocução “garanto” no *Corpus do Português*.

user: Mr. Renato Ribeiro corpus: brasileiro	
Concordance Word List Corpus Info ?	Query prometo , V 404 (0.4 per million)
	Page 1 of 21 Go Next Last
	doc#4 Acresce a necessidade de complementação do prometo institucional-peda gógico da Congregação
	doc#4 ou no fracasso escolar dos alunos ; com - prometo a comunidade local , tornando-a responsável
	doc#4 26/87 , o Relator vo ta pela aprovação do prometo do curso de Administração com habilitação
	doc#4 grande rei gritou o ratinho Deixe-me ir e prometo não incomodá-lo mais . Talvez um dia eu
	doc#4 gaguejo : Q , Q , Q , Q , Q ! Amiga Q , eu prometo que toda vez que você precisar sair para
	doc#4 lambreta , quando o fiscal propôs : Eu prometo à senhora que deixo a senhora passar ,
	doc#4 bola . -- Claro ! Se me trouxer a bola , prometo tudo isso ! -- respondeu prontamente a
	doc#4 matérias saíram da pauta completamente e prometo reescreve-las o mais breve possível . Jornalismo
Save View options KWIC Sentence Sort Left Right Node Shuffle Sample Filter Overlaps 1st hit in doc Frequency Node tags Node forms Text Types Collocations	doc#4 bola . -- Claro ! Se me trouxer a bola , prometo tudo isso ! -- respondeu prontamente a
	doc#4 divulgar , mas as farei noutra entrevista , que prometo para breve . INFORMAÇÃO DO ESTRANGEIRO
	doc#4 fundamentais para se independência de Portugal. prometo casar de novo com ela \ " . compreender
	doc#4 lambreta , quando o fiscal propôs : - Eu prometo à senhora que deixo a senhora passar ,
	doc#4 grande rei - gritou o ratinho Deixe-me ir e prometo incomodá-lo mais . Talvez um dia eu possa
	doc#4 Constituição espúria seria inimaginável , mas prometo que votarei a favor de sua iniciativa ,
	doc#4 ingressar no corpo de alunas desta escola , prometo : compreender , respeitar e honrar a bela
	doc#4 precisamos nos ver por aí . Pra semana , prometo , talvez nos vejamos , quem sabe ? Quanto
	doc#6 nas mãos e tentar alcançar o título . Não prometo fazer nenhum milagre tático para me promover
	doc#6 e pintura . Vou esperar a sua carta e eu prometo respondê-la . Saudações da Argentina !
	doc#6 aquilo a gente faz o que pode . O que eu prometo é que eu vou fazer , vou entregar , não
	doc#6 que jogou machucado anteontem \ " \ " Eu não prometo repetir isso . Mas gostaria de ver todos

Figura 15: nº de ocorrências totais da ilocução “prometo” no *Corpus do Português*.

Query **asseguro**, V 157 (0.1 per million)

Page 1 of 8 [Go](#) [Next](#) | [Last](#)

doc#4 estruturada com base em departamentos , **asseguro** a participação da comunidade acadêmica
 doc#4 plantação , (C) contar uma piada . Eu te **asseguro** não chore não , viu ? (D) contar uma
 doc#4 meu agradecimento pela sua colaboração e **asseguro** que as informações do questionário serão
 doc#4 se compreenderás minha intenção , mas te **asseguro** que vou me sentir muito melhor depois de
 doc#4 questões sejam respondidas . Para tanto , **asseguro** a você que as informações contidas neste
 doc#4 de bom êxito , então perante todos vós eu **asseguro** que me esgotarei para não decepcionar aos
 doc#4 Agri- l u s t r a m o s . E u d e m i m v o s **asseguro** colas , em nosso meio , d u r a n t e o s
 doc#4 de bom êxito , então perante todos vós eu **asseguro** que me esgotarei para não decepcionar aos
 doc#4 posso dizer que prejudicou o trabalho , mas **asseguro** que a discrimi- nação dos TAEs da DEMEC
 doc#4 expresso em suas palavras finais : \ " Eu vos **asseguro** , se- nhor diretor e senhores professores
 doc#6 . Disse que tinha sabido pela imprensa e **asseguro** que ele ficou contrariado . Folha - Logo
 doc#6 Meus atletas ainda são inexperientes . Mas **asseguro** ao povo chinês que três medalhas de ouro
 doc#6 o primeiro Campeonato Mundial . Aliás , **asseguro** que sou um péssimo torcedor até mesmo do
 doc#6 Aos empresários com mandados de prisão , **asseguro** : não vai acontecer quase nada . Fiquem
 doc#6 ter à disposição as terapias gênicas e lhe **asseguro** que eles acorrerão em massa a essas clínicas
 doc#6 inexistentes . Se isso é viver muitas vidas , **asseguro** que não é um privilégio dos atores . A
 doc#6 Com condições para voltar ao Paraguai , **asseguro** que ganho as eleições contra todos juntos
 doc#6 não vão subir os preços ? _ Não . Eu lhe **asseguro** que , um ano depois de ter sido feita a
 doc#6 acham que há situações exageradas , mas eu **asseguro** que tudo é verdade . Obra é \ " Beleza Americana
 doc#6 com o presidente da República , mas eu lhe **asseguro** que isso não afeta em nada o julgamento

Page 1 of 8 [Go](#) [Next](#) | [Last](#)

Figura 16: nº de ocorrências totais da ilocução “asseguro” no *Corpus do Português*.

Query **comprometo**, V 71 (0.1 per million)

Page 1 of 4 [Go](#) [Next](#) | [Last](#)

doc#4 nenhuma palavra para lhe dar . Mas eu me **comprometo** com V. Ex. a em pessoa , com a opção desta
 doc#4 Compromisso do Solicitante Pelo presente me **comprometo** a : 1. Dedicar todo meu tempo ao programa
 doc#4 Compromisso do Solicitante Pelo presente me **comprometo** a : 1) Apresentar um atestado médico quando
 doc#6 dos outros não estou prometendo . Só me **comprometo** com aquilo que depende exclusivamente de
 doc#6 federal Ceci Cunha (PSDB-AL) . \ " Eu me **comprometo** a encaminhar qualquer providência pedida
 doc#6 Prestígio . FRASES 01/07/94 FRASES \ " Eu **comprometo** a minha credibilidade de que essas metas
 doc#6 científica \ " . Assumem as máximas : não me **comprometo** , não te comprometes . Não me comprometa
 doc#6 Exemplo : compro uma área de terra e me **comprometo** a pagar o preço ajustado em tantos meses
 doc#6 cesta de tarifas com base nisso . E eu me **comprometo** : vai haver tarifas promocionais mesmo
 doc#6 , e anunciou a mudança . \ " \ " Mas eu me **comprometo** aqui , na frente de todos vocês , a escalar
 doc#6 quiser olhar minhas declarações , eu me **comprometo** a trazê-las . Eu não me sinto usuário do
 doc#6 apóia o vencedor da prévia ? Requião - Eu me **comprometo** a aceitar o resultado . Não é possível
 doc#6 novembro de 1971 , ela escreveu : \ " Eu me **comprometo** , de livre e espontânea vontade , a colaborar
 doc#6 ao Senado . Essa é a razão pela qual me **comprometo** a fazer um acordo razoável . Eu ainda acredito
 doc#6 nunca pode ser quebrado . Para tanto me **comprometo** a : - Respeitar e honrar o ideal de pureza
 doc#6 pretende fazer a respeito disso ? Almunia - Me **comprometo** a submeter esse projeto ao Parlamento para
 doc#6 e inovadoras . Assim que ocorrerem , me **comprometo** em registrá-la para os leitores da Folha
 doc#6 produção e investimentos ? Maluf - Eu me **comprometo** a receber a direção da Fiesp para discutirmos
 doc#6 indicando que você pagou pelo jornal . Eu me **comprometo** , por meio desta , a lhe devolver o recibo
 doc#6 fragmentária que sou e precários os momentos _ só me **comprometo** com vida que nasça com o tempo e com ele

Page 1 of 4 [Go](#) [Next](#) | [Last](#)

Figura 17: nº de ocorrências totais da ilocução “comprometo” no *Corpus do Português*.

Anexo 2. Proposta de construção de corpora apresentada no relatório de qualificação

Neste anexo, ofereceram-se as discussões e procedimentos adotados na confecção de *corpora* a partir de comentários de internautas a matérias jornalísticas publicadas online, tal como contido no relatório de qualificação – com exceção de alguns poucos trechos ou aperfeiçoados a partir de apontamentos feitos pela banca, ou suprimidos por fugirem aos atuais propósitos.

*

Confecção dos corpora

Hipertexto como *corpus*: comentários de internautas

De modo a se obter *corpora* a partir dos quais se possa sugerir possibilidades de estabelecimento de espécies de força ilocucionária do português-brasileiro e do italiano, optou-se pela coleta de *comentários* feitos por leitores (internautas) de matérias jornalísticas publicadas em mídias brasileiras e italianas *online*. Sobre os gêneros textuais eletrônicos, estes são tidos como “altamente interativo[s]” (MARCUSCHI, 2005: 33), tanto que, em muitos casos, “a escrita empregada reproduz estratégias da língua falada” (INGLEZ, 2009: 186), devido talvez à sua tendência à “informalidade” (XAVIER; SANTOS: 2005, 37) – ainda que, como diálogos escritos, sejam às vezes fragmentários e desrespeitem as características rituais dos intercâmbios orais (HERRERO CECÍLIA, 2006: 143) (embora se possa dizer que justamente por essa fragmentação assemelham-se à oralidade, uma vez que esta última não é tão linear quanto se exige da língua escrita). Neste sentido, considera-se possível apontar nos gêneros textuais eletrônicos (mais propriamente, nos *comentários*) algumas características da língua oral usada para a comunicação ordinária, tais como interatividade, informalidade, espontaneidade e flexibilidade.

Procedimentos adotados para coleta e confecção dos *corpora*

Antes de explicitar e justificar os critérios adotados para coleta e confecção dos *corpora* do português-brasileiro e do italiano, serão mostrados passo a passo todos os procedimentos realizados para a consecução dessas tarefas, de modo a facilitar a compreensão.

Em primeiro lugar, pensou-se em coletar comentários de sites jornalísticos e/ ou de sites que contenham notícias jornalísticas mais acessados do Brasil e da Itália. Para tal, baseou-se na pesquisa realizada pela *Alexa Internet Inc.* (<http://www.alexa.com/>), que mede os sites mais acessados da internet por país. No dia 07 de janeiro de 2013, acessou-se o site *Alexa*, clicou-se em *Top sites*, em seguida no link *By country* e em seguida escolheu-se *Brazil*. Após a busca, o *Alexa* trouxe os 500 sites mais visitados deste país, sendo que dentre os primeiros 25 sites, pode-se destacar os seis sites jornalísticos mais acessados que oferecem opção de comentários de leitores (internautas), como se pode ver na tabela 2:

Nome	Endereço	Posição - BR	Posição - Mundo
UOL	uol.com.br	5	99
Globo.com	globo.com	7	111
Yahoo.br	yahoo.com	9	4
Terra	terra.com.br	13	340
Ig	ig.com.br	14	380
Folha Online	folha.uol.com.br	19	606

Tabela 22: os seis sites que fornecem notícias jornalísticas e opção para comentários mais acessados do Brasil. Fonte: <http://www.alexa.com/topsites/countries/BR>. Acesso em: 07.01.13.

No caso da Itália, o resultado foi o seguinte, segundo tabela 3:

Nome	Endereço	Posição - IT	Posição - Mundo
Libero	Libero.it	6	256
Yahoo.it	Yahoo.com	4	4
La Repubblica	repubblica.it	12	483
Virgilio	virgilio.it	14	555
Il Corriere della Sera	corriere.it	17	642
Tiscali	www.tiscali.it	28	1377

Tabela 23: os seis sites que fornecem notícias jornalísticas e opção para comentários mais acessados da Itália. Fonte: <http://www.alexa.com/topsites/countries/IT>. Acesso em: 07.01.13.

Feito isso, o passo seguinte foi, evidentemente, acessar todos os sites de todas essas doze mídias, de modo a realizar a coleta. Por exemplo, acessou-se, como se vê na figura 3, a *homepage* da mídia Folha Online, uma das quais foram coletados comentários:

Figura 18: página principal da mídia Folha Online. Disponível em : <http://www.folha.uol.com.br/>. Acesso: 02.09.2013.

Note-se que nesta página inicial – tal como nas páginas iniciais das outras mídias – são disponibilizadas categorias correspondentes a assuntos determinados. Assim, considerando o interesse por coletar comentários referentes a política, o procedimento utilizado foi o de acessar a categoria ‘Política’ disposta na página inicial de cada uma das mídias; em seguida, considerando o interesse por coletar comentários referentes a esportes, o procedimento para tal foi o de acessar a categoria ‘Esporte’, e assim sucessivamente. De modo geral, o procedimento adotado foi o mesmo: acesso a cada uma das categorias correspondentes aos assuntos determinados de cada uma das mídias escolhidas, como forma de se poder escolher as matérias das quais se coletar comentários.

Figura 19: página principal da mídia Folha Online contendo categorias de assuntos, dentre as quais ‘Política’. Disponível em : <http://www.folha.uol.com.br/>. Acesso: 02.09.2013.

Assim, ao clicar na categoria 'Política', exposta na figura 4, abriu-se a seguinte página, contendo matérias relacionadas a este assunto, tal como indicada pela figura 5:

Figura 20: página principal da categoria 'Política' da mídia Folha Online, com matérias passíveis de serem coletadas. Disponível em: <http://www1.folha.uol.com.br/poder/>. Acesso: 02.09.2013.

Considerando-se a página principal da categoria 'Política' da mídia Folha Online, a coleta das matérias de política se deu a partir do acesso a cada uma delas ao se clicar sobre a manchete de cada uma, de modo a adentrar na página de conteúdo da notícia:

Figura 21: matéria contida na categoria 'Política' da mídia Folha Online. Disponível em: <http://www1.folha.uol.com.br/poder/2013/09/1335461-dilma-aumenta-a-seguranca-no-7-de-setembro-por-temer-violencia.shtml/>. Acesso: 02.09.2013.

Para a coleta da matéria em questão foram necessários os seguintes procedimentos: (...) Ctrl+C (...) Ctrl+V (...) para colá-la em um arquivo do software ‘Bloco de Notas’. A passagem do site para o ‘Bloco de Notas’ pode ser visualizada na figura abaixo:

Figura 22: cópia e colagem de uma matéria contida na categoria ‘Política’ da mídia Folha Online para o software ‘Bloco de Notas’.

Nota-se que no arquivo do ‘Bloco de Notas’, acima da matéria salva consta um asterisco (‘*’) e um número (‘1’), que servem para indicar o número da matéria coletada e facilitar uma eventual busca pela mesma, caso necessário durante a análise dos comentários. No caso, note-se que a matéria em questão exposta na figura 7 é a matéria de nº 1. Assim, esse procedimento será realizado com todas as matérias de todas as mídias, sendo que todas as matérias serão armazenadas num único arquivo do ‘Bloco de Notas’, separadas e individualizadas pelos sinais acima indicados.

O próximo passo foi a coleta de comentários. Tomando-se como exemplo a matéria exposta nas figuras 6 e 7, abaixo desta constavam comentários de leitores (internautas) – ou seja, comentários dispostos na própria página da matéria correspondente, tal como indicados pela figura 8:

www1.folha.uol.com.br/poder/2013/09/1335461-dilma-aumenta-a-seguranca-no-7-de-setembro-por-temer-violencia.shtml

Apto de 181m² 4 dorms (2 suítes) vista panorâmica! Conheça mais
www.Gafisa.com.br/Scena_Santana

Anuncie aqui

Comentar esta reportagem **Ver todos os comentários (34)**

fora corruPTos (1662) (08h21) há 2 horas

f t g+ Responder 1 16

Fora P T esa quadrilha de corruPTos incomPTentes !!!!

O comentário não representa a opinião do jornal; a responsabilidade é do autor da mensagem

Comentários da matéria 1

M.Mig (3438) (07h49) há 2 horas

f t g+ Responder 1 14

Os protestos de meses atrás começaram promovidos pelo MPL por causa de 20 centavos e rapidamente saíram do controle do MPL e passaram a exigir mudanças do governo federal. As manifestações pararam e promessas federais foram esquecidas. A dilmagda não está com medo de violência, ela está com medo é das dimensões dessa onda de manifestações.Pelo jeito as manifestações pelo impeachment de Collor (que foi instigada pelo petê, entre outros) deverá ficar chinelo...

O comentário não representa a opinião do jornal; a responsabilidade é do autor da mensagem

Luiz Sampaio (5077) (08h07) há 2 horas

f t g+ Responder 0 8

As manifestações tem que acontecer nas PROXIMIDADES DAS ELEIÇÕES DE 2014, desta forma até parece que esses movimentos estão NEGOCIANDO A PAZ E A CONVIVÊNCIA ATÉ AS PRÓXIMAS ELEIÇÕES. As manifestações pacíficas durante uma campanha eleitoral pode ser de grande valor o debate e UMA DISCUSSÃO COM PROPOSTAS FACTÍVEIS, tanto por parte dos interesses com da sociedade como dos políticos que pretendem representar a sociedade.

Figura 23: comentários de uma matéria contida na categoria 'Política' da mídia Folha Online. Disponível em: <http://www1.folha.uol.com.br/poder/2013/09/1335461-dilma-aumenta-a-seguranca-no-7-de-setembro-por-temer-violencia.shtml/>. Acesso: 02.09.2013.

A coleta dos comentários se deu a partir do mesmo procedimento usado para coletar matérias:

Figura 24: cópia e colagem de comentários de uma matéria contida na categoria 'Política' da mídia Folha Online para o software 'Bloco de Notas'.

Contudo, pode-se dizer que este arquivo do 'Bloco de Notas' está "sujo", uma vez que contém dados que não interessam à pesquisa – p. ex., dados oferecidos pela mídia em questão em termos de redes sociais, o número de comentários que cada usuário fez desde que se cadastrou na mídia etc. Tais dados foram, então, apagados:

Figura 25: dados que não interessam à pesquisa e que foram apagados.

Neste sentido, os únicos dados que compõem os *corpora* são: o comentário dos leitores (internautas) e o nome do leitor (internauta).

Figura 26: dados que interessam à pesquisa e que serão mantidos.

Note-se que acima e abaixo dos comentários salvos no arquivo do ‘Bloco de Notas’ encontram-se asteriscos (*) ao lado de números (1, 2 etc). Tais sinais servem para separar os comentários correspondentes a cada matéria, de modo que possam ser facilmente reconhecidos em relação a essas, caso necessário. Assim, os comentários abaixo de “*1” referem-se aos comentários da matéria 1, os abaixo de “*2” referem-se aos comentários da matéria 2, e assim sucessivamente. Importante dizer que todos os comentários serão salvos num único arquivo do ‘Bloco de Notas’.

Assim, dois arquivos do software ‘Bloco de Notas’ foram criados: um contendo as matérias; outro contendo os comentários das matérias.

Figura 27: arquivos do ‘Bloco de Notas’, um contendo as matérias e outro os comentários das matérias.

Crterios adotados para coleta e confeco dos *corpora*

Os *corpora* italiano e portugus-brasileiro foram compostos durante o ano de 2013, sobretudo de fevereiro a maro e junho deste ano, a partir da coleta de comentrios de leitores (internautas) de matrias jornalsticas publicadas em mdias brasileiras e italianas *online* deste perodo¹⁴⁶. Vale dizer que o critrio para escolha de comentrios recentes se deu pelo fato de poder, com eles, analisar o italiano e o portugus-brasileiro contemporneos, j que o objetivo geral a produo de dicionrios bilngues e/ ou hbridos – logo, instrumentos que devem ser compostos por dados e anlises recentes e atuais de lnguas naturais.

O primeiro critrio foi o da escolha das mdias das quais coletar comentrios. Tendo-se por norte dicionrios bilngues e/ ou hbridos italiano/ portugus-brasileiro, objetivou-se confeccionar os *corpora* de modo a preterir ao mximo variantes restritas da lngua em uso prprias a um ou outro grupo de falantes contidos no interior de cada uma das comunidades de falantes de tais lnguas naturais, uma vez que, se assim no fosse, os resultados da anlise destes *corpora* (os padres sugeridos) seguramente seriam influenciados por variantes restritas a tais grupos especficos. Logo, para escolha das mdias pautou-se pela quantidade de acessos de tais sites, optando-se pelos mais acessados em cada pas, por inferir disso a possibilidade de que tais sites so aqueles em que mais falantes de tais lnguas naturais (e mais falantes de diversas partes de cada um dos pases, e porque no residentes tambm em outros pases) fazem comentrios, abrangendo assim uma comunidade ampla de falantes de cada uma dessas lnguas naturais e garantindo assim, mais amplamente, aqueles fenmenos caractersticos da linguagem ordinria – uma vez que, sendo falantes de lnguas naturais os mais diversos, certa flexibilidade de comunicao deve existir, alm do fato de que, por se tratar das mdias mais acessadas, pode-se pensar que se trata das mdias mais acessveis (no sentido de estarem presente no cotidiano) a um maior nmero de indivduos, proporcionando a existncia de um maior grau de interao. Fora isso, a razo pela qual coletar comentrios de mais de uma mdia se deu pela ideia de que *corpora* formados por dados coletados de uma nica fonte poderiam conter particularidades dessa prpria, distanciando-se de um uso geral da lngua pertencente a um pblico mais amplo – p.

¹⁴⁶ Privilegiaram-se notcias desse perodo; entretanto, na falta dessas, notcias do ano de 2012 foram coletadas, uma vez que so temporalmente prximas ao presente momento e no influenciam negativamente os propsitos deste trabalho.

ex., a coleta de comentários de uma só mídia jornalística poderia representar determinado grupo social leitor dessa única mídia.

Isso não significa que variantes regionais não interessam a dicionários bilíngues e/ ou híbridos – uma vez que estes podem sim conter este tipo de informação, podendo ser úteis a estudantes, tradutores etc. Contudo, considera-se que esta seria uma tarefa particular que não poderá ser discutida aqui. Ainda que um dicionário bilíngue/ híbrido possa conter informações regionais, não sendo este seu objetivo principal é preciso que o lexicógrafo delimite critérios para a escolha de quais regionalismos serão admitidos e/ ou preteridos. Logo, mais do que variantes regionais, relevantes são formas padronizadas da língua *standard*, já que – acredita-se – em primeiro lugar esta forma *standard* deve estar bem assimilada pelo estudante antes deste se aventurar em conhecer regionalismos.

Em segundo lugar, pensou-se em *corpora* que contemplassem assuntos dos mais variados, já que um tema específico poderia constituir uma linguagem particular¹⁴⁷. Neste sentido, pensou-se em coletar dessas doze mídias, seis brasileiras e seis italianas, comentários relacionados a cinco assuntos diversos (política, esporte, celebridades, cotidiano e mundo), por se inferir disso que, desta forma, os proferimentos de ambas as línguas não estariam restritos a um único assunto – uma vez que, não sendo assuntos por demais delimitados, uma ampla camada da população falante de determinada língua natural poderia sentir-se capaz de opinar, comentar etc, já que não se exige amplo conhecimento técnico prévio de tais assuntos.

Entretanto, julgou-se necessário também um critério capaz de possibilitar a coleta de comentários relacionados a esses assuntos. Neste sentido, pensou-se em coletar comentários a partir da delimitação quantitativa de matérias jornalísticas referentes a tais assuntos (cinco matérias de cada mídia correspondentes a cada um dos assuntos), de modo a facilitar a coleta e a proporcionar certo equilíbrio entre os assuntos que compõem os *corpora*. Importante frisar que as matérias não compõem os *corpora* de análise, mas funcionam como “guias de assuntos” (...).

¹⁴⁷ Essa é mais uma razão pela qual optou-se pelos comentários, ao invés de coletar conversas dispostas em fóruns de discussão, p. ex., uma vez que, em geral, os fóruns são organizados a partir de determinado tema/ assunto/ interesse, o que poderia restringir os resultados acerca de padrões de uso geral da língua. É verdade que uma possibilidade seria confeccionar *corpora* a partir da coleta de comentários de vários fóruns, de vários assuntos; no entanto, em geral os fóruns existentes são aqueles destinados a temas específicos, sendo difícil encontrar aqueles de temas corriqueiros.

Pelo assunto ‘política’, entendem-se matérias¹⁴⁸ acerca de qualquer aspecto da política local, a saber, do Brasil e da Itália. Por ‘esporte’, matérias sobre esportes em geral, praticados em qualquer local do planeta. Por ‘celebridades’, notícias sobre artistas e indivíduos famosos em geral, também oriundos de qualquer parte do planeta. Por ‘cotidiano’, notícias relacionadas a eventos/ acontecimentos locais, sobretudo relacionados a temas sociais, tais como educação, cultura e segurança. Por fim, pela categoria ‘internacional’ pensa-se em matérias cujo foco seja, sobretudo, algum acontecimento político/ econômico fora da Itália e do Brasil. Sabe-se da dificuldade de se classificar notícias segundo categorias amplas como essas – p. ex., uma notícia sobre a política econômica da Argentina deveria ser inserida em ‘Política’, ‘Internacional’ ou ‘Economia’? Entretanto, o mesmo ocorre na comunicação cotidiana, pois se um vizinho pergunta a outro “você leu sobre a reforma fiscal da Cristina Kirchner?”, o primeiro está interessado em discutir sobre assuntos internacionais, sobre economia ou sobre a política argentina? E o que o primeiro pretendeu com tal pergunta – interagir com o vizinho, passar o tempo ou constrangê-lo por saber previamente de seu desinteresse por tais temas?

Neste sentido, as definições acima esboçadas são apenas diretrizes, pautadas mais pela noção wittgensteiniana de “semelhanças de família” do que por categorias rígidas e inflexíveis. E não seria absurdo dizer que a noção de ‘semelhanças de família’ existe na própria constituição dos jornais *online* – e porque não dizer, da internet. Por exemplo, não é incomum clicar no link ‘Política’ de determinado jornal, escolher determinada matéria a partir de determinada manchete e, ao clicar sobre esta manchete, notar que a matéria em questão foi indexada sob outra categoria que não aquela ‘Política’ – p. ex., ‘Economia’ ou ‘Cotidiano’. Logo, importante dizer que ainda que se tenha clicado rigorosamente em cada um dos links correspondentes àquelas categorias (política, esporte, celebridades, cotidiano e mundo), aconteceram casos em que a matéria escolhida estava indexada sob outra categoria. Mesmo assim, optou-se por coletá-la, uma vez que, como se disse, tal delimitação é mais uma diretriz do que uma norma.

Outra questão importante é que a escolha das matérias se deu por estas estarem presentes na página principal de sua categoria correspondente. Por exemplo, se ao clicar no link ‘Política’, a página aberta continha cinco notícias sobre política, estas cinco foram privilegiadas – desde que contivessem comentários de leitores (internautas).

¹⁴⁸ Matéria e notícia são aqui tidas como sinônimos.

Somente no caso de uma ou mais dessas matérias não conter comentários é que se processou à busca pelo interior do site – busca por notícias antigas, busca pelas notícias mais lidas, busca pelas notícias mais comentadas etc. A tabela 4 mostra como cada mídia caracteriza suas matérias, segundo as categorias previamente definidas para o presente trabalho:

Assunto/ Mídia	Política	Cotidiano	Internacional	Esporte	Celebridades
UOL	Política	Cotidiano	Internacional	Esporte	Celebridades
Globo.com	Política	País	Mundo	Esportes	Ego
Yahoo.br	Política/ Brasil	Brasil	Mundo	Esportes	OMG!
Terra	Política/ Brasil	Brasil/ Cidades	Mundo	Esportes	Diversão> Gente
Ig	Política	Brasil	Mundo	Esportes	Gente
Folha Online	Política/ Poder	Cotidiano	Mundo	Esporte	F5
Libero	Politica	Milano/ Roma	Politica/ Esteri	Sport	Gossip
Yahoo.it	Politica	Italia	Mondo	Sport	Celebrity
La Repubblica	Pubblico/ Politica	Cronaca	Esteri	Sport	Persone
Virgilio	Politica	Cronaca	Esteri	Sport	Gossip
Il Corriere della Sera	Politica	Cronache/ locali	Esteri	Sport	Gossip
Tiscali	Politica	Cronaca	Esteri	Sport	Lifestyle/ Celebrity

Tabela 24: categorização dos temas segundo os doze sites pesquisados.

Tal como o critério adotado anteriormente em relação à escolha das mídias, optou-se por coletar apenas os comentários dispostos na página da própria notícia, localizados abaixo desta. Assim, se uma dada matéria apresenta cinco comentários, localizados abaixo de si própria, serão estes cinco os que serão coletados; se outra mídia, p. ex., tem

o hábito de só dispor três comentários abaixo de cada matéria, então só esses três é que serão coletados. Não serão coletados, assim, outros comentários – p. ex., obtidos a partir de links tais como ‘página seguinte’, ‘mais comentários’, ‘todos comentários’, ‘ver todos os comentários’ etc. A razão para tal não é simplesmente a de facilitar a coleta, mas também um modo de garantir que a coleta não seja necessariamente rígida, já que uma das características da linguagem ordinária é a flexibilidade. Ou seja: de modo a compor *corpora* os mais espontâneos e flexíveis possíveis, optou-se por este método (ainda que, pensando-se num *corpus* capaz de servir de base para a elaboração de dicionários, o maior número de comentários possíveis deve ser coletado).

Fora isso, relevante dizer que junto ao comentário do leitor (internauta) o nome e/ou *nickname* (todos serão tomados como *nickname*, já que não se pode garantir se tratar do nome real do leitor) do próprio foi preservado nos *corpora*, de modo a verificar, se necessário, se determinada ocorrência se deu em grande quantidade devido ao uso por uma só pessoa ou se por um público amplo, influenciando assim sua adoção ou não como padrão a ser dicionarizado.

Por sua vez, finalizada a coleta todos os comentários coletados foram salvos conjuntamente, novamente, num arquivo do *software* ‘Bloco de Notas’, de modo a poder ser analisado pelo software de Linguística de *Corpus* chamado *AntConc*, versão 3.2.4w, desenvolvido por Lawrence Anthony, professor-diretor do CELESE (*Center for English Language Education*), ligado à Faculdade de Ciências e Engenharia da *Waseda University*, localizada em Tóquio, Japão.

Segundo Serra (2011: 164), o *AntConc*

(...) é um extrator de termos semelhantes com o pacote do conhecido programa Word Smith Tools. Além de fazer a lista de possíveis termos, o programa permite ao usuário encontrar contextos e fazer colocação, em porcentagem e em números exatos, da frequência de um termo em um texto (*corpora*) especializado(s).

Ainda com o autor (2011: 164-165), as principais ferramentas disponibilizadas pelo *AntConc* são: *Concordance*, que permite a visualização dos contextos em que aparece determinada palavra; *Cluster*, que permite a visualização de sintagmas lexicais; *WordList*, que é a lista de todas as palavras presentes no *corpus*.

Vale comentar também, antes de finalizar este tópico, algumas dificuldades encontradas durante a coleta. Por exemplo, pode-se verificar desde o início da coleta que nem todas as mídias oferecem possibilidades de comentários a todas as notícias. Isso prejudicou de certa forma a coleta, porque, como se verá mais à frente, a princípio

não foi possível coletar o mesmo número de matérias/ comentários de todas as mídias de modo quantitativamente equânime.

Outro problema foi o número desigual de comentários dispostos pelas mídias na página da própria notícia. Por exemplo, se a Folha Online disponibiliza, na própria página da notícia, três comentários, o Yahoo Brasil não disponibiliza um número definido, podendo disponibilizar de dez a vinte comentários por notícia. Isso pode ter acarretado, de certa forma, uma predominância de comentários de leitores do Yahoo Brasil, em detrimento de leitores da Folha Online. Contudo, decidiu-se manter os *corpora* coletados por se considerar que essa disparidade não necessariamente representa um problema, afinal está-se interessado no *conteúdo* dos comentários – e não num número preciso deles (...).

Dados obtidos

Primeira coleta

Como se disse anteriormente foram escolhidas as seis mídias jornalísticas do Brasil e da Itália mais acessadas, objetivando com isso coletar comentários de cinco notícias de cada um dos cinco assuntos escolhidos. Neste sentido, a tabela 5 mostra o que se conseguiu coletar, a princípio:

	Nº de matérias por assunto				
Idioma	Política	Cotidiano	Internacional	Esportes	Celebridades
PT-BR	30	30	30	30	30
IT	25	24	25	30	20

Tabela 25: nº de matérias por assunto/ idioma.

Segundo esta tabela, nota-se que se conseguiu coletar de todas as mídias do Brasil (UOL, Yahoo.br, O Globo, IG, Folha Online, Terra) matérias de cinco notícias de cinco assuntos diferentes, totalizando trinta matérias de cada um. O mesmo não se deu com as mídias italianas; veja-se na tabela 6 o caso italiano em mais detalhe:

	Nº de matérias por mídia italiana/ assunto				
Mídias	Política	Cotidiano	Internacional	Esportes	Celebridades
Libero	0	0	0	5	0
Yahoo.it	5	5	5	5	5
La Repubblica	5	4	5	5	0

Virgilio	5	5	5	5	5
Il Corriere della Sera	5	5	5	5	5
Tiscali	5	5	5	5	5

Tabela 26: nº de matérias por assunto/ idioma do italiano.

Isso se deu porque tais mídias não disponibilizaram comentários em todas as matérias de todos os assuntos pretendidos. Da mídia *Libero*, p. ex., só se conseguiu coletar comentários das cinco notícias de esporte que se pretendeu, sendo que dos outros assuntos a mídia não disponibilizou comentários. O mesmo aconteceu com o *La Repubblica*, pois deste faltou a coleta de uma notícia do cotidiano e todas de celebridade – até pelo fato deste jornal apresentar poucas notícias acerca deste último assunto.

Diante disto, pensou-se em realizar uma nova etapa de busca para coleta; entretanto, como o objetivo aqui não é coletar matérias (que servem apenas como guias), mas comentários, preferiu-se realizar uma primeira investigação quantitativa acerca destes até agora coletados, antes de se pensar em tentar equivaler esses números. Curiosamente, viu-se que o *corpus* de comentários do italiano, apesar de ter sido composto a partir da coleta de um número inferior de matérias, se comparado com o *corpus* de comentários do português-brasileiro, possui um maior número de comentários do que este, como se nota na tabela 7:

Idioma	Política	Cotidiano	Internacional	Esportes	Celebridades
PT-BR	179	173	151	181	198
IT	308	226	148	201	150

Tabela 27: nº de comentários por assunto/ idioma.

Entretanto, ao se dizer que o objetivo é coletar comentários não se pretende com isso dizer que se buscou coletar um número preciso de comentários – uma vez que o que se pretende, de fato, é analisar o *conteúdo pragmático* contido num *corpus* único de cada língua, composto pelos mais variados exemplares de comentários. Neste sentido, para se ter uma noção quantitativa mais exata dos *corpora* do português-brasileiro e do italiano, processou-se uma nova pesquisa estatística, de modo a descobrir o número de *tokens* e *types* de cada *corpus*. Para tal, utilizou-se o *software AntConc*; na aba *WordList* oferecida por este, é possível visualizar o número de *tokens* e *types* contidas num *corpus* selecionado.

Em Linguística de *Corpus*, segundo Ribeiro (2004: 170), os *tokens* correspondem ao número total de itens e/ ou palavras, considerando as repetições dos mesmos, enquanto que os *types* correspondem a cada item ou palavra, desconsiderando as repetições. Por exemplo, imaginando um comentário do tipo “minha casa é amarela, sua casa é azul”, tal comentário possui oito *tokens* (‘minha’, ‘casa’, ‘é’, ‘amarela’, ‘sua’, ‘casa’, ‘é’, ‘azul’) e seis *types* (‘minha’, ‘casa’, ‘é’, ‘amarela’, ‘sua’, ‘azul’). Neste sentido, os dados que se obtêm são os seguintes, primeiro do português-brasileiro:

Idioma	Política	Cotidiano	Internacional	Esportes	Celebridades
<i>Tokens</i>	6109	6215	4767	5224	5136
<i>Types</i>	2541	2726	2161	2125	2289

Tabela 28: nº de *tokens/ types* por assunto, segundo *corpus* do português-brasileiro.

Abaixo, do italiano:

Idioma	Política	Cotidiano	Internacional	Esportes	Celebridades
<i>Tokens</i>	15008	10013	5778	6595	6649
<i>Types</i>	4553	3532	2494	2487	2699

Tabela 29: nº de *tokens/ types* por assunto, segundo *corpus* do italiano.

A comparação entre o número de *tokens/ types* de cada um dos *corpora* é o seguinte:

	<i>Tokens</i>	<i>Types</i>
PT-BR	27451	8265
IT	44033	10725

Tabela 30: nº de *tokens/ types* dos *corpora* do português-brasileiro e do italiano.

Importante dizer que, tal como não é objetivo aqui coletar um número específico de matérias, nem um número específico de comentários, não é também coletar um número preciso de *tokens* e *types*. Entretanto, como esta última pesquisa quantitativa mostrou um *corpus* quantitativamente maior do italiano em relação ao do português-brasileiro, pensou-se em tentar equivaler este número – já que o que se busca são palavras (e/ ou expressões) passíveis de se tornarem verbetes de dicionários. É verdade que a força ilocucionária não está presa às palavras e/ ou a um número específico dessas; por outro lado, como o que se pensa é sugerir padrões a serem dicionarizados, esta sugestão poderia ser prejudicada em função de *corpora* desiguais em termos quantitativos, uma vez que o estabelecimento de uma equivalência ou não-equivalência entre itens lexicais de duas línguas poderia ser prejudicado. Assim, propõe-se a coleta

de mais um número de matérias/ comentários, de modo a equiparar ao máximo ambos os *corpora*, já que um número idêntico é difícil (se não impossível).

Como critério para saber de qual ou de quais mídias seriam coletadas novas matérias/ comentários, realizou-se a seguinte estatística, também com o *AntConc*, de modo a verificar o quanto cada uma até aqui contribuiu, em termos de *types/ tokens*:

Mídia	Yahoo.br	UOL	Globo.com	Terra	IG	Folha Online
<i>Tokens</i>	4500	7504	2012	5036	6438	1961
<i>Types</i>	2141	2892	1115	2260	2779	1016

Tabela 31: nº de *tokens/ types* por mídia do *corpus* português-brasileiro.

Neste sentido, nota-se que Folha Online e Globo.com são as mídias das quais menos se coletou e/ ou as que menos ofereceram comentários – em função de apresentarem poucos comentários, ou de terem menos comentários por notícias, ou por seus comentários coletados terem sido mais sucintos etc. A mídia Globo.com, contudo, apresenta um número maior (2012 *tokens/ 1115 types*) ainda do que a Folha Online (1961 *tokens/ 1016 types*); fora isso, como mostram as tabelas 8 e 9, segundo as quais as maiores diferenças quantitativas entre o *corpus* italiano e português-brasileiro são as matérias de política e cotidiano, pensou-se então em privilegiar a coleta destas nessa nova leva, da seguinte forma: mais três matérias de política e cotidiano, duas de internacional, e uma de esporte e celebridade para a mídia Globo.com; para a Folha Online, o dobro disso (seis de política e cotidiano, quatro de internacional, duas de esporte e celebridade).

Acréscimo

Feito esse acréscimo, os novos dados são os seguintes:

Idioma	Política	Cotidiano	Internacional	Esportes	Celebridades
PT-BR	39	39	36	33	33
IT	25	24	25	30	20

Tabela 32: nº de matérias por assunto do *corpus* português-brasileiro.

Fez-se o acréscimo ao *corpus* do português-brasileiro, enquanto que o do italiano permaneceu inalterado, já que este último era maior quantitativamente do que o primeiro. Em termos de comentário por assunto, vejamos os novos números:

Idioma	Política	Cotidiano	Internacional	Esportes	Celebridades
PT-BR	205	200	167	189	209
IT	308	226	148	201	150

Tabela 33: nº de comentários por assunto do *corpus* português-brasileiro.

Por fim, e o que de fato interessa, considerando-se a tabela 9, que indica o *corpus* do italiano dividido por assunto, em termos de *types/ tokens*, deve-se compará-la com o novo *corpus* do português-brasileiro:

	Política	Cotidiano	Internacional	Esportes	Celebridades
<i>Tokens</i>	7216	6215	5372	5472	5415
<i>Types</i>	3036	2726	2371	2277	2386

Tabela 34: nº de *tokens/ types* por assunto, segundo *corpus* do português-brasileiro.

Sendo, então, que o novo total é o seguinte:

	<i>Tokens</i>	<i>Types</i>
PT-BR	30390	8938
IT	44033	10725

Tabela 35: nº de *tokens/ types* dos *corpora* do português-brasileiro e do italiano.

Observa-se que ainda existem diferenças quantitativas em termos de *tokens* e *types* entre o *corpus* do italiano e do português-brasileiro, o que talvez até demandasse uma nova tentativa de equivalência estatística; contudo, essa demanda talvez seja infinita, pois dificilmente dois *corpora* comparados serão numericamente equivalentes em termos de *types/ tokens*. Afora isso – e talvez até mais importante – considerando que o objetivo aqui é sugerir padrões linguísticos de natureza pragmática, quando então o critério qualitativo é mais relevante ou tão relevante quanto o critério quantitativo (sendo que tais padrões são visados de modo a serem dicionarizados), mais do que o número total de palavras (*tokens*) é relevante aqui saber *quais* as palavras (*types*) que ocorrem nesses *corpora* – e, neste sentido, nota-se que a diferença não é demasiada entre os dois *corpora*.

Assim, salienta-se ao pesquisador que quiser trabalhar com forças ilocucionárias direcionadas a dicionários de línguas naturais com o auxílio da LC que é necessário criar mecanismos capazes de manter essa proporção entre assunto/ mídia-fonte/ comentários, de modo a produzir resultados satisfatórios¹⁴⁹.

¹⁴⁹ Em seguida a tais etapas de coleta e confecção de *corpora*, foram realizadas discussões acerca de “procedimentos” para reconhecimento e extração de ilocuições via tais *corpora*. Ao todo foram cinco

“procedimentos”, tais como apontados na Metodologia (ver IV), apresentando resultados insatisfatórios (não porque a metodologia de coleta e confecção aqui apresentada foi incoerente, e sim porque os “procedimentos” de investigação, como se viu no tópico mencionado, foram equivocados). Afora isso, os *corpora* confeccionados são pequenos, o que pode ter também contribuído com os poucos resultados. Em função disso tudo, optou-se por *corpora* bem mais extensos (*Paisà* e *do Português*) e aprimoraram-se os critérios para busca e extração de ilocuções (ver 4.1). Entretanto, julga-se que, para a investigação de aspectos da dimensão pragmática da língua, tanto quanto os *corpora* orais os *corpora* escritos confeccionados a partir comentários de internautas, blogs, fóruns virtuais etc podem ser de grande valia.