



**UNIVERSIDADE ESTADUAL DO CEARÁ
CENTRO DE HUMANIDADES
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA APLICADA
DOUTORADO EM LINGUÍSTICA APLICADA**

SARAH VIRGINIA CARVALHO RIBEIRO

**AVALIAÇÃO DA PERCEPÇÃO DE TRADUTORES AUTOMÁTICOS *ONLINE* A
EXPRESSÕES IDIOMÁTICAS E SUAS METÁFORAS CONCEITUAIS
SUBJACENTES**

**FORTALEZA – CEARÁ
2018**

SARAH VIRGINIA CARVALHO RIBEIRO

AVALIAÇÃO DA PERCEPÇÃO DE TRADUTORES AUTOMÁTICOS *ONLINE* A
EXPRESSÕES IDIOMÁTICAS E SUAS METÁFORAS CONCEITUAIS
SUBJACENTES

Tese apresentada ao Curso de Doutorado em Linguística Aplicada do Programa de Pós-graduação em Linguística Aplicada do Centro de Humanidades da Universidade Estadual do Ceará, como requisito parcial para obtenção do título de Doutora em Linguística Aplicada.
Área de Concentração: Linguagem e Interação

Orientadora: Prof^a. Dr^a. Paula Lenz Costa Lima.

FORTALEZA – CEARÁ

2018

Dados Internacionais de Catalogação na Publicação

Universidade Estadual do Ceará

Sistema de Bibliotecas

Ribeiro, Sarah Virginia Carvalho.

Avaliação da percepção de tradutores automáticos online a expressões idiomáticas e suas metáforas conceituais subjacentes [recurso eletrônico] / Sarah Virginia Carvalho Ribeiro. - 2018.

1 CD-ROM: il.; 4 1/2 pol.

CD-ROM contendo o arquivo no formato PDF do trabalho acadêmico com 215 folhas, acondicionado em caixa de DVD Slim (19 x 14 cm x 7 mm).

Tese (doutorado) - Universidade Estadual do Ceará, Centro de Humanidades, Programa de Pós-Graduação em Linguística Aplicada, Fortaleza, 2018.

Área de concentração: Linguagem e Interação.

Orientação: Prof.^a Dra. Paula Lenz Costa Lima.

1. Fraseologismo. 2. Métricas de Avaliação de Tradução Automática. 3. Tradução Semântica. I. Título.

SARAH VIRGINIA CARVALHO RIBEIRO

AVALIAÇÃO DA PERCEPÇÃO DE TRADUTORES AUTOMÁTICOS ONLINE A EXPRESSÕES IDIOMÁTICAS E SUAS METÁFORAS CONCEITUAIS SUBJACENTES.

Tese apresentada ao Curso de Doutorado em Linguística Aplicada do Programa de Pós-Graduação em Linguística Aplicada do Centro de Humanidades da Universidade Estadual do Ceará, como requisito parcial para a obtenção do título de Doutora em Linguística Aplicada.

Área de Concentração: Linguagem e Interação

Aprovada em: 28 de fevereiro de 2018.

BANCA EXAMINADORA



Profa. Dra. Paula Lenz Costa Lima (Orientadora)
Universidade Estadual do Ceará – UECE



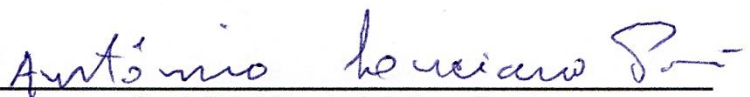
Prof. Dr. Tiago Martins da Cunha
Universidade da Integração Internacional da Lusofonia Afro-Brasileira – UNILAB



Profa. Dra. Silvia Malena Modesto Monteiro
Graduação/Universidade Estadual do Ceará – UECE



Prof. Dr. Wilson Júnior de Araújo Carvalho
Universidade Estadual do Ceará – UECE



Prof. Dr. Antonio Luciano Pontes
Universidade Estadual do Ceará – UECE

Para meus maiores tesouros: Arthur, Raul
e Lara.

AGRADECIMENTOS

Primeiramente, a Deus, que com bondade e misericórdia, concedeu-me mais esta bênção na vida.

Aos meus pais, que com dedicação, investiram sempre em educação acadêmica e moral, dando exemplo de perseverança e fé a mim e a minhas irmãs.

Ao meu esposo, Arthur, e meus filhos, Raul e Lara, pela compreensão e carinho nos momentos de minha ausência em nossa convivência diária.

Aos familiares e amigos, pela força e estímulo no caminho árduo de estudo até o fim.

Aos padres redentoristas da Paróquia de São Raimundo Nonato, pelo apoio com orações, desde a fase da seleção até o momento da defesa, e que plantaram a semente do amor pelos estudos, quando aluna do Colégio Redentorista.

À Profa. Paula Lenz, por orientações compartilhadas com seu saber, material e imaterial, em diversos momentos desta caminhada.

Aos colegas participantes do Grupo COMETA, pela amizade construída, compartilhando experiências e angústias.

Aos componentes da Banca Examinadora e suas contribuições, nas diferentes fases do processo, em especial ao Prof. Tiago Cunha, na ajuda da escolha da temática pesquisada e empréstimo de material de leitura.

À coordenação, professores e servidores do Pós-LA, que contribuíram de maneira ímpar, no acompanhamento de todo o processo cognitivo.

Aos colegas do Pós-LA, pelos momentos de construção de conhecimento e amizade.

Aos colegas do IFCE, que com a ajuda de redução de minhas horas em sala de aula, propiciaram maior dedicação aos estudos. Agradecimento especial ao colega Jean Lima, que auxiliou na construção de meu projeto de tese para a seleção do programa.

Aos tradutores profissionais participantes, sem os quais esta pesquisa não poderia acontecer.

“A leitura, como a comida, não alimenta
senão digerida.”

(Marquês de Maricá)

*“Cooking is all about people. Food is
maybe the only universal thing that really
has the power to bring everyone together.
No matter what culture, everywhere around
the world, people get together to eat.”*

(Guy Fieri)

*“Food is not rational. Food is culture, habit,
craving and identity.”*

(Jonathan Safran Foer)

*“Food for the body is not enough. There
must be food for the soul.”*

(Dorothy Day)

RESUMO

Esta pesquisa avaliou a percepção de tradutores automáticos *online* (TAO's) gratuitos em relação às expressões idiomáticas (EI's) e suas metáforas conceituais subjacentes. Desde sua origem, a Tradução Automática tem gerado grandes expectativas em relação à sua qualidade, intercalando momentos de grande interesse e investimentos, com outros de quase nenhum. Por outro lado, as EI's (fraseologismos) e outras linguagens figuradas têm despertado cada vez mais o interesse dos estudiosos, devido à forma sistemática e frequente em que ocorrem em todas as línguas, nos diferentes gêneros discursivos. Embora bastante comuns, as EI's parecem particularmente problemáticas de serem traduzidas por TAO's gratuitos. A análise de traduções de EI's sob a perspectiva da Teoria da Metáfora Conceitual, pode contribuir para o aperfeiçoamento dessas ferramentas. Nessa abordagem, as EI's podem ser licenciadas por metáforas conceituais e, portanto, compartilhadas por várias línguas. Isto explicaria, por exemplo, porque várias traduções de EI's são também EI's. Avaliamos 43 EI's com termos de alimento em inglês e suas respectivas traduções em português, coletadas em dois dicionários de EI's. Inicialmente, as expressões foram analisadas para se verificar possíveis (a) interferências culturais em suas composições e (b) licenciamentos por metáforas conceituais. Em seguida, utilizando o *Google Books N-gram Viewer* e o concordanciador *online Webcorp*, verificamos a frequência de uso das EI's e selecionamos os excertos a serem traduzidos por três TAO's gratuitos e três tradutores humanos profissionais (THP's). Todas as traduções foram avaliadas quantitativamente, utilizando as métricas *F-measure*, BLEU e METEOR, e qualitativamente, utilizando uma tradução de referência previamente estabelecida e as metáforas conceituais subjacentes às EI's analisadas. Os resultados mostraram que (1) há influência cultural na escolha dos alimentos que compõem as EI's; (2) a maior parte das EI's foi gerada por metáforas conceituais; (3) majoritariamente, as traduções elaboradas pelos TAO's não perceberam as EI's e suas metáforas conceituais, comprometendo sua compreensão; (4) os THP's, majoritariamente, acessaram a metáfora conceitual subjacente às EI's e produziram melhores traduções; e, finalmente, (5) tanto qualitativamente quanto quantitativamente, os TAO's obtiveram resultados aquém dos THP's. Considerando todas essas informações, verificou-se que, apesar de toda a evolução dos sistemas de tradução automática, estes ainda realizam poucas análises

semânticas, não reconhecendo as EI's nem na perspectiva clássica, como expressões congeladas, e muito menos na visão de vanguarda das metáforas conceituais, o que prejudica sensivelmente a compreensão da tradução automática de textos contendo EI's. A apropriação da Teoria da Metáfora Conceitual pela inteligência artificial pode provocar grandes avanços para a área de Tradução Automática.

Palavras-chave: Fraseologismo. Métricas de Avaliação de Tradução Automática. Tradução Semântica.

ABSTRACT

This research evaluated the perception of free online machine translators (TAO) in relation to idioms and their underlying conceptual metaphors. Since its inception, Machine Translation has generated great expectations regarding its quality, interspersing moments of great interest and investments with others of almost none. On the other hand, idioms (phraseologisms) and other figurative languages have increasingly aroused the attention of scholars, due to the systematic and frequent way in which they occur in all languages, within different discursive genres. Although quite common, they seem to be particularly problematic when translated by TAO. The analysis of idiom translations from the perspective of the Conceptual Metaphor Theory can contribute to the improvement of those tools. In this approach, idioms can be licensed by conceptual metaphors and, thus, shared by several languages. This would explain, for example, why many translations of idioms are also idioms. We evaluated 43 idioms with food terms in English and their respective translations in Portuguese, collected from two dictionaries of idioms. Initially, the expressions were analyzed to identify possible (a) cultural interferences in their compositions and (b) licensing by conceptual metaphors. Then, using the Google Books N-gram Viewer and the online concordance Webcorp, we checked their frequency of use and selected the excerpts, which were translated by three free TAO and three professional human translators (THP). All translations were quantitatively evaluated, using the metrics *F-measure*, BLEU and METEOR, and qualitatively, using a previously established reference translation and the conceptual metaphors underlying the idioms analyzed. The results showed that (1) there was cultural influence in the choice of the food that composed the idioms; (2) most idioms were licensed by conceptual metaphors; (3) generally, the TAO translations did not perceive the idioms and their conceptual metaphors, compromising their understanding; (4) the THP mostly accessed the conceptual metaphor underlying the idioms and produced better translations; and finally, (5) both qualitatively and quantitatively, the TAO obtained lower results than the THP. Considering all this information, we verified that, despite all the evolution of the TAO systems, they still do little semantic analysis, not recognizing the idioms, neither in the classical perspective as frozen expressions, nor in the vanguard view of conceptual metaphors, which significantly impairs the understanding of automatic translated texts

with idioms. The appropriation of the Conceptual Metaphor Theory by artificial intelligence may lead to great advances in the Machine Translation field.

Keywords: Phraseology. Metrics for Machine Translation Evaluation. Semantic Translation.

LISTA DE FIGURAS

Figura 1 –	Níveis de convencionalidade.....	54
Figura 2 –	<i>Continuum</i> entre literalidade e figuratividade.....	63
Figura 3 –	Avanços da Tradução Automática nas últimas décadas.....	68
Figura 4 –	O triângulo de Vauquois.....	72
Figura 5 –	<i>Precision e recall</i>	81
Figura 6 –	Exemplo da métrica METEOR.....	87
Figura 7 –	Interface do usuário do <i>Bing</i> Tradutor.....	95
Figura 8 –	<i>Bing</i> Tradutor com a tradução para <i>a piece of cake</i> em contexto.....	96
Figura 9 –	Página de comparação entre os dois tipos de tradução disponíveis no <i>Bing</i> Tradutor.....	97
Figura 10 –	Resultado da comparação entre as técnicas estatística e neural do <i>Bing</i>	98
Figura 11 –	Página inicial do tradutor de conversas do <i>Bing</i>	99
Figura 12 –	Interface do usuário do <i>Google Translate</i>	101
Figura 13 –	Interface do usuário do <i>Google Translate</i> com o dicionário de expressões aberto.....	102
Figura 14 –	<i>Google Translate</i> com a tradução para <i>a piece of cake</i> em contexto.....	103
Figura 15 –	<i>Google Translate</i> estatístico com a expressão <i>a piece of cake</i> em evidência.....	104
Figura 16 –	<i>Google Translate</i> para a tradução de ‘bolo’.....	104
Figura 17 –	<i>Google Translate</i> neural com a expressão <i>a piece of cake</i> em evidência.....	105
Figura 18 –	<i>Google Translate</i> para a palavra <i>cake</i>	106
Figura 19 –	<i>Google Translate</i> com <i>a piece of cake</i> selecionado.....	106
Figura 20 –	<i>Google Translate</i> com tradução para <i>a piece of cake</i>	107
Figura 21 –	Interface do usuário do SYSTRAN com instruções de uso....	109
Figura 22 –	SYSTRAN com a tradução para <i>a piece of cake</i> em contexto.	109
Figura 23 –	Dicionários disponibilizados pelo SYSTRAN.....	110
Figura 24 –	Resultado do dicionário do SYSTRAN para <i>a piece of cake</i> ...	111

Figura 25 – SYSTRAN neural com a tradução para <i>a piece of cake</i> em contexto.....	112
Figura 26 – Palavras em evidência na versão neural do SYSTRAN.....	113
Figura 27 – Participação do usuário na melhoria das traduções na versão neural do SYSTRAN.....	114
Figura 28 – Interface do usuário em outras línguas na versão neural do SYSTRAN.....	114
Figura 29 – Versão <i>demo</i> do SYSTRAN em francês.....	115
Figura 30 – Opções de configuração da versão neural <i>demo</i> do SYSTRAN.....	115
Figura 31 – <i>Print</i> do <i>Google Translate</i>	153
Figura 32 – <i>Print</i> do <i>Google Translate</i> após escolhida a melhor opção....	153

LISTAS DE QUADROS

Quadro 1 –	Objetivos secundários, questões e hipóteses da pesquisa.	33
Quadro 2 –	Graus de Idiomaticidade.....	61
Quadro 3 –	Fatos tipológicos que assemelham e diferem línguas.....	65
Quadro 4 –	Resumo comparativo de três técnicas de Tradução Automática.....	76
Quadro 5 –	Dez traduções para uma sentença em chinês - exemplo típico do conjunto de avaliações do NIST 2001.....	78
Quadro 6 –	Ferramentas de avaliação para anotadores humanos para julgamento de adequação e fluência.....	79
Quadro 7 –	Possibilidades de combinações de n-gramas no exemplo dado.....	85
Quadro 8 –	Línguas com suporte total da métrica METEOR.....	89
Quadro 9 –	Línguas com suporte parcial da métrica METEOR.....	90
Quadro 10 –	El's em língua inglesa e língua portuguesa e seus significados.....	118
Quadro 11 –	El's classificadas dentro do <i>continuum</i> de idiomaticidade..	120
Quadro 12 –	Metáforas conceituais subjacentes identificadas.....	136
Quadro 13 –	Traduções dos excertos pelos TAO's.....	143
Quadro 14 –	Traduções dos excertos pelos TAO's na segunda coleta....	149
Quadro 15 –	Tradução dos excertos pelos THP's.....	157

LISTA DE GRÁFICOS

Gráfico 1 –	Desempenho dos TAO's na elaboração das traduções da primeira coleta.....	147
Gráfico 2 –	Desempenho dos TAO's na elaboração das traduções na segunda coleta.....	156
Gráfico 3 –	Desempenho dos THP's na elaboração das traduções da única coleta.....	161
Gráfico 4 –	Resultados da métrica BLEU para os dois tipos de tradutores.....	164
Gráfico 5 –	Resultados da métrica <i>F-measure</i> para os dois tipos de tradutores.....	166
Gráfico 6 –	Resultados da métrica METEOR para os dois tipos de tradutores.....	168
Gráfico 7 –	Resultados da métrica BLEU para os TAO's na segunda coleta.....	171
Gráfico 8 –	Resultados da métrica <i>F-measure</i> para os TAO's na segunda coleta.....	172
Gráfico 9 –	Resultados da métrica METEOR para os TAO's na segunda coleta.....	173
Gráfico 10 –	Desempenho do TAO1 nas duas coletas para a métrica BLEU.....	174
Gráfico 11 –	Desempenho do TAO2 nas duas coletas para a métrica BLEU.....	175
Gráfico 12 –	Desempenho do TAO3 nas duas coletas para a métrica BLEU.....	176
Gráfico 13 –	Desempenho do TAO1 nas duas coletas para a métrica <i>F-measure</i>	177
Gráfico 14 –	Desempenho do TAO2 nas duas coletas para a métrica <i>F-measure</i>	178
Gráfico 15 –	Desempenho do TAO3 nas duas coletas para a métrica <i>F-measure</i>	179
Gráfico 16 –	Desempenho do TAO1 nas duas coletas para a métrica METEOR.....	180

**Gráfico 17 – Desempenho do TAO2 nas duas coletas para a métrica
METEOR.....**

LISTA DE TABELAS

Tabela 1 –	Desempenho dos TAO's na elaboração das traduções da primeira coleta.....	146
Tabela 2 –	Desempenho dos TAO's na elaboração das traduções da segunda coleta.....	154
Tabela 3 –	Resumo dos percentuais de desempenho dos TAO's na elaboração das traduções na segunda coleta.....	155
Tabela 4 –	Desempenho dos THP's na elaboração das traduções.....	160
Tabela 5 –	Resumo dos percentuais de desempenho dos THP's na elaboração das traduções.....	161
Tabela 6 –	Resumo dos percentuais de desempenho dos dois tipos de tradutores na elaboração das traduções.....	162
Tabela 7 –	Resultados da métrica BLEU para os TAO's e THP's.....	163
Tabela 8 –	Resultados da métrica BLEU para os TAO's.....	164
Tabela 9 –	Resultados da métrica BLEU para os THP's.....	165
Tabela 10 –	Resultados da métrica <i>F-measure</i> para os TAO's e THP's.....	165
Tabela 11 –	Resultados da métrica <i>F-measure</i> para os TAO's.....	166
Tabela 12 –	Resultados da métrica <i>F-measure</i> para os THP's.....	167
Tabela 13 –	Resultados da métrica METEOR para os TAO's e THP's.....	168
Tabela 14 –	Resultados da métrica METEOR para os TAO's.....	169
Tabela 15 –	Resultados da métrica METEOR para os THP's.....	169
Tabela 16 –	Resultados da métrica BLEU para os TAO's na segunda coleta.....	170
Tabela 17 –	Resultados da métrica <i>F-measure</i> para os TAO's na segunda coleta.....	171
Tabela 18 –	Resultados da métrica METEOR para os TAO's na segunda coleta.....	172
Tabela 19 –	Comparação dos resultados da métrica BLEU para o TAO1..	173
Tabela 20 –	Comparação dos resultados da métrica BLEU para o TAO2..	174
Tabela 21 –	Comparação dos resultados da métrica BLEU para o TAO3..	175
Tabela 22 –	Comparação dos resultados da métrica <i>F-measure</i> para o TAO1.....	177

Tabela 23 – Comparação dos resultados da métrica <i>F-measure</i> para o TAO2.....	178
Tabela 24 – Comparação dos resultados da métrica <i>F-measure</i> para o TAO3.....	179
Tabela 25 – Comparação dos resultados da métrica METEOR para o TAO1.....	179
Tabela 26 – Comparação dos resultados da métrica METEOR para o TAO2.....	180
Tabela 27 – Comparação dos resultados da métrica METEOR para o TAO3.....	181

LISTA DE ABREVIATURAS E SIGLAS

ALPAC	<i>Automatic Language Processing Advisory Committee</i>
API	<i>Application Programming Interface</i>
BLEU	<i>Bilingual Evaluation Understudy</i>
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
DDR&E	<i>Director of Defense Research and Engineering</i>
DTD	<i>Document Type Definition</i>
EI('s)	Expressão(ões) Idiomática(s)
IBM	<i>International Business Machines</i>
LA	Língua Alvo
LF	Língua Fonte
MIT	<i>Massachusetts Institute of Technology</i>
MOSES	<i>Open-source statistical MT engine</i>
NIST	<i>National Institute of Standards and Technology</i>
PLN	Processamento de Linguagem Natural
SMS	<i>Short Message Service</i> ou mensagem de texto
SYSTRAN	<i>System Analysis Translator</i>
TA	Tradução Automática
TAO('s)	Tradutor(es) Automático(s) <i>Online</i>
TER	<i>Translation Edit Rate</i>
THP('s)	Tradutor(es) Humano(s) Profissional(is)
TXT	<i>Text File</i> (extensão de nome de arquivo)
XML	<i>Extensible Markup Language</i> (extensão de nome de arquivo)

SUMÁRIO

1	INTRODUÇÃO – A HARD NUT TO CRACK.....	22
2	FUNDAMENTAÇÃO TEÓRICA – SEPARATING THE WHEAT FROM THE CHAFF.....	35
2.1	METÁFORAS E EXPRESSÕES IDIOMÁTICAS.....	35
2.1.1	Literal versus Metafórico.....	36
2.1.2	Metáforas.....	40
2.1.2.1	Teoria da Metáfora Conceitual.....	41
2.1.3	Fraseologismo.....	50
2.1.3.1	Expressão Convencional versus Expressão Idiomática.....	53
2.1.3.2	Níveis de Convencionalidade.....	53
2.1.3.3	(In)Decomponibilidade das Expressões Idiomáticas.....	55
2.1.3.4	Graus de Idiomaticidade.....	61
2.2	TRADUÇÃO AUTOMÁTICA.....	63
2.2.1	Histórico.....	69
2.2.2	Principais arquiteturas.....	72
2.2.3	Paradigmas de tradução.....	73
2.2.4	Avaliação de Tradução Automática.....	77
2.2.4.1	Métrica <i>F-Measure</i>	80
2.2.4.2	Métrica BLEU.....	83
2.2.4.3	Métrica METEOR.....	86
3	PERCURSO METODOLÓGICO – MAKING NO BONES ABOUT IT... 	91
3.1	TIPO DE PESQUISA.....	91
3.2	CONTEXTO DA PESQUISA.....	92
3.2.1	Tradutores Automáticos Online.....	92
3.2.1.1	<i>Bing</i> Tradutor.....	94
3.2.1.2	<i>Google Translate</i>	100
3.2.1.3	Tradutor SYSTRAN.....	107
3.3	PROCEDIMENTOS METODOLÓGICOS.....	116
3.3.1	Elaboração da lista e excertos com as EI's.....	117
3.3.2	Submissão do projeto de tese ao Comitê de Ética.....	124
3.4	DESENHO EXPERIMENTAL.....	125

3.4.1	Tradução pelos TAO's e pelos THP's.....	125
3.4.1.1	Traduções candidatas.....	126
3.4.1.2	Traduções de referência.....	127
3.4.1.3	Perfil dos THP's.....	127
3.4.2	Submissão às métricas de avaliação de Tradução Automática.....	129
3.5	DESCRIÇÃO DAS ANÁLISES.....	129
3.5.1	Identificação de aspectos culturais relacionados aos alimentos que compõem as EI's em estudo.....	129
3.5.2	Identificação das metáforas conceituais subjacentes às EI's.....	130
3.5.3	Avaliação das traduções das EI's produzidas pelos TAO's, considerando a análise das metáforas conceituais.....	130
3.5.4	Avaliação das traduções das EI's produzidas pelos THP's, considerando a análise das metáforas conceituais.....	131
3.5.5	Análise das métricas de avaliação de Tradução Automática.....	132
4	ANÁLISE E DISCUSSÃO – <i>HANDLING THE HOT POTATO</i>.....	133
4.1	IDENTIFICAÇÃO DE ASPECTOS CULTURAIS RELACIONADOS AOS ALIMENTOS QUE COMPÕEM AS EI'S EM ESTUDO.....	133
4.2	METÁFORAS CONCEITUAIS SUBJACENTES ÀS EXPRESSÕES IDIOMÁTICAS.....	135
4.3	AVALIAÇÃO DAS TRADUÇÕES DAS EI'S PRODUZIDAS PELOS TAO'S, CONSIDERANDO A ANÁLISE DAS METÁFORAS CONCEITUAIS.....	143
4.4	AVALIAÇÃO DAS TRADUÇÕES DAS EI'S PRODUZIDAS PELOS THP'S, CONSIDERANDO A ANÁLISE DAS METÁFORAS CONCEITUAIS.....	157
4.5	ANÁLISE DAS MÉTRICAS DE AVALIAÇÃO DE TRADUÇÃO AUTOMÁTICA.....	162
4.5.1	Métrica BLEU.....	162
4.5.2	Métrica <i>F-measure</i>.....	165
4.5.3	Métrica METEOR.....	167
4.5.4	Análise da evolução dos TAO's.....	170
5	CONCLUSÃO – <i>POURING OIL ON TROUBLED WATERS</i>.....	183
	REFERÊNCIAS.....	189

APÊNDICES	194
APÊNDICE A – APROVAÇÃO PELO COMITÊ DE ÉTICA EM PESQUISA DA UECE.....	195
APÊNDICE B – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO.....	196
APÊNDICE C – PERFIL DO TRADUTOR HUMANO.....	197
APÊNDICE D – EXCERTOS TRADUZIDOS COM INSTRUÇÕES PARA OS TRADUTORES HUMANOS.....	198
APÊNDICE E – TRADUÇÕES DE REFERÊNCIA.....	201
ANEXOS	203
ANEXO A – EI's COM ALIMENTOS EM PORTUGUÊS E SUAS RESPECTIVAS TRADUÇÕES EM INGLÊS.....	204
ANEXO B – EI's COM ALIMENTOS EM INGLÊS E SUAS RESPECTIVAS TRADUÇÕES EM PORTUGUÊS.....	209

1 INTRODUÇÃO – A HARD NUT TO CRACK



O surgimento da Tradução Automática, no início dos anos 50, gerou grandes expectativas em relação à qualidade por parte de governos e empresas, intercalando momentos de grande interesse e investimentos com outros de completo desinteresse. O Relatório da ALPAC¹ (1966) foi um dos motivos mais notórios para o cancelamento de investimentos nessa área na época. A causa está ligada, principalmente, ao fato de a Tradução Automática deixar a desejar no quesito qualidade se comparada à tradução elaborada por um tradutor humano.

Com os avanços tecnológicos cada vez maiores na área da Inteligência Artificial (IA) e sua presença em microcomputadores e aparelhos de comunicação móveis, tais como *tablets* e celulares, os tradutores automáticos *online* (doravante TAO's) têm se tornado uma ferramenta de uso fácil e corriqueiro. Em um mundo globalizado como o atual, não são poucas as situações em que eles podem ser aplicados. Citamos aqui somente algumas, como comunicação em viagens internacionais, leituras acadêmicas e de lazer, compras feitas pela *Internet*, entre outras. Os aplicativos que podem traduzir textos e imagens em línguas diferentes² são exemplos mais recentes de como a tecnologia trabalha para solucionar essas necessidades. Outro exemplo foi o lançamento, em 2016, de um aparelho auditivo que traduz automaticamente, quase que em tempo real, conversas entre duas pessoas que falam línguas diferentes³. Para este aparelho, ainda é necessário o uso do celular

¹ Disponível em: <<https://www.nap.edu/read/9547/chapter/1>>. Último acesso em: 02 jan. 2018.

² O *Google Translate App* permite que você utilize a câmera do seu celular para traduzir textos ao seu redor, tais como placas, anúncios ou anotações. Conf. Site <https://support.google.com/translate/answer/6142483?hl=en&ref_topic=7011659>. Acesso em: 04 abr. 2017.

³ O nome do invento é Pilot, da empresa Waverly Labs, que promete fazer traduções entre 2 línguas diferentes automaticamente em linguagem oral a partir de pequenos dispositivos colocados no ouvido de cada falante. Conf. Site <https://www.youtube.com/watch?v=NjjQ5cH_YzI>. Acesso em: 01 abr. 2017.

com o aplicativo, sendo mais prático, porém, pelo fato de o usuário ouvir a tradução diretamente em seu ouvido (com o dispositivo) e não ter que posicionar o aparelho celular entre as pessoas que o estão utilizando. Esses exemplos representam avanços para a Tradução Automática e, dependendo do objetivo da tradução que está sendo realizada, o produto tradutório pode ser aceitável, independentemente de sua acurácia. No vídeo de propaganda do referido aparelho, a tradução parece perfeita, mas é questionável se o resultado seria o mesmo para qualquer tipo de sentença, independente da estrutura, do léxico utilizados ou com uso de linguagem figurada, por exemplo.

A linguagem figurada está inserida no universo de qualquer língua. Comunicar-se através de metáforas, metonímias e expressões idiomáticas (doravante EI's) é algo bastante comum, e esse parece ser um aspecto particularmente desafiador de ser traduzido por tradutores automáticos. Quando se aprende uma nova língua, é inevitável deparar-se com esses fenômenos linguísticos que, em algumas ocasiões, parecem óbvios, porque podem ser semelhantes aos de nossa língua materna. Entretanto, em outras ocasiões, muitas vezes não é possível identificar o seu sentido. Na sentença seguinte, por exemplo: "*The child's toy was **as flat as a pancake**⁴ after the car drove over it*" (O brinquedo da criança ficou achatado feito uma panqueca depois que o carro passou sobre ele⁵), o aprendiz poderá fazer a associação entre o estado atual do brinquedo e uma panqueca, pois há certa transparência no sentido da EI. Contudo, em "*I **was cheesed off** that I would not be able to go away for the weekend*" (Eu fiquei chateado porque não conseguiria sair no final de semana), dificilmente o aprendiz associará a EI ao seu significado, uma vez que não há uma ligação óbvia entre o significado e as palavras que compõem essa EI.

Sobre o uso de EI's, Ortiz Alvarez (2006) afirma que:

...pelo fato de a língua estar intimamente ligada à cultura de cada povo, dificilmente um mesmo mecanismo de encadeamento de sentidos é usado em línguas diferentes e quando acontece, pode ser fonte de erro, às vezes até engraçados. (ORTIZ ALVAREZ, 2006, p. 159-160)

⁴ Como forma de identificar mais facilmente as EI, elas estarão sempre em negrito, em ambas as línguas.

⁵ Todas as traduções dos exemplos inseridos no texto foram feitas pela autora.

Percebe-se claramente a influência cultural na sentença “*The Math test was a piece of cake*” e a tradução para o português: “A prova de matemática foi fácil/**mamão com mel/mamão com açúcar/sopa no mel**”. Não convém traduzir como “A prova de matemática foi um pedaço de bolo”. “Cada língua tem a sua maneira de ver o mundo, e a partir desses olhares, brinca com as figuras de linguagem” (ORTIZ ALVAREZ, 2006, p. 160).

Acredita-se que a influência cultural tem um papel fundamental na criação das EI's e isso pode explicar o fato de estas, apesar de possuírem metáforas conceituais subjacentes semelhantes, se utilizam de esquemas imagéticos diferentes se compararmos línguas variadas. Nos exemplos citados, tanto a EI ***a piece of cake*** quanto **mamão com mel/açúcar** e **sopa no mel** expressam a ideia de ‘facilidade’, representada por alimentos diferentes, por cada língua (o português e o inglês). É interessante, contudo, que ambas as línguas se utilizam de alimentos doces para expressar facilidade, pois o açúcar parece estar relacionado à sensação de prazer ao paladar.

Tseng (2017), expressa a importância de se conhecer essas diferenças culturais quando lidamos com expressões fixas e idiomáticas, pois, apesar da metáfora subjacente ser a mesma, as variações existem devido às diferenças culturais ou diferentes concepções de alimentos que cada cultura tem, o que corrobora para a justificativa do presente estudo.

Diversos trabalhos de pesquisa com Linguística de *Corpus* também têm evidenciado a importância das EI's na comunicação. A partir desta área de estudo, podemos ter acesso a *corpora* cada vez maiores e mais específicos, o que facilita a identificação e análise das EI's. Diversas ferramentas digitais têm sido desenvolvidas para a análise desses *corpora*, dentre as quais está o *Webcorp*, um concordanceador que utiliza a *Web* como *corpus*.

A partir do que foi exposto e com o propósito de verificar a funcionalidade da Tradução Automática, para o aspecto das EI's da língua, foi realizado um pequeno teste utilizando a EI ***a piece of cake*** e três TAO's de acesso livre, escolhidos dentre os vários que estão disponíveis na *Web*. A mesma sentença foi traduzida pelos três TAO's e obtivemos resultados diferentes:

Texto original: *Mary told me her Math test was a piece of cake. She'll get an A.*

Tradutor A: Mary me contou sua matemática teste foi um pedaço de bolo. Ela vai ter um.

Tradutor B: Mary me disse que seu teste de Matemática foi um pedaço de bolo. Ela vai ter um A.

Tradutor C: Mary disse-me que seu teste da matemática era um pedaço de bolo. Obterá um A.

Ao ler as sentenças, percebe-se facilmente que as traduções obtidas são de qualidade questionável, principalmente porque nenhum dos três sistemas de tradução automática *online* identificou que na sentença da língua de partida, o trecho **a piece of cake** estava sendo utilizado como EI, em sentido figurado, interferindo na compreensão do significado pretendido pelo texto da língua de chegada.

Além deste problema, ressalte-se que há outros problemas nas traduções produzidas pelos três TAO utilizados:

- a) Na tradução realizada pelo Tradutor A, a ordem de algumas palavras não está apropriada e houve a omissão de resultados para 'A' em *She'll get an A*;
- b) A linguagem utilizada pelo Tradutor C parece ser muito formal. Além disso, o trecho 'da matemática' soa estranho para o falante do português do Brasil.

Entretanto, os problemas acima citados, não interferem na compreensão do texto como a tradução inapropriada da EI **a piece of cake**. Ao finalizar o teste, imagina-se a possibilidade de ter sido um resultado isolado ou algo que se repetiria para as EI's, de maneira geral, quando traduzidas por tradutores eletrônicos.

Existe um universo imenso de EI, o qual será explicitado mais adiante. Portanto, para o desenvolvimento da atual pesquisa, escolheu-se analisar especificamente, as EI's que contêm um alimento em sua estrutura lexical. O alimento faz parte do dia a dia da população e está intimamente ligado à sobrevivência da humanidade. A coocorrência dos alimentos com outros fatores significativos, certamente, é uma fonte potencial de metáforas e expressões. Além disso, alimentos envolvem aspectos físicos e culturais, nas diferentes formas que são utilizados e nas diversas maneiras que são manuseados e preparados. Acredita-se que, apesar de a conceitualização ser universal, os alimentos variam entre as culturas, podendo gerar expressões metafóricas diferentes para a mesma experiência corpórea.

O foco da pesquisa é, portanto, a análise do processo de tradução de EI's, com um alimento em sua composição lexical, realizado por TAO's, à luz da teoria das metáforas conceituais. Para o leitor iniciante no assunto, a Teoria da Metáfora Conceitual preconiza que o sistema conceitual é fundamentalmente metafórico e que as metáforas não se encontram na linguagem, mas sim no pensamento inconsciente. Na seção 2, o tema será visto com mais detalhes.

A presente pesquisa é motivada pela importância dos diferentes elementos que a compõem: EI's, metáforas conceituais e Tradução Automática. A seguir, serão descritos quais aspectos de cada um desses elementos justifica sua elaboração.

As EI's são compostas por diferentes itens lexicais e estes pertencem a vários domínios, como partes do corpo humano (**dos pés à cabeça**), animais (**a vaca foi pro brejo**) ou alimentos (**o caldo entornou**), entre outros.

Tradicionalmente, as EI's são consideradas expressões fixas, cujos significados não podem ser inferidos pela análise das palavras que as compõem. Alguns autores, inclusive, consideram as EI's como metáforas mortas, cujo interesse por suas origens foi se perdendo. De fato, por um determinado tempo, elas foram consideradas quase como 'acessórios' para as línguas e, portanto, não importantes de serem estudadas. Contudo, essa visão foi se modificando nos últimos anos e elas passaram a ser de interesse de estudiosos da linguagem, por meio dos estudos da fraseologia. A fraseologia "... ocupa-se das combinações estáveis de unidades léxicas, constituídas por mais de duas palavras gráficas."⁶ Segundo Ellis (2008), em uma pesquisa ProQuest-CSA, em publicações sobre Ciências Sociais nos últimos 50 anos, identificou-se que o número de trabalhos relacionados à fraseologia eram apenas 53 nos anos 60 e 199 nos anos 70. Nos anos 80, no entanto, o número aumentou para 529 e para 709 nos anos 90. A razão para esse interesse ter mudado está no argumento de que os mecanismos gramaticais utilizados para descrever os fenômenos da periferia gramatical (elementos e construções marcados que não são muito conhecidos), também poderiam dar conta dos fenômenos principais (aqueles que podem ser descritos pelos princípios parametrizados da Gramática Universal). Houve também uma mudança na nomenclatura das EI's. Nos anos 50, segundo Ellis

⁶ E-Dicionário de termos literários. Carlos Ceia. Disponível em: <<http://edtl.fcsh.unl.pt/business-directory/6197/fraseologia/>>. Acesso em: 23 jan. 2018.

(2008, p. 2), elas eram conhecidas como padrões estruturais, hoje 'atendem' pelo nome de 'construções' ou 'fraseologismos'.

Novas teorias sobre a compreensão da metáfora trouxeram luz aos estudos sobre EI's, encorajando diferentes perspectivas que enfatizam mais sua essência cognitiva do que sua origem semântica. Estudiosos como Lakoff e Johnson (1980), Grady (1997), Gibbs (1993) e Giora (1997), dentre outros, trouxeram *insights* importantes sobre os mecanismos de compreensão das EI's (LIU; SHEN, 2013). Transparência, decomponibilidade, saliência e convencionalidade são aspectos que têm um papel importante para determinar a compreensão das EI's. Familiaridade é outro aspecto que está diretamente ligado à frequência de uso desse tipo de linguagem.

Sendo construções comuns e seu conhecimento essencial para a compreensão de qualquer língua, as EI's necessitam de análise em seus diferentes contextos. Um deles é o da tradução. De que maneira essas unidades lexicais são tratadas por TAO's, é o principal campo de investigação deste estudo. Por meio dos achados pesquisados, pode-se sugerir a alimentação de dados que facilitem a identificação das EI's, ora em evidência.

Um aspecto importante a considerar, é que existe um grande número de EI's em todas as línguas e que uma grande parte dessas expressões pode ser traduzida de uma língua para a outra por meio de uma EI também. Ora, se as EI's são estruturas fixas, mortas e congeladas, como apregoado pela visão clássica, como se explica que as mais diversas línguas utilizem também uma EI para os mesmos contextos discursivos pragmáticos e muitas destas não se configurem exatamente como empréstimos ou decalques uma das outras (LIMA, 2017).

O segundo elemento do estudo são as metáforas conceituais. Considera-se relevante estudar a natureza das metáforas que licenciam as EI's, pois este pode ser um ponto fundamental para entender o fenômeno mencionado no parágrafo anterior e, portanto, para a construção do sistema de reconhecimento dessas expressões pela máquina. De acordo com a Teoria da Metáfora Conceitual, diferentemente da visão clássica, a metáfora não é meramente um recurso estilístico, mas parte natural da forma como pensamos, racionamos e, portanto, nos expressamos (LAKOFF; JOHNSON, 1980, 1999). A referida teoria parte da observação de que a metáfora é utilizada com grande frequência na linguagem como um todo (linguagem cotidiana, científica nas diversas áreas do conhecimento,

jornalística, técnica, dentre outras) e de forma sistêmica. Por exemplo, frases como “Eu não consigo engolir esse conceito” e “Ainda estou digerindo essa nova abordagem”, ou ainda, “Ele ainda está ruminando aquela ideia” indicam que conceitualizamos ‘ideias’ como ‘alimentos’. Essa forma de expressar o que pensamos, não é arbitrária e conscientemente construída no momento da elocução, mas naturalmente licenciada pela metáfora conceitual IDEIAS SÃO ALIMENTOS (LIMA, 2005).

Berrada (2007, p. 8), defende a necessidade de haver mais estudos interlinguísticos para averiguar a universalidade da metáfora conceitual IDEIAS SÃO ALIMENTOS, ou seja, questionar se esta metáfora realmente existe em diferentes culturas, ou se é fruto de empréstimos entre diferentes línguas. Como suposição para seu estudo, a autora afirma que

...devido à proeminente importância dos alimentos em nossa vida como fonte de sustento e prazer, é suscetível que o alimento seja usado de forma generalizada em várias culturas e línguas relacionadas ou não, como domínio-fonte mapeando ideias, virtude, conduta, disposições humanas, desejo sexual e uma série de outros domínios-alvo menos delineados.⁷ (BERRADA, 2007, p. 8, tradução nossa).

Ao analisar El’s que contêm um alimento em sua composição lexical nesta pesquisa, julga-se estar contribuindo para os estudos interlinguísticos supracitados, envolvendo as línguas inglesa e portuguesa.

Nota-se, no entanto, que como explica Lima (2017), sendo a metáfora um mecanismo conceitual natural, subjacente à forma como nos expressamos, não é de se estranhar que as línguas possam ter estruturas semelhantes, inclusive de empréstimos e decalques, para comunicar as mesmas ideias. Além disso, considerando-se que a metáfora é gerada, segundo a teoria, por mecanismos corpóreos experienciais, também não é de se estranhar que elementos culturais estejam envolvidos no uso das metáforas em si e das expressões metafóricas licenciadas por elas.

Spulber (2008), exemplifica os diferentes aspectos da vida social e cultural que são expressados em termos de alimentos: família, religião, sexo, gênero, posição

⁷ ... owing to the prominent importance of food in our life as a source of sustenance and pleasure, food is liable to be pervasively used in various related or unrelated cultures and languages as a source domain mapping ideas, virtue, conduct, human dispositions, sexual desire, and a host of other less delineated target domains.

social e identidade grupal, dentre outros. Nesse sentido, a metáfora conceitual é um importante mecanismo para a compreensão da linguagem e, portanto, deve ser contemplada por qualquer ferramenta que tem a linguagem como seu objeto, incluindo a Tradução Automática, o terceiro elemento desta pesquisa.

Apesar do desenvolvimenro de vários estudos sobre metáforas conceituais, alimentos e El's, percebe-se que nenhum desses estudos se debruçou, especificamente, sobre as El's compostas por alimentos, nem tampouco, como suas traduções são realizadas por tradutores automáticos. Dentre a bibliografia estudada, temos Tseng (2017), que analisa a metáfora conceitual PENSAMENTO É ALIMENTO por meio do ato real da experiência de comer no evento de *design* de alimentos intitulado EATAIPEI, no *London Design Festival* em 2015. Tseng analisou como as metáforas multimodais de alimentos foram criativamente representadas e elaboradas dentro do festival, mais especificamente tratando-as como metáforas complexas que resultam das interações de metáforas primárias, esquemas proposicionais e performance intercultural, adaptando para públicos internacionais conhecimento compartilhado ou cultural-específico sobre alimentos. Abreu (2010), estudou a motivação cognitiva para o significado das expressões idiomáticas, na tentativa de apresentar estratégias mais eficazes para sua compreensão, quanto à possibilidade das partes constituintes das expressões se relacionarem conceptualmente ao significado idiomático e poderem ser mapeadas cognitivamente com base na nossa experiência de mundo. O estudo confirmou que, as partes literais constituintes das expressões idiomáticas carregam parte do significado idiomático e podem ser mapeadas cognitivamente e sugere que a aprendizagem de expressões pode ser favorecida através da motivação das relações entre os significados literais e os figurados. Tam (2009), analisou o conjunto de correspondências dos mapeamentos IDEIA-COMO-ALIMENTO e FAZER SEXO-É-COMER, em expressões linguísticas nas línguas inglesa e vietnamita. Berrada (2007), demonstra em estudo, a penetração da metáfora com alimentos em interações na língua árabe marroquina e sua exploração limitada no discurso do Alcorão do árabe clássico.

Os tradutores eletrônicos são o foco da presente pesquisa, pois trata-se de ferramentas de uso cada dia mais frequente, sendo utilizadas por usuários de diferentes níveis de conhecimento e com os mais variados objetivos, conforme mencionado anteriormente, desde uma simples tradução de letras de uma música, até os resumos de trabalhos acadêmicos. Apesar de os TAO's já terem evoluído muito

desde sua invenção, Monteiro (2002), encontrou resultados positivos para o uso da Tradução Automática na leitura e compreensão de textos por alunos de mestrado de diferentes áreas, pontuando que o objetivo do usuário pode exigir um nível de acurácia mais alto ou não – eles continuam sendo criticados em relação à qualidade dos resultados obtidos. Portanto, um banco de dados contendo as metáforas que licenciam as EI's analisadas, pode alimentar os sistemas de Tradução Automática, trazendo maior acurácia às suas traduções.

O *MetaNet*⁸ é um exemplo de banco de dados e frames de metáforas, atualmente hospedado no Instituto Internacional de Ciências da Computação da Universidade de Berkeley na Califórnia, que pode ser acessado a partir do endereço eletrônico: <https://metaphor.icsi.berkeley.edu/pub/en/index.php/Category:Metaphor>. Analisando-o, encontram-se listadas as seguintes metáforas dentro do campo semântico de alimentos (tradução própria)⁹:

- a) CONSUMIR MUITOS RECURSOS É COMER MUITA COMIDA,
- b) IDEIAS PERTURBADORAS SÃO ALIMENTOS REPUGNANTES,
- c) IDEIAS SÃO ALIMENTOS,
- d) INTERESSE EM IDEIAS É APETITE POR ALIMENTO,
- e) IDEIAS INTERESSANTES SÃO ALIMENTOS APETITOSOS,
- f) PREPARAR IDEIAS PARA SEREM COMPREENDIDAS É PREPARAR ALIMENTOS,
- g) APRENDER É INGERIR ALIMENTO NUTRITIVO,
- h) ADQUIRIR IDEIAS É COMER,
- i) CONSUMIR RECURSOS É COMER,
- j) GANÂNCIA É FOME,
- k) DESEJAR É TER FOME, e
- l) RACIOCINAR É COMER.

As metáforas identificadas e analisadas nesta pesquisa, podem suprir este banco de dados. Além disso, a pesquisa pretende contribuir para os estudos da

⁸ Disponível em: <<https://metanet.icsi.berkeley.edu/metanet/>>. Acesso em: 08 fev. 2017.

⁹ *CONSUMING TOO MANY RESOURCES IS EATING TOO MUCH FOOD; DISTURBING IDEAS ARE DISGUSTING FOODS; IDEAS ARE FOOD; INTEREST IN IDEAS IS APPETITE FOR FOOD; INTERESTING IDEAS ARE APPETIZING FOODS; LEARNING IS GETTING NUTRITIOUS FOOD; PREPARING IDEAS TO BE UNDERSTOOD IS FOOD PREPARATION; LEARNING IS GETTING NUTRITIOUS FOOD; ACQUIRING IDEAS IS EATING; CONSUMING RESOURCES IS EATING; GREED IS HUNGER; DESIRE IS HUNGER; and COGNIZING IS EATING.*

relação entre linguagem e pensamento em si, ou seja, das metáforas conceituais e suas realizações linguísticas.

No escopo da Tradução Automática de EI's, pode-se justificar que, através de uma análise preliminar da presente pesquisa, utilizando apenas o *Google Translate*, evidenciou que a tradução da grande maioria das EI's avaliadas, do português para o inglês e do inglês para o português, foi realizada palavra-por-palavra, em seu sentido NÃO metafórico.

Por outro lado, sabe-se que as tecnologias digitais são atualizadas e aperfeiçoadas em uma dinâmica rápida. Pesquisas sincrônicas, como a presente, são primordiais para que se possa registrar dados e resultados obtidos para um determinado período, de forma que estes sejam comparados aos anteriores (se houver) e que também sirvam de referência para os que serão levantados no futuro. Durante o processo de elaboração da pesquisa aqui relatada, verificou-se mudanças significativas nos recursos digitais utilizados, tanto para os TAO's, quanto para o *Webcorp*.

Apresenta-se como objetivo geral para esta pesquisa, a avaliação da percepção dos TAO's, as EI's e suas metáforas subjacentes. Para o estudo, optou-se avaliar as EI's que contêm um alimento em sua composição, acreditando-se porém, que quaisquer outras EI's poderiam ser utilizadas. As línguas analisadas foram o inglês e o português do Brasil.

Com o propósito de alcançar o objetivo principal, foram traçados os objetivos secundários. Para cada um desses objetivos secundários foram elaboradas perguntas e hipóteses relacionadas:

- a) O primeiro objetivo secundário está relacionado à questão da possibilidade de interferência cultural na composição das EI's. Portanto, é necessário identificar se há aspectos culturais envolvidos no tipo de alimento que as compõem. Apenas verificando as EI's e seus correspondentes nas duas línguas pesquisadas, pode-se perceber que há diferenças, conforme mencionado anteriormente. A hipótese para este objetivo é que a cultura tenha um papel importante na composição das EI's.
- b) O segundo objetivo secundário diz respeito às metáforas conceituais subjacentes às EI's. Faz-se necessário que sejam identificadas, para que seja possível verificar, se os TAO's as reconhecem quando

elaboram as traduções. A literatura da área defende que EI's são metáforas convencionalizadas. Portanto, a hipótese é de que seja possível identificar as metáforas conceituais da maioria das EI's analisadas.

- c) O terceiro objetivo secundário é o cerne da pesquisa, pois faz-se necessário produzir as traduções elaboradas pelos TAO's e analisá-las, de forma que possamos identificar de que maneira os TAO's percebem as EI's e suas metáforas conceituais. Levando em consideração a análise preliminar, têm-se como hipótese que os TAO's não percebem as EI's e suas metáforas correspondentes, e que na maioria das vezes, elaboram traduções não metafóricas.
- d) O quarto objetivo secundário está diretamente relacionado ao anterior. Considerando que a literatura na área defende que a tradução elaborada por humanos é o parâmetro de avaliação para as traduções automáticas, faz-se necessário avaliar também traduções produzidas por tradutores humanos profissionais (doravante THP's), a fim de serem comparadas àquelas produzidas pelos TAO's, quanto à sua percepção às EI's e metáforas conceituais subjacentes. Dessa forma, a hipótese é de que os THP's reconhecerão as EI's e suas metáforas conceituais e farão suas traduções utilizando linguagem metafórica, mas não necessariamente, utilizarão uma EI com alimento.
- e) O quinto e último objetivo secundário visa estabelecer uma comparação entre os resultados das traduções realizadas pelos TAO's e THP's, utilizando métricas de avaliação de Tradução Automática conhecidas e também uma métrica semântica. Entende-se como métrica semântica, a análise manual, de elaboração própria, na qual o aspecto lexical e os domínios conceituais metafóricos são levados em consideração. Uma tradução de referência elaborada foi utilizada como parâmetro para tal. Essa comparação é importante para que se possa verificar qual tipo de tradutor obtém melhores resultados nas métricas escolhidas. A hipótese é de que os THP's obterão melhores resultados que os TAO's, apesar do conhecimento de que as métricas de avaliação de Tradução Automática, em geral, penalizam as variações lexicais que são comuns de serem realizadas por tradutores humanos.

A fim de apresentar os objetivos secundários, as questões e hipóteses de pesquisa de uma forma mais clara, apresenta-se um resumo no Quadro 1:

Quadro 1 – Objetivos secundários, questões e hipóteses da pesquisa

Item	Objetivos Secundários	Questões de Pesquisa	Hipóteses
a)	Identificar se há aspectos culturais envolvidos no tipo de alimento que compõe as EI's em estudo.	Há interferência cultural no tipo de alimento que compõe as EI's em estudo?	A cultura tem papel importante na composição das EI's em estudo.
b)	Verificar se há metáforas conceituais subjacentes a todas às EI's em estudo.	As EI's em estudo são licenciadas por metáforas conceituais?	A maioria das EI's analisadas é licenciada por metáforas conceituais.
c)	Avaliar as traduções de EI's com alimentos, produzidas por TAO's.	Quão perceptíveis são os sistemas de tradução automática <i>online</i> às EI's e suas metáforas conceituais subjacentes?	Os sistemas de tradução automática <i>online</i> não percebem as EI's nem as suas metáforas conceituais subjacentes, e farão, na maioria das vezes, as traduções de forma não metafórica.
d)	Avaliar traduções de EI's com alimentos, produzidas por THP's, via processo não supervisionado.	Quão perceptíveis são os THP's às EI's e suas metáforas conceituais subjacentes?	Os THP's reconhecerão as EI's e farão suas traduções utilizando linguagem metafórica, mas não necessariamente utilizando uma EI com alimento.
e)	Comparar as traduções realizadas pelos TAO's e THP's com uma tradução de referência estabelecida, utilizando métricas de avaliação de Tradução Automática e uma métrica semântica própria, ou seja, uma métrica que leva em consideração as metáforas conceituais subjacentes.	Qual tradução obterá melhor resultado na avaliação realizada pelas métricas quando comparada à tradução de referência estabelecida?	Ao ser comparada com a tradução de referência estabelecida, a maioria das traduções produzidas pelos TAO's obterá resultados mais baixos que as traduções elaboradas pelos THP's.

Fonte: Elaborado pela autora.

Em geral, a avaliação de traduções automáticas é realizada utilizando como parâmetro traduções produzidas por humanos. A mesma linha foi seguida, com uma diferença: a elaboração de uma tradução em português, que foi utilizada como a tradução de referência pelas métricas de avaliação de tradução, escolhidas para esta pesquisa, sendo solicitados THP's que traduzissem os mesmos excertos submetidos aos TAO's, para verificar a percepção dos profissionais em relação às EI's e suas metáforas subjacentes, uma vez que esta não é uma informação dada. Portanto, apesar de, neste estudo, o foco ser os TAO's, fez-se imperativo avaliar traduções

realizadas tanto pelos TAO's quanto pelos THP's. A fim de quantificar o desempenho de ambos os tipos de tradutores, aplicou-se métricas automáticas.

Para esta pesquisa, tomou-se como modelo algumas das métricas utilizadas por Cunha (2013) na análise de sintagmas nominais com o caso genitivo (BLEU e *F-measure*) e acrescentou-se outra que aceita algumas diferenças lexicais (METEOR). Entretanto, como nenhuma dessas métricas parece avaliar a questão semântica, utilizou-se uma métrica semântica própria, que leva em consideração a metáfora conceitual subjacente às EI's, como um parâmetro de avaliação do sucesso da tradução.

A presente pesquisa expõe a literatura sobre EI's e metáforas, bem como sobre Tradução Automática e avaliação de sistemas de Tradução Automática. A aplicação da metodologia tem um caráter exploratório com procedimentos experimentais. Foram selecionadas 43 EI's, depois reduzidas para 31, que, tanto em inglês quanto na sua forma correspondente em português, são compostas por um léxico de alimento. Excertos, em língua inglesa, ilustrativos do uso das EI's, foram retirados do *Webcorp* e submetidos aos TAO's e THP's, para que o produto tradutório em língua portuguesa fosse analisado.

Esta pesquisa está dividida em cinco seções, iniciando pela introdução. Na seção de número dois, apresentam-se os estudos e a fundamentação teórica que a embasaram. A metodologia, ferramentas utilizadas para a coleta de dados, a construção dos *corpora* para as análises e todos os passos do processo de desenvolvimento da pesquisa, são apresentados na seção de número três, seguidas pelas análises e discussão na seção de número quatro. A seção cinco traz as conclusões e sugestões de trabalhos futuros. Por fim, apresentam-se as referências bibliográficas, os apêndices e os anexos.

2 FUNDAMENTAÇÃO TEÓRICA – *SEPARATING THE WHEAT FROM THE CHAFF*



Nesta seção, apresentamos a fundamentação teórica, norteadora da pesquisa, com a revisão literária das áreas que a fundamentam. Suas subseções estão divididas nas duas grandes áreas envolvidas, Linguística Cognitiva e Linguística Computacional.

Na primeira subseção (2.1), abordamos a relação entre metáforas e El's, estabelecendo o que consideramos como sentido literal e figurado; falamos sobre o tipo de metáfora estudado; explanamos, mais especificamente, a Teoria da Metáfora Conceitual, em suas diferentes versões, o fraseologismo, a (in) decomponibilidade das El's, a diferença entre expressão convencional e idiomática, bem como, os níveis de convencionalidade e os graus de idiomaticidade. Na segunda, subseção 2.2, apresentamos informações gerais sobre Tradução Automática; pontos importantes de seu histórico; as principais arquiteturas e paradigmas de tradução; e a avaliação de tradução, com descrição das métricas de avaliação escolhidas para este estudo (*F-Measure*, BLEU e METEOR).

2.1 METÁFORAS E EXPRESSÕES IDIOMÁTICAS

Metáforas e El's fazem parte de uma variedade de estruturas linguísticas, denominadas linguagem figurada. Sobre a linguagem figurada, Gibbs e Colston (2012) iniciam o Capítulo 6 de seu livro, indagando aos leitores, de maneira geral, se entendem o rico conjunto de significados metafóricos, que um determinado artigo de jornal contém. Os autores defendem, que interpretar significado figurado, não exige um tipo único de processo psicológico, que produz os mesmos significados, automaticamente, da mesma forma e em todos os contextos e que a gama de

significados inferidos, quando as pessoas se deparam com linguagem figurada e a quantidade necessária de esforço para extrair essas inferências depende de uma variedade de fatores, que interagem entre si.

Os autores destacam quatro fatores que, segundo eles, e baseados em experimentos de diversos outros pesquisadores, moldam a compreensão da linguagem figurada. São eles: 1) as pessoas; 2) os materiais linguísticos; 3) o objetivo ou a tarefa de compreensão; e 4) os métodos usados para avaliar o desenvolvimento figurativo que, por sua vez, contêm vários subitens.

O uso de linguagem figurada é, talvez, tão antigo quanto o uso da própria linguagem. Metáforas, metonímias e símiles, dentre outras, são tradicionalmente vistas como artifícios para tornar a comunicação mais rica e estão presentes no falar cotidiano das pessoas, independente do local que vivam ou da língua que falem. Contudo, estas estão tão presentes em nosso pensamento que o que antes era considerado literal ou metafórico mudou a partir de estudos sob a perspectiva da Teoria da Metáfora Conceitual.

2.1.1 Literal versus Metafórico

Definir o que é literal é uma tarefa difícil. Lakoff (1986, p. 296) finaliza seu artigo, *The Meanings of Literal*, afirmando que assim como a ciência tornou o uso da palavra 'mãe' mais difícil, assim também é o uso da palavra 'literal'. Em um mundo e em um modelo de linguagem idealizados e super-simplificados, ambas as definições seriam fáceis de serem concebidas, mas “num mundo onde existem bebês de proveta, barrigas de aluguel, adoção por casais homoafetivos, assim como caprichos da vida moderna, os critérios utilizados para definir 'mãe' podem divergir drasticamente” (1986, p 291). Em relação ao uso de 'literal', a ciência a qual o autor se refere é a Teoria da Metáfora Conceitual, que modificou a forma de descrever o que são linguagens literal e metafórica.

Após apresentar quatro conceitos diferentes e, comumente aceitos, para definir o que é 'literal', Lakoff explica e, posteriormente, refuta três deles, bem como algumas suposições, que são desdobramentos dessas definições. Vejamos os quatro conceitos e as suposições às quais o autor se refere: (1986, p. 292)

Literal 1 ou literalidade convencional – a linguagem convencional comum, ou seja, a linguagem não poética, sem uso de linguagem figurada como, por exemplo, hipérbole, aproximação, embelezamento, polidez excessiva, desvio e outros.

Literal 2 ou literalidade de assunto – a linguagem que é comumente usada para falar sobre um determinado domínio de assunto.

Literal 3 ou literalidade não metafórica – a linguagem, que é diretamente significativa, ou aquela que é compreendida, mesmo que parcialmente, em termos de outro significado.

Literal 4 ou literalidade de verdade condicional – a linguagem que é capaz de ‘se ajustar ao mundo’, que se refere a coisas objetivamente existentes ou a linguagem, que pode ser objetivamente verdadeira ou falsa.

O autor continua (1986, p. 292) esclarecendo, que esses quatro conceitos não são os únicos, mas são relevantes para discussões teóricas sobre metáforas. Também afirma que, tanto a semântica linguística quanto a lógica filosófica, aceitam que eles convergem entre si. As suposições que são decorrentes dessa convergência são:

- a) Literal 1 = Literal 3 (a linguagem convencional comum é diretamente significativa e, portanto, não metafórica)
- b) Literal 1 = Literal 4 (toda linguagem convencional comum é capaz de se referir à realidade objetiva, ou de ser objetivamente verdadeira ou falsa)
- c) Literal 2 = Literal 4 (há apenas uma forma correta e objetiva de se compreender cada assunto; dessa forma, a linguagem convencional, usada para falar sobre um assunto, é capaz de ser falsa ou verdadeira)

Lakoff (1986, p. 293), contudo, refuta os sentidos Literal 1, 2 e 4 e as suposições a), b) e c), acima citadas, e reconhece que, ao fazê-lo, causa mal-estar e confusão, pois questiona visões de senso comum usuais sobre a linguagem. O autor se queixa da inexistência em inglês de uma palavra que possa significar cada um dos quatro usos de ‘literal’ e sugere outros termos para substituir Literal 1 (linguagem convencional ou cotidiana) e Literal 3 (linguagem diretamente significativa ou linguagem não metafórica). Mas diz não ter encontrado outras formas de descrever, sem rodeios, Literal 2 e 4 (1986, p. 293).

A fim de explicar o porquê de refutar os sentidos de Literal 1, 2 e 4, Lakoff (1986) apresenta o que ele e Mark L. Johnson consideram metafórico, em sua teoria

exposta no livro *Metaphors We Live By*, e confessa que se tivesse feito essa distinção, na época em que o livro foi publicado, teria evitado muita confusão:

Metáforas são primariamente formas de pensar sobre algo, e dessa forma, elas podem ser convencionais ou novas. A linguagem metafórica surge quando uma palavra para algo no domínio fonte é também usada para o elemento correspondente do domínio alvo.¹⁰ (LAKOFF, 1986, p. 294, tradução nossa)

Se uma metáfora é convencional, então a linguagem, que expressa essa figura, não pode ser Literal 3, que é definido como diretamente significante e não metafórico. Sendo assim, considerada Literal 1, por ser convencional. Dessa forma, a suposição a) Literal 1 = Literal 3 se torna falsa.

Vejamos melhor através do exemplo dado pelo autor. É convencional compreender a morte por meio das metáforas A VIDA ESTÁ LOCALIZADA AQUI, enquanto que A MORTE É IR PARA UM LOCAL LONGE DAQUI. As expressões metafóricas ‘Ele partiu’, ‘Ele ainda está entre nós’ ou ‘Ele fez a passagem’ são formas convencionais de se falar sobre a morte. Portanto, se é convencional, Literal 1. Mas, se considerarmos que a palavra ‘partir’ tem seu sentido primário, no domínio espacial, e seu sentido metafórico, no domínio vida-e-morte, logo não será Literal 3. Dessa forma, a suposição a) torna-se falsa, pois Literal 1, neste caso, não pode ser igual a Literal 3.

Continuando o mesmo raciocínio, se ‘Ele partiu’ é uma forma convencional de se falar sobre o assunto ‘morte’, sendo também Literal 2. Dado seu significado metafórico como ‘Ele está morto’, a expressão pode ser falsa ou verdadeira, e, portanto, Literal 4.

Um segundo exemplo é utilizado por Lakoff para refutar as suposições a), b) e c) – a forma como compreendemos e raciocinamos sobre eletricidade: como fluxo de fluidos e movimento de multidões de Gentner e Gentner (1982).

As sentenças ‘Um resistor impedirá o fluxo de eletricidade’ e ‘Um resistor é um portão estreito, que só deixará passar um certo número de elétrons por vez, e quando há uma grande quantidade de elétrons, eles ficam armazenados no portão’ são formas convencionais de se pensar e falar sobre eletricidade. Se convencionais,

¹⁰ *Metaphors are primarily ways of thinking about something, and as such they can be conventional or novel. Metaphorical language arises when a word for something in the source domain is also used for the corresponding element of the target domain.*

são Literal 1. Há, contudo, duas metáforas envolvidas (de fluidos e de multidões) e, portanto, se são metafóricas, não serão Literal 3. Novamente, refuta-se a suposição a) Literal 1 = Literal 3. Em se tratando de formas convencionais de se falar sobre o assunto eletricidade, então, aceita-se ambas como Literal 2. Mas, como não temos uma forma objetiva de saber o que a eletricidade ‘realmente é’, impossibilita-se a aceitação de serem Literal 4. Se ambas assumem a forma de Literal 2 e não aceitáveis como Literal 4, a suposição c) Literal 2 = Literal 4 é uma falácia. Além do mais, as duas sentenças são formas convencionais, ou seja, Literal 1, não podendo assumir a forma de Literal 4, então a suposição b) Literal 1 = Literal 4, também é falsa.

Após mostrar as diferenças do uso de literal dentro da Teoria da Metáfora Conceitual, o autor nos aconselha evitar essa palavra o máximo possível, quando se discute sobre metáforas, mas admite que, ironicamente, já é esperado que a palavra ‘literal’ seja usada para contrastar o que é ‘metafórico’.

Deveria estar claro agora porque não é uma boa ideia usar a palavra *literal* em discussões sobre metáfora, ou se alguém o fizer, limitá-la claramente ao uso Literal 3. Somente o Literal 3 opõe-se à *metafórico*. Ironicamente, o objetivo de se usar *literal* em discussões sobre metáfora é geralmente apresentar um contraste para *metafórico*¹¹. (LAKOFF, 1986, p. 295, tradução nossa)

Saiba-se que ‘metafórico’ para este estudo, não o que tradicionalmente é definido como figura de linguagem, que usamos quando explicamos, descrevemos ou caracterizamos algo ou alguém, em termos de um ato ou uma pessoa, para embelezar ou tornar o texto mais rebuscado. Aquilo que o dicionário Aurélio (2009, p. 551), define como: “Tropo em que a significação natural duma palavra é substituída por outra que tem relação de semelhança”, e a exemplifica com a palavra ‘raposa’, que seria usada, metaforicamente, para descrever uma pessoa ‘astuta, sendo esta a visão restrita da metáfora. Necessita-se de uma visão mais abrangente e que inclua as El’s, formas cotidianas de se expressar, como representações convencionais e metafóricas.

Na seção seguinte, trataremos mais detalhadamente sobre metáforas e a Teoria da Metáfora Conceitual que baseará nosso estudo.

¹¹ *It should be clear now why it is not a very good idea to use the word literal in discussions of metaphor, or if one does, to limit it clearly to Literal 3. Only Literal 3 contrasts with metaphorical. Ironically, the purpose of using literal in discussions of metaphor is usually to provide a contrast with metaphorical.*

2.1.2 Metáforas

Uma vez que a presente pesquisa se propôs a trabalhar com metáforas e máquinas (programas de computadores), julgamos ser bastante apropriado iniciarmos esta seção com a citação sobre metáfora de Sardinha (2007, p. 12):

A metáfora é um recurso tão humano que talvez seja a última coisa que os robôs do futuro entendam. Provavelmente eles responderão 'não tem registro, não tem registro', como aquele robô de *Perdidos no espaço*, quando ouvirem uma metáfora.

A metáfora é tão relevante para nossas vidas, que Sardinha (2007, p. 16) destaca algumas razões para seu estudo. Sendo assim:

- Entender melhor como conceitualizamos o mundo, as pessoas, os sentimentos, os conceitos mais profundos e duradouros na humanidade;
- Enxergar criticamente como grupos sociais e ideologias enquadram o mundo e que tipos de mensagens querem passar;
- Perceber como conceitualizamos o mundo, individual e sócio-historicamente;
- Detectar o estilo de escritores, políticos e outros profissionais; e
- Dar-nos conta de que tudo isso é feito pela linguagem.

O número de estudos e pesquisas sobre metáforas é bastante significativo na atualidade. Apenas fazendo uma busca por títulos no *site* de Periódicos da CAPES usando a palavra 'metáfora', encontramos 2.187 resultados para o período de 1952 a 2016 (64 anos). Restringindo a pesquisa para o período de 2006 a 2015, vemos que o número de resultados é de 1.552. Ou seja, em 10 anos foram produzidas mais de 70% das pesquisas em que a metáfora aparece citada de alguma forma ou é o objeto principal de estudo.

Existem diferentes teorias ou abordagens relacionadas às metáforas – a visão mais tradicional, a conceitual, a sistemática, a gramatical etc. Dentre essas, restringimos nosso foco para a Teoria da Metáfora Conceitual e, mais especificamente, para a Hipótese da Metáfora Primária de Grady (1997) e as contribuições de Lima (2003 e 2006), no que se refere à influência da cultura para as metáforas e cenas primárias. Esta escolha se justifica por ser a teoria que mais se adéqua ao estudo de EI's, pois poderá dar uma rica base sobre os processos envolvidos na produção e compreensão de metáforas e, em especial, para as EI's, nosso objeto de estudo.

2.1.2.1 Teoria da Metáfora Conceitual

A Teoria da Metáfora Conceitual¹² foi iniciada por George Lakoff e Mark L. Johnson, tendo como marco a publicação de *Metaphors We Live By* em 1980, conforme mencionado anteriormente. Esta foi a primeira versão dos autores sobre este tema. Para evidenciar sua importância, digitamos o termo ‘Teoria da Metáfora Conceitual’ no motor de busca do *Google*, em maio de 2016 e o termo foi referenciado 194.000 vezes. Repetimos o mesmo procedimento em abril de 2017, e encontramos 206.000 referências. Em dezembro de 2017, refizemos a mesma busca e obtivemos 219.000. Um acréscimo de, aproximadamente, 12.000 referências dentro do período de um ano, e de 13.000 referências de abril a dezembro de 2017. Isso evidencia a importância dessa teoria para os estudos da metáfora no Brasil. Fazendo a mesma busca, com o termo em inglês *Conceptual Metaphor Theory*, encontramos 12.200.000 ocorrências em maio de 2016 e 44.500.000, em dezembro de 2017, o que nos leva a perceber sua importância de forma global. Essa teoria evidencia e defende que a metáfora é um fenômeno cognitivo, culturalmente construído.

De acordo com a Teoria da Metáfora Conceitual, encontramos metáforas na linguagem comumente usadas em nosso cotidiano, tais como nas seguintes sentenças: “Estou perdendo meu tempo”, “Nossa história juntos chegou ao fim” ou “Explodi de tanta raiva”. Podemos considerá-las metáforas, pois unem conceitos muito díspares (domínios), como: tempo e dinheiro, namoro e viagem e emoção e pressão em um recipiente, respectivamente.

Nessa teoria, metáfora significa ‘metáfora conceitual’, sendo a representação da relação entre a língua e a cognição humana, e não apenas uma figura de linguagem, em que aspectos do domínio fonte são evidenciados no domínio alvo, “como um mecanismo imaginativo da razão que se configura como um modo fundamental de compreendermos e fazermos emergir sentidos” (LIMA, FELTES e MACEDO, 2008, p. 130). Em suma, a metáfora deixa de ser considerada um fenômeno apenas linguístico e evolui para uma operação cognitiva.

¹² Alguns autores preferem utilizar a tradução conceptual. Aqui optamos por usar a forma conceitual. Somente usaremos conceptual se estiver grafada desta forma em citações diretas.

Os estudiosos da área convencionaram que as metáforas conceituais são sempre escritas em caixa alta, como no exemplo: IDEIAS SÃO ALIMENTOS. A realização linguística das metáforas conceituais é denominada expressão metafórica. Os exemplos supracitados “Estou perdendo meu tempo”, “Nossa história juntos chegou ao fim” e “Explodi de tanta raiva” são, portanto, expressões metafóricas licenciadas pelas metáforas conceituais TEMPO É DINHEIRO¹³, O AMOR É UMA VIAGEM, e EMOÇÃO É UM FLUIDO EM UM RECIPIENTE, respectivamente.

Mencionamos os domínios fonte e alvo. Estes mostram as áreas do conhecimento ou da experiência humana, que são relacionadas, quando usamos metáforas. Sardinha (2007, p. 31) os define da seguinte forma:

Domínio-fonte é aquele a partir do qual conceitualizamos alguma coisa metaforicamente... geralmente é algo concreto, advindo da experiência. O domínio-alvo é aquele que desejamos conceitualizar; esse é o domínio abstrato.... Um mesmo domínio-fonte pode servir a vários domínios-alvo; por exemplo, VIAGEM é fonte do conceito AMOR, mas também de VIDA.

As relações entre os domínios fonte (**A**) e alvo (**B**) são chamadas de mapeamentos ou projeções metafóricas.

A projeção metafórica de **A** para **B** é motivada naturalmente por uma correlação estrutural regular que associa **A** a **B**. Os detalhes do mapeamento entre **A** e **B** são motivados pelos detalhes da correlação estrutural, sendo a relação específica de **A** para **B**. (LIMA; FELTES; MACEDO, 2008, p. 138)

Para a metáfora conceitual, O AMOR É UMA VIAGEM são feitos, por exemplo, os seguintes mapeamentos (Sardinha, 2007, p. 31):

- Viajantes: amantes ou marido e mulher;
- Mapa da viagem: planos futuros da vida a dois;
- Destino da viagem: relação feliz a dois;
- Deslocamento tranquilo na viagem: relação sem problemas;
- Deslocamento contínuo e previsível na viagem: relação com problemas, devido à monotonia;
- Pegar carona na viagem: ter um caso fora do relacionamento.

Outro conceito relacionado a Teoria da Metáfora Conceitual é o dos desdobramentos ou acarretamentos. Segundo Sardinha (2007, p. 32), são as

¹³ TEMPO É DINHEIRO é um dos raros casos em que uma metáfora conceitual é também usada como expressão metafórica.

inferências que podemos fazer a partir da metáfora conceitual. Voltando ao exemplo da metáfora AMOR É UMA VIAGEM, podemos inferir que, se uma viagem longa, pode tornar-se monótona e cansativa, então, um casal, que convive há muito tempo, pode achar o relacionamento monótono e enfadonho.

Lakoff (1980, p. 461-464), define-se três tipos de metáforas: as Orientacionais, as Ontológicas e as Estruturais. Ele as classifica de acordo com os domínios físico, cultural e intelectual, porém sua delimitação é muito tênue e elas interagem entre si de maneira significativa. Vejamos o que caracteriza cada uma delas:

- a) Metáforas Orientacionais: não estruturam um conceito em termos de outro, mas organizam todo um sistema de conceitos relacionados uns com os outros. São chamadas Orientacionais devido, a maioria, está relacionada a orientações espaciais – noções como em cima-em baixo, dentro-fora, frente-trás, centro-periferia. As metáforas Orientacionais dão uma orientação espacial a um conceito. Elas representam como se estabelecem estruturas entre os diferentes domínios de conceito físico, cultural e intelectual. As orientações metafóricas não são arbitrárias, mas têm uma base em nossa experiência física e cultural. Apesar das oposições polares mencionadas acima (cima-baixo etc) e de serem físicas em natureza, podem variar de uma cultura para outra. Existem, por exemplo, culturas que orientam o futuro para frente, outras para trás. Isso será um fator relevante para a determinação das metáforas Orientacionais dentro de cada domínio. Tomemos como exemplos de metáforas Orientacionais ALEGRIA É PARA CIMA (e.g. *"I'm feeling up"*; *"That boosted my spirits"*; e *"My spirits rose"*.) e TRISTEZA É PARA BAIXO (e.g. *"I'm feeling down"*; *"I'm depressed"*; e *"He's really low these days"*.). A base física para estas metáforas seria uma postura curva, com ombros tipicamente caídos, que adotamos quando nos sentimos tristes e uma postura ereta, para um estado emocional positivo.
- b) Metáforas Ontológicas: emergem de nossa experiência com objetos e substâncias físicas. Implicam em projetar características de entidade ou substância sobre algo que não tem essas características de maneira inerente. Elas nos permitem visualizar eventos, atividades, emoções, ideias etc, como entidades para diferentes objetivos (para referenciá-las,

categorizá-las, agrupá-las ou quantificá-las). Esses tipos de metáforas dificilmente são percebidos, porque são muito básicos em nossa conceitualização e nosso funcionamento diário. Contudo, são metáforas convencionais, uma vez que entendemos algo não físico (abstrato) ou com limites não claramente evidenciados. Na maioria das vezes, envolvem o uso de um conceito do domínio físico para estruturar um conceito dos domínios cultural ou intelectual. Exemplo: INFLAÇÃO É UMA ENTIDADE (e.g. *"Inflation is lowering our standard of living"*; *"If there's much more inflation, we'll never survive"*; e *"We need to combat inflation"*.) A personificação é um tipo de metáfora Ontológica em que a entidade é especificada como sendo uma pessoa, o que nos permite compreender experiências não-humanas através de atividades, motivações e características que são essencialmente humanas. A metáfora UMA TEORIA É UMA PESSOA é um exemplo (e.g. *"His theory explained to me the behavior of chickens raised in factories"*; *"This fact argues against the standard theories"*; e *"Life has cheated on me"*.)

- c) Metáforas Estruturais: resultam de mapeamentos complexos. Exemplo: AMOR É UMA VIAGEM (e.g. *"Look how far we've come"*; *"We're at a crossroads"*; e *"We'll just have to go our separate ways"*.), formada por mapeamentos entre viajante e amante, caminho percorrido e vida a dois, destino da viagem e felicidade etc. Essas são as prototípicas metáforas conceituais. Casos em que um conceito é metaforicamente estruturado em termo de outro. DISCUSSÃO é estruturado em termos de GUERRA. As metáforas estruturais, frequentemente, envolvem o uso do conceito de um domínio (GUERRA como um fenômeno físico ou cultural) para estruturar um conceito de outro domínio (DISCUSSÃO como sendo primariamente intelectual, mas com conteúdo cultural).

Nem todos os aspectos do domínio fonte são utilizados metaforicamente no domínio alvo. Por exemplo, se analisarmos a metáfora conceitual TEORIAS SÃO EDIFÍCIOS, veremos que nem todas as partes do conceito de um edifício são encontradas nas expressões metafóricas licenciadas por essa metáfora. Em geral, somente a estrutura e a parte exterior de edifícios são usadas. A parte interior de um edifício não é normalmente usada, digamos as escadas, corredores, cômodos e teto. Então, o conceito metafórico TEORIAS SÃO EDIFÍCIOS tem somente uma parte que

é usada. As palavras ‘construir’ e ‘base’ são exemplos da parte ‘usada’ e são facilmente encontradas em nossa linguagem literal cotidiana sobre teorias. Exemplos sobre a parte ‘não usada’ seriam, de acordo com Lakoff e Johnson (1980, p. 471), os seguintes: “Sua teoria tem milhares de pequenos quartos e corredores longos e sinuosos” ou “Suas teorias são sempre barrocas”, ou ainda “Teorias complexas geralmente têm problemas como o encanamento”. Essas sentenças saem do domínio de linguagem literal usual e são chamadas de linguagem ‘figurativa’ ou ‘imaginativa’. Portanto, expressões literais e expressões imaginativas podem ser exemplos da mesma metáfora.

Metáfora Nova é aquela que não é, normalmente, usada para estruturar parte de nosso sistema conceitual, mas uma forma nova, como o próprio nome diz, de pensar sobre algo como, por exemplo, “As teorias clássicas são como patriarcas que geram muitos filhos, a maioria dos quais brigam incessantemente”. Ela é um dos três subtipos de metáfora imaginativa ou não literal apresentados por Lakoff e Johnson (1980, p. 472), como aqueles que estão fora da parte ‘usada’ de um conceito metafórico, que estrutura nosso sistema conceitual normal. Os outros dois subtipos são extensões da parte ‘usada’ da metáfora como, por exemplo, “Estes fatos são os tijolos e argamassa da minha teoria” e exemplos da parte ‘não usada’ da metáfora literal “Sua teoria tem milhares de pequenos quartos e corredores longos e sinuosos”.

Outro tipo de expressões metafóricas apresentado pelos autores são as expressões metafóricas idiossincráticas. Exemplos dessas são “pé de serra”, “cabeça de alho” e “perna da mesa”. Elas são formadas por um, dois ou três, no máximo, de uma parte usada do conceito metafórico e podem ser interessantes para formar metáforas novas ou piadas. Esclarecendo, a expressão “o pé da montanha” é a única parte do conceito metafórico UMA MONTANHA É UMA PESSOA usado para sua criação. Ou seja, “são idiossincráticas, não-sistemáticas e isoladas. Elas não interagem com outras metáforas, não têm um papel importante em nosso sistema conceitual e, portanto, não são metáforas pelas quais vivemos”. (LAKOFF E JOHNSON, 1980, p. 472). Essas são as únicas expressões, na concepção dos autores, que poderiam ser consideradas ‘mortas’, se alguma expressão metafórica merecesse assim ser chamada.

É importante mencionar que Lima (1995, p. 15) discorda da afirmação acima feita por Lakoff e Johnson, em relação a não sistematicidade dessas expressões:

Achamos que essa metáfora é interessante porque revela algo além dessa primeira metáfora, O PREGO É UMA PESSOA. Se o prego, como pessoa, só tem cabeça é porque a grande metáfora esta na cabeça e não no prego propriamente dito. Lakoff & Johnson falam em sistemas conceituais, organizados a partir da experiência do homem com o seu próprio corpo e com o mundo. Termos como as partes do corpo, por exemplo, têm ampla aplicação na nossa linguagem do dia-a-dia e na forma de nomearmos as coisas. Um único domínio lexical, como a cabeça, pode ser estruturado no nosso sistema conceitual de várias formas e inclui conceitos concretos e abstratos.

O estudo de Lima classificou nove metáforas conceituais com a palavra ‘cabeça’, em inglês e português, e 311 grupos de expressões, evidenciando a grande sistematicidade conceitual que há entre elas.

A primeira versão da Teoria da Metáfora Conceitual, apesar de ter sido seminal, deixava lacunas e pedia “uma teoria mais robusta, que explicasse mais que a simples existência de sistemas metafóricos na língua” (LIMA; FELTES; MACEDO, 2008, p. 144). Alguns dos questionamentos realizados eram, por exemplo, o porquê de ocorrer somente parte de um domínio no mapeamento de algumas metáforas; ou porque algumas metáforas tinham uma base experiencial corpórea direta; ou ainda, explicar a relação entre várias metáforas, algumas por compartilharem grande parte de suas estruturas e conteúdos, outras, por abrangerem contradições nas projeções de um mesmo domínio.

Lakoff e Johnson (1999, p. 45-46) declaram ter respostas parciais para perguntas nessa perspectiva: “por que existe uma variedade tão grande de metáforas conceituais; como são aprendidas e quais são os detalhes precisos; qual o mecanismo pelo qual raciocinamos metaforicamente; e quais metáforas são universais (ou pelo menos, generalizadas) e por quê”.

As respostas são apresentadas em uma nova versão da Teoria da Metáfora Conceitual que integra quatro teorias: a) a teoria da fusão de Christopher Johnson, b) a teoria da metáfora primária de Grady, c) a teoria neural da metáfora de Narayanan e d) a teoria da mesclagem conceitual de Fauconnier e Turner (Lakoff e Johnson, *ibid*, p. 46-47):

- a) A teoria da fusão de Christopher Johnson concerne ao processo de aprendizagem e inicia-se na infância. A palavra fusão se refere à associação entre as experiências e os julgamentos subjetivos (não sensório-motores) e as experiências sensório-motoras. O exemplo dado

pelos autores facilita a compreensão: crianças pequenas não conseguem distinguir as experiências sensório-motoras (e.g. a sensação de calor ao serem abraçadas) das não sensório-motoras (e.g. a experiência subjetiva do afeto) quando ocorrem juntas e as associações são automaticamente construídas entre os dois domínios. Somente posteriormente, no período da diferenciação, é que as crianças conseguem separar os domínios, mas as associações construídas entre os domínios persistem, formando os mapeamentos das metáforas conceituais, que licenciarão expressões nessa continuidade “um sorriso caloroso”, “um grande problema” e “um amigo próximo”.

- b) A teoria da metáfora primária de Grady defende que todas as metáforas complexas são ‘moleculares’, formadas por partes metafóricas ‘atômicas’, as metáforas primárias. Estas têm uma estrutura mínima, fluem natural, inconsciente e automaticamente por meio das fusões. As metáforas complexas são formadas por mesclas conceituais. Ainda segundo esta teoria, as primeiras experiências universais nos levam às fusões gerais, que se desenvolvem como metáforas conceituais convencionais universais (ou generalizadas). Um exemplo de metáfora complexa seria a metáfora TEORIAS SÃO EDIFÍCIOS, composta pelas metáforas primárias ORGANIZAÇÃO É ESTRUTURA FÍSICA e PERSISTIR É PERMANECER ERETO.
- c) A teoria neural da metáfora de Narayanan afirma que as ‘associações’ realizadas no período da fusão são neuralmente percebidas em ativações simultâneas, que resultam em conexões neurais permanentes, estabelecidas entre as redes neurais, que definem os domínios conceituais. A base anatômica das ativações fonte-para-alvo é formada por essas conexões. As ativações constituem os desdobramentos metafóricos. De acordo com esta teoria, um desdobramento em nível neural ocorre quando uma sequência de ativações neurais, ‘A’, resulta em uma ativação neural ‘B’. Se ‘B’ está conectado a um grupo neural, ‘C’, em uma rede, que caracteriza outro domínio conceitual, então ‘B’ pode ativar ‘C’. Isso constitui, portanto, um desdobramento metafórico, pois a ativação de ‘B’ é um desdobramento literal; ‘C’ está

‘metaforicamente’ ligado a ‘B’, uma vez que está em outro domínio conceitual; portanto, a ativação de ‘C’ é um desdobramento metafórico.

- d) A teoria da mesclagem conceitual de Fauconnier e Turner estabelece que domínios conceituais distintos podem ser ativados simultaneamente e, sob determinadas condições, podem ser formadas conexões entre esses domínios, levando a novas inferências. Tais ‘mesclas conceituais’ podem ser de dois tipos: convencionais ou totalmente originais. Essas mesclas convencionais são os mecanismos pelos quais, de acordo com a teoria de Grady, as metáforas primárias se juntam para formar as metáforas complexas.

A junção dessas quatro teorias implica na formação de um sistema amplo de metáforas primárias, de forma automática e inconsciente, desde nossa infância, simplesmente por estarmos vivos e realizando atividades cotidianas. Segundo Lakoff e Johnson (1999, p. 47), não temos escolha nesse processo “devido à maneira como as conexões neurais são formadas durante o período de fusão, nós naturalmente pensamos usando centenas de metáforas primárias”.

A Hipótese da Metáfora Primária será alvo, a partir desta seção, de nossa pesquisa. Essa hipótese nos interessa particularmente, pois faz referências à influência cultural na formação das metáforas. Sobre a Hipótese da Metáfora Primária, Lima (2003, p. 17) explica que há dois tipos de metáforas, que se subdividem em outros tipos:

De um lado, há metáforas geradas por correlações entre experiências corpóreas de níveis distintos – as primárias ou compostas de primárias – e, do outro, aquelas que envolvem percepção de semelhanças entre objetos ou algumas de suas características – as metáforas de imagem, de semelhança e de nível genérico.

Inicialmente, dedicaremos-nos às metáforas do primeiro tipo, as correlacionais. A Hipótese da Metáfora Primária de Grady defende que todas as metáforas complexas (ou compostas) são ‘moleculares’, formadas por partes metafóricas ‘atômicas’, ou seja, as metáforas primárias (ou primitivas). Estas têm uma estrutura mínima, fluem natural, automática e inconscientemente por meio de mesclas conceituais, fusões estabelecidas, que se desenvolvem como metáforas conceituais convencionais universais (ou generalizadas), melhor explicadas na teoria de Christopher Johnson (item ‘a’ supracitado). Um exemplo de metáfora complexa seria

a metáfora VIDA COM PROPÓSITOS É UMA VIAGEM, composta pelas metáforas primárias PROPÓSITOS SÃO DESTINOS e AÇÕES SÃO MOVIMENTOS.

Sobre os domínios fonte e alvo, a hipótese de Grady afirma serem mais restritos, com experiência mais localizada e características bastante específicas. O domínio fonte e alvo se diferenciam em sua natureza: o primeiro é definido por conteúdo sensorial (imagético) e o segundo é uma resposta para este estímulo/*input* sensorial e não tem conteúdo imagético. Relacionam-se devido uma correlação estreita em suas cenas primárias, sendo esse o construto fundamental da hipótese. Um claro exemplo é a metáfora primária DESEJAR É TER FOME, onde a percepção recorrente de fome (domínio fonte) é experienciada de forma coocorrente à resposta a esse *input* que é o desejo por comida (domínio alvo). A partir do mapeamento das cenas primárias, há o licenciamento das expressões metafóricas, havendo a previsibilidade para uma grande parte delas.

Sobre os aspectos culturais, para as metáforas primárias, Grady acredita que eles não estão envolvidos e devem ser comuns a todas as línguas. Já as metáforas complexas devem ter mais aspectos culturais envolvidos. Lima (2006) percebe e questiona que as experiências corporais podem ser percebidas de formas diferentes em culturas distintas, podendo gerar tanto metáforas primárias similares, quanto ligeiramente ou completamente díspares. Esse questionamento nos direciona para o presente estudo.

O segundo tipo de metáforas especificado por Grady (1997) é as de percepção de semelhança, que se subdivide em metáforas de imagem (um vestido verde limão), de semelhança (Aquiles é um leão) e de relação de instanciação (ARRISCAR É UM JOGO DE AZAR). De acordo com Lima (2003), essas são metáforas bidirecionais e ilimitadas e que precisam ser melhor especificadas.

Apesar de a Teoria da Metáfora Conceitual e a Hipótese da Metáfora Primária preconizarem que as metáforas primárias são universais e, portanto, não deveriam sofrer influência cultural, outros estudos mostram que essa influência pode existir. Siqueira e Parente (2006, p. 1821), por exemplo, tentam averiguar essa influência a partir da compreensão de expressões linguísticas oriundas de mapeamentos primários por indivíduos de culturas distintas, no caso, estudantes da 5ª série do ensino fundamental, oriundos de três comunidades distintas: 1- filhos de pequenos agricultores; 2- crianças do movimento sem-terra; e 3- moradores de região metropolitana. Seus resultados sugerem que

... é interessante levar em consideração aspectos culturais nos estudos sobre as metáforas, sejam elas primárias ou não, uma vez que nossas experiências corpóreas são levadas a cabo em uma cultura específica, que tem práticas mais ou menos universais. (SIQUEIRA E PARENTE, 2006, p. 1821).

A influência cultural é, portanto, um fator que deve ser levado em consideração ao analisarmos qualquer tipo de metáfora. Abordaremos na subseção seguinte a estreita relação entre metáforas e fraseologismo.

2.1.3 Fraseologismo

A fraseologia, ou o estudo das unidades lexicais constituídas por mais de duas palavras gráficas, tornou-se uma área de estudo tão importante e ampla que poderia ser tratada em uma seção específica. Contudo, optamos por inseri-la como uma subseção nesta pesquisa, uma vez que o objeto do presente estudo inclui somente um tipo de fraseologismo, ou seja, as EI's, e não todos os aspectos desse rico campo de investigação.

As EI's figuram no grupo das expressões consideradas como metafóricas. Conforme mencionamos anteriormente, apesar de alguns linguistas as considerarem como metáforas mortas, outros argumentam de forma contrária. Para Kövecses (2010, p. 238) metáforas e EI's estão estreitamente relacionadas:

o que determina o significado geral de uma expressão idiomática (*i.e.* a que conceito ela se relaciona) é o domínio alvo da metáfora conceitual que se aplica a expressão idiomática em questão, e que seu significado mais preciso depende do mapeamento conceitual que a ela se aplica. (tradução nossa)¹⁴

Kövecses defende que a maioria das EI's é conceitual, e não linguística, por natureza, pois são produtos de nosso sistema conceitual, não apenas uma questão lexical. Esta visão se contrapõe à visão tradicional, que divorcia o significado linguístico do sistema conceitual e do conhecimento enciclopédico, que os falantes de uma língua compartilham, o que compactua com a Teoria da Metáfora Conceitual e a Hipótese da Metáfora Primária. Vemos aqui um portão de entrada para a influência

¹⁴ ... what determines the general meaning of an idiom (*i.e.*, what concept it has to do with) is the target domain of the conceptual metaphor that is applicable to the idiom at hand, and that the more precise meaning of the idiom depends on the particular conceptual mapping that applies to the idiom.

que a cultura pode desempenhar na produção de EI's. Contudo, o autor (2010, p. 233) deixa claro que, em alguns casos, o significado das EI's não tem motivação conceitual, como podemos ver no clássico exemplo da EI ***to kick the bucket*** (**bater as botas**), que representa plenamente a visão tradicional.

Um dos grandes estudiosos da linguagem figurada, Gibbs (1993, p. 74), em sua visão conceitual sobre EI's, defende que muitas dessas expressões podem ser compreendidas como parcialmente motivadas por esquemas figurativos de pensamento, que estão vivos e fazem parte de nossos pensamentos e raciocínios diários.

Gibbs (1993, p. 57) aponta algumas das possíveis formas identificadas por estudiosos para compreender o significado de EI's. A obtenção de seus significados estipulados ocorreria de três formas: a) a partir do léxico, considerado rejeitado ou inapropriado, após análise de seus significados literais; b) em paralelo ao processamento de seus significados literais; ou c) diretamente sem qualquer análise de seus significados literais.

A teoria defendida pelo autor é a de que “os falantes compreendem as expressões idiomáticas devido ao conhecimento metafórico que motiva os significados figurativos dessas frases”, GIBBS (1993, p. 57). Essa concepção alternativa de idiomaticidade é importante para nosso entendimento do uso e da compreensão de linguagem figurada, e, mais significativamente, para ilustrar a natureza figurada do pensamento cotidiano.

A possível razão para que alguns estudiosos tratem as EI's como metáforas mortas é justificada pela confusão entre o que são metáforas (históricas) mortas e metáforas convencionais. Gibbs (1993, p. 58) esclarece com o exemplo: “*He's almost gone*”, sentença normalmente usada em nossa linguagem cotidiana para se referir a uma pessoa que está ‘à beira da morte’. Nesse caso, devido à convencionalidade do uso de ‘*gone*’ como ‘*dead*’ (morto), então este significado já deve(ria) ter sido incorporado como um dos significados literais (não metafóricos) para essa palavra.

Gibbs (1993, p. 60) cita Lakoff & Johnson (1980) para explicar o motivo do insucesso ao distinguir entre metáforas convencionais, que são parte de nosso sistema conceitual, e metáforas históricas, que já morreram:

O erro procede da suposição que as coisas em nossa cognição que estão mais vivas e mais ativas são aquelas que são conscientes. Pelo contrário, aquelas que estão mais vivas e mais profundamente entranhadas, eficientes

e poderosas são aquelas que são tão automáticas que se tornam inconscientes e sem esforço. (tradução nossa)¹⁵

A fim de determinar se uma EI está morta ou simplesmente inconscientemente convencional, faz-se necessária (entre outras coisas) uma busca por sua manifestação sistemática na linguagem como um todo e em nossos padrões de raciocínio cotidianos (cf. Lakoff & Johnson, 1980). Como exemplo dado pelos autores de metáfora conceitual temos AMOR COMO NUTRIENTE, que licencia expressões nessa sequência “Eu estou louco por ela”, “A magia acabou” ou “Nós somos um”. Essas expressões são consideradas quase literais, pois não possuem mais qualquer grau de metafóricidade. Mas, devido à sua grande sistematicidade de uso, refletem formas metafóricas sobre como nós pensamos, frequentemente, sobre o amor.

A arbitrariedade do significado da EI defendida por alguns estudiosos se justifica, provavelmente, pela falta de conhecimento dos falantes atuais em relação às raízes metafóricas originais de algumas expressões, como é o caso de ***kick the bucket*** (bater as botas) e ***spill the beans*** (dar com a língua nos dentes). Pensamos ser esse fator relevante e pertinente. Verificando o primeiro registro de algumas das EI's em inglês¹⁶ do nosso *corpus* no *Google Books N-gram Viewer*, pode-se verificar que a grande maioria já aparece em livros ainda no início do século XIX. Dentre as EI's pesquisadas, o primeiro registro mais recente é de 1958, ou seja, pelo menos 60 anos atrás. Existem vários livros, glossários e *sites* em que a origem dessas EI's são averiguadas. É importante salientar, contudo, que muitas podem não ser verídicas.

Ainda sobre a origem das expressões, Kövecses (2010) nos alerta para o fato de que a metáfora que usamos hoje pode não refletir a compreensão atual sobre a nossa cultura. O pesquisador afirma que muitas das expressões metafóricas foram geradas a partir de determinadas situações históricas e, na medida em que elas se fossilizam, sua motivação fica, de certa maneira, pouco transparente para os falantes de uma língua.

¹⁵ *The mistake derives from an assumption that things in our cognition that are most alive and most active are those that are conscious. On the contrary, those that are most alive and most deeply entrenched, efficient, and powerful are those that are so automatic as to be unconscious and effortless.*

¹⁶ Explicando a grosso modo, o *Google Books N-gram Viewer* realiza pesquisas de frases (até o máximo de cinco palavras) em um *corpus* de, aproximadamente, cinco milhões de livros, publicados entre 1500 e 2008, e apresenta o registro (percentual) da ocorrência dessas frases em forma de gráfico. O *Google Books N-gram Viewer* ainda não permite pesquisas em língua portuguesa.

Muitas expressões são convencionalizadas, mas nem todas são El's. Vejamos a diferença no próximo subitem.

2.1.3.1 Expressão Convencional *versus* Expressão Idiomática

Há diferentes formas de expressões lexicais fixas. Faz-se necessário, portanto, distinguir entre expressão convencional e idiomática. O termo 'expressões convencionais' é autoexplicativo, ou seja, convencionou-se usar as mesmas palavras juntas para uma determinada situação, não ocorrendo mudança em seu significado. Com as El's, as palavras também são usadas juntas, mas seus significados sofrem alteração no momento do uso. Tagnin (2013, p. 21) mostra essa diferença com dois exemplos em português – as expressões 'Feliz Natal' e '**mundos e fundos**'. A autora explica que a expressão 'Feliz Natal' está ligada intimamente a um fato social, a comemoração do Natal. Já em '**mundos e fundos**', considera-se uma expressão, devido à sua forma. Convencionou-se então, usar as palavras 'mundos' e 'fundos' juntas, nessa ordem e ligadas pela conjunção 'e'. Explicando mais claramente:

No momento em que a convenção passa para o nível do significado entramos no campo da *idiomaticidade*. Dizemos que uma expressão é idiomática apenas quando seu significado não é *transparente*, isto é, quando o significado da expressão toda não corresponde à somatória do significado de cada um de seus elementos. (TAGNIN, 2013, p. 22)

Podemos perceber, portanto, que toda expressão idiomática é também uma expressão convencional, mas nem toda expressão convencional é idiomática, e que há níveis diferentes de convencionalidade e graus de idiomaticidade, os quais explicaremos mais detalhadamente nas seções seguintes.

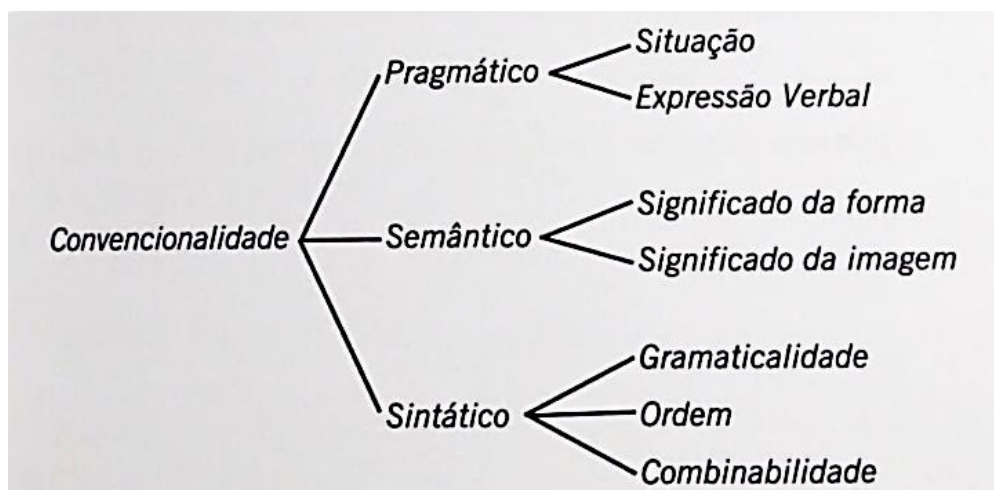
2.1.3.2 Níveis de Convencionalidade

Existem diferentes níveis de convencionalidade. Um nível mais básico, segundo Tagnin (2013, p. 25), seria a relação arbitrária que existe entre uma palavra e seu significado. Por exemplo, não há razão aparente para que uma cadeira seja chamada de 'cadeira'.

Além do nível básico, existem outros níveis – os níveis fonológico, morfológico, sintático, semântico e pragmático. Esses três últimos são os que dizem

respeito às expressões convencionais e idiomáticas. O foco de nossa pesquisa será o nível semântico, pois é “neste nível que se observa a convencionalidade da relação não motivada entre uma expressão e seu significado” (TAGNIN, 2013, p. 26). No nível semântico, tanto o significado de uma expressão linguística, quanto o de uma imagem podem ser convencionalizados. São as metáforas que permeiam a linguagem (Lakoff & Johnson, 2002, *apud* Tagnin, 2013, p. 26-27). Cada um desses níveis compreende subdivisões, que podem ser visualizadas na Figura 1:

Figura 1 – Níveis de convencionalidade



Fonte: Tagnin (2013, p. 27).

Conforme mencionamos anteriormente, uma EI é definida pela convencionalidade de seu significado, ou seja, estamos no nível semântico, isso implica que o seu significado foi convencionalizado, não sua forma. É possível que ela tenha sido originalmente uma expressão metafórica, mas cuja imagem já se perdeu no presente, e passou a ser codificada como um todo. Não podemos inferir seu significado partindo dos significados das palavras que a compõem.

Dentro do nível semântico, diferentes aspectos podem ser convencionalizados. A forma linguística é um exemplo. Se analisarmos as expressões **pay the piper** (arcar com as consequências/pagar o pato) e **in the soup** (estar frito ou estar em uma confusão), veremos que ambas são construídas dentro das regras gramaticais (*pay*-verbo + *the piper*-objeto e *in*-preposição + *the*-artigo + *soup*-substantivo). Portanto, não convencionalizadas sintaticamente, pois não há quebra de nenhuma regra sintática. Ambas as expressões são semanticamente convencionalizadas, pois, dependendo do contexto, seus significados não

correspondem mais à soma do significado individual de cada uma de suas partes, ou seja, 'pagar o cachimbo' e 'na sopa', respectivamente.

Outro aspecto, que pode ser convencionalizado, é a imagem. Por exemplo, as metáforas conceituais BOM É PARA CIMA e RUIM É PARA BAIXO podem transparecer essa mesma imagem em unidades lexicais, tais como **altos e baixos**, **cair de cama e estar na fossa**.

É relevante mencionar que a mesma imagem pode ter significados diferentes em culturas diferentes, como é o caso do gesto feito com o encontro do polegar com o indicador. Para os falantes de língua inglesa significa 'OK', mas para nós, brasileiros, significa um gesto obsceno.

Outro fator controverso no estudo das EI's é a possibilidade, ou não, das mesmas serem decomponíveis. É possível inferir seus significados a partir das partes que a compõem? No próximo subitem, discutiremos um pouco sobre as diferentes visões acerca desse tema.

2.1.3.3 (In)Decomponibilidade das Expressões Idiomáticas

Se presumirmos que as EI's são metáforas mortas, a ideia de que elas são indecomponíveis é aceitável. Contudo, de acordo com Gibbs (1993, p. 61), se acreditamos que os significados figurados não estão mortos, mas metaforicamente convencionalizados, isto significa que, de alguma forma, conseguimos encontrar um *insight* nos significados e formações lexicais das EI's. A teoria da indecomponibilidade das EI's parece ir de encontro a esta visão, pois ela defende que os significados individuais das palavras não contribuem para o significado figurado total dessas expressões. Se pegarmos o exemplo "*John kicked the bucket*" (John **bateu as botas**/morreu), a teoria da indecomponibilidade se aplica perfeitamente. Há outros argumentos, que reforçam esta teoria, tais como a impossibilidade de realizar algumas operações sintáticas nessas expressões sem que seu significado figurado seja perdido (e.g. usar a passiva: *the bucket was kicked by John* – o balde foi chutado por John) e a aparente fixação de suas formas lexicais (e.g. *John kicked the pail* – John **chutou o balde**¹⁷).

¹⁷ Neste caso, temos uma expressão idiomática (popular) na Língua Portuguesa que tem significado totalmente diferente de 'bater as botas/morrer': 'perder o controle', 'desistir de tudo' ou 'abrir mão de algo', como uma resposta a um sentimento de raiva ou ira.

Gibbs (1993, p. 62) afirma que parte dos problemas relacionados à idiomaticidade das expressões está no fato de apenas alguns exemplos serem estudados, como é o caso de **to kick the bucket**, e muitas outras expressões serem deixadas de lado, causando uma teorização incompleta sobre a natureza da idiomaticidade. O autor cita diferentes pesquisas, que mostram que “palavras individuais em várias expressões idiomáticas sistematicamente contribuem para as interpretações figuradas gerais dessas frases, contrária à visão não decomposicional das expressões idiomáticas”, a partir das quais foram identificados três tipos de EI’s: as Decomponíveis; as Indecomponíveis; e as Decomponíveis de forma anômala. Vejamos os exemplos com sua caracterização:

- a) Decomponíveis – consideradas EI’s semanticamente decomponíveis, porque cada um dos seus componentes, obviamente, contribui para a compreensão de seu sentido figurado geral. São exemplos: **to spill the beans** (dar com a língua nos dentes) e **to pop the question** (pedir em casamento).
- b) Indecomponíveis – consideradas EI’s semanticamente não decomponíveis, porque seus componentes não contribuem para a compreensão de seu sentido figurado geral. O exemplo clássico é **to kick the bucket**, mas também podemos citar **to shoot the breeze** (jogar papo fora) e **to put oneself out** (se mostrar, ser inconveniente ou dar o melhor de si).
- c) Decomponíveis de forma anômala – consideradas semanticamente decomponíveis, mas de forma anormal, pois seus componentes individuais se relacionam de forma diferente com seus referentes idiomáticos. Neste caso, o significado metafórico (não literal) dos componentes individuais da EI são ‘acessados’ para auxiliar na compreensão figurada geral. Em **to carry a torch** (ter uma paixão platônica ou continuar sofrendo por um amor perdido), sendo possível identificar seu referente figurado através do conhecimento da palavra *torch* (tocha ou lanterna), metaforicamente convencionalizada para descrever ‘sentimentos calorosos’. Outro exemplo que temos, **to hit the panic button** (entrar em pânico), em que **hitting some buttons** (apertar alguns botões), que metaforicamente convencionalizada, expressa atitudes em circunstâncias extremas.

A partir dessas considerações, vemos que a decomponibilidade de uma EI não se manifesta em sua formação, mas nos significados de suas partes, sejam eles metafóricos ou não, de forma que contribuam, independentemente, para sua compreensão figurada geral. Sua possível análise é uma questão de grau de saliência de suas partes individuais. “Quanto mais saliente é uma palavra dentro de uma expressão idiomática, mais provável será que ela compartilhe o mesmo campo semântico que seu referente idiomático” (GIBBS, 1993, p. 63, tradução nossa).¹⁸

Nossa intuição sobre a analisibilidade de uma EI desempenha um papel importante para determinar sua produtividade sintática, flexibilidade lexical, facilidade de compreensão e facilidade de aprendizagem. A produtividade das EI's se relaciona ao campo da sintaxe, sendo consideradas produtivas aquelas que mantêm seus significados figurados, mesmo que em diferentes construções sintáticas. Por outro lado, são consideradas improdutivas aquelas que têm forma 'fixa/congelada' e perdem sua interpretação figurada, se forem sintaticamente alteradas.

Segundo Gibbs, (1993, p. 63), os resultados de vários experimentos mostraram que a intuição das pessoas sobre a versatilidade sintática das EI's é afetada por sua analisibilidade ou decomponibilidade. As EI's, normalmente decomponíveis (**to pop the question**), parecem ser sintaticamente mais produtivas que as EI's semanticamente indecomponíveis (**to chew the fat – jogar conversa fora/bater papo**). Enquanto que as EI's, anormalmente decomponíveis, não parecem ser sintaticamente produtivas, pois cada parte que as compõem não se refere a algum componente do referente idiomático, mas apenas a uma relação metafórica entre a parte específica e o referente. Assim, ao lermos uma sentença do tipo *A torch for Sally was carried by Jim* (Uma tocha para Sally foi carregada por Jim), dificilmente, reconheceremos algum significado figurado ou idiomático. Na verdade, parece uma sentença incompreensível até em seu significado não figurado.

Outro papel importante que a analisibilidade das EI's desempenha é em suas interpretações *online* imediatas. Gibbs (1993, p. 64) sugere que o processamento das EI's decomponíveis parece ser feito de forma composicional, onde as representações semânticas das partes, que formam a EI são acessadas e combinadas com as regras sintáticas da língua. Por exemplo, na EI **to lay down the**

¹⁸ *The more salient some word is within an idiom, the more likely that the word shares the same semantic field as does its idiomatic referent.* (destaque do original)

law (ditar as regras), as duas partes que a compõem (*lay down* e *the law*) têm um significado que contribui para a interpretação figurada completa da EI. Por outro lado, as EI's semanticamente indecomponíveis nos dão pouca informação quando as analisamos composicionalmente e nos levam, posteriormente, a tentar recuperar os significados diretamente estipulados do nosso léxico mental, o que, provavelmente, exige um esforço de processamento adicional.

O papel que o significado literal das partes que compõem uma EI tem é algo bastante debatido nas pesquisas sobre compreensão de EI's. Algumas pesquisas mostram que não precisamos analisar primeiro o significado literal e depois o figurado. Outras mostram que as pessoas podem acessar os significados literais de cada palavra simultaneamente, o que nos leva a crer que as EI's decomponíveis devem ser mais facilmente compreendidas do que as indecomponíveis, pois a análise do significado literal das partes que compõem as EI's indecomponíveis não nos traria muita informação e poderia até interferir em sua compreensão.

Giora (1997, p. 183-184) declara que “pesquisas contemporâneas sobre linguagem figurada têm afirmado que compreender metáforas é essencialmente idêntico a compreender linguagem literal”¹⁹, e que essa afirmação tem dupla implicação: 1) a linguagem literal não tem prioridade sobre a linguagem não literal (o processamento do significado não literal é direto, não sendo necessário processar o significado literal primeiro e sem interferência do mesmo) e 2) compreender linguagem literal e não literal envolve os mesmos processos complexos e informação contextual. De maneira mais específica, estudiosos a favor dessa visão defendem que as interpretações metafórica e literal são funcionalmente equivalentes: a) que ambos os significados literal e metafórico são computados automaticamente de forma obrigatória e b) que a interpretação de linguagem metafórica não exige uma condição de acionamento. Ou seja, que é tão fácil compreender uma metáfora quanto a linguagem literal.

Essa visão contrasta com a visão tradicional a qual defende que a compreensão de linguagem metafórica e literal diverge em três aspectos: a) a interpretação literal tem prioridade incondicional, com o significado literal sendo sempre ativado anteriormente a qualquer outro significado; b) a interpretação de

¹⁹ *Contemporary research on figurative language has claimed that understanding metaphor is essentially identical to understanding literal language.*

linguagem metafórica exige uma condição de acionamento, podendo ser, por exemplo, a violação de uma regra; e c) significados metafóricos devem ser mais difíceis de serem compreendidos, possivelmente envolvendo um processo sequencial, e, portanto, devem exigir um suporte contextual diferente e maior para sua derivação.

A autora critica as visões tradicional e a contemporânea, alegando que ambas se detiveram em apenas alguns achados, e defende sua própria teoria. De acordo com Giora, a saliência é quem governa a compreensão de linguagem literal e metafórica, isto é, os significados mais salientes são processados primeiramente. Nesse contexto, saliência de palavras e enunciados se define como:

...uma função de sua convencionalidade (e.g., Gibbs 1980), familiaridade (e.g., Blasko and Connine 1993), frequência (e.g., Hogaboam and Perfetti 1975; Neill, Hilliard, e Cooper 1988), ou *status* de inquestionabilidade em um determinado contexto (linguístico e não linguístico)²⁰. (GIORA, 1997, p. 185)

O contexto, segundo a autora, tem um papel importante para a saliência de significados convencionais. Para o caso de palavras que têm mais de um significado possível, o mais popular, ou mais prototípico, ou mais frequentemente usado em uma comunidade, ou aquele com o qual estamos mais familiarizados, ou que aprendemos recentemente, ou ainda, que foi ativado ou feito previsível por contexto anterior será o mais saliente.

O exemplo dado pela autora mostra a importância do contexto para a compreensão do trocadilho: “*Two men walk into a bar, and a third man ducks*”. A convencionalidade da palavra *bar* dentro do contexto da primeira sentença nos leva a interpretá-la, automaticamente, como ‘bar’ o local de venda de bebidas alcoólicas e não ‘barra’, que seria outro significado possível e o mais apropriado neste caso. Somente ao ler a segunda parte da oração é que vemos não haver sentido se utilizarmos o significado mais convencional de *bar*. A tradução para o português (Dois homens esbarram em uma barra, e um terceiro homem baixa a cabeça rapidamente) perde a comicidade que é possível no inglês, justamente pelo uso do jogo de palavras. A segunda parte traz o inesperado, causando para o interlocutor a necessidade de

²⁰ *The saliency of a word or an utterance is a function of its conventionality (e.g., Gibbs 1980), familiarity (e.g., Blasko and Connine 1993), frequency (e.g., Hogaboam and Perfetti 1975; Neill, Hilliard, and Cooper 1988), or givenness status in a certain (linguistic and nonlinguistic) context.*

rever sua primeira compreensão do trecho “*Two men walk into a bar*” e reprocessar os significados para que haja sentido no que está sendo dito.

Embora importante, o contexto tem um papel limitado, pois “apesar de facilitar o significado de uma palavra, não inibe a ativação de significados salientes (e.g. frequente)” GIORA (1997, p. 186), mesmo quando o contexto anterior está altamente favorável ao significado menos saliente (e.g. menos frequente). A evidência vem do experimento com a fixação do olhar de Rayner, Pacht e Duify (1994, *apud* Giora, 1997), que sugere a ativação do significado saliente da palavra, apesar da predisposição do contexto anterior. Mesmo quando o contexto dá suporte ao significado menos saliente, o significado saliente não consegue ser ignorado. Se não houver o conhecimento do significado menos saliente, a compreensão ficará incompleta. No exemplo supracitado (verbo *to duck* – baixar a cabeça rapidamente), foi necessário recorrermos ao dicionário para verificar outras possibilidades de significado, de forma que pudéssemos compreender a sentença.

A saliência de significado é uma questão de graduação, em que a interpretação mais provável é o significado mais saliente de uma palavra ou sentença em um determinado contexto. A hipótese da saliência gradual faz previsões somente em relação à ativação do significado e não sobre quais significados ativados devem ser retidos ou suprimidos e descartados como irrelevantes (Gernsbacher 1990, *apud* Giora, 1997).

Segundo a hipótese da saliência gradual, as visões tradicional e contemporânea citadas anteriormente estão incorretas e deveriam ser modificadas para:

- a. A interpretação saliente (e.g., convencional) tem prioridade incondicional sobre a interpretação menos saliente (e.g., nova): O significado saliente de uma palavra ou enunciado sempre é ativado.
- b. A interpretação nova de um significado saliente envolve um processo sequencial, onde o significado saliente é processado primeiro, rejeitado como o significado pretendido, e reinterpretado. Quanto mais saliente é a linguagem (reinterpretada), mais difícil é rejeitar como o significado pretendido.
- c. A interpretação nova deve ser a mais difícil de deduzir; deve exigir um suporte contextual diferente e maior para sua dedução. (GIORA, 1997, p. 186, tradução nossa)²¹

²¹ a. *Salient (e.g., conventional) interpretation has unconditional priority over less salient (e.g., novel) Interpretation: The salient meaning of a word or an utterance is always activated. b. Novel Interpretation of a salient meaning involves a sequential process, whereby the salient meaning is processed first, rejected as the intended meaning, and reinterpreted. The more salient the (reinterpreted) language, the more difficult it is to reject as the intended meaning. c. Novel*

Giora defende que a dicotomia literal/metafórica deveria ser evitada e, em seu lugar, aplicada a teoria da saliência gradual, que concilia o debate sobre o processamento direto, paralelo e sequencial: significados salientes, literais ou figurados, devem ser processados primeiro; interpretações alternativas de grau de saliência similar devem ser processados em paralelo; e usos novos envolvem um processo sequencial.

Diretamente ligado à decomponibilidade ou ausência das EI's está o grau de idiomaticidade das mesmas. Vejamos com mais detalhes no subitem que segue:

2.1.3.4 Graus de Idiomaticidade

Tagnin (2013, p. 99) esclarece que existem diferentes graus de idiomaticidade. Por exemplo, em **white lie** (mentira branca), somente o adjetivo é idiomático. Trata-se, portanto, de uma expressão com idiomaticidade parcial. O Quadro 2 ilustra os diferentes graus de idiomaticidade.

Quadro 2 – Graus de Idiomaticidade

Grau	Expressão Idiomática – Inglês	Expressão Idiomática – Português	Significado	Explicação
10	<i>to put oneself out</i>	se virar do avesso para fazer algo	empregar muito esforço para realizar algo	Totalmente idiomáticas – mesmo depois de aprender seu sentido, o leitor não encontra relação entre o significado e as partes
9	<i>to chew the fat</i>	bater papo	conversar à toa, sem profundidade; em geral, rapidamente	
8	<i>to burn the candle at both ends</i>	trabalhar dia e noite	trabalhar demasiado	Difícil entendimento – alusão não muito clara – depois de aprender o significado, o leitor poderá perceber alguma relação entre o significado e as partes
7	<i>to beat the bush</i>	fazer rodeios	falar de forma vaga ou com eufemismos sobre um tópico desagradável	
6	<i>to bark up the wrong tree</i>	bater à porta errada	estar em/escolher o caminho errado; pedir algo a alguém que não pode te atender	Imagem conhecida pelo leitor brasileiro - <i>wrong</i> dá a indicação do 'alvo errado' – dedução do significado
5	<i>to pour oil on troubled waters</i>	por água na fervera	tentar acalmar uma situação difícil	Expressões metafóricas que o leitor poderá

Interpretation must be more difficult to derive; it should require more and different contextual support for its derivation.

4	<i>to keep one's head above water</i>	estar com água até o pescoço	sobreviver (financeiramente)	compreender desde que conheça a imagem aludida
3	<i>to have the ball at one's feet</i>	ter/estar com faca e o queijo na mão	estar em uma situação favorável	
2	<i>to speak one's mind</i>	dizer o que pensa	falar sem se preocupar com as consequências	Apenas <i>mind</i> é idiomático – não é uma EI em Português
1	<i>to hold one's head up</i>	andar de cabeça erguida	sem perder o orgulho/autoconfiança	Fácil entendimento - mesma imagem consagrada em ambas as línguas

Fonte: Adaptado de Tagnin (2013, p. 104).

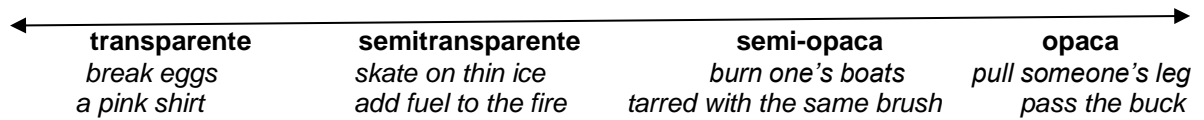
Na parte mais baixa, vemos as expressões menos idiomáticas, ou seja, aquelas em que apenas um ou alguns de seus elementos são idiomáticos ou as expressões metafóricas cuja imagem é facilmente decodificada. As expressões totalmente idiomáticas são aquelas em que nenhum de seus constituintes contribui para o significado total da expressão. Se fizermos um paralelo entre os graus de idiomaticidade e a (in)decomponibilidade das EI's, pode-se perceber que as mais idiomáticas são as indecomponíveis, enquanto que as menos idiomáticas são aquelas normalmente decomponíveis. As que se encontram em um grau médio de idiomaticidade seriam as anormalmente decomponíveis.

Dessa forma, acredita-se que o grau de idiomaticidade relaciona-se à transparência do significado das EI's. Quanto mais transparente uma expressão se apresenta, menor o seu grau de idiomaticidade e, conseqüentemente, quanto maior for seu grau de idiomaticidade, menos transparente será.

Pode-se incluir perfeitamente no Quadro 2 os exemplos citados anteriormente “*There were too many facts in the paper for me to digest them all*” (Havia muitos fatos no jornal para eu digerir tudo) ou “*The argument smells fishy*” (O argumento cheira a trapaça/encrenca) no grupo das expressões menos idiomáticas. Já a EI *to kick the bucket* ficaria no topo do referido quadro, pois nenhum de seus elementos contribui para facilitar o entendimento de seu significado metafórico.

Uma forma de entender o grau de idiomaticidade e a transparência das EI's é apresentado por Fernando e Flavell (1981, p. 28), através de um *continuum* do significado, que engloba desde o significado não figurado e mais transparente, até o nível mais opaco, conforme apresentamos na Figura 2:

Figura 2 – Continuum entre literalidade e figuratividade



Fonte: Adaptado de Fernando e Flavell (1981).

Classificamos dentro desse *continuum* as 43 El's analisadas em seu *corpus* e as apresentamos na seção 3.3.1, Quadro 11. O referido quadro servirá de referência para se averiguar se as El's mais transparentes ou semitransparentes obtiveram melhores traduções do que as opacas e semi-opacas.

A subseção seguinte trata das informações relativas à Tradução Automática e avaliação de tradução, as quais se apresentam como objetos de estudo desta tese.

2.2 TRADUÇÃO AUTOMÁTICA

Para que a presente pesquisa se realizasse, foi necessário buscar sua fundamentação em conhecimentos oriundos da Linguística Computacional que “é um campo de estudos que agrupa os conhecimentos da Linguística, Informática e Inteligência Artificial para o processamento computacional da linguagem humana”. (CUNHA, 2013a, p. 57)

A Tradução Automática é uma modalidade de tradução viabilizada pelo uso de computadores, celulares ou outros dispositivos eletrônicos digitais, cujo objetivo principal é traduzir textos escritos de uma língua para outra, existindo também a possibilidade de a tradução ser feita de uma modalidade para outra. Por exemplo, texto oral para texto escrito e vice-versa.

A Tradução Automática se representa como uma das aplicações do Processamento de Linguagem Natural (doravante PLN) que se mostra diretamente ligada à linguística e à Inteligência Artificial, que busca simular a performance da inteligência humana, utilizando mecanismos ou *software*. Ambas estudam os problemas de análise e geração automática de línguas humanas naturais. Os sistemas de geração de linguagem natural convertem informações ou dados formais legíveis ao computador em linguagem humana seguindo orientações, regras ou algoritmos, e os sistemas de compreensão de linguagem natural convertem

ocorrências em representações formais ou dados legíveis ao computador. As formas de *input* e *output* de linguagem são realizadas através da fala e/ou de textos escritos.

O PLN é composto pela Análise de Linguagem Natural e a Geração de Linguagem Natural. O primeiro é responsável por mapear o *input* dado em representações úteis e analisar os diferentes aspectos da língua. A análise tem como desafio lidar com a polissemia, pois a Linguagem Natural tem formas e estruturas ricas e ambiguidades lexicais, sintáticas e referenciais – um *input* pode significar várias coisas e vários *inputs* podem significar uma mesma coisa. O segundo componente se torna responsável pela produção do texto, pelo planejamento da sentença e pela realização do texto e/ou discurso.

A tradução feita por humanos se configura como uma tarefa bastante complexa, que envolve processos cognitivos e linguísticos. De acordo com Cunha (2013a, p. 52),

A competência linguística deve ser um componente básico para o tradutor. O conhecimento sobre a teoria, abordagens e métodos semânticos reforçam a competência semântica do tradutor e contribuem com seu ato tradutório. ... A competência tradutória pode ser compreendida como um somatório de competências relacionadas aos conhecimentos linguísticos com as habilidades de codificação e decodificação textual.

Ao transferir essa tarefa para um computador, imagina-se que processos semelhantes sejam necessários. Um processo de imitação da inteligência humana.

Apesar de existirem similaridades entre as línguas, alguns padrões e aspectos que parecem universais, tais como palavras para descrever pessoas (homem, mulher etc) e ações (comer e beber), ser polido ou mais agressivo, acha-se também as muitas diferenças com as quais a tradução precisa lidar. Dentre elas, as diferenças em tipologia linguística²², tradução de El's e o isolamento de anomalias. Jurafsky e Martin (2009, p. 898-202) descrevem alguns dos fatos tipológicos, que assemelham ou diferem as línguas, os quais resumimos no Quadro 3.

²² Tipologia Linguística – o estudo das similaridades e diferenças sistemáticas entre línguas (Croft, 1990; Comrie, 1989 *apud* Jurafsky e Martin (2009, p. 898)

Quadro 3 – Fatos tipológicos que assemelham e diferem línguas

Morfologicamente 2 dimensões	Número de morfemas por palavra – de isolados (vietnamita e cantonês) a polissintéticos (yupik siberiano - esquimó)
	Grau no qual os morfemas são segmentáveis – de aglutinados (língua turca) a fusionados (russo)
Sintaticamente diferença na sequência	SVO (sujeito+verbo+objeto) – geralmente têm preposições
	SOV (sujeito+objeto+verbo) – geralmente têm posposições (<i>postpositions</i>)
Estrutura e ligação do argumento dos predicados com seus argumentos	<i>Head-marking</i> – tendem a marcar a relação entre o núcleo e seus dependentes no núcleo Ex: inglês - <i>the man's house</i>
	<i>Dependent-marking</i> – tendem a marcar a relação nos não-núcleos (dependentes) Ex: <i>húngaro – az ember ház-a</i> = inglês: <i>the man house-his</i>
Direção do movimento ou o modo do movimento marcados no verbo ou nos 'satélites'	Línguas <i>verb-framed</i> – marcam a direção de movimento no verbo, deixando os satélites para marcar o modo de movimento. Ex: espanhol – <i>acercar-se, alcanzar, entrar, salir</i>
	Línguas <i>satellite-framed</i> – marcam a direção do movimento no satélite, deixando o verbo para marcar o modo do movimento. Ex: inglês – <i>crawl out, float off, jump down, run after</i>
O que pode ser omitido - pronomes	<i>Pro-drop</i> – podem omitir os pronomes
	<i>Non-pro-drop</i> – não permitem a omissão de pronomes

Fonte: Elaborado pela autora, baseado em Jurafsky e Martin (2009, p. 898-902).

Essas diferenças tipológicas entre línguas podem causar problemas para a tradução. Outras divergências estruturais, mencionadas pelos autores, configuram diferenças idiossincráticas características de línguas específicas, tais como: sequência de adjetivos e substantivos; a estrutura de *relative clauses*; construções sintáticas idiossincráticas (como por exemplo, em inglês, o *there* no início de sentenças, que introduz novas cenas em uma história); o formato de datas e diferentes calendários; ou ainda, o uso do apóstrofo para o caso genitivo, como na língua inglesa.

As divergências lexicais demonstram outro fator que causa grandes dificuldades à tradução, fazendo com que a necessidade de desambiguação esteja frequentemente presente. A polissemia evoca apenas um exemplo de divergência lexical. Em algumas situações, o tradutor também poderá se deparar com um 'hiato lexical' (*lexical gap*), no qual não existe uma palavra ou frase correspondente na língua-alvo. Um exemplo, no português, seria a palavra 'saúde', ou as EI's, nosso objeto de estudo.

Há um conjunto de aplicações computacionais de manipulação da língua que podem facilitar o trabalho de THP e incorporado às suas estações de trabalho. Cunha (2013a, p. 57-58) lista:

- Corretores ortográficos;
 - Corretores gramaticais;
 - Dicionários e glossários digitais;
 - Bancos de dados terminológicos;
 - Bancos de dados de estruturas sintáticas com análise superficial e profunda;
 - Tradução automática e memórias de tradução.
- Estes recursos são produtos de estudos relacionados a diversas disciplinas teóricas da Linguística aplicados à LC.

Algumas dessas aplicações estão disponíveis em computadores pessoais, como é o caso dos corretores ortográficos e gramaticais, e são utilizadas, obviamente, não somente por THP, mas por todos os usuários. A Tradução Automática encontra-se dentre as aplicações mencionadas e se mostra, nesse caso, não como o produto final, mas um facilitador que deve passar por pós-edição, com o objetivo de reduzir o tempo de trabalho do usuário.

A memória de tradução configura outro recurso bastante utilizado por muitos tradutores, assim como os bancos de dados terminológicos, dicionários e glossários. Sobre a memória de tradução, Cunha (2013a, p. 84) afirma que

Os sistemas de MT²³ realizam o armazenamento dos textos traduzidos por tradutores humanos e os apresenta quando são solicitados. Para melhor ilustrar, imagine que um tradutor profissional realiza a tradução manual de um livro. Ao usar um sistema de MT, em toda situação que o texto na LF for repetido, as memórias armazenadas anteriormente serão solicitadas e sugeridas ao usuário. Esse tipo de ferramenta possibilita uma maior organização terminológica da tradução. Por sua eficiência e organização, este tipo de ferramenta ganhou espaço junto à comunidade de tradutores profissionais.

A maioria dos *softwares* de Tradução Automática permite a personalização por domínio ou área do conteúdo textual. Um exemplo de resultados positivos da Tradução Automática se mostra em áreas como a previsão do tempo ou em documentos legais e de governo, em que o número de substituições necessárias se faz limitado, devido ao uso de linguagem formal ou formuláica. Para a comunicação mais coloquial, os resultados produzidos pela Tradução Automática nem sempre usáveis de imediato, necessitando passar por pós-edição.

Em geral, temos os seguintes passos para a geração de Tradução Automática:

- Criação de dicionários ou vocabulários;

²³ Por ter sido usada somente nesta citação direta, a sigla MT (Memória de Tradução) não foi incluída na lista de siglas desta tese, para evitar ser confundida com *Machine Translation*.

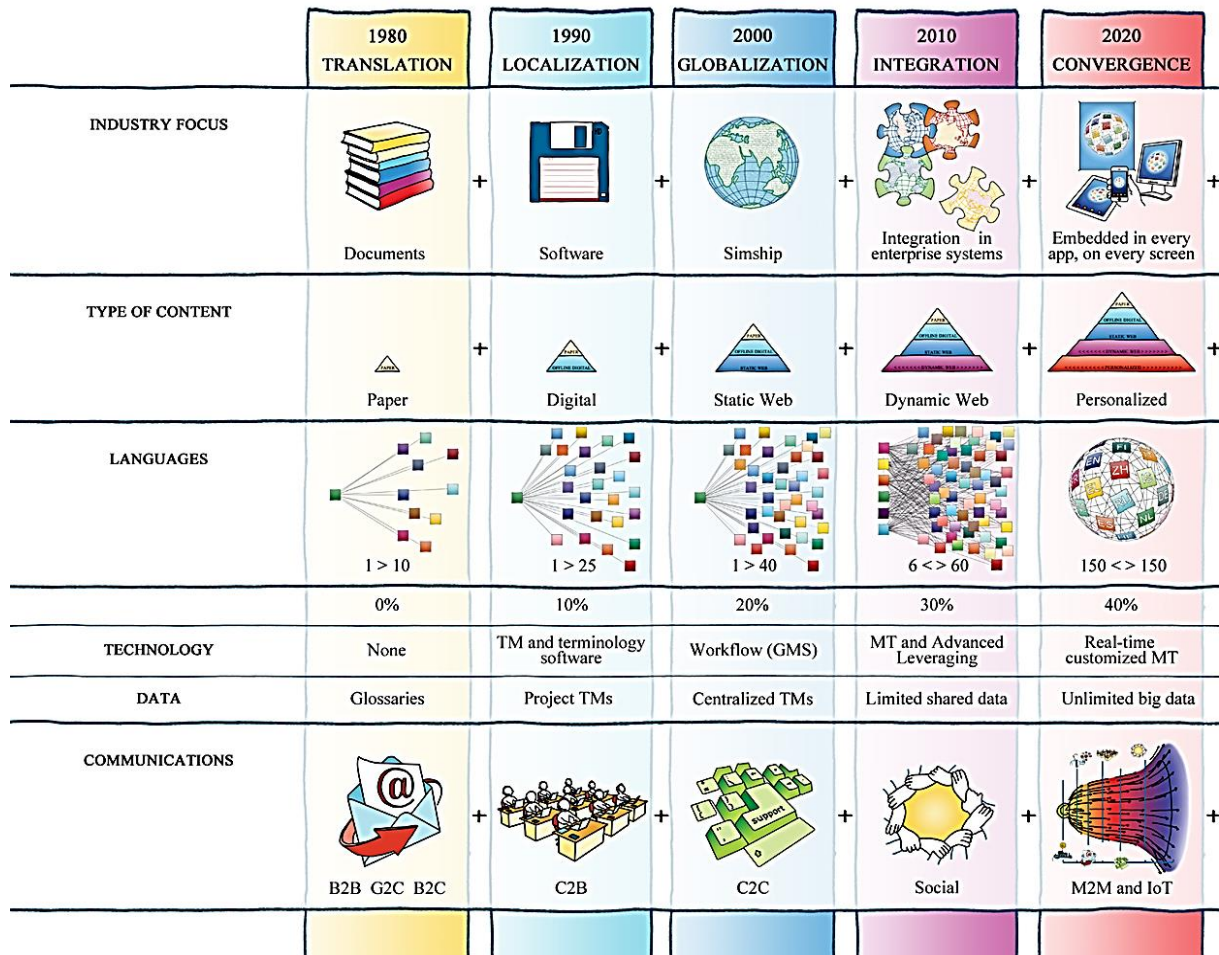
- Gramáticas formais (Gramáticas geradas seguindo formalismo como CFG, DCG, LFG ou HPSG²⁴);
- *Parsing* de gramáticas (Processo computacional de execução de análises vinculadas à gramática); e
- Verificação semântica (Consulta ou filtro em bancos de dados ontológicos, verificação de traços sintático-semânticos).

Mais recentemente, outros procedimentos estão sendo incorporados, à medida que sistemas híbridos estão sendo desenvolvidos. Que serão abordados no subitem 2.2.3.

A evolução na Tradução Automática pode ser vista na Figura 3, que sintetiza seus avanços nas últimas décadas e faz previsões para um futuro próximo. Van der Meer (2015) prevê condições perfeitas para uma enxurrada de desenvolvimentos e oportunidades:

²⁴ CFG = *Context-free Grammar*, DCG = *Definite Clause Grammar*, HPSG = *Head-driven Phrase Structure Grammar*, e LFG = *Lexical Functional Grammar*.

Figura 3 – Avanços da Tradução Automática nas últimas décadas



Fonte: Van der Meer (2015)²⁵.

Legenda:

- TM – *translation memories* (memórias de tradução)
- B2B – *business-to-business* (empresa para empresa)
- G2C – *government-to-citizen* (governo para cidadão)
- B2C – *business-to-consumer* (empresa para consumidor)
- C2B – *consumer-to-business* (consumidor para empresa)
- C2C – *consumer-to-consumer* (consumidor para consumidor)
- M2M – *machine-to-machine* (máquina para máquina)
- IoT – *Internet of Things* (Internet das Coisas)

A Figura 3 apresenta cada década, dentro de um período de 50 anos, de 1980 a 2020. É possível ver o acréscimo de novas ferramentas, de acordo com o foco da indústria, os tipos de conteúdos disponibilizados, o aumento no número de línguas contempladas, a tecnologia empregada, o volume de dados e a comunicação estabelecida. De acordo com o autor supracitado, estamos no período da

²⁵ Disponível em: <<https://taus.net/think-tank/articles/perfect-storm-conditions-for-machine-translation>>. Acesso em: 05 jan. 2018.

convergência, no qual a Tradução Automática se disponibiliza em todos os aplicativos e telas, com conteúdos cada vez mais personalizados; a possibilidade de trabalhar com, aproximadamente, 150 línguas, de forma bidirecional, as quais representam 40% do número de línguas faladas no planeta; utilizando uma tecnologia customizada e em tempo real; com acesso a dados ilimitados; e comunicação de máquina para máquina e *Internet das Coisas*.

Contudo, até chegar ao atual estado de desenvolvimento, a Tradução Automática passou por percalços na sua história, os quais apresentamos, resumidamente, na próxima seção.

2.2.1 Histórico

Apresentamos, de forma resumida, alguns dos principais eventos, que contribuíram para o desenvolvimento da Tradução Automática. De forma geral, os conflitos entre países foram (e continuam sendo) grandes promotores dos estudos e projetos para o crescimento e aperfeiçoamento da Tradução Automática. Iniciaram-se nos anos 50 e intencionava a criação de sistemas, que traduzissem textos automaticamente, com alta qualidade, semelhantes às traduções feitas por humanos, principalmente, para fins de espionagem.

Nos anos 50, o filósofo Yehosha Bar-Hillel foi pioneiro no questionamento sobre a ausência do conhecimento de mundo necessário para um computador traduzir a sentença "*Little John was looking for his toy box. Finally he found it. The box was in the pen. John was very happy*". (O pequeno John estava procurando sua caixa de brinquedo. Finalmente a encontrou. A caixa estava no cercado. Este ficou muito feliz.). Para o filósofo, não seria possível atribuir a um sistema automático o conhecimento de *pen*, como uma abreviação de *play-pen*. Se reconhecer a variação de significados de uma palavra era algo impossível para um computador, o que dizer sobre o reconhecimento das *El's*? Contudo, o próprio Bar-Hillel (1952) fala sobre o assunto, afirmando que seria necessário primeiro definir 'idiomático' para um computador. Ele menciona três métodos possíveis de serem implementados, todos relacionados ao uso de dicionários. Nesse artigo (não publicado), ele discute possibilidades futuras para o desenvolvimento dos sistemas de Tradução Automática em prática hoje, que convergem para a unidirecionalidade ou bidirecionalidade da tradução de textos.

Em 1962, forma-se a Associação para Tradução Automática e Linguística Computacional, nos Estados Unidos. Dois anos depois, a Academia Nacional de Ciências criou o ALPAC (*Automatic Language Processing Advisory Committee – Comitê de Aconselhamento sobre o Processamento Automático da Linguagem*). Contudo, em 1966, este mesmo comitê apresentou um relatório com informações que desacreditavam o potencial da Tradução Automática em produzir traduções de boa qualidade, o que levou à redução de fundos para pesquisas na área.

No início dos anos 70, com o conflito entre Estados Unidos e Vietnã, a tradução de manuais militares restabelece a produção de Tradução Automática em larga escala com um relatório do DDR&E (*Director of Defense Research and Engineering*). Outro uso de Tradução Automática, nessa década, relaciona-se a Xerox, que passou a utilizar o SYSTRAN para traduzir seus manuais. O SYSTRAN, mesmo sem o apoio financeiro do governo, continuou a realizar manualmente as traduções e catalogá-las no seu sistema, tornando-se, posteriormente, um parâmetro de comparação entre Tradução Automática.

Outro momento relevante para a história da Tradução Automática, aconteceu no final dos anos 80, com o crescente interesse por modelos estatísticos de Tradução Automática, devido o aumento da força computacional e na redução dos preços dos computadores, bem como o aumento do valor do trabalho humano. Várias empresas de Tradução Automática surgiram, incluindo a TRADOS, primeira empresa a desenvolver e comercializar a tecnologia de memória de tradução.

Os anos 90 foram especialmente importantes para o desenvolvimento de teorias e estudos da Tradução Automática, devido às necessidades de tradução de documentos em diferentes línguas com a criação da União Europeia.

A Tradução Automática na *Web* (rede mundial de computadores), iniciou-se com o SYSTRAN oferecendo traduções grátis de pequenos textos em 1996, seguida por outros, como o *Babel Fish* (financiado pela YAHOO) e o *Globalink* (*Software* embarcado em computadores com sistema operacional *Microsoft Windows*). A necessidade de comunicação entre povos de diferentes línguas somente cresceu nos anos 2000, fazendo com que os sistemas de Tradução Automática se tornassem cada vez mais utilizados e popularizados.

Em 2007, um projeto de pareamento e alinhamento de palavras, bem como expressões, gerou o motor de busca aberto para Tradução Automática Baseada em

Estatística: o MOSES. Atualmente, de alguma forma, quase todos os tradutores automáticos se utilizam dessa técnica.

No final dos anos 2000, uma nova tecnologia de *machine learning*, chamada *deep learning* ou *deep neural networks*, começou a ser desenvolvida e somente muito recentemente começou a ser utilizada com a Tradução Automática, produzindo traduções de maior qualidade, apesar de ainda estar em um estágio inicial de desenvolvimento, segundo o *site* do *Bing Tradutor*.

Mais recentemente, Van der Meer (2015) afirma que as empresas começaram a considerar a tradução como uma estratégia para ampliar os negócios, ou seja, uma oportunidade: “compreender e falar a língua dos clientes em mercados globais que se expandem leva à mais renda e mais ideias sobre como desenvolver melhores produtos e serviços”²⁶. Os chamados negócios da Era da Convergência, para o autor, exploram como o conhecimento linguístico ou a tradução podem estar incorporados a produtos e serviços, tornando a Tradução Automática parte integral da solução de negócios. Após consultar mais de 50 operadores de tecnologia de Tradução Automática, a conclusão chegada pela TAUS²⁷ consiste em uma adoção massiva dos sistemas de Tradução Automática. Para essa organização, a Tradução Automática foi incorporada à produção de tradução elaborada por humanos e, nos próximos cinco anos, a pós-edição de produtos elaborados por Tradução Automática superará o uso da memória de tradução, transformando-se no principal ambiente de produção da tradução industrial. Outra tendência é que a própria Tradução Automática está no processo de se tornar uma *commodity*, na qual os dados são a maior riqueza, isto é, os dados de linguagem que treinam os motores (pares de língua, falados e monolíngues).

A próxima subseção descreve as principais arquiteturas utilizadas pelos sistemas de Tradução Automática.

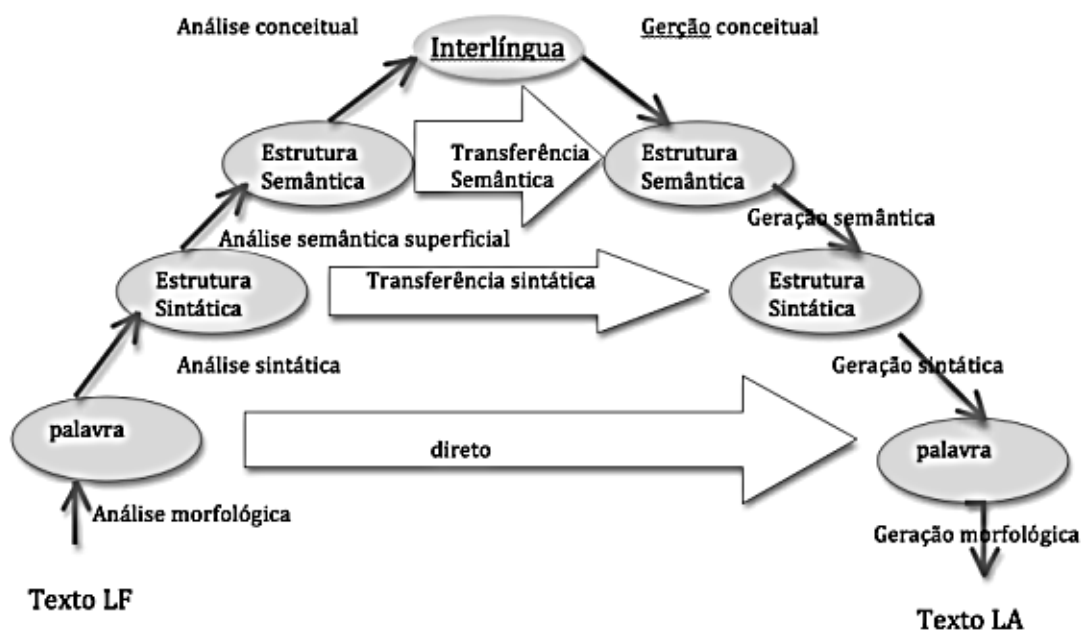
²⁶ *Understanding and speaking the languages of customers in expanding global markets leads to more revenue and more insights in how to develop better products and services.* (tradução nossa)

²⁷ TAUS é o acrônimo para *Translation Automation User Society* (Sociedade dos Usuários de Automação de Tradução), uma organização industrial neutra e independente que desenvolve comunidades através de eventos e grupos de usuários *online*, compartilhando conhecimentos, métricas e dados que auxiliam todos os investidores da indústria da tradução a desenvolver melhores serviços.

2.2.2 Principais arquiteturas

O Triângulo de Vauquois consiste numa ilustração clássica (de 1968) das principais arquiteturas utilizadas na Tradução Automática. Existem dois tipos de mecanismos de Tradução Automática. O primeiro se refere à arquitetura direta, e o segundo se subdivide em arquitetura de transferência e de interlíngua. Essas arquiteturas correspondem a diferentes níveis de transferências, de acordo com a análise realizada pelo sistema de tradução.

Figura 4 – O triângulo de Vauquois



Fonte: Jurafsky; Martin (2009, p. 903).

A forma mais primitiva de transferência é a da arquitetura direta e, portanto, se encontra na base da pirâmide. De acordo com Jurafsky e Martin (2009, p. 904-910), consiste na substituição de cada palavra da língua fonte (doravante LF) por outra da língua alvo (doravante LA). Este tipo de tradução se baseia em um grande dicionário bilíngue, em que cada entrada corresponde a um pequeno programa, cuja tarefa é traduzir uma palavra. Após a tradução, podem ser aplicadas regras simples de reordenação das palavras como, por exemplo, a sequência de adjetivos e substantivos entre o inglês e o francês ou o português. A fundamentação deste processo consiste na tradução incrementada do texto da LF em texto da LA. Este

método não é mais usado em sua forma pura, porém a mesma intuição transformacional que o baseia, subjaz em todos os sistemas modernos de Tradução Automática, tanto nos estatísticos como nos não estatísticos.

O segundo tipo de arquitetura é a de transferência, que fica no centro da pirâmide. A maioria dos sistemas de tradução atuais possui esta arquitetura variando de uma análise sintática rasa a uma análise semântica profunda. O conhecimento contrastivo entre as LF e LA demonstra sua aplicabilidade. Este modelo pressupõe uma análise da LF, que é seguido por uma fase de geração para, de fato, a sentença resultante ser gerada. Portanto, há três fases envolvidas: análise, transferência e geração.

No topo do triângulo de Vauquois, mostra-se a arquitetura interlíngua, que tem como intuição tratar a tradução como um processo de extração do significado do *input* e expressá-lo na LA. Este esquema pressupõe a existência de uma representação do significado (interlíngua), em uma forma canônica independente de qualquer língua. A intenção é representar da mesma forma todas as sentenças com significados iguais/similares, independente da língua usada.

Os sistemas de Tradução Automática estatísticos partem de outro princípio, de acordo com Jurafsky e Martin (2009, p. 910), focalizam o resultado e não se detêm no processo. Portanto, fazem uso da probabilidade matemática para a determinação dos resultados e se beneficiam das possibilidades mais recentes de uso de grandes *corpora* bilíngues. Além disso, utilizam um modelo estatístico (Bayesiano), que permite à máquina aprender modelos de língua e de tradução a partir de inferências probabilísticas. Para esse tipo de sistema de Tradução Automática, três premissas se fazem essenciais: os modelos de línguas; o modelo de tradução; e um decodificador para produzir os resultados.

No subitem a seguir, descrevemos os paradigmas de pesquisa para a elaboração das traduções, que são outra forma de comparar os sistemas de Tradução Automática.

2.2.3 Paradigmas de Tradução

Além da arquitetura do sistema, os paradigmas de pesquisa representam uma outra forma de classificação dos sistemas de Tradução Automática. Enquanto a arquitetura diz respeito ao método de processamento, o paradigma de pesquisa se

refere aos componentes da informação, que auxiliam o processamento. Contudo, um não pressupõe o outro. (TODARO, 2013).

Os paradigmas da Tradução Automática se dividem em três tipos (baseados ou não baseados em linguística) e a junção desses tipos (híbridos), que forma a tendência mais recente (TODARO, 2013), como podemos verificar no esquema que segue:

a) Paradigmas baseados em linguística: aplicam seleções lexicais, semânticas e de sintaxe para produzir na LA uma percepção apropriada da sentença na LF:

- Técnica das restrições: utiliza restrições na combinação de itens lexicais.

- Técnica do léxico: qualquer sistema que forneça regras para relacionar as entradas lexicais de um idioma às de outro.

- Técnica baseada em regras: faz parte de sistemas que adotaram níveis linguísticos diversos (informações morfológicas, sintáticas e semânticas) para efetuar tradução entre pares de idiomas. O objetivo é converter estruturas da LF em estruturas da LA.

- Técnica 'agitar e assar': seleciona palavras no dicionário léxico da LA e tenta realizar diferentes ordenações dessas palavras até produzir uma sentença que satisfaça todas as restrições sintáticas.

b) Paradigmas não baseados em linguística: o rápido desenvolvimento do poder computacional e a disponibilidade de dicionários e *corpora* monolíngues e bilíngues permitiu aos pesquisadores investigar estes paradigmas. Os *corpora* são essencialmente lexicais, e por isso acarretam problemas de adaptabilidade para a Tradução Automática.

- Técnica baseada em exemplo: faz traduções 'por analogia', ou seja, emula a tradução humana ao reconhecer a semelhança de uma frase na LF com traduções previamente feitas.

- Técnica baseada em estatística: análises estatísticas de *corpora* bilíngues paralelos, pesadamente usadas sem a utilização de regras gramaticais. O contexto é levado em consideração, mas apenas uma janela pequena (entre três e cinco palavras). Técnicas de modelagem estatísticas e algoritmos eficientes ajudam o computador a lidar com o problema de detectar as correspondências entre as LA e LF, no

treinamento de dados e encontrar a melhor tradução para uma sentença nova.

- Técnica baseada em rede neural: a tradução leva em conta o contexto de toda a sentença e produz traduções, que parecem mais fluidas, semelhantes às produzidas por humanos. É baseada em treinamento de rede neural, fazendo uso de representações, de características específicas dentro de cada par de língua. A rede neural define quais dimensões devem ser levadas em conta (gênero, nível de formalidade, classe gramatical e quaisquer outras características não óbvias derivadas do treinamento).

-Técnica baseada em conhecimento: associa conceitos mais profundos a um léxico, que representa um domínio ou ontologias, tais como informações morfológicas, sintáticas e semânticas (TRIPAPHI; SARKHEL, 2010).

- c) Paradigmas híbridos: misto de paradigmas e de arquiteturas da Tradução Automática. A abordagem mais comum é usar métodos linguísticos para obter análises do texto fonte e usar técnicas estatísticas ou de exemplo para efetuar as traduções.

Os sistemas de Tradução Automática baseados em estatística e em redes neurais têm dois elementos em comum: 1- ambos exigem quantidades elevadas de conteúdos traduzidos previamente por humanos, para treinar os sistemas; e 2- nenhum funciona como dicionário bilíngue traduzindo palavras. Baseam-se em uma lista de traduções potenciais, contudo elaboram a tradução pelo contexto da palavra usada na sentença.

O estudo de Leal, Maia e Maia (2014) comparou três sistemas de Tradução Automática, cada um utilizando técnicas diferentes, Tradução Automática Baseada em Regras (*Rule Based Machine Translation* – RBMT), Tradução Automática Baseada em Conhecimento (*Knowledge Based Machine Translation* – KBMT) e Tradução Automática Baseada em Estatísticas (*Statistic Based Machine Translation* – SBMT). A partir de critérios estabelecidos, os autores verificaram as vantagens, desvantagens e recursos necessários para cada sistema de Tradução Automática analisado. O Quadro 4 apresenta os resultados obtidos:

Quadro 4 – Resumo comparativo de três técnicas de Tradução Automática

Critérios de Comparação	RBMT	KBMT	SBMT
Disponibilidade de literatura	média	baixa	alta
Amadurecimento da tecnologia	atual	médio	atual
Complexidade de implementação	alta	média	baixa
Custo de manutenção	alto	baixo	baixo
Desempenho	alto	médio	médio

Fonte: Leal, Maia e Maia (2014).

Analisando o Quadro 4, com base nos critérios de disponibilidade de literatura, desempenho e amadurecimento da tecnologia, a técnica que se destaca é a baseada em regras. No entanto, se forem observados os critérios de disponibilidade de literatura em conjunto com o amadurecimento da tecnologia, a complexidade de implementação e o custo de manutenção, sobressai-se a técnica baseada em estatística, seguida pela baseada em conhecimento.

Segundo Cunha (2013a, p. 12-13), os sistemas de Tradução Automática, que têm seu mecanismo de tradução baseados em regras gramaticais manualmente estipuladas são eficientes na análise e geração de estruturas para a tradução, contudo pecam pela falta de robustez no fornecimento de respostas quando uma situação textual submetida ao sistema não é prevista pelas regras. Para tentar melhorar e solucionar a eficiência deste processo, a Tradução Automática Baseada em Estatísticas foi desenvolvida, popularizando os sistemas. Apesar de sua eficiência e pontos positivos

... ainda apresentam limitações e causam inseguranças. Uma das inseguranças geradas por estes sistemas é a imprevisibilidade dos resultados. Esta imprevisibilidade é causada pela consulta ao corpus quando este é muito extenso ou de um gênero textual diferente. CUNHA (2013a, p. 13)

A criação de sistemas híbridos combina o que há de melhor em cada técnica e paradigma de tradução.

Independente da técnica ou paradigma utilizado, avaliam-se os sistemas de Tradução Automática por meio de métricas automáticas, necessárias para que o produto tradutório ou as traduções elaboradas, sejam analisadas em relação à sua qualidade. Na seção seguinte, descrevemos seu funcionamento e detalhamos aquelas que foram utilizadas nesta pesquisa.

2.2.4 Avaliação de Tradução Automática

Ao analisar o histórico da Tradução Automática, percebe-se claramente que a qualidade das traduções produzidas tem sido o ponto crucial para sua aceitação ou rejeição, seja por parte dos investidores e organismos governamentais, no passado, ou por parte dos milhões de usuários atuais. Segundo Dorr B., Olive J., McCary J., Christianson C. (2011, p. 803), em 1966, no Apêndice 10 da ALPAC “experimentos foram reportados com avaliações humanas de inteligibilidade, assim como a informatividade de uma tradução humana vista depois de estudar a tradução automática”²⁸. Os mesmos autores apontam a importância da avaliação dos sistemas de Tradução Automática (2011, p. 801):

A avaliação de sistemas de tradução automática (TA) é um campo de pesquisa vital, tanto para determinar a efetividade dos sistemas de TA existentes quanto para otimizar o desempenho dos sistemas de TA.²⁹ (tradução nossa).

Existem ainda muitos questionamentos no que se refere ao uso das métricas automáticas de avaliação de Tradução Automática. Dentre os quais, se a diferença encontrada nos resultados das métricas é estatisticamente significativa e o que os números encontrados simbolizam. É senso comum que os resultados são considerados importantes quando compatíveis com aqueles obtidos por seres humanos. Isso nos leva a outro desafio inerente à tradução: se pedirmos a diferentes tradutores humanos para traduzir uma sentença simples, provavelmente, teremos diferentes traduções como resposta. Segundo Koehn (2010, p. 217), a razão para tal é não haver uma resposta única e exata para comparar com os resultados de uma tradução automática, como podemos verificar no Quadro 5.

²⁸ ...experiments were reported with human ratings of intelligibility, as well as the informativeness of a human translation seen after studying the machine translation.

²⁹ The evaluation of machine translation (MT) systems is a vital field of research, both for determining the effectiveness of existing MT systems and for optimizing the performance of MT systems.

Quadro 5 – Dez traduções para uma sentença em chinês - exemplo típico do conjunto de avaliações do NIST 2001

这个机场的安全工作由以色列方面负责。

Israeli officials are responsible for airport security.

Israel is in charge of the security at this airport.

The security work for this airport is the responsibility of the Israel government.

Israeli side was in charge of the security of this airport.

Israel is responsible for the airport's security.

Israel is responsible for safety work at this airport.

Israel presides over the security of the airport.

Israel took charge of the airport security.

The safety of this airport is taken charge of by Israel.

This airport's security is the responsibility of the Israeli security officials.

Fonte: Koehn (2010, p. 218).

Então, como aferir a qualidade de uma tradução feita por computadores? Pode-se solicitar que humanos julgem essa qualidade ou se comparar similaridades entre traduções automáticas e humanas.

A avaliação de sistemas de Tradução Automática “é um campo de pesquisa bastante ativo, e um assunto arduamente debatido” KOEHN (2010, p. 217). Usualmente, a avaliação de sistemas de Tradução Automática se realiza de duas formas: manual ou automática. As avaliações manuais se mostram extensas, caras e demandam mais tempo. Portanto, alguns dos recursos que se poupa quando se usa os TAO's. A fim de minimizar a demanda por esses recursos, algumas métricas computadorizadas foram criadas para desempenhar essa tarefa. De acordo com Koehn (2010, p. 222), o objetivo da avaliação automática de traduções, feitas por sistemas de Tradução Automática é a obtenção de um programa de computador, que de forma rápida, mostre a resposta sobre o êxito ou insucesso do sistema após algum ajuste ter sido realizado.

Para a avaliação realizada por humanos, usar somente o critério ‘certo x errado’ parece ser muito vago. Normalmente, utiliza-se dois critérios: o de fluência e adequação, com notas entre um e cinco, onde um é a menor pontuação e cinco a maior. Como podemos verificar no Quadro 6, essa pontuação funciona como uma ferramenta para auxiliar os anotadores (avaliadores) nessa tarefa.

Quadro 6 – Ferramentas de avaliação para anotadores humanos para julgamento de adequação e fluência

Adequacy (Adequação)		Fluency (Fluência)	
5	<i>all meaning</i> (todo o significado)	5	<i>flawless English</i> (inglês impecável)
4	<i>most meaning</i> (a maior parte do significado)	4	<i>good English</i> (inglês bom)
3	<i>much meaning</i> (muito do significado)	3	<i>non-native English</i> (inglês não nativo)
2	<i>little meaning</i> (pouco do significado)	2	<i>disfluent English</i> (inglês não fluente)
1	<i>None</i> (nada)	1	<i>Incomprehensible</i> (incompreensível)

Fonte: Koehn (2010, p. 219).

No critério da fluência se verifica se a tradução apresenta a LA de forma fluente e boa, envolvendo tanto a correção da gramática, quanto as escolhas de palavras idiomáticas. Já no da adequação se observa se a tradução transmite o mesmo significado que o texto de origem, ou se alguma parte da mensagem foi perdida, acrescentada ou distorcida.

Contudo, essas definições ainda se mostram muito vagas e subjetivas, pois um anotador pode ser mais condescendente ou rigoroso que outro, ou ainda, manter a consistência de seu julgamento para todas as sentenças analisadas é uma tarefa árdua. Ao invés de julgar utilizando esses critérios, a tendência é ranquear dois ou três sistemas, comparando-os entre si, sentença por sentença. Desta forma, uma resposta mais consistente e justa pode ser dada pelos anotadores humanos: “se a tradução do sistema A é melhor, pior ou indistinguível a do sistema B” (KOEHN, 2010, p. 220).

Os objetivos para se usar métricas automáticas de avaliação são: o baixo custo, ser ajustável, ter significância, a consistência, a concordância entre anotadores, estabilidade e a correção. Outros critérios, ainda levados em conta, são de avaliação da velocidade, o tamanho do sistema, a integração e a customização.

Ainda segundo Koehn (2010, p. 222), todas as métricas automáticas de avaliação usam o mesmo truque – cada tradução elaborada automaticamente é comparada com uma ou mais traduções humanas da mesma sentença, ou seja, as traduções de referência, a qual explicaremos melhor no item 3.3.3.2. Apesar da improbabilidade de que duas traduções sejam exatamente iguais ao serem comparadas, sejam elas feitas por humanos ou sistemas automáticos de tradução, a lógica é que uma TA muito similar a uma tradução humana tem a maior possibilidade de estar correta do que uma substancialmente diferente. O desafio para as métricas de avaliação se encontra, justamente, numa boa medida de similaridade.

Existem diferentes métricas desenvolvidas para avaliar a qualidade de traduções. Utilizamos, para avaliar as traduções produzidas pelos três TAO's e THP's, duas que também foram utilizadas por Cunha (2013b), *F-measure* e BLEU, e acrescentamos a métrica METEOR, as quais descrevemos a seguir.

2.2.4.1 Métrica *F-Measure*

Essa métrica foi desenvolvida pela Universidade de Nova Iorque e associa conjuntos de coocorrências entre traduções candidatas e de referência. Conhecida também como *F1 score* ou *F score*. Ela representa a média harmônica de *precision*³⁰ (precisão) e *recall* (cobertura), em que o melhor resultado alcançado seria 1 e o pior 0.

Para ilustrar o que significam os valores de *precision* e *recall*, traz-se um exemplo de Koehn (2010, p. 223):

SISTEMA A: Israeli officials ~~responsibility of~~ airport safety
 REFERÊNCIA: Israeli officials ~~are responsible for~~ airport security
 SISTEMA B: airport security ~~Israeli officials are~~ responsible

Resultado:

Métrica	SISTEMA A	SISTEMA B
<i>Precision</i>	50%	100%
<i>Recall</i>	43%	86%
<i>f-measure</i>	46%	92%

Em um sistema de avaliação, baseado em 'correspondência de palavras', temos a tradução do SISTEMA A, a tradução de referência e a tradução do SISTEMA B. O SISTEMA A traz uma tradução com seis palavras, onde três delas (50%) correspondem às palavras da tradução de referência. O SISTEMA B oferece uma tradução com seis palavras também e todas elas (100%) têm uma correspondente na tradução de referência. Esse tipo de métrica chama-se *precision*, ou seja, a correspondência entre o número de palavras da tradução candidata comparada à tradução de referência.

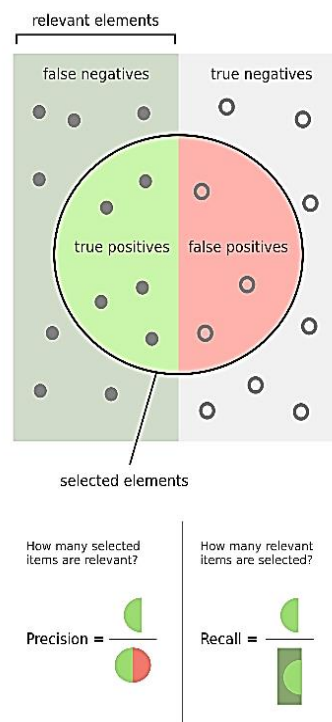
³⁰ Utilizaremos os termos *precision* e *recall* em língua inglesa, para evitar ambiguidades com seus referentes em português.

Contudo, podemos perceber que mesmo a tradução do SISTEMA B, que tem 100% de *precision*, ainda traz problemas em relação à tradução de referência, tais como a ordem das palavras e o fato de a palavra *for* não aparecer na tradução produzida pelo SISTEMA B. Tudo isso deve ser levado em consideração quando se mede a acurácia de uma tradução. Portanto, ao invés de computar quantas palavras corretas um sistema produz, deve-se medir quantas das palavras, que um sistema deve gerar, estão corretas. Essa métrica é chamada de *recall*. Temos assim, as seguintes fórmulas:

$$precision = \frac{correct}{output - length} \quad recall = \frac{correct}{reference - length}$$

Outra forma de explicar os valores de *precision* e *recall* associa-se à noção de relevância, ou seja, em relação a valores positivos e negativos: *precision* (valor positivo previsto) é, então, a fração de exemplos relevantes, dentre todos os exemplos gerados, enquanto *recall* (sensibilidade), mostra a fração dos exemplos relevantes, que foram gerados sobre o valor total dos exemplos relevantes.

Figura 5 – Precision e recall



Fonte: <https://en.wikipedia.org/wiki/Precision_and_recall>.

Nesse contexto, temos as seguintes fórmulas:

$$precision = \frac{TP}{(TP+FP)} \quad recall = \frac{TP}{(TP+FN)}$$

Legenda:

TP (*true positives*) – verdadeiros positivos

FP (*false positives*) – falsos positivos

TN (*true negatives*) – verdadeiros negativos

FN (*false negatives*) – falsos negativos

Para a avaliação de Tradução Automática, os verdadeiros positivos são as palavras das sentenças de referência, que também apareceram nas sentenças candidatas (itens que julgamos relevantes e que são corretamente identificados como relevantes). Os falsos positivos representam um erro, simbolizados pelas palavras que apareceram nas sentenças candidatas, mas não estão nas sentenças de referência (itens irrelevantes, que são incorretamente identificados como relevantes). Os falsos negativos também representam erro, pois indicam palavras da tradução de referência, que não aparecem nas traduções candidatas (itens relevantes, que são incorretamente identificados como irrelevantes). Os verdadeiros negativos mostram as palavras que não aparecem nem na tradução de referência e nem na candidata (itens que são corretamente considerados como irrelevantes).

Pode-se, deliberadamente, manipular os valores de *precision* e *recall*, contudo, se traduzirmos somente as palavras que conhecemos, o resultado será mais curto, com alto valor de *precision* e baixo valor de *recall*. Ou o oposto, se traduzirmos todos os tipos de palavras, aumentando a chance de produzir correspondentes com a tradução de referência, mas com baixa *precision*. Essas são métricas importantes em PLN, e em algumas aplicações, uma pode ser mais importante do que a outra, mas tratando-se de Tradução Automática, ambas são tipicamente importantes – “não queremos produzir palavras erradas, mas também não queremos deixar nenhuma informação de fora” (KOEHN, 2010, p. 224). Uma forma para combiná-los é a métrica *F-measure*:

$$f\text{-measure} = \frac{precision \times recall}{(precision + recall)/2}$$

Ou reformulada como:

$$f\text{-measure} = \frac{\text{correct}}{(\text{output} - \text{length} + \text{reference} - \text{length})/2}$$

Os resultados gerados pelo programa mostram os valores de *precision*, *recall* e o valor *F-measure* desejado.

Apesar de ser importante, essa métrica ainda não resolve todos os problemas relacionados à avaliação de Tradução Automática. Um deles é a questão da ordem das palavras e outro, mais importante para nós, é a tradução correta dos significados. A seguir, apresentaremos outra métrica reconhecida para a avaliação de Tradução Automática, a BLEU.

2.2.4.2 Métrica BLEU

BLEU indica um acrônimo para *Bilingual Evaluation Understudy* e mostra uma das métricas mais utilizadas para a avaliação de Tradução Automática. Ela pode ser definida da seguinte maneira:

BLEU é uma métrica desenvolvida pela *International Business Machines* – IBM, com a finalidade de avaliar sistemas de tradução automática com maior economia, rapidez e independência de linguagens do que avaliações realizadas manualmente. É fundamentada na proximidade entre a tradução automática e a tradução realizada por uma pessoa qualificada em traduções profissionais, chamada de tradução de referência. MELO, MATOS e DIAS (2014, p. 36)

De acordo com Koehn (2010, p. 226), esta métrica é uma solução para a questão da ordem das palavras e, de forma similar, para o índice de erros de palavras com posição independente, pois considera a combinação de n-gramas maiores com a tradução de referência. N-gramas são as palavras que compõem as sentenças candidatas e de referência³¹. Um unigrama é uma única palavra; bigrama corresponde

³¹ Explicando de forma geral, as sentenças candidatas são aquelas elaboradas pelos TAO's e que serão analisadas pelas métricas de tradução. Sentença de referência é aquela considerada como o melhor padrão (*gold standard*) e é utilizada para ser comparada às sentenças candidatas. Dependendo da métrica, pode-se usar mais de uma sentença de referência. Elas serão melhor explicadas na seção de número 3.

a duas, trígama à três e assim sucessivamente. O quadrigrama (BLEU-4) é o maior número sugerido para uso, mas nada impede que um número maior seja usado.

Essa implementação de avaliação de algoritmo de Tradução Automática pode ser executada por meio de um *script Unix shell* feito a partir do terminal de computador. Esse *script* também aplica automaticamente as medidas nas traduções geradas e de referência (CUNHA, 2013b).

Ainda segundo Koehn (2010), a métrica BLEU introduz o conceito de ‘penalidade por brevidade’ para tentar solucionar a questão da falta de penalidade para as traduções, que omitem ou diminuem o número de palavras dos textos traduzidos. A penalidade reduz a pontuação se o produto tradutório for muito curto. Essa métrica é, tipicamente, analisada sobre todo o *corpus*, e não sentenças isoladas.

Se utilizarmos o mesmo exemplo, que apresentamos para a métrica *F-measure*, temos a seguinte análise:

SISTEMA A: Israeli officials responsibility of airport safety
 2-GRAM MATCH 1-GRAM MATCH

REFERÊNCIA: *Israeli officials are responsible for airport security*

SISTEMA B: airport security Israeli officials are responsible
 2-GRAM MATCH 4-GRAM MATCH

A penalidade por brevidade funciona da seguinte forma: realiza-se uma comparação entre o número de palavras da sentença candidata do SISTEMA A (seis palavras) e da sentença de referência (sete palavras), para verificar se o número de n-gramas é similar. Neste caso, o SISTEMA A tem uma palavra a menos do que a sentença de referência. O mesmo se realiza para o SISTEMA B. Coincidentemente, neste caso, o SISTEMA B também elaborou uma tradução com seis palavras. Portanto, tanto o SISTEMA A quanto o SISTEMA B perdem pontuação, porque suas traduções são um pouco menores do que a sentença de referência (um n-grama a menos). Contudo, apesar de o número de palavras destas traduções candidatas ser bem próximo ao da tradução de referência, os vocábulos não são necessariamente os mesmos e nem estão dispostos na mesma sequência da sentença de referência. É feita, então, a associação dos n-gramas para verificar o número de acertos dentro das possibilidades de combinações das sentenças candidatas em comparação à(s)

de referência(s):

Quadro 7 – Possibilidades de combinações de n-gramas no exemplo dado

Unigramas	
REFERÊNCIA	<i>Israeli / officials / are / responsible / for / airport / security</i>
SISTEMA A (3/6)	Israeli / officials / responsibility / of / airport / safety
SISTEMA B (6/6)	airport / security / Israeli / officials / are / responsible
Bigramas	
REFERÊNCIA	<i>Israeli officials / officials are / are responsible / responsible for / for airport / airport security</i>
SISTEMA A (1/5)	Israeli officials / officials responsibility / responsibility of / of airport / airport safety
SISTEMA B (4/5)	airport security / security Israeli / Israeli officials / officials are / are responsible
Trigramas	
REFERÊNCIA	<i>Israeli officials are / officials are responsible / are responsible for / responsible for airport / for airport security</i>
SISTEMA A (0/4)	<i>Israeli officials responsibility / officials responsibility of / responsibility of airport / of airport safety</i>
SISTEMA B (2/4)	airport security Israeli / security Israeli officials / Israeli officials are / officials are responsible
Quadrigramas	
REFERÊNCIA	<i>Israeli officials are responsible / officials are responsible for / are responsible for airport / responsible for airport security</i>
SISTEMA A (0/3)	<i>Israeli officials responsibility of / officials responsibility of airport / responsibility of airport safety</i>
SISTEMA B (1/3)	airport security Israeli officials / security Israeli officials are / Israeli officials are responsible

Fonte: Elaborado pela autora.

Os n-gramas em negrito para os sistemas A e B representam aqueles que encontram um par exato (*match*) de n-gramas na sentença de referência. Dentro do número de palavras geradas nas sentenças candidatas se faz combinações de um a quatro n-gramas, de acordo com a ordem dos mesmos n-gramas na sentença de referência. Por exemplo, para os unigramas, há sete n-gramas (palavras) na sentença de referência e seis unigramas gerados tanto na tradução candidata do SISTEMA A quanto do SISTEMA B. Das seis possibilidades de acerto do SISTEMA A, três foram exitosas (**Israeli/officials/responsibility/of/airport/safety**). Já do SISTEMA B, todas foram exitosas, ou seja, todas têm um par exato, na sentença de referência (**airport/security/Israeli/officials/are/responsible**). Para os bigramas, há a possibilidade de cinco acertos. O SISTEMA A obteve apenas um (**Israeli officials/officials responsibility/responsibility of/of airport/airport safety**), enquanto que o SISTEMA B obteve quatro (**airport security/security Israeli/Israeli officials/**

officials are/are responsible). O mesmo processo se realiza para trigramas e quadrigramas. O resumo desses resultados se apresenta de duas formas:

- a) O número de acertos de cada sentença, incluindo a penalidade por brevidade e o valor geral obtido por cada sistema:

Métrica	Sistema A	Sistema B
<i>precision</i> (1gram)	3/6	6/6
<i>precision</i> (2gram)	1/5	4/5
<i>precision</i> (3gram)	0/4	2/4
<i>precision</i> (4gram)	0/3	1/3
<i>brevity penalty</i>	6/7	6/7
BLEU	0%	52%

- b) O escore em percentual obtido para cada sentença dos dois sistemas:

Métrica	Sistema A	Sistema B
BLEU-1	42%	86%
BLEU-2	9%	69%
BLEU-3	0%	34%
BLEU-4	0%	11%

Por meio da métrica BLEU, podemos verificar que o SISTEMA B obteve melhores resultados do que o SISTEMA A para todos os n-gramas possíveis. Indicando que o SISTEMA B produziu uma tradução candidata bem mais semelhante à tradução de referência e que o SISTEMA A precisa de ajustes para obter resultados mais próximos aos da tradução de referência.

Outra inovação da métrica BLEU é a possibilidade de se usar múltiplas traduções de referência, o que aumenta as chances de se encontrar traduções aceitáveis. Na avaliação aqui proposta, teremos apenas uma tradução de referência (aquela elaborada por nós), mas múltiplas traduções candidatas (aquelas elaboradas pelos três TAO's e pelos três THP's).

2.2.4.3 Métrica *METEOR*

O METEOR é o acrônimo para *Metric for Evaluation of Translation with Explicit Ordering*. Métrica desenvolvida pela Universidade Carnegie Mellon que promete corrigir algumas fraquezas presentes em outras métricas, como a BLEU, uma vez que demonstra ter altos níveis de correlação com os julgamentos humanos de qualidade de tradução (LAVIE; AGARWAL, 2007). Uma de suas vantagens, segundo

os autores, é que os resultados podem ser melhor vistos em nível de sentenças (segmentos).

O METEOR avalia uma tradução computando uma nota, que se baseia em associações corretas explícitas, palavra-por-palavra, entre a tradução candidata e a de referência, podendo haver mais de uma tradução de referência (Figura 6). A caixa verde com o ponto escuro indica a semelhança entre as palavras e a sua posição na sentença. A caixa amarela, com um círculo não preenchido, indica que as palavras são sinônimas (*related* e *linked*). Quando não há qualquer tipo de correspondência, a caixa fica em branco. Apresentam-se ainda os resultados de *precision* (P), *recall* (R), penalidade por fragmentação e o resultado final da métrica.

Figura 6 – Exemplo da métrica METEOR

	these	include	activities	linked	to	energy	and	,	in	particular	,	energy	efficiency	.
these	●													
are		○												
the														
activities			●											
related				○										
to					●									
energy						●								
,											●			
and							●							
in								●						
particular									●					
to														
energy												●		
efficiency													●	
.														●

Segment 2022

P: 0.897

R: 0.907

Frag: 0.514

Score: 0.440

Fonte: <<http://www.cs.cmu.edu/~alavie/METEOR/examples.html>>.

Dado um par de sentenças, essa métrica cria um alinhamento (mapeamento entre palavras) entre as sentenças, de forma que cada palavra de uma sentença só pode ser mapeada com o máximo de uma palavra na outra sentença. A partir desse alinhamento, uma sequência de módulos produz nesta ordem (*default* da métrica): módulo *exact* (mapeia duas palavras, que são exatamente iguais), módulo *porter stemmer* (mapeia duas palavras, que se tornam idênticas após o uso do algoritmo *Porter stemmer*) e módulo *WN synonymy* (mapeiam duas palavras se elas

são consideradas sinônimas, ou seja, pertencendo ao mesmo *synset*³² na *WordNet*®³³). No exemplo supracitado, as palavras *linked* e *related* são consideradas sinônimas.

Os módulos de mapeamento de palavras identificam todas as associações corretas possíveis entre as duas sentenças e depois mostram o maior subconjunto desses mapeamentos, de forma que o conjunto resultante constitui um alinhamento como definimos acima. Se houver mais de um alinhamento cardinal máximo, a métrica escolhe aquele em que a ordem de palavras, entre os pares de sentenças, for a mais próxima possível.

A métrica computa a pontuação usando a seguinte fórmula:

$$F_{mean} = \frac{P \cdot N}{\alpha \cdot P + (1 - \alpha) \cdot R}$$

Baseando-se no número de unigramas mapeados encontrados entre as duas sentenças, o total de unigramas na tradução e o total de unigramas na sentença de referência, calcula-se *precision* (P) e *recall* (R). Posteriormente, calcula-se uma média harmônica parametrizada de P e R. Há preferência por *recall* sobre *precision*; de escolha de palavras sobre a ordem das palavras; e por palavras de conteúdo corretas sobre palavras de ligação corretas.

Precision, *recall* e a média harmônica se baseiam em associações corretas de uma palavra. Para dar conta da extensão na qual os unigramas corretos estão na mesma ordem nas duas sentenças, a métrica computa uma penalidade por fragmentação, na qual a sequência de unigramas com associações corretas entre as duas sentenças se divide no maior número possível de segmentos (*chunks*), de modo que os unigramas corretos, em cada segmento, estejam adjacentes em ambas as sentenças e em ordem idêntica de palavras. Com o número de segmentos (ch) e o número de acertos (m) se calcula a fração de fragmentação $\text{frag} = \text{ch}/m$. A penalidade é então computada com a seguinte fórmula:

³² De acordo com o *WordNet*, *synset* é um grupo de um ou mais sinônimos que podem ser utilizados, de forma intercambiável em determinados contextos, sem o alterar o valor da proposição em que estão inseridos.

³³ *WordNet*® é uma base de dados lexicais da língua inglesa cuja estrutura a torna em uma ferramenta útil, de acesso livre, para a linguística computacional e PLN. Disponível em: Princeton University "About WordNet." *WordNet*. Princeton University. 2010 < <http://wordnet.princeton.edu>>. Acesso em: 08 jan. 2018.

$$Pen = \gamma \cdot frag^{\beta}$$

O valor de γ determina a penalidade máxima ($0 \leq \gamma \leq 1$). O valor de β determina a relação funcional entre a fragmentação e a penalidade. Calcula-se, ao final, a pontuação da métrica para o par de sentenças:

$$score = (1 - Pen) \cdot F_{mean}$$

A versão utilizada para nossa pesquisa é a mais atual, Version 1.5 (WMT 2014)³⁴, que permite alinhamentos baseados em associações exatas, *stem* (radicais), *synonym* (sinônimos) e *paráfrase* (paráfrase) entre as palavras e frases. Contudo, alguns parâmetros, que podem ser ajustados para emular várias tarefas de julgamento humano, só se utilizam para algumas línguas (inglês, tcheco, alemão, francês, espanhol e árabe). O português está na lista de línguas que têm, apenas, algumas das funções disponíveis, conforme podemos ver nos Quadro 8 e 9:

Quadro 8 – Línguas com suporte total da métrica METEOR

Language	Exact Match	Stem Match	Synonym Match	Paraphrase Match	Tuned Parameters
English	Yes	Yes	Yes	Yes	Yes
Arabic	Yes	No	No	Yes	Yes
Czech	Yes	No	No	Yes	Yes
French	Yes	Yes	No	Yes	Yes
German	Yes	Yes	No	Yes	Yes
Spanish	Yes	Yes	No	Yes	Yes

Fonte: <<http://www.cs.cmu.edu/~alavie/METEOR/README.html>>.

³⁴ Disponível em: <<http://www.cs.cmu.edu/~alavie/METEOR/index.html#About>>. Acesso em: 14 dez. 2017.

Quadro 9 – Línguas com suporte parcial da métrica METEOR

Language	Exact Match	Stem Match	Synonym Match	Paraphrase Match	Tuned Parameters
Danish	Yes	Yes	No	No	LI
Dutch	Yes	Yes	No	No	LI
Finnish	Yes	Yes	No	No	LI
Hungarian	Yes	Yes	No	No	LI
Italian	Yes	Yes	No	No	LI
Norwegian	Yes	Yes	No	No	LI
Portuguese	Yes	Yes	No	No	LI
Romanian	Yes	Yes	No	No	LI
Russian	Yes	Yes	No	No	LI
Swedish	Yes	Yes	No	No	LI
Turkish	Yes	Yes	No	No	LI

Fonte: <<http://www.cs.cmu.edu/~alavie/METEOR/README.html>>.

As línguas que têm suporte total incluem associações corretas flexíveis de palavras e frases (pelo menos um tipo de associação correta, que não seja a exata – *stem*, *synonym* ou *paraphrase*) e parâmetros específicos para línguas ajustados para maximizar a correlação entre a pontuação do METEOR e julgamentos humanos de qualidade de tradução. As línguas, cujo suporte é parcial, incluem associações de palavras flexíveis (*stem*) e usam parâmetros independentes da língua (LI) escolhidos.

A métrica METEOR se implementa em Java e não requer instalação ou dependências para avaliar as traduções elaboradas.

Nesta seção, descrevemos a fundamentação teórica que norteou este trabalho, bem como, revisão literária das áreas que o fundamentam dentro das duas grandes áreas envolvidas, Linguística Cognitiva e Linguística Computacional. A próxima seção trará todos os detalhes sobre o percurso metodológico da pesquisa em questão.

3 PERCURSO METODOLÓGICO – *MAKING NO BONES ABOUT IT*



Descreveremos nesta seção o percurso metodológico traçado para a coleta e análise de dados desta pesquisa, assim como os aspectos metodológicos que a nortearam, tais como o tipo de pesquisa, seu contexto e as etapas de elaboração dos *corpora*, descrição dos TAO's analisados e coleta dos dados. Por último, são descritos os procedimentos que foram utilizados na análise dos dados coletados.

3.1 TIPO DE PESQUISA

A presente pesquisa é de caráter exploratório-descritivo de base quali-quantitativa, baseada em *corpora*, na modalidade aplicada. Sua análise qualitativa descritiva foi realizada através da identificação das metáforas conceituais subjacentes às EI, tendo sido necessária pesquisa bibliográfica para averiguar o estado da arte dos diversos campos de estudo envolvidos. Há também uma vertente experimental quanto à percepção dos TAO's em relação às metáforas conceituais subjacentes às EI's em suas traduções. O produto tradutório desse experimento gerou as sentenças que foram analisadas por três métricas reconhecidas como apropriadas para a avaliação de traduções. A partir dos resultados obtidos pelas métricas de avaliação de tradução, pudemos elaborar as análises quantitativas. Neste estudo há apenas uma variável independente: o tipo de tradutor, ou seja, se a tradução foi elaborada por um TAO ou por um THP. A Linguística de *Corpus* foi fundamental para a construção das sentenças candidatas e de referência que foram analisadas. Procurou-se elaborar um desenho metodológico que pudesse responder às questões e hipóteses levantadas, a fim de alcançar os objetivos pretendidos.

3.2 CONTEXTO DA PESQUISA

Esta pesquisa de doutorado está vinculada ao Programa de Pós-Graduação em Linguística Aplicada (Pos-LA) da UECE e se enquadra entre os estudos de metáforas conceituais, fraseologismos, estudos tradutórios, Linguística de *Corpus* e Linguística Computacional.

Com a finalidade de construir o embasamento teórico, pudemos encontrar estudos relacionados a metáforas com diferentes objetivos: a) sobre metáforas conceituais primárias e seus mapeamentos – dentre esses, alguns sobre a identificação das metáforas relacionadas ao ensino, à sala de aula, ao aprendizado, etc, como por exemplo, parte de um processo reflexivo que ajude os professores a reconceitualizar seus processos de ensino e aprendizagem pela autocrítica e pela apreciação de perspectivas múltiplas – e b) estudos comparativos entre metáforas comuns em diferentes culturas (e línguas) que têm o potencial de revelar como nossa experiência corporificada e nossa neurobiologia determinam nossa visão global e nosso uso da língua (Tam, 2009). Estes são os que nos interessam.

Em relação às pesquisas sobre Tradução Automática, encontramos algumas sob diferentes perspectivas. Algumas são estudos sobre a acurácia nas traduções de determinadas partes da oração ou relação entre palavras, outras são estudos comparativos para verificar a qualidade da tradução entre dois ou mais TAO's utilizando uma ou mais métricas de avaliação, mas nenhuma relacionada à Tradução Automática de El's com alimentos em sua composição lexical.

Antes de detalharmos os procedimentos metodológicos utilizados nesta pesquisa, descreveremos os três TAO's analisados. Esse relato é relevante pois registra como cada um deles se caracterizava no momento em que o estudo foi realizado. É notório que as mudanças no mundo da informática são rápidas e nem sempre perceptíveis ao olho do usuário casual. Durante o período de elaboração desta tese, pudemos registrar diversas atualizações e modificações nas ferramentas e TAO's utilizados que interferiram diretamente em seu resultado.

3.2.1 Tradutores Automáticos *Online*

O objetivo principal da TA é traduzir textos escritos de uma língua para outra, existindo também a possibilidade de a tradução ser feita de uma modalidade

para outra. Por exemplo, texto oral para texto escrito e vice-versa. A modalidade de Tradução Automática que analisamos nesta pesquisa foi a de textos escritos entre dois idiomas diferentes – os TAO's e THP's traduziram excertos de textos escritos, do inglês para o português. Os TAO's escolhidos foram o *Bing Tradutor*, o *Google Translate* e o SYSTRAN (sempre citados em ordem alfabética). A escolha desses TAO's é justificada pelo fato de todos disponibilizarem vários pares de línguas, incluindo os de nosso interesse para esta pesquisa, além da facilidade e gratuidade de seu acesso, a importância histórica do SYSTRAN e a expressividade do número de traduções do *Google Translate*. Outro fator para a escolha destes sistemas de tradução automática *online* foi a diferença no tipo de técnica utilizada por cada um³⁵. O *Bing Tradutor* (ou *Microsoft Translator*) e o *Google Translate*, recentemente, iniciaram o uso de tecnologia de rede neural. Contudo, o SYSTRAN é, atualmente, um sistema híbrido que combina técnicas baseadas em regras e estatísticas. Ambos o *Bing Tradutor* e o *Google Translate* oferecem possibilidade de colaboração por parte dos usuários para dar *feedback* sobre as traduções realizadas.

Atualmente existem vários TAO's disponíveis (*Bing Tradutor*, *Google Translate*, SYSTRAN, *Babylon*, *World Lingo*, *Babel Fish*, *Lingoes* etc). Alguns são pagos, mas também há os de acesso livre, a principal diferença entre os dois está no suporte para correções que é disponibilizado pelos tradutores pagos, resultando em uma melhor qualidade do produto tradutório. Alguns são grátis até um determinado número de palavras traduzidas dentro de um período específico. O número de línguas a que cada um dá suporte varia bastante. A língua portuguesa falada no Brasil está inclusa em muitos desses *software*, que são uma opção cada vez mais utilizada para traduções rápidas de textos e também de páginas acessadas pela *Internet*.

Conforme mencionado anteriormente, escolhemos para avaliar três desses programas – o *Bing Tradutor*, o *Google Translate* e o SYSTRAN – por razões também já citadas, tais como todos serem de acesso livre, sua popularidade entre os usuários, o fato de eles utilizarem paradigmas de tradução diferentes e, de modo mais específico, a importância histórica do SYSTRAN e a expressividade do número de traduções realizadas pelo *Google Translate* diariamente. Descrevemos melhor cada um deles a seguir.

³⁵ Quando iniciamos esta pesquisa, as técnicas utilizadas por cada um dos TAO's eram diferentes, conforme será melhor descrito no item 2.2.4 Tradutores Automáticos *Online* e em seus subitens.

3.2.1.1 *Bing* Tradutor³⁶

O *Bing* Tradutor é um portal de TAO provido pela *Microsoft*, que, atualmente (desde o final de 2016), dá suporte a mais de 60 línguas para tradução de textos; 10 para tradução de fala; 18 de texto para fala; e 21 com tradução neural. Dentre as línguas disponíveis, algumas têm suporte do *Cognitive Services Custom Speech Service* (Serviço de Fala Personalizado dos Serviços Cognitivos), que permite aos usuários personalizar o modelo de língua do reconhecimento de fala, customizando-o ao seu vocabulário e estilo de falar. Há ainda suporte para que seja realizado o *download* com um pacote de muitas dessas línguas (44 para os sistemas *iOS*, *Android* e *Amazon Fire* e 19 para o *Windows*) para serem trabalhadas *off-line* em situações nas quais não há acesso à Internet ou o custo desse acesso é muito alto, tais como em viagens internacionais. Para telefones celulares, há suporte para a tradução de imagens em 20 línguas, também utilizando os sistemas acima citados. Os pares de línguas traduzidas são alimentados pelo *Microsoft Translator Text API*, um serviço de Tradução Automática em nuvem, que faz parte da coleção *Microsoft Cognitive Services API*. O português do Brasil está incluído em todas as opções disponíveis atualmente.

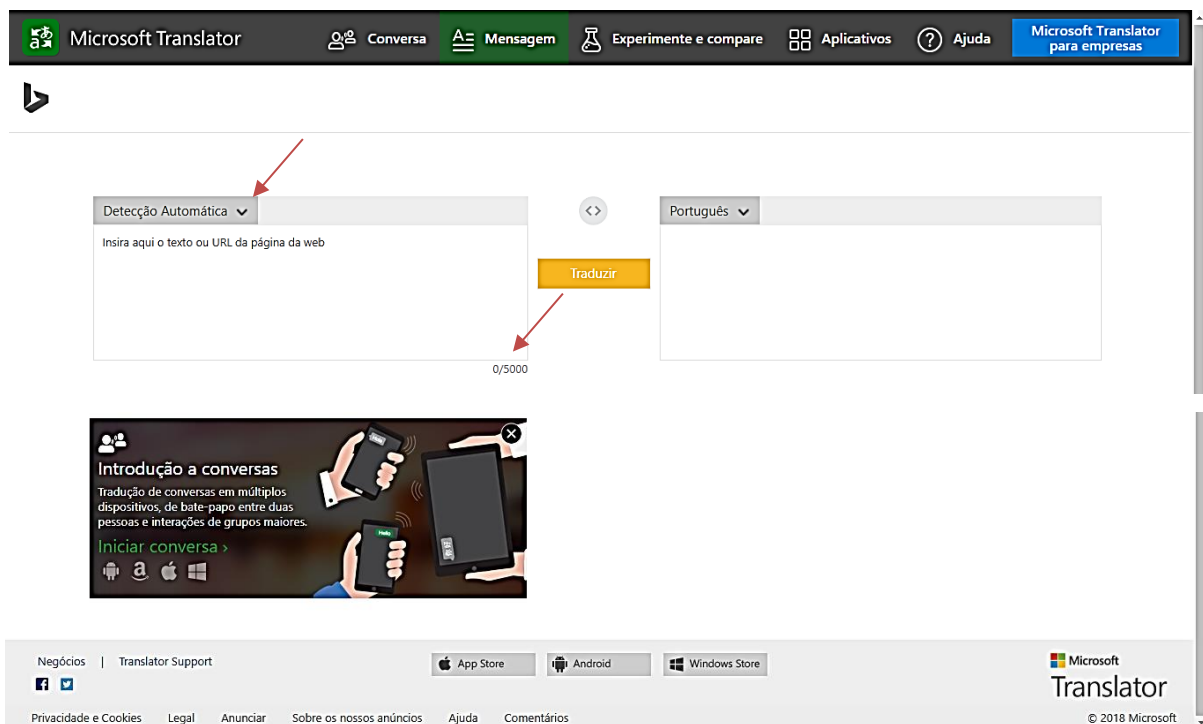
Em relação à técnica utilizada, o *Microsoft Translator* é um sistema híbrido que une a força dos métodos estatísticos com a informação linguística para produzir modelos que fazem melhores generalizações, levando a traduções mais compreensíveis. Esse método não prescinde de dicionários ou regras gramaticais, mas utiliza o contexto em que uma palavra se encontra, ao invés de fazer traduções de cada palavra individualmente. Essa é uma característica que, teoricamente, deve favorecer a tradução de *El's*. O *Bing* sugere que traduções de palavras isoladas sejam elaboradas a partir de dicionários bilíngues desenvolvidos, que podem ser acessados em <http://www.bing.com/translator>.

O *Bing* Tradutor permite inserir até 5000 caracteres em texto digitado pelo usuário em sua interface (Figura 7) ou acessar um *link* para uma página e traduzi-la

³⁶ Todas as informações sobre esse TAO foram retiradas de sua página oficial, a partir dos sites: <http://www.microsofttranslator.com> ou <http://www.bing.com/translator>. Foram realizados diferentes acessos em diferentes datas ao longo da pesquisa. O acesso mais recente foi realizado em 26 de janeiro de 2018.

por inteiro. Nesse caso, deve-se inserir o *URL* da página da *Web*. São disponibilizados quatro *layouts* para o Visualizador Bilíngue: lado a lado, um acima do outro, original com foco (flutuando) na tradução e tradução com foco (flutuando) no original.

Figura 7 – Interface do usuário do *Bing* Tradutor



Fonte: <<http://www.microsofttranslator.com>> ou <<http://www.bing.com/translator>>.

Logo que o texto de origem é inserido, a função de reconhecimento automático do idioma (Detecção Automática) é ativada e a tela já mostra o par de línguas que será utilizado, bem como outros ícones (Figura 7): as funções de áudio no par de línguas, com a opção de voz (masculina ou feminina) e variação da língua, quando aplicável. Por exemplo, para o inglês, há as opções Estados Unidos, Austrália, Canadá, Reino Unido e Índia. Já para o português, Brasil e Portugal; há a possibilidade de enviar o resultado via redes sociais (*Facebook*, *Twitter*, *LinkedIn*) ou e-mail; há a opção de buscar resultados na *Web* para o texto traduzido; e dar o *feedback* to texto traduzido através dos ícones 👍 (esta é uma boa tradução) ou 👎 (relatar a tradução como incorreta). O usuário também pode percorrer o *mouse* sobre as palavras em ambas as línguas. Ao passar o *mouse* sobre as palavras do texto original ou do traduzido, as palavras equivalentes são marcadas/sombreadas e são mostrados outros significados para cada uma, quando possível, incluindo uma barra de

frequência de uso. Na Figura 8, por exemplo, vemos os possíveis significados para o verbo *get*. As barras que mostram a frequência de cada significado evidenciam a técnica estatística utilizada por esse tradutor automático³⁷.

Figura 8 – Bing Tradutor com a tradução para *a piece of cake* em contexto

The screenshot shows the Microsoft Translator interface. The source text in English is "Mary told me her Math test was a piece of cake. She'll get an A." and the target text in Portuguese is "A Mary disse-me que o teste de matemática dela foi canja. Ela vai ter um a." Below the translation, there is a section for "Alternative Translations" for the word "get".

Verbs	Portuguese	English
obter	get, obtain, getting, gain, achieve	
começ	get, getting, got, gets, gotten	
conseguir	get, achieve, getting, able, manage, succeed	
chegar	get, reach, arrive, come, coming, getting, comes	
pegar	get, take, pick up, catch, grab, taking	
receber	receive, get, getting, welcome, collect	
ficar	stay, be, get, keep, getting, staying, stand	
tirar	take, get, taking, draw, remove, pull	
levar	take, get, lead, bring, taking, carry, drive	
ter	have, having, be, 've, get, take, make	
se	if, get	

Fonte: <<http://www.microsofttranslator.com>>ou <<http://www.bing.com/translator>>.

É possível fazer uma comparação entre os dois tipos de tradução, atualmente, disponíveis no *Bing Tradutor* (estatística e neural) utilizando a barra de ferramentas encontrada no topo da página. Ao clicar em 'Experimente e Compare', uma nova janela é aberta com a seguinte configuração (Figura 9):

³⁷ A fim de mostrar o estado da arte dos TAO's analisados nesta pesquisa, a interface do usuário disponibilizada nesta figura é a de janeiro de 2018. Devemos salientar que, uma vez que a técnica utilizada pelo TAO em questão foi modificada, a tradução atual difere daquela elaborada no pequeno teste preliminar realizado e que foi apresentada na Introdução desta tese (p. 22). O resultado agora traz uma tradução excelente, pois, além da *a piece of cake* ter sido reconhecida pelo tradutor na LF, seu resultado também apresenta um alimento na LA.

Figura 9 – Página de comparação entre os dois tipos de tradução disponíveis no *Bing Tradutor*

The screenshot displays the Microsoft Translator 'Try & Compare' interface. At the top, the navigation bar includes 'Microsoft Translator', 'Conversation', 'Text', 'Try & Compare' (highlighted with a red arrow), 'Apps', 'Help', and 'Microsoft Translator for Business'. The main heading is 'Try & Compare' with a flask icon and the text 'Artificial Intelligence, powered by neural networks'. A text input field is set to 'English' and contains the sentence: 'With a minimal amount of planning and practice, you can get good at it, even if you are a relatively novice cook|'. Below the input is a green 'Translate & Compare!' button. Underneath, two empty text boxes for Portuguese translations are shown, each with a '✓ Is this better?' prompt. A footer contains social media links, help, and terms of use.

Fonte: <<https://translator.microsoft.com/neural>>.

Uma caixa de texto é aberta para que o excerto em LF seja inserido e, logo abaixo, surgem duas caixas de texto menores, nas quais são apresentadas as duas traduções elaboradas, com a opção '*Is this better?*', para que o usuário possa escolher qual resultado lhe parece melhor. Feita a escolha, automaticamente, o tradutor revela a técnica de tradução para cada uma das traduções apresentadas. Fizemos alguns testes com nossos excertos e pudemos perceber que as traduções de nossa segunda coleta de dados foram todas realizadas com a técnica neural, mas nem sempre os resultados para as EI's foram melhores que aqueles elaborados com a técnica estatística da primeira coleta. Contudo, a fluidez do texto nos pareceu sempre melhor, se levarmos em conta outros aspectos. Apresentamos um exemplo na Figura 10.

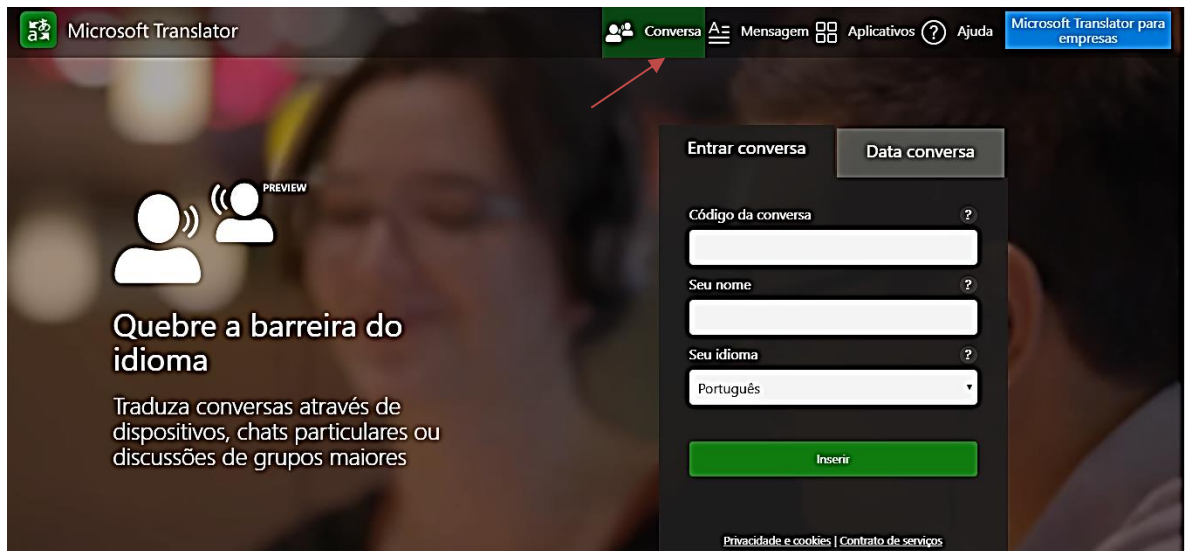
Figura 10 – Resultado da comparação entre as técnicas estatística e neural do Bing

The screenshot displays the Microsoft Translator 'Try & Compare' feature. At the top, a navigation bar includes 'Conversation', 'Text', 'Try & Compare', 'Apps', 'Help', and 'Microsoft Translator for Business'. The main area is titled 'Try & Compare' and 'Artificial Intelligence, powered by neural networks'. It shows the source text in English: 'Mary told me her Math test was a piece of cake. She'll get an A.' Below this is a green 'Translate & Compare!' button. The target language is set to Portuguese. Two translation results are shown side-by-side: 'Statistical' (brown box) and 'Neural' (dark blue box). The Statistical result is 'Mary me contou sua matemática teste foi um pedaço de bolo. Ela vai ter um.' and the Neural result is 'A Mary disse-me que o teste de matemática dela foi canja. Ela vai ter um a.' A note at the bottom states: 'These test sentences are randomly populated from our data set.' and a link to 'Find out more about Microsoft's Neural Machine Translation system' is provided. The footer contains social media links, help information, and the Microsoft Translator logo.

Fonte: <<https://translator.microsoft.com/neural>>.

Utilizando ainda a barra de ferramentas na parte superior da interface do usuário do *Bing*, é possível baixar aplicativos, retornar à página de tradução de textos, obter ajuda, acessar o *Microsoft Translator for Business* e acessar a página inicial do tradutor de conversas, com as instruções para o usuário (Figura 11). A barra inferior traz outras informações técnicas sobre esse TAO (ajuda, contato, privacidade e *cookies*, termos de uso, marcas registradas e anúncios) e *links* para o *Facebook*, *Twitter* e o blog do *Microsoft Translator*.

Figura 11 – Página inicial do tradutor de conversas do Bing



Como funciona



1. Iniciar

Clique em "Iniciar conversa", autentique-se e insira seu nome e idioma.



2.

Compartilhar

Compartilhe o código da conversa com outros participantes, que podem associar-se por meio do aplicativo Translator ou do site



3. Speak

Fale ou digite em sua língua para se comunicar com outros participantes, que verão suas mensagens em suas próprias línguas.

Fonte: <<https://translator.microsoft.com/>>.

É necessário fazer o *download* do aplicativo para que essa função do *Bing* seja utilizada. O usuário que inicia a conversa recebe um código com cinco letras que é compartilhado com os outros participantes. Cada participante insere esse código que é autenticado. Cada participante também insere um nome, pelo qual será reconhecido pelos outros participantes durante a conversa e escolhe a língua em que se comunicará. Conforme mencionamos anteriormente, há disponibilidade de 10 línguas para texto e fala, mas o usuário pode escolher uma língua que tem apenas suporte para texto e se comunicar com o grupo utilizando somente esse recurso.

Após termos descrito o *Bing Tradutor*, descrevemos, na seção seguinte, o segundo TAO analisado, o *Google Translate*.

3.2.1.2 Google Translate³⁸

O *Google Translate* é um serviço de TAO multilíngue de acesso livre para textos escritos e falados, imagens ou vídeos em tempo real, que pertence à empresa *Google*. Esse sistema oferece suporte para 103 línguas, em diferentes níveis, e serve mais de 500 milhões de pessoas diariamente. Em 2007, a empresa introduziu o seu próprio *software* de tradução, uma vez que, até então, utilizava o tradutor SYSTRAN, ainda utilizado por outros serviços de tradução como o *Babel Fish* (do Altavista), o tradutor do AOL, do *Yahoo* e do MSN (*Wikipedia*).

Desde do início do uso de seu próprio *software* até recentemente, o *Google Translate* utilizava a técnica estatística para a elaboração de suas traduções, tendo como dados linguísticos os transcritos realizados por humanos para as Nações Unidas e para o Parlamento Europeu. São utilizados milhões de documentos para procurar os padrões que ajudam a decidir qual a melhor tradução. Apesar de ser muitas vezes criticado e até ridicularizado, é o tradutor mais popular.

Da mesma forma que o *Bing Tradutor*, no final de 2016, o *Google Translate* se transformou em um sistema de tradução neural, o *Google Neural Machine Translation (GNMT)*, o qual traduz sentenças completas de uma vez, ao invés de partes por partes, uma primícia dos sistemas neurais, e que utiliza o contexto maior para ajudar a identificar qual a tradução mais relevante, que é então reorganizada e ajustada para parecer mais com a linguagem humana. O sistema neural do *Google* utiliza a técnica baseada em exemplos, descrita anteriormente e pode desenvolver uma tradução automática de interlíngua ao codificar a semântica da sentença, ao invés da memorização de traduções de frases para frases. Sua arquitetura proposta para a aprendizagem do sistema foi testada em todas as línguas a que o *Google Translate* dá suporte. A tradução neural do *Google Translate* está disponível do inglês para todas essas línguas e vice-versa.

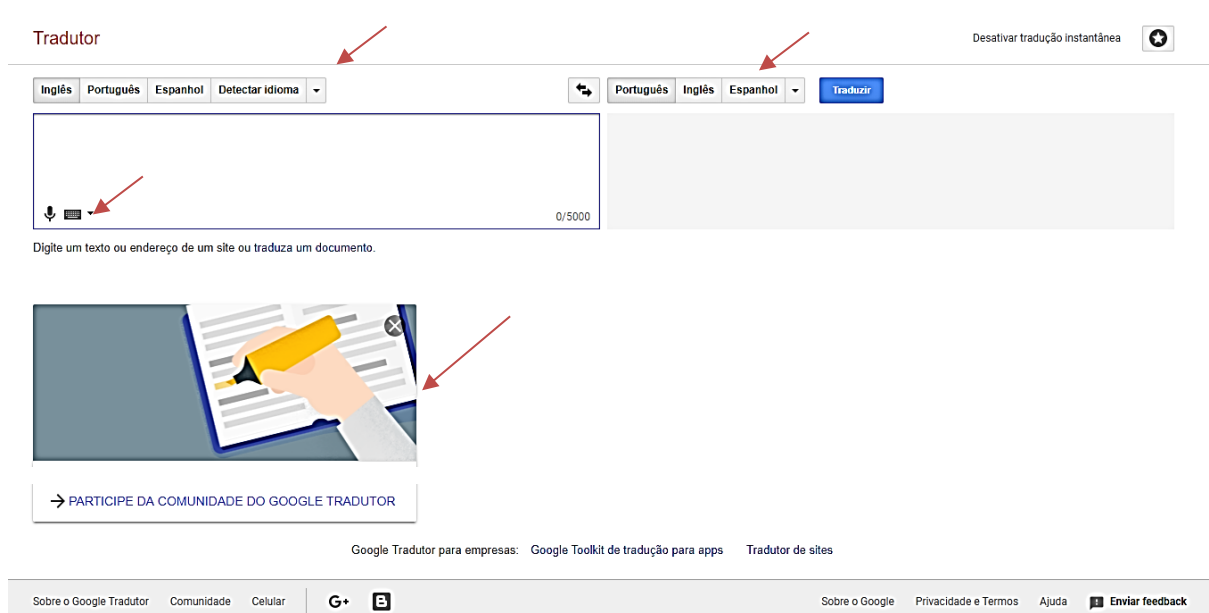
Até a introdução do sistema neural, se a LA não fosse o inglês, o *Google Translate* primeiro traduzia da LF para o inglês e depois para a LA. Contudo, foi

³⁸ As diversas informações sobre esse TAO foram retiradas de sua página oficial, a partir do *site*: <https://translate.google.com>. Foram realizados diferentes acessos em diferentes datas ao longo da pesquisa. O acesso mais recente foi realizado em 28 de janeiro de 2018.

anunciado que, com esse avanço de tecnologia, as traduções podem ser feitas diretamente de qualquer LF para qualquer LA disponibilizada por esse TAO.

A interface de usuário do *Google Translate* (Figura 12) traz a possibilidade de escolher os diferentes idiomas ofertados e o modo de *input* do texto através dos ícones ‘ativar entrada de fala’ e ‘ativar teclado virtual’ ou digitados normalmente pelo teclado do computador. Caso o usuário não saiba em qual língua está o texto original, há a função de detecção automática também. A interface estimula a colaboração do usuário na validação ou correção das traduções produzidas, como pode ser visto na ilustração da parte inferior esquerda da Figura 12.

Figura 12 – Interface do usuário do *Google Translate*

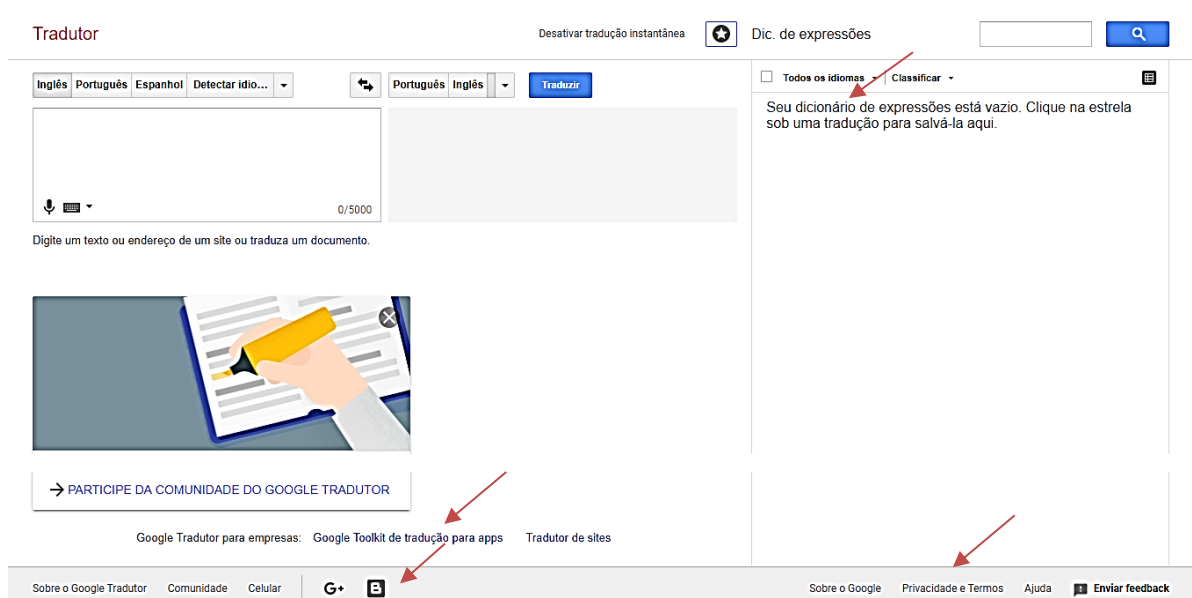


Fonte: <<https://translate.google.com>>.

Ainda na página inicial do *Google Translate* há, no canto superior direito, a função de desativar a tradução automática e uma estrela. Clicando nessa estrela, você tem acesso a um dicionário de expressões, o qual deve ser alimentado pelo próprio usuário, conforme pode ser visto na Figura 13. Esta é uma função bastante interessante, pois o usuário vai construindo seu próprio dicionário, o que facilita futuras traduções com as mesmas expressões. A barra de tarefas na parte inferior traz *links* para outros serviços de tradução do *Google* (para empresas, para *apps* e tradução de *sites*), bem como *links* para informações mais detalhadas sobre o *Google Translate*, para a comunidade de tradutores, para *download* de aplicativos para celulares e para

o *blog* do *Google Translate*, além de informações técnicas (sobre o *Google*, ajuda, termos de privacidade e *feedback*).

Figura 13 – Interface do usuário do *GoogleTranslate* com o dicionário de expressões aberto

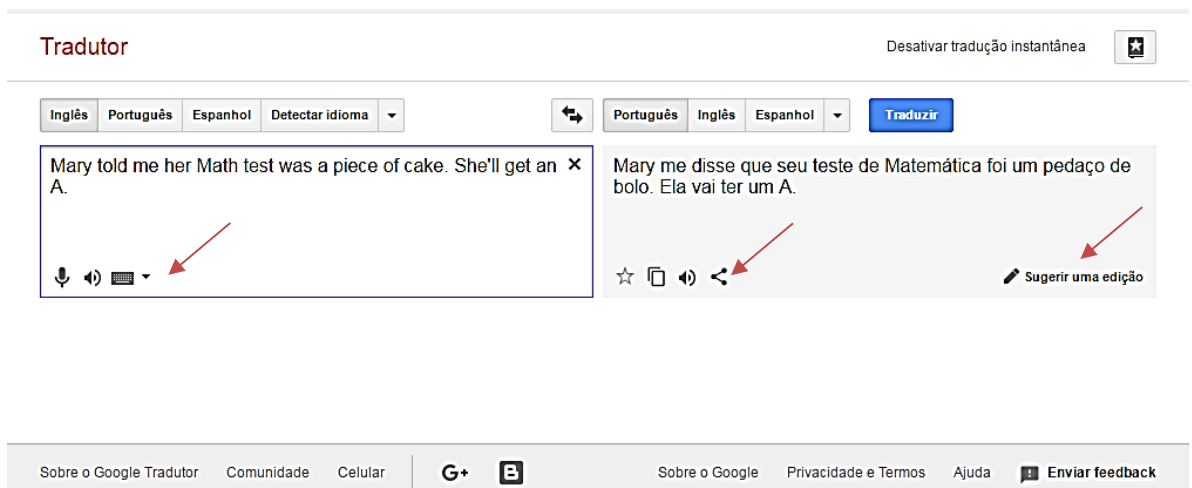


Fonte: <<http://translate.google.com>>.

Tão logo o texto é traduzido, a ilustração é retirada e o endereço também é alterado. Na caixa do texto original, agora, além dos métodos de *input* mencionados anteriormente, temos o ícone do áudio do texto a ser traduzido. Não há opção pelo tipo de voz – a voz disponibilizada é somente feminina. Outra mudança, se comparada com a tela anterior, é que ao clicar no ícone do teclado virtual, abre-se uma janela com três opções de teclados: inglês, inglês - DVORAK³⁹ e inglês - pesquisa escrita. Na caixa do texto traduzido, aparecem as opções de salvar, copiar, ouvir o áudio da tradução e compartilhá-la por meio de *e-mail*, *Twitter* ou *Google +*. Há ainda a possibilidade de sugerir uma edição para a tradução ao se clicar no ícone ‘sugerir uma edição’. (Figura 14)

³⁹ DVORAK é um tipo de teclado com layout simplificado que foi desenvolvido pelos *designers* August Dvorak e William Dealey em 1920 e 1930 como uma alternativa para o teclado QWERTY que é o mais comum. Disponível em: <https://pt.wikipedia.org/wiki/Teclado_Simplificado_Dvorak>. Acesso em: 02 abr.2017.

Figura 14 – GoogleTranslate com a tradução para a *piece of cake* em contexto



Fonte: <<https://translate.google.com/#pt/en/Mary%20told%20me%20her%20Math%20test%20was%20a%20piece%20of%20cake.%20She'll%20get%20an%20A>>.

Na versão estatística do *Google Translate*, ao passar o *mouse* sobre o texto de origem, não havia marcação nos itens correspondentes na tradução. Porém, ao passar o *mouse* sobre a tradução, as palavras correspondentes eram marcadas/sombreadas no texto de origem. O *GoogleTranslate* mostrava palavras isoladas ou blocos de palavras, o que evidenciava que esse tradutor reconhecia determinadas palavras como um bloco inteiro (*chunk*) traduzindo-as em conjunto. Clicando nas palavras ou frases, abria-se uma janela que permitia a substituição por outras traduções quando existentes, ou ainda, melhorar a tradução. Abaixo da caixa do texto original, aparecia a palavra ou frase em evidência com o ícone de áudio e abaixo dela(s), outra opção de pesquisa intitulada 'veja também', que traz as mesmas palavras ou outras relacionadas.

Figura 15 – GoogleTranslate estatístico com a expressão *a piece of cake* em evidência

The screenshot shows the Google Translate interface. The source text is "Mary told me her Math test was a **piece of cake**. She'll get an A." The translated text is "Mary me disse que seu teste de Matemática foi um **pedaço de bolo**. Ela vai ter um A." A red arrow points to the highlighted phrase "pedaço de bolo" in the translation. Below the translation, there is a section titled "Veja também" (See also) with a red arrow pointing to the text "de, bolo, pedaço". The interface includes language selection menus for Portuguese, English, and Spanish, and a "Traduzir" button.

Fonte: <<https://translate.google.com/#pt/en/Mary%20told%20me%20her%20Math%20test%20was%20a%20piece%20of%20cake.%20She'll%20get%20an%20A>>.

Na Figura 15, por exemplo, vemos a opção para 'pedaço de bolo' na versão estatística. Nesse caso, podíamos clicar em qualquer uma das opções 'de', 'bolo' ou 'pedaço'. Ao fazê-lo, era criada uma nova busca como pode ser verificada na Figura 16. Novas informações podiam ser encontradas dentre diferentes opções de significados e sua frequência de uso, através de barras ao lado de cada um. As barras evidenciavam a técnica estatística utilizada por esse TAO.

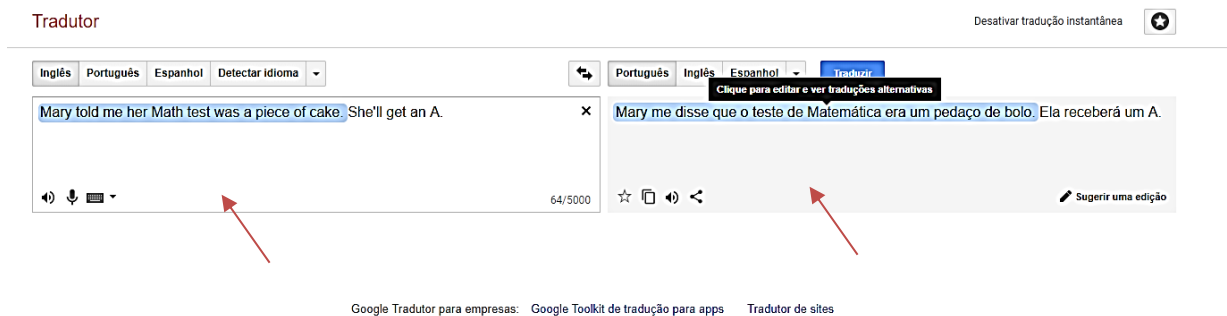
Figura 16 – GoogleTranslate para a tradução de 'bolo'

The screenshot shows the Google Translate interface with the word "bolo" entered in the source field and "cake" in the target field. A red arrow points to the word "bolo" in the source field. Below the translation, there are two sections: "Sinônimos de bolo" (Synonyms of bolo) and "Traduções de bolo" (Translations of bolo). The synonyms section lists "panqueca, panqueca" under the category "substantivo". The translations section lists "cake", "cookie", "pellet", and "stake" under the category "substantivo", with corresponding Portuguese translations: "bolo, torta, pastel, pastilha, torrão, talbete", "biscoito, bolo, biscoitinho", "pastilha, pílula, bolinha, bolo, grão de chumbo miúdo, pelouro", and "estaca, poste, parada, suporte, bolada, bolo". A red arrow points to the "Traduções de bolo" section. The interface includes language selection menus and a "Traduzir" button.

Fonte: <<https://translate.google.com/#pt/en/bolo>>.

Na versão neural, atualmente em uso, quando se passa o *mouse* sobre o texto de origem, não há marcação nos itens correspondentes na tradução. Porém, ao passar o *mouse* sobre a tradução, toda a sentença fica em evidência agora e não apenas uma palavra ou um *chunk* (Figura 17).

Figura 17 – Google Translate neural com a expressão *a piece of cake* em evidência



Fonte: <<https://translate.google.com/#en/pt/Mary%20told%20me%20her%20Math%20test%20was%20a%20piece%20of%20cake.%20She'll%20get%20an%20A>>.

Não há mais a possibilidade de selecionar palavras no texto traduzido, mas apenas no texto de origem. Ao selecionar uma palavra, diferentes informações relacionadas a ela são apresentadas, tais como possíveis definições, classes gramaticais, traduções, frequência de uso (barras) e pronúncia, tanto a representação fonética quanto a possibilidade de ouvir (Figura 18).

Figura 18 – GoogleTranslate para a palavra *cake*

Tradutor Desativar tradução instantânea

Inglês Português Espanhol Detectar idioma Português Inglês Espanhol Traduzir

Mary told me her Math test was a piece of **cake**. She'll get an A. 64/5000

Mary me disse que o teste de Matemática era um pedaço de bolo. Ela receberá um A. Sugerir uma edição

cake kak < > x

Definições de cake

substantivo
an item of soft, sweet food made from a mixture of flour, shortening, eggs, sugar, and other ingredients, baked and often decorated.
"a carrot cake"

verbo
(of a thick or sticky substance that hardens when dry) cover and become encrusted on (the surface of an object).
"a pair of boots caked with mud"
sinónimos: coat, encrust, plaster, cover

Veja também
chocolate cake, a piece of cake, birthday cake, piece of cake, bake a cake, rice cake, wedding cake, carrot cake, cake shop, moon cake

Traduções de cake

substantivo

- o bolo
- a torta
- o pastel
- a pastilha
- o torrão
- o taibete

verbo

- endurecer
- solidificar
- tornar em massa

cake, cookie, pellet, stake
pie, cake, dumpling, flan
pastel, pastry, pie, cake, tartlet
tablet, pellet, wafer, lozenge, pastille, cake
clod, divot, turf, dollop, cake, cob
cake

harden, stiffen, toughen, cake, set, sear
solidify, set, congeal, freeze, cake
cake

Google Tradutor para empresas: Google Toolkit de tradução para apps Tradutor de sites

Sobre o Google Tradutor Comunidade Celular Sobre o Google Privacidade e Termos Ajuda Enviar feedback

Fonte: <<https://translate.google.com/#en/pt/Mary%20told%20me%20her%20Math%20test%20was%20a%20piece%20of%20cake.%20She'll%20get%20an%20A>>.

No caso de ser selecionado um grupo de palavras, como no exemplo *a piece of cake*, são apresentadas diferentes possibilidades de pesquisa com as palavras que o compõem (Figura 19).

Figura 19 – GoogleTranslate com *a piece of cake* selecionado

Tradutor Desativar tradução instantânea

Inglês Português Espanhol Detectar idioma Português Inglês Espanhol Traduzir

Mary told me her Math test was **a piece of cake**. She'll get an A. 64/5000

Mary me disse que o teste de Matemática era um pedaço de bolo. Ela receberá um A. Sugerir uma edição

a piece of cake < > x

Veja também
a, of, piece, cake, a piece, piece of cake, a piece of, it's a piece of cake, as a piece of cake

Google Tradutor para empresas: Google Toolkit de tradução para apps Tradutor de sites

Sobre o Google Tradutor Comunidade Celular Sobre o Google Privacidade e Termos Ajuda Enviar feedback

Fonte: <<https://translate.google.com/#en/pt/Mary%20told%20me%20her%20Math%20test%20was%20a%20piece%20of%20cake.%20She'll%20get%20an%20A>>.

Nesse caso, se clicarmos em ***piece of cake, it's a piece of cake*** ou ainda ***as a piece of cake***, teremos o resultado esperado para a EI, incluindo possíveis sinônimos na língua de origem (Figura 20). Contudo, se o usuário não souber que esse grupo de palavras representa uma EI, ele, provavelmente, não saberá como chegar ao seu significado metafórico.

Figura 20 – GoogleTranslate com tradução para a *piece of cake*



Fonte: <<https://translate.google.com/#en/pt/piece%20of%20cake>>.

O *Google Translate* está disponível em dispositivos móveis para os sistemas *Android* (a partir de 2010) e *iOS* (a partir de 2011). Além da tradução de textos, ele disponibiliza também a tradução de conversas (a partir de 2011) e de imagens/placas (a partir de 2015) para todas as línguas. Esses recursos também podem ser acessados *off-line*, bastando que sejam feitos os *downloads* dos dicionários das línguas escolhidas⁴⁰.

Após termos descrito o *Google Translate* e suas funções, descreveremos, no próximo subitem, o terceiro e último TAO analisado nesta pesquisa, o SYSTRAN.

3.2.1.3 Tradutor SYSTRAN

Considerado por muitos como parâmetro de qualidade, quando se trata de Tradução Automática, o SYSTRAN é um dos sistemas de tradução automática mais

⁴⁰ Disponível em: <<https://translate.google.com/intl/pt-BR/about/>>. Acesso em: 29 jan.2018.

antigos, tendo suas atividades iniciadas em 1968 por seu fundador, Dr. Peter Toma. Essa empresa (SYSTRAN) tem prestado um extenso serviço ao Departamento de Defesa dos Estados Unidos e a Comissão das Comunidades Europeias (*Wikipedia*).

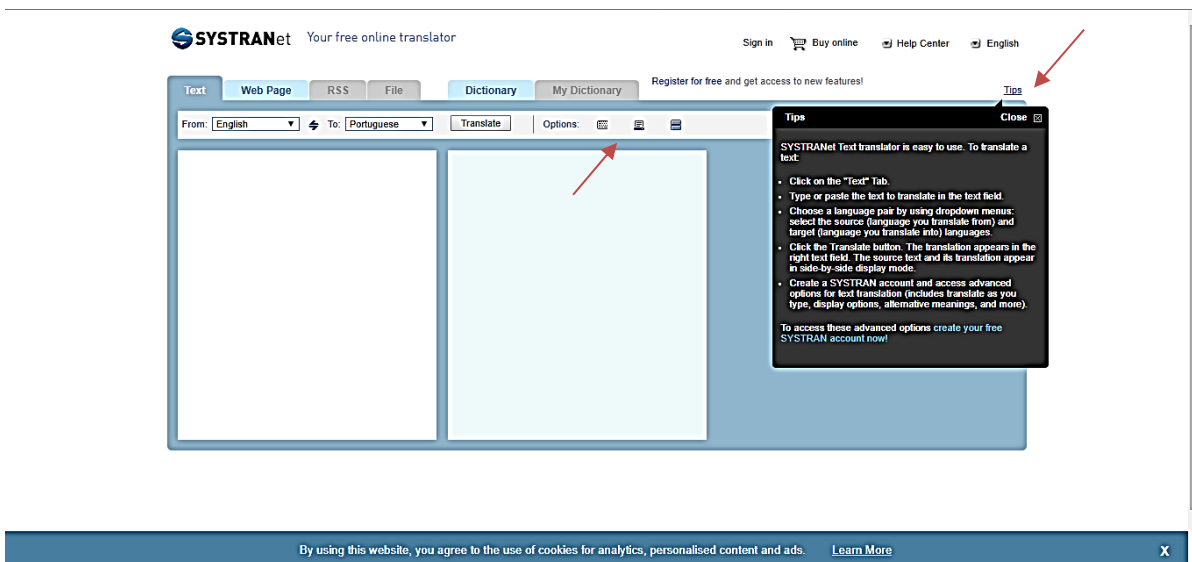
O SYSTRAN foi historicamente um sistema de Tradução Automática baseado em regras, mas, em 2010, inovou com um sistema híbrido que, de acordo com o *site* oficial da empresa, combina as fortalezas das técnicas de Tradução Automática Baseada em Regras com a Baseada em Estatísticas, a fim de produzir tradução de alta qualidade para qualquer domínio.

O *software* baseado em regras é a espinha dorsal desse novo sistema híbrido e provê uma estrutura de tradução sólida, enquanto as novas técnicas estatísticas aprendem dos dados monolíngues e bilíngues já existentes para melhorar cada fase do processo de tradução e aumentar o processo de customização.

Ainda segundo o *site* do SYSTRAN, cada par de línguas é construído no mesmo motor de tradução e integra as mais novas tecnologias de processamento natural de linguagem. Ao todo, são 52 pares.

Versões comerciais do SYSTRAN podem ser utilizadas no *Microsoft Windows*, incluindo o *Windows Mobile*, no Linux e no Solaris. A versão de acesso grátis *online* avaliada traz a interface de usuário com algumas abas (Figura 21) que podem facilitar o acesso a páginas da *web*, escolher um arquivo pessoal ou inserir o próprio texto com um máximo de 3000 caracteres. Não há ícones para ouvir o áudio dos textos, nem de compartilhar os resultados obtidos. Porém, há um ícone para imprimir, outro para inserir caracteres especiais e outro para modificar o *layout* de apresentação dos textos original e traduzido (lado a lado ou um acima do outro). Há também um botão com dicas (*tips*) para o usuário. Em versões anteriores desse sistema havia propagandas, mas, aparentemente, foram retiradas nessa versão mais recente. Outra diferença encontrada se mostrou na solicitação opcional de registro para ter acesso a outras funções.

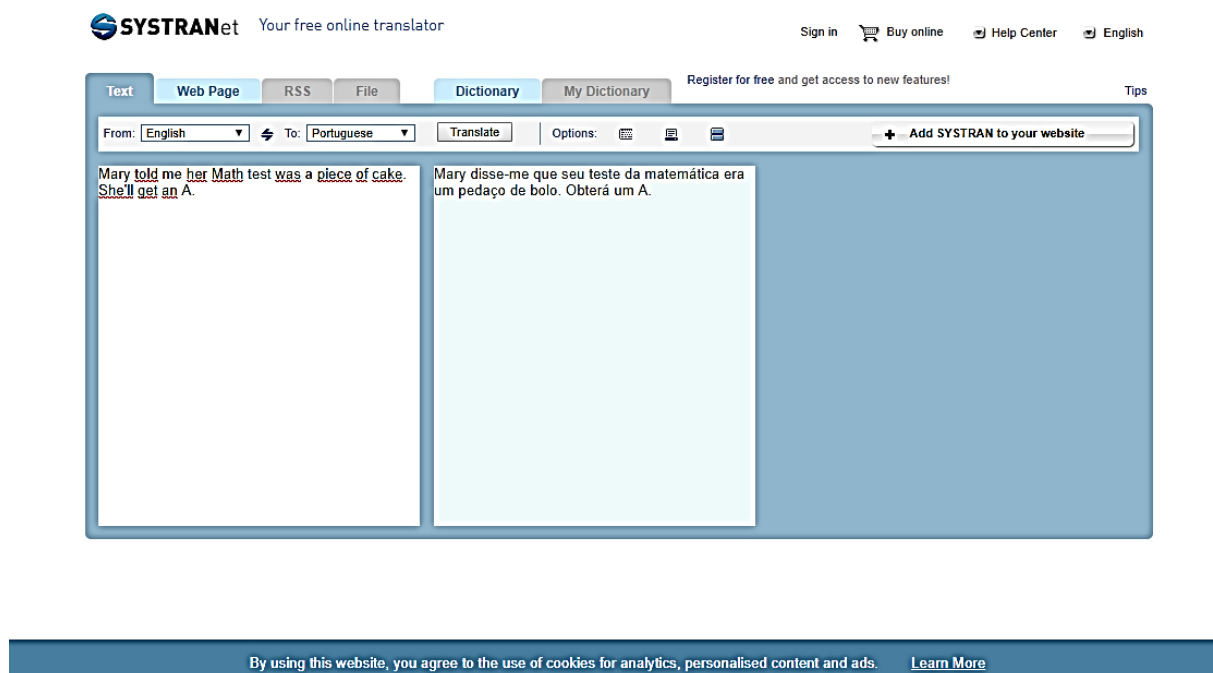
Figura 21 – Interface do usuário do SYSTRAN com instruções de uso



Fonte: <<http://www.systranet.com/translate>>.

Ao inserir o texto de origem, esse TAO faz o reconhecimento automático e elabora a tradução, disponibilizando-a no quadro ao lado ou abaixo, dependendo da configuração escolhida. A tradução do excerto com a *El a piece of cake* é apresentada na Figura 22, como exemplo.

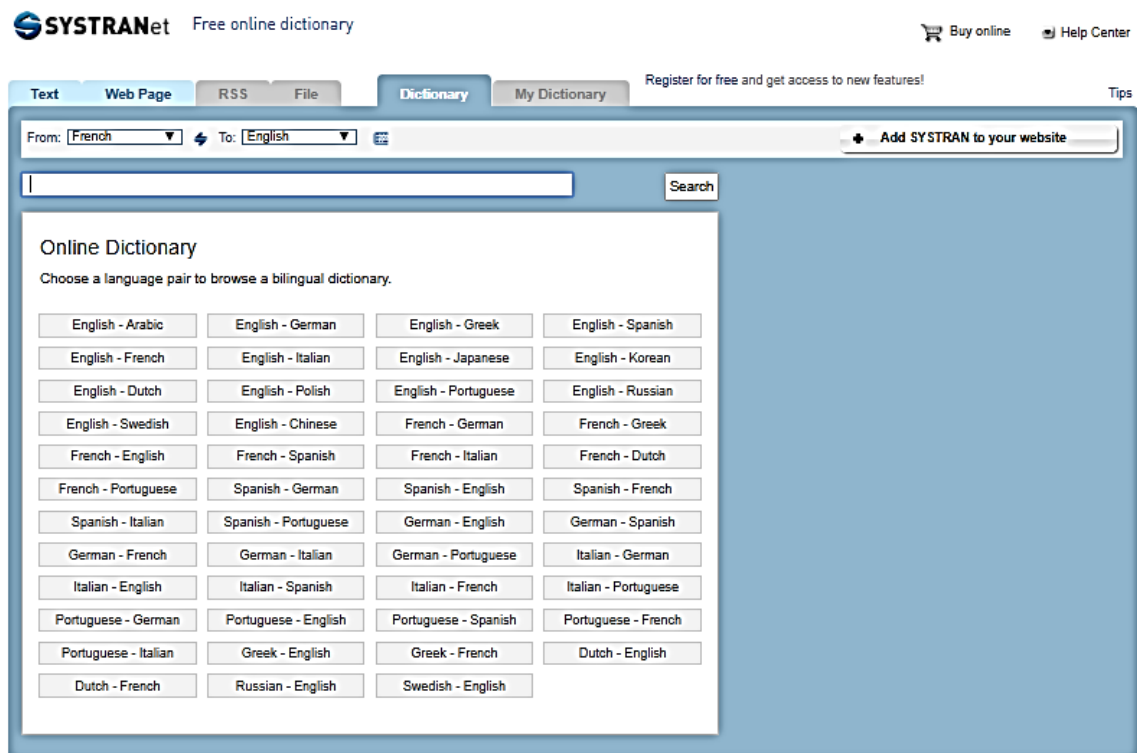
Figura 22 – SYSTRAN com a tradução para *a piece of cake* em contexto



Fonte: <<http://www.systranet.com/translate>>.

Nesta versão, as palavras são marcadas/sombreadas ao se passar o *mouse* sobre os textos originais e traduzidos, mas não aparecem opções de significados. Contudo, há a possibilidade de acessar um dicionário geral (Figura 23) e um customizado do usuário, caso ele tenha uma conta registrada.

Figura 23 – Dicionários disponibilizados pelo SYSTRAN



By using this website, you agree to the use of cookies for analytics, personalised content and ads. [Learn More](#)

Fonte: <<http://dictionary.systranet.com/>>.

Apresentamos, na Figura 24, as informações para a El ***a piece of cake*** no dicionário bilíngue inglês-português do SYSTRAN.

Figura 24 – Resultado do dicionário do SYSTRAN para *a piece of cake*

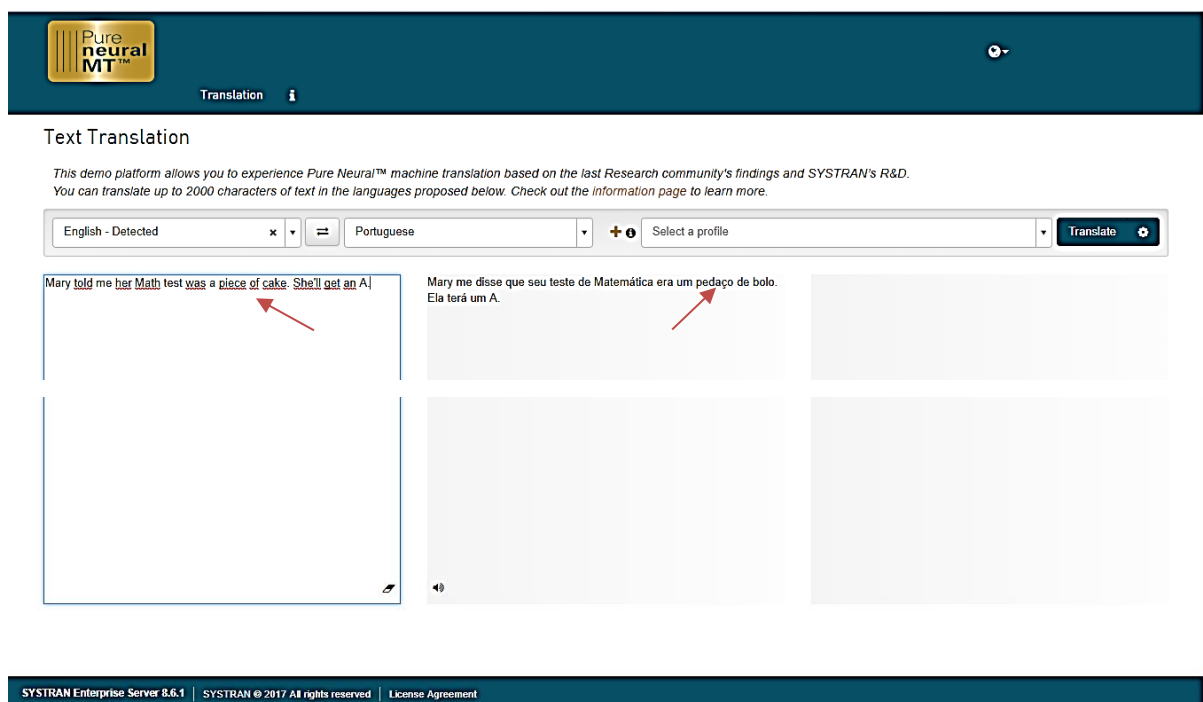
The screenshot shows the SYSTRANet website interface. At the top, there is a navigation bar with the SYSTRANet logo and the text 'Your free online translator'. On the right, there are links for 'Sign in', 'Buy online', 'Help Center', and 'English'. Below this, there are tabs for 'Text', 'Web Page', 'RSS', 'File', 'Dictionary', and 'My Dictionary'. The 'Dictionary' tab is selected. A search bar at the top of the dictionary section shows 'From: English' and 'To: Portuguese'. The search term 'piece of cake' is entered in the search box. The results are displayed under the heading 'English to Portuguese' and 'SYSTRAN Dictionary'. The main entry is '1. piece noun', which includes a list of Portuguese equivalents: 'parte', 'bocado', 'fragmento', 'naco', 'pedacinho', and 'pedaço'. It also provides specific translations for 'piece of cake', 'piece of land', and 'piece of paper'. Other entries include '2. piece adj' and '3. piece verb'.

Fonte: <<http://www.systranet.com/dictionary/english-portuguese/piece%20of%20cake?>>.

No final de 2016, ainda em fase de teste, a tecnologia neural da empresa SYSTRAN foi lançada. Intitulada *Pure Neural™ Machine Translation* ou *PNMT™*, ela possibilita a aprendizagem, geração de regras de linguagem a partir de um texto traduzido e a produção de uma tradução, que pode atingir um nível acima do atual estado da arte e melhor que aquela elaborada por um falante não nativo⁴¹. A versão *demo* pode ser acessada em <https://demo-pnmt.systran.net/information#/view> disponibilizando para o usuário um limite de até 2000 caracteres, segundo as instruções que aparecem na interface do usuário dessa versão (Figura 25).

⁴¹ Disponível em: <<http://www.systransoft.com/systran/translation-technology/pure-neural-machine-translation/#lets goback>>. Acesso em: 31 jan. 2018.

Figura 25 – SYSTRAN neural com a tradução para *a piece of cake* em contexto

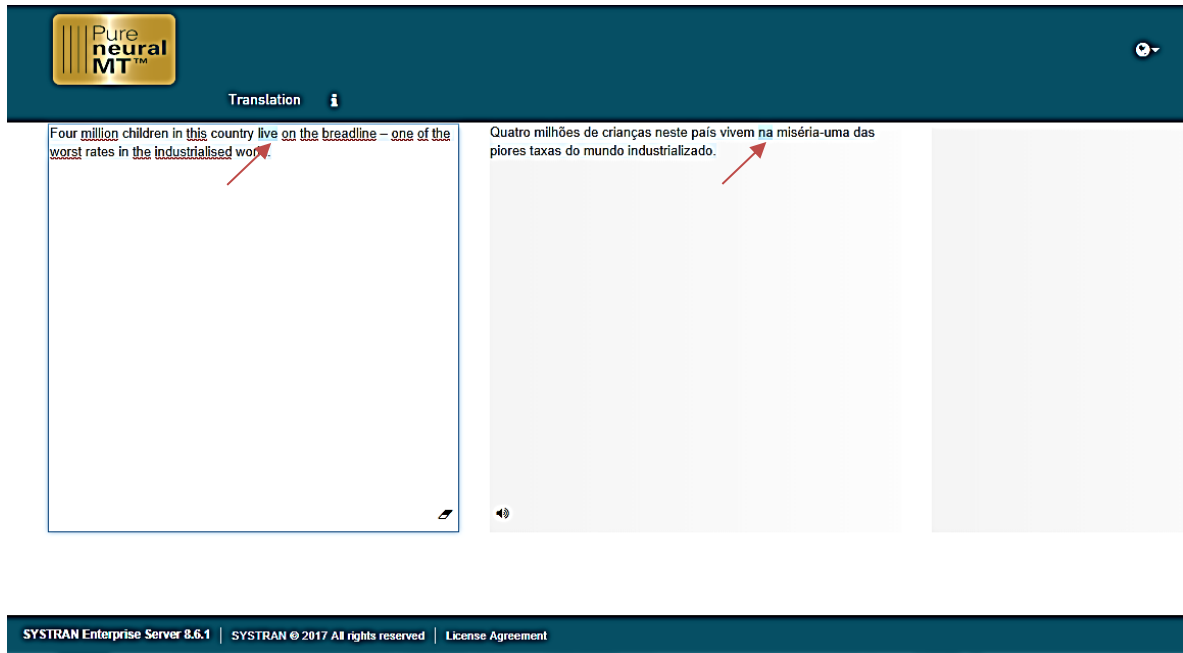


Fonte: <<https://demo-pnmt.systran.net/production#/translation>>.

Nota-se que a versão neural também não consegue detectar a EI **a piece of cake**, traduzindo-a literalmente. Na verdade, se compararmos com a tradução elaborada pela versão padrão, praticamente, não há diferenças: ‘disse-me’ – ‘me disse’ e ‘Obterá um A’ – ‘Ela terá um A’.

Na versão neural (*demo*), há uma curiosidade, ao passar o *mouse* sobre as palavras no texto de origem, colocando-as em evidência, algumas palavras também são marcadas no texto traduzido. Contudo, percebemos que não são as palavras que correspondem ao significado correto. Por exemplo, evidenciamos o verbo *live* no texto em inglês e a preposição ‘na’ é que ficou em evidência no texto traduzido (Figura 26). Isto é algo que deve ser aprimorado em versões futuras desse TAO.

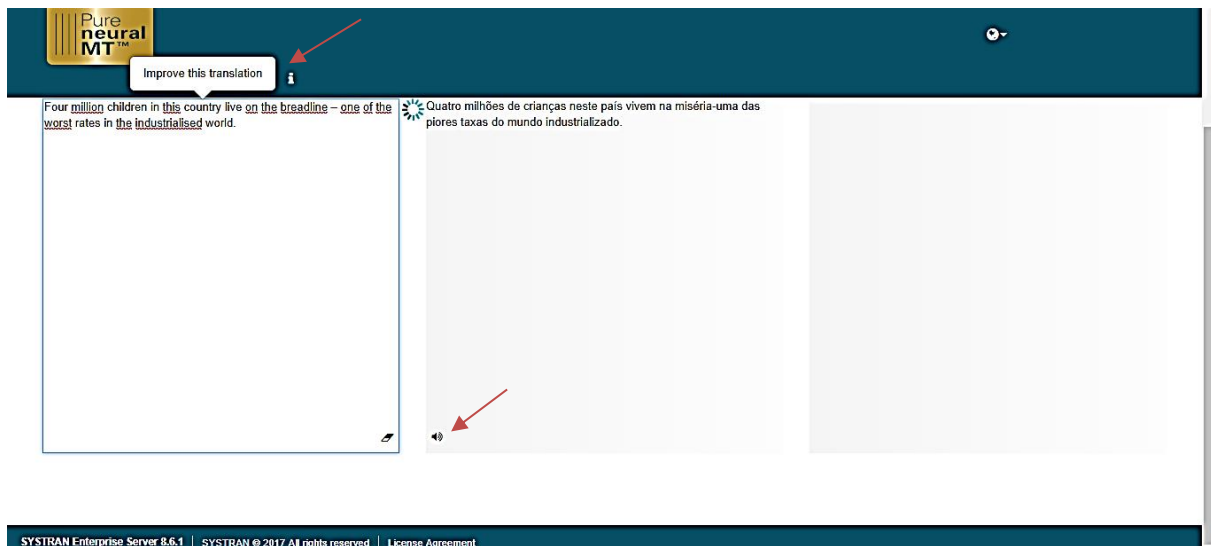
Figura 26 – Palavras em evidência na versão neural do SYSTRAN



Fonte: <<https://demo-pnmt.systran.net/production#/translation>>.

Para a versão neural de demonstração do SYSTRAN, há a possibilidade de sugerir melhorias na tradução nas duas línguas em trabalho. Clicando em qualquer parte do texto de origem ou traduzido, uma aba com a inscrição *improve this translation* (melhore esta tradução) aparece. Outro aprimoramento dessa versão, em relação à padrão, é que há a possibilidade de ouvir o texto traduzido, clicando no ícone do autofalante no canto inferior esquerdo do quadro do texto traduzido (Figura 27).

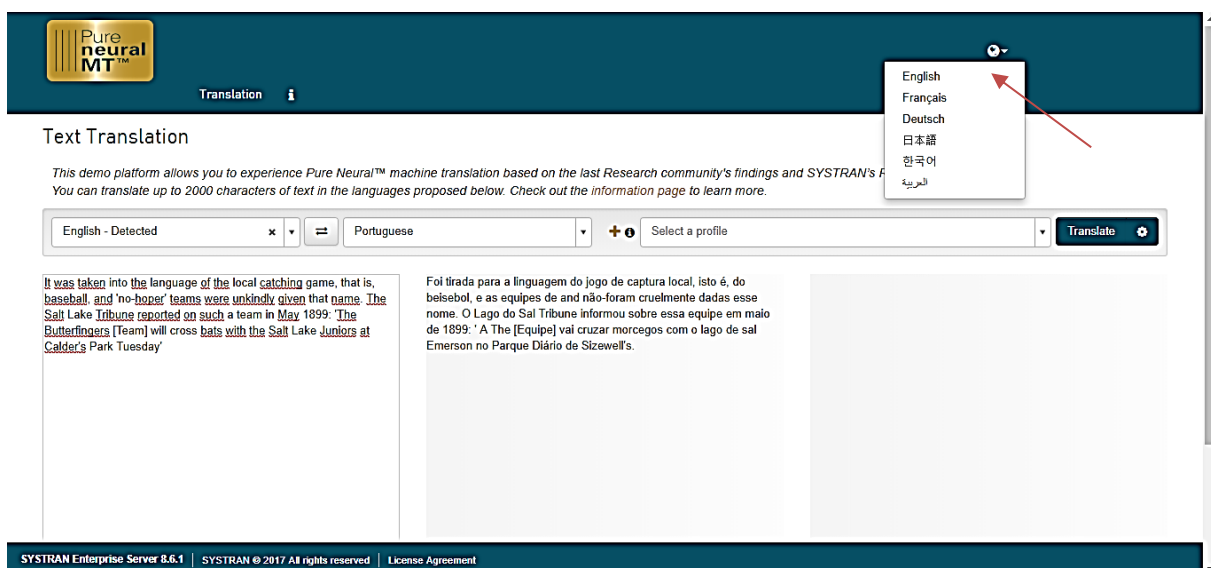
Figura 27 – Participação do usuário na melhoria das traduções na versão neural do SYSTRAN



Fonte: <<https://demo-pnmt.systran.net/production#/translation>>.

A versão *demo* oferece a possibilidade da interface do usuário estar em seis línguas diferentes, como pode ser visto na Figura 28.

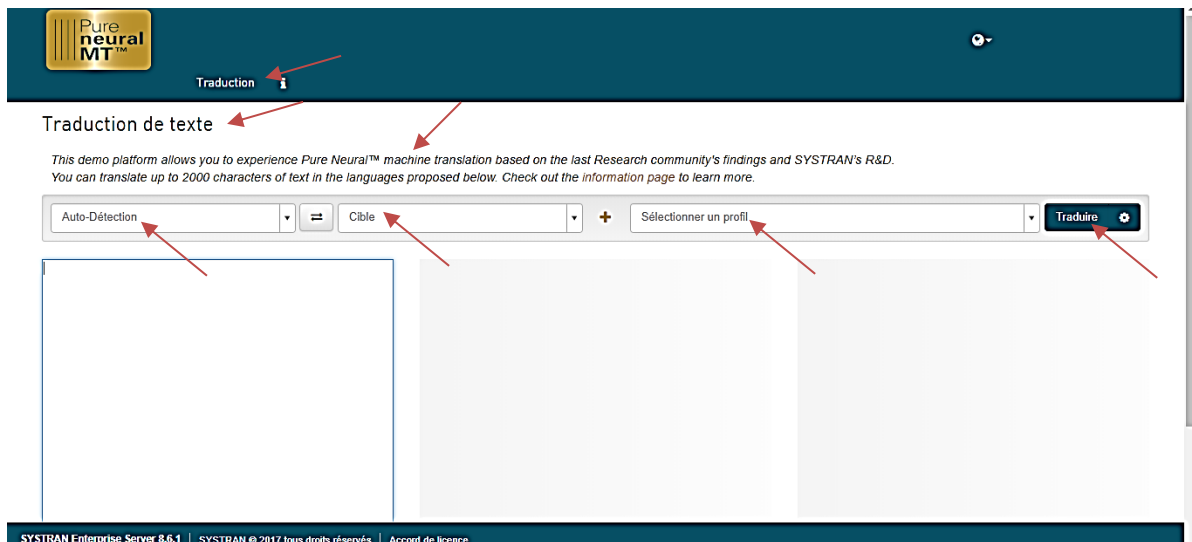
Figura 28 – Interface do usuário em outras línguas na versão neural do SYSTRAN



Fonte: <<https://demo-pnmt.systran.net/production#/translation>>.

Clicando na língua escolhida, o menu e títulos são, automaticamente, traduzidos, mas as instruções principais continuam em inglês (Figura 29).

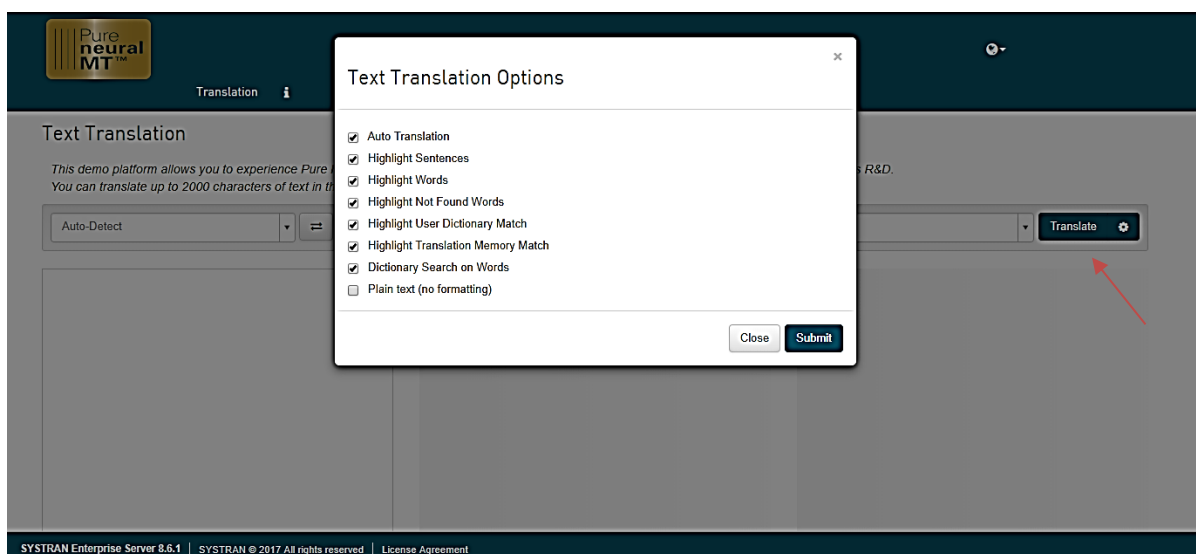
Figura 29 – Versão *demo* do SYSTRAN em francês



Fonte: <<https://demo-pnmt.systran.net/production#/translation>>.

Para finalizar a descrição dessa versão neural *demo*, apresentamos as possibilidades de configuração, ofertadas ao clicar no botão ‘traduzir’, localizado à direita da interface do usuário (Figura 30).

Figura 30 – Opções de configuração da versão neural *demo* do SYSTRAN



Fonte: <<https://demo-pnmt.systran.net/production#/translation>>.

Entendemos que as características dos três TAO's ora descritos sofrerão novas modificações com avanços tecnológicos vindouros. Contudo, o presente

registro é relevante para que se possa fazer comparações em estudos futuros, nos quais os mesmos TAO's sejam analisados.

Exploraremos na próxima subseção os procedimentos metodológicos utilizados para este estudo.

3.3 PROCEDIMENTOS METODOLÓGICOS

A metodologia pretendida para a pesquisa avaliou os três sistemas de tradução automática *online* escolhidos, que estão disponíveis gratuitamente na *Internet*, quanto à sua percepção em reconhecer e traduzir, do inglês para o português, EI's que contenham um alimento em sua composição lexical, bem como suas metáforas conceituais subjacentes. Os TAO's escolhidos foram o *Bing* Tradutor, o *Google Translate* e o SYSTRAN, descritos no subitem 3.2.1. Até o momento, a documentação dos sistemas supracitados não menciona o uso de recursos linguísticos para o produto de suas análises, ou seja, as traduções.

Descrevemos de forma sucinta as etapas que foram necessárias para a coleta e análise dos dados desta pesquisa e, nos subitens seguintes, o passo a passo de cada uma: a primeira etapa foi elaborar a lista das EI's em inglês que foram analisadas, retirando seus excertos do *Webcorp*, realizando o tratamento necessário para que, posteriormente, fossem traduzidas e, então, classificá-las em relação ao tipo de metáfora conceitual subjacente. Paralela à primeira etapa, a segunda etapa foi submeter o presente projeto à análise do Comitê de Ética em Pesquisa da Universidade Estadual do Ceará (UECE), uma vez que THP's participaram da parte experimental de nossa pesquisa. Enquanto aguardávamos a devida aprovação do Comitê de Ética (conf. APÊNDICE A), iniciamos a terceira etapa com a elaboração das traduções das EI's com alimentos utilizando os três TAO's. Esta etapa teve continuidade, após aprovação pelo Comitê de Ética, com a solicitação que THP's traduzissem as EI's com alimentos selecionadas. Ainda nesta etapa, aplicou-se um questionário para traçar o perfil dos THP's. Na quarta etapa, as sentenças candidatas foram submetidas às três métricas de avaliação de Tradução Automática. Finalmente, foram feitas avaliações qualitativas e quantitativas dos produtos tradutórios dos TAO's e THP's. A partir da terceira etapa, iniciamos a vertente experimental do presente estudo. Portanto, essas etapas estão incluídas em nosso desenho experimental, subitem 3.4.

Descrevemos, nas próximas subseções, como cada etapa do processo se desenvolveu:

3.3.1 Elaboração da lista e excertos com as EI's

Primeiramente, salienta-se que os alimentos que compõem lexicalmente as EI's que foram analisadas para esta pesquisa estão em sua grande maioria em sintagmas nominais, com apenas algumas expressões, em que o alimento funciona sintaticamente como adjetivo ou verbo.

A lista de EI's com alimentos em sua composição lexical, a princípio, era composta de 80 EI's do português para o inglês, retiradas do Dicionário de Expressões Idiomáticas Metafóricas Português-Inglês (Anexo A) e de 92 EI's em inglês retiradas do *Dictionary of Metaphoric Idioms Portuguese-English* (Anexo B), ambos de Camargo e Steinberg (1989 e 1990, respectivamente). Nossa análise preliminar, citada na introdução, nos mostrou a problemática relação entre Tradução Automática e EI's. Nesse primeiro estudo, realizamos uma análise por amostra das EI's com alimentos, somente com o TAO *Google Translate*, tanto do inglês para o português como vice-versa, estando estas fora de um contexto. A maioria das EI's não foi reconhecida pelo *Google Translate*. Somente quatro EI's, do inglês para o português, de um total de 27 EI's pesquisadas e seis do português para o inglês, de um total de 34 EI's pesquisadas. Esses resultados puderam nos auxiliar a desenvolver melhor os objetivos e questões da presente pesquisa, bem como delimitar o universo de EI's analisadas.

A partir da coleta das EI's com alimentos nos dois dicionários, verificou-se que algumas delas tinham um termo de alimento em ambas as línguas. Definimos, então, que restringiríamos nossa análise para somente essas, a fim de facilitar a análise das metáforas subjacentes nas duas línguas de interesse para este estudo. Esta decisão nos deixou com 43 EI's, que agrupamos em inglês como LF, suas respectivas traduções para o português e significados, os quais listamos no Quadro 10:

Quadro 10 – El's em língua inglesa e língua portuguesa e seus significados

Nº	EXPRESSÃO IDIOMÁTICA EM INGLÊS	EXPRESSÃO IDIOMÁTICA EM PORTUGUÊS	SIGNIFICADO
01	To get out of a jam	Descascar o abacaxi ; ter/estar com um abacaxi nas mãos	Conseguir resolver/lidar com um problema
02	To handle the/a hot potato	Descascar o abacaxi ; ter/estar com um abacaxi nas mãos	Conseguir resolver/lidar com um problema
03	To be left with the hot potato	Ficar com o abacaxi	Ficar com um problema nas mãos
04	To give sm the hot potato	Passar o abacaxi	Passar o problema para outra pessoa
05	To have a hot potato	Estar com uma batata quente nas mãos	Ter um problema nas mãos
06	To be/have a hard nut to crack	Ser/Ter um osso duro de roer	Ter que resolver uma situação difícil
07	To have a lemon on your hands	Descascar o abacaxi ; ter/estar com um abacaxi nas mãos	Ter um problema nas mãos
08	A (pretty) kettle of fish	Um angu de caroço	Uma situação confusa, desorganizada
09	To upset the/sm's apple cart	Entornar o caldo	Atrapalhar os planos de alguém, bagunçar
10	To be butterfingers	Ter mão(s) de manteiga	Ser desastrado
11	Not to cut the mustard any more	Queimar óleo quarenta	Não ser capaz de fazer algo, frequentemente relacionado ao sexo
12	To pull the chestnuts out of/from the fire	Tirar a castanha/sardinha com a mão do gato	Colocar outra pessoa em uma situação perigosa, a fim de obter vantagens
13	To walk on eggs/eggshells	Pisar em/sobre ovos	Ter cautela
14	To be no piece of cake	Não ser bolinho	Não ser fácil
15	A piece of cake	Mamão com mel/açúcar	Algo fácil, sem complicações
16	To be a small beer	Ser café pequeno	Algo rápido de ser feito ou resolvido
17	To sell like hot cakes/hotcakes	Vender como água	Vender facilmente/rapidamente
18	To be a cake-eater/cake eater	Ser arroz de festa	Algo fácil de ser encontrado/almofadinha/homem afeminado/bom <i>vivan</i>
19	To pour oil on troubled waters	Jogar/Lançar/Por água (fria) na fervura	Tentar acalmar uma situação difícil/complicada
20	To throw cold water	Jogar/Lançar/Por água (fria) na fervura	Tentar acalmar uma situação difícil/complicada
21	To pour cold water	Jogar/Lançar/Por água (fria) na fervura	Tentar acalmar uma situação difícil/complicada
22	To swallow the bait	Morder a isca	Cair em uma armadilha
23	To rise to the bait	Morder a isca	Cair em uma armadilha
24	To take the bait	Morder a isca	Cair em uma armadilha
25	To kill the goose that lays the golden eggs	Matar a galinha dos ovos de ouro	Tomar uma atitude que traz uma consequência negativa para você próprio

26	To be cheese paring	Ser pão duro	Ser avarento
27	To count one's chickens before they hatch	Contar com o ovo na galinha	Contar com algo incerto
28	To cry over spilled/spilt milk	Chorar (sobre) o leite derramado	Arreponder-se
29	To live off the fat of the land	Viver à tripa forra	Viver confortavelmente
30	To have a finger in the pie	Tirar uma casquinha	Tirar vantagem de uma situação; estar envolvido em algo
31	To earn one's daily bread	Ganhar o pão nosso de cada dia	Ganhar o seu sustento
32	To be our daily bread	Ser o pão nosso de cada dia	Ser a atividade cotidiana/rotineira
33	To take the bread out of sm's mouth	Tirar o pão da boca de a.	Tomar de alguém algo que era tido como certo ou seu sustento
34	To live on the breadline	Passar a pão e água	Viver de forma miserável
35	Milk and water	Sem sal	Algo insípido/sem graça/sem atrativos
36	To be all talk and no cider	Ser mais as vozes que as nozes	Falar muito e não agir, não resolver nada
37	To waffle	Encher linguiça	Enrolar/Embromar
38	To be neither fish nor flesh (nor good red Herring)	Não ser carne nem peixe	Não tem lado definido; ficar em cima do muro; não contribuir nem para uma coisa nem outra
39	All one's geese to be swans	Comer frango e arrotar peru/faisão	Exagerar sobre algo para mostrar que é melhor
40	To imbibe sth with one's mother's milk	Beber a. c. com o leite materno	Adquirir algo com a convivência, criação; de berço
41	To have nothing to do with the price of tea in China	Não ter nada a ver com o peixe	Eximir-se da responsabilidade sobre algo
42	To separate the wheat from the chaff	Separar o joio do trigo	Diferenciar o bem do mal
43	To make no bones about sth	Não ter papas na língua	Falar abertamente sobre algo; não fazer segredo

Fonte: Elaborado pela autora.

A partir desta lista, classificamos as EI's do nosso *corpus* dentro do *continuum* entre literalidade e figuratividade de Fernando e Flavell (1981) no Quadro 11. Esclarecemos que essa classificação foi elaborada a partir de uma percepção individual e que pode haver divergências, uma vez que o limite entre uma categoria e outra é bastante tênue. O próprio nome (*continuum*) pressupõe essa característica.

Quadro 11 – EI's classificadas dentro do *continuum* de idiomaticidade

Língua Inglesa			
Transparentes (14)	Semitransparentes (9)	Semi-opacas (12)	Opacas (8)
To walk on eggs/eggshells To sell like hot cakes/hotcakes To throw cold water To pour cold water To pour oil on troubled waters To swallow the bait To take the bait To kill the goose that lays the golden eggs To count one's chickens before they hatch To cry over spilled/spilt milk To earn one's daily bread To be our daily bread To take the bread out of sm's mouth To separate the wheat from the chaff	To handle the/a hot potato To be left with the hot potato To give sm the hot potato To have a hot potato To be/have a hard nut to crack To rise to the bait To live on the breadline All one's geese to be swans To be neither fish nor flesh	To get out of a jam To have a lemon on your hands To upset the/sm's apple cart To be butterfingers To pull the chestnuts out of/from the fire To be no piece of cake A piece of cake To be a small beer To be a cake-eater/cake eater To live off the fat of the land To have a finger in the pie To be all talk and no cider	A (pretty) kettle of fish Not to cut the mustard any more To be cheese paring Milk and water To have nothing to do with the price of tea in China To make no bones about sth To imbibe sth with one's mother's milk To waffle
Língua Portuguesa			
Transparentes (14)	Semitransparentes (7)	Semi-opacas (7)	Opacas (10)
Descascar o abacaxi Ter/estar com um abacaxi nas mãos Ficar com o abacaxi Passar o abacaxi Pisar em/sobre ovos Vender como água Morder a isca Matar a galinha dos ovos de ouro Contar com o ovo na galinha Chorar sobre o leite derramado Ganhar o pão nosso de cada dia Ser o pão nosso de cada dia Passar a pão e água Separar o joio do trigo	Estar com uma batata quente nas mãos Ser/Ter um osso duro de roer Ter mão(s) de manteiga Jogar/Lançar/Por água (fria) na fervura Tirar o pão da boca de a. Sem sal Não ser carne nem peixe	Entornar o caldo Mamão com mel/açúcar Não ser bolinho Comer frango e arrotar peru/faisão Ser pão duro Encher linguiça Viver à tripa forra	Beber a. c. com o leite materno Não ter nada a ver com o peixe Tirar a castanha/sardinha com a mão do gato Um angu de carço Queimar óleo quarenta Ser café pequeno Ser arroz de festa Tirar uma casquinha Ser mais as vezes que as nozes Não ter papas na língua

Fonte: Elaborado pela autora.

Verifica-se que um número maior das EI's está no grupo de idiomaticidade transparente e semitransparente, aproximadamente, 56% em ambas as línguas. O

inglês apresenta um número maior de EI's semi-opacas (12) do que o português (7). A diferença entre o número de EI's semitransparentes e opacas é semelhante, com o inglês apresentando um número um pouco maior de expressões semitransparentes e o português um número um pouco maior de expressões opacas. O número de exemplos para cada uma das classificações desse *continuum* nos parece ser uma amostragem suficiente para verificar a percepção dos diferentes graus de idiomaticidade das EI's pelos TAO's e THP's.

Após a classificação dentro do *continuum*, verificou-se o uso e a frequência dessas 43 EI's, assim como definiu-se os excertos a serem traduzidos pelos TAO's e THP's. Nessa verificação se utilizou o *Webcorp*, uma vez que os dicionários utilizados como fonte foram publicados há mais de 25 anos, de forma a identificar se as EI's com alimentos estão em uso atualmente. A ferramenta *Webcorp* utiliza a *Web* como *corpus* e, portanto, considerada apropriada para trabalhar com EI's, uma vez que possibilita mais amostragens de linguagem do uso cotidiano e coloquial, as quais EI's são, tipicamente, mais utilizadas.

Os filtros utilizados pelo *Webcorp* para as expressões em inglês foram: não distinguir maiúsculas e minúsculas (*case insensitive*); utilizar o motor de busca do *Bing - Azure (using the Bing (Azure) API)*; e pesquisar em *sites* jornalísticos e acadêmicos dos Estados Unidos, Canadá e Reino Unido (*from sites: thetimes.co.uk telegraph.co.uk theguardian.com independent.co.uk mirror.co.uk thesun.co.uk dailystar.co.uk nytimes.com washingtonpost.com latimes.com suntimes.com bbc.co.uk/news .edu .ac.uk .ca .uk*). Sabendo-se que tanto o inglês quanto o português são línguas faladas em diversos países, restringimos nosso campo de pesquisa para o português falado no Brasil, o inglês norte-americano (EUA e Canadá) e o britânico, que são os que mais têm ingresso em nosso país. Os filtros utilizados para as EI's em português foram os mesmos utilizados para as EI's em inglês – não distinguir maiúsculas e minúsculas (*case insensitive*) e utilizar o motor de busca do *Bing - Azure (using the Bing (Azure) API)*. Contudo, há um número bem mais reduzido de *sites* em português disponibilizados pelo *Webcorp* se comparado com o inglês. Na verdade, há apenas uma opção (*from sites: .br*).

Cada busca no *Webcorp* pode gerar até 50 *sites*, em que o texto pesquisado⁴² está inserido, neste caso a EI. Contudo, o número de concordâncias é ilimitado, pois cada *site* pode trazer o mesmo texto pesquisado diversas vezes. As pesquisas no *Webcorp* para geração dos excertos com as EI's para este estudo foram realizadas em agosto de 2016. Foram geradas, no total, 2.932 ocorrências em inglês e 2.708 em português.

Analisamos individualmente cada ocorrência gerada para todas as EI's em inglês, a fim de verificar se os excertos traziam a EI em sentido figurado ou não, a data de uso, em que parte do texto (título ou corpo), o gênero textual e o assunto ou tópico. Em português, verificamos se houve ocorrência de uso somente. Houve uma variação bastante relevante no número de ocorrências geradas – algumas EI's geraram mais de 300 ocorrências, enquanto outras geraram apenas uma ou nenhuma. Registramos que, para algumas EI's, foi necessário realizar mais de uma busca, pois não foi possível utilizar o asterisco (* - *wild card*)⁴³ para oportunizar diferentes resultados.

Os alimentos fazem parte do cotidiano de toda cultura e, obviamente, são usados não somente nas EI's, apresentando uma frequência de uso imensa em seu significado não metafórico. A partir dos resultados obtidos em nossa análise preliminar, em que as EI's foram traduzidas isoladamente, fora de um contexto, e com o objetivo de evitar a ambiguidade no uso dessas expressões, se no seu sentido metafórico ou não, para esta pesquisa, as EI's traduzidas pelos TAO's e pelos THP's precisavam estar inseridas em um contexto, dentro de uma sentença ou um pequeno texto que foram retirados das ocorrências geradas pela pesquisa no *Webcorp*.

Outro mecanismo que utilizamos para verificar se as EI's ainda estão em uso, foi o *Google Books N-gram Viewer*. Todas as 43 EI's foram pesquisadas em língua inglesa somente, uma vez que esta ferramenta não está disponível em língua portuguesa ainda. A partir do percentual de frequência de cada EI em 2008, último

⁴² Em uma primeira busca, a fim de nos familiarizarmos com o *Webcorp*, o motor de busca utilizado era o *Google* e essa ferramenta gerava até 64 *sites*. Atualmente (dezembro de 2017), o motor de busca mudou para o *Bing (Cognitive)* ou *FAROO*, diferentes daquele disponível em agosto de 2016 – *Bing (Azure)*. Isso nos mostra quão dinâmico é o mundo virtual, fazendo-se necessário que pesquisas registrem essas mudanças.

⁴³ *Wild card* – carta curinga. Recurso utilizado em buscadores de *Internet* possibilitando a ocorrência de qualquer palavra para substituir o asterisco (e.g. 'upset * apple cart' – upset **his/her/my/the company's** apple cart). No português, facilita encontrar diferentes ocorrências de tempos verbais (e.g. 'entorn* o caldo – entorn**ar**/entorn**ei**/entorn**ou** o caldo').

ano disponibilizado para pesquisa, pudemos verificar que praticamente todas estão em uso. As exceções, ou seja, as EI's que obtiveram 0% de frequência em 2008 foram ***not cut the mustard anymore*** e ***a cake-eater*** (com esta grafia somente). O *Google Books N-gram Viewer* não disponibilizou qualquer resultado para a EI ***give * a hot potato***. Apesar de haver resultados para *milk and water* e *a small beer*, esses não estavam em seu uso metafórico. Os resultados encontrados foram compatíveis àqueles do *Webcorp*, portanto.

Ilustramos a necessidade das EI's estarem dentro de um contexto com o seguinte exemplo: ao inserirmos somente a EI ***a piece of cake*** para ser traduzida pelos TAO's fora de uma sentença ou contexto, provavelmente seria traduzida como 'um pedaço de bolo' que é sua tradução não metafórica, totalmente possível em vários contextos, tal como em "A menina comeu mais um pedaço de bolo". Contudo, ao inserirmos dentro de um contexto do tipo "*Mary told me her Math test was a piece of cake. She'll get an A*", como fizemos em nosso pequeno teste relatado na seção de Introdução desta pesquisa, seria equivocado um ser humano traduzir a EI ***a piece of cake*** como 'um pedaço de bolo'. Nesse caso, esperávamos que os TAO's também reconhecessem o uso da EI e fizessem a tradução como "Maria me disse que seu teste de matemática foi fácil" ou melhor, utilizando uma EI equivalente no português que também tivesse um alimento em sua composição lexical, poderia traduzir, neste caso, como "Maria me disse que seu teste de matemática foi ***mamão com açúcar***", por exemplo. Dessa forma, pudemos avaliar se os TAO's e THP's reconheceram as EI's e suas metáforas conceituais subjacentes, traduzindo-as apropriadamente.⁴⁴

Para a nossa análise final, restringimos nossa lista a 31 excertos contendo as EI's, englobando diferentes gêneros textuais e datas de uso, mas todas retiradas da parte do texto das ocorrências geradas pelo *Webcorp*. Nenhum título foi utilizado. A justificativa para esta decisão foi que um dos critérios definidos pelo grupo de estudos em Linguística Cognitiva liderado pela Profa. Paula Lenz requeria que o contexto fosse suficiente para que os tradutores pudessem identificar que o excerto

⁴⁴ Refizemos esse pequeno teste ao final da pesquisa e verificamos que somente o *Bing*, em sua versão mais recente, reconheceu a EI ***a piece of cake***. Apesar do *Google Translate* apresentar o significado metafórico desta EI fora do contexto (como uma das opções de tradução), quando dentro do excerto, a EI foi traduzida de forma não metafórica. Da mesma forma, o SYSTRAN não reconheceu esta EI em nenhuma de suas versões (padrão e neural *demo*). Contudo, nem o dicionário desse TAO apresentou seu significado metafórico.

continha uma EI e não palavras em seu uso não figurado, algo que um título, talvez, não pudesse trazer. A razão para a redução no número de EI's foi o fato de algumas delas serem muito similares em seu significado, mudando somente um verbo ou outro termo (***rise to the bait / swallow the bait / take the bait***, por exemplo), contendo a mesma metáfora conceitual subjacente. Essa medida precisou ser adotada também para facilitar a tradução dos THP's e não dar pistas sobre o objeto da pesquisa nos excertos traduzidos por eles. Outra razão foi o fato do *Webcorp* não ter gerado nenhuma concordância em seu uso metafórico para algumas EI's (***milk and water***) ou de forma alguma para outras (**queimar o óleo quarenta**). Para as EI's ***to make no bones about something*** (não ter papas na língua) e ***to have nothing to do with the price of tea in China*** (não ter nada a ver com o peixe) a razão para a exclusão foi o fato de não identificarmos metáforas conceituais subjacentes para ambas.

Além dos critérios mencionados, outros critérios foram estabelecidos para a escolha dos excertos a serem traduzidos:

- a) serem exemplos de fala real que contenham uma EI metafórica com um alimento em sua composição lexical, retirados do *Webcorp*;
- b) não serem demasiadamente longos;
- c) não conterem ambiguidade de sentido entre o uso metafórico e não metafórico da EI; e
- d) terem suficiência de contexto, isto é, que o contexto apresentado seja suficiente para o tradutor compreender a EI.

Nossos *corpora* finais foram compostos pelas traduções dos 31 excertos elaborados pelos três TAO's na primeira coleta (31 x 3 = 93), adicionadas da segunda coleta (31 x 4 = 124) e pelos três THP's (35 x 3 = 93), totalizando 310 traduções candidatas, ou seja, as traduções que foram submetidas às três métricas de avaliação de tradução, gerando, pelo menos, 930 resultados a serem analisados qualitativa e quantitativamente. Na seção 4, apresentamos e discutimos detalhadamente esses resultados.

3.3.2 Submissão do projeto de tese ao Comitê de Ética

Durante a seleção das EI's e dos seus respectivos excertos, conforme supracitado, elaboramos o Projeto de Pesquisa para ser submetido ao Comitê de Ética em Pesquisa da Universidade Estadual do Ceará (UECE), o qual foi submetido em

outubro de 2016. A participação dos THP's na parte experimental da presente pesquisa requereu esse procedimento. Juntamente com o projeto, foram elaborados e submetidos o Termo de Consentimento Livre e Esclarecido (conf. APÊNDICE B) e o formulário do Perfil do Tradutor Humano (conf. APÊNDICE C). O referido projeto foi aprovado em dezembro de 2016. Enquanto aguardávamos a aprovação, iniciamos a tradução dos excertos pelos TAO's.

As próximas etapas fazem parte da vertente experimental da pesquisa e são melhor detalhadas na subseção seguinte.

3.4 DESENHO EXPERIMENTAL

3.4.1 Tradução pelos TAO's e pelos THP's

Os sistemas de tradução automática *online* utilizados foram os mesmos avaliados por Cunha (2013b), que investigou a eficiência do *Bing Tradutor*, *GoogleTranslate* e SYSTRAN na tradução de expressões nominais com o caso genitivo e gerúndio que, por sua vez, se baseou em Toral *et al* (2011) e Chen *et al* (2012). Levou-se também em consideração o fato de serem bastante conhecidos e utilizados, provavelmente, por serem de acesso fácil e gratuitos, assim como disponibilizarem vários pares de línguas. O fato de eles utilizarem paradigmas de tradução diferentes é bastante relevante. Dentre os sistemas que compuseram este estudo, dois utilizam paradigmas não linguísticos para a Tradução Automática, o *Bing Tradutor* e o *GoogleTranslate*, que usam técnicas baseadas em rede neural. Já o SYSTRAN usa um paradigma híbrido, com técnicas baseadas em regras e técnicas estatísticas. A avaliação dos diferentes sistemas poderá evidenciar se a técnica utilizada influencia nos resultados obtidos e, portanto, qual a mais apropriada para a tradução automática de EI's, objeto de nosso estudo.

O período de acesso para a tradução das EI's foi o mesmo para os três sistemas de tradução automática *online* por nós avaliados. As traduções automáticas para a primeira coleta foram elaboradas no dia 21 de março de 2017. As traduções produzidas por cada um deles compõem nosso primeiro *corpus* de traduções candidatas, sobre as quais explanamos melhor na subseção seguinte.

3.4.1.1 Traduções Candidatas

Traduções candidatas são aquelas que são realizadas por ferramentas de Tradução Automática, quando comparadas com traduções realizadas por um especialista em traduções (tradução de referência). Em uma tradução ideal, a sentença candidata deve ser semelhante à de referência em tamanho, escolha e ordem de palavras. (MELO, MATOS e DIAS, 2014, p. 36). As traduções candidatas produzidas pelos sistemas de tradução automática *online* pesquisados são parte do nosso primeiro *corpus* avaliado e foram coletadas a partir da tradução dos excertos. Foram produzidas 31 traduções candidatas por cada um dos três TAO's avaliados, totalizando 93 traduções candidatas para esse *corpus*.

Também foram consideradas traduções candidatas (nosso segundo *corpus*) aquelas produzidas pelos três THP's. Cada THP foi contatado individualmente para confirmar sua participação no experimento e assinar o Termo de Consentimento Livre e Esclarecido. Juntamente com o Termo, enviamos, via e-mail, um pequeno questionário com perguntas para ser preenchido pelos participantes e devolvido à pesquisadora responsável. Este questionário visava à coleta de informações pessoais, profissionais e acadêmicas que ajudaram a identificar o perfil de cada participante. Também foram enviados aos THP's os excertos a serem traduzidos (APÊNDICE D), assim como as instruções necessárias para a elaboração das traduções e o procedimento a ser adotado ao finalizá-las, a fim de enviar os excertos traduzidos para a pesquisadora. Cada tradutor elaborou a tradução das sentenças via processo não supervisionado, em seu próprio domicílio ou ambiente de trabalho, de acordo com sua disponibilidade de tempo. Apesar de haver sido estabelecido o prazo de 30 dias, todos o extrapolaram.

Antes de finalizarmos nossas análises e submeter as traduções candidatas às métricas escolhidas, decidimos fazer uma segunda coleta de traduções candidatas elaboradas pelos TAO's (nosso terceiro *corpus*), a qual foi realizada em dezembro de 2017. O objetivo dessa segunda coleta foi verificar se houve alguma mudança nas traduções produzidas, uma vez que esses sistemas podem ser atualizados frequentemente, possuem memória de tradução e recebem sugestões de aprimoramento dos próprios usuários. O procedimento foi o mesmo utilizado na primeira coleta, com exceção para o *Google Translate*, do qual coletamos duas opções de tradução disponíveis nessa etapa.

Apesar de os TAO's e THP's terem traduzido o excerto completo, somente a EI de cada excerto foi submetida às métricas. Antes de descrever o procedimento para a avaliação das traduções, explicaremos como foram elaboradas as traduções de referência das EI's.

3.4.1.2 Traduções de Referência

Com o objetivo de avaliar os sistemas de tradução automática *online* escolhidos, precisamos de elementos de referência. Normalmente, as sentenças traduzidas pelos THP's seriam consideradas como nossas traduções de referência, uma vez que a tradução elaborada por humanos é considerada o parâmetro para as métricas de avaliação de tradução. Porém, para esta pesquisa, as traduções de referência são as traduções dos 31 excertos retirados do *Webcorp* realizadas pela pesquisadora e referendadas pelo grupo de estudos sobre Metáforas e Cognição, orientado pela Profa. Dra. Paula Lenz Costa Lima (conf. APÊNDICE E). A tradução que o *Dictionary of Metaphoric Idioms Portuguese-English* traz para cada uma das EI's com alimentos estudadas também foi utilizada para a composição das traduções de referência. Esse procedimento foi adotado porque nossos objetivos incluíam comparar a análise da percepção das metáforas subjacentes às EI's idiomáticas tanto pelos TAO's quanto pelos THP's.

Tanto as traduções produzidas pelos TAO's (primeira e segunda coletas) quanto aquelas produzidas pelos THP's foram submetidas às métricas escolhidas, utilizando nossas traduções de referência como parâmetro. Contudo, antes de explorarmos essa etapa, abriremos parênteses para descrever o perfil dos THP's delineado a partir do formulário Perfil do Tradutor Humano (conf. APÊNDICE C).

3.4.1.3 Perfil dos THP's

Talvez a maior dificuldade encontrada para a execução da pesquisa tenha sido encontrar THP's que aceitassem participar da tarefa proposta. Foram contatados, via *e-mail*, todos os tradutores juramentados que estão na lista da Junta Comercial do Estado do Ceará. Não obtivemos retorno da maioria. Alguns responderam declinando a solicitação. Contudo, uns poucos, por serem mais próximos, aceitaram participar. Enviamos o material para ser traduzido, mas nem todos completaram a tarefa. Ao

final, tivemos apenas três, dos cinco THP's que havíamos planejado compor este estudo. Acreditamos que, em parte, o motivo para a recusa ou não interesse em participar foi o número de palavras que compõem os excertos a serem traduzidos, 1.739 no total. Com o intuito de dar mais contexto e de não deixar as EI's muito explícitas para os THP's, acabamos pecando nesse aspecto, deixando os excertos um pouco longo. Imaginávamos que o tempo dado para a conclusão da tarefa, 30 dias, seria suficiente para que a tradução fosse elaborada sem se configurar em um fardo para os participantes. Contudo, todos extrapolaram esse prazo.

O formulário para compor o perfil geral dos THP's participantes foi enviado juntamente com os excertos para serem traduzidos e o Termo de Consentimento Livre e Esclarecido (TCLE). Seu objetivo era identificar o grau de experiência dos profissionais, assim como conhecer se os THP's fazem uso de ferramentas que os auxiliem no processo tradutório, mais especificamente, se utilizam os TAO's que estamos avaliando.

Os três participantes são jovens entre 33 e 41 anos de idade, dois do sexo feminino e um do masculino. Todos brasileiros concursados pela Junta Comercial do Estado do Ceará, habilitados para elaborarem traduções da língua inglesa. Todos concluíram sua graduação entre 2000 e 2005. Dos três, dois são graduados em Letras (Português e Inglês) e um em Publicidade e Propaganda. Todos também possuem pós-graduação, dois deles em Tradução. Um dos THP's tem mais de 15 anos de experiência profissional em tradução de texto e os outros dois entre 6 e 10 anos.

Os recursos normalmente utilizados pelos três THP's em sua atuação profissional são os corretores ortográficos (*Word*) e diversos dicionários e glossários digitais. Porém, um deles declarou também fazer uso de dicionários impressos. Os corretores gramaticais (*Word*) são utilizados por dois deles, bem como memória de tradução (*Wordfast*, *SDL Trados* e *MemoQ*) e tradutores automáticos. Nesse caso, eles especificaram que fazem uso do *Google Translate*. Apenas um dos THP's afirmou utilizar bancos de dados terminológicos (*PROZ*) e outras ferramentas não listadas por nós: conversores de formato de texto, alinhadores de texto, aglutinadores de texto, editores de PDF, ferramentas de análise de *corpus* – diversos. Nenhum dos THP's utiliza bancos de dados de estruturas sintáticas.

A partir do que expomos acima, podemos concluir que os três THP's são experientes e estão perfeitamente habilitados para cumprir a tarefa a eles proposta. É relevante frisar que dois dos três fazem uso do TAO *Google Translate* para auxiliá-los

em suas atividades como tradutores profissionais e que o uso de ferramentas digitais é algo comum a todos.

3.4.2 Submissão às métricas de avaliação de Tradução Automática

Para avaliar nossas traduções candidatas elaboradas pelos TAO's e THP's, usamos as três métricas descritas no item 2.2.4: a implementação de Zwart da IBM BLEU, *F-measure* e METEOR. São métricas já bastante conhecidas pelos pesquisadores de avaliação de Tradução Automática, mas que, segundo Cunha (2013b), ainda não são muito citadas pela literatura dessa área em português. Conforme mencionamos anteriormente, somente a parte que continha a EI foi submetida às métricas e não todo o excerto traduzido. As EI's foram inseridas em um documento em formato .txt, de acordo com a formatação exigida por cada uma das métricas. Os resultados foram limpos e tabulados para facilitar sua visualização e análise, a fim de atingir nossos objetivos secundários: somente a pontuação dos TAO's, somente a pontuação dos THP's e, finalmente, a comparação entre ambos. Esses resultados, bem como os dos outros objetivos secundários são apresentados na seção 4. Antes de apresentá-los, descrevemos como as análises foram elaboradas.

3.5 DESCRIÇÃO DAS ANÁLISES

Apresentamos o passo-a-passo das análises organizado de acordo com os cinco objetivos secundários desta pesquisa. O primeiro deles é relacionado aos aspectos culturais que podem influenciar qual termo de alimento compõe as EI's.

3.5.1 Identificação de aspectos culturais relacionados aos alimentos que compõem as EI's em estudo

O primeiro passo foi listar as 43 EI's selecionadas e suas respectivas traduções. Depois, identificamos se o mesmo alimento era utilizado nas duas línguas estudadas (inglês e português). O passo seguinte foi tentar traçar a origem das EI's, com o objetivo de verificar sua motivação, de forma que pudessemos identificar se havia influência cultural para o fato de algumas EI's terem o mesmo alimento e outras

não. O passo seguinte foi analisar se o tipo de alimento (doce, salgado etc) exercia alguma influencia em sua escolha para compor a EI.

O próximo passo do processo diz respeito ao nosso segundo objetivo secundário:

3.5.2 Identificação das metáforas conceituais subjacentes às EI

Para identificar as metáforas conceituais subjacentes às EI's, analisamos cada EI e mapeamos os aspectos do domínio alvo (quais características dos alimentos eram mapeadas) para tentar associá-los a um domínio fonte. Pesquisamos bancos de metáforas conceituais existentes, dentre eles o *MetaNet*, a fim de verificar se as EI's eram exemplos de metáforas conceituais já identificadas por outros estudiosos. Para aquelas que não pudemos associar à uma metáfora já identificada, classificamos em grupos de acordo com novas metáforas conceituais identificadas por nós. Quando não foi possível identificar a metáfora conceitual subjacente, registramos o fato. Para visualizar melhor os grupos de EI's e suas respectivas metáforas subjacentes, organizamos em um quadro que é apresentado no subitem 4.2.

3.5.3 Avaliação das traduções das EI's produzidas por TAO's, considerando a análise das metáforas conceituais

Nosso terceiro objetivo secundário era avaliar as traduções das EI's com alimentos produzidas pelos três TAO's escolhidos, considerando sua percepção às metáforas conceituais subjacentes. O primeiro passo para esse objetivo foi estabelecer uma tradução de referência para os excertos traduzidos. Essa tradução serviu de parâmetro para as métricas automáticas de avaliação e também para nossa métrica semântica. Nesse momento, restringimos o número de EI's de 43 para 31, conforme anteriormente explicado (subitem 3.3.1).

Os 31 excertos foram submetidos aos três TAO's e o produto tradutório resultante foi organizado em um quadro. Elaboramos nossa análise semântica de cada uma das EI's utilizando a tradução de referência e os seguintes critérios:

- a) TAO reconheceu a EI e utilizou a tradução de referência/alimento;
- b) TAO reconheceu a EI, mas a tradução não contém um alimento;
- c) TAO fez a tradução não figurada, mas é possível compreender a EI;

- d) TAO não reconheceu a EI; e
- e) TAO não traduziu ou deixou parte em branco.

Os resultados foram interpretados e colocados em uma tabela, de forma que pudessem ser melhor visualizados. Essa tradução elaborada pelos TAO's em 21 de março de 2017 foi nomeada primeira coleta. Produzimos uma segunda coleta em dezembro de 2017, na qual todos os passos foram repetidos, com o acréscimo de uma segunda opção de tradução que agora era disponibilizada pelo TAO2 (*Google Translate*). Ao final, os dados das duas coletas foram consolidados em um gráfico, de forma que pudemos verificar o desempenho dos três TAO's em relação à sua percepção das metáforas conceituais subjacentes às EI's analisadas.

O passo seguinte está relacionado ao quarto objetivo secundário proposto para a presente pesquisa: a avaliação das traduções das EI's produzidas pelos três THP's, a qual é melhor detalhada no subitem seguinte.

3.5.4 Avaliação das traduções das EI's produzidas por THP's, considerando a análise das metáforas conceituais

O primeiro passo desta parte da pesquisa teve início com a submissão de seu projeto ao Comitê de Ética da Universidade Estadual do Ceará (UECE). Essa etapa aconteceu concomitante à elaboração das traduções pelos TAO's. Enquanto aguardávamos a aprovação pelo referido comitê, demos início as análises das traduções elaboradas pelos três TAO's, etapa descrita no subitem anterior. Tão logo obtivemos a aprovação, entramos em contato com os THP's concursados pela Junta Comercial do Estado do Ceará, para que eles pudessem participar da pesquisa. O retorno não foi o esperado e tivemos que reduzir o número previsto de participantes de cinco para três.

Confirmados os participantes, solicitamos que eles assinassem o Termo de Consentimento Livre e Esclarecido, preenchessem o questionário do Perfil do Tradutor Humano e realizassem a mesma tarefa que os TAO's, ou seja, traduzir os 31 excertos selecionados do *Webcorp*. Após o prazo de 30 dias, que havia sido concedido para a realização da tarefa, não obtivemos resposta e precisamos ampliar esse prazo para que todos pudessem finalizar as traduções.

Após recebermos as traduções elaboradas, os passos seguintes foram os mesmo que desenvolvemos com o produto tradutório dos TAO's, ou seja, os excertos

traduzidos foram organizados em um quadro e elaboramos nossa análise semântica de cada uma das EI's utilizando a tradução de referência e os seguintes critérios (os mesmos utilizados para analisar os TAO's):

- a) THP reconheceu a EI e utilizou a tradução de referência/alimento;
- b) THP reconheceu a EI, mas a tradução não contém um alimento;
- c) THP fez a tradução não figurada, mas é possível compreender a EI;
- d) THP não reconheceu a EI; e
- e) THP não traduziu ou deixou parte em branco.

Os resultados foram interpretados e colocados em uma tabela, de forma que pudessem ser melhor visualizados. Não houve uma segunda coleta para os THP's, pois não se fazia necessária.

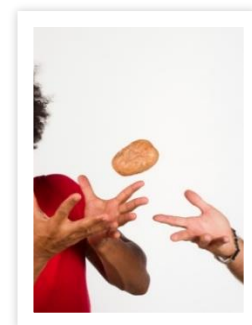
Uma vez que todas as traduções candidatas haviam sido elaboradas (nossos *corpora*), o passo seguinte foi fazer o tratamento necessário para aplicar as métricas de avaliação de tradução escolhidas, o quinto e último objetivo secundário de nosso estudo.

3.5.5 Análise das métricas de avaliação de tradução automática

A última etapa da pesquisa foi a análise das métricas de avaliação de tradução. Todas as traduções elaboradas pelos TAO's e THP's (traduções candidatas) e a tradução de referência foram tratadas para serem processadas pelas três métricas escolhidas BLEU, *F-measure* e METEOR. As traduções das EI's foram isoladas dos excertos e salvas em documentos em formato .txt, um para cada coleta de cada tipo de tradutor (três para a primeira coleta dos TAO; quatro para a segunda coleta dos TAO; e três para a tradução dos THP's), perfazendo um total de 10 documentos.

Após serem processados pelas métricas, os resultados foram limpos e organizados em tabelas para facilitar sua visualização e para que as análises fossem realizadas. Elaboramos tabelas dos resultados de todos os tradutores (TAO's e THP's) e separadamente para cada tipo, de forma que pudéssemos identificar os maiores e menores resultados gerais e de cada tipo de tradutor. Foram também elaborados gráficos. Apresentamos todos os resultados e análises na seção seguinte.

4 ANÁLISE E DISCUSSÃO – *HANDLING THE HOT POTATO*



Esta seção traz a análise e a discussão dos dados coletados, para que os objetivos da presente pesquisa fossem alcançados. Ela está dividida em cinco subitens, cada um relacionado, respectivamente, aos cinco objetivos secundários propostos na Introdução, Quadro 1 (p. 31). São eles:

- a) Identificar se há aspectos culturais envolvidos no tipo de alimento, que compõe as EI's em estudo;
- b) Verificar se há metáforas conceituais subjacentes a todas as EI's em estudo;
- c) Avaliar as traduções de EI's com alimentos produzidas por TAO's;
- d) Avaliar traduções de EI's com alimentos produzidas por THP's, via processo não supervisionado; e
- e) Comparar as traduções realizadas pelos TAO's e THP's com uma tradução de referência estabelecida, utilizando métricas de avaliação de Tradução Automática e de metáforas conceituais subjacentes.

4.1 IDENTIFICAÇÃO DE ASPECTOS CULTURAIS RELACIONADOS AOS ALIMENTOS QUE COMPÕEM AS EI'S EM ESTUDO

A fim de atingir nosso primeiro objetivo secundário, estudamos as 43 EI's selecionadas e suas respectivas traduções, de forma que pudéssemos identificar suas metáforas conceituais subjacentes e responder à questão levantada: como os alimentos são usados metaforicamente na construção das EI's. Esta análise evidenciou que as metáforas subjacentes estão, na maioria das vezes, ligadas a facilidades ou dificuldades, aspectos positivos ou negativos, relacionados aos alimentos que as compõem. Exemplificamos com as EI's ***have a hot potato, have a***

lemon e **a hard nut to crack** (estar com uma batata quente nas mãos, ter/estar com um abacaxi nas mãos e ser/ter um osso duro de roer, respectivamente) para alimentos, que são de difícil manuseio ou ingestão, que causam um certo desconforto, enquanto que expressões como **sell like hotcakes** e **a piece of cake** (vender como água e mamão com mel/açúcar, respectivamente) se utilizam de alimentos que são de fácil ingestão ou consumo.

Outro aspecto que se constata é a influência cultural na escolha dos alimentos, que compõem as EI's. Os alimentos que trazem esses aspectos positivos ou negativos para cada EI são, geralmente, diferentes nas duas línguas estudadas. Se em inglês, a EI se constitui por '*a hard nut*', em português, temos 'um osso duro'. Uma amêndoa (*nut*) não é um alimento muito comum ou popular no Brasil. Já um osso, degusta-se em cozidos e frangos, dentre outros alimentos. Outro exemplo: **to be a small beer** e **ser café pequeno**. Sendo o café uma bebida muito mais típica do Brasil do que uma cerveja (*beer*).

Contudo, encontramos o mesmo alimento nas duas línguas, como em '*a hot potato*' e 'uma batata quente'. Apesar de haver também um uso bastante expressivo de EI's com 'abacaxi', que é um fruto tropical, menos comum nos países de língua inglesa, se comparados ao Brasil. Já a batata é um alimento básico para esses povos. A mesma expressão com 'abacaxi' é também utilizada, como correspondente em português, para as EI's **get out of a jam** e **have a lemon on your hands**.

Algumas EI's são praticamente iguais em ambos os idiomas estudados nesta pesquisa. As EI's, que possuem 'pão' (*bread*) em sua composição, são exemplos disso: **to be our daily bread**, **to earn one's daily bread**, **to take the bread out of someone's mouth**, e **to live on the breadline**. Atribuímos essa semelhança à importância que esse alimento tem em várias culturas. Outra possível explicação seria seu uso na cultura cristã, em diversas passagens da Bíblia e na oração do Pai Nosso (**our daily bread**). O mesmo, pensamos, se aplica para a EI **to separate the wheat from the chaff** (separar o joio do trigo), que vem de uma passagem bíblica.

Outras EI's que têm traduções bastante semelhantes em inglês e português são as que contêm 'egg' (ovo) e 'chicken' (galinha) em sua composição lexical: **to kill the goose that lays the golden eggs**, **to count one's chickens before they hatch**, **to walk on eggs(hells)**. Galinhas e ovos são alimentos bastante comuns, tanto no Brasil, quanto em países como os EUA, Canadá e Reino Unido.

Para algumas EI's, o que as determina é a atividade relacionada ao alimento e não o próprio alimento. Analisando as correspondências entre as línguas portuguesa e inglesa para as EI's de nosso estudo, identifica-se que as atividades contidas nas EI's também são culturalmente dependentes. Quando as atividades são comuns nas duas culturas estudadas, permanecem. Do contrário, se utiliza outra atividade ou ação, por exemplo, para as EI's ***to swallow the bait, take the bait, rise to the bait***. Peixes mordem iscas em qualquer cultura, portanto o equivalente em português também faz uso da mesma atividade ***morder a isca***. Já em ***to upset the/sm's apple cart***, em português é ***entornar o caldo***. Carros de maçãs não se mostram comuns no Brasil, mas caldos fazem parte do nosso universo cultural e popular.

As informações relatadas corroboram nossa revisão literária, no que se refere à origem de algumas EI's serem culturalmente influenciadas. Da mesma forma, levam-nos a acreditar que há empréstimo de uma língua para outra, provavelmente provenientes de traduções de livros mundialmente utilizados, como é o caso da Bíblia e das fábulas de Esopo (***tirar a castanha/sardinha com a mão do gato***). Alimentos comuns em diferentes culturas, como pão, ovo e galinha, tendem a ser utilizados nas EI's em ambas as línguas neste estudo. Contudo, aqueles que não fazem parte do universo cotidiano, tendem a ser substituídos por outros, que são mais típicos daquele lugar, como é o caso de *geese* e *swans* (gansos e cisnes) na EI ***All one's geese to be swans*** cujo equivalente em português é ***comer frango e arrotar peru/faisão***. Frangos e perus figuram mais como alimentos próximos de nossa cultura alimentar do que gansos e cisnes.

Feita a análise da influência cultural na formação das EI's estudadas, passamos para nosso segundo objetivo secundário que é a identificação das metáforas conceituais subjacentes.

4.2 METÁFORAS CONCEITUAIS SUBJACENTES ÀS EI'S

O segundo objetivo secundário tinha a seguinte pergunta como ponto de partida: qual metáfora conceitual subjaz cada EI analisada? O Quadro 12 apresenta as metáforas conceituais identificadas por nós para cada uma delas:

Quadro 12 – Metáforas conceituais subjacentes identificadas

Nº	EXPRESSÃO IDIOMÁTICA INGLÊS	METÁFORA	EXPRESSÃO IDIOMÁTICA PORTUGUÊS
01	To get out of a jam	DIFICULDADE É UM ALIMENTO DIFÍCIL DE MANIPULAR/INGERIR	Descascar o abacaxi ; ter/estar com um abacaxi nas mãos
02	To handle the/a hot potato		Descascar o abacaxi ; ter/estar com um abacaxi nas mãos
03	To be left with the hot potato		Ficar com o abacaxi
04	To give sm the hot potato		Passar o abacaxi
05	To have a hot potato		Estar com uma batata quente nas mãos
06	To be/have a hard nut to crack		Ser/Ter um osso duro de roer
07	To have a lemon on your hands		Descascar o abacaxi ; ter/estar com um abacaxi nas mãos
08	A (pretty) kettle of fish		Um angu de caroço
09	To upset the/sm's apple cart		Entornar o caldo
10	To be butterfingers		Ter mão(s) de manteiga
11	Not to cut the mustard any more		Queimar óleo quarenta
12	To pull the chestnuts out of/from the fire		Tirar a castanha/sardinha com a mão do gato
13	To walk on eggs/eggshells		Pisar em/sobre ovos
14	To be no piece of cake	FACILIDADE É UM ALIMENTO FÁCIL DE MANIPULAR/INGERIR	Não ser bolinho
15	A piece of cake		Mamão com mel/açúcar
16	To be a small beer		Ser café pequeno
17	To sell like hot cakes/hotcakes		Vender como água
18	To be a cake -eater/cake eater		Ser arroz de festa
19	To be cheese paring	RECURSOS SÃO ALIMENTOS (STATUS SOCIAL É UM ALIMENTO CARO/BARATO)	Ser pão duro
20	To count one's chickens before they hatch		Contar com o ovo na galinha
21	To cry over spilled/spilt milk		Chorar (sobre) o leite derramado
22	To live off the fat of the land		Viver à tripa forra
23	To have a finger in the pie		Tirar uma casquinha
24	To earn one's daily bread		Ganhar o pão nosso de cada dia
25	To be our daily bread		Ser o pão nosso de cada dia
26	To take the bread out of sm's mouth		Tirar o pão da boca de a.
27	To live on the breadline		Passar a pão e água
28	All one's geese to be swans	Comer frango e arrotar peru/faisão	
29	To swallow the bait	TER GANÂNCIA É QUERER UM ALIMENTO	Morder a isca
30	To rise to the bait		Morder a isca
31	To take the bait		Morder a isca
32	To kill the goose that lays the golden eggs		Matar a galinha dos ovos de ouro
33	To separate the wheat from the chaff	BOAS IDEIAS SÃO BONS ALIMENTOS/IDEIAS RUINS SÃO ALIMENTOS RUINS	Separar o joio do trigo
34	Milk and water		Sem sal

35	To be neither fish nor flesh	IDEIAS DESINTERESSANTES SÃO ALIMENTOS SEM GOSTO (DEFINIDO)/IDEIAS INTERESSANTES SÃO ALIMENTOS BONS	Não ser carne nem peixe
36	To waffle	EMBROMAÇÃO É ALIMENTO DE LONGA PREPARAÇÃO	Encher linguiça
37	To be all talk and no cider		Ser mais as vezes que as nozes
38	To imbibe sth with one's mother's milk	APRENDER É INGERIR ALIMENTO NUTRITIVO	Beber algo com o leite materno
39	To make no bones about sth	Não identificada	Não ter papas na língua
40	To have nothing to do with the price of tea in China	Não identificada	Não ter nada a ver com o peixe
41	To pour oil on troubled waters	RAIVA É UM LÍQUIDO QUENTE EM UM RECIPIENTE/JOGAR LÍQUIDO NA PANELA FERVENTE É DIMINUIR A RAIVA ⁴⁵	Jogar/Lançar/Por água (fria) na fervura
42	To throw cold water		Jogar/Lançar/Por água (fria) na fervura
43	To pour cold water		Jogar/Lançar/Por água (fria) na fervura

Fonte: Elaborado pela autora.

A metáfora identificada, que licencia o maior número de EI's é DIFICULDADE/FACILIDADE É UM ALIMENTO DIFÍCIL/FÁCIL DE MANIPULAR/INGERIR – 18 ao todo. A fim de deixar mais evidente a visualização dessa oposição entre dificuldade (DIFICULDADE É UM ALIMENTO DIFÍCIL DE MANIPULAR/INGERIR) e facilidade (FACILIDADE É UM ALIMENTO FÁCIL DE MANIPULAR/INGERIR), separamos essa metáfora em dois grupos no Quadro 12. Contudo, entendemos que fazem parte de uma única metáfora. Nossa análise evidenciou que os mapeamentos das metáforas subjacentes estão ligados às experiências corpóreas com os alimentos que funcionam como domínio fonte, seja por aspectos positivos ou negativos. Porém, pudemos perceber que o número de EI's que expressam dificuldade (13) é bem maior do que o número de EI's que expressam facilidade (5). Alguns exemplos, que expressam dificuldade são **have a hot potato**, **have a lemon** e **a hard nut to crack** (estar com uma batata quente nas mãos, ter/estar com um abacaxi nas mãos e ser/ter um osso duro de roer, respectivamente). Ressaltamos que a EI **no piece of cake** poderia ser acrescentada a essa lista, uma vez que faz a negação da facilidade.

⁴⁵ Ao analisarmos melhor essas EI's, percebemos que, apesar de serem compostas pelo alimento água, neste caso, elas são licenciadas pelas metáforas conceituais RAIVA É UM LÍQUIDO QUENTE EM UM RECIPIENTE e JOGAR LÍQUIDO NA PANELA FERVENTE É DIMINUIR A RAIVA.

Os alimentos encontrados causam uma experiência corpórea de desconforto por diferentes razões: características eventuais (estarem quentes - *hot potato*/batata quente); paladar (acidez – *lemon*/abacaxi); textura (espinhosa – abacaxi); ou consistência (rigidez – *nut*/osso). Essas sensações é que são mapeadas, originando assim as metáforas.

As características que evidenciam a facilidade na experiência corpórea com os alimentos podem ser semelhantes às da dificuldade como, por exemplo, estarem quentes (*hotcakes* – nesse caso, evidenciando o frescor do alimento, que acabou de ser preparado). Outras características encontradas foram o paladar (doçura – mel/açúcar); a consistência (maciez – bolo/mamão); o tamanho (*small beer*/café pequeno/bolinho) e a frequência de degustação do alimento (arroz de festa).

Em relação às motivações para o aparecimento das metáforas, no caso da EI **a kettle of fish**, há o esquema de imagem da situação caótica e desorganizada, bem como a experiência corpórea de desconforto gerada por um *picnic* às margens de um corpo de água, após uma pescaria, o que originou a EI em inglês. Já para **um angu de caroço**, seu correspondente em português, a motivação é possivelmente histórica⁴⁶, com origem no tempo da escravidão quando, na hora da alimentação, os escravos recebiam seu alimento em uma cuia, o qual era, normalmente, uma porção de angu de fubá. A escrava que os servia, sempre que podia, escondia algum pedaço de carne, ou algum torresmo, debaixo do angu. Quando outro escravo percebia isso, costumava comentar o fato com o seu companheiro, comumente usando a expressão: "debaixo desse angu tem caroço". Daí o significado de que alguém está escondendo alguma coisa. Contudo, atualmente, a EI tem sido mais utilizada em situações onde algo dá resultado contrário ao previsto; desordem; complicação ou confusão. Talvez, por ser esse o provável resultado, caso os escravos fossem pegos em tal prática ou a própria dificuldade de encontrar o pedaço de carne sob o angu.

Também situações de esquema de imagem motivam as EI's **upset someone's apple cart** e **entornar o caldo**⁴⁷. A EI em inglês tem origem romana e, em português, há registro de que, originalmente, era **entornar o carro**, bem mais próxima de seu correspondente em língua inglesa. Ambas significam complicar a

⁴⁶Disponível em: <<https://www.geledes.org.br/9-expressoes-populares-com-origens-ligadas-escravidao-e-voce-nem-imaginava/>>. Acesso em: 15 out. 2017.

⁴⁷ Disponível em: <<http://ditoscuriosos.blogspot.com.br/2011/09/entornar-o-caldo.html>>. Acesso em: 15 out. 2017.

situação, estragar uma situação com um ato desastrado ou intempestivo. É fácil remeter à situação de dificuldade, quando um carro de maçãs (ou frutas) tomba e o esforço necessário para coletar todas de volta. Da mesma forma, uma panela com caldo. Esses esquemas de imagem facilitam a compreensão de ambas EI's, tornando seus níveis de idiomaticidade mais baixos.

A EI ***to pull the chestnuts out of/from the fire*** (tirar a castanha/sardinha com a mão do gato) tem provável origem na fábula O Macaco e o Gato, de Esopo. Sendo ele um autor mundialmente conhecido, que teve suas fábulas traduzidas para várias línguas, acreditamos que essa seja a explicação para que o mesmo alimento tenha permanecido no correspondente em português (castanha), bem como uma possível adaptação para um alimento mais tipicamente brasileiro (sardinha).

A EI ***not to cut the mustard any more*** e seu correspondente **queimar óleo quarenta** representam um desafio em relação à identificação da motivação de suas origens, pois não há informações precisas. Apesar de terem sido incluídas nesta análise de identificação das metáforas subjacentes, foram excluídas da lista de EI's traduzidas pelos TAO's e THP's, uma vez que o *Webcorp* gerou pouquíssimos exemplos de uso para esta EI em inglês e nenhum para seu correspondente em português.

Por outro lado, os alimentos encontrados, que causam uma experiência corpórea de facilidade, para a mesma metáfora conceitual, se utilizam de alimentos em pequenas quantidades, encontrados com frequência e que são de fácil ingestão ou consumo, nesse caso, evidenciando características positivas, como estar quente (bolo) ou a consistência macia (bolo/mamão) ou ainda o sabor doce de todos esses alimentos – ***sell like hotcakes*** e ***a piece of cake*** (vender como água e mamão com mel/açúcar/sopa no mel, respectivamente). É interessante notar que para bolos 'estar quente' é uma característica positiva, pelo menos em nossa cultura, pois mostra o frescor do alimento, que acabou de sair do forno, diferentemente de uma batata quente.

O uso dos alimentos *cake* e *pie* como símbolo de 'facilidade' e 'deleite' pode ser encontrado em outras expressões na língua inglesa, tais como ***as easy as pie***, ***a cake-walk*** e ***that takes the cake/biscuit***.⁴⁸ Da mesma forma, a negação da facilidade

⁴⁸ Disponível em: <<https://www.phrases.org.uk/meanings/piece-of-cake.html>>. Acesso em: 17 out. 2017.

se encontra em *to be no picnic*, *to be no piece of cake* e seu equivalente em português **não ser bolinho**.

A EI *to be a cake-eater/cake eater* tem diferentes significados figurados, quando se refere somente ao sexo masculino (*playboy*, *bon-vivant* ou um homem afeminado) e, também quando não faz diferença entre gêneros, pois mapeia a facilidade da vida ou a doçura dos modos. O significado equivalente em português, **ser arroz de festa**⁴⁹ tem origem histórica, remontando a uma sobremesa (arroz-doce), que era frequentemente servida em festas, fazendo a associação entre a infalibilidade/facilidade da presença desse alimento em festas, tanto quanto a presença da pessoa.

Sobre o número de ocorrências geradas para as EI's, pesquisadas no *Webcorp*, houve uma variação bastante relevante – algumas EI's geraram mais de 300 ocorrências, enquanto outras, gerou apenas uma. Eliminamos do total as concordâncias repetidas e todas nas quais as EI's era o título de livros, músicas, nomes de restaurantes etc. Houve um pequeno número de concordâncias em que não foi possível identificar se a expressão estava em seu sentido metafórico ou não.

Dentre as concordâncias consideradas válidas, geradas pelo *Webcorp*, ou seja, após a eliminação supracitada, a EI *a small beer* foi a que obteve o maior número de ocorrências com significado não metafórico. Contudo, ainda metonímico (47 de 49). Outras EI's, que geraram um número alto de concordâncias não metafóricas foram *have a hot potato* (3 de 7); *a pretty kettle of fish* (23 de 40); e *walk on eggs* (22 de 38). Faz-se interessante perceber que a variação *a kettle of fish* gerou um número bem maior de ocorrências metafóricas (20 de 24). As EI's que tiveram todas as suas concordância metafóricas foram *walk on eggshells* (47); *handle the/a hot potato* (8); *be left with the hot potato* (4); *give sm the hot potato* (1); *be/have a hard nut to crack* (44); *have a lemon on your hands* (4); *be butterfingers* (9); *not to cut the mustard anymore* (12); *pull the chestnuts out of the fire* (34); *be no piece of cake* (40); *sell like hot cakes* (28) /*hotcakes* (30); e *be a cake-eater* (2) /*cake eater* (26). As EI's remanescentes não apresentaram um número alto de ocorrências não metafóricas, no máximo quatro.

Em nossa análise dos exemplos gerados pelo *Webcorp*, houve dúvidas em alguns momentos sobre se o texto em questão estava sendo usado em seu sentido

⁴⁹ Disponível em: <<http://www.historiadetudo.com/arroz-de-festa>>. Acesso em: 17 out. 2017.

metafórico ou não, uma vez que a EI utilizada estava dentro do mesmo campo semântico do texto em questão. Percebemos essa tendência, principalmente, em títulos e textos jornalísticos (e.g. *upsetting Apple's cart*, referindo-se a empresa de tecnologia), assim como em comentários e opiniões do leitor ou usuário de algum serviço. Entendemos que o uso desse recurso é para chamar a atenção do leitor. Sobre esse aspecto Gibbs *et al* (1997, p. 147) nos apontam uma possível explicação:

... alguns usos literais de expressões idiomáticas na verdade expressam ambos os significados literal e figurado. Portanto, quando alguém está literalmente patinando em gelo fino, está também em uma situação perigosa (i.e. o significado figurado da expressão patinar em gelo fino) (tradução nossa)⁵⁰.

Identificamos também a metáfora RECURSOS SÃO ALIMENTOS, que subjaz várias das EI's analisadas. Os aspectos emprestados do domínio fonte são economia de recursos (*to be cheese paring* – **ser pão duro**), desperdiçar recursos (*to cry over spilled milk* – **chorar sobre o leite derramado**), ter abundância de recursos (*to live off the fat of the land* – **viver à tripa forra**), contar com recursos (*to count one's chicken before they hatch* – **contar com o ovo na galinha**), tirar vantagem (*to have a finger in the pie* – **tirar uma casquinha**) e identificar o que é recurso (*to separate the wheat from the chaff* – **separar o joio do trigo**).

Um acarretamento dessa metáfora é STATUS SOCIAL É UM ALIMENTO CARO/BARATO. Percebe-se que várias EI's têm *bread* (pão) em sua composição, alimento básico, que se encontra diariamente na mesa de muitos povos. Quanto mais básico o alimento, provavelmente mais baixa será a condição financeira, daí a associação de pão à miséria. Alguns exemplos dessa metáfora são as EI's *to live on the breadline* e *all one's geese to be swans* (**passar a pão e água e comer frango e arrotar peru/faisão**, respectivamente).

Para as EI's *to swallow the bait*, *take the bait*, *rise to the bait* (**morder a isca**) e *to kill the chicken that lay the golden eggs* (**matar a galinha dos ovos de ouro**) identificamos a metáfora conceitual TER GANÂNCIA É QUERER UM ALIMENTO, sendo um acarretamento da metáfora conceitual GANÂNCIA É FOME.

⁵⁰ ... some literal uses of idioms actually express both literal and figurative meanings. Thus, when one is literally skating on thin ice, one is also in a dangerous situation (i.e., the figurative meaning of the phrase skating on thin ice).

Identificamos alguns acarretamentos para a metáfora conceitual IDEIAS SÃO ALIMENTOS: BOAS IDEIAS SÃO BONS ALIMENTOS e, semelhantemente, sua oposição IDEIAS RUINS SÃO ALIMENTOS RUINS, sendo essa a motivação para a EI ***to separate the wheat from the chaff*** (separar o joio do trigo); IDEIAS DESINTERESSANTES SÃO ALIMENTOS SEM GOSTO (DEFINIDO) é outro acarretamento de IDEIAS SÃO ALIMENTOS, que motiva as EI's ***milk and water*** (sem sal) e ***to be neither fish nor flesh*** (não ser carne nem peixe). É importante mencionar que esse acarretamento já foi estendido para objetos, pessoas, situações, como podemos ver em expressões do tipo ***isto é só o filé; é mesmo que queijo; é o milho debulhado***.

A metáfora conceitual EMBROMAÇÃO É UM ALIMENTO DE LONGA PREPARAÇÃO subjaz as EI's ***to waffle*** (encher linguiça) e ***to be all talk and no cider*** (ser mais as vozes que as nozes). A experiência corporificada de dedicar muito tempo no preparo do alimento motiva a noção de perda de tempo que encontramos no significado dessas expressões. Essa metáfora também subjaz as expressões ***está me cozinhando*** (em banho maria), por exemplo.

A metáfora conceitual CONHECIMENTO É ALIMENTO acarreta APRENDER É INGERIR ALIMENTO NUTRITIVO, a qual motiva as EI's ***to imbibe something with one's mother milk*** (beber algo com o leite materno).

Não identificamos metáforas para as EI's ***to make no bones about something*** (não ter papas na língua) e ***to have nothing to do with the price of tea in China*** (não ter nada a ver com o peixe). Portanto, essas EI's foram retiradas da parte experimental da pesquisa, conforme mencionado anteriormente.

A hipótese para o segundo objetivo secundário, ou seja, que a maioria das EI's analisadas com alimentos seria licenciada por metáforas conceituais, foi confirmada, uma vez que identificamos metáforas conceituais para a maioria das EI's. Também é importante se salientar que algumas das metáforas conceituais identificadas não estão listadas no banco de dados da *MetaNet*, o que nos leva a crer que nossa contribuição para essa ferramenta será bastante relevante.

4.3 AVALIAÇÃO DAS TRADUÇÕES DAS EI'S PRODUZIDAS POR TAO'S, CONSIDERANDO A ANÁLISE DAS METÁFORAS CONCEITUAIS

Para nosso terceiro objetivo secundário ser atingido (avaliar traduções de EI's com alimentos produzidas por TAO's, considerando a análise das metáforas conceituais), apenas 31 dos 43 excertos selecionados do *Webcorp* foram submetidos aos três TAO's, conforme explicado na seção 3, Percurso Metodológico.

Os 31 excertos foram submetidos aos três TAO's: *Bing Tradutor* (TAO1), *Google Translate* (TAO2) e *SYSTRAN* (TAO3). O produto tradutório nos trouxe resultados, em que há vários problemas de tradução, seja por uma escolha lexical inapropriada, por uma sequência incorreta de palavras ou até pela não tradução, a qual ocorreu de duas formas 1) trazendo a palavra não traduzida na língua de origem ou 2) a omissão de palavras na língua de chegada (\emptyset). No entanto, não daremos ênfase a esses problemas, por não serem foco desta pesquisa e, portanto, não serão relatados de forma mais detalhada aqui. Restringiremo-nos apenas aos resultados encontrados, relacionados ao (não) reconhecimento das EI's e às escolhas lexicais, caso elas interfiram na compreensão do significado das EI's. Para tal, apresentamos o quadro resumo com as traduções elaboradas por cada TAO:

Quadro 13 – Traduções dos excertos pelos TAO's

Nº	Excerto	TAO1 (<i>Bing</i>)	TAO2 (<i>Google</i>)	TAO3 (<i>SYSTRAN</i>)
01	So the current stance of fiscal policy may well have been seriously inappropriate before the Brexit vote upset the apple cart .	chateado o carrinho de maçã	virar o carrinho de maçã	virado o carro da maçã
02	People (who earn) over £100,000 will actually be paying more, I make no bones about that	não faço nenhum osso sobre isso	não fazer nenhum osso sobre isso	fazem-no a nenhuns ossos sobre aquele
03	Four million children in this country live on the breadline	vivem a miséria	vivem \emptyset	vivem no breadline
04	The Butterfingers [Team] will cross bats with the Salt Lake Juniors at Calder's Park Tuesday	Butterfingers	Butterfingers	Butterfingers
05	I am a cake-eater and have a sedentary job.	sou um comedor de bolo	sou um bolo-comedor	sou um bolo-comedor
06	Hobby's Toskana Exclusive D750 H UCC may not have made much impression on German showgoers but it is sure to sell like hot cakes here	vendem como bolos quentes	para vender como bolos quentes	vender como bolos quentes

07	As usual, it is the business of the British Government to pull the chestnuts out of the fire.	para tirar as castanhas do fogo	para puxar as castanhas para fora do fogo	para puxar as castanhas fora do fogo
08	You know what happens when you count your chickens before they hatch.	contar suas galinhas antes que choquem	conta suas galinhas antes que elas chocam	conta suas galinhas antes que choquem
09	In a more typical situation, one does not have to walk on eggs all the time.	andar sobre ovos	andar Ø ovos	andar em ovos
10	How can you live off the fat of the land when you make no contribution to the country?	viver da gordura da terra	viver da gordura da terra	viver fora da gordura da terra
11	The muddle is all the fault of the Johnny-come-lately politicians who like to have a finger in the pie.	ter um dedo no bolo	ter um dedo na torta	ter um dedo na torta
12	The movie is neither fish nor flesh	é nem carne nem peixe	não é peixe nem carne	é nem peixe nem carne
13	Joe Mercer once said of Bill Shankly that he 'thinks all his geese are swans' .	pensa que todos os gansos são cisnes	pensa que todos seus gansos são cisnes	pensa que todos seus gansos são cisnes
14	Theme park owners need to wise-up to this or they'll kill the goose that lays the golden eggs.	matar a galinha dos ovos de ouro	matar o ganso que coloca os ovos de ouro	matarão o ganso que coloca os ovos dourados
15	Borrowers think they just need a little more time to get out of a jam.	para conseguir sair de uma encrenca	para sair de um congestionamento	de sair de um doce
16	But sorting out Scotland is going to cost a pretty penny, and pretty much everything could turn into a pretty kettle of fish.	uma linda chaleira de peixes	uma jarra de peixe bonita	uma chaleira de peixes bonita
17	I know, I should have planned to fly to southern Spain "first" and then on to Morocco, but it's too late to "cry over spilled milk" .	chorar sobre o leite derramado	chorar sobre o leite derramado	grita leite sobre derramado
18	For he was more learned than any of his predecessors, Abbots of Croyland, having imbibed literature of every description with his mother's milk from his very cradle.	embebidas literatura de todas as descrições com leite da mãe	embebido literatura de toda descrição com o leite de sua mãe	absorvendo a literatura de cada descrição com seu leite de mãe
19	It is a commonplace of philosophy that the notion of free will is a hard nut to crack.	é um osso duro de roer	é uma loucura difícil de quebrar	é uma porca dura a rachar-se
20	Baker attempted to pour oil on troubled waters after a prosecutor lost her temper with a defence barrister in court.	derramar óleo sobre águas turvas	derramar óleo em águas turbulentas	derramar o óleo em águas incomodadas
21	It's a sign of another thing: that liberal men are wimps who can't handle the hot potato that is a combination of feminine sexuality and female political brilliance.	consegue lidar com a batata quente	segurar a batata quente	segurar a batata quente
22	When Mark Messier was captain for a period, before and rumors swirled about the Moose being gay, Bure wanted gone. But that all should have nothing to do with the price of tea in China.	ter nada a ver com o preço do chá na China	ter nada a ver com o preço do chá na China	ter nada fazer com o preço do chá em China
23	I place a lot more emphasis on soft skills now. For me, they separate the wheat from the chaff.	separam o trigo do joio	separam o trigo da palha	separam o trigo do debulho

24	But surely you can have all those things and still have a lemon on your hands ? You don't need those things to make a good Bond film.	tem um limão em suas mãos	tem um limão em suas mãos	comer um limão em suas mãos
25	It's a rare exercise class that can cater for mixed abilities and I can vouch that it's no piece of cake for fit people	não é nenhum pedaço de bolo	não é nenhum pedaço de bolo	não é nenhum pedaço de bolo
26	By delivering a 1p cut in the price of a pint this one deserves to be called the small beer budget.	pequena esplanado	cerveja pequena	pequeno da cerveja
27	Be prepared with what to expect in such an interview and ensure that you never rise to the bait no matter how important it seems.	se levanta para a isca	subir para a isca	aumente à isca
28	A student in the audience for Thursday night's televised debate on the EU referendum caused a stir on social media after accusing David Cameron of " waffling ".	devaneios	waffling	Waffling
29	It would be sheer madness to contemplate even for a moment giving up Britain's independent nuclear deterrent. And there is no room for compromise, and no room for cheese paring .	emparelhamento de queijo	queijo paring	casca do queijo
30	Attention to detail is our daily bread and we love it.	é o nosso pão de cada dia	é o nosso pão diário	é nosso pão diário
31	now let's set about to do something, not all talk and no cider , but let's see to it that something is done along these lines.	não só conversa e nenhuma cidra	nem todos falar e não cidra	não todos falam e nenhuma cidra

Legenda:

	TAO reconheceu a EI e utilizou a tradução de referência/com alimento
	TAO reconheceu a EI, mas a tradução não contém um alimento
	TAO fez a tradução não figurada, mas é possível compreender a EI
	TAO não reconheceu a EI
	TAO não traduziu ou deixou parte em branco

Fonte: Elaborado pela autora.

Analisando de forma geral, dentre as 31 EI's traduzidas pelos três TAO's, 12 não foram percebidas por nenhum deles, e portanto, traduzidas de forma não figurada, comprometendo sua compreensão; outras cinco EI's também não foram percebidas pelos três TAO's, mas sua tradução não figurada ainda permite a sua compreensão, devido à sua transparência; uma EI não foi traduzida por nenhum deles (**Butterfingers**) – acreditamos que foi percebida como um nome próprio (do time de beisebol); as 13 EI's restantes foram tratadas de diferentes formas pelos TAO's.

Comparando as traduções elaboradas pelos três TAO's e o Quadro 11 (EI's classificadas dentro do *continuum* de idiomaticidade), verifica-se que as EI's, que foram classificadas em língua inglesa como transparentes, em sua maioria, foram traduzidas para o português de forma não figurada, mas ainda assim eram

compreensíveis. As EI's classificadas como semitransparentes foram percebidas de diferentes formas: traduzidas de forma figurada, traduzidas de forma não figurada, não traduzidas e não reconhecidas. Todas as EI's classificadas como semi-opacas não foram reconhecidas pelos TAO's, exceto **Butterfingers** que não foi traduzida de forma alguma, conforme supracitado. As EI's classificadas como opacas, em sua maioria, também não foram reconhecidas pelos TAO's ou foram traduzidas de forma não figurada, sendo uma delas não traduzida (**waffle**). Isso mostra que o grau de idiomaticidade, de fato, tem influência na percepção dos TAO's em relação às EI's.

A partir do Quadro 13 e utilizando o mesmo esquema de cores de sua legenda, elaboramos a Tabela 1, que nos permite visualizar de forma mais clara o desempenho dos três TAO's em relação à sua percepção das EI's e, posteriormente, descrever o desempenho de cada um individualmente.

Tabela 1 – Desempenho dos TAO's na elaboração das traduções da primeira coleta

TAO	Reconheceu e utilizou	Reconheceu, mas não utilizou	Tradução não figurada	Não reconheceu	Não traduziu
TAO1	4	3	10	13	1
TAO2	2	-	9	15	5
TAO3	1	-	8	19	3

Fonte: Elaborada pela autora.

O TAO1 (*Bing Tradutor*) reconheceu e traduziu sete EI's em seu sentido metafórico, dentre as quais quatro também continham uma EI com alimentos para a tradução em português. As outras três traduções mostravam um reconhecimento da metáfora conceitual subjacente, trazendo um sentido metafórico, mas não utilizavam uma EI com alimentos. Por exemplo, para a EI **live on the breadline**, a tradução para o português foi 'vivem a miséria'; apesar de ter traduzido de forma metafórica, não fez uso da EI **passam a pão e água**, que também contém um alimento em sua composição lexical e poderia ter sido usada neste caso.

É importante salientar que estas traduções são consideradas adequadas e tão importantes quanto as que utilizaram nossa tradução de referência, especialmente quando o grau de transparência da EI é menor, como é o caso de **waffling** e **get out of a jam**. O fato do TAO reconhecê-las é mais notável do que reconhecer, por

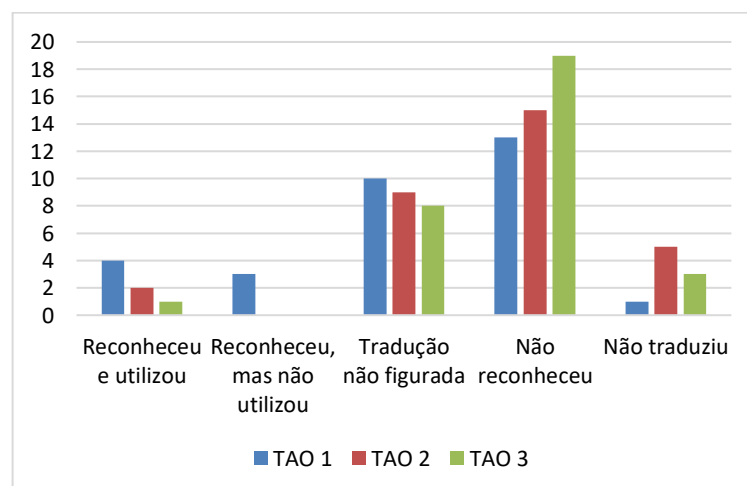
exemplo, ***cry over spilled milk***, que é compreensível com sua tradução não figurada. Se levarmos em consideração as traduções elaboradas com sentido não figurado (10), mas que são compreensíveis devido o seu grau de transparência, o TAO1 foi o tradutor que obteve o maior desempenho de todos, considerando que dentre as 31 EI's, 17 seriam possíveis de serem compreendidas pelo usuário e 14 não. Em termos proporcionais isso representa 54,8%.

O TAO2 (*Google Translate*) reconheceu e traduziu em sentido metafórico apenas duas das EI's. Porém, ambas são facilmente compreensíveis, se for utilizada também uma tradução não figurada: ***cry over spilled milk*** e ***handle the hot potato***. Se somarmos as traduções metafóricas com alimentos (2) e as traduzidas não metafóricamente (9), teremos um total de apenas 11 EI's compreensíveis para esse TAO. Isso representa 35,4%.

O TAO3 (SYSTRAN) foi o que reconheceu o menor número de EI's e suas metáforas subjacentes, apenas uma: ***handle the hot potato***. Conforme mencionamos, uma EI facilmente compreensível, se for utilizada uma tradução não figurada. Somando as traduções metafóricas com alimentos (1) e as traduzidas não metafóricamente (8), teremos um total de apenas nove EI's compreensíveis para esse TAO. Isso representa apenas 29%, sendo o TAO com menor percentual de percepção das EI's, dentre os pesquisados.

Outra forma de mostrar os resultados obtidos pode ser observada no Gráfico 1.

Gráfico 1 – Desempenho dos TAO's na elaboração das traduções da primeira coleta



Fonte: Elaborado pela autora.

Visualiza-se que os maiores resultados estão nos três últimos tipos de tradução realizada pelos TAO's (tradução com linguagem não figurada, não reconheceu a EI e não a traduziu). A partir desses resultados, evidencia-se que os TAO's não percebem, na maioria das vezes, as metáforas conceituais que licenciam as EI's. Essa era nossa segunda hipótese, a qual acreditamos ter sido corroborada pela presente pesquisa.

Contudo, no mundo digital, tudo se renova com muita rapidez. Comprovando-se diversas vezes no decorrer de nossa pesquisa: a mudança do motor de busca do *Webcorp* por duas vezes, a transformação na técnica de Tradução Automática do *Bing Tradutor* e *Google Translate* de estatística para neural, ou a mudança do sistema do SYSTRAN de baseado em regras para híbrido e mais recentemente a introdução da versão neural *demo*; o número de línguas traduzidas pelos diferentes sistemas, as possibilidades de tradução de imagens e conversas, dentre outras. Levando-se em consideração que a coleta das traduções candidatas dos TAO's aconteceu em março de 2017 e que os TAO's, em geral, contam com a colaboração de seus usuários para aperfeiçoar as traduções realizadas, bem como modificaram suas técnicas de tradução, conduzimos uma nova coleta em dezembro de 2017, para uma melhor verificação de resultados, os quais apresentamos no Quadro 14:





Quadro 14 – Traduções dos excertos pelos TAO's na segunda coleta

Nº	Excerto	TAO1 (Bing)	TAO2 (Google)	TAO2 (Google - 2ª opção)	TAO3 (SYSTRAN)
01	So the current stance of fiscal policy may well have been seriously inappropriate before the Brexit vote upset the apple cart .	chateado o carrinho de maçã	abalou o carrinho da maçã	virar a mesa	virado o carro da maçã
02	People (who earn) over £100,000 will actually be paying more, I make no bones about that	não faço nenhum osso sobre isso	não faço nenhum osso sobre isso	não fazer nenhum osso sobre isso	fazem-no a nenhuns ossos sobre aquele
03	Four million children in this country live on the breadline	vivem na padaria	vivem na linha do pão	vivem na miséria	vivem no breadline
04	The Butterfingers [Team] will cross bats with the Salt Lake Juniors at Calder's Park Tuesday	os dedos de manteiga	The Butterfingers	Os Butterfingers	O Butterfingers
05	I am a cake-eater and have a sedentary job.	Sou um comedor de bolos	sou um comedor de bolo	sou um bolo-comedor	sou um bolo-comedor
06	Hobby's Toskana Exclusive D750 H UCC may not have made much impression on German showgoers but it is sure to sell like hot cakes here	vender como bolos quentes	vende como bolos quentes	vender como pão quente	vender como bolos quentes
07	As usual, it is the business of the British Government to pull the chestnuts out of the fire .	puxar as castanhas para fora do fogo	puxar as castanhas para fora do fogo	puxar as castanhas do fogo	puxar as castanhas fora do fogo
08	You know what happens when you count your chickens before they hatch .	conta suas galinhas antes de eclodirem	conta suas galinhas antes de chocar	contar as galinhas antes que choquem	conta suas galinhas antes que choquem
09	In a more typical situation, one does not have to walk on eggs all the time.	andar em ovos	andar com ovos	andar em ovos	andar em ovos
10	How can you live off the fat of the land when you make no contribution to the country?	viver fora da gordura da terra	viver da gordura da terra	viver fora da gordura da terra	viver fora da gordura da terra
11	The muddle is all the fault of the Johnny-come-lately politicians who like to have a finger in the pie .	ter um dedo na torta	ter um dedo na torta	ter um dedo na torta	ter um dedo na torta
12	The movie is neither fish nor flesh	não é nem peixe nem carne	não é peixe nem carne	não é nem peixe nem carne	é nem peixe nem carne

13	Joe Mercer once said of Bill Shankly that he 'thinks all his geese are swans'.	todos os seus gansos são cisnes	todos os seus gansos são cisnes	todos os seus gansos são cisnes	todos seus gansos são cisnes
14	Theme park owners need to wise-up to this or they'll kill the goose that lays the golden eggs.	vão matar o ganso que coloca os ovos de ouro	matarão o ganso que coloca os ovos de ouro	vão matar a galinha dos ovos de ouro	matarão o ganso que coloca os ovos dourados
15	Borrowers think they just need a little more time to get out of a jam.	sair de uma Jam	sair de uma merda	sair de uma enrascada	sair de um doce
16	But sorting out Scotland is going to cost a pretty penny, and pretty much everything could turn into a pretty kettle of fish.	uma chaleira bonita dos peixes	uma bela chaleira de peixe	uma chaleira muito de peixe	uma chaleira de peixes bonita
17	I know, I should have planned to fly to southern Spain "first" and then on to Morocco, but it's too late to cry over spilled milk".	chorar sobre o leite derramado	chorar sobre o leite derramado	chorar sobre o leite derramado	grita leite sobre derramado
18	For he was more learned than any of his predecessors, Abbots of Croyland, having imbibed literature of every description with his mother's milk from his very cradle.	absorvido a literatura de todas as descrições com o leite de sua mãe	absorvido literatura de todas as descrições com o leite de sua mãe	absorvido literatura de cada descrição com o leite de sua mãe	absorvendo a literatura de cada descrição com seu leite de mãe
19	It is a commonplace of philosophy that the notion of free will is a hard nut to crack.	é um osso duro de roer	é uma porca dura para se quebrar	é um osso duro de roer	é uma porca dura a rachar-se
20	Baker attempted to pour oil on troubled waters after a prosecutor lost her temper with a defence barrister in court.	derramar petróleo em águas turbulentas	despejar óleo em águas turbulentas	deitar água na fervura	derramar o óleo em águas incomodadas
21	It's a sign of another thing: that liberal men are wimps who can't handle the hot potato that is a combination of feminine sexuality and female political brilliance.	lidar com a batata quente	lidar com a batata quente	lidar com a batata quente	segurar a batata quente
22	When Mark Messier was captain for a period, before and rumors swirled about the Moose being gay, Bure wanted gone. But that all should have nothing to do with the price of tea in China.	ter nada a ver com o preço do chá na China	ter nada a ver com o preço do chá na China	ter nada a ver com o preço do chá na China	ter nada fazer com o preço do chá em China
23	I place a lot more emphasis on soft skills now. For me, they separate the wheat from the chaff.	separam o trigo do joio	separam o trigo da palha	separar o trigo do joio	separam o trigo do debulho

24	But surely you can have all those things and still have a lemon on your hands ? You don't need those things to make a good Bond film.	ter um limão em suas mãos	tem um limão em suas mãos	tem um limão em suas mãos	comer um limão em suas mãos
25	It's a rare exercise class that can cater for mixed abilities and I can vouch that it's no piece of cake for fit people	não é pedaço de bolo	não seja um pedaço de bolo	não é nenhum pedaço de bolo	não é nenhum pedaço de bolo
26	By delivering a 1p cut in the price of a pint this one deserves to be called the small beer budget.	cerveja pequena	cerveja pequena	cerveja pequena	pequeno da cerveja
27	Be prepared with what to expect in such an interview and ensure that you never rise to the bait no matter how important it seems.	se levante à isca	se levante para a isca	mordeu a isca	aumente à isca
28	A student in the audience for Thursday night's televised debate on the EU referendum caused a stir on social media after accusing David Cameron of " waffling ".	waffle	waffling	Waffling	waffling
29	It would be sheer madness to contemplate even for a moment giving up Britain's independent nuclear deterrent. And there is no room for compromise, and no room for cheese paring .	queijo paring	separação do queijo	queijo emparelhamento	casca do queijo
30	Attention to detail is our daily bread and we love it.	é o nosso pão de cada dia	é o nosso pão diário	é o nosso pão de cada dia	é nosso pão diário
31	now let's set about to do something, not all talk and no cider , but let's see to it that something is done along these lines.	nem toda a conversa e nenhuma cidra	nem toda conversa e sem cidra	não só conversa e nenhuma cidra	não todos falam e nenhuma cidra

Legenda:

	TAO aprimorou a tradução
	TAO deixou a tradução mais incompreensível
	TAO fez pequenas alterações que não ajudaram na compreensão
	TAO não fez qualquer alteração na tradução

Fonte: Elaborado pela autora.

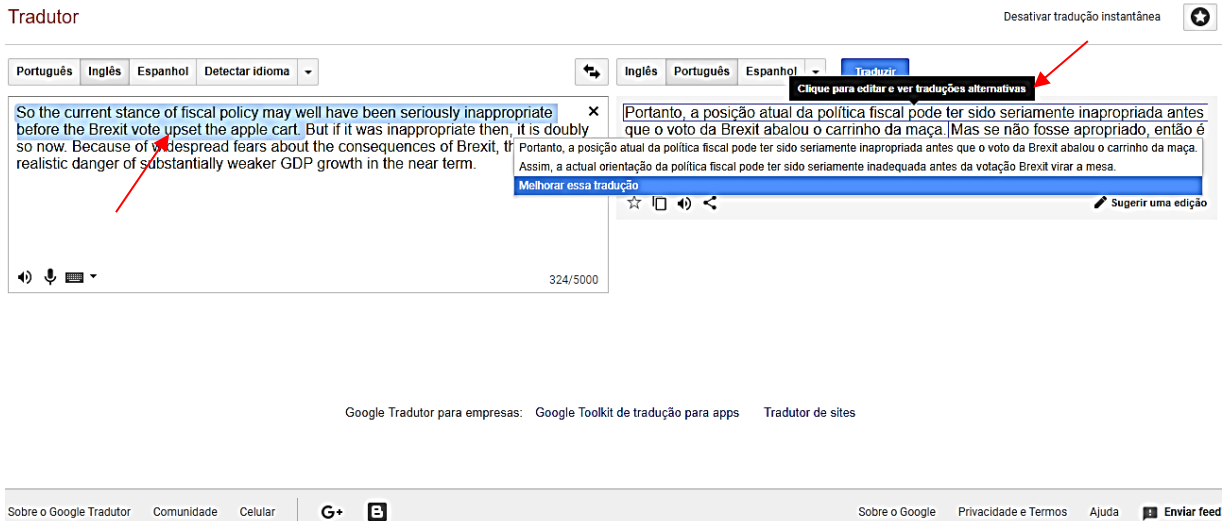
O TAO1 (*Bing*) produziu 10 traduções iguais e 21 diferentes das elaboradas na primeira coleta. A maioria das modificações consistiu em pequenas mudanças de tempos verbais (vendem > vender como bolos quentes), preposições (andar sobre > em ovos); mudança de palavras (contar suas galinhas antes que choquem > de eclodirem); ou elaboração da tradução, quando antes não havia sido realizada (queijo *paring* > emparelhamento de queijo).

Contudo, surpreendeu-nos o fato de seis traduções, outrora consideradas mais próximas da tradução de referência, tornarem-se mais incompreensíveis nessa segunda coleta (vivem a miséria > vivem na padaria; viver da gordura da terra > viver fora da gordura da terra; matar a galinha dos ovos de ouro > matar o ganso que coloca os ovos de ouro; para conseguir sair de uma encrenca > sair de uma Jam; derramar óleo sobre águas turvas > derramar petróleo em águas turbulentas; e devaneios > *waffle*. Houve duas traduções que melhoraram, equiparando-se às traduções de referência (*Butterfingers* > **os dedos de manteiga**; e nem carne nem peixe > **não é nem peixe nem carne**).

O TAO2 (*Google*) produziu 11 traduções iguais as que foram coletadas em março e 10 diferentes. Da mesma forma que o TAO1, o TAO2 fez modificações de algumas palavras, pequenas mudanças de tempos verbais e preposições, que não influenciaram muito no resultado. Contudo, aprimorou uma delas, apesar de utilizar uma linguagem chula (para sair de um congestionamento > sair de uma merda). Por outro lado, deixou outra mais incompreensível (não é nenhum pedaço de bolo > não seja um pedaço de bolo).

A segunda coleta do TAO2 trouxe a possibilidade de duas opções de tradução para cada excerto ou de seus diferentes segmentos, como podemos visualizar na Figura 31:

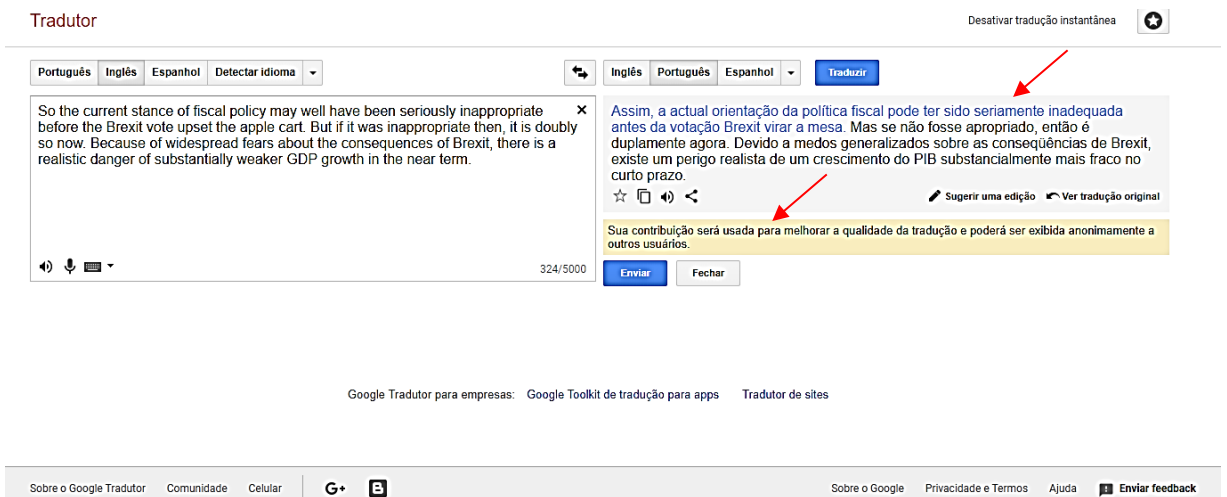
Figura 31 – Print do Google Translate



Fonte: Página do *Google Translate*.



Ao clicar na parte do texto, cuja opção de tradução você quer verificar, uma janela se abre, com a primeira tradução elaborada e a segunda opção. É possível ainda fazer uma sugestão de tradução. Caso você escolha a segunda opção, ela substitui a anterior no texto traduzido e aparece a opção de 'ver a tradução original', além do aviso de que "sua contribuição será usada para melhorar a qualidade da tradução e poderá ser exibida anonimamente a outros usuários".

Figura 32 – Print do Google Translate após escolhida a melhor opção



Fonte: Página do *Google Translate*.

As traduções elaboradas para essa segunda opção do TAO2 se apresentaram melhores do que aquelas produzidas em março e também do que as da primeira opção desta coleta, apesar de 10 dos 31 excertos ainda serem idênticos aos da primeira coleta. Dos 21 restantes, nove apresentaram algumas pequenas alterações que, como já citado, não ajudaram a compreensão das EI's; um teve sua compreensão ainda mais prejudicada (viver da gordura da terra > viver fora da gordura da terra); mas 11 tiveram sua tradução largamente aprimorada, muitas delas se equiparando às da tradução de referência.

O TAO3 (SYSTRAN) não apresentou qualquer mudança entre as traduções elaboradas nas duas coletas. Esse resultado era, de certa forma previsível, por ser o único, dentre os três TAO's analisados, que não oferece opção de participação do usuário para o aprimoramento das traduções. O TAO1 traz, abaixo da janela do texto traduzido, a opção de mostrar satisfação, com o resultado obtido, por meio dos ícones  e . O TAO2 é o mais aberto a sugestões do usuário. Na verdade, incentiva-o a fazer por meio do botão 'sugerir uma edição', que fica no canto inferior da caixa do texto traduzido.

Resumimos essas informações na Tabela 2, de forma que o desempenho dos TAO's, nessa segunda coleta fica mais perceptível.

Tabela 2 – Desempenho dos TAO's na elaboração das traduções da segunda coleta

TAO	Reconheceu e utilizou	Reconheceu, mas não utilizou	Tradução não figurada	Não reconheceu	Não traduziu
TAO1	5	-	8	16	2
TAO2	1	1	10	17	2
TAO2 2ª	9	3	7	10	2
TAO3	1	-	8	19	3

Fonte: Elaborada pela autora.

O TAO1 (*Bing Tradutor*) reconheceu e traduziu cinco EI's em seu sentido metafórico e todas continham uma EI com alimentos na tradução elaborada. Se levarmos em consideração as traduções com sentido não figurado (8), mas que são compreensíveis devido o seu grau de transparência, temos um total de 13 EI's possíveis de compreensão pelo usuário e 18 não compreensíveis. Em termos

proporcionais isso representa 41,9%, implica-se que houve uma redução no desempenho desse TAO, pois na primeira coleta o percentual fora de 54,8%.

Os resultados da primeira opção de tradução do TAO2 (*Google Translate*) mostram que ele reconheceu e traduziu, em sentido metafórico, apenas duas das EI's, para uma delas ele utilizou também uma expressão com alimentos. Se somarmos as traduções metafóricas com e sem alimentos (2) e as traduzidas não metaforicamente, mas compreensíveis (10), teremos um total de 12 EI's compreensíveis elaboradas por esse TAO, mostrando que houve um pequeno aumento de 34,4%, na primeira coleta, para 38,7%, mas ainda abaixo do desempenho do TAO1.

Contudo, a segunda opção de tradução disponibilizada pelo TAO2 na segunda coleta, obteve um maior desempenho. Na verdade, o maior de todos os TAO's analisados. Foram 12 EI's traduzidas metaforicamente, dentre as quais nove com alimentos. Somando as EI's àquelas que foram traduzidas não metaforicamente, mas que ainda assim são compreensíveis, temos um total de 19 EI's. Isso representa 61,2%. Um resultado bem acima dos 34,4% da primeira opção dessa coleta e também dos 38,7% da primeira coleta para esse TAO.

O TAO3 (SYSTRAN) não apresentou qualquer modificação nas traduções da primeira para a segunda coleta. Ou seja, continuou com apenas 29% de EI's compreensíveis e, novamente, o TAO com o menor percentual de percepção das EI's dentre os pesquisados. Os percentuais obtidos por cada TAO, em ambas as coletas, resumam-se na Tabela 3.

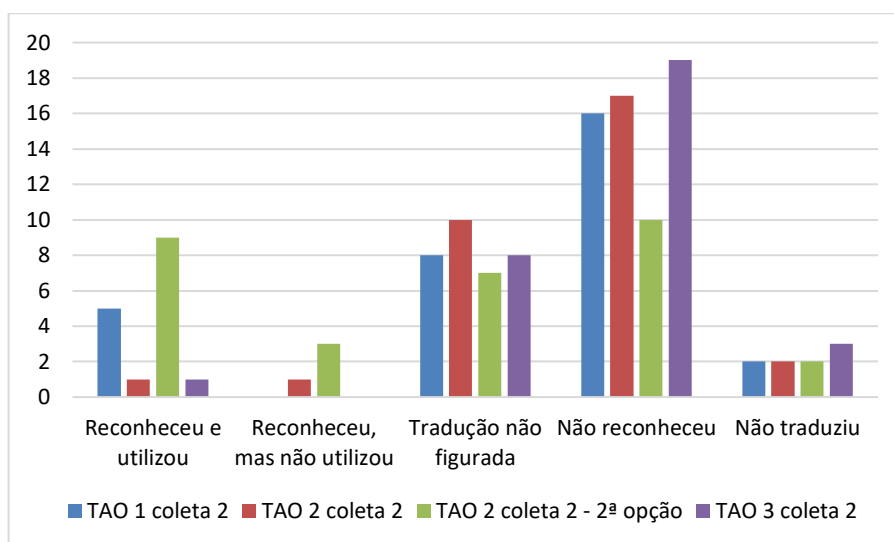
Tabela 3 – Resumo dos percentuais de desempenho dos TAO's na elaboração das traduções na segunda coleta

Tradutor:	1ª coleta	2ª coleta	2ª coleta – 2ª opção
TAO1	54,8%	41,9%	-
TAO2	35,4%	38,7%	61,2%
TAO3	29,0%	29,0%	-

Fonte: Elaborada pela autora.

Outra forma de mostrar os resultados obtidos pode ser observada no Gráfico 2.

Gráfico 2 – Desempenho dos TAO's na elaboração das traduções na segunda coleta



Fonte: Elaborado pela autora.

A partir desse gráfico, podemos visualizar claramente que o desempenho do TAO2, na segunda coleta, elevou-se com relação aos demais TAO's. Contudo, de modo geral, os três últimos tipos de traduções elaboradas (tradução com linguagem não figurada, não reconheceu a EI e não traduziu a EI) ainda representam os maiores valores obtidos para os TAO's.

Esses resultados nos levam a crer que a percepção dos TAO's às EI's e suas metáforas subjacentes, ainda são imprevisíveis, elaborando traduções não muito confiáveis e que precisam passar por uma pós-edição, especialmente, dependendo do objetivo da tradução, corroborando novamente nossa hipótese de número 2 (que os sistemas de tradução automática *online* não percebem as EI's nem às suas metáforas conceituais subjacentes e que fariam, na maioria das vezes, as traduções de forma não metafórica). A avaliação dos diferentes sistemas evidenciou que a técnica utilizada por cada TAO não parece ter influência nos resultados obtidos, uma vez que, apesar dos três sistemas utilizarem a técnica neural, a mais avançada atualmente, a percepção dos sistemas ainda é muito baixa. Mostra-nos também que é necessário que o usuário esteja atento às possíveis opções que os tradutores automáticos podem disponibilizar, pois elas podem passar despercebidas ao usuário, não conhecedor das possibilidades ofertadas.

4.4. AVALIAÇÃO DAS TRADUÇÕES DAS EI'S PRODUZIDAS POR THP'S, CONSIDERANDO A ANÁLISE DAS METÁFORAS CONCEITUAIS

Para atingir nosso quarto objetivo específico (avaliar traduções de EI's com alimentos produzidas por THP's, considerando a análise das metáforas conceituais), solicitamos que três THP's realizassem a mesma tarefa que os TAO's realizaram, ou seja, traduzir os 31 excertos selecionados do *Webcorp*, conforme explicado na seção 3, Percurso Metodológico.

O produto tradutório nos trouxe resultados muito mais próximos da nossa tradução de referência. Contudo, houve alguns problemas semelhantes aos obtidos pelos TAO's, porém em menor quantidade, tais como escolha lexical inapropriada, sequência não usual de palavras, não tradução, tanto trazendo a palavra na língua de origem ou nada (\emptyset). Da mesma forma que tratamos as traduções elaboradas pelos TAO's, versaremos pelas produzidas pelos THP's. Por questões éticas, a identidade dos THP's participantes, a quem mais uma vez agradecemos sua imensa contribuição para esta pesquisa, foi preservada e eles foram apenas numerados de 1 a 3. Apresentamos o Quadro 15 com o resumo das traduções elaboradas por cada um:

Quadro 15 – Traduções dos excertos pelos THP's

Nº	Excerto	THP1	THP2	THP3
01	So the current stance of fiscal policy may well have been seriously inappropriate before the Brexit vote upset the apple cart .	abalasse a política da Apple	confusão causada	estragar os planos feitos
02	People (who earn) over £100,000 will actually be paying more, I make no bones about that	não discuto isso...	pra ser sincero	não tenho a menor dúvida disso
03	Four million children in this country live on the breadline	vivem na linha da pobreza	vivem em extrema pobreza	vivem na miséria
04	The Butterfingers [Team] will cross bats with the Salt Lake Juniors at Calder's Park Tuesday	Butterfingers	Dedos de Manteiga	butterfingers (desengonçados)
05	I am a cake-eater and have a sedentary job.	sou um comedor de bolo	Sou um comedor de bolo	Sou um playboy
06	Hobby's Toskana Exclusive D750 H UCC may not have made much impression on German showgoers but it is sure to sell like hot cakes here	vender como pão quente	vender aqui como água no deserto	vender como água
07	As usual, it is the business of the British Government to pull the chestnuts out of the fire .	acalmar a situação	resolver o problema dos outros	apagar o fogo
08	You know what happens when you count your chickens before they hatch .	conta com os ovos antes que	conta com ovo "na cloaca" da galinha	conta com o que ainda não aconteceu

		as galinhas os ponham		
09	In a more typical situation, one does not have to walk on eggs all the time.	andar pisando em ovos	andar "pisando em ovos"	andar pisando em ovos
10	How can you live off the fat of the land when you make no contribution to the country?	viver dos recursos do país	viver tão confortavelmente	viver da fartura da terra
11	The muddle is all the fault of the Johnny-come-lately politicians who like to have a finger in the pie .	ter sua fatia do bolo	querem se intrometer	gostam de se meter
12	The movie is neither fish nor flesh	não é peixe nem carne	nem fede nem cheira	não é nem uma coisa nem outra
13	Joe Mercer once said of Bill Shankly that he 'thinks all his geese are swans' .	acha que todos os seus gansos são cisnes	todos os seus gansos são cisnes	é muito exagerado
14	Theme park owners need to wise-up to this or they'll kill the goose that lays the golden eggs .	matarão a galinha dos ovos de ouro	vão matar a sua galinha dos ovos de ouro	perderão a galinha que põe seus ovos de ouro
15	Borrowers think they just need a little more time to get out of a jam .	sair de uma situação crítica	sair do sufoco	sair do buraco
16	But sorting out Scotland is going to cost a pretty penny, and pretty much everything could turn into a pretty kettle of fish .	uma bela confusão	uma bela confusão	uma situação embaraçosa
17	I know, I should have planned to fly to southern Spain "first" and then on to Morocco, but it's too late to 'cry over spilled milk' .	chorar sobre o leite derramado	chorar sobre o leite derramado	chorar sobre o leite derramado
18	For he was more learned than any of his predecessors, Abbots of Croyland, having imbibed literature of every description with his mother's milk from his very cradle.	bebido da literatura de todas as origens com o leite de sua mãe	bebido de literatura de toda espécie junto com o leite de sua mãe	assimilado literatura de todas as definições Ø
19	It is a commonplace of philosophy that the notion of free will is a hard nut to crack .	é difícil de explicar	é um problema difícil de resolver	é uma tarefa difícil de cumprir
20	Baker attempted to pour oil on troubled waters after a prosecutor lost her temper with a defence barrister in court.	complicar uma situação turbulenta	acalmar as coisas	amenizar a situação
21	It's a sign of another thing: that liberal men are wimps who can't handle the hot potato that is a combination of feminine sexuality and female political brilliance.	lidar com a batata quente	segurar a batata quente	lidar com a batata quente
22	When Mark Messier was captain for a period, before and rumors swirled about the Moose being gay, Bure wanted gone. But that all should have nothing to do with the price of tea in China .	não deve ter nada a ver Ø	não devia ter a ver com o preço do chá na China	não tinha nada a ver com o assunto
23	I place a lot more emphasis on soft skills now. For me, they separate the wheat from the chaff .	separam o joio do trigo	separam o joio do trigo	separam o joio do trigo
24	But surely you can have all those things and still have a lemon on your hands ? You don't need those things to make a good Bond film.	ter um desastre nas mãos	ter uma porcaria nas mãos	ter um problema nas mãos

25	It's a rare exercise class that can cater for mixed abilities and I can vouch that it's no piece of cake for fit people	seria muito fácil	acharia demais fácil	não é fácil agradar
26	By delivering a 1p cut in the price of a pint this one deserves to be called the small beer budget.	cerveja pequena	cervejinha	orçamento trivial
27	Be prepared with what to expect in such an interview and ensure that you never rise to the bait no matter how important it seems.	morda a isca	cair na provocação	morder a isca
28	A student in the audience for Thursday night's televised debate on the EU referendum caused a stir on social media after accusing David Cameron of " waffling ".	"enrolar"	"embromação"	"ficar em cima do muro"
29	It would be sheer madness to contemplate even for a moment giving up Britain's independent nuclear deterrent. And there is no room for compromise, and no room for cheese paring .	para divisões	para avareza	compromissos possíveis Ø
30	Attention to detail is our daily bread and we love it.	é o nosso pão de cada dia	é o nosso ganha-pão	é nosso pão de cada dia
31	now let's set about to do something, not all talk and no cider , but let's see to it that something is done along these lines.	só conversa e nada de trabalho	só muita conversa e pouca ação	só papo, mas vamos garantir que algo é feito

Legenda:

	THP reconheceu a EI e utilizou a tradução de referência/com alimento
	THP reconheceu a EI, mas a tradução não contém um alimento
	THP fez a tradução não figurada, mas é possível compreender a EI
	THP não reconheceu a EI
	THP não traduziu ou deixou parte em branco

Fonte: Elaborado pela autora.

Em termos gerais, dentre as 31 EI's traduzidas pelos três THP's, 30 foram reconhecidas, se não por todos, por pelo menos um deles; 16 das EI's tiveram suas metáforas reconhecidas por todos os três THP's, das quais sete foram traduzidas com a tradução de referência ou com o mesmo alimento e as outras nove com outras escolhas lexicais, mas que tinham o significado correspondente. Apesar do três THP's terem reconhecido a EI **make no bones about that**, entendemos que o significado traduzido não é o mesmo da nossa tradução de referência. Percebemos que as traduções com linguagem não figuradas, apresentam-se em número bastante reduzido, principalmente se comparado com o número produzido pelos TAO's. Quando ocorreram, a maioria tornou-se compreensível, devido à sua transparência. Apenas um THP não traduziu a EI *Butterfingers* – acreditamos que foi reconhecida como um nome próprio (do time de beisebol). As EI's restantes foram interpretadas de diferentes formas pelos THP's.

A partir do Quadro 15 e utilizando o mesmo esquema de cores da legenda, elaboramos a Tabela 4, que nos permite visualizar de forma mais clara o desempenho dos três THP's em relação à sua percepção das EI's e, posteriormente, descrevemos o desempenho de cada um.

Tabela 4 – Desempenho dos THP na elaboração das traduções

THP	Reconheceu e utilizou	Reconheceu, mas não utilizou	Tradução não figurada	Não reconheceu	Não traduziu
THP1	12	9	4	4	2
THP2	10	16	4	1	-
THP3	8	19	-	2	2

Fonte: Elaborada pela autora.

O THP1 reconheceu a metáfora subjacente a 21 EI's, traduzindo 12 delas com EI's semelhantes às da tradução de referência ou outra expressão com o mesmo alimento. As outras nove continham correspondentes lexicais sem um alimento em sua composição. Dentre as 10 EI's restantes, quatro foram traduzidas em sentido não figurado, mas sendo possível compreender seu significado; outras quatro EI's, apesar de terem sido traduzidas, seus significados não estão compatíveis com aquele identificado por nós para essas EI's; uma EI foi parcialmente traduzida (**não deve ter nada a ver Ø**) e uma não foi traduzida, sendo mantida na língua de origem (**Butterfingers**). Se somarmos as traduções metafóricas com e sem alimentos (21) e as traduzidas não metaforicamente (4), teremos um total de 25 EI's compreensíveis para esse THP, totalizando um percentual de 80,6%.

O THP2 foi o que obteve o desempenho mais alto, pois reconheceu a metáfora subjacente em 26 EI's, traduzindo 10 delas com traduções semelhantes às da tradução de referência, ou com outro alimento, e as outras 16 com correspondentes lexicais sem utilizar um alimento em sua composição. Dentre as cinco EI's restantes, três foram traduzidas com significado não metafórico, mas sendo possível compreender seu significado; e somente uma não teve seu significado compatível com aquele identificado por nós para a EI em questão. Esse THP não deixou de traduzir nenhuma EI. Se somarmos todas as traduções possíveis de serem compreendidas pelo usuário, teremos um total de 30 EI's, o que significa um percentual de 96,7%.

O THP3 reconheceu a metáfora subjacente a 27 EI's, traduzindo oito delas com nossas traduções de referência ou outra com o mesmo alimento. As outras 19 continham correspondentes lexicais sem um alimento em sua composição. Dentre as quatro EI's restantes, duas não tiveram seu significado compatível com aquele identificado por nós para essas EI's; e duas tiveram parte da EI não traduzida. Contudo, é possível compreender o significado de uma delas (**assimilado literatura de todas as definições Ø**). Somando todas as traduções das EI's com compreensão possível, temos 28, totalizando 90,3%.

Os percentuais obtidos por cada THP estão sumarizados na Tabela 5.

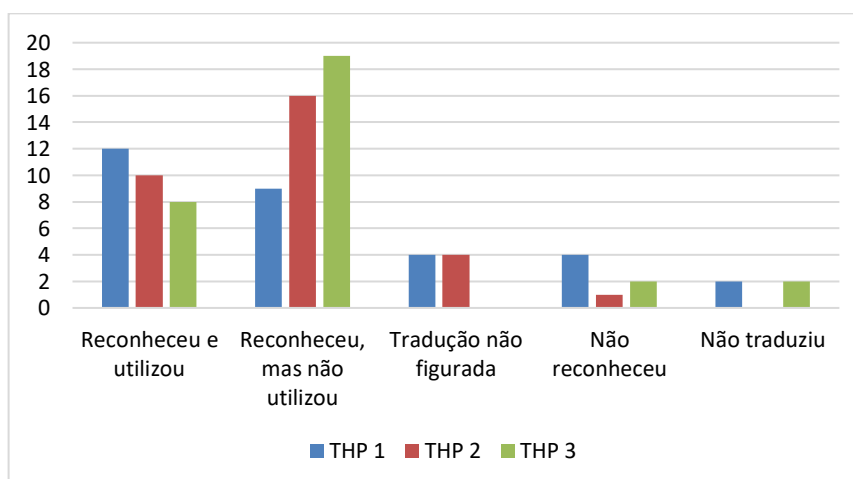
Tabela 5 – Resumo dos percentuais de desempenho dos THP's na elaboração das traduções

Tradutor	Percentual
THP1	80,6%
THP2	96,7%
THP3	90,3%

Fonte: Elaborada pela autora.

Outra forma de verificar o desempenho dos THP's é mostrada no Gráfico 3, no qual visualizamos claramente os melhores desempenhos obtidos pelo THP2 e THP3, com a maior parte dos resultados concentrados nos primeiros dois tipos de tradução elaborada (reconheceu a EI e utilizou um alimento; e reconheceu a EI, mas não utilizou um alimento).

Gráfico 3 – Desempenho dos THP's na elaboração das traduções na única coleta



Fonte: Elaborado pela autora.

Sumarizando o percentual de desempenho de todos os tradutores (TAO's e THP's), podemos verificar que o desempenho dos THP's está muito acima dos TAO's, como podemos melhor visualizar na Tabela 6.

Tabela 6 – Resumo dos percentuais de desempenho dos dois tipos de tradutores na elaboração das traduções

Tradutor:	1ª coleta	2ª coleta	2ª coleta – 2ª opção
TAO1	54,8%	41,9%	-
TAO2	35,4%	38,7%	61,2%
TAO3	29,0%	29,0%	-
THP1	80,6%	-	-
THP2	96,7%	-	-
THP3	90,3%	-	-

Fonte: Elaborada pela autora.

De acordo com os percentuais obtidos a partir de nossa análise semântica, evidencia-se que os THP's obtiveram, em geral, um resultado acima dos TAO's. O THP2 obteve o resultado mais alto de todos (96,7%), enquanto o TAO3 obteve o mais baixo (29%), em ambas as coletas realizadas. Nossa hipótese que os THP's reconheceriam as EI's e fariam suas traduções utilizando linguagem metafórica, mas não necessariamente utilizando uma EI com alimento, foi, portanto, corroborada.

Após utilizarmos nossa métrica semântica para analisar todas as sentenças candidatas, verificamos como as métricas automáticas as analisaram. Os resultados e discussão são apresentados no subitem seguinte.

4.5. ANÁLISE DAS MÉTRICAS DE AVALIAÇÃO DE TRADUÇÃO AUTOMÁTICA

O quinto e último objetivo secundário a ser alcançado diz respeito às análises das métricas de avaliação de Tradução Automática. Para esta pesquisa, escolhemos trabalhar com três métricas diferentes: BLEU, *F-measure* e METEOR. Apresentamos, nos próximos subitens, os resultados obtidos para cada uma delas.

4.5.1 Métrica BLEU

BLEU é o acrônimo para *Bilingual Evaluation Understudy* e é uma métrica que foi desenvolvida com a finalidade de avaliar sistemas de Tradução Automática

com maior economia, rapidez e independência de linguagens do que as avaliações realizadas manualmente. Essa métrica é fundamentada na proximidade entre a tradução realizada por computadores e a tradução realizada por uma pessoa qualificada em traduções profissionais, ou seja, a tradução de referência. Todos os acertos recebem pesos iguais. Isso significa que, tanto as palavras de ligação quanto as de conteúdo valem o mesmo.

A escolha e a ordem correta das palavras na sentença candidata, em comparação com a sentença de referência, são recompensadas tendo como base a combinação de palavras (n-gramas), partindo de um até o máximo de quatro, em geral. Há uma penalidade quando o número de palavras da sentença candidata é menor do que o da de referência (*brevity penalty*). Para essa métrica, o melhor resultado é 1, e o pior, 0.

Apresentamos na Tabela 7 os resultados obtidos pela métrica BLEU em nossa primeira coleta e os valores de *precision* para os diferentes n-gramas:

Tabela 7 – Resultados da métrica BLEU para os TAO's e THP's⁵¹

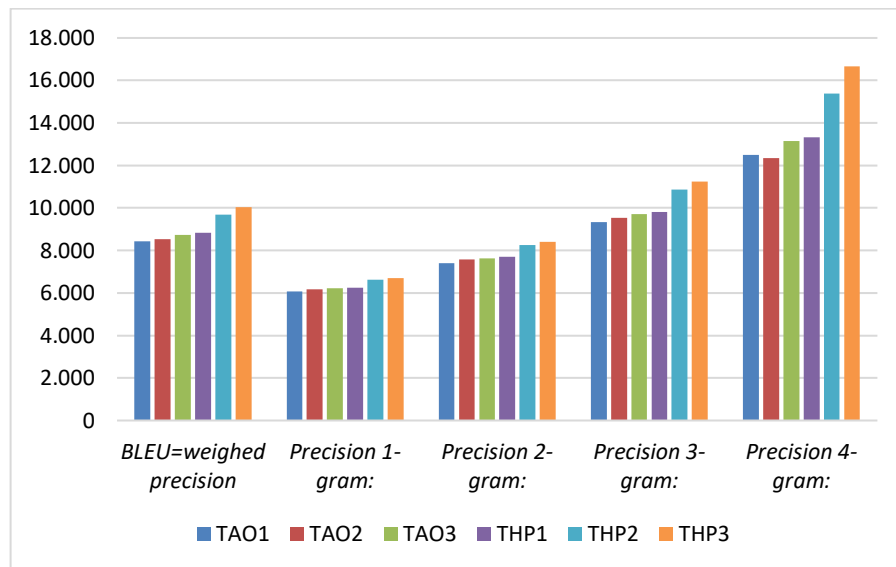
	TAO1	TAO2	TAO3	THP1	THP2	THP3
BLEU = weighed precision	0.008436	0.008535	0.008742	0.008822	0.009681	0.010030
<i>Precision 1-gram:</i>	0.006061	0.006173	0.006211	0.006250	0.006623	0.006711
<i>Precision 2-gram:</i>	0.007407	0.007576	0.007634	0.007692	0.008264	0.008403
<i>Precision 3-gram:</i>	0.009346	0.009524	0.009709	0.009804	0.010870	0.011236
<i>Precision 4-gram:</i>	0.012500	0.012346	0.013158	0.013333	0.015385	0.016667

Fonte: Elaborada pela autora.

Os resultados trazem frações decimais muito extensas que dificultam perceber a diferença entre os valores obtidos. Multiplicamos todos os resultados por 1.000 (hum mil), de forma que pudéssemos representar graficamente e visualizar de forma mais didática, apresentados no Gráfico 4.

⁵¹ Para esta e demais tabelas desta seção, marcamos os maiores resultados em azul e os menores em vermelho.

Gráfico 4 – Resultados da métrica BLEU para os dois tipos de tradutores



Fonte: Elaborado pela autora.

Para esta métrica, se levarmos em consideração ambos os tipos de tradutores, podemos verificar que o THP3 obteve os maiores resultados dentre todos. Já o TAO1 obteve todos os menores resultados, exceto para o quadrigrama, que foi obtido pelo TAO2. Contudo, se analisarmos separadamente os TAO's (Tabela 8) e os THP (Tabela 9), podemos verificar que o TAO3 obteve os maiores resultados e o THP1 obteve os resultados mais baixos. A penalidade por brevidade foi a mesma para todos: 1.000000.

Tabela 8 – Resultados da métrica BLEU para os TAO's

	TAO1	TAO2	TAO3
BLEU = weighed precision	0.008436	0.008535	0.008742
<i>Precision 1-gram:</i>	0.006061	0.006173	0.006211
<i>Precision 2-gram:</i>	0.007407	0.007576	0.007634
<i>Precision 3-gram:</i>	0.009346	0.009524	0.009709
<i>Precision 4-gram:</i>	0.012500	0.012346	0.013158

Fonte: Elaborada pela autora.

O fato de o TAO3 obter os maiores resultados e o TAO1 quase todos os piores na métrica BLEU deve ser salientado, pois vai totalmente de encontro aos resultados de nossa avaliação semântica. Contudo, essa métrica, apesar de ser

amplamente utilizada, já foi questionada por aferir resultados baixos para tradutores humanos, devido à variedade lexical utilizada por estes.

Tabela 9 – Resultados da métrica BLEU para os THP's

	THP1	THP2	THP3
BLEU = weighed precision	0.008822	0.009681	0.010030
<i>Precision 1-gram:</i>	0.006250	0.006623	0.006711
<i>Precision 2-gram:</i>	0.007692	0.008264	0.008403
<i>Precision 3-gram:</i>	0.009804	0.010870	0.011236
<i>Precision 4-gram:</i>	0.013333	0.015385	0.016667

Fonte: Elaborada pela autora.

Já para os THP, os resultados são mais próximos daqueles obtidos em nossa análise semântica, na qual o THP2 obteve o resultado mais alto, seguido pelo THP3 e, por último, o THP1. Apesar de haver a inversão entre os dois THP com resultados maiores, a diferença entre ambos foi de apenas 6,4%. A diferença de resultados entre esses THP's para a métrica BLEU (*weighed precision*) também é bastante baixa, somente 0,000349. Vejamos a seguir, os resultados da métrica *F-measure*.

4.5.2 Métrica *F-measure*

F-measure, ou *F1 score*, ou ainda *F score*, é uma métrica desenvolvida pela Universidade de Nova Iorque que associa conjuntos de coocorrências entre traduções candidatas e de referência. Ela representa a média harmônica de *precision* e *recall*, em que o melhor resultado alcançado é 1 e o pior 0.

Apresentamos, na Tabela 10, os valores de *precision* e *recall* e os resultados obtidos para a métrica *F-measure*, para cada um dos tradutores, na primeira coleta de dados realizada para os TAO e na única coleta para os THP.

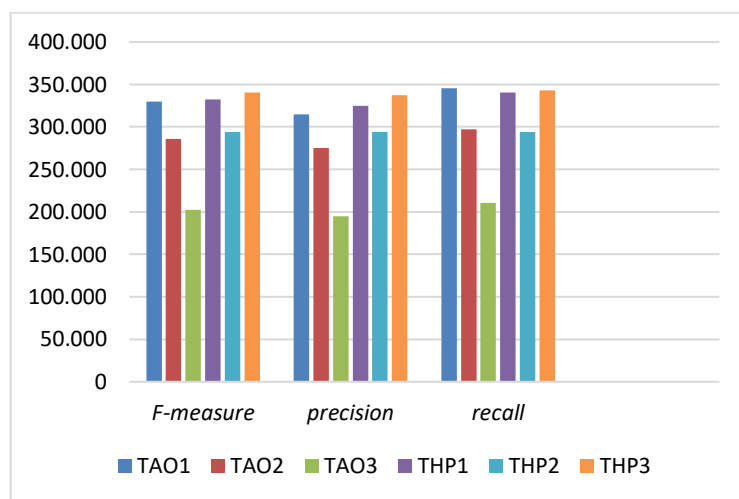
Tabela 10 – Resultados da métrica *F-measure* para os TAO's e THP's

	TAO1	TAO2	TAO3	THP1	THP2	THP3
<i>F-measure</i>	0.329610	0.285833	0.202254	0.332443	0.293985	0.340165
<i>precision</i>	0.315042	0.275395	0.194869	0.324756	0.293985	0.337127
<i>recall</i>	0.345591	0.297093	0.210222	0.340502	0.293985	0.343257

Fonte: Elaborada pela autora.

Assim como para a métrica BLEU, os resultados trazem frações decimais muito extensas que dificultam perceber a diferença entre os valores obtidos. Adotamos o mesmo procedimento e multiplicamos todos os resultados por 1.000 (hum mil), de forma que pudéssemos representar graficamente e visualizar de forma mais didática, os quais apresentamos no Gráfico 5.

Gráfico 5 – Resultados da métrica *F-measure* para os dois tipos de tradutores



Fonte: Elaborado pela autora.

Analisando de maneira geral os TAO's e os THP's, verifica-se que o THP3 obteve o maior resultado de *precision* e *F-measure*. Já o maior resultado de *recall* foi obtido pelo TAO1. Os menores resultados foram todos obtidos pelo TAO3. Apresentamos os resultados obtidos separadamente pelos TAO's na Tabela 11 e pelos THP's na Tabela 12:

Tabela 11 – Resultados da métrica *F-measure* para os TAO's

	TAO1	TAO2	TAO3
<i>F-measure</i>	0.329610	0.285833	0.202254
<i>precision</i>	0.315042	0.275395	0.194869
<i>recall</i>	0.345591	0.297093	0.210222

Fonte: Elaborada pela autora.

Diferentemente dos resultados obtidos pela métrica BLEU para os TAO's, esses resultados são mais próximos daqueles obtidos em nossa análise semântica das traduções elaboradas pelos TAO's, no qual o TAO1 obteve os maiores resultados, gerando traduções mais próximas de nossas traduções de referência, e o TAO3

obteve os menores resultados. Vejamos a seguir os resultados isolados alcançados pelos THP's.

Tabela 12 – Resultados da métrica *F-measure* para os THP's

	THP1	THP2	THP3
<i>F-measure</i>	0.332443	0.293985	0.340165
<i>precision</i>	0.324756	0.293985	0.337127
<i>recall</i>	0.340502	0.293985	0.343257

Fonte: Elaborada pela autora.

Assim como para a métrica BLEU, o THP3 obteve a maior pontuação nessa métrica. Já o THP2 obteve a pontuação mais baixa, sendo este um resultado totalmente oposto ao que obtivemos em nossa análise semântica, conforme mencionado anteriormente. A diferença no resultado da métrica *F-measure* entre esses dois THP's foi de 0,04618.

Mostramos os resultados da métrica METEOR na próxima subseção.

4.5.3 Métrica METEOR

METEOR é o acrônimo para *Metric for Evaluation of Translation with Explicit ORdering*. Essa métrica foi desenvolvida pela Universidade Carnegie Mellon e foi escolhida porque demonstra ter altos níveis de correlação com os julgamentos humanos de qualidade de tradução (LAVIE; AGARWAL, 2007). Nessa métrica, os resultados podem ser vistos em nível de sentenças ou segmentos (*chunks*) e há uma penalidade pela quebra desses segmentos (*fragmentation penalty*). A avaliação é realizada computando-se uma nota que se baseia em associações corretas explícitas, palavra-por-palavra, entre a tradução candidata e a de referência. Assim como as métricas anteriores, o melhor resultado é 1 e o pior, 0.

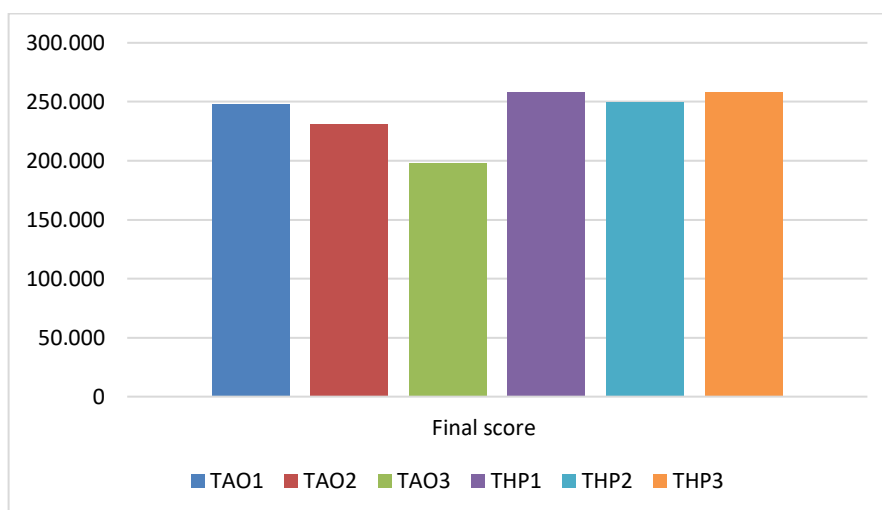
Apresentamos a pontuação obtida utilizando a métrica METEOR. Para cada um dos tradutores (TAO's e THP's), essa métrica disponibiliza o número de palavras testadas das traduções candidatas e de referência (neste caso, 198 para todos), o número de segmentos (*chunks*), *precision*, *recall*, *f1*, *fmean*, *fragmentation penalty* e a pontuação final (*final score*). Na Tabela 13, apresentamos apenas o resultado final (*final score*). Os outros valores são apresentados posteriormente para cada tipo de tradutor separadamente:

Tabela 13 – Resultados da métrica METEOR para os TAO's e THP's

METEOR	TAO1	TAO2	TAO3	THP1	THP2	THP3
<i>Final score</i>	0.24815606 91980535	0.23124835 93822616	0.19840393 760634323	0.25808706 4926912	0.24988915 56437419	0.25781939 80931749

Fonte: Elaborada pela autora.

O mesmo procedimento que adotamos para as métricas BLEU e *F-measure* foi adotado para a métrica METEOR. Multiplicamos todos os resultados por 1.000 (hum mil). Além disso, restringimos o número de casas decimais para o máximo de três, de forma que pudéssemos representar graficamente e visualizar de forma mais didática esses valores, os quais apresentamos no Gráfico 6.

Gráfico 6 – Resultados da métrica METEOR para os dois tipos de tradutores

Fonte: Elaborado pela autora.

Em uma visão geral, vemos que os resultados mais baixos foram novamente obtidos pelos TAO's, enquanto que os maiores, obviamente, foram obtidos pelos THP's, resultado semelhante ao das métricas BLEU e *F-measure*, os quais podemos verificar com mais detalhes quando apresentamos de acordo como o tipo de tradutor (Tabelas 14 e 15).

Tabela 14 – Resultados da métrica METEOR para os TAO's

METEOR	TAO1	TAO2	TAO3
<i>Test words</i>	212	208	208
<i>Reference words</i>	198	198	198
<i>Chunks</i>	52	46	53
<i>Precision</i>	0.49	0.4558441558441559	0.41633986928104577
<i>Recall</i>	0.5237931034482759	0.48413793103448277	0.43931034482758624
<i>f1</i>	0.5063333333333333	0.4695652173913044	0.42751677852349
<i>fMean</i>	0.5184300341296929	0.4796720191322173	0.43570451436388513
<i>Fragmentation penalty</i>	0.5213316110926289	0.5179031710029347	0.5446364885706848
<i>Final score</i>	0.2481560691980535	0.2312483593822616	0.19840393760634323

Fonte: Elaborada pela autora.

A Tabela 14 evidencia, claramente, que o TAO1 obteve todas as maiores pontuações, exceto para a penalidade por fragmentação, cujo melhor resultado foi obtido pelo TAO2. O TAO3 obteve o menor resultado em todos os itens, o que é consistente com nossa análise semântica. Vejamos os resultados da métrica METEOR para os THP's isoladamente na Tabela 15.

Tabela 15 – Resultados da métrica METEOR para os THP's

METEOR	THP1	THP2	THP3
<i>Test words</i>	200	193	194
<i>Reference words</i>	198	198	198
<i>Chunks</i>	46	45	41
<i>Precision</i>	0.5222602739726028	0.518804920913884	0.5265017667844523
<i>Recall</i>	0.5258620689655172	0.5089655172413793	0.5137931034482759
<i>f1</i>	0.5240549828178694	0.5138381201044386	0.5200698080279232
<i>fMean</i>	0.5253186358939029	0.5104175672170831	0.5156601488146738
<i>Fragmentation penalty</i>	0.5087037708309341	0.5104221098693831	0.5000207041676317
<i>Final score</i>	0.258087064926912	0.2498891556437419	0.2578193980931749

Fonte: Elaborada pela autora.

Da mesma forma que na métrica *F-measure*, o THP2 obteve a menor pontuação dentre os THP's para a métrica METEOR e, surpreendentemente, o THP1 obteve o maior resultado final. Isso representa um resultado bastante diferente daquele obtido em nossa análise semântica, na qual o THP2 obteve os maiores resultados, seguido pelo THP3 e o THP1. É, contudo, relevante perceber que a diferença entre o maior e menor resultados obtidos nesta métrica é mínima, apenas 0,0081979092.

Com essas informações finalizamos as análises dos dados obtidos na primeira coleta pelos TAO's e THP's. Para tal, utilizamos as três métricas selecionadas para este estudo: BLEU, *F-measure* e METEOR. No subitem seguinte, apresentamos

as análises para a segunda coleta de traduções candidatas que foi realizada pelos TAO's somente, uma vez que não havia necessidade de solicitar uma segunda coleta aos THP's.

4.5.4. Análise da evolução dos TAO's

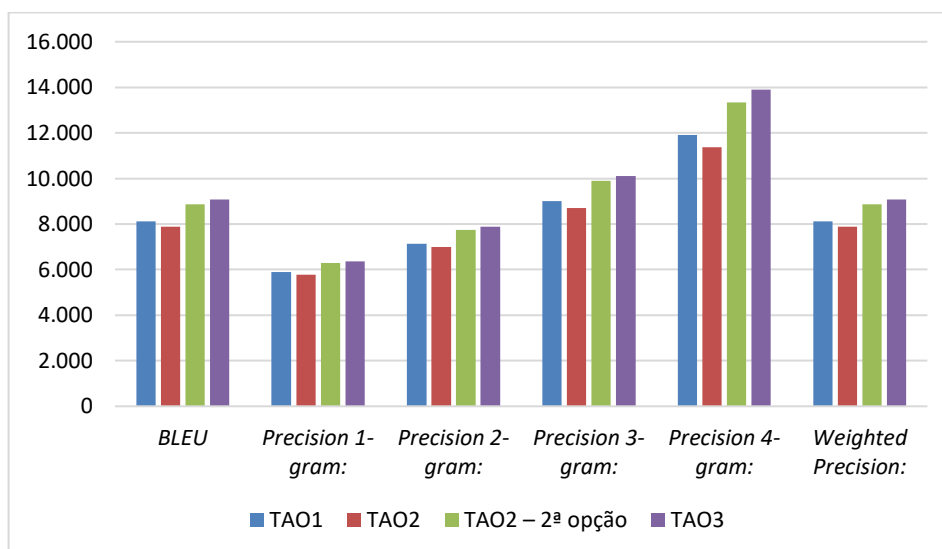
Apresentamos a seguir, os resultados obtidos a partir das traduções candidatas elaboradas pelos TAO's na segunda coleta (dezembro de 2017), após serem submetidos às mesmas três métricas. O objetivo dessa segunda análise foi verificar se as modificações ocorridas nos TAO's produziram melhores resultados do que os anteriores. Lembramos que foram registradas duas opções de tradução para o TAO2 nessa coleta. A tradução de referência utilizada para essa análise foi a mesma da primeira. Iniciamos a apresentação com os resultados da métrica BLEU, na Tabela 16.

Tabela 16 – Resultados da métrica BLEU para os TAO's na segunda coleta

	TAO1	TAO2	TAO2 – 2ª opção	TAO3
BLEU = <i>weighted precision</i>	0.008124	0.007885	0.008874	0.009072
<i>Precision 1-gram:</i>	0.005882	0.005780	0.006289	0.006369
<i>Precision 2-gram:</i>	0.007143	0.006993	0.007752	0.007874
<i>Precision 3-gram:</i>	0.009009	0.008696	0.009901	0.010101
<i>Precision 4-gram:</i>	0.011905	0.011364	0.013333	0.013889

Fonte: Elaborada pela autora.

Da mesma forma que procedemos com os dados da primeira coleta, procedemos com os da segunda – multiplicamos todos os resultados por 1.000 (hum mil). Isso nos permitiu elaborar o Gráfico 7, que nos ajuda a visualizar melhor os resultados obtidos por cada TAO.

Gráfico 7 – Resultados da métrica BLEU para os TAO's na segunda coleta

Fonte: Elaborado pela autora.

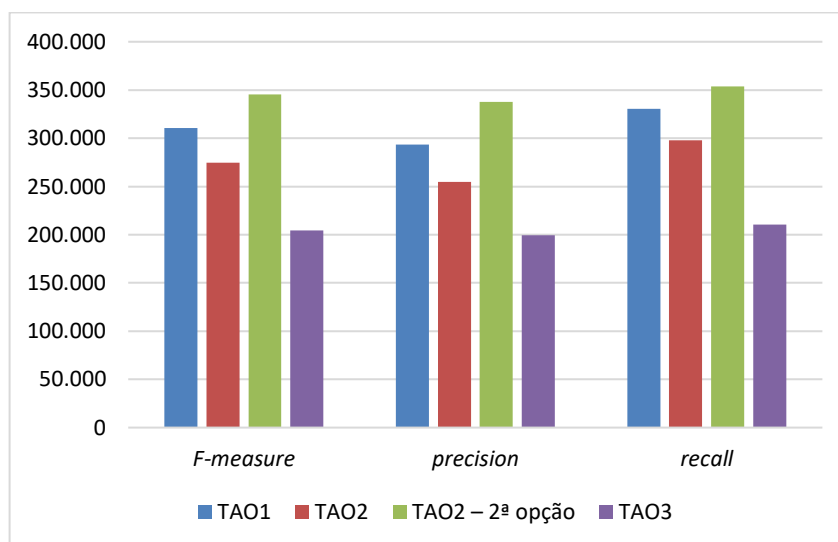
Novamente, diferente dos resultados encontrados por nossa métrica semântica, o TAO3 obteve os maiores resultados para a métrica BLEU. Já os menores resultados foram obtidos pelo TAO2 em sua primeira opção de tradução, o que corrobora nossa avaliação semântica. Vejamos na Tabela 17 os resultados para a métrica *F-measure* para essa segunda coleta de dados dos TAO's.

Tabela 17 – Resultados da métrica *F-measure* para os TAO's na segunda coleta

	TAO1	TAO2	TAO2 – 2ª opção	TAO3
<i>F-measure</i>	0.310741	0.274514	0.345416	0.204641
<i>precision</i>	0.293200	0.254601	0.337430	0.199348
<i>recall</i>	0.330516	0.297806	0.353790	0.210222

Fonte: Elaborada pela autora.

Mais uma vez, multiplicamos por 1.000 (hum mil) os resultados obtidos nessa métrica, para facilitar a elaboração do Gráfico 8.

Gráfico 8 – Resultados da métrica *F-measure* para os TAO's na segunda coleta

Fonte: Elaborado pela autora.

Os resultados obtidos na métrica *F-measure*, tanto os maiores quanto os menores, corroboram os resultados de nossa métrica semântica para todos os TAO's avaliados: os maiores resultados foram obtidos pelo TAO2 em sua segunda opção, e os mais baixos foram obtidos pelo TAO3. A Tabela 18 nos mostra os resultados aferidos pela métrica METEOR para os TAO's em sua segunda coleta de dados.

Tabela 18 – Resultados da métrica METEOR para os TAO's na segunda coleta

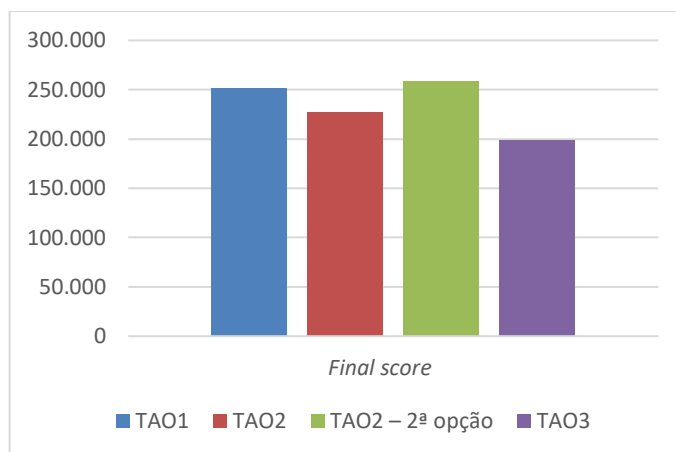
METEOR	TAO1	TAO2	TAO2 – 2ª opção	TAO3
<i>Test words</i>	219	214	206	208
<i>Reference words</i>	198	198	198	198
<i>Chunks</i>	49	48	49	53
<i>Precision</i>	0.4756589147286 822	0.44240506329113 927	0.5132450331125 827	0.41633986928104 577
<i>Recall</i>	0.5289655172413 793	0.48206896551724 143	0.5344827586206 896	0.43931034482758 624
<i>f1</i>	0.5008979591836 735	0.46138613861386 14	0.5236486486486 486	0.42751677852349
<i>fMean</i>	0.5202204323866 045	0.47567199727798 576	0.5311857436600 41	0.43570451436388 513
<i>Fragmentation penalty</i>	0.5161593496765 684	0.52233033797767 45	0.5141967097393 753	0.54463648857068 48
<i>Final score</i>	0.2517037923174 7153	0.22721408217326	0.2580517820095 8473	0.19840393760634 323

Fonte: Elaborada pela autora.

Para essa métrica, apresentamos todos os dados obtidos na Tabela 18, mas somente o resultado final (*final score*) para o Gráfico 9. Novamente, multiplicamos

os valores por 1.000 (hum mil), de forma que possamos apresentá-los mais didaticamente.

Gráfico 9 – Resultados da métrica METEOR para os TAO's na segunda coleta



Fonte: Elaborado pela autora.

Também para essa métrica, os resultados obtidos, tanto os maiores quanto os menores, corroboram nossa avaliação semântica para todos os TAO's avaliados, ou seja, os maiores resultados foram obtidos pelo TAO2 em sua segunda opção, e TAO3 obteve os resultados mais baixos.

Finalizada a análise da segunda coleta dos TAO's entre si, fizemos as análises dos resultados obtidos individualmente por cada TAO para as duas coletas, a fim de explicitar se houve aprimoramento do sistema entre uma coleta e outra. Para tal, utilizamos os resultados obtidos em cada métrica. Apresentamos tabelas individuais para cada um dos três⁵² TAO's:

Tabela 19 – Comparação dos resultados da métrica BLEU para o TAO1

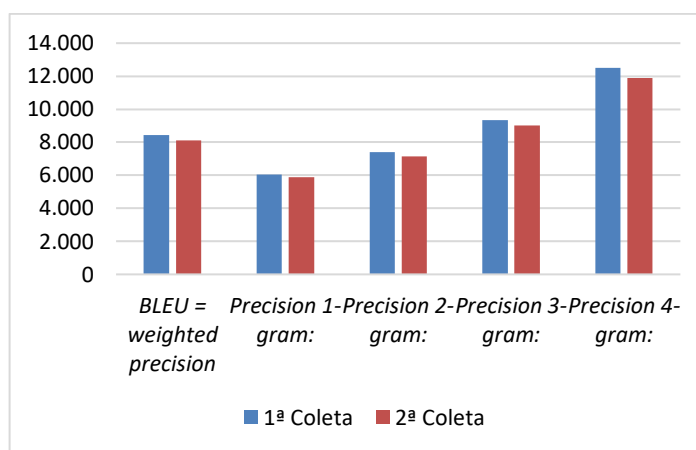
	1ª Coleta	2ª Coleta
BLEU = <i>weighted precision</i>	0.008436	0.008124
<i>Precision 1-gram:</i>	0.006061	0.005882
<i>Precision 2-gram:</i>	0.007407	0.007143
<i>Precision 3-gram:</i>	0.009346	0.009009
<i>Precision 4-gram:</i>	0.012500	0.011905

Fonte: Elaborada pela autora.

⁵² Para todos os gráficos apresentados a partir deste ponto seguimos o mesmo procedimento adotado para os gráficos anteriores, ou seja, multiplicamos os valores obtidos por cada TAO analisado por 1.000 (hum mil), de forma que possamos apresentá-los mais didaticamente.

Esses resultados corroboram nossa avaliação qualitativa, na qual observamos que o número de traduções compreensíveis elaboradas na primeira coleta pelo TAO1 foi superior ao da segunda coleta. Algumas das traduções que havíamos considerado compreensíveis na primeira coleta, sofreram alterações na segunda coleta que deixaram a compreensão das EI's prejudicada. Essa diferença é melhor visualizada no Gráfico 10.

Gráfico 10 – Desempenho do TAO1 nas duas coletas para a métrica BLEU



Fonte: Elaborado pela autora.

Também neste gráfico fica mais visível o desempenho do TAO1 em relação ao número de n-gramas. Percebemos que os valores mais altos foram para os tetragramas, seguido pelos trigramas.

Na Tabela 20, podemos verificar os resultados para o TAO2 isoladamente. Infelizmente, a segunda opção de tradução não foi coletada em março de 2017, quando realizamos a primeira coleta. Acreditamos que a segunda opção de tradução somente foi disponibilizada para esse TAO com a mudança para o sistema neural, mas não podemos afirmar definitivamente se não estava disponível à época da primeira coleta ou se foi um erro no procedimento metodológico de nossa pesquisa.

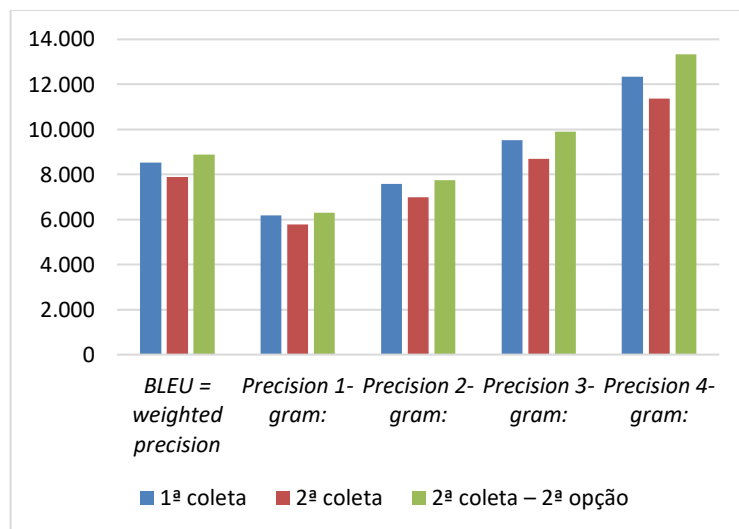
Tabela 20 – Comparação dos resultados da métrica BLEU para o TAO2

	1ª coleta	2ª coleta	2ª coleta – 2ª opção
BLEU = weighted precision	0.008535	0.007885	0.008874
<i>Precision 1-gram:</i>	0.006173	0.005780	0.006289
<i>Precision 2-gram:</i>	0.007576	0.006993	0.007752
<i>Precision 3-gram:</i>	0.009524	0.008696	0.009901
<i>Precision 4-gram:</i>	0.012346	0.011364	0.013333

Fonte: Elaborada pela autora.

Também para esse TAO, elaboramos um gráfico (Gráfico 11), que nos permite visualizar melhor o desempenho alcançado entre as coletas.

Gráfico 11 – Desempenho do TAO2 nas duas coletas para a métrica BLEU



Fonte: Elaborado pela autora.

Para esse TAO, temos três valores: um da primeira coleta e dois da segunda. Isso nos permite perceber que os valores de tetragrama e trigrama também foram mais altos para o TAO2, bem como a regressão na percepção das EI's da primeira coleta para a primeira opção da segunda coleta e, posteriormente, a progressão se compararmos essas com a segunda opção da segunda coleta, resultados que são corroborados por nossa avaliação semântica.

Apresentamos na Tabela 21 os valores obtidos para o TAO3 na avaliação da métrica BLEU:

Tabela 21 – Comparação dos resultados da métrica BLEU para o TAO3

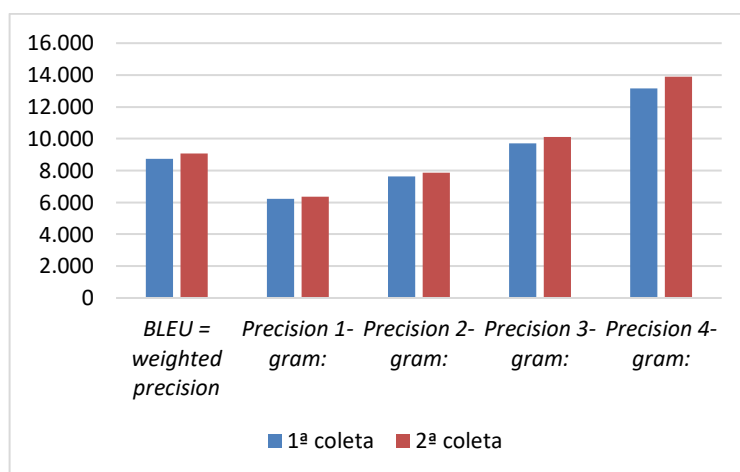
	1ª coleta	2ª coleta
BLEU = weighted precision	0.008742	0.009072
<i>Precision 1-gram:</i>	0.006211	0.006369
<i>Precision 2-gram:</i>	0.007634	0.007874
<i>Precision 3-gram:</i>	0.009709	0.010101
<i>Precision 4-gram:</i>	0.013158	0.013889

Fonte: Elaborada pela autora.

É interessante perceber que, apesar de não ter havido qualquer mudança nas traduções elaboradas por esse TAO, os valores obtidos na segunda coleta foram maiores do que os da primeira. Acreditamos que esses resultados tenham sido em decorrência de uma pequena modificação, em algumas sentenças candidatas, quando omitimos o verbo anterior à EI analisada (pensa que todos os seus gansos são...). Portanto, na verdade, não houve aprimoramento do TAO, mas uma modificação nos dados. Vale salientar que essa modificação não invalida os resultados obtidos para a segunda coleta, uma vez que foi realizada em todas as métricas.

Da mesma forma que elaboramos um gráfico para os TAO1 e TAO2, elaboramos o Gráfico 12 para esse TAO.

Gráfico 12 – Desempenho do TAO3 nas duas coletas para a métrica BLEU



Fonte: Elaborado pela autora.

Pode-se perceber que os maiores valores obtidos pelo TAO3 também foram de tetragramas e trigramas, havendo um pequeno aumento no desempenho da primeira para a segunda coleta, conforme justificamos acima. Isso mostra quão sensíveis são as métricas a pequenas mudanças nos dados.

Finalizadas as comparações no desempenho de cada um dos TAO's, analisados pela métrica BLEU, nas tabelas seguintes, apresentaremos as comparações individuais para a métrica *F-measure*.

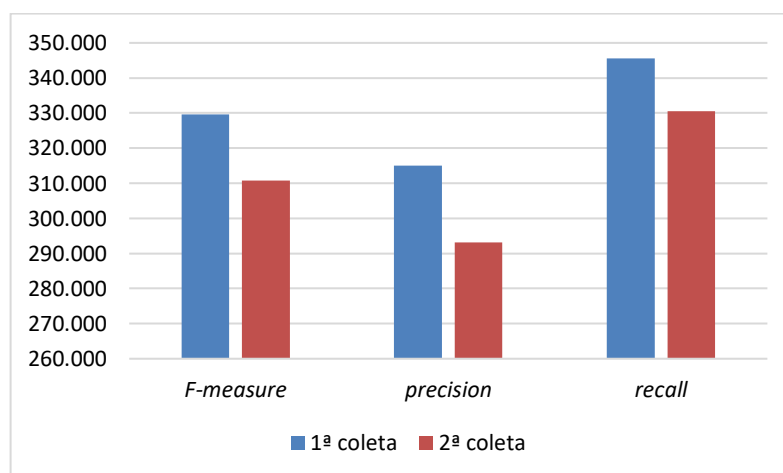
Tabela 22 – Comparação dos resultados da métrica *F-measure* para o TAO1

	1ª coleta	2ª coleta
<i>F-measure</i>	0.329610	0.310741
<i>precision</i>	0.315042	0.293200
<i>recall</i>	0.345591	0.330516

Fonte: Elaborada pela autora.

Similarmente ao que vimos nos resultados da métrica BLEU, houve redução nos resultados obtidos pelo TAO1 na segunda coleta, o que contraria a lógica, pois espera-se que esses sistemas sejam aprimorados com o tempo. Não o oposto. Esses resultados são melhor visualizados no Gráfico 13.

Gráfico 13 – Desempenho do TAO1 nas duas coletas para a métrica *F-measure*



Fonte: Elaborado pela autora.

Percebe-se claramente que os resultados obtidos pelo TAO1 na avaliação da métrica *F-measure* corroboram aqueles elaborados pela métrica BLEU e também por nossa avaliação semântica, ou seja, a primeira coleta trouxe resultados mais altos que a segunda. Para esse TAO, é ainda possível perceber que os resultados de *recall* são maiores que os de *precision* em ambas as coletas.

Na Tabela 23, vemos os resultados obtidos pelo TAO2 na avaliação da métrica *F-measure* em ambas as coletas, incluindo as duas opções de tradução da segunda coleta.

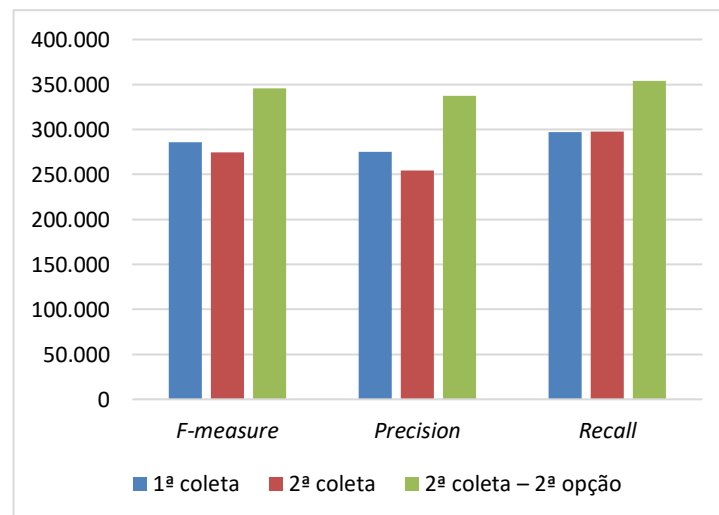
Tabela 23 – Comparação dos resultados da métrica *F-measure* para o TAO2

	1ª coleta	2ª coleta	2ª coleta – 2ª opção
<i>F-measure</i>	0.285833	0.274514	0.345416
<i>Precision</i>	0.275395	0.254601	0.337430
<i>Recall</i>	0.297093	0.297806	0.353790

Fonte: Elaborada pela autora.

Da mesma forma que se verificou com a métrica BLEU, a segunda opção da coleta mais recente obteve os maiores resultados também na avaliação dessa métrica. A primeira coleta teve a avaliação mais baixa no *recall*, enquanto que a primeira opção da segunda coleta teve resultados mais baixos de *precision* e *F-measure*. Resultado bastante semelhante ao da métrica BLEU e ao de nossa avaliação qualitativa, conforme mencionamos anteriormente. Podemos visualizar melhor por meio do Gráfico 14.

Gráfico 14 – Desempenho do TAO2 nas duas coletas para a métrica *F-measure*



Fonte: Elaborado pela autora.

Percebe-se facilmente que os resultados para a segunda opção da segunda coleta são os mais altos, conforme já mencionamos e corroboram as avaliações da métrica BLEU e nossa avaliação semântica. Os resultados de *precision* e *recall* também são os maiores para a segunda opção da segunda coleta. É possível perceber mais uma vez que a primeira coleta obteve resultados um pouco maiores do que os da primeira opção da segunda coleta, mas resultados de *recall* iguais.

Na Tabela 24, temos os resultados obtidos pelo TAO3 na avaliação da métrica *F-measure*:

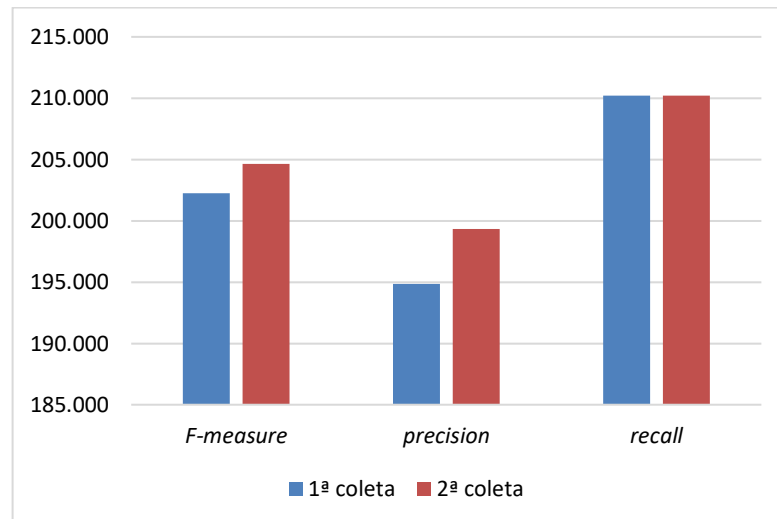
Tabela 24 – Comparação dos resultados da métrica *F-measure* para o TAO3

	1ª coleta	2ª coleta
<i>F-measure</i>	0.202254	0.204641
<i>precision</i>	0.194869	0.199348
<i>recall</i>	0.210222	0.210222

Fonte: Elaborada pela autora.

Os resultados obtidos pelo TAO3 mostram semelhanças aos obtidos na avaliação da métrica BLEU, observando que a pontuação de *precision* para a segunda coleta foi um pouco maior do que a da primeira. Porém, os resultados para *recall* foram iguais em ambas as coletas.

Gráfico 15 – Desempenho do TAO3 nas duas coletas para a métrica *F-measure*



Fonte: Elaborado pela autora.

Finalizadas as comparações de desempenho isolados para cada um dos TAO's analisados pela métrica *F-measure*, os próximos resultados apresentados são a comparação da avaliação da métrica METEOR.

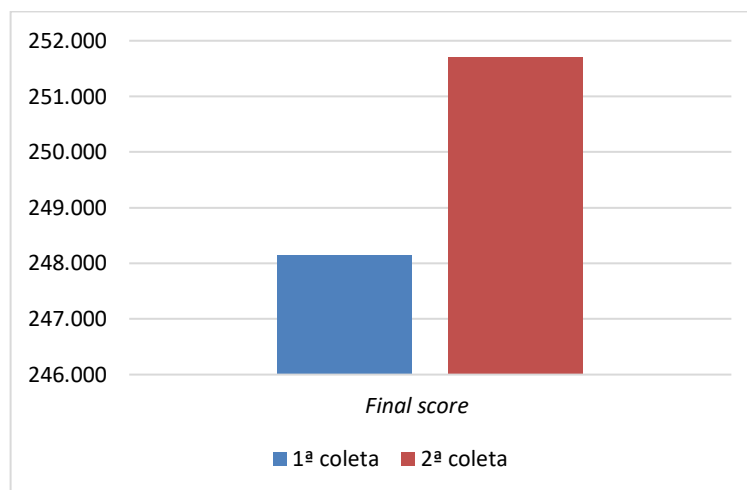
Tabela 25 – Comparação dos resultados da métrica METEOR para o TAO1

METEOR	1ª coleta	2ª coleta
<i>Final score</i>	0.2481560691980535	0.25170379231747153

Fonte: Elaborada pela autora.

Apesar de ter obtido os resultados mais baixos quando comparados aos outros TAO's, essa métrica evidencia que houve uma melhora no resultado final da primeira para a segunda coleta do TAO1.

Gráfico 16 – Desempenho do TAO1 nas duas coletas para a métrica METEOR



Fonte: Elaborado pela autora.

O Gráfico 16 evidencia uma melhor avaliação das traduções candidatas pela métrica METEOR na segunda coleta. Esse resultado é contrário ao que verificamos em nossa análise semântica, ou seja, a redução de 54,8% para 41,9% no número de traduções compreensíveis.

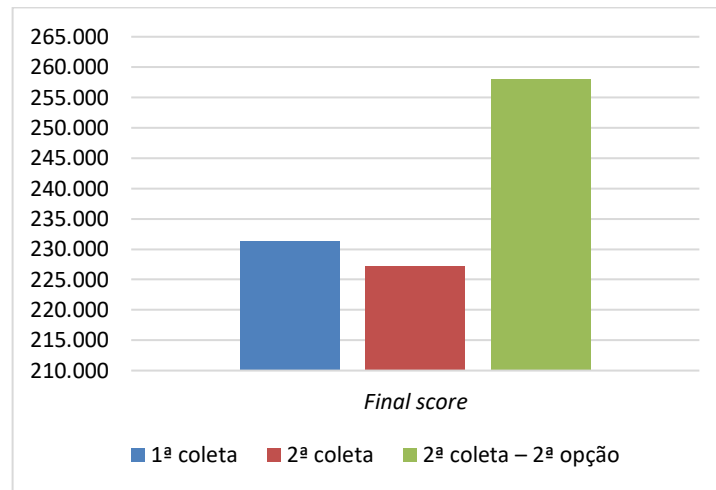
Os resultados obtidos pelo TAO2 nas duas coletas são mostrados na Tabela 26.

Tabela 26 – Comparação dos resultados da métrica METEOR para o TAO2

METEOR	1ª coleta	2ª coleta	2ª coleta – 2ª opção
<i>Final score</i>	0.2312483593822616	0.22721408217326	0.25805178200958473

Fonte: Elaborada pela autora.

Não houve aprimoramento na elaboração das traduções do TAO2 na segunda coleta se compararmos com a primeira opção de tradução desse TAO. Contudo, se compararmos com a segunda opção disponibilizada, há uma melhora em todos os resultados obtidos, conforme podemos visualizar no Gráfico 17.

Gráfico 17 – Desempenho do TAO2 nas duas coletas para a métrica METEOR

Fonte: Elaborado pela autora.

Esses resultados corroboram aqueles obtidos pelas métricas BLEU e *F-measure* e também por nossa avaliação semântica. É importante, portanto, que o usuário desse TAO conheça todas as opções de tradução disponibilizadas, a fim de que possa obter traduções mais compreensíveis.

Em seguida, temos os resultados da avaliação do TAO3 pela métrica METEOR.

Tabela 27 – Comparação dos resultados da métrica METEOR para o TAO3

METEOR	1ª coleta	2ª coleta
<i>Final score</i>	0.19840393760634323	0.19840393760634323

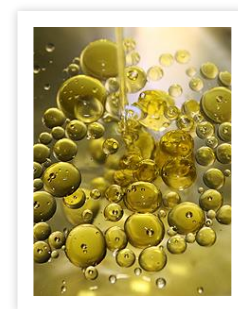
Fonte: Elaborada pela autora.

Não houve diferença nos resultados obtidos pelo TAO3 na métrica METEOR. Portanto, não há necessidade de elaboração de gráficos.

Após analisar todos os dados, tanto entre TAO's e THP's de maneira geral, quanto do desempenho de cada TAO isoladamente, vimos que os THP's obtiveram os maiores resultados, o que corrobora nossa hipótese inicial de que, ao serem comparadas com a tradução de referência estabelecida, as traduções candidatas produzidas pelos TAO's obteriam, em sua maioria, resultados mais baixos do que os das traduções elaboradas pelos THP's.

Verificou-se também que mesmo os TPH's obtendo resultados mais altos na avaliação das três métricas, consideramos que eles são baixos para a qualidade da tradução elaborada. Portanto, as críticas quanto ao uso dessas métricas para a avaliação de traduções produzidas por humanos parecem ser pertinentes: mesmo aquela que tem mais opções lexicais, parece não perceber a variação semântica empregada pelos THP's.

5 CONCLUSÃO – *POURING OIL ON TROUBLED WATERS*



A presente pesquisa teve como objetivo principal avaliar a percepção de tradutores automáticos *online* em relação às expressões idiomáticas e suas metáforas conceituais subjacentes. Refere-se a uma tarefa que envolve dois campos de estudo diferentes, o da Linguística Cognitiva, pois tem como objeto de estudo a linguagem figurada e as metáforas conceituais (fazendo um adendo para incluir também a Fraseologia que já se estabeleceu como um campo de pesquisa em si), e o da Linguística Computacional, uma vez que foram analisadas traduções elaboradas por tradutores automáticos e aplicadas métricas automáticas para avaliação de tradução.

Por se tratar de um estudo sobre metáforas conceituais, a Teoria da Metáfora Conceitual de Lakoff e Johnson (1980), a Hipótese da Metáfora Primária de Grady (1997) e os estudos de Lima (2003 e 2006), compuseram a fundamentação teórica desse aspecto da pesquisa. A compreensão e o processamento da linguagem figurada foram baseados na teoria da saliência do significado, defendida por Giora (1997), decomponibilidade e idiomaticidade por Gibbs (1993), Gibbs e Colston (2012) e Tagnin, sobre *El's*, metáfora e cultura, Kövecses (2005 e 2010).

A fundamentação teórica sobre Tradução Automática e as métricas de avaliação de tradução foram baseadas em Bar-Hillel (1952), Koehn (2010), Jurafsky e Martin (2009) e Cunha (2013), dentre outros.

Apesar de existirem diferentes *El's*, optou-se por estudar aquelas que têm em sua composição lexical, um alimento. O ser humano necessita de alimento para sua sobrevivência. A experiência corpórea com esse elemento fundamental para a vida, é frequente e constante. Até a falta do alimento causa experiências e sensações corpóreas que, conseqüentemente, se tornam uma fonte riquíssima de metáforas conceituais. As línguas de interesse para a elaboração das traduções foram o inglês britânico e norte-americano e o português do Brasil.

O objetivo geral e os objetivos secundários, no total de cinco, exigiram uma pesquisa de caráter exploratório-descritivo de base quali-quantitativa, baseada em *corpora*, na modalidade aplicada. A análise qualitativa descritiva foi realizada por meio da identificação dos alimentos que compõem as EI's analisadas e a possível interferência cultural para essa escolha. Também foi feita uma análise das metáforas conceituais subjacentes às EI's, tendo sido necessária uma pesquisa bibliográfica para averiguar o estado da arte dos diversos campos de estudo envolvidos supracitados.

A pesquisa tem uma vertente experimental que objetivou avaliar a percepção dos TAO's em relação às metáforas conceituais subjacentes às EI's ao elaborarem suas traduções. O produto tradutório desse experimento gerou excertos que foram analisados e comparados com uma tradução de referência estabelecida previamente, além da tradução dos mesmos excertos elaboradas por três THP's. Ao final, os dados coletados foram submetidos a três métricas reconhecidas como apropriadas para a avaliação de traduções. A partir dos resultados obtidos pelas métricas de avaliação de tradução, foi possível elaborar análises quantitativas e compará-las à análise semântica realizada no atual estudo.

O tipo de tradutor, se um TAO ou um THP é a única variável independente desta pesquisa. A Linguística de *Corpus* foi fundamental para a construção das sentenças candidatas e de referência que foram analisadas. Foi utilizado o *Webcorp*, uma ferramenta *online* que utiliza a *Web* como *corpus* e de onde foram retirados todos os excertos traduzidos. Essa ferramenta também foi utilizada para verificar a atual frequência de uso das EI's. Outra ferramenta utilizada, foi o *Google Books N-gram Viewer*, que faz pesquisas de até, no máximo, pentagramas em mais de cinco milhões de livros, publicados entre 1500 e 2008, e disponibiliza o percentual de ocorrências dessas frases em forma de gráfico.

Apresenta-se a seguir, os resultados encontrados para cada objetivo secundário traçado, após análise e discussão dos dados.

O primeiro objetivo era verificar como os alimentos que compõem as EI's eram utilizados, especificamente, para verificar se havia influência cultural nesse uso. Foram selecionadas 43 EI's de dois dicionários de EI's, nas duas línguas de interesse para esta análise.

Foram encontradas evidências de influência cultural na realização de várias das EI's estudadas, em uma língua e na outra. Para expressar a facilidade com

que se faz algo, por exemplo, usa-se em inglês, **a piece of cake**, enquanto em português temos **mamão com mel/açúcar**, alimentos facilmente encontrados na mesa do brasileiro, bem mais do que um bolo que, apesar de atualmente ser comum, provavelmente não era na época em que a EI foi criada. As EI's que contêm o mesmo alimento nas duas línguas são, provavelmente, fruto de empréstimo entre línguas para a tradução de livros mundialmente conhecidos, como a Bíblia e as fábulas de Esopo. A hipótese para este objetivo de que a cultura tem um papel importante na composição das EI's, foi corroborada.

O segundo objetivo secundário diz respeito às metáforas conceituais subjacentes às EI's. Para atingir esse objetivo, foi necessário analisar e classificar as 43 EI's em grupos, de acordo com a metáfora conceitual que as licenciava, para que fosse possível verificar se os TAO's as reconheceriam quando elaborassem suas traduções. Identificou-se diversas metáforas, tais como DIFICULDADE/FACILIDADE É UM ALIMENTO DIFÍCIL/FÁCIL DE MANIPULAR/INGERIR, que é uma metáfora primária que licencia 18 das 43 EI's analisadas.

Outras metáforas conceituais identificadas foram RECURSOS SÃO ALIMENTOS e o acarretamento STATUS SOCIAL É UM ALIMENTO CARO/BARATO; TER GANÂNCIA É QUERER UM ALIMENTO, um acarretamento da metáfora conceitual GANÂNCIA É FOME; BOAS IDEIAS SÃO BONS ALIMENTOS e sua oposição IDEIAS RUINS SÃO ALIMENTOS RUINS; IDEIAS DESINTERESSANTES SÃO ALIMENTOS SEM GOSTO (DEFINIDO) e sua oposição IDEIAS INTERESSANTES SÃO ALIMENTOS BONS; EMBROMAÇÃO É ALIMENTO DE LONGA PREPARAÇÃO; e APRENDER É INGERIR ALIMENTO NUTRITIVO.

Não foi possível identificar com sucesso as metáforas conceituais de dois pares de EI's **to make no bones about sth** - não ter papas na língua e **to have nothing to do with the price of tea in China** - não ter nada a ver com o peixe.

Para as EI's **to pour oil on troubled waters**, **to throw cold water** e **to pour cold water** (jogar/lançar/por água (fria) na fervura) identificou-se as metáforas conceituais RAIVA É UM LÍQUIDO QUENTE EM UM RECIPIENTE e JOGAR LÍQUIDO NA PANELA FERVENTE É DIMINUIR A RAIVA. Portanto, nesse caso, não estão relacionadas a alimentos. A literatura sobre EI's defende que elas são metáforas convencionalizadas, visão compartilhada na pesquisa e que sustenta a hipótese de que seria possível identificar as metáforas conceituais da maioria das EI's analisadas. Portanto, a segunda hipótese também foi corroborada.

Várias das metáforas identificadas no estudo não estão listadas no banco de dados da *MetaNet*, o que leva a crer que uma futura inclusão já é uma contribuição relevante da pesquisa para essa ferramenta.

O terceiro objetivo secundário, que compõe o cerne da pesquisa, foi verificar quão perceptíveis os TAO's são em relação às metáforas subjacentes às EI's analisadas. A partir de uma avaliação semântica das traduções elaboradas pelos TAO's em duas coletas diferentes (março e dezembro de 2017), pode-se verificar que estes ainda pouco percebem as metáforas, mesmo tendo sido dado contexto suficiente para que elas fossem identificadas. Essa foi a hipótese inicial que foi também corroborada.

Apesar de os TAO's terem tido suas técnicas de tradução aprimoradas com a inclusão de rede neural, não houve grandes avanços no resultado das traduções elaboradas na segunda coleta realizada. Na verdade, identificou-se uma melhor fluidez no texto, mas a percepção das EI's, de maneira geral, continuou bastante baixa, dentro do período pesquisado. Na verdade, para os TAO's 1 e 2, houve uma regressão em seu desempenho. Para o TAO3, não houve qualquer alteração nas traduções elaboradas.

É importante ressaltar, contudo, que, levando-se em consideração a segunda opção de tradução disponibilizada pelo TAO2, foi possível ver uma grande melhoria nos resultados obtidos. Portanto, também para a segunda coleta, a hipótese inicial de que os TAO's não perceberiam as EI's e as suas metáforas correspondentes, e que na maioria das vezes, elaborariam traduções não metafóricas, foi corroborada. Deve-se aqui abrir parênteses para lembrar que os resultados poderão ser diferentes se a mesma pesquisa for elaborada utilizando a versão paga (quando disponibilizada) dos mesmos TAO's.

O quarto objetivo secundário era bastante similar ao terceiro, mudando apenas o tipo de tradutor, de automático para humano. Considerando que a literatura da área defende que a tradução elaborada por humanos é o parâmetro de avaliação para a Tradução Automática, foi necessário avaliar também traduções produzidas por THP's, a fim de serem comparadas àquelas produzidas pelos TAO's quanto à percepção das EI's e às metáforas conceituais subjacentes. Dessa forma, a hipótese era de que os THP's reconheceriam as EI's e suas metáforas conceituais, além de fazer as traduções utilizando linguagem metafórica, mas não necessariamente, utilizariam a mesma EI com alimento da tradução de referência. A partir de uma

análise qualitativa, a qual chamamos de métrica semântica, que levou em consideração as escolhas semânticas dos THP's, foi possível verificar que a tradução elaborada por eles, era mais perceptível às metáforas conceituais identificadas, resultando em traduções mais compreensíveis para a maioria dos excertos, o que corrobora a hipótese inicial.

O quinto e último objetivo secundário visava estabelecer uma comparação entre os resultados das traduções realizadas por ambos os TAO's e os THP's, utilizando métricas de avaliação de Tradução Automática conhecidas e também nossa avaliação semântica. Uma tradução de referência elaborada no atual estudo foi utilizada como parâmetro. Essa comparação foi relevante para que fosse verificado qual tipo de tradutor obteria melhores resultados nas métricas escolhidas. A hipótese era de que os THP's obteriam melhores resultados do que os TAO's, apesar do conhecimento de que as métricas de avaliação de Tradução Automática, em geral, penalizam as variações lexicais que são comuns de serem utilizadas pelos THP's.

As três métricas de avaliação de Tradução Automática utilizadas foram BLEU, *F-measure* e METEOR. Conforme previsto em hipótese, os THP's obtiveram resultados melhores em todas as métricas, exceto para a pontuação de *recall* da métrica *F-measure*, cujo maior resultado foi obtido pelo TAO1. Apesar de os resultados favorecerem os THP's, a diferença de pontuação entre eles e os TAO's é muito pequena, o que leva a crer que o uso de métricas automáticas para avaliação de tradução, ainda tem um longo caminho para percorrer antes de atingir o desempenho esperado, ou pelo menos, similar ao dos humanos. A hipótese inicial para esse objetivo também foi corroborada pelos resultados obtidos.

Considerando todos os resultados, é possível concluir que os TAO's, na maioria das vezes, ainda têm uma baixa percepção das metáforas conceituais subjacentes e EI's pesquisadas, pois pouco realizam análises semânticas, não reconhecendo as EI's nem na perspectiva clássica, como expressões congeladas (*chunks*), e muito menos na visão de vanguarda das metáforas conceituais, o que compromete sensivelmente a compreensão da tradução de textos com EI's.

Acredita-se que esses resultados forneçam subsídios que possam aperfeiçoar os TAO's, tornando-os ferramentas cada vez mais eficientes para os usuários. A apropriação de estudos da Linguística Cognitiva, como a Teoria da Metáfora Conceitual, pela inteligência artificial, poderia provocar grandes avanços

para a área de Tradução Automática. Espera-se também que os resultados encontrados possam ser incluídos em disciplinas na formação de futuros THP's.

Para trabalhos futuros, sugere-se criar uma métrica automática que possa incorporar, de forma mais eficaz, as escolhas semânticas realizadas pelos TAO's e THP's; investigar o aprendizado (*machine learning*) dos TAO's neurais em relação à linguagem figurada – uma forma de realizar isso, seria alimentar os TAO's que disponibilizam a opção 'melhorar a tradução' com traduções de referência e, após um prazo estabelecido, verificar se elas foram incorporadas e aceitas pelos sistemas; investigar a percepção dos TAO's pagos em relação a essa e outras formas de linguagem figurada; e realizar pesquisa do português para o inglês, para verificar se os resultados seriam semelhantes.

REFERÊNCIAS

- ABREU, D.T.B. Expressões idiomáticas: um estudo sob a perspectiva da linguística cognitiva. **Signo**. Santa Cruz do Sul, v. 35, n. 59, p. 92-104, jul./dez., 2010. Disponível em: <<http://online.unisc.br/seer/index.php/signo/index>>. Acesso em: 05 out. 2017.
- BAR-HILLEL, Y. The treatment of "idioms" by a translating machine. In: CONFERENCE ON MECHANICAL TRANSLATION at Massachusetts Institute of Technology, 1952, not published. **Proceedings...** Disponível em: <<http://www.mit-archive.info/MIT-1952-Bar-Hillel-2.pdf>>. Acesso em: 07 dez. 2017.
- BERRADA, K. Food metaphors: a contrastive approach. **The Metaphorik**. Deutschland, v. 13, p. 7-38, 2007. Disponível em: <<http://www.metaphorik.de/13/berrada.pdf>>. Acesso em: 31 ago. 2013.
- CAMARGO, S.; STEINBERG, M. **Dicionário de expressões idiomáticas metafóricas português-inglês**. São Paulo: E.P.U., 1989.
- _____. **Dictionary of metaphoric idioms English-Portuguese**. São Paulo: E.P.U., 1990.
- CHEN, J.; DING, R.; JIANG, S.; KNUDSON, R. A preliminary evaluation of metadata records machine translation. **The Electronic Library**. Bingley, UK. v. 30, n. 2, p. 264-277, 2012. Disponível em: <<https://pdfs.semanticscholar.org/f18d/b5330a210369fcac3ab8dbfac5d8821d3caf.pdf>>. Acesso em: 05 jul. 2017.
- CUNHA, T. M. da. **A criação de um sistema híbrido de tradução automática para a conversão de expressões nominais da língua inglesa para a língua portuguesa**. 2013. 163 f. Tese (Doutorado em Linguística) – Programa de Pós-Graduação em Linguística, Universidade Federal do Ceará, Centro de Humanidades, Departamento de Letras Vernáculas, Fortaleza, 2013a.
- _____. **Automatic evaluation of machine translation from English into Portuguese: the DP problem**. Fortaleza. 2013b. Não publicado.
- DORR, B.; OLIVE, J.; McCARY, J.; CHRISTIANSON, C. Machine translation evaluation and optimization. In: OLIVE, J.; CHRISTIANSON, C.; McCARY, J. (Eds.). **Handbook of natural language processing and machine translation**. New York, NY: Springer, 2011. Disponível em: <https://www.researchgate.net/publication/234194809_Machine_Translation_Evaluation_and_Optimization>. Acesso em: 31 dez 2017.
- ELLIS, N. C. Phraseology: the periphery and the heart of language. In: MEUNIER, F.; GRANGER, S. (Eds.). **Phraseology in language learning and teaching**. Amsterdam: John Benjamins. 2008. p. 1-13.

FERNANDO, C.; FLAVELL, R. **On idiom: critical views and perspectives**. Exeter, UK: University of Exeter Press, Exeter Linguistic Studies. 1981. 94p. v. 5.

GIBBS JR. R. W. Taking metaphor out of our heads and putting it into the cultural world. In: GIBBS, R.; STEEN, G. **Metaphor in cognitive linguistics**. Amsterdam: John Benjamins. 1999.

_____. Why idioms are not dead metaphors. In: CACCIARI, C.; TABOSSI P. (Eds.). **Idioms: processing, structure and interpretation**. Hillsdale, NJ: Erlbaum. 1993. p. 57-78.

GIBBS JR, R. W.; COLSTON, H. W. **Interpreting figurative meaning**. New York: Cambridge University Press. 2012. 390 p.

GIORA, R. Understanding figurative and literal language: the graded salience hypothesis. In: DIVJAK, D.; NEWMAN, J. (Eds.) **Cognitive linguistics**. United Kingdom, v. 8. n. 3, p. 183-206, 1997. Disponível em: https://www.tau.ac.il/~giorar/files/giora1997_graded_salience_hypothesis.pdf. Acesso em: 29 maio 2016.

GRADY, Joseph. **Foundations of meaning: primary metaphors and primary scenes**. 1997. 307f. Tese (Doutorado em Linguística). Department of Linguistics, Graduate Division of the University of California in Berkeley, Berkeley, 1997. Disponível em: <http://escholarship.org/uc/item/3g9427m2>>. Acesso em: 18 out. 2017.

JURAFSKY; MARTIN. **Speech and language processing: An Introduction to natural language processing, computational linguistics, and speech recognition**. New Jersey: Pearson Prentice Hall, 2009. 1032 p.

KOEHN, P. **Statistical machine translation**. Cambridge: Cambridge University Press. 2010. 488 p.

KÖVECSES, Z. **Metaphor: a practical introduction**. New York: Oxford University Press, 2010. 400 p.

_____. **Metaphor in culture: universality and variation**. New York: Cambridge University Press. 2005. 314 p.

LAKOFF, G. The meanings of literal. **Metaphor and symbolic activity**. University of California, Lawrence Erlbaum Associates, Inc., v. 1. n. 4, p. 291-296, 1986. Disponível em: https://www.tandfonline.com/doi/abs/10.1207/s15327868ms0104_3>. Acesso em: 02 maio 2016.

LAKOFF, G; JOHNSON, M. Conceptual metaphor in everyday language. **The Journal of Philosophy**. Chicago, v. 77, Aug., p. 453-486, 1980a. Disponível em: https://www.pdcnet.org/jphil/content/jphil_1980_0077_0008_0453_0486>. Acesso em: 25 maio 2016.

_____. **Metaphors we live by**. Chicago: University of Chicago Press, 1980. 276 p.

_____. **Philosophy in the flesh: the embodied mind and its challenge to western thought**. New York: Basic Books, 1999. 624 p.

LAVIE A.; AGARWAL A. METEOR: An automatic metric for MT evaluation with high levels of correlation with human judgments. WORKSHOP ON STATISTICAL MACHINE TRANSLATION, 2.; MEETING OF THE ASSOCIATION FOR COMPUTATIONAL LINGUISTICS, 45., 2007, Prague. **Proceedings...** Prague: 2007, p. 228-231. Disponível em: <<https://www.cs.cmu.edu/~alavie/METEOR/pdf/Lavie-Agarwal-2007-METEOR.pdf>>. Acesso em: 07 jan. 2018.

LEAL, M. A. O. C.; MAIA, A. L. L. M.; MAIA, T. D'M. Um estudo comparativo de três técnicas de tradução automática para processamento de textos. In: ESCOLA REGIONAL DE COMPUTAÇÃO BAHIA ALAGOAS SERGIPE 14., 2014, Feira de Santana. **Anais eletrônicos...** Feira de Santana: 2014. Conference Paper. Disponível em: <<http://erbase2014.uefs.br/informacao-geral/publicacoes>>. Acesso em: 21 jan. 2018.

LIMA, P. L. C.; FELTES, H. P. M.; MACEDO, A. C. P. Teoria da metáfora conceitual: construção da teoria. In: MACEDO, A. C. P. de; FELTES, H. P. de M.; FARIAS, E. M. P. (Org.). **Cognição e lingüística: territórios, mapeamentos e percursos**. Porto Alegre, Caxias do Sul: EDPUCRS, EDUCS, 2008, p. 127-165. v. 1.

LIMA, P. L. C. A nova tipologia da metáfora conceitual. **Revista de humanidades e ciências sociais da UECE**, Fortaleza, v. 5, n.2, p. 17-26, 2003. Disponível em: <http://www.leffa.pro.br/tela4/Textos/Textos/Anais/Textos_Em_Psicolin/Artigos/A%20nova%20tipologia%20da%20met%20C3%A1fora%20conceitual.pdf>. Acesso em: 29 nov. 2017.

_____. About primary metaphors. **D.E.L.T.A.**. São Paulo, v. 22: n. especial, p. 109-122, 2006. Disponível em: <https://www.researchgate.net/publication/250982716_About_primary_metaphors>. Acesso em: 10 out. 2017.

_____. Metáfora e ensino/aprendizagem de língua estrangeira. In: _____; ARAÚJO A. D. (Org.). **Questões de lingüística aplicada**. Fortaleza: EdUECE, 2005, p. 97-123. v. 1.

_____. **Desejar é ter fome: novas idéias sobre antigas metáforas conceituais**. 1999. 213f. Tese (Doutorado em Lingüística) – Curso de Lingüística do Instituto de Estudos da Linguagem. Universidade Estadual de Campinas, Campinas, SP. 1999.

_____. **Usando a cabeça: Um estudo da representação do substantivo cabeça no sistema conceitual das línguas inglesa e portuguesa, através de expressões metafóricas convencionais**. 1995. Dissertação (Mestrado em Letras – Área de Concentração em Língua Inglesa) – Centro de Humanidades. Universidade Estadual do Ceará, Fortaleza, 1995. 174p.

_____. **Comunicação pessoal**: encontros de orientação. Fortaleza. 2017. Não publicado.

LIU, W.; SHEN, H. CiteSpace II: Idiom studies development trends. **Journal of arts and humanities (JAH)**, Maryland, USA, v. 2, n. 2, Mar., p. 85-97, 2013. Disponível em: < <https://www.theartsjournal.org/index.php/site/article/view/72/71>>. Acesso em: 12 nov. 2017.

MELO, F. R.; MATOS, H. C. O.; DIAS, E. R. B. Aplicação da métrica BLEU para avaliação comparativa dos tradutores automáticos Bing tradutor e Google tradutor. **Revista e-escrita - Revista do Curso de Letras da UNIABEU**, Nilópolis, v.5, n. 3, set/dez. 2014. Disponível em: <<http://revista.uniabeu.edu.br/index.php/RE/article/view/1719>>. Acesso em: 06 dez. 2017.

MONTEIRO, S. M. M. Desmistificando a tradução automática: um estudo sobre seu uso e sua recepção. In: JORNADA NACIONAL DE ESTUDOS LINGÜÍSTICOS, 19., 2002, Fortaleza. **Anais eletrônicos...** Fortaleza: 2002. Disponível em: <http://www.nilc.icmc.usp.br/til/til2003_English/poster/Monteiro_22.pdf>. Acesso em: 02 jan. 2018.

ORTIZ ALVAREZ, M. L. Metáforas em uso: a riqueza da linguagem metafórica. In: SILVA, M. do P. S. C. da S. (Org.). **Linguagens**: estudos interdisciplinares e multiculturais. As interfaces dos estudos linguísticos. Belém do Pará: Editora da Universidade da Amazônia UNAMA, v. 4, p. 149-170, 2006.

SARDINHA, T. B. **Metáfora**. São Paulo: Parábola Editorial. 2007. 167 p.

SIQUEIRA, M. S. G.; PARENTE, M. A. P. Metáforas primárias e influências culturais. In: SIMPÓSIO NACIONAL, 11.; SIMPÓSIO INTERNACIONAL DE LETRAS E LINGÜÍSTICA – SILEL, 1., 2006, Uberlândia. **Anais eletrônicos...** Uberlândia: ILEEL, 2006. Disponível em: <http://www.filologia.org.br/ileel/artigos/artigo_466.pdf>. Acesso em: 30 nov. 2017.

SPULBER, A. Food-based conceptual metaphors in English. In: MEETING OF STYLES ACROSS CULTURES, 7., 2008, Bacau, **Presentation...** Bacau: 2008. Disponível em: < <https://pt.scribd.com/presentation/51655850/CONCEPTUAL-FOOD-METAPHORS-finalllll>>. Acesso em: 31 maio 2016.

TAGNIN, S. E. O. **O jeito que a gente diz**: combinações consagradas em inglês e português. Barueri, SP: DISAL. 2013. 223 p.

TAM, D. T. P. **Food metaphor in English and Vietnamese** – A contrastive analysis. HCMC University of Pedagogy. 2009. Disponível em: <http://wisevn.com/ed/CA2009/4C06_DoThiPhuongTam_FoodMetaphorinEnglishandVietnamese.doc>. Acesso em: 31 ago. 2013.

TODARO, M. S. F. **Tradução por software**: estado da arte e análise comparativa no contexto do Falibras. 2013. 47f. Trabalho de Conclusão de Curso (Bacharelado em Sistemas de Informação) – Pró-reitoria de Graduação, Instituto de Computação

Universidade Federal de Alagoas, Maceió, 2013. Disponível em:
<<http://www.ufal.edu.br/unidadeacademica/ic/graduacao/sistemas-de-informacao/banco-de-monografias/2013/traducao-por-software-estado-da-arte-e-analise-comparativa-no-contexto-do-falibras>>. Acesso em: 19 fev. 2016.

TORAL, A.; GASPARI, F.; NASKAR, S. K.; WAY, A. Comparative evaluation of research vs. online MT systems. In: INTERNATIONAL CONFERENCE OF THE EUROPEAN ASSOCIATION FOR MACHINE TRANSLATION 15., 2011, Leuven, Belgium. **Proceedings...** Leuven, Belgium: 2011. p. 13-20. Disponível em:
<<https://pdfs.semanticscholar.org/327f/30d868a4fcd813353bc1acda1ac0726271eb.pdf>>. Acesso em: 05 jul. 2017.

TRIPATHI, S.; SARKHEL, J. K. Approaches to machine translation. **Annals of library and information studies**. New Delhi, v. 57, Dec., p. 388-393, 2010. Disponível em:
<[http://nopr.niscair.res.in/bitstream/123456789/11057/1/ALIS%2057\(4\)%20388-393.pdf](http://nopr.niscair.res.in/bitstream/123456789/11057/1/ALIS%2057(4)%20388-393.pdf)>. Acesso em: 11 abr. 2016.

TSENG, M-Y. Primary metaphors and multimodal metaphors of food: examples from an intercultural food design event. **Metaphor and Symbol**. United Kingdom, v. 32, n. 3, p. 211-229, 2017. Disponível em:
<<https://doi.org/10.1080/10926488.2017.1338027>>. Acesso em: 01 out. 2017.

APÊNDICES

APÊNDICE A – Aprovação pelo Comitê de Ética em Pesquisa da UECE

Saúde
Ministério da Saúde

Público
Pesquisador
Alterar Meus Dados

Cadastrados

 SARAH VIRGINIA CARVALHO RIBEIRO - Pesquisador | V3.2
Sua sessão expira em: 30min 21

DETALHAR PROJETO DE PESQUISA

DADOS DA VERSÃO DO PROJETO DE PESQUISA

Título da Pesquisa: Expressões idiomáticas com alimentos em português e inglês: uma avaliação de tradutores automáticos gratuitos online, à luz da Teoria das Metáforas Conceituais

Pesquisador Responsável: SARAH VIRGINIA CARVALHO RIBEIRO

Área Temática: Versão: 1

CAAE: 61420216.4.0000.5534

Submetido em: 21/10/2016

Instituição Proponente: Programa de Pós-Graduação em Linguística Aplicada

Situação da Versão do Projeto: Aprovado

Localização atual da Versão do Projeto: Pesquisador Responsável

Patrocinador Principal: Financiamento Próprio

Comprovante de Recepção: PB_COMPROVANTE_RECEPCAO_612376

DOCUMENTOS DO PROJETO DE PESQUISA

- ↳ Versão Atual Aprovada (PO) - Versão 1
 - ↳ Projeto Original (PO) - Versão 1
 - ↳ Documentos do Projeto
 - ↳ Comprovante de Recepção - Submissão
 - ↳ Folha de Rosto - Submissão 1
 - ↳ Informações Básicas do Projeto - Subm
 - ↳ Outros - Submissão 1
 - ↳ Projeto Detalhado / Brochura Investiga
 - ↳ TCLE / Termos de Assentimento / Justif
 - ↳ Apreciação 1 - UNIVERSIDADE ESTADUA
 - ↳ Projeto Completo

Tipo de Documento	Situação	Arquivo	Postagem	Ações

LISTA DE APRECIÇÕES DO PROJETO

Apreciação	Pesquisador Responsável	Versão	Submissão	Modificação	Situação	Exclusiva do Centro Coord.	Ações
PO	SARAH VIRGINIA CARVALHO RIBEIRO	1	21/10/2016	19/12/2016	Aprovado	Não	

Expressões idiomáticas com alimentos em português e inglês: uma avaliação de tradutores automáticos gratuitos online, à luz da Teoria das Metáforas Conceituais

HISTÓRICO DE TRÂMITES

Apreciação	Data/Hora	Tipo Trâmite	Versão	Perfil	Origem	Destino	Informações
PO	19/12/2016 16:39:48	Parecer liberado	1	Coordenador	UNIVERSIDADE ESTADUAL DO CEARÁ - UECE	PESQUISADOR	
PO	07/12/2016 13:18:24	Parecer do colegiado emitido	1	Coordenador	UNIVERSIDADE ESTADUAL DO CEARÁ - UECE	UNIVERSIDADE ESTADUAL DO CEARÁ - UECE	
PO	04/12/2016 16:47:17	Parecer do relator emitido	1	Membro do CEP	UNIVERSIDADE ESTADUAL DO CEARÁ - UECE	UNIVERSIDADE ESTADUAL DO CEARÁ - UECE	
PO	04/12/2016 16:37:30	Aceitação de Elaboração de Relatoria	1	Membro do CEP	UNIVERSIDADE ESTADUAL DO CEARÁ - UECE	UNIVERSIDADE ESTADUAL DO CEARÁ - UECE	
PO	24/11/2016 13:21:23	Confirmação de Indicação de Relatoria	1	Coordenador	UNIVERSIDADE ESTADUAL DO CEARÁ - UECE	UNIVERSIDADE ESTADUAL DO CEARÁ - UECE	
PO	23/11/2016 12:38:26	Indicação de Relatoria	1	Coordenador	UNIVERSIDADE ESTADUAL DO CEARÁ - UECE	UNIVERSIDADE ESTADUAL DO CEARÁ - UECE	
PO	27/10/2016 11:47:57	Aceitação do PP	1	Coordenador	UNIVERSIDADE ESTADUAL DO CEARÁ - UECE	UNIVERSIDADE ESTADUAL DO CEARÁ - UECE	
PO	21/10/2016 15:14:14	Submetido para avaliação do CEP	1	Pesquisador Principal	PESQUISADOR	UNIVERSIDADE ESTADUAL DO CEARÁ - UECE	

LEGENDA:

(*) Apreciação

PO = Projeto Original de Centro Coordenador	POp = Projeto Original de Centro Participante	POc = Projeto Original de Centro Coparticipante
E = Emenda de Centro Coordenador	Ep = Emenda de Centro Participante	Ec = Emenda de Centro Coparticipante
N = Notificação de Centro Coordenador	Np = Notificação de Centro Participante	Nc = Notificação de Centro Coparticipante

(*) Formação do CAAE

Ano de submissão do Projeto					Tipo do centro		Código do Comitê que está analisando o projeto													
n	n	n	n	n	a	a	.	d	v	.	t	x	x	x	.	i	i	i	i	i
Sequencial para todos os Projetos submetidos para apreciação					Dígito verificador		Sequencial quando estudo possui Centro(s) Participante(s) e/ou Coparticipante(s)													

Voltar

Este sistema foi desenvolvido para os navegadores Internet Explorer (versão 7 ou superior), ou Mozilla Firefox (versão 9 ou superior).

APÊNDICE B – Termo de Consentimento Livre e Esclarecido

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Você está sendo convidado(a) a participar de uma pesquisa que tem como objetivo avaliar o potencial de alguns mecanismos automáticos na tradução de textos em inglês para o português. Nesse contexto, a tradução realizada por humanos é um parâmetro fundamental de comparação, inclusive porque o processo tradutório é uma atividade subjetiva, onde diferentes escolhas lexicais e de estrutura podem ser realizadas, não caracterizando, portanto, existir uma única tradução correta. Leia, por gentileza, este Termo até o final e, ao concordar em participar, assine-o, preencha as informações sobre seu perfil no questionário em anexo, e envie ambos à pesquisadora responsável, abaixo subscrita, por e-mail. Logo após sua aceitação, a pesquisadora lhe enviará a tarefa, que consiste na tradução, para o português, de 40 excertos em inglês. O processo não será supervisionado, uma vez que será realizado em seu próprio espaço de trabalho ou domicílio, via uso de computador. Para minimizar possíveis desconfortos, a tarefa de tradução pode ser realizada em diferentes etapas e de acordo com sua disponibilidade de tempo, dentro do prazo de 30 dias, a partir do recebimento dos excertos. Ao finalizar, o arquivo com as traduções deverá ser encaminhado por e-mail à pesquisadora responsável. É importante registrar que seu nome ou quaisquer informações que possam identificá-lo(a) não serão divulgados sob nenhuma hipótese. Esclarecemos que a pesquisa não lhe trará nenhum ônus além do tempo empregado em sua participação, e que você tem a liberdade de participar ou não da mesma, bem como retirar o seu consentimento a qualquer momento. Esclarecemos também que você não receberá qualquer pagamento, reembolso ou gratificação por sua participação neste estudo, mas terá o agradecimento profundo e sincero da pesquisadora responsável e da comunidade acadêmica, por sua enorme contribuição aos estudos da linguagem. Para esclarecimento de eventuais dúvidas, entre em contato com a profissional responsável pela pesquisa, a doutoranda Sarah Virginia Carvalho Ribeiro, que pode ser encontrada pelos telefones (85) 99983-0404, 98897-0520 e e-mail: sarah.virginia@aluno.uece.br. Se você tiver alguma consideração ou dúvida sobre a ética da pesquisa, entre em contato com o Comitê de Ética em Pesquisa – CEP/UECE, através dos seguintes contatos: Av. Dr. Silas Munguba, 1700, Campus do *Itaperi*, Fortaleza-CE, CEP 60.714.903, telefone (85) 3101 – 9890 ou e-mail: cep@uece.br.

Assinatura da Responsável pela Pesquisa

De posse das informações sobre a pesquisa acima referenciada, concordo voluntariamente em participar da mesma, de forma livre e esclarecida.

Nome: _____

Assinatura: _____

Fortaleza -CE, _____ de _____ de _____.

APÊNDICE C – Perfil do Tradutor Humano

PERFIL DO TRADUTOR HUMANO**1. Identificação:**

Nome: _____

Sexo: _____

Idade: _____

2. Formação acadêmica:

Curso: _____

Período: _____

Universidade: _____

Pós-graduação: sim () não ()

Em tradução: sim () não ()

3. Experiência profissional em tradução de textos:

() De 1 a 5 ano () De 6 a 10 anos

() De 11 a 15 anos () mais de 15 anos

4. Marque os recursos que você utiliza profissionalmente para traduzir, citando-os na linha ao lado:

() Corretores ortográficos _____

() Corretores gramaticais _____

() Dicionários impressos _____

() Dicionários e glossários digitais _____

() Tradutor automático _____

() Memória de tradução _____

() Bancos de dados terminológicos _____

() Bancos de dados de estruturas sintáticas _____

() Outros _____

APÊNDICE D – Excertos Traduzidos com Instruções para os Tradutores Humanos Profissionais

Obrigada por participar desta pesquisa. Sua tarefa consiste em traduzir, para o português, os 35 excertos em inglês que se encontram na tabela a seguir. Você pode realizar esta tarefa em seu próprio espaço de trabalho ou domicílio, via uso de computador, e em diferentes etapas, de acordo com sua disponibilidade de tempo. Você tem o prazo de 30 dias para completá-la, a partir do recebimento dos excertos. Ao finalizar, pedimos a gentileza de enviar o arquivo com as traduções para a pesquisadora responsável, através do e-mail sarah.virginia@aluno.uece.br. Para fins de esclarecimento, lembramos que não existe uma única tradução correta para cada excerto, uma vez que o processo tradutório é uma atividade subjetiva, onde diferentes escolhas lexicais e de estrutura podem ser realizadas. Agradecemos mais uma vez sua participação. Bom trabalho.

Nº	Excertos	Tradução
1	So the current stance of fiscal policy may well have been seriously inappropriate before the Brexit vote upset the apple cart. But if it was inappropriate then, it is doubly so now. Because of widespread fears about the consequences of Brexit, there is a realistic danger of substantially weaker GDP growth in the near term.	
2	It's the Iraq war. What was Blair thinking? He must have no backbone to have let the US convince him to go.	
3	People (who earn) over £100,000 will actually be paying more, I make no bones about that, in three years' time, when the changes come through,' he said.	
4	Four million children in this country live on the breadline – one of the worst rates in the industrialised world.	
5	It was taken into the language of the local catching game, that is, baseball, and 'no-hoper' teams were unkindly given that name. The Salt Lake Tribune reported on such a team in May 1899: 'The Butterfingers [Team] will cross bats with the Salt Lake Juniors at Calder's Park Tuesday'.	
6	Let me firstly establish that I dislike exercise. I am a cake-eater and have a sedentary job. Exercise is a chore.	
7	Hobby's Toskana Exclusive D750 H UCC may not have made much impression on German showgoers but it is sure to sell like hot cakes here – complete with palatial rear lounge in leather and a habitation door on the UK nearside.	
8	Mr. Henderson stood firm, as we know, and whatever form future collaboration on specific European problems may take, its sting has been drawn, and the League's pre-eminence has been asserted. As usual, it is the business of the British Government to pull the chestnuts out of the fire. We have our reward, however, in the notable appreciation by all those who have the best interests of the League at heart.	
9	I thought that BTN had to beat Horsham 1st. You know what happens when you count your chickens before they hatch. Remember this is football.	
10	There is no frame of reference for what it is like to be in a normal household. You also have no frame of reference for what is O.K. to say and feel. In a more typical situation, one does not have to walk on eggs all the time.	
11	There should be a three strikes and you're out policy. I would think three chances to be a decent human being isn't too much of a stretch. I also believe that if your child is excluded or suspended, the family's benefits and housing should be removed or suspended. How can you live off the fat of the land when you make no contribution to the country?"	
12	Other veterans are disgruntled by the Government's handling of arrangements. Ron Eastwood, 70, said: 'The muddle is all the fault of the Johnny-come-lately politicians who like to have a finger in the pie. They are dealing with something they know nothing about.	

13	The movie is neither fish nor flesh: neither a Russian play nor a Fellini experience. Mastroianni is charming, but the performance is derivative of his performances for Fellini. You needn't buy nor view this movie.	
14	He would speak of his Ipswich players like a man reciting a litany: Burley, Mills, Beattie, Wark, Talbot. Joe Mercer once said of Bill Shankly that he 'thinks all his geese are swans'. So it was with Bobby. In later years, he would nurture some of the best that the game has known, from Shilton and Gascoigne, Lineker and Bryan Robson to Romario, Ronaldo, Figo and Shearer. But it was the Ipswich lads he would remember with most affection, for they helped him learn his trade and make his name.	
15	We managed 3 rides in the day, because the park was jammed packed. The population of the UK is growing rapidly. As a result attractions such as theme parks are far busier than they used to be. This trend will probably continue. Theme park owners need to wise-up to this or they'll kill the goose that lays the golden eggs. Queuing over 60 mins for a ride is not acceptable, given that you've paid a lot of money to get into the park.	
16	Many people see payday loans as a temporary financial fix. Borrowers think they just need a little more time to get out of a jam. But often that's not what happens.	
17	So the election has produced a clear answer and everything is straightforward now? You thought we were sitting pretty with the future pretty as a picture. But sorting out Scotland is going to cost a pretty penny, and pretty much everything could turn into a pretty kettle of fish. Let's hope for the future.	
18	If you're looking for a new position to try with your other half, why not try the Milk and Water Embrace position? If you're used to keeping things in the bedroom when it comes to making love, then branching out and getting a chair involved could be a huge rush for the both of you.	
19	I will arrive in Casablanca on Sunday afternoon, Nov. 6, 2016, and could be free to travel to Pileta Caves in southern Spain that evening if I could find a "connection". I would have to be back in Casablanca by noon on Tuesday, Nov. 8. Anyone think I could "pull this off"? What city would I fly "to"? I'm assuming I would have to rent a car and drive to where the caves are. I welcome any ideas! (I know, I should have planned to fly to southern Spain "first" and then on to Morocco, but it's too late to "cry over spilled milk".) Thanks!!	
20	For he was more learned than any of his predecessors, abbots of Croyland, having imbibed literature of every description with his mother's milk from his very cradle.	
21	The advent of High Definition television in the UK caused a right old stir among make-up artists, actresses and TV presenters. This new technology meant that old make-up techniques and old products simply would not cut the mustard any more. Every change in texture would now be visible.	
22	It is a commonplace of philosophy that the notion of free will is a hard nut to crack. A simple, compelling argument can be made to show that behavior for which an agent is morally responsible cannot be the outcome of prior determining causal factors.	
23	On another day, Baker attempted to pour oil on troubled waters after a prosecutor lost her temper with a defence barrister in court. 'These things can get terribly heated,' he said during a break after the flare-up. 'I know what I'll do; when we are all sat down and ready to begin again, I'll beam at her.'	
24	"It's a sign of another thing: that liberal men are wimps who can't handle the hot potato that is a combination of feminine sexuality and female political brilliance."	
25	When Mark Messier was captain for a period, before and rumors swirled about the Moose being gay, Bure wanted gone. But that all should have nothing to do with the price of tea in China. Linden eventually had his jersey retired. So why not retire No. 10?	
26	In a speech to the Global Warming Policy Foundation, Mr Paterson will highlight recent studies and temperature records that he says pour cold water on many of the most alarmist forecasts. He will say: 'I readily accept the main points of the greenhouse theory. Other things being equal, carbon dioxide emissions will	

	produce some warming. 'The question always has been: how much? On that, there is considerable uncertainty.'	
27	As the owner and MD of my own company, employing freelancers and as a 42 year old consumer, with years and year of experience of bad customer service, I place a lot more emphasis on soft skills now. For me, they separate the wheat from the chaff. Admittedly, their importance will vary depending on the role required of the individual.	
28	But surely you can have all those things and still have a lemon on your hands? You don't need those things to make a good Bond film. In fact, if you get to the end of the film and think you've missed them, it's going to be the least of your worries, because putting them in won't make the film any better.	
29	It's a rare exercise class that can cater for mixed abilities and I can vouch that it's no piece of cake for fit people. I have to admit that I thought I'd find it too easy, but I'm certainly feeling the effects after pushing myself on all the moves.	
30	All budgets end up being tagged. Last year's was the pasty and caravan budget. By delivering a 1p cut in the price of a pint this one deserves to be called the small beer budget. And from a tax campaigner's point of view that is exactly what it was.	
31	The purpose of this interview is to determine how well you react under stress. The interviewer is likely to be rude or sarcastic and is likely to cast aspersions on your skills and/or experience. Be prepared with what to expect in such an interview and ensure that you never rise to the bait no matter how important it seems. Stay calm and collected, but ensure that you are assertive in the point you are making.	
32	A student in the audience for Thursday night's televised debate on the EU referendum caused a stir on social media after accusing David Cameron of "waffling". Soraya Bouazzaoui, an English Literature student at Southampton Solent University, pressed the Prime Minister over whether the UK should be working with Turkey and about the lack of positive campaigning by the Remain group, despite being a Remain voter herself.	
33	It would be sheer madness to contemplate even for a moment giving up Britain's independent nuclear deterrent. And there is no room for compromise, and no room for cheese paring. We need a full fleet of four submarines, capable between them of providing what the military call 'Continuous At Sea Deterrence', or permanent, around-the-clock cover.	
34	We always create a romantic environment for all of our hotel guests and club owners. Attention to detail is our daily bread and we love it.	
35	They can meet and discuss these various topics and outline a plan of procedure and now let's set about to do something, not all talk and no cider, but let's see to it that something is done along these lines.	
Os excertos sombreados em cinza foram retirados da análise por diferentes motivos explicados no texto da tese		

APÊNDICE E – Traduções de Referência

Neste quadro encontram-se as traduções de referência somente dos 35 excertos que foram traduzidos pelos TAO's e pelos THP's. Em vermelho, destacamos as EI's em negro.

Nº	Inglês	Português
1	So the current stance of fiscal policy may well have been seriously inappropriate before the Brexit vote upset the apple cart .	Então, a postura atual de política fiscal pode ter sido seriamente inapropriada antes da votação do Brexit entornar o caldo .
2	He must have no backbone to have let the US convince him to go.	Ele não deve ter tutano para ter deixado os EU convencê-lo a ir.
3	People (who earn) over £100,000 will actually be paying more, I make no bones about that , in three years' time, when the changes come through,' he said.	Pessoas (que ganham) acima de £100,000 pagarão mais, na verdade, eu não tenho papas na língua a esse respeito, em três anos, quando as mudanças vierem,' ele disse.
4	Four million children in this country live on the breadline .	Quatro milhões de crianças neste país passam a pão e água .
5	The Butterfingers [Team] will cross bats with the Salt Lake Juniors at Calder's Park Tuesday'.	O (time dos) Mãos de manteiga vai cruzar os bastões com os Salt Lake Juniors no Calder's Park na terça-feira.
6	I am a cake-eater and have a sedentary job.	Eu sou arroz de festa e tenho um trabalho sedentário.
7	it is sure to sell like hot cakes here.	é certeza vender como água aqui.
8	As usual, it is the business of the British Government to pull the chestnuts out of the fire .	Como de costume, o negócio do Governo Britânico é tirar a sardinha com a mão do gato .
9	You know what happens when you count your chickens before they hatch .	Você sabe o que acontece quando você conta com os ovos dentro da galinha .
10	In a more typical situation, one does not have to walk on eggs all the time.	Em uma situação mais típica, não é necessário pisar sobre ovos o tempo inteiro.
11	How can you live off the fat of the land when you make no contribution to the country?"	Como se pode viver à tripa forra quando não se contribui para o país?
12	The muddle is all the fault of the Johnny-come-lately politicians who like to have a finger in the pie .	A confusão é toda culpa dos políticos que mal chegaram, que gostam de tirar uma casquinha .
13	The movie is neither fish nor flesh .	O filme não é carne nem peixe .
14	Joe Mercer once said of Bill Shankly that he 'thinks all his geese are swans' .	Joe Mercer uma vez falou sobre Bill Shankly que ele 'Come frango e arrotta peru' .
15	Theme park owners need to wise-up to this or they'll kill the goose that lays the golden eggs .	Os donos de parques temáticos precisam estar atentos a isso ou eles irão matar a galinha dos ovos de ouro .
16	Borrowers think they just need a little more time to get out of a jam .	Os mutuários pensam que apenas precisam um pouco de tempo para descascar o abacaxi .
17	and pretty much everything could turn into a pretty kettle of fish .	e tudo pode muito bem se transformar em um angu de caroço .
18	If you're looking for a new position to try with your other half, why not try the Milk and Water Embrace position?	Se você está procurando por uma nova posição para experimentar com sua metade da laranja, por que não tentar a posição do Abraço Sem sal?
19	but it's too late to "cry over spilled milk" .	mas é tarde demais para "chorar sobre o leite derramado" .
20	having imbibed literature of every description with his mother's milk from his very cradle.	tendo bebido literatura de todos os tipos com o leite materno desde o berço.

21	This new technology meant that old make-up techniques and old products simply would not cut the mustard any more.	Esta nova tecnologia significava que as antigas técnicas de maquiagem e os velhos produtos simplesmente não queimavam mais o óleo quarenta.
22	It is a commonplace of philosophy that the notion of free will is a hard nut to crack.	É lugar-comum na filosofia que a noção de livre arbítrio é um osso duro de roer.
23	On another day, Baker attempted to pour oil on troubled waters after a prosecutor lost her temper with a defense barrister in court.	Outro dia, Baker tentou jogar água fria na fervura depois que um promotor público perdeu a paciência com um advogado de defesa no tribunal.
24	liberal men are wimps who can't handle the hot potato that is a combination of feminine sexuality and female political brilliance.	homens liberais são covardes que não conseguem segurar a batata quente que é uma combinação de sexualidade feminina e brilhantismo político feminino.
25	But that all should have nothing to do with the price of tea in China.	Mas tudo isso não deveria ter nada a ver com o peixe.
26	Mr Paterson will highlight recent studies and temperature records that he says pour cold water on many of the most alarmist forecasts.	O Sr. Paterson enfatizará estudos recentes e recordes de temperaturas que ele diz jogam água fria na fervura em muitas das previsões mais alarmistas.
27	For me, they separate the wheat from the chaff.	Para mim, elas separam o joio do trigo.
28	But surely you can have all those things and still have a lemon on your hands?	Mas, certamente, você pode ter todas aquelas coisas e ainda ter um abacaxi nas mãos?
29	I can vouch that it's no piece of cake for fit people.	Eu posso confirmar que não é bolinho para pessoas que estão em forma.
30	By delivering a 1p cut in the price of a pint this one deserves to be called the small beer budget.	Ao trazer o corte de 1p no preço de um quartilho, esse orçamento merece ser chamado de orçamento docafé pequeno.
31	ensure that you never rise to the bait no matter how important it seems.	certifique-se de nunca morder a isca , não importa quão importante pareça.
32	A student in the audience for Thursday night's televised debate on the EU referendum caused a stir on social media after accusing David Cameron of " waffling ".	Uma aluna na plateia do debate sobre o referendun da EU que foi televisionado quinta à noite, causou uma celeuma nas media sociais depois de acusar David Cameron de " encher linguça ".
33	And there is no room for compromise, and no room for cheese paring.	E não há espaço para concessões, e nem espaço para ser pão duro.
34	Attention to detail is our daily bread and we love it.	Atenção aos detalhes é o pão nosso de cada dia e nós amamos isso.
35	now let's set about to do something, not all talk and no cider	agora vamos definir algo para ser feito, e não ser mais vozes do que nozes

Os excertos sombreados em cinza foram retirados da análise por diferentes motivos explicados no texto da tese

ANEXOS

ANEXO A – El's com alimentos em português e suas respectivas traduções em inglês

01	Descascar o abacaxi <i>Não sei como descascar este abacaxi.</i>	To get out of a jam ; to handle the hot potato ; to solve a rough problem <i>I don't know how to get out of this jam.</i>
02	Ficar com o abacaxi <i>Todos tiveram férias no escritório e eu fiquei com o abacaxi.</i>	To be left holding the baby (B)/the bag (A); to be left with the hot potato <i>Everybody at the office went off on holidays and I was left holding the baby.</i>
03	Passar o abacaxi <i>Meu chefe não quis fazer o trabalho e me passou o abacaxi.</i>	To give sm the hot potato ; to pass the buck <i>My boss didn't want to do the job and gave me the hot potato to handle.</i>
04	Beber água de chocalho <i>Não suporto aquela menina, parece que engoliu água de chocalho.</i>	To babble like a brook <i>I can't stand that girl – she babbles like a brook.</i>
05 06	Cozinhar em água morna = Cozinhar em banho-maria = Cozinhar o galo <i>Estão cozinhando nosso abono de Natal em água morna.</i>	To put on the back burner (A); to procrastinate, to postpone the solution <i>They have put our Christmas bonus on the back burner.</i>
07	Jogar/Lançar/Por água (fria) na fervura <i>Meus filhos estavam brigando sem parar, mas eu consegui por água na fervura.</i>	To pour oil on troubled waters; to throw cold water /to put a damper on things <i>My sons were having an endless quarrel, but I managed to pour oil on troubled water.</i>
08	Por mais água no feijão <i>É claro que seu irmão também pode vir para o almoço, vou por mais água no feijão.</i>	To stretch the <u>meal</u> <i>Sure your brother can come for lunch. I'll stretch the meal.</i>
09	Confundir alhos com bugalhos <i>Você está confundindo alhos com bugalhos porque eu não disse nada disso.</i>	To make black of white/ducks and drakes of everything <i>You are making ducks and drakes of everything. I didn't say anything of the kind.</i>
10	Um angu de carçoço <i>Eles fizeram um angu de carçoço da situação, e agora ninguém entende nada.</i>	A pretty hash; A pretty kettle of fish <i>They made a pretty hash of the situation and now no one understands anything.</i>
11	Ser arroz de festa <i>Se Irene não foi ao baile deve estar doente, pois ela é arroz de festa.</i>	To be a cake -eater/a standing feature <i>If Irene didn't go to the dancing party she must be sick because she's a cake-eater.</i>
12	Por azeitona na empada de a. <i>Só consegui terminar o trabalho porque ele pôs azeitona na minha empada.</i>	To lend sm a hand <i>I was only able to finish the work because he lent me a hand.</i>
13	Dar uma banana para <i>Na minha idade você pode dar uma banana para o que pensam de você.</i>	To thumb a nose to; to cock a snook; to fly in the face/teeth of <i>At my age you can thumb a nose at what people think of you.</i>
14	Ser um banana <i>Não creio que ele realize alguma coisa, ele não passa de uma banana</i>	To be a wishy-washy/timid person <i>I don't believe he will accomplish anything, he's just a wishy-washy person.</i>
15	Estar com uma batata quente nas mãos <i>Estou com uma batata quente nas mãos: tenho que negociar com os grevistas.</i>	To have a hot potato /troublesome problem; to be under perplexing circumstances <i>I have a hot potato in my hands: I must negotiate with the strikers.</i>
16	Mandar a. plantar batatas = mandar a. tomar banho <i>Se ele me encher de novo vou mandá-lo tomar banho.</i>	To send sm packing; to tell sm to drop dread <i>If he bothers me again I'll send him packing.</i>
17	Não ser bolinho <i>Trabalhar para ele não é bolinho.</i>	To be no piece of cake /no pic-nic <i>To work for him is no pic-nic.</i>
18	Ser de fritar bolinho = não ser de nada <i>Não conte com o João, ele não é de nada.</i>	To be wishy-washy person

		<i>Don't count on João, he's just a wishy-washy person.</i>
19	Dar/Levar o bolo – dar/levar o cano <i>Marina estava triste porque o namorado lhe deu o cano.</i>	To stand sm up; to walk out on sm; to fail to show up <i>Marina was sad because her boyfriend walked out on her.</i>
20	Ganhar o bolo <i>Você sabe quem ganhou o bolo?</i>	To sweep the stakes <i>Do you know who swept the stakes?</i>
21	Puxar a brasa para a sua sardinha <i>Ele só trabalha quando pode puxar a brasa para a sua sardinha.</i>	To feather one's own nest <i>He works only if he can feather his own nest.</i>
22	Ser café pequeno <i>Esta tarefa é café pequeno.</i>	To be a trifle/nothing to worry about/a small beer (B) <i>This assignment is a trifle.</i>
23	Não ganhar nem para o cafezinho <i>Se você não ganha nem para o cafezinho, como é que vai comprar um carro?</i>	To earn a pittance; to make next to nothing <i>If you earn a pittance how are you going to buy a car?</i>
24	Engrossar o caldo <i>É melhor irmos embora antes que o caldo engrosse.</i>	The plot thickens <i>We'd better leave before the plot thickens.</i>
25	Entornar o caldo <i>A intervenção dele entornou o caldo: o trabalho não foi para frente.</i>	To upset the apple cart <i>His intervention upset the apple cart – the work didn't proceed.</i>
26	Não ser carne nem peixe <i>Ela não é carne nem peixe quanto a política.</i>	To be neither fish nor flesh (nor good red Hering) (B) <i>She's neither fish nor flesh as far as politics is concerned.</i>
27	Tratar a. com casca e tudo <i>Eu estava furioso com a atitude dele e tratei-o com casca e tudo.</i>	To handle sm without gloves; to ride roughshod over sm <i>I was mad with his attitude and handled him without gloves.</i>
28	Tirar uma casquinha <i>O negócio é rendoso, vou ver se tiro uma casquinha.</i>	To improve the occasion; to profit by; to have a finger in the pie <i>The business is profitable, I'll try to have a finger in the pie.</i>
29	Tirar a castanha/sardinha com a mão do gato <i>Ele recebe um bom dinheiro pelo seu trabalho, mas sempre tira a castanha com a mão do gato.</i>	To be made a cat's paw; to pull the chestnuts out of the fire <i>He gets quite a lot of money for his work, but always finds someone to pull the chestnuts out of the fire.</i>
30	Tomar chá de sumiço <i>Não conseguiram encontrá-lo em nenhum lugar, ele tinha tomado chá de sumiço.</i>	To go to earth/ground; to make oneself scarce; to disappear <i>They couldn't find him anywhere, he had gone to earth.</i>
31	Dar/Tomar chá de cadeira <i>O dentista me deu um chá de cadeira por mais de uma hora.</i>	To kick/cool one's heels; to be a wall flower [at a dance] <i>The dentist mad me kick my heels for more than one hour.</i>
32	Tomar chá de trepadeira <i>Seu filho tomou chá de trepadeira!</i>	To grow like a weed [of a child] <i>Your son is growing like a weed!</i>
33	Não ter tomado chá de pequeno <i>Ele nunca agradece porque não tomou chá de pequeno.</i>	To be ill-mannered/ill- bread /impolite; to be brought up badly. <i>He never says thank you because he was brought up badly.</i>
34 35	Ser farinha do mesmo saco=ser vinho da mesma pipa <i>Eles se entendem bem porque são farinha do mesmo saco: os dois são desonestos.</i>	To be tarred with the same brush; to be cast in the same mould/mold; to be cut of the same cloth <i>They get along well because they are both tarred with the same brush: they are both dishonest.</i>

36	Vender farinha <i>Maria, você está vendendo farinha.</i>	To be snowing down south (A) <i>Maria, it is snowing down south.</i>
37	Mardar às favas <i>O menino me aborreceu tanto que no fim mandei-o às favas.</i>	To send sm to hell; tell sm to go fly a kite/jump in the lake <i>The boy annoyed me so much that I finally sent him to hell.</i>
38	Comer frango e arrotar peru/faisão <i>Ele é um João Ninguém, mas gosta de comer frango e arrotar peru.</i>	All one's chickens to be goose ; all one's geese to be swans (B) <i>He's a nobody but all his geese are swans.</i>
39	Comer/comprar/passar/vender gato por lebre <i>Ele não é um vendedor de confiança, sempre tenta vender gato por lebre.</i>	To buy/to sell/to eat a bill of goods (A); to sell/to be sold a pup (B) <i>He's not an honest salesman, he always tries to sell a bill of goods.</i>
40	Morder a isca <i>Tentamos de tudo, mas ele não mordeu a isca.</i>	To swallow the bait ; to rise to the bait (B) <i>We tried hard but he didn't swallow the bait.</i>
41	Separar o joio do trigo <i>Nem todos são corruptos, é preciso separar o joio do trigo.</i>	To separate the wheat from the chaff <i>They are not all corrupt, one must separate the wheat from the chaff.</i>
42	Esconder o leite <i>Ele tem muita informação sobre o assunto mas está escondendo o leite.</i>	To be holding sth back; to be keeping sth from sm; to hold out on sm (B) <i>He has a lot of information about the subject but he's holding it back.</i>
43	Jogar a leite de pato <i>Não me incomodo de perder, só joga a leite de pato.</i>	To play for love/no stakes <i>I don't care if I lose, I only play for love.</i>
44	Tirar leite de pedra <i>É mais fácil tirar leite de pedra do que conseguir uma doação dele.</i>	To get blood out of/from a stone <i>It's easier to get blood out of a stone than a donation from him.</i>
45	Encher linguiça <i>O menino não tinha estudado e só encheu linguiça no exame.</i>	To drag out a story; to fill up; to waffle (B) <i>The boy hadn't studied and just filled up the exam.</i>
46	Ser manteiga derretida <i>Minha irmã é manteiga derretida, chora em qualquer filme triste.</i>	To be crybaby <i>My sister is crybaby, she cries during sad movies.</i>
47	Estar com a mão na massa <i>Deixe que eu termino já estou com a mão na massa.</i>	To be at it <i>Let me finish it since I am at it.</i>
48	Ter mão de manteiga <i>Minha empregada tem mão de manteiga, deixou cair outro copo.</i>	To be butterfingers <i>My maid is such a butterfingers, she dropped another glass.</i>
49	Catar milho [em máquina de escrever] <i>Por que você não aprende a datilografar em vez de ficar catando milho?</i>	To hunt and peck [on the typewriter] <i>Why don't you learn how to type instead of hunting and pecking?</i>
50	Cozinhar os miolos <i>Ela cozinhou os miolos para resolver o problema.</i>	To tax one's brains (A); to rack/to pick/to cudgel one's brains; to beat one's brain out (B) <i>She taxed her brains to solve the problem.</i>
51	Que nem mosca no mel = cair/ser sopa no mel <i>Aquele dinheiro extra foi a sopa no mel, agora posso comprar um carro novo.</i>	To turn out just right; to come in handy <i>That extra money came in handy, now I can buy a new car.</i>
52		
53	Ser um/comprar nabo(s) no saco <i>Comprei estes sapatos pelo reembolso postal e foi um nabo no saco.</i>	To buy a pig in a poke <i>I bought these shoes by mail order and it was a pig in a poke</i>
54	Queimar óleo quarenta <i>Ele já está queimando óleo quarenta mas certamente não parece.</i>	To be over the hill; not to cut the mustard any more <i>He's already over the hill but certainly doesn't look it.</i>
55	Estar roendo um osso <i>Coitado dele! Está roendo um osso com todos aqueles filhos para alimentar.</i>	To have a hard/a thin time <i>Poor him! He's having a thin time with all those children to feed.</i>

56	Ser/Ter um osso duro de roer <i>Fazer a revisão daquele livro foi um osso duro de roer.</i>	To be/To have a hard job/hard nut to crack <i>To revise that book was a hard nut to crack.</i>
57	Estar/Ficar de ovo virado = estar/ficar de lua <i>Não adianta falar com ele hoje, ele está de lua.</i>	To be out of sorts <i>No use talking to him today, he's out of sorts.</i>
58	Contar com o ovo na galinha <i>Você está contando com o ovo na galinha mas o negócio pode não dar certo.</i>	To count one's chickens before they hatch <i>You are counting your chickens before they hatch but the business may not come out right.</i>
59	No frigrir dos ovos <i>No frigrir dos ovos o lucro ficou com o corretor.</i>	When all is/was said and done; in the end <i>When all was said and done the profit went to the broker.</i>
60	Pisar em/sobre ovos 1. <i>Este é um assunto delicado e discuti-lo é pisar sobre ovos.</i> 2. <i>Eu estava pisando sobre ovos porque o chão estava cheio de cacos de vidro.</i>	1. to walk on eggs , to skate on thin ice 2. to pick one's steps <i>This is a delicate subject and to discuss it is to walk on eggs.</i> <i>I was picking my steps because the floor was full of fragments of glass.</i>
61	Comer o pão que o diabo amassou <i>Ela comeu o pão que o diabo amassou para poder criar aqueles filhos.</i>	To go through hell; to have a hard life; to have it rough <i>She went through hell to rear those children.</i>
62	Dizer pão-pão queijo-queijo <i>Eu sempre digo pão-pão queijo-queijo e nem todos gostam de minha franqueza.</i>	To call a spade a spade <i>I always call a spade a spade and not everybody likes my sincerity.</i>
63	Ganhar o pão nosso de cada dia <i>O Carlos está muito doente e não consegue ganhar o pão nosso de cada dia.</i>	To earn one's daily bread <i>Carlos is very ill and incapable of earning his daily bread.</i>
64	Passar a pão e água <i>Ela está passando a pão e água desde que o marido morreu.</i>	To live on the breadline (B) <i>She's living on the breadline ever since her husband ever since her husband died.</i>
65	Ser pão duro = ser unha de fome <i>Nem espere uma contribuição do Pedro, ele é unha de fome.</i>	To be tightfisted/tight as tick/ cheese paring/a tightwad/a skinflint/a scrooge; to scrimp and scrape/save <i>Don't expect a contribution from Pedro, he's tightwad.</i>
66	Ser o pão nosso de cada dia <i>Não gosto do meu emprego, mas é o pão nosso de cada dia.</i>	To be our daily bread <i>I don't like my job but it's our daily bread.</i>
67	Tirar o pão da boca de a. <i>Ele abriu uma loja ao lado da minha e me tirou o pão da boca.</i>	To take the bread out of sm's mouth <i>He opened a shop next door to mine and took the bread out of my mouth.</i>
68	Não ter papas na língua <i>Elza não tem papas na língua e disse tudo o que ele precisava ouvir.</i>	To make no bones about sth; not to mince one's words; not to put too fine a point on it <i>Elza didn't mince her words and told him everything he should hear.</i>
69	Chorar as pitangas <i>Ela vire chorando as pitangas mas não faz nada para melhorar a situação.</i>	To spill one's guts <i>She spills her guts but does nothing to improve her situation.</i>
70	A preço de banana <i>Ele comprou aquela casa a preço de banana.</i>	For a song/a shoestring <i>He bought that house for a song.</i>
71	Virar presunto <i>O bandido se entregou porque não queria virar presunto.</i>	To croak; to become a corpse <i>The gangster surrendered because he didn't want to become a corpse.</i>
72	Ter comido queijo <i>Ele nunca sabe onde estão as coisas porque comeu muito queijo.</i>	To have a head/a mind/a memory like a sieve; to be forgetful <i>He never knows where things are because he has a head like a sieve.</i>
73	Entregar a rapadura = entregar os pontos <i>Antônio é muito persistente e não vai entregar os pontos tão facilmente.</i>	To throw in/up the sponge; to give in; to hand down/lower one's colours <i>Antonio is very persevering, he won't throw up the sponge so easily.</i>

74	Dar (um) rolo/ bolo <i>Deu rolo na reunião.</i>	There to be/to result in a row/browl/riot <i>There was a riot during the meeting.</i>
75	Por sal na moleira <i>É um negócio arriscado, é melhor você por sal na moleira.</i>	To pull in one's neck <i>It's a risky business, you'd better pull in your neck.</i>
76	Do tempo em que se amarrava cachorro com linguiça = no/do tempo do onça <i>Este meu vestido é do tempo do onça.</i>	(in) the good old days; a long time ago; (from) ages ago; (from) before the flood <i>This dress of mine is from ages ago.</i>
77	Comer à tripa forra <i>Não almocei, mas no jantar comi à tripa forra.</i>	To eat one's heart content; to fill one's stomach; to make a pig of oneself (B) <i>I didn't have any lunch but at dinner I ate to my heart's content.</i>
78	Viver à tripa forra <i>Com o dinheiro herdado ela vai poder viver à tripa forra.</i>	To live off the fat of the land <i>With the money she inherited she can live off the fat of the land.</i>
79	Não ter tutano = não ter fibra <i>Ele não vai se desincumbir desta tarefa, ele não tem fibra.</i>	To have no backbone <i>He won't accomplish this task, he has no backbone.</i>
80	Ser mais as vozes que as nozes <i>Nossas reuniões são mais as vozes que as nozes.</i>	To be more shadow than substance/all talk and no cider ; more cry than wool <i>Our meetings are all talk and no cider.</i>

Fonte: Camargo; Steinberg (1989).

(A) Inglês Americano

(B) Inglês Britânico

ANEXO B – EIs com alimentos em inglês e suas respectivas traduções em português

01	To apple polish <i>There goes John apple polishing his teacher again.</i>	Puxar o saco de a. <i>Lá está o John puxando o saco da professora de novo.</i>
02	To upset the/sm's apple -cart <i>All was going well, but suddenly someone upset the apple-cart.</i>	Jogar areia; atrapalhar os planos de a.; entornar o caldo <i>Tudo ia indo muito bem, mas de repente alguém entornou o caldo.</i>
03 04	As sure as God made little apples (B) = as sure as two times two makes four = As sure as eggs are eggs <i>You may be sure as two times two makes four that he will already be there.</i>	Certo como dois e dois são quatro <i>É certo como dois e dois são quatro que ele já estará lá.</i>
05	To have no backbone <i>He will never accomplish anything, He has no backbone.</i>	Não ter fibra/ tutano <i>Ele nunca irá realizar nada; ele não tem fibra.</i>
06	Not to have a bean <i>He pretends to be rich but actually he does not have a bean.</i>	Não ter um tostão furado/um gato morto para puxar pelo rabo; estar sem dinheiro/duro <i>Ele finge que é rico, mas na verdade não tem um tostão furado.</i>
07	Not to know beans about sth <i>The new assistant does not know beans about the job.</i>	Não saber nada/bulhufas <i>O novo assistente não sabe bulhufas do seu trabalho.</i>
08	To give sm beans <i>Grandma was furious with the children and gave them beans.</i>	Passar um sabão em a.; ralhar com a.; castigar a. <i>A vovó estava furiosa com as crianças e lhes passou um sabão.</i>
09	To spill the beans <i>We meant it to be a secret, but someone spilled the beans.</i>	Dar com a língua nos dentes <i>Nós queríamos que fosse um segredo, mas alguém deu com a língua nos dentes.</i>
10 11 12	To take the biscuit (B) = To take the bun = To take the cake (A) <i>With this remark, you take the biscuit.</i>	Levar a palma; ultrapassar/sobrepular tudo; ser inacreditável <i>Com esta observação, você leva a palma.</i>
13	Two bites at the cherry <i>It's too small a task for two bites at the cherry.</i>	Duas tentativas para alcançar a mesma meta; fazer a. c. em duas etapas <i>É um trabalho pequeno demais para ser feito em duas etapas.</i>
14	To make no bones about sth <i>He made no bones about belonging to the Communist Party.</i>	Não ter papas na língua; não esconder; não hesitar; não temer as consequências de a. c. <i>Ele não escondia o fato de pertencer ao Partido Comunista.</i>
15	To live on bread and cheese and kisses <i>She eats so little, she must live on bread and cheese and kisses.</i>	Viver de brisa <i>Ela come tão pouco, deve viver de brisa.</i>
16	To know which side one's bread is <u>buttered</u> <i>Johnny is so smart, he always knows which side his bread is buttered.</i>	Saber qual é o lado mais vantajoso; ser oportunista <i>Johnny é esperto, ele sempre sabe qual é o lado mais vantajoso.</i>
17	To take the bread of sm's mouth <i>When they opened the supermarket they took the bread out of the grocer's mouth.</i>	Tirar o pão da boca de a. <i>Quando abriram o super-mercado, eles tiraram o pão da boca do dono da mercearia.</i>
18	To cast/To throw one's bread upon the waters <i>She is a good soul, always casting bread upon the waters.</i>	Fazer uma boa ação desinteressadamente <i>Ela é uma boa alma, sempre fazendo uma boa ação desinteressadamente.</i>
19	To live on the breadline (B) <i>She lost her job and now is living on the breadline.</i>	Passar a pão e água <i>Ela perdeu o emprego e está agora passando a pão e água.</i>

20 21	to be with a bun in the oven = to be in the family way = To be in the pudding club <i>My daughter-in-law just called to say that she is in the family way.</i>	Estar esperando nenê/grávida <i>Minha nora acaba de me telefonar dizendo que está esperando nenê.</i>
22	To look as if butter will/would not melt in sm's mouth <i>At first we all believed him, he looked as if butter wouldn't melt in his mouth.</i>	Com aparência de inocente; não ser flor que se cheire; ser santo do pau oco <i>No começo acreditávamos nele, ele tinha aparência de inocente.</i>
23	To be butterfingers <i>I would not trust her with that crystal vase, because she is butterfingers.</i>	Ter mão de manteiga <i>Eu não lhe confiaria aquele vaso de cristal, ela tem mão de manteiga.</i>
24	To boil one's cabbage twice <i>I already heard you, please do not boil your cabbage twice!</i>	Ser relógio de repetição; cantar a mesma ladainha/cantilena <i>Eu já escutei, por favor não cante a mesma ladainha de novo.</i>
25	To be a cake -eater <i>If Irene did not go to the party she must be ill because she is a cake-eater.</i>	Ser arroz de festa <i>Se Irene não foi à festa ela deve estar doente porque ela é arroz de festa.</i>
26	Cakes and <u>ale</u> <i>Our gathering was all cakes and ale.</i>	Folia; divertimento; divertidos <i>Nosso encontro foi muito divertido.</i>
27	To sell like hot cakes <i>We hope our books will sell like hot cakes.</i>	Vender como água /rapidamente; ter boa saída <i>Esperamos que nosso livro venda como água.</i>
28	To be caviare to the general <i>Lodging them in this luxurious mansion would be caviare to the general.</i>	Jogar/Lançar pérolas aos porcos; ser bom demais para [a. que não sabe apreciar] <i>Alojá-los naquela mansão de luxo seria lançar pérolas aos porcos.</i>
29	To pull the chestnuts out of the fire <i>He is very successful, but he always finds someone to pull the chestnuts out of the fire for him.</i>	Tirar castanha com a mão do gato <i>Ele tem muito sucesso, mas ele sempre acha alguém para tirar a castanha com a mão do gato.</i>
30	To count one's chickens before they hatch <i>You're counting your chickens before they hatch, but do not forget that your aunt may not send you the money.</i>	Contar com o ovo na <u>galinha</u> <i>Você está contando com o ovo na galinha, mas não se esqueça que sua tia pode não mandar o dinheiro.</i>
31	To flap one's chops (A) <i>They have nothing to do, they can afford flapping their chops idly all day long.</i>	Bater (um) papo/(uma) caixa <i>Eles não têm o que fazer e ficam batendo caixa o dia inteiro.</i>
32	To lick one's chops = to lick/to smack one's lips <i>It smells so good, I am already smacking my lips.</i>	Lamber os beiços; demonstrar satisfação antecipada <i>O cheiro está tão bom, já estou lambendo os beiços.</i>
33	To be/To live in clover <i>No wonder he is living in clover now, with all that money he inherited.</i>	Levar uma vida de rei; viver no bem bom; viver em luxo e conforto <i>Não é à toa que ele está levando uma vida de rei, com a herança que recebeu.</i>
34	To warm the cockle(s) of one's heart <i>His words warmed the cockles of my heart.</i>	Tocar/Emocionar profundamente <i>Suas palavras me emocionaram profundamente.</i>
35	To sit upon hot cockles = to be on pins and needles <i>You'd better go see a doctor if you still feel pins and needles in your arm.</i>	Estar com formigamento em; estar sobre brasas/impaciente <i>É melhor você consultar um médico se você ainda está com formigamento no braço.</i>
36	To acknowledge the corn <i>With all that evidence he had to acknowledge the corn.</i>	Admitir seu próprio erro; dar a mão à palmatória <i>Com todas aquelas provas ele teve que admitir seu erro.</i>
37	To treat on sm's corns <i>I didn't mean to tread on his corns when I criticized his sister.</i>	Pisar nos calos de a.; magoar <i>Eu não quis magoá-lo quando critiquei sua irmã.</i>
38	To chew the cud <i>It was gratifying to sit there and chew the cud of our childhood.</i>	Fazer a hora da saudade; ruminar; refletir; ponderar

		<i>Foi gratificante ficar sentado ali fazendo a hora da saudade de nossa infância.</i>
39	Not to be one's cup of tea <i>Modern poetry is not my cup of tea.</i>	Não ser o fraco/da predileção de a. <i>Poesia moderna não é o meu fraco.</i>
40	To marry into dough = to marry money <i>He didn't love her, his intention was to marry money.</i>	Dar o golpe do baú <i>Ele não gostava dela, sua intenção era dar o golpe do baú.</i>
41	To have egg on one's face = to hang one's head <i>When I realized my mistake I hung my head.</i>	Ficar com a cara no chão; murchar a orelha <i>Fiquei com a cara no chão quando me dei conta de meu engano.</i>
42	To be in the egg <i>There is nothing to say about the research yet, it is still in the egg.</i>	Estar em estado embrionário/no início <i>Ainda não há nada para dizer da pesquisa, ela ainda está no início.</i>
43	To put all one's eggs in one basket <i>My grandfather was left penniless, for he put all his eggs in one basket, and when the bank closed he lost everything.</i>	Apostar num cavalo só, arriscar tudo num só negócio <i>Meu avô ficou sem um vintém, porque ele apostou num cavalo só, e quando o banco fechou ele perdeu tudo.</i>
44	To have eggs on the spit (A) <i>I cannot talk to you now, I have eggs on the spit.</i>	Não ter tempo nem para se coçar <i>Não posso falar com você agora, não tenho tempo nem para me coçar.</i>
45	To bring one's eggs to a bad/the wrong market <i>She worked so hard, but brought eggs to the wrong market.</i>	Dar com os burros n'água; bater em porta errada <i>Ela trabalhou tanto, mas deu com os burros n'água.</i>
46	To walk on eggs <i>To discuss this subject with him is to walk on eggs.</i>	Pisar sobre ovos <i>Discutir esse assunto com ele é pisar sobre ovos.</i>
47	The fat is/will be in the fire <i>He exchanged harsh words with his neighbour, the fat is in the fire.</i>	Vai dar confusão/bode; as coisas estão mal paradas <i>Ele bateu boca com o vizinho, vai dar bode!</i>
48	To live off the fat of the land <i>John is doing very well, he's living off the fat of the land.</i>	Viver à tripa forra; ter do bom e do melhor <i>John está muito bem, vivendo à tripa forra.</i>
49	To chew the fat <i>Stop chewing the fat! There's nothing we can do about it.</i> <i>He spends his evenings chewing the fat with his friends.</i>	Falar entre os dentes; ficar resmungando; reclamar (B); bater papo (A) <i>Pare de resmungar, não há nada que possamos fazer!</i> <i>Ele passa as tardes batendo papo com os amigos.</i>
50	To run to fat <i>Some people run to fat after a certain age.</i>	Ter tendência para engordar <i>Algumas pessoas têm tendência para engordar depois de uma certa idade.</i>
51	To be in a ferment <i>Now that the elections are close again, the whole country seems to be in a ferment.</i>	Encontrar-se num estado de efervescência/agitação/tumulto <i>Agora que as eleições estão aí novamente, o país inteiro encontra-se num estado de efervescência.</i>
52	To have a finger in the pie That project sounds interesting; I would like to have finger in the pie.	Estar envolvido em a. c.; participar de tarefa/projeto; ter a. c. a ver com a. c.; tirar uma casquinha <i>Esse projeto parece interessante, eu gostaria de participar dele.</i>
53	Neither fish , flesh , nor good red herring <i>I couldn't make out her point of view, it was neither fish, flesh nor good red herring.</i>	Não ser carne nem peixe /nem uma coisa, nem outra; ser de natureza ambígua/indefinida <i>Não consegui entender o ponto de vista dela, não era ma coisa nem outra.</i>
54	All one's geese are swans <i>Her car is so old, and her house is falling to pieces, but all her geese are swans.</i>	Comer frango e arrotar peru/faisão ; sobrestimar; exagerar tudo.

		<i>O carro dela está velho, a casa está caindo aos pedaços, mas ela come frango e arrota faisão.</i>
55	To take the gilt off the gingerbread <i>With that nasty remark he took the gilt off the gingerbread.</i>	Empanar o brilho de a. c.; estragar o prazer/tudo <i>Com aquele comentário maldoso ele estragou tudo.</i>
56	To kill the goose that lays the golden eggs <i>Trying to make his grandfather sign that will kill the goose that lays the golden eggs, for the old man disinherited him.</i>	Matar a galinha dos ovos de ouro; acabar com a fonte de riqueza <i>Ao tentar fazer o avô assinar aquele testamento ele matou a galinha dos ovos de ouro pois o avô o deserdou.</i>
57	To cook sm's goose <i>I was hoping to get the position but the manager's recently graduated brother cooked my goose.</i>	Acabar com as esperanças de a.; estragar os planos de a. <i>Eu estava com a esperança de obter o cargo, mas o irmão recém-graduado do gerente acabou com as minhas esperanças.</i>
58	To be in a jam <i>I cannot help you now, I am myself in a jam.</i>	Ver-se/Estar num aperto/em apuros/numa pior <i>Não posso ajudá-lo agora, eu mesmo me encontro num aperto.</i>
59	To get into a jam <i>I accepted his partnership and got into a jam.</i>	Meter-se numa fria/em apuros <i>Eu aceitei a sua sociedade e agora me meti numa fria.</i>
60	To get out of a jam <i>She created the problem and it is for me to get out of the jam.</i>	Descascar o abacaxi <i>Ela criou o problema e eu é que tenho que descascar o abacaxi.</i>
61	To stew in one's own juice <i>She insisted on marrying that good-for-nothing, and now she can stew in her own juice.</i>	Deitar na cama que fez; sofrer as consequências dos próprios atos <i>Ela insistiu em se casar com aquele sujeito que não vale nada e agora está deitada na cama que fez.</i>
62	A pretty kettle of fish <i>No one knew what was going on, it was a pretty kettle of fish.</i>	Um angu de caroço <i>Ninguém entendia o que estava acontecendo, era um angu de caroço.</i>
63	Milk and water <i>I can't stand his books, they are all milk and water.</i>	Linguagem insípida/sem graça/sem sal <i>Eu não suporto os livros dele, são todos sem graça.</i>
64	To cry over spilled milk <i>Crying over spilled milk will not correct your mistake.</i>	Chorar sobre leite derramado/a morte da bezerra <i>Chorar sobre leite derramado não vai corrigir o seu erro.</i>
65	To imbibe sth with one's mother's milk <i>He imbibed his taste for music with his mother's milk.</i>	Beber a. c. com o leite materno <i>Ele bebeu com o leite materno o seu gosto pela música.</i>
66	To make mincemeat of <i>The clever prosecuting attorney made mincemeat of the false witness.</i>	Acabar com a raça de a.; triturar a. (numa discussão) <i>O esperto promotor acabou com a raça da falsa testemunha.</i>
67	To go back to the salt mine <i>My vacations are over, and now I must go back to the salt mine.</i>	Pegar no batente de novo <i>Minhas férias acabaram, agora preciso pegar no batente de novo.</i>
68	(not) to cut the mustard <i>We'll have to fire him, he does not cut the mustard.</i>	(não) dar conta do recado; (não) dar no couro <i>Vamos ter que despedi-lo, ele não dá conta do recado.</i>
69	To use one's noodle = to use ones wits <i>Use your wits and you will find the answer.</i>	Usar a cabeça <i>Use a cabeça e você encontrará a resposta.</i>
70	A hard nut to crack <i>The old man is a hard nut to crack, but his grand-daughter will Bend him.</i>	Um osso duro de roer; uma palavra dura; duro na queda <i>O velho é um osso duro de roer, mas a sua neta irá dobrá-lo.</i>
71	To be nuts about/over sm/sth <i>She is so fat because she is nuts about ice-cream.</i>	Estar/Ser louco por a./a. c. <i>Ela é gorda desse jeito porque é louca por sorvete.</i>

72	To feel one's oats <i>He is almost seventy, but still feels his oats.</i>	Estar radiante/cheio de vida; agir com audácia <i>Ele tem quase setenta anos, mas ainda está cheio de vida.</i>
73	To sow one' wild oats <i>Nowadays most youngsters sow their wild oats, but settle down when they get married.</i>	Gozar a mocidade; fazer as estrepolias da mocidade <i>Hoje em dia os jovens gozam a mocidade mas sossegam quando se casam.</i>
74	To pour oil on troubled waters <i>My sons were having an endless quarrel, but I managed to pour oil on troubled waters.</i>	Por água na fervura; acabar com uma briga <i>Meus filhos estavam brigando sem parar, mas consegui por água na fervura.</i>
75	To know one's onions <i>I trust him entirely, he knows his onions.</i>	Entender do riscado; ser competente <i>Confio nele inteiramente, ele entende do riscado.</i>
76	(in) apple-pie order <i>He just moved in, and his books are already in apple-pie order.</i>	(em) ordem perfeita <i>Ele acaba de se mudar e seus livros já estão na mais perfeita ordem.</i>
77	To be/to have a pea in the shoe <i>Linda won't listen to reason, she really is a pea in her mother's shoe.</i>	Ser uma/Estar com uma pedra no sapato <i>Linda se recusa a ouvir a voz da razão, ela é realmente uma pedra no sapato da mãe.</i>
78	To eat humble pie <i>He behaved arrogantly, but when I reminded him of his origins, he ate humble pie.</i>	Por o rabo entre as pernas/a viola no saco <i>Ele se comportou de maneira arrogante, mas quando o lembrei de suas origens ele os o rabo entre as pernas.</i>
79	To gild/To sugar /To sweeten the pill <i>No use trying to sweeten the pill, it won't make matters any better.</i>	Dourar/Adoçar a pílula; tornar a. c. menos desagradável. <i>Não adianta tentar adoçar a pílula, não vai melhorar as coisas.</i>
80	To take (statement etc.) with a pinch of salt = to take (statement etc.) with a grain of salt <i>There was not sufficient evidence, so I took the news about his return with a grain of salt.</i>	Aceitar [uma afirmação etc.] com uma ponta de dúvida/desconfiança/reserva; dar um desconto pelo exagero <i>Não havia provas suficientes, de modo que aceitei a notícia de sua volta com uma ponta de dúvida.</i>
81	to handle a hot potato <i>I have a hot potato to handle: I must negotiate with the strikers and please the boss.</i>	Descascar um abacaxi ; ter/estar com um abacaxi nas mãos <i>Estou com um abacaxi nas mãos; preciso negociar com os grevistas e agradar o patrão.</i>
82	To have nothing to do with the price of tea in China <i>Don't try to involve me in that affair, I have nothing to do with the price of tea in China.</i>	Não ter nada a ver com peixe <i>Não tente me envolver nessa história, não tenho nada a ver com o peixe.</i>
83	To be a dripping roast <i>The investments his father left him are a dripping roast of income.</i>	Ser uma fonte inesgotável (de renda) <i>Os investimentos que o pai lhe deixou são uma fonte inesgotável de renda.</i>
84	To have a rod in pickle for sm <i>The boy did not want to go back home; he knew his father had a rod in pickle for him.</i>	Ter um castigo preparado para a. <i>O menino não queria ir para casa; sabia que seu pai tinha um castigo para ele.</i>
85	To salt sth away <i>He lives so modestly because he is salting his money away.</i>	Fazer um pé de meia; poupar/guardar dinheiro para o futuro <i>Ele vive tão modestamente porque está fazendo um pé de meia.</i>
86	Not/Hardly to be worth one's salt <i>I keep my maid because she is trustworthy, but actually she is not worth her salt.</i>	Que não vale o que come; de pouco/nenhum valor; incompetente <i>Eu fico com a minha empregada porque ela é de confiança, mas na realidade ela não vale o que come.</i>
87	To be in the soup <i>It's because he acted so foolishly that he is now in the soup.</i>	Estar em maus lençóis; em apuros; em dificuldades; numa enrascada <i>É porque ele agiu de maneira tão tola que ele está agora em maus lençóis.</i>

88 89	To be in/to get into a stew <i>He lost all his documents and now he is in a stew.</i>	Estar/Ficar em maus lençóis/no mato sem cachorro <i>Ele perdeu todos os documentos e agora está no mato sem cachorro.</i>
90	To pour/To throw cold water <i>I was planning a trip to the Far East, but the high prices threw cold water on my projects.</i>	Jogar água fria em; desencorajar; desestimular <i>Eu estava planejando uma viagem ao Extremo Oriente, mas os altos preços jogaram água fria nos meus projetos.</i>
91	To be in/to get into hot water <i>Because of our solicitor's foolish advice we got into hot water.</i>	Estar/meter-se numa fria/enrascada/em dificuldades [por causa de uma bobagem cometida] <i>Estamos em dificuldades por causa do conselho insensato de nosso advogado.</i>
92	To separate the wheat from the chaff <i>You'd better separate the wheat from the chaff before casting judgement on his words.</i>	Separar o joio do trigo <i>É melhor você separar o joio do trigo antes de emitir um julgamento sobre as palavras dele.</i>

Fonte: Camargo; Steinberg (1990).

(A) Inglês Americano

(B) Inglês Britânico