

O discurso do norte-americano em livros de culinária brasileira: uma proposta de análise baseada em linguística de corpus

Rozane Rodrigues Rebechi
(Universidade de São Paulo)

Resumo: A Linguística de Corpus tem tido papel fundamental em pesquisas terminológicas e lexicográficas. No entanto, ainda tem sido pouco explorada como metodologia para a Análise do Discurso. Sob o enfoque da Linguística de Corpus, o presente trabalho tem por objetivo propor uma metodologia de análise do discurso do norte-americano em relação ao Brasil e ao brasileiro, tendo como corpus de estudo livros de culinária brasileira escritos originalmente em inglês, por norte-americanos. Os resultados obtidos mostraram que métodos semiautomáticos de análise ajudam a identificar padrões de forma mais prática e confiável do que seria possível a partir de uma leitura sequencial de textos individuais.

Palavras-chave: Linguística de Corpus; Culinária Brasileira; Análise do Discurso.

Uma simples pesquisa na Internet revela uma série de problemas envolvendo a tradução e/ou definição de termos da culinária brasileira em língua inglesa: palavras com grafia incorreta, falta de padronização na tradução/definição dos termos (cf. Costa, 2006) e confusão entre termos do espanhol e do português são alguns exemplos do que pode ser encontrado em *sites* de língua inglesa quando o assunto é culinária brasileira. Aliás, parece importante esclarecer que problemas de tradução envolvendo os pares de língua inglês-português não são exclusivos da culinária típica brasileira, mas atingem a culinária como um todo (cf. Teixeira, 2004).

Em relação aos problemas mencionados, poder-se-ia argumentar que, na Internet, encontra-se qualquer coisa, inclusive “variações” ortográficas não dicionarizadas, afirmações que não correspondem à verdade, textos de autoria não comprovada etc. Enfim, a rede mundial de computadores divulga toda sorte de informações e, muitas vezes, fica difícil separar o que é confiável do que não é. Portanto, com o objetivo de investigar, em uma pesquisa em nível de mestrado, como se dá a tradução dos termos referentes à culinária brasileira para a língua

inglesa, optei por analisar apenas textos impressos e publicados, pois, teoricamente, estariam “acima de qualquer suspeita”, uma vez que, em relação a obras impressas, a expectativa é que:

- a. o autor tenha feito uma pesquisa aprofundada acerca do tema que vai desenvolver;
- b. antes da publicação, o livro tenha sido submetido a uma rigorosa revisão; e
- c. obras impressas servem como referência da língua culta.

1 O corpus

Para construir o corpus de estudo, foram selecionados oito livros de receitas brasileiras disponíveis para compra no mercado norte-americano no período de 2007 a 2008. Ao dar início à análise das receitas presentes nesses livros, foi surpreendente descobrir que os “mal-entendidos” em relação à tradução da culinária típica brasileira não estão restritos ao que se divulga em *sites* da Internet. As obras impressas que compõem o corpus também revelaram problemas de tradução/definição de termos, tais como:

- Definições incorretas:
 Ex.: *Petiscos* → *candies made from grated hard cheese, sugar, egg yolks and lemon*.
- Substituições que acarretam apagamento de marcas culturais:
 Ex.: *1 cup cachaça or white rum* [receita de caipirinha].
- Tradução do termo por palavra não dicionarizada na língua de chegada:
 Ex.: *chuchu* → *chayotes or chocho*¹.
- Erros de ortografia:
 Ex.: *Lagarta* ao mostarda e champagne [lagarto ao molho mostarda e champanhe].
- Confusão entre termos da língua portuguesa e do espanhol falado nos outros países da América Latina:

¹ Não foi encontrada a palavra *chocho* nos dicionários pesquisados (*Cambridge International Dictionary of English*, *Macmillan English Dictionary for Advanced Learners of American English* e *Collins Cobuild English Language Dictionary*).

Ex: açougue → *a butcher shop; also called carniceria.*

Essas descobertas tornam evidente que mesmo os autores das obras impressas, pessoas que, em sua maioria, moraram no Brasil por algum tempo e demonstraram interesse em compartilhar essa vivência com seus compatriotas, muitas vezes desconhecem aspectos que se propõem a discutir e acabam levando ao leitor informações nem sempre confiáveis.

2 A tradução culinária

É indiscutível que a escassez de obras terminográficas bilíngues inglês-português na área da culinária dificulta muito a atuação do tradutor e/ou do escritor que trabalha com o tema (cf. Teixeira, 2004). No entanto, uma simples pesquisa na Internet poderia ter evitado algumas das confusões como as citadas acima. Além disso, ao se dispor a “levar” para a cultura norte-americana o que é típico do Brasil, o que leva um autor/tradutor a substituir ingredientes essenciais para a receita? Ao substituir “cachaça” por *rum*, por exemplo, o autor não está simplesmente “facilitando” a atuação do consulente/cozinheiro, mas sim descaracterizando uma bebida que é parte da cultura brasileira.

A partir dessas constatações em relação às receitas, surgiu o interesse em verificar o que existe em comum entre o que esses autores “falam” sobre o Brasil e o povo brasileiro nas introduções para os livros e para as receitas, pois os livros que compõem o corpus não se limitam a fornecer receitas de pratos típicos brasileiros. Em maior ou menor grau, todos oferecem textos introdutórios à obra e/ou às receitas, nos quais o(s) autor(es) relatam experiências vividas no Brasil, curiosidades, fatos históricos etc. Enfim, muito mais do que obras de referência culinária, os livros analisados pretendem também levar à cultura norte-americana fatos da cultura brasileira. Essas informações serviram de fonte para a presente pesquisa.

3 Textos que compõem o corpus: originais ou traduções?

É importante ressaltar que, entre os livros que compõem o corpus, nenhum faz menção ao fato de as receitas e os textos informativos serem traduções. Aliás, quando o tema é culinária, raramente é feita

qualquer menção ao texto-fonte. Em geral, receitas são trocadas, modificadas e ‘traduzidas’ sem levar em conta questões de direitos autorais.

No caso dos livros analisados, alguns autores comentam nas introduções sobre o primeiro contato que tiveram com o prato ou sobre quem forneceu a receita, mas alguns fazem questão de esclarecer que tomaram a liberdade de adaptar a receita, seja a fim de adequá-la ao paladar do norte-americano, seja pelo fato de levar ingredientes dificilmente encontrados naquele país. No entanto, parece óbvio que todas as receitas partiram de uma (ou mais) receita “original” em português, independentemente de se fazer ou não menção a esse processo.

Em relação aos textos informativos que formam o corpus deste estudo, também não se fala em tradução. Alguns autores apenas fornecem, ao final, uma lista de obras consultadas. Contudo, o fato de o texto de chegada não fazer menção a um ou mais textos de partida não significa que o processo tradutório não esteja envolvido. Em seu estudo sobre a tradução de textos jornalísticos, Zipser (2002) afirma que um texto da língua de partida pode servir como insumo para sua tradução, sem que a autoria esteja, necessariamente, explícita. Afinal, ao passar de uma língua/cultura para outra, o texto jornalístico muitas vezes sofre alterações necessárias. Autora denomina esse processo de tradução como representação cultural.

Para Lefevere (1992), a tradução (assim como a crítica literária, a antologização, a historiografia e a editoração) é sempre uma forma de reescrita do texto original, uma vez que leva em consideração aspectos (ideológicos do tradutor e receptor; restrições etc.) de recepção do texto traduzido. No caso da culinária, essas adaptações também são feitas, não só na transposição de uma língua/cultura para outra, mas mesmo dentro de uma mesma língua/cultura.

4 A Linguística de Corpus e a Análise do Discurso

A Linguística de Corpus (LC) que, segundo definição de Berber-Sardinha (2004, p. 3), “ocupa-se da coleta e da exploração de corpora, ou conjuntos de dados linguísticos textuais coletados criteriosamente, com o propósito de servirem para a pesquisa de uma língua ou variedade linguística”, tem tido papel fundamental em pesquisas terminológicas e lexicográficas, como pode ser observado em Sinclair (1991 e 2004), Mahlberg e O’Donnell (2008) e Stubbs (2002), entre muitos outros. No entanto, essa metodologia ainda tem sido pouco explorada em Análise

do Discurso (AD), só para citar a área de estudo em que esta pesquisa se insere.

Há, no entanto, indícios de que essa situação esteja mudando (cf. Orpin, 2005; Berber-Sardinha; Barbara, 2008), talvez impulsionada por Sinclair (2004, p. 10), que defende a junção das duas áreas por enxergá-las como “pilares gêmeos da pesquisa linguística”, uma vez que a tecnologia moderna ajuda a encontrar e lidar com as evidências, fator essencial na formulação de novas hipóteses.

5 Metodologia

Utilizando os pressupostos da Linguística de Corpus, esta pesquisa vem propor uma metodologia de análise do discurso do sujeito norte-americano (representado, aqui, pelos autores dos livros que compõem o corpus) em relação ao Brasil e ao brasileiro, por meio dos textos informativos presentes nos livros de receitas brasileiras escritos em língua inglesa.

A fim de comprovar a nacionalidade dos autores, o corpus de estudo foi compilado a partir de oito livros de receitas publicados originalmente em língua inglesa, que foram digitalizados e salvos em formato .txt, para serem explorados de maneira semiautomática com ajuda da ferramenta WordSmith Tools 5.0 (Scott, 2007) (doravante, WST5). Em seguida, esse corpus foi dividido em dois subcorpora: um subcorpus “receitas” e um subcorpus “textos informativos”. Para a presente pesquisa, foi utilizado apenas o subcorpus “textos informativos”, aqui considerado, portanto, como o corpus de estudo. A Tabela 1 mostra o número de palavras de cada livro que compõe o corpus:

Tabela 1: Número de palavras dos textos informativos dos livros que compõem o corpus de estudo

WordList	Livro	Nº Palavras
WLatbINF	<i>A Taste of Brazil</i> (Neeleman; Neeleman, 2007)	8.339
WLbclINF	<i>Brazilian Cooking</i> (Leroux, 1980)	1.475
WLbcjINF	<i>Brazil: A Culinary Journey</i> (Hamilton, 2005)	7.390
WLbctINF	<i>Brazil: A Cook's Tour</i> (Idone, 1995)	32.732
WLctbwINF	<i>Cooking the Brazilian Way</i> (Behnke; Duro, 2004)	5.879
WLdbcINF	<i>Delightful Brazilian Cooking</i> (Ang, 1993)	650
WLesbINF	<i>Eat Smart in Brazil</i> (Peterson; Peterson, 1995)	16.309
WLtcobINF	<i>The Cooking of Brazil</i> (Locricchio, 2005)	3.884
Total de palavras		76.657

Conforme pode ser verificado na Tabela 1, o tamanho dos textos informativos de cada livro varia bastante, e o número total de palavras do corpus soma 76.657. Esse número de palavras seria considerado insuficiente, por exemplo, no caso de uma pesquisa com fins lexicográficos. Porém, o critério de tamanho deve estar diretamente ligado à noção de representatividade do corpus (Aijmer; Altenberg, 1991), e, lembrando que esta pesquisa visa encontrar características em comum no que os autores dizem sobre o brasileiro nos livros de culinária, acreditamos que o corpus seja representativo para esse fim.

Por meio da função WordList do WST5, foram geradas as listas de todas as palavras dos textos de cada livro que compõe o corpus. As palavras de conteúdo entre as palavras mais frequentes dão uma pista em relação ao tema central do corpus, mas não são suficientes para evidenciar o que é característico dele. Segundo Stubbs (2002), são as palavras mais raras que costumam fornecer informações mais relevantes sobre o gênero textual estudado. Assim, o passo seguinte foi gerar uma lista de palavras-chave dos textos do corpus, usando como referência a lista de palavras do BNC, corpus de inglês britânico escrito e falado que possui mais de 100 milhões de palavras.

A lista de palavras-chave já evidencia muitas palavras de conteúdo do corpus. No entanto, em meio a palavras que remetem à culinária e à cultura brasileiras, encontram-se, também, palavras muito próprias de cada obra. Como o objetivo deste estudo é analisar o que é recorrente nos diferentes livros, escritos por diferentes autores para, possivelmente, identificar características comuns a eles, as listas de palavras-chave foram comparadas entre si (por meio da opção Database do WST 5.0)², evidenciando uma lista de palavras-chave-chave (para esta análise, foi estipulado que, para ser considerada chave-chave, a palavra deveria ser chave em um mínimo de cinco livros), mostrada na Tabela 2.

Apesar de haver um “empate” entre as palavras “Brazil”, “Brazilian” e “cooking” em relação à chavidade, uma vez que são chave nos oito livros do corpus, se levarmos em consideração que é possível se referir aos brasileiros como “the Brazilian people”, “the Brazilian”, além de “the Brazilians”, essa lematização provavelmente resultaria em uma chavidade ainda maior para a(s) palavra(s) usadas para se referir aos brasileiros em geral. Entretanto, a princípio, isso é

² Para instruções detalhadas sobre como utilizar a função Database do programa, ver Help do WST5.

apenas uma suposição, pois esse levantamento não foi realizado. De qualquer forma, acreditamos que a palavra de busca “Brazilians” pode revelar resultados bastante relevantes para esta análise. Portanto, foi a escolhida para darmos continuidade à análise.

Tabela 2: Primeiras 40 palavras-chave-chave do corpus, em ordem decrescente

N	KW	Texts	%	Over- all Freq.	N	KW	Texts	%	Over- all Freq.
1	BRAZIL	8	100.00	314	21	MEAL	6	75.00	9
2	BRAZILIAN	8	100.00	342	22	MEAT	6	75.00	25
3	COOKING	8	100.00	142	23	MEATS	6	75.00	9
4	AND	7	87.00	3,113	24	NATIVE	6	75.00	0
5	DISHES	7	87.00	177	25	OIL	6	75.00	21
6	FOOD	7	87.00	204	26	PORTUGUESE	6	75.00	73
7	RECIPES	7	87.00	91	27	RICE	6	75.00	5
8	AFRICAN	6	75.00	99	28	SHRIMP	6	75.00	0
9	AMAZON	6	75.00	55	29	BAHIA	5	62.00	9
10	BEANS	6	75.00	79	30	COCONUT	5	62.00	3
11	BEEF	6	75.00	57	31	COOKED	5	62.00	0
12	BRAZILIANS	6	75.00	76	32	CORN	5	62.00	4
13	BRAZIL'S	6	75.00	52	33	CUISINE	5	62.00	8
14	CHEESE	6	75.00	51	34	DE	5	62.00	74
15	HICKE	6	75.00	57	35	LICIOUS	5	62.00	3
16	CULINARY	6	75.00	46	36	DESSERTS	5	62.00	3
17	DISH	6	75.00	100	37	DRIED	5	62.00	8
18	FISH	6	75.00	179	38	FAVORITE	5	62.00	3
19	FRUITS	6	75.00	103	39	FEIJOADA	5	62.00	1
20	INGREDIENTS	6	75.00	68	40	FLAVOR	5	62.00	4

A etapa seguinte do estudo deu-se a partir das linhas de concordância em que a palavra de busca (“Brazilians”) está inserida. Essa função possibilita a visualização da palavra em seu contexto (Key Word in Context). Usando o utilitário Concord do WST5, foram geradas as linhas de concordância com todas as ocorrências da palavra de busca “Brazilians” no corpus:

N	Concordance	Set	Tag	Word #	t. #	os.	#	os.	#	os.	t. #	os.	File
1	the state of Bahia, however, that many Brazilians associate with outstanding			1,774	94	3%	0	1%			0	1%	tcob06_inf.txt
2	and once you taste it you'll know why Brazilians love it. As you travel across			1,062	56	9%	0	9%			0	9%	tcob06_inf.txt
3	believe originated in Minas Gerais. Brazilians call their chicken soup, canja,			1,048	56	1%	0	8%			0	8%	tcob06_inf.txt
4	festive atmosphere is it any wonder that Brazilians call their hometown Cidade			420	18	0%	0	9%			0	9%	tcob06_inf.txt
5	from the dorsal hump of the zebu steer. Brazilians have been known to walk out			6,095	295	2%	0	0%			0	0%	esb10_inf.txt
6	know as guava is called goiaba by the Brazilians , and in the form of a sweet			1,076	54	0%	0	4%			0	4%	esb10_inf.txt
7	can be an overwhelming experience. Brazilians use fruits in many ways. They			1,006	51	9%	0	3%			0	3%	esb10_inf.txt
8	and include a monthly newsletter, The Brazilians , which contains articles about			596	39	8%	0	1%			0	1%	esb08_inf.txt
9	which has the largest population of Brazilians in the United States. Stores			76	2	4%	0	0%			0	0%	esb08_inf.txt
10	from 7-11 PM. In metropolitan areas Brazilians dine late. If you arrive much			238	13	1%	0	1%			0	1%	esb06_inf.txt
11	will be plenty of items to identify. While Brazilians eat a light breakfast, the			156	8	9%	0	0%			0	0%	esb06_inf.txt
12	the beverage of choice for most Brazilians , replacing long-established			3,609	162	5%	0	1%			0	1%	esb05_inf.txt
13	that a recent guide book says that Brazilians do not do very much with			2,180	95	0%	0	5%			0	5%	esb05_inf.txt
14	diet. To this day the diet of many Brazilians includes few leafy green			782	34	4%	0	0%			0	0%	esb05_inf.txt
15	they quickly learned from the native Brazilians how to make manic meal			633	28	8%	0	6%			0	6%	esb05_inf.txt
16	which is as basic to the diet of Brazilians today as it was to the early			440	18	4%	0	1%			0	1%	esb05_inf.txt
17	65). The vendors, and many of the other Brazilians around you, will be happy to			240	12	6%	0	7%			0	7%	esb02_inf.txt
18	associated with particular celebrations, Brazilians also enjoy them throughout			98	6	1%	0	6%			0	6%	ctbw_06inf.txt
19	are popular on hot summer days. Brazilians enjoy thick fruit shakes and			120	7	7%	0	1%			0	1%	ctbw_05inf.txt
20	Ever since the arrival of the Portuguese, Brazilians have loved sweets. The			13	0	1%	0	0%			0	0%	ctbw_05inf.txt

Figura 1: Primeiras 20 linhas de concordância com a palavra de busca “Brazilians”

A ferramenta retornou 76 linhas de concordância, resultado perfeitamente passível de análise manual. No entanto, o número de linhas de concordância pode ser muito maior no caso de outras palavras de busca ou de *corpora* mais extensos. Além disso, o uso de ferramentas computacionais na exploração do *corpus* também revela informações que poderiam ficar ‘escondidas’ durante a leitura sequencial, visto que esta pode ‘direcionar’ o pesquisador a encontrar a informação que está procurando para provar algo que já intuía, impedindo-o de “enxergar” dados novos para a pesquisa.

Como a presente pesquisa visa identificar o que é recorrente nos textos, o passo seguinte foi verificar os principais colocados da palavra de busca, ou seja, palavras que ocorrem ao redor da palavra de busca. Palavras isoladas só podem servir como ponto de partida, pois são as colocações que criam denotações (Stubbs, 2005). Como esta pesquisa

não tem pretensões lexicográficas ou terminográficas, o objetivo não está na busca por colocações, ou seja, tendência de duas ou mais palavras co-ocorrerem (cf. Hunston, 2002), mas sim na busca por palavras associadas a “Brazilians”, na tentativa de identificar características comumente atribuídas ao brasileiro para, possivelmente, chegar a uma categorização.

A partir das linhas de concordância, clicando em Collocates, visualizamos os colocados da palavra de busca:

Tabela 3: Colocados da palavra de busca “Brazilians”, em ordem decrescente de frequência da associação

N	Word	With	Relation	Texts	Total
1	BRAZILIANS	brazilians	0	25	77
2	THE	brazilians	0	21	47
3	OF	brazilians	0	14	18
4	AND	brazilians	0	11	16
5	TO	brazilians	0	8	14
6	A	brazilians	0	10	14
7	IN	brazilians	0	9	11
8	IS	brazilians	0	8	10
9	THAT	brazilians	0	7	9
10	FOR	brazilians	0	7	8
11	MANY	brazilians	0	7	8
12	IT	brazilians	0	6	8
13	ALL	brazilians	0	5	7
14	HAVE	brazilians	0	6	7
15	ENJOY	brazilians	0	5	6
16	THEIR	brazilians	0	4	6
17	AS	brazilians	0	5	6
18	NOT	brazilians	0	5	5
19	WHO	brazilians	0	4	5
20	DO	brazilians	0	4	5
21	LOVE	brazilians	0	3	4
22	ARE	brazilians	0	4	4
23	USE	brazilians	0	4	4
24	WITH	brazilians	0	4	4
25	BY	brazilians	0	4	4
26	CALL	brazilians	0	3	4
27	WERE	brazilians	0	3	3
28	WHAT	brazilians	0	3	3
29	BASIC	brazilians	0	3	3
30	BUT	brazilians	0	3	3
31	ALSO	brazilians	0	3	3
32	ALWAYS	brazilians	0	3	3
33	DAY	brazilians	0	3	3
34	FROM	brazilians	0	3	3
35	LIKE	brazilians	0	3	3
36	DIET	brazilians	0	2	3
37	EAT	brazilians	0	3	3

A busca acima leva em consideração apenas a frequência em que uma determinada palavra ocorre no entorno³ da palavra de busca. Podemos observar na Tabela 3 que o colocado mais frequente (o programa sempre mostra a palavra de busca em primeiro lugar na contagem dos colocados, portanto o colocado mais frequente é o que aparece em segundo lugar e assim sucessivamente) com a palavra de busca é o artigo “the”, que acompanha a palavra de busca em 49% de suas ocorrências. No entanto, essas palavras gramaticais não fornecem muita informação em relação às características atribuídas aos brasileiros; portanto, a pesquisa de colocados teve que ser refinada.

Para gerar a lista de colocados com a palavra de busca, deve-se, antes mesmo de gerar as linhas de concordância, fazer os ajustes necessários. Por meio da função Settings, do Concord, foi estabelecido que o colocado deveria aparecer pelo menos três vezes com a palavra de busca (como forma de descartar escolhas aleatórias) e em uma janela de quatro palavras à direita e quatro à esquerda dela⁴. A medida de associação utilizada foi *Mutual Information* (MI), visto que os resultados gerados a partir desse índice independem do tamanho do corpus (cf. Hunston, 2002).

Uma associação não aleatória é aquela que é mais comum do que o esperado (Berber-Sardinha, 2004, p. 201). O cálculo MI leva em consideração três aspectos:

- a. a frequência com que cada colocado aparece no corpus de estudo;
- b. a frequência com que a palavra de busca aparece no corpus de estudo; e
- c. a frequência com que aparecem juntos na janela selecionada (Scott, 2007).

A Tabela 4 mostra os principais colocados com a palavra de busca, usando o índice MI para calcular a força de atração (resultados em ordem decrescente):

³ Em Settings, é possível ajustar a ferramenta para que considere como colocado uma palavra que esteja dentro de uma determinada janela, à direita e/ou à esquerda, da palavra de busca.

⁴ Não existe unanimidade em se tratando de horizonte (distância à direita ou esquerda) entre a palavra de busca e o colocado, mas Sinclair (1995 *apud* Berber-Sardinha, 2004) considera que essa janela não deve ser maior do que quatro palavras à direita e quatro à esquerda.

**Tabela 4: Colocados com a palavra de busca “Brazilians”,
 utilizando o cálculo *Mutual Information*.**

N	Word	With	Relation	Texts	Total
1	BRAZILIANS	brazilians	9,959	25	77
2	ENJOY	brazilians	7,9	5	6
3	BASIC	brazilians	7,844	3	3
4	LOVE	brazilians	7,204	3	4
5	CALL	brazilians	6,915	3	4
6	DIET	brazilians	6,335	2	3
7	USE	brazilians	6,178	4	4
8	ALWAYS	brazilians	6,052	3	3
9	EAT	brazilians	6,021	3	3
10	ALL	brazilians	5,587	5	7
11	DO	brazilians	5,513	4	5
12	MANY	brazilians	5,337	7	8
13	WHAT	brazilians	5,315	3	3
14	WHO	brazilians	5,237	4	5
15	DAY	brazilians	5,118	3	3
16	HAVE	brazilians	5,108	6	7
17	NOT	brazilians	5,052	5	5
18	THEIR	brazilians	4,425	4	6
19	BUT	brazilians	4,415	3	3
20	LIKE	brazilians	4,364	3	3
21	THAT	brazilians	4,358	7	9
22	IT	brazilians	4,235	6	8
23	WERE	brazilians	4,101	3	3
24	ALSO	brazilians	4,085	3	3
25	FOR	brazilians	3,716	7	8
26	BY	brazilians	3,66	4	4
27	AS	brazilians	3,446	5	6
28	IS	brazilians	3,267	8	10
29	THE	brazilians	3,213	21	47
30	TO	brazilians	3,168	8	14
31	IN	brazilians	2,888	9	11
32	A	brazilians	2,818	10	14
33	ARE	brazilians	2,767	4	4
34	OF	brazilians	2,759	14	18
35	FROM	brazilians	2,693	3	3
36	AND	brazilians	2,34	11	16
37	WITH	brazilians	2,252	4	4

A Tabela 4 mostra os mesmos colocados com a palavra de busca “Brazilians”; porém, em ordem decrescente em relação à força de associação (*relation*). Church e Hanks (1990 *apud* Dayrell, 2007) sugerem um MI mínimo de três para identificar padrões lexicais, pois, quanto

maior o MI, menor a chance de a co-ocorrência da palavra de busca com o colocado se dever ao acaso (Sinclair, 1991). Apesar de este estudo não ter por objetivo a identificação de padrões lexicais, também acreditamos que itens lexicais fornecem informações mais relevantes para a continuidade da análise. Estabelecendo três como número de corte, poder-se-iam descartar sete colocados (“in”, “a”, “are”, “of”, “from”, “and” e “with”).

Entre os 26 colocados com MI maior do que três, nove (34,6%) são verbos (“enjoy”, “love”, “call”, “use”, “eat”, “do”, “have”, “like” e “were”). Apesar de não estar entre os colocados com MI maior do que três, ficou decidido que a análise abrangeria o verbo “are”, pois pode revelar informações interessantes para a análise. Naturalmente, isso é intuitivo, mas, segundo Stubbs (1995), nenhum procedimento pode ser completamente automático, pois a intuição sobre o que é interessante pesquisar é o início de toda pesquisa.

A etapa seguinte da pesquisa visa analisar o contexto em que a palavra de busca (“Brazilians”) ocorre com cada colocado principal (no caso, os verbos). Feitos os ajustes necessários, a ferramenta Concord revelou as linhas de concordância em que a palavra de busca aparece com o colocado:

N Concordance	
1	associated with particular celebrations, Brazilians also enjoy them throughout
2	are popular on hot summer days. Brazilians enjoy thick fruit shakes and
3	With a day off from work and school, Brazilians enjoy picnics during pleasant
4	continue late into the evening as the Brazilians enjoy their tasty food and
5	and dinner. Nowadays, the upper class Brazilians enjoy going out to dinner in
6	oil and vinegar and chopped cilantro. Brazilians traditionally enjoy beef tongue,

Figura 2: Linhas de concordância em que a palavra de busca “Brazilians” aparece com o colocado “enjoy”.

Em seguida, foi analisado o contexto completo em que a palavra de busca aparece com cada colocado, possibilitando o agrupamento de algumas características atribuídas ao povo brasileiro:

- Festeiros, extrovertidos:
 - “With a day off from work and school, **Brazilians** **enjoy** picnics during pleasant weather, or they may join friends and family for meals at restaurants.”

- “All **Brazilians** love a party. If no reason can be found for a celebration, Brazilians will manage to invent one.”⁵
- Adeptos de hábitos alimentares pouco saudáveis (ingerem uma quantidade muito pequena de alimentos naturais, privilegiando os doces em sua dieta):
 - *Generally, **Brazilians are** not great salad eaters, and this display of simply prepared salads is a nice surprise.*
 - *The **Brazilians**, who **have** never heard of caffeine-free coffee and who still manage to sleep eight hours per night, drink numerous small cups of coffee during the day, either in the office or in the botequim on the corner: “Um cafezinho por favor!”*
- Místicos:
 - ***Brazilians are** above all people with a belief in the supernatural; every so often one feels the touch of strange influences, and I occasionally hear stories of apparitions, ghosts, and other enchantments.*
- Amadores, em se tratando de gastronomia:
 - *I learned to cook by feel like the **Brazilians do**.*
 - *It is incredible that a recent guide book says that **Brazilians** do not **do** very much with meat, other than just cooking it!*

Conforme mencionado, esta pesquisa não tem por objetivo uma análise aprofundada do discurso do norte-americano em relação ao Brasil e ao brasileiro⁶, mas sim propor uma metodologia de análise baseada em Linguística de Corpus. Porém, acreditamos que os resultados mostrados acima serviriam de insumo para um analista. Por meio dos contextos obtidos com ajuda de ferramenta computacional (WST5), foi possível chegar a algumas características que os autores dos livros que compõem o corpus atribuem ao brasileiro.

Alguns autores fazem questão de ressaltar que pretendem, por meio da obra, transmitir “sua” visão em relação ao Brasil ao relatar “suas” experiências pessoais. Mas será possível fazer uma análise unicamente pessoal de um país e seu povo? O sujeito do enunciado nunca é neutro, pois está sempre inserido em uma história, em uma cultura (Hall, 1990 *apud* Woodward, 2000). Apesar de ter a ilusão de di-

⁵ Todas as ênfases (palavras em negrito) são minhas.

⁶ Uma análise mais aprofundada do discurso do norte-americano nos textos informativos do corpus pode ser verificada em Rebechi (2010).

zer o que quer e de que a interpretação por parte do interlocutor só pode ser uma, o autor transfere para a obra escolhas ideológicas compartilhadas pelo grupo ao qual pertence, fato que impossibilita um olhar puramente pessoal (Orlandi, 2001), e não existe sobreposição perfeita entre o que é dito e o que é compreendido. Portanto, pode-se concluir que o que está sendo analisado não é a “opinião”, “visão”, “interpretação” de um sujeito autor norte-americano em relação ao sujeito brasileiro, mas sim do norte-americano em geral, representado por alguns “sujeitos”.

6 Considerações finais

Com este estudo, procuramos mostrar que, apesar de a intuição sobre o que é importante analisar estar presente em vários momentos, métodos semiautomáticos ajudam a identificar padrões e podem documentar o uso característico de palavras de forma muito mais confiável e completa do que seria possível a partir de uma leitura intuitiva de textos individuais. O método convencional de Análise do Discurso pode levar o analista a buscar no texto “evidências” que comprovem suas hipóteses, deixando de lado informações que possam ser contraditórias a essas hipóteses. Uma análise baseada em Linguística de Corpus vem comprovar que textos individuais só podem ser explicados se comparados ao que é considerado normal e esperado em relação ao uso geral da língua, e essa é precisamente a informação comparativa que dados quantitativos de corpus podem fornecer. As ferramentas computacionais oferecem várias técnicas de exploração para investigar características de textos e corpora (Stubbs, 2005), e não é preciso saber a intenção de um autor para interpretar seu texto. O texto é autônomo; o autor é irrelevante para a interpretação (Stubbs, 1996, p. 6).

Referências bibliográficas

- AIJMER, K.; ALTENBERG, B. (Org.). *English corpus linguistics: studies in honour of Jan Svartvik*. Londres: Longman, 1991.
- ANG, E. T. *Delightful Brazilian cooking*. Seattle: Ambrosia, 1993.
- BEHNKE, A.; DURO, K. L. *Cooking the Brazilian way*. Minneapolis: Lerner Publications Company, 2004.
- BERBER-SARDINHA, A. P. *Linguística de corpus*. Barueri: Manole, 2004.

BERBER SARDINHA, A. P.; BARBARA, L. Linguística de *corpus* e análise do discurso. In: CALDAS-COULTHARD, C. R.; SCLiar-CABRAL, L. *Desvendando discursos: conceitos básicos*. Florianópolis: UFSC, 2008.

COSTA, A. T. P. *Brasil mostrando a sua cara: estratégias de tradução no material de divulgação cultural – um estudo baseado em corpus*. Dissertação (Mestrado em Linguística Aplicada) – Departamento de Línguas Estrangeiras e Tradução, Universidade de Brasília, Brasília, 2006.

DAYRELL, C. A quantitative approach to compare collocational patterns in translated and non-translated texts. *International Journal of Corpus Linguistics*, v. 12, n. 3. p. 375-414. 2007.

HAMILTON, C. *Brazil: a culinary journey*. Nova Iorque: Hippocrene Books, 2005.

HUNSTON, S. Methods in corpus linguistics: Beyond the concordance line. In: HUNSTON, S. *Corpora in applied linguistics*. Cambridge: CUP, 2002. p. 67-95.

IDONE, C. *Brazil: a cook's tour*. New York: Clarkson/Potter, 1995.

LEFEVERE, A. *Translation, rewriting, and the manipulation of literary fame*. Londres: Routledge, 1992.

LEROUX, G. *Brazilian cooking*. [s.l.]: Les Éditions du Pacifique, 1980.

LOCRICCHIO, M. *The cooking of Brazil*. Tarrytown: Benchmark, 2005.

MAHLBERG, M.; O'DONNELL, M. B.. A fresh view of the structure of hard news stories. In: EUROPEAN SYSTEMIC FUNCTIONAL LINGUISTICS CONFERENCE AND WORKSHOP, 19., 2008, Liverpool. *Proceeding...* Liverpool: University of Liverpool, 2008. Acesso em: 10 set. 2009.

NEELEMAN, G.; NEELEMAN, R. *A taste of Brazil*. São Paulo: Marco, 2007.

ORLANDI, E. P. *Discurso e texto: formulação e circulação dos sentidos*. Campinas: Pontes, 2001.

ORPIN, D. Corpus linguistics and critical discourse analysis: examining the ideology of sleaze. *International Journal of corpus linguistics*, v. 10, n. 1, p. 37-61, 2005.

PETERSON, J.; PETERSON, D. *Eat smart in Brazil: how to decipher the menu, know the market foods & embarking on a tasting adventure*. Wisconsin: Ginkgo, 1995.

REBECHI, R. R. O discurso do norte-americano em livros de culinária brasileira: um estudo baseado em linguística de corpus. *Revista Veredas*, v. 13, n. 2, p. 117-133, 2009.

SCOTT, M. *WordSmith tools, version 5*. Oxford: OUP, 2007. Disponível em: <<http://www.lexically.net/wordsmith/>>. Acesso em: 6 jul. 2008.

SINCLAIR, J. *Corpus, concordance, collocation*. Oxford: OUP, 1991.

SINCLAIR, J. *Trust the text*. Londres: Routledge, 2004.

STUBBS, M. Collocations and semantic profiles: on the cause of the trouble with quantitative studies. *Functions of Language*, v. 2, n. 1, 1995. Disponível em: <<http://www.uni-trier.de/uni/fb2/anglistik/Projekte/stubbs/cause.htm>>. Acesso em: 02 mar. 2007.

STUBBS, M. *Text and Corpus analysis: computer-assisted studies of language and culture*. Oxford: Blackwell Publishers, 1996.

STUBBS, M. *Words and phrases: corpus studies of lexical semantics*. Oxford: Blackwell, 2002.

STUBBS, M. Conrad in the computer: examples of quantitative stylistic methods. *Language and Literature*, v. 14, n. 1, p. 5-24, 2005. Disponível em: <<http://www.uni-trier.de/uni/fb2/anglistik/Projekte/stubbs/conrad.htm>>. Acesso em: 3 abr. 2009.

TEIXEIRA, E. D. *Receita qualquer um traduz. Será? – a Culinária como área técnica de tradução*. Dissertação (Mestrado em Estudos Linguísticos e Literários em Inglês) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2v, 2004.

WOODWARD, K. Identidade e diferença: uma introdução teórica e conceitual. In: SILVA, T. T. (Org.). *Identidade e diferença: a perspectiva dos estudos culturais*. Petrópolis: Vozes, 2000. p. 7-72.

ZIPSER, M. E. *Do fato a reportagem: as diferenças de enfoque e a tradução como representação cultural*. Tese (Doutorado em Língua e Literatura Alemã) – Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, São Paulo, 2002.