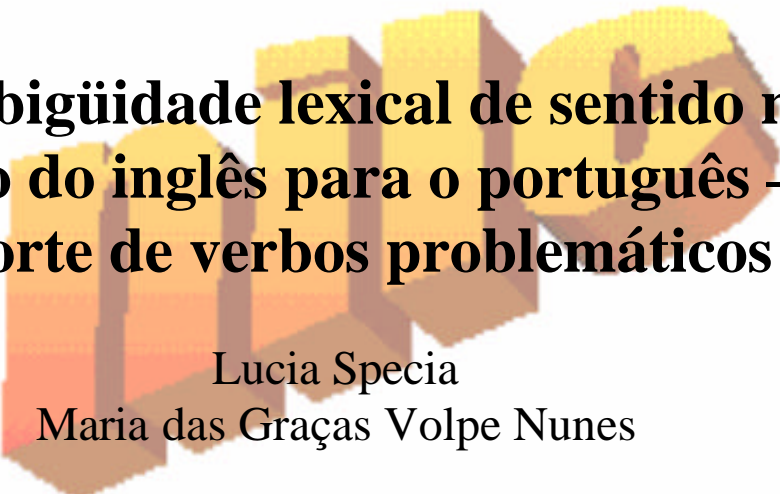


Universidade de São Paulo - USP  
Universidade Federal de São Carlos - UFSCar  
Universidade Estadual Paulista - UNESP



**A ambigüidade lexical de sentido na  
tradução do inglês para o português – um  
recorte de verbos problemáticos**

Lucia Specia  
Maria das Graças Volpe Nunes

**NILC-TR-04-01**

Março, 2004

Série de Relatórios do Núcleo Interinstitucional de Lingüística Computacional  
NILC - ICMC-USP, Caixa Postal 668, 13560-970 São Carlos, SP, Brasil

## **Resumo**

Este relatório apresenta um estudo sobre o problema da ambigüidade lexical de sentido na tradução automática de verbos do inglês para o português, realizado com base em sentenças do cópús BNC (*British National Corpus*), visando selecionar um subconjunto de verbos freqüentes nesse cópús, que apresentam ambigüidade na tradução e para os quais a ambigüidade não recebe tratamento adequado por parte dos sistemas de tradução comumente utilizados atualmente.

# Índice

1	Introdução .....	1
2	A ambigüidade lexical e a tarefa de DLS .....	2
3	O cópús BNC.....	3
3.1	Composição do cópús.....	4
3.2	Codificação dos textos .....	5
3.3	Etiquetagem gramatical.....	6
4	Análise do cópús BNC .....	7
4.1	Seleção dos verbos.....	7
4.2	Seleção das sentenças .....	9
4.3	Escolha dos sistemas de tradução e realização da tradução .....	10
4.4	Análise das traduções .....	12
4.5	Compilação e resultados da análise .....	18
5	Considerações Finais .....	29
	Agradecimentos .....	29
	Referências Bibliográficas .....	30

## Tabelas

Tabela 1	– Possíveis etiquetas para verbos do <i>tagset C7</i> .....	6
Tabela 2	– Exemplos de sentenças selecionadas para alguns verbos .....	10
Tabela 3	– Exemplos de traduções das sentenças da Tabela 2 .....	11
Tabela 4	– Análise das traduções das sentenças do verbo <i>to get</i> .....	12
Tabela 5	– Sentidos dos verbos nas sentenças selecionadas da Tabela 2 .....	17
Tabela 6	– Compilação da análise da ambigüidade dos verbos.....	19
Tabela 7	– Detalhamento da análise do verbo <i>to be</i> .....	19
Tabela 8	– Detalhamento da análise do verbo <i>to have</i> .....	20
Tabela 9	– Detalhamento da análise do verbo <i>to do</i> .....	20
Tabela 10	– Detalhamento da análise do verbo <i>to say</i> .....	20
Tabela 11	– Detalhamento da análise do verbo <i>to go</i> .....	20
Tabela 12	– Exemplo de problema com o uso do verbo <i>to go</i> no futuro imediato.....	21
Tabela 13	– Exemplos de problemas com o uso do verbo <i>to go</i> em <i>phrasal verbs</i> ou expressões .....	21
Tabela 14	– Detalhamento da análise do verbo <i>to get</i> .....	22
Tabela 15	– Detalhamento da análise do verbo <i>to make</i> .....	22
Tabela 16	– Exemplos de problemas com o uso do verbo <i>to make</i> .....	23
Tabela 17	– Detalhamento da análise do verbo <i>to see</i> .....	23
Tabela 18	– Detalhamento da análise do verbo <i>to know</i> .....	23
Tabela 19	– Exemplos de problemas com o uso do verbo <i>to know</i> .....	23
Tabela 20	– Detalhamento da análise do verbo <i>to take</i> .....	24
Tabela 21	– Exemplos de problemas com o uso do verbo <i>to take</i> .....	24
Tabela 22	– Detalhamento da análise do verbo <i>to think</i> .....	25
Tabela 23	– Exemplos de problemas com o uso do verbo <i>to think</i> .....	25
Tabela 24	– Detalhamento da análise do verbo <i>to come</i> .....	25
Tabela 25	– Exemplo de problema com o uso do verbo <i>to come</i> seguido de preposição .....	26
Tabela 26	– Exemplos de problemas com o uso do verbo <i>to come</i> .....	26
Tabela 27	– Detalhamento da análise do verbo <i>to give</i> .....	27
Tabela 28	– Exemplos de problemas com o uso do verbo <i>to give</i> .....	27
Tabela 29	– Detalhamento da análise do verbo <i>to look</i> .....	27

Tabela 30 – Exemplos de problemas com o uso do verbo <i>to look</i> .....	28
Tabela 31 – Detalhamento da análise do verbo <i>to use</i> .....	28
Tabela 32 – Exemplos de problemas com o uso do verbo <i>to use</i> .....	28

# 1 Introdução

No cenário de globalização atual, a disseminação cada vez maior de informações multilíngües, principalmente em meio eletrônico, como a *web*, evidencia a necessidade de traduções rápidas, eficientes e baratas para facilitar a comunicação e o compartilhamento de informações. Nesse contexto, é crescente o interesse por sistemas de Tradução Automática (TA), ou seja, por sistemas que permitam a tradução por computador de textos de uma língua natural (LF – Língua Fonte) para outra (LA – Língua Alvo).

Muito embora seja uma das áreas mais antigas do Processamento das Línguas Naturais (PLN), a TA ainda apresenta muitos problemas. Grande parte desses problemas está relacionada às diferenças lexicais entre as línguas, ou seja, aos diferentes modos pelos quais as línguas caracterizam ou classificam o mundo, quais conceitos elas escolhem para expressar pelas palavras, e quais elas decidem não utilizar.

Em função dessas diferentes classificações do mundo, uma palavra na LF pode ter vários usos e realizações na LA, dando origem ao problema da ambigüidade na seleção lexical, ou seja, na escolha pela palavra mais adequada, na LA, para traduzir uma palavra da LF, quando a LA oferece mais de uma opção para a tradução. Os casos mais graves desse problema ocorrem quando são identificadas apenas variações de significado (ou sentido) nas opções de tradução na LF, ou seja, quando essas opções são todas da mesma categoria gramatical. Esse tipo de problema é chamado de **ambigüidade lexical de sentido** e a área que se ocupa do seu tratamento, de **Desambiguação do Sentido das Palavras** (DLS), do inglês *Word Sense Disambiguation*.

Além da TA, os mecanismos de DLS se aplicam a várias outras áreas, como a recuperação de informações, o processamento de textos e de fala, a análise de conteúdo, etc. Na TA, em especial, a ambigüidade de sentido na seleção lexical é, reconhecidamente, um problema de difícil solução, que caracteriza sempre uma escolha imprescindível e cujos efeitos podem ser extremamente prejudiciais à tradução. Em se tratando da TA para o Português do Brasil como LA, particularmente, a falta de mecanismos efetivos de DLS é certamente um dos principais motivos para os resultados insatisfatórios dos sistemas existentes<sup>1</sup>.

Existem diferentes abordagens para a DLS na TA. Geralmente, essas abordagens delimitam o problema para um recorte da língua, com base em suas características sintáticas e/ou semânticas comuns, ou com base no uso das palavras, por exemplo, a partir de análises de frequência e relevância em um *corpus*. Essa delimitação, empregada neste trabalho, permite a construção incremental de modelos de DLS e é essencial em abordagens que utilizam conhecimento lingüístico e/ou extralingüístico especificado parcial ou totalmente de forma manual, uma vez que esse processo demanda tempo e conhecimento altamente especializado.

O estudo apresentado neste relatório está incorporado a um trabalho mais amplo, que visa à criação de um modelo de DLS a ser utilizado em um sistema de TA do inglês para o português do Brasil. O objetivo geral desse estudo é, com base na língua em uso (em um *corpus*), analisar o problema da ambigüidade lexical de sentido na tradução do inglês para o português, considerando-se, inicialmente, apenas a tradução de alguns verbos. Como objetivos específicos, pretende-se: (1) identificar um conjunto de verbos mais problemáticos no que se refere à ambigüidade lexical de sentido na TA, levando em conta a sua frequência em um *corpus*, a quantidade de acepções possíveis para a sua tradução em uma amostra das sentenças desse *corpus* e a inexistência de um tratamento adequado para esse fenômeno por parte dos tradutores disponíveis atualmente. (2) obter subsídios para analisar, posteriormente, a influência de características morfológicas, sintáticas, semânticas e/ou pragmáticas dos verbos (e das sentenças) para a sua tradução adequada, bem como para a tradução adequada da sentença completa; e (3) obter informações que possam ser usadas para identificar, também posteriormente, regularidades na tradução desse conjunto de verbos que permitam tratar a ambigüidade lexical de sentido de maneira sistemática. Para tanto, foram selecionadas sentenças do inglês contendo os verbos mais frequentes do *corpus* BNC (*British National Corpus*) (Burnard, 2000) e analisadas as suas traduções para o português em três

---

<sup>1</sup> Oliveira Jr. et al. (2000) relatam uma análise sobre a qualidade dos sistemas de TA para o português, considerando diversos aspectos da tradução.

sistemas de tradução disponíveis atualmente: Systran, FreeTranslation e Globalink L&H Power Translator® Pro.

Visando contextualizar esse estudo, o problema da ambigüidade lexical de sentido, bem como a tarefa de DLS, são descritos com mais detalhes na Seção 2. O corpus BNC, com base no qual o estudo é realizado, é apresentado na Seção 3. Na Seção 4 é descrita a análise das sentenças do corpus BNC e das suas traduções, propriamente dita, desde a seleção dos verbos e das sentenças até a análise das suas traduções e a compilação dos resultados. Na Seção 5 são discutidas algumas conclusões deste trabalho.

## 2 A ambigüidade lexical e a tarefa de DLS

Os problemas da tradução se devem, em grande parte, às diferenças lexicais existentes entre as línguas. Dentre os aspectos lexicais nos quais as línguas divergem estão os diferentes níveis de granularidade das distinções que elas fazem e também a base para essas distinções. Por exemplo, o francês possui três distinções para um termo único do inglês, *leg* (perna), quais sejam: *jambe* (perna de um humano), *patte* (perna de um animal) e *pied* (perna de uma mesa). Com relação à base das distinções, um exemplo é que o inglês escolhe diferentes verbos para a ação “vestir” (*putting on*) e o estado “usar” (*wearing*), enquanto o japonês escolhe os diferentes verbos de acordo com o objeto que está sendo vestido (sapatos, óculos, chapéu, blusa, etc.).

Em função dessas diferentes classificações do mundo, uma palavra na LF pode ter um uso e uma realização muito diferentes que sua equivalente na LA, ocasionando, entre outros problemas, a ambigüidade lexical, isto é, a ambigüidade na seleção lexical da palavra da LA para traduzir uma palavra da LF. Essa ambigüidade é causada, fundamentalmente, pela existência de algumas relações semânticas interlexicais entre as línguas, principalmente a polissemia, a homonímia e a sinonímia.

A **polissemia** diz respeito à diversificação do sentido das palavras, que faz com que a uma palavra possam corresponder diversos equivalentes, todos de significado relacionado, conforme o contexto. Na tradução, a polissemia faz com que a uma palavra da LF correspondam duas ou mais palavras da LA, com diferentes significados relacionados entre si. Por exemplo, à palavra do inglês *know* podem corresponder duas palavras relacionadas no português, “saber” e “conhecer”. Da mesma forma, à palavra do francês *simple* podem corresponder as palavras “simples” e “singelo”. A diferença entre os sentidos, embora seja, muitas vezes sutil (como no caso de *simple*), se ignorada na tradução pode fazer com que somente a forma mais comum seja utilizada, implicando uma simplificação da língua.

Na relação de **homonímia**, a uma mesma palavra correspondem diversos significados não relacionados entre si, por exemplo, a palavra “banco”, que pode significar “instituição financeira” ou “assento”. A homonímia entre duas línguas faz com que a uma palavra da LF correspondam duas ou mais palavras da LA, com diferentes significados não relacionados entre si. Por exemplo, a palavra inglesa *light*, que pode ser traduzida como “leve” ou “luz”, no português.

A **sinonímia** ocorre quando da existência de palavras de sentido idêntico, ou quase idêntico. Os sinônimos idênticos, chamados de sinônimos perfeitos, não representam problemas para a tradução, uma vez que qualquer um dos sinônimos pode ser escolhido como correspondente para uma dada palavra. No entanto, casos de sinônimos perfeitos são muito raros. Geralmente, os significados dos sinônimos diferem entre si, ainda que sutilmente e, portanto, eles não são permutáveis em todos os enunciados possíveis. Por exemplo, “pai” e “papai” são considerados sinônimos, mas não é comum dizer “João é papai de três filhos”. Na tradução, esse problema ocorre quando a uma palavra da LF podem corresponder vários sinônimos da LA. Por exemplo, para traduzir *dog*, do inglês, é preciso escolher entre “cachorro” e “cão”. Decisões como essa podem envolver, além das diferenças de significado, questões estilísticas. Esse tipo de relação não será considerado neste trabalho.

Vale ressaltar que a diferença entre esses três tipos de relação, principalmente entre a polissemia e a homonímia, que são os casos mais graves de ambigüidade e serão contempladas neste trabalho, não é, muitas vezes, clara. Somente parte dos autores considera como diferencial o fato de os significados serem ou não relacionados. Mesmo aqueles que consideram esse critério de diferenciação afirmam que há uma grande dificuldade em se identificar quais significados são relacionados e quais não são. Por exemplo,

determinados autores consideram a etimologia das acepções para verificar a existência do relacionamento, assumindo que, se a origem das acepções é a mesma, elas são relacionadas. Contudo, nem sempre se pode verificar a verdadeira origem das palavras e, além disso, muitas palavras tiveram seu significado modificado no decorrer do tempo, com o seu uso. Por essas razões, neste trabalho, a diferenciação entre tais relações não é relevante, pretende-se analisar ambos os fenômenos indistintamente.

O problema da ambigüidade lexical na tradução pode, ainda, ser classificado como **ambigüidade categorial** ou **ambigüidade de sentido** (Jurafsky & Martin, 2000). A ambigüidade categorial ocorre quando as duas ou mais opções de tradução de uma dada palavra da LF têm diferentes categorias gramaticais na LA. Um exemplo de ambigüidade categorial causada pela relação de homonímia é a palavra do inglês *field*, que pode ser traduzida para as palavras “campo” (substantivo) ou “interceptar” (verbo), no português. Já um exemplo de ambigüidade categorial derivada da relação de polissemia é a palavra do inglês *eats*, que pode ser traduzida no português como “mantimentos, víveres, gêneros alimentícios” (substantivos) ou “come” (verbo “comer” conjugado na terceira pessoa singular, presente do indicativo). Esse tipo de ambigüidade não ocorre no caso da relação de sinonímia. A ambigüidade de sentido, por sua vez, ocorre quando as duas ou mais opções de tradução de uma dada palavra da LF têm a mesma categoria gramatical na LA. Nesse caso, valem os exemplos já citados anteriormente: *know*, traduzido como “saber” ou “conhecer”, como exemplo de polissemia; *light*, traduzido como “leve” ou “luz”, como exemplo de homonímia; e *dog*, traduzido como “cachorro” e “cão”, como um exemplo de sinonímia.

A ambigüidade categorial é, em geral, muito mais simples que a de sentido, uma vez que pode ser resolvida, na maioria das vezes, pela análise das características sintáticas das palavras (realizada por procedimentos de etiquetagem gramatical ou análise sintática). A resolução da ambigüidade de sentido, por sua vez, exige a análise da semântica das palavras e, eventualmente, a análise do uso de tais palavras nas línguas fonte e alvo (realizadas por procedimentos de análise semântica e pragmática, por exemplo). Atualmente, em função dos avanços do PLN, de fato, a ambigüidade categorial praticamente não representa um grande problema para a TA. O foco da maioria dos trabalhos para o tratamento da ambigüidade lexical é, portanto, o problema da ambigüidade de sentido, isto é, a tarefa de DLS.

A tarefa de DLS constitui-se, basicamente, das seguintes etapas:

1. Especificação das palavras a serem desambiguadas, isto é, todas as palavras ambíguas de um texto a ser traduzido ou, como normalmente ocorre, apenas determinados grupos de palavras, estabelecidos com base na similaridade sintático-semântica das palavras, na frequência dessas palavras em um *corpus*, etc.;
2. Identificação de todos os diferentes sentidos de cada palavra ambígua selecionada. Para tanto, podem ser utilizados recursos como dicionários bilíngües, tradutores humanos, *corpus* paralelos, etc.;
3. Definição de um mecanismo para atribuir a cada ocorrência da palavra o sentido mais apropriado, isto é, criação do modelo de WSD, propriamente dito;
4. Incorporação desse modelo a um sistema de TA.

O foco deste relatório, conforme mencionado, está na primeira etapa, de especificação de um grupo de palavras para delimitar o modelo de DLS a ser inicialmente proposto. Essa etapa será realizada por meio de um estudo com base na análise de uma amostra do *corpus* BNC, brevemente descrito a seguir.

### 3 O *corpus* BNC

O BNC (*British National Corpus*)<sup>2</sup> (Burnard, 2000) é um *corpus* de 100 milhões de palavras distribuídas em amostras de textos na língua inglesa escrita e falada de uma grande variedade de fontes, de diferentes gêneros e domínios. A sua construção foi realizada de 1991 a 1994, por um consórcio

---

<sup>2</sup> <http://www.natcorp.ox.ac.uk>

formado pela união da *Oxford University Press* com algumas editoras de dicionários e alguns centros acadêmicos de pesquisa e bibliotecas.

A parte do *corp*us referente à língua escrita (90%) inclui, por exemplo, extratos de jornais regionais e nacionais, periódicos especializados e revistas voltadas para todas as idades e interesses, livros acadêmicos e de ficção, cartas e memorandos publicados e não publicados, redações de escolas e universidades, entre outros tipos de texto. A parte referente à língua falada (10%) inclui uma grande quantidade de conversas informais, gravadas por voluntários selecionados de diferentes idades, regiões e classes sociais, de maneira demograficamente balanceada. Inclui, também falas coletadas a partir de vários tipos de contextos, desde reuniões de negócios formais, informais ou governamentais a shows e programas de rádio e de televisão.

Ao todo, o *corp*us contém 4.124 textos, dos quais 863 são transcritos de conversas ou monólogos falados. Cada texto é segmentado em sentenças, sendo que para cada palavra da sentença é atribuído um código que representa sua categoria gramatical. Há 6.250.000 sentenças no *corp*us todo. A segmentação e a classificação gramatical foram realizadas automaticamente pelo etiquetador estocástico CLAWS4 (Marshall, 1983; Garside et al., 1987), desenvolvido na Universidade de Lancaster.

O *corp*us BNC foi projetado para representar significativamente a língua inglesa britânica moderna, nas suas formas escrita e falada, em vários gêneros e domínios. É amplamente utilizado em trabalhos de várias áreas relacionadas ao estudo e processamento da língua inglesa, incluindo muitos trabalhos de PLN, uma vez que apresenta características que facilitam esse tipo de trabalho, como a segmentação em sentenças e palavras e a etiquetagem gramatical. Por exemplo, no trabalho do qual o estudo apresentado neste relatório faz parte, isto é, o desenvolvimento de um modelo de DLS, pretende-se utilizar esse *corp*us. Por essas razões, ele foi escolhido para ser utilizado neste estudo.

### 3.1 Composição do *corp*us

As principais características da composição do *corp*us BNC são as seguintes:

- O *corp*us é **monolingüe**, consistindo apenas do inglês britânico moderno. No entanto, palavras em inglês não britânico e em línguas estrangeiras podem ocasionalmente ocorrer.
- O *corp*us é **sincrônico**, cobrindo apenas o inglês britânico do século XX, e não o desenvolvimento histórico que o produziu.
- O *corp*us é **genérico**, incluindo muitos estilos diferentes e não é limitado a um domínio, um gênero ou um registro particular.
- Com relação à escolha das **amostras** do *corp*us, para as fontes escritas, amostras de 45.000 palavras foram extraídas de diversas partes de textos de autores individuais. Textos menores ou textos com vários autores foram integralmente incluídos. A técnica de amostragem serviu para permitir uma cobertura maior do *corp*us no limite de 100 milhões de palavras e evitar a representação excessiva de textos específicos.

Com relação aos textos referentes à língua escrita, eles foram selecionados de acordo com 3 critérios de seleção independentes: domínio (o tipo de texto), período (a data de publicação do texto) e meio (o tipo de publicação na qual o texto ocorre). As proporções foram definidas para cada um desses critérios, conforme segue:

#### **Domínio:**

- 75% dos textos são *informativos*, dos quais foram selecionadas quantidades iguais dos campos de ciências aplicadas, artes, crenças & pensamento, comércio & fianças, lazer, ciência pura & natural, ciência social, assuntos mundiais;
- 25% dos textos escritos são *imaginativos*, ou seja, de trabalhos literários e criativos.



**Meio:**

- 60% dos textos escritos provêm de livros;
- 25% provêm de periódicos (jornais, etc.);
- entre 5% e 10% provêm de outros tipos variados de materiais publicados (panfletos, folhetos de propaganda, etc.);
- entre 5% e 10% provêm de material escrito não publicado, como cartas e diários pessoais, redações e memorandos, etc.;
- menos de 5% provêm de material escrito para ser utilizado oralmente (por exemplo, discursos políticos, textos de peças teatrais, etc.).

**Período:** Foram selecionados apenas textos posteriores a 1964.

Para a documentação do *cópus* foram definidos vários critérios de classificação para os seus textos. Não foram definidas proporções fixas de textos que deveriam respeitar cada um desses critérios, a intenção era simplesmente garantir um nível apropriado de variação em cada critério. Tais critérios incluem:

- Número de palavras e extensão da amostra;
- Tópico ou assunto do texto;
- Nome, idade do autor, gênero, região de origem e domicílio;
- Faixa etária e gênero alvos;
- “Nível” da escrita (uma medida subjetiva da dificuldade da leitura): quanto mais literário ou técnico é o texto, mais alto o seu nível.

Já com relação aos textos referentes à língua falada, há duas partes de igual tamanho: uma parte **demográfica**, contendo transcrições de conversas naturais e espontâneas realizadas por pessoas comuns, e uma parte **governada pelo contexto**, contendo transcrições de gravações feitas em categorias específicas de reuniões e eventos.

Para a parte demográfica, foram recrutados 124 voluntários, incluindo-se, de maneira balanceada, pessoas dos sexos masculino e feminino, de várias idades, moradores de 38 lugares diferentes do Reino Unido e pertencentes a diferentes classes sociais. Foram gravadas todas as conversas dos voluntários reunidos por dois ou três dias.

Na parte governada pelo contexto, a intenção era coletar quantidades iguais de textos falados em cada uma das seguintes categorias de contextos sociais:

- Eventos educacionais e informativos, como palestras, noticiários, discussões em salas de aula, tutoriais, etc.
- Eventos de negócios, como demonstrações de vendas, encontros de comércio, entrevistas, etc.
- Eventos institucionais e públicos, como sermões em igrejas, discursos políticos, procedimentos parlamentares, etc.
- Eventos de lazer, como comentários esportivos, encontros em clubes, etc.

## 3.2 Codificação dos textos

A marcação introduzida nos textos do BNC indica uma série de informações importantes, incluindo:

- O limite e a categoria gramatical de cada palavra;
- Os limites de cada sentença;
- Limites dos parágrafos, seções e cabeçalhos nos textos escritos;
- Entonações, pausas e características para-lingüísticas, tais como risos, nos textos falados;
- Informações meta-textuais sobre a fonte ou a codificação dos textos.

O formato para expressar essas características é chamado *Corpus Document Interchange Format* (CDIF). CDIF é uma aplicação do SGML (*Standard Generalized Markup Language*).

### 3.3 Etiquetação gramatical

As 100 milhões de palavras do BNC, bem como os sinais de pontuação, foram anotadas com suas etiquetas gramaticais. As etiquetas gramaticais foram definidas pela *Unit for Computer Research on the English Language* (UCREL) da Universidade de Lancaster.

O corpus de 100 milhões de palavras foi etiquetado usando um conjunto de etiquetas (C5) chamado *BNC Basic Tagset*. Adicionalmente, um subconjunto de 2 milhões do BNC, o “Córpus Central” (*Core Corpus*), foi etiquetado usando um conjunto de etiquetas mais detalhado (C7), chamado *BNC Enriched Tagset*. Esse conjunto prevê as etiquetas da Tabela 1 para os verbos da língua inglesa. Vale notar que as etiquetas da classe “geral” podem se aplicar a todos os verbos, enquanto as demais só se aplicam a determinados verbos, conforme mostra a Tabela.

Tabela 1 – Possíveis etiquetas para verbos do *tagset C7*

geral	<b>VV0</b> – a forma básica do verbo, como a forma finita (em sentenças declarativas e passivas) (e.x. <i>give, find, look, receive</i> )
	<b>VVD</b> – a forma do verbo no passado simples (e.x. <i>gave, found, looked, received</i> )
	<b>VVG</b> – a forma <i>-ing</i> do verbo (e.x. <i>giving, finding, looking, receiving</i> )
	<b>VVGK</b> – a forma <i>-ing</i> de um verbo concatenativo (e.x. <i>going</i> em <i>be going to</i> )
	<b>VVI</b> – a forma básica do verbo como um infinitivo (e.x. <i>give, find, look, receive</i> )
	<b>VVN</b> – a forma do verbo no passado particípio (e.x. <i>given, found, looked, received</i> )
	<b>VVNK</b> – o passado particípio de um verbo concatenativo (e.x. <i>bound</i> em <i>be bound to</i> )
	<b>VVZ</b> – a forma <i>-s</i> do verbo (e.x. <i>gives, finds, looks, receives</i> )
be	<b>VBDR</b> – <i>were</i>
	<b>VBDZ</b> – <i>was</i>
	<b>VBG</b> – <i>being</i>
	<b>VBI</b> – <i>be</i> como forma infinitiva
	<b>VBM</b> – <i>am, 'm</i>
	<b>VBN</b> – <i>been</i>
	<b>VBR</b> – <i>are, 're</i>
	<b>VBZ</b> – <i>is, 's</i>
do	<b>VD0</b> – <i>do</i> como forma infinitiva (em sentenças declarativas e imperativas)
	<b>VDD</b> – <i>did</i>
	<b>VDG</b> – <i>doing</i>
	<b>VDI</b> – <i>do</i> como forma infinitiva
	<b>VDN</b> – <i>done</i>
	<b>VDZ</b> – <i>does, 's</i>
have	<b>VH0</b> – <i>have, 've</i> como forma infinitiva (em sentenças declarativas e imperativas)
	<b>VHD</b> – <i>had, 'd</i> como tempo passado
	<b>VHG</b> – <i>doing</i>
	<b>VHI</b> – <i>have</i> como forma infinitiva
	<b>VHN</b> – <i>had</i> como passado particípio
	<b>VHZ</b> – <i>has, 's</i>
modal	<b>VM</b> – auxiliar modal (e.x. <i>can, could, will, would, must</i> )
	<b>VMK</b> – auxiliar modal auxiliar concatenativo (i.e. <i>ought</i> e <i>used</i> quando seguidos pelo marcador de infinitivo <i>to</i> )

O corpus de 100 milhões de palavras do BNC foi etiquetado automaticamente, usando o etiquetador CLAWS4. Em função do tamanho do cópús, não havia possibilidade de realizar a pós-edição e a correção da etiquetação, assim, os erros (1,7% do total de palavras, excluindo-se os sinais de

pontuação) permaneceram no *cópus*. Além disso, esse *corpus* contém etiquetas ambíguas, tais como VVD-VVN. Aproximadamente 4.7% das etiquetas (excluindo-se as etiquetas de pontuação) são ambíguas. Contudo, a etiquetagem das 2 milhões de palavras do *Cópus Central* foi manualmente pós-editada e corrigida. Por essa razão, esse subconjunto do *cópus* apresenta baixas taxas de erro: menos que 0.3%.

O objetivo quando da criação do *Cópus Central* de 2 milhões de palavras era obter um subconjunto representativo do *cópus* completo (de 100 milhões de palavras), no sentido de que ele contém exemplos de todas as maiores subdivisões do BNC completo, aproximadamente nas mesmas proporções. Contudo, há uma diferença essencial entre os dois: enquanto no BNC completo somente 10 milhões de palavras (10% do *cópus*) consiste de textos falados, o *Corpus Central* apresenta aproximadamente a mesma quantidade de textos escritos e falados (1 milhão de palavras cada).

## 4 Análise do *cópus* BNC

Para a realização do estudo sobre o problema da ambigüidade lexical de sentido na TA de verbos do inglês para o português, com base nas sentenças do BNC, a metodologia seguida envolveu as atividades descritas no decorrer desta seção.

### 4.1 Seleção dos verbos

Conforme mencionado, o recorte do trabalho para a análise apenas dos verbos se deve a uma delimitação inicial da proposta do modelo de DLS. Posteriormente, pretende-se estender o trabalho a palavras de outras classes gramaticais. Além desta, outra delimitação foi estabelecida com relação ao conjunto de verbos a serem analisados nesta etapa inicial: apenas os verbos mais frequentes do BNC. Para tanto, foram selecionados os verbos com mais de 100 mil ocorrências no *cópus*, a partir de uma lista de frequência lematizada levantada por Adam Kilgarriff (<ftp://ftp.itri.bton.ac.uk/pub/bnc/>). Esse critério resultou na seleção dos seguintes 15 verbos: *to be, to have, to do, to say, to go, to get, to make, to see, to know, to take, to think, to come, to give, to look* e *to use*.

Antes de analisar a ambigüidade dos verbos selecionados com relação ao seu uso do *cópus*, procurou-se identificar aqueles que apresentavam ambigüidade em potencial, isto é, aqueles verbos para os quais mais de uma tradução, com diferentes sentidos, mas da mesma categoria gramatical (verbo, neste caso), é encontrada em recursos multilingües. Verbos que apresentam apenas uma possível tradução, ou cujas traduções variam somente no caso do seu uso em expressões de sentido figurado, não seriam considerados verbos “problemáticos” para a tradução do ponto de vista do fenômeno da ambigüidade lexical de sentido e, portanto, seriam eliminados da análise em questão.

Para analisar as possíveis acepções (traduções) de cada um dos 15 verbos, foram consultadas as suas entradas, na categoria gramatical “verbo” no dicionário eletrônico DTS DIC Prático Michaelis® 5.1. As traduções transcritas abaixo são apresentadas nesse dicionário. Vale notar que também foram consideradas acepções possíveis aquelas relativas ao uso do verbo (no seu sentido denotativo) em expressões, as quais não estão transcritas aqui devido a sua grande quantidade e à possibilidade de variação da sua tradução, de acordo com as variações nas demais palavras da expressão. Por outro lado, traduções de expressões considerando o uso dos verbos no sentido figurado não foram consideradas.

**To be:** 1 ser, existir, viver, ser realidade. 2 ter lugar, acontecer, realizar-se. 3 permanecer, ficar, continuar. 4 igualar, representar, significar. 5 estar, encontrar-se. 6 visitar.

**To have:** 1 ter, haver, possuir. 2 sofrer. 3 manter, reter (na memória). 4 conter, compreender. 5 ter de, precisar, ser obrigado a, dever. 6 permitir, tolerar, deixar. 7 tomar, beber. 8 comer.

**To do:** 1 fazer, executar, agir, atuar, efetuar, trabalhar. 2 acabar, pôr fim a, concluir, completar. 3 preparar, arranjar. 4 interpretar, representar, desempenhar o papel de. 5 criar, produzir. 6 causar, levar a efeito ou a termo. 7 render, prestar. 8 haver-se, portar-se, atuar, proceder. 9 estar ou passar bem ou mal de saúde. 10 tratar com, ocupar-se de, acabar com. 11 servir, bastar, ser suficiente ou satisfatório, convir. 12 cozer, assar. 13 percorrer, cobrir. 14 enganar, lograr, trapacear. 15 matar, liquidar, arruinar. 16 Coloq. acolher, entreter, sustentar. 17 visitar lugares interessantes. 18 cumprir. 19 esgotar-se, gastar-se. 20 esforçar-se. 21 traduzir, reter em, modificar. 22 Gír. consumir drogas.

**To say:** 1 falar, dizer, afirmar. 2 exprimir, declarar, anunciar, pôr em palavras. 3 recitar, repetir. 4 supor, dar como exemplo. 5 explicar, dar uma opinião.

**To go:** 1 ir, seguir, prosseguir, andar. 2 sair, partir, deixar, ir embora. 3 estar em movimento, andar, trabalhar (máquinas), soar. 4 ficar, tornar-se, vir a ser. 5 estar, ser. 6 começar, empreender. 7 proceder, avançar. 8 correr, vagar, estar em uso corrente. 9 meter-se, intrometer-se. 10 estender-se, alcançar. 11 passar. 12 ser vendido, ser entregue. 13 tender, levar, conduzir. 14 resultar, redundar. 15 pertencer, caber. 16 combinar, harmonizar. 17 explodir, estourar. 18 deixar de existir, perder, gastar. 19 morrer. 20 afrouxar.

**To get:** receber, obter, ganhar, alcançar. 2 ficar, tornar-se, vir a ser. 3 aprender, decorar. 4 adquirir, contrair, apanhar, pegar. 5 suceder, conseguir. 6 buscar, pegar, arranjar, procurar. 7 tomar, comer. 8 induzir, persuadir. 9 causar, motivar. 10 mandar, mandar fazer. 11 criar, dar a luz. 12 mover, trazer, tirar. 13 Coloq. ser obrigado a, ter de. 14 chegar, vir, ir, partir, alcançar. 15 Gír. bater, surrar, matar. 16 compreender, entender. 17 ter, possuir. 18 engendrar, procriar. 19 Coloq. ser bem-sucedido. 20 Coloq. levar a melhor. 21 achar. 22 pegar, colher, surpreender. 23 comover. 24 Am. Gír. assassinar. 25 tomar, comer. 26 tratar de, encarregar-se de. 27 providenciar, mandar fazer. 28 irritar, confundir. 29 transportar, levar, trazer. 30 persuadir, convencer.

**To make:** 1 fazer, fabricar. 2 construir. 3 criar. 4 elaborar. 5 compor. 6 efetuar. 7. causar, motivar. 8 executar, representar. 9 resultar. 10 dispor. 11 preparar. 12 determinar, promulgar. 13 promover. 14 constituir. 15 ganhar, lograr. 16 forçar, induzir, compelir. 17 dirigir-se. 18 marcar pontos (jogos). 19 surtir efeito. 20 deduzir. 21 julgar. 22 percorrer. 23 servir. 24 atingir. 25 receber, ganhar. 26 refrear.

**To see:** 1 ver, perceber, olhar. 2 examinar com os olhos, observar, espiar. 3 perceber, compreender, aprender. 4 descobrir, verificar. 5 pensar, considerar. 6 atender. 7 encontrar, conversar com. 8 procurar, consultar. 9 visitar. 10 assistir, cuidar. 11 acompanhar, escoltar. 12 passar por, ter experiência com.

**To know:** 1 saber, conhecer, entender. 2 reconhecer, indentificar (for como). 3 estar ciente, estar informado. 4 ter a certeza. 5 estar relacionado, conhecer pessoalmente. 6 estar habilitado, ter experiência, ser hábil, destro. 7 distinguir (from de). 8 arc. Bíblia conhecer, ter relações sexuais com.

**To take:** 1 tomar, pegar. 2 alcançar, agarrar, prender, capturar, apropriar-se. 3 arrebatado, arrancar, levar. 4 receber (como pagamento), aceitar, obter, adquirir. 5 tomar, receber como esposo ou esposa, tomar alojamento ou pensão. 6 suportar, receber, acolher. 7 tomar, comer, beber, engolir, consumir. 8 ganhar. 9 apanhar, contrair (doença). 10 ocupar. 11 usar, tomar (um veículo). 12 aproveitar (oportunidade). 13 tirar, tomar (férias). 14 submeter-se, sofrer, aguentar. 15 necessitar, requerer, exigir. 16 abranger, ocupar, consumir (tempo). 17 escolher, selecionar. 18 afastar, remover (por morte). 19 diminuir, prejudicar. 20 subtrair, extrair, extorquir, fraudar, lesar, roubar. 21 guiar, levar. 22 acompanhar, escoltar, levar. 23 carregar, transportar. 24 fazer, tirar (fotografia), ser fotogênico. 25 sentir (orgulho). 26 determinar, verificar. 27 agir, ter efeito. 28 compreender. 29 assumir, supor. 30 considerar. 31 assumir (responsabilidade), adotar. 32 alugar, empregar. 33 escrever, anotar. 34 tomar assinatura, assinar. 35 Gram. ser usado com. 36 agradar, atrair, encantar. 37 prender (a atenção). 38 ir, andar. 39 vencer,

conquistar, tomar (obstáculo). 40 ficar, tornar-se. 41 pegar, começar a crescer. 42 ser afetado por, sentir, experimentar, contrair. 43 Fig. manietar, subjugar. 44 assaltar, surpreender. 45 necessitar, custar. 46 colher. 47 fazer (viagem). 48 conquistar. 49 enveredar. 50 deduzir. 51 inflamar-se. 52 conceber, emprenhar. 53 morder a isca. 54 preparar-se, empreender.

**To think:** 1 conceber, formar na mente, imaginar. 2 pensar, idear, cogitar. 3 considerar, julgar. 4 crer, supor, opinar. 5 refletir, meditar, considerar, estudar. 6 imaginar, especular, ponderar. 7 lembrar, recordar. 8 considerar, formar opinião sobre.

**To come:** 1 vir, aproximar(-se). 2 chegar. 3 surgir. 4 alcançar, atingir. 5 acontecer, ocorrer. 6 resultar, redundar, advir. 7 nascer, proceder, emanar. 8 ficar, tornar-se, vir a ser. 9 passar, entrar. 10 andar, percorrer. 11 ocorrer, ser lembrado. 12 ser obténível, estar disponível. 13 importar em, custar, perfazer. 14 chegar a. 15 formar-se, tomar forma ou feitio. 16 estender-se, avançar, ir até. 17 tratar-se de, referir-se a. 18 ser levado a, vir a. 19 Coloq. ter um orgasmo, gozar.

**To give:** 1 dar, presentear, conceder. 2 entregar, oferecer, ceder. 3 propor, oferecer. 4 fornecer, prover. 5 apresentar, mostrar, notificar, transmitir. 6 aplicar, administrar. 7 prestar. 8 conferir, atribuir, confiar, incumbir. 10 proferir, dizer, contar, exalar. 11 aplicar-se, dedicar-se, esforçar-se. 12 degelar (solo). 13 abrir, dar passagem, dá vista. 14 Gír. acontecer.

**To look:** 1 olhar. 2 contemplar, observar. 3 considerar. 4 prestar atenção. 5 ter vista para. 6 parecer. 7 inspecionar, examinar.

**To use:** 1 usar. 2 praticar. 3 habituar, acostumar, costumar. 4 aproveitar(-se), servir(-se), utilizar(-se), explorar. 5 gastar, consumir, esgotar. 6 aplicar, empregar. 7 manusear. 8 tratar.

Vale observar que há duas diferentes divisões para as possíveis traduções dos verbos: diversos grupos de acepções, identificados por diferentes números, e diversas acepções separadas por vírgula e identificadas pelo mesmo número. Por exemplo, o primeiro grupo para o verbo “be” indica as traduções “1 ser, existir, viver, ser realidade”, enquanto que o segundo grupo indica as traduções “2 ter lugar, acontecer, realizar-se”. A maioria dos dicionários costuma agrupar traduções que representam acepções polissêmicas de uma palavra, ou seja, acepções relacionadas entre si. As traduções em diferentes grupos, neste caso, representariam casos de homonímia, ou seja, de acepções não relacionadas. Dessa maneira, haveria um relacionamento intragrupo, mas não entre intergrupos. Contudo, conforme mencionado, não há consenso, mesmo na área de Lexicografia, no que diz respeito aos critérios que definem a diferenciação entre homonímia e polissemia e, portanto, esses fenômenos serão analisados indistintamente.

Como pode ser observado, todos os verbos podem apresentar diferentes traduções, seja com significados relacionados ou totalmente distintos. Assim, considerou-se que os 15 verbos pré-selecionados podem causar problemas de ambigüidade na tradução e, portanto, todos eles foram mantidos para análise.

## 4.2 Seleção das sentenças

Para cada um dos 15 verbos considerados, foram selecionadas, aleatoriamente, sentenças do Córpus Central do BNC, isto é, da amostra do córpus cuja etiquetagem gramatical foi revisada manualmente. Somente essa porção do córpus foi utilizada para a seleção, nesta etapa, para evitar que erros de etiquetagem pudessem interferir na diversidade das sentenças. Uma vez que o número de sentenças selecionadas é relativamente pequeno, problemas com a etiquetagem poderiam fazer com que não fossem incluídos exemplos dos verbos com variações no tempo, pessoa e modo, indicadas por

determinadas etiquetas gramaticais. A seleção das sentenças foi feita a partir da interface de acesso ao BNC, denominada SARA<sup>3</sup>, a qual possui, entre suas ferramentas e recursos concordanciador.

O número total de sentenças selecionadas para cada verbo variou de acordo com as possíveis etiquetas do verbo. Em princípio, seriam selecionadas cinco sentenças para cada etiqueta de cada verbo (Tabela 1). Contudo, para alguns verbos, o corpus não possuía exemplos de sentenças nas quais o verbo se apresentava em todas as suas possíveis variações ou, em alguns casos, possuía menos de cinco sentenças para cada variação do verbo. Dessa maneira, o número de sentenças variou de 30 a 51. Por exemplo, para o verbo *to be* foram selecionadas 51 sentenças, para o verbo *to take*, 35 sentenças, e para o verbo *to do*, 30 sentenças.

Alguns exemplos de sentenças selecionadas para diferentes verbos com suas respectivas etiquetas são ilustrados na Tabela 2.

Tabela 2 – Exemplos de sentenças selecionadas para alguns verbos

Verbo	Sentença	Etiqueta do verbo
<i>to be</i>	<b>Be</b> so good as to enter.	VBO
<i>to be</i>	<b>Be</b> ready to stake any which show signs of flopping.	VBO
<i>to be</i>	I <b>am</b> 18 years old.	VBM
<i>to be</i>	And he warned that if there <b>was</b> not an honourable settlement, that would explode the whole region.	VBDZ
<i>to come</i>	Most nuns <b>come</b> from the Tibetan countryside and may join the nunnery in their early teens or even when they are as young as nine or 10.	VVO
<i>to come</i>	Well he <b>come</b> into garage the other day and told me his wife had kicked him out!	VVD
<i>to come</i>	"This city has suddenly <b>come</b> alive," said her husband, an off-duty border guard.	VVN
<i>to come</i>	Mr Gonzalez has also <b>come</b> in for criticism from within his own party.	VVN
<i>to get</i>	It's best to be alone when the noises <b>get</b> this loud.	VVO
<i>to get</i>	"You <b>get</b> three clear benefits out of this," he explained.	VVO
<i>to get</i>	Policemen struggled briefly with the women as they tried to <b>get</b> into the embassy.	VVI
<i>to get</i>	I was quite sick on the train before we <b>got</b> to Euston, and have not liked long train journeys since.	VVD
<i>to give</i>	Let those who <b>give</b> us such advice sign it themselves.	VVO
<i>to give</i>	He arrived late yesterday afternoon, it seems, and he has offered to <b>give</b> a little recital to anyone on the island who wishes to attend.	VVI
<i>to give</i>	The queen broached the matter to me and ostensibly I am now <b>giving</b> thought to my answer — what would you say to going in my place?	VVG
<i>to know</i>	These systems work well if they <b>know</b> the parameters of Soviet radar very well.	VVO
<i>to know</i>	But I suppose you're right: the public isn't going to notice and those who <b>know</b> me.	VVO

### 4.3 Escolha dos sistemas de tradução e realização da tradução

Atualmente há vários sistemas de TA disponíveis na *web*, os quais podem ser utilizados por usuários *on-line*. Tais usuários submetem os textos em um língua e estes são traduzidos para outra língua. Para realizar a tradução, o usuário normalmente não tem acesso a opções de configuração.

A princípio, foram escolhidos cinco sistemas de TA com opção para a tradução do inglês para o português, em função da sua disponibilidade gratuita e/ou da sua ampla utilização: Globalink L&H Power Translator® Pro, Systran<sup>4</sup>, FreeTranslation<sup>5</sup>, E-Translation Server<sup>6</sup> e Amikai<sup>7</sup>. Em uma análise preliminar, contudo, as versões *on-line* utilizadas dos sistemas E-Translation Server e Amikai não

<sup>3</sup> <http://www.natcorp.ox.ac.uk/SARA/index.htm>

<sup>4</sup> <http://www.systransoft.com>

<sup>5</sup> <http://www.freetranslation.com>

<sup>6</sup> <http://www.linguatex.net/online/ptwebtext>

<sup>7</sup> <http://www.amikai.com/demo.jsp>

permitiram a tradução de sentenças longas (muito frequentes no corpus) e foram, por essa razão, desconsiderados. Portanto, somente os três primeiros sistemas foram efetivamente utilizados para a análise em questão, a fim de que pudessem ser obtidas três traduções para todas as sentenças. Os três sistemas escolhidos pretendem ser independentes de gênero e domínio, isto é, não possuem características específicas para a tradução de textos de gêneros e/ou domínios específicos. O primeiro deles é um sistema comercial cujo desenvolvimento vem sendo aprimorado desde 1993, enquanto que os demais são sistemas voltados especificamente para o acesso gratuito via *web* ou possuem versões disponíveis gratuitamente para esse tipo de acesso.

Em um processo manual, as sentenças selecionadas (descritas na seção anterior) foram submetidas aos três sistemas de TA e as suas respectivas traduções foram esquematizadas em tabelas para a análise a ser realizada posteriormente. Exemplos de traduções dos diferentes sistemas para algumas sentenças da Tabela 2 são ilustrados na Tabela 3.

Tabela 3 – Exemplos de traduções das sentenças da Tabela 2

<b>Be so good as to enter.</b>	
FreeTranslation	Seja tão bom como entrar.
Systran	Seja assim bom a respeito de entram.
TranslatorPro	Seja tão bom sobre entre.
<b>Be ready to stake any which show signs of flopping.</b>	
FreeTranslation	Esteja pronto a estaca qualquer que mostra sinais de cair.
Systran	Esteja pronto para estacar alguns que mostrarem sinais de flopping.
TranslatorPro	Esteja pronto para apostar qualquer que sinais de espetáculo de baquear.
<b>I am 18 years old.</b>	
FreeTranslation	Tenho 18 anos de idade.
Systran	Eu tenho 18 anos velho.
TranslatorPro	Eu tenho 18 anos.
<b>Most nuns come from the Tibetan countryside and may join the nunnery in their early teens or even when they are as young as nine or 10.</b>	
FreeTranslation	A maioria de freiras vem da campina de Tibetan e pode unir-se o nunnery em seu cedo adolescente ou mesmo quando são tão jovem quanto nove ou 10.
Systran	A maioria de nuns vêm do campo tibetano e podem juntar o nunnery em seus teens adiantados ou mesmo quando são tão novos quanto nove ou 10.
TranslatorPro	A maioria das freiras vem da zona rural de Tibetan e pode se juntar ao convento nas adolescências cedo deles/delas ou até mesmo quando eles são tão jovens quanto nove ou 10.
<b>Well he come into garage the other day and told me his wife had kicked him out!</b>	
FreeTranslation	Bem vem em garagem o outro dia e me contou sua esposa tinha chutado-o para fora!
Systran	Poço vem na garagem o outro dia e dito me sua esposa tinha-o retrocedido para fora!
TranslatorPro	Bem ele entra em garagem o outro dia e me falou para a esposa dele tinha o expulsado!
<b>"This city has suddenly come alive," said her husband, an off-duty border guard.</b>	
FreeTranslation	"Esta cidade repentinamente veio vivo," disse seu marido, um guarda de fronteira de fora-dever.
Systran	"esta cidade tem vivo de repente vindo," dito seu marido, um protetor off-duty da beira.
TranslatorPro	"Esta cidade veio viva" de repente, disse o marido dela, um guarda de borda de fora-dever.
<b>It's best to be alone when the noises get this loud.</b>	
FreeTranslation	Está melhor estar só quando os barulhos recebem este alto.
Systran	É o mais melhor estar sozinho quando os ruídos começam este alto.
TranslatorPro	É melhor para estar só quando os barulhos adquirirem este alto.
<b>"You get three clear benefits out of this," he explained.</b>	
FreeTranslation	"Recebe três remover benefícios de isto," explicou
Systran	"Você começa três benefícios desobstruídos fora deste," ele explicou
TranslatorPro	" Você sai três benefícios claros disto," ele explicou.
<b>Let those who give us such advice sign it themselves.</b>	
FreeTranslation	Deixe esses quem nos dão tal conselho assinaá-lo sees.
Systran	Deixe aqueles que nos dão tal sinal do conselho ele eles mesmos.
TranslatorPro	Deixe esses que nos dão tal sinal de conselho isto eles.
<b>He arrived late yesterday afternoon, it seems, and he has offered to give a little recital to anyone on the island who wishes to attend.</b>	
FreeTranslation	Chegou tarde ontem tarde, parece, e ofereceu dar um recital pequeno a ela na ilha que deseja assistir.

Systran	Chegou tarde atrasada do ontem, parece, e ofereceu dar um recital pequeno a qualquer um no console que deseja atender.
TranslatorPro	Ele chegou ontem tarde tarde, parece, e ele ofereceu a dar um pequeno recital a qualquer um na ilha que deseja assistir.
These systems work well if they <b>know</b> the parameters of Soviet radar very well.	
FreeTranslation	Estes sistemas trabalham bem se sabem os parâmetros de radar Soviético muito bem.
Systran	Estes sistemas trabalham bem se souberem os parâmetros do radar soviético muito bem.
TranslatorPro	Estes sistemas trabalham bem se eles souberem muito bem os parâmetros de radar soviético.
But I suppose you're right: the public isn't going to notice and those who <b>know</b> me.	
FreeTranslation	Mas eu o suponho são direito: o público não notará e esses quem me sabem.
Systran	Mas eu s uponho que você é direito: o público não está indo observar e aqueles que me conhecem.
TranslatorPro	Mas eu suponho você tem razão: o público não vai notar e esses que me conhecem.

#### 4.4 Análise das traduções

O objetivo da análise das traduções, nesta fase do experimento, era verificar as acepções de cada verbo nas diferentes sentenças e o comportamento dos sistemas de tradução selecionados diante desses verbos. Para tanto, cada tabela com as sentenças de origem de um verbo, juntamente com suas respectivas traduções nos sistemas, foi analisada manualmente, com o auxílio de dicionários bilíngües e outros recursos como a WordNet, por uma tradutora inglês-português profissional. Nessa análise, a tradutora acrescentou, para cada sentença de cada tabela, a tradução mais adequada (considerada ideal) do verbo em questão, a identificação dos sistemas que levavam a essa tradução ou a traduções alternativas, mantendo a mesma acepção, e comentários gerais, quando julgados importantes. Um exemplo de tabela com sua respectiva análise, para o verbo *to get*, é ilustrado na Tabela 4.

Vale ressaltar que essa análise focalizou apenas a tradução do verbo em questão, e não da sentença completa. Vale notar também que, desde que fosse possível identificar a acepção do verbo, problemas de concordância, tempo, modo ou mesmo de ortografia<sup>8</sup> na tradução do verbo não foram considerados erros.

Tabela 4 – Análise das traduções das sentenças do verbo *to get*

GET – VVO	
1. Soviet strikes point to crisis of morale that threatens future of perestroika Broadcasts of parliament <b>get</b> more viewers than Agatha Christie films It is the first November 7 since the revolution that coal miners have been on strike.	
Acepção correta: ter – tem / obtem – obtém	
FreeTranslation	Greves soviéticas apontam a crise de moral que ameaça futuro de Transmissões de perestroika de parlamento recebe mais observadores que Christie de Agatha filma é o primeiro novembro 7 desde que a revolução aquele mineiros de carvão estiveram em greve.
Systran	O soviete golpeia o ponto à crise do morale que ameaça o futuro de transmissões do perestroika do parliament começa mais visores do que películas que de Agatha Christie é o primeiro novembro 7 desde que a volta que os mineiros de carvão estiveram na batida.
TranslatorPro	Ponto de greves soviético para crise de moral que ameaça futuro de Radiodifusões de perestroika de parlamento adquire mais espectadores que Agatha Christie filma que é no primeiro 7 de novembro desde a revolução que os mineiros de carvão estiveram em greve.
Sistemas que traduziram corretamente: Nenhum	
Comentários: Não dá para saber quem é o sujeito da sentença e, portanto, se o verbo deve estar na 3ps ou 3pp.	
2. It's best to be alone when the noises <b>get</b> this loud.	
Acepção correta: ficar – fica	
FreeTranslation	Está melhor estar só quando os barulhos recebem este alto.
Systran	É o mais melhor estar sozinho quando os ruídos começam este alto.
TranslatorPro	É melhor para estar só quando os barulhos adquirirem este alto.
Sistemas que traduziram corretamente: Nenhum	
Comentários: essa sentença tb está mal escrita. O correto, gramaticalmente, seria: "It is better to be alone...."	

<sup>8</sup> O sistema Systran apresentou alguns erros de ortografia na tradução de alguns verbos, por exemplo, o verbo *to make*, na sentença "We shall be **making** music." é traduzido como "fazemos".



<b>3. "You get three clear benefits out of this," he explained.</b>	
Acepção correta: adquirir – adquirir / receber – recebe	
FreeTranslation	"Recebe três remover benefícios de isto," explicou
Systran	"Você começa três benefícios desobstruídos fora deste," ele explicou
TranslatorPro	" Você sai três benefícios claros disto, " ele explicou.
Sistemas que traduziram corretamente: FreeTranslation	
Comentários:	
<b>4. Planning a new border is always an exciting proposition but before you start, follow Phillipa Lambert's preparation tips to ensure that your plants get off to a flying start</b>	
Acepção correta: começar – comecem	
FreeTranslation	O planejamento uma nova fronteira é sempre uma proposta animadora mas antes de você começa, segue pontas de preparação do Lambert de Phillipa assegurar que suas plantas descem a um voar começa
Systran	Planejar uma beira nova é sempre um proposition emocionante mas antes que você comece, segue pontas da preparação de Phillipa Lambert para assegurar-se de que suas plantas comecem fora a um começo do vôo
TranslatorPro	Planejando uma borda nova sempre é uma proposição excitante mas antes de você começar, siga a preparação de Phillipa Lambert inclina para assegurar que suas plantas descem para um começo voador
Sistemas que traduziram corretamente: Nenhum	
Comentários:	
<b>5. Get back to your roots and solve your beauty problems the natural way</b>	
Acepção correta: voltar – volte	
FreeTranslation	Afaste-se a suas raízes e resolve seus problemas de beleza o meio natural
Systran	Comece para trás a suas raízes e resolva seus problemas da beleza a maneira natural
TranslatorPro	Volte a suas raízes e resolva seus problemas de beleza o modo natural
Sistemas que traduziram corretamente: TranslatorPro	
Comentários:	
<b>GET – VVI</b>	
<b>1. Even if Swapo fails to get the two-thirds of the vote it needs to write its own constitution, it will have little difficulty in persuading one or another of the smaller parties to work with it.</b>	
Acepção correta: conseguir / consiga (mesmo / ainda que X não consiga – mesmo que X não conseguir....)	
FreeTranslation	Mesmo que Swapo não consegue para receber os dois -terceiros do voto que necessita escrever o própria constituição, terá dificuldade pequena em convencer um ou outro dos partidos menores trabalhar com ele.
Systran	Mesmo se Swapo não começa o dois terços do voto necessita escrever seu próprio constitution, terá pouca dificuldade em persuadir u m ou outro dos partidos menores para trabalhar com ele.
TranslatorPro	Até mesmo se Swapo não adquire os dois -terços do voto que precisa escrever sua própria constituição, terá pequena dificuldade persuadindo um ou outras das festas menores para trabalhar com isto.
Sistemas que traduziram corretamente: Systran	
Comentários: Errou o tempo verbal.	
<b>2. Policemen struggled briefly with the women as they tried to get into the embassy.</b>	
Acepção correta: entrar	
FreeTranslation	Os policiais lutaram em resumo com as mulheres como tentaram fazer-se parte de a embaixada.
Systran	Os polícias esforçaram-se momentaneamente com as mulheres enquanto tentaram começar no embassy.
TranslatorPro	Os policial lutaram brevemente com as mulheres como eles tentaram entrar na embaixada.
Sistemas que traduziram corretamente: TranslatorPro	
Comentários:	
<b>3. I'd like to get this underbrush cleaned out now.</b>	
Acepção correta: ter /ser (Eu gostaria que esse mato fosse cortado agora)	
FreeTranslation	Gostaria de receber este underbrush limpar agora.
Systran	Eu gostaria de começar este underbrush limpado para fora agora.
TranslatorPro	Eu gostaria de adquirir este underbrush limpado agora fora.
Sistemas que traduziram corretamente: Nenhum	
Comentários:	
<b>4. The Sandinista campaign slogan, " Everything will get better", seems in this context an admission of defeat.</b>	
Acepção correta: ficar (melhor) / melhorar (get better) – vai ficar/se tornar / melhor / melhorará etc	
FreeTranslation	O lema de campanha de Sandinista, " Tudo melhorará", parece neste contexto uma admissão de derrota.
Systran	O slogan da campanha de Sandinista, "tudo começará mais melhor", parece neste contexto uma admissão da derrota.
TranslatorPro	O "slogan de campanha de Sandinista, Tudo melhorará, parece neste contexto uma admissão de derrota.

Sistemas que traduziram corretamente: TranslatorPro e FreeTranslation	
Comentários:	
5. A lot of international help will be needed to <b>get</b> things moving.	
Acepção correta: fazer	
FreeTranslation	Muita ajuda internacional será necessitada receber coisas movendo.
Systran	Os muitos da ajuda internacional serão needed começar mover-se das coisas.
TranslatorPro	Será precisada muita ajuda internacional adquirir mudança de coisas.
Sistemas que traduziram corretamente: Nenhum	
Comentários:	
<b>GETS – VVZ</b>	
1. My body <b>gets</b> itself into a state of refusal.	
Acepção correta: colocar – coloca-se	
FreeTranslation	O meu corpo recebe se num estado de recusa.
Systran	Meu corpo começa-se em um estado da recusa.
TranslatorPro	Meu corpo se entra em um estado de recusa.
Sistemas que traduziram corretamente: Nenhum	
Comentários:	
2. The Palace Garden contains the wells from which the city <b>gets</b> its name.	
Acepção correta: receber – recebeu	
FreeTranslation	O Jardim de Palácio contem o bem de que a cidade recebe seu nome.
Systran	O jardim do palácio contem os poços de que a cidade começa seu nome.
TranslatorPro	O Jardim de Palácio contém os poços dos quais a cidade obtém seu nome.
Sistemas que traduziram corretamente: FreeTranslation	
Comentários:	
3. It is only at this point that the style <b>gets</b> locked in.	
Acepção correta: ficar – fica / tornar – torna-se	
FreeTranslation	É só a este ponto que o estilo recebe encerrou.
Systran	É somente neste momento que o estilo começa locked dentro.
TranslatorPro	Só é neste momento que o estilo é prendido.
Sistemas que traduziram corretamente: Nenhum	
Comentários:	
4. If the business is a goer, the entrepreneur moves on to the full Enterprise Allowance system and <b>gets</b> back any surplus funds.	
Acepção correta: receber – recebe (de volta)	
FreeTranslation	Se o negócio é um goer, o empresário move em ao pleno sistema de Concessão de Empresa e afasta-se qualquer fundos de sobra.
Systran	Se o negócio for um goer, o empreendedor move-se sobre para o sistema cheio da permissão da empresa e começa-se para trás todos os fundos em excesso.
TranslatorPro	Se o negócio for um goer, o empresário passa para o sistema de Mesada de Empreendimento cheio e volta qualquer fundo de excesso.
Sistemas que traduziram corretamente: Nenhum	
Comentários:	
5. Music day <b>gets</b> deeply dippy start.	
Acepção correta: ter – tem	
FreeTranslation	O dia de música recebe profundamente dippy começa.
Systran	O dia da música começa o começo profundamente dippy.
TranslatorPro	Dia de música adquire começo de dippy profundamente.
Sistemas que traduziram corretamente: Nenhum	
Comentários:	
<b>GOT – VVD</b>	
1. When they <b>got</b> through into Lenin Square, the leadership of the republic climbed down from their viewing podium and ran off.	
Acepção correta: entrar – entraram	
FreeTranslation	Quando receberam por em Quadrado de Lenin, a liderança da república desceu de seu está vendo podium e correu fora.
Systran	Quando começaram completamente no quadrado de Lenin, a liderança da república escalada para baixo de seu podium da visão e funcionou fora.

TranslatorPro	Quando eles entraram por em Lenin Square, a liderança da república desceu do pódio vendo deles/delas e escapou.
Sistemas que traduziram corretamente: TranslatorPro	
Comentários:	
<b>2. But when she <b>got</b> outside and began to rearrange her belongings to fit more neatly in the bicycle basket, Elisabeth Danziger found a bar of fleur-de-lis soap and a cube of rose geranium bath crystals in her hand.</b>	
Acepção correta: ir (para fora) – quando ela foi para fora / saiu	
FreeTranslation	Mas quando recebeu fora de e começou a rearrange seu pertences assentar mais arrumadamente na cesta de bicicleta, Danziger de Elisabeth achou uma barra de fleur-de-sabão de lis e um cubo de cristais de banho de geranium de rosa na sua mão.
Systran	Mas quando começou fora e começou a rearranjar seus pertences para caber mais ordenadamente na cesta da bicicleta, Elisabeth Danziger encontrou uma barra do sabão da flor de lis e de um cubo de cristais cor-de-rosa do banho do geranium em sua mão.
TranslatorPro	Mas quando ela adquiriu fora e começou a rearranjar os pertences dela para ajustar mais nitidamente na cesta de bicicleta, Elisabeth Danziger achou uma barra de sabão de flor-de-lis e um cubo de rosa gerânio banho cristais na mão dela.
Sistemas que traduziram corretamente: Nenhum	
Comentários:	
<b>3. I hated it, and Dr Kelleher <b>got</b> some of them taken down so that I could at least see the daylight.</b>	
Acepção correta: Algo como “ter” (algum deles rebaixados) – Fulano teve algum deles rebaixados. Se o “them” forem as caixas que estavam na frente da janela, a tradução mais adequada seria: “Fulano abaixou/rebaixou/desceu (ou até retirou) algumas delas (as caixas).”	
FreeTranslation	Odiei-o, e Kelleher de Dr recebeu algum deles tomados para baixo de modo que podia pelo menos vejo a luz do dia.
Systran	Eu odiei-o, e o Dr. Kelleher começou alguma deles feita exame para baixo de modo que eu pudesse ao menos ver a luz do dia.
TranslatorPro	Eu odiei isto, e Dr Kelleher adquiriu alguns deles tirados de forma que mim poderiam ver a luz do dia pelo menos.
Sistemas que traduziram corretamente: Nenhum	
Comentários:	
<b>4. Our death list <b>got</b> shorter.</b>	
Acepção correta: diminuir – diminuiu / ficou mais curta / encurtou...	
FreeTranslation	Nossa lista de morte ficou mais curto.
Systran	Nossa lista da morte começou mais curta.
TranslatorPro	Nossa lista de morte se pôs mais curta.
Sistemas que traduziram corretamente: FreeTranslation	
Comentários: Não é um phrasal verb.	
<b>5. I was quite sick on the train before we <b>got</b> to Euston, and have not liked long train journeys since.</b>	
Acepção correta: chegar – chegamos	
FreeTranslation	Estava bastante doente no trem antes de nós recebeu a Euston, e não gostou de viagens longas de trem desde que.
Systran	Eu fui completamente doente no trem antes que nós começamos a Euston, e não tenho gostado de viagens longas do trem desde.
TranslatorPro	Eu estava bastante doente no trem antes de nós adquirimos a Euston, e não gostou trem longo viaja desde então
Sistemas que traduziram corretamente: Nenhum	
Comentários:	
<b>GOT – VVN</b>	
<b>1. They were often together trying to think how the babies had <b>got</b> into this state.</b>	
Acepção correta: ficar – ficaram / tinham ficado	
FreeTranslation	Eram tentar freqüentemente junto pensar como que os bebês tinha feito-se parte de este estado.
Systran	Freqüentemente estavam tentando junto pensar de como os bebês tinham começado neste estado.
TranslatorPro	Eles estavam tentando freqüentemente junto para pensar como os bebês tiveram neste estado.
Sistemas que traduziram corretamente: Nenhum	
Comentários:	
<b>2. A whole load of French soldiers have <b>got</b> away from Dunkirk too, and we have got to take the lot.</b>	
Acepção correta: afastar – afastaram-se / tinham se afastado	
FreeTranslation	Uma carga inteira de soldados franceses escapou de Dunkirk demais, e nós recebemos tomar o lote.

Systran	Uma carga inteira de soldados franceses começou afastado de Dunkirk demasiado, e nós começamos fazer exame do lote.
TranslatorPro	Uma carga inteira de soldados franceses também tem longe de Dunkirk, e nós temos que levar o lote.
Sistemas que traduziram corretamente: Nenhum	
Comentários:	
<b>3. Here, I've got you a paper to prove it.</b>	
Acepção correta: conseguir – consegui	
FreeTranslation	Aqui, recebi-o um papel provaá-lo.
Systran	Aqui, eu have got o um papel para prová-lo.
TranslatorPro	Aqui, eu o tenho um papel provar isto.
Sistemas que traduziram corretamente: Nenhum	
Comentários:	
<b>4. When they had got beyond Stamford and to their destination, the wolf suddenly disappeared and the flames on the lances were extinguished.</b>	
Acepção correta: vejamos: ir / foram	
FreeTranslation	Quando tinham recebido além de Stamford e a seu destino, o lobo repentinamente desapareceu e as chamas nas lanças foram extinguidas.
Systran	Quando tinham começado além de Stamford e a seu destino, o lobo desapareceu de repente e as flamas nos lances foram extinguidas.
TranslatorPro	Quando eles tiveram além de Stamford e ao destino deles/delas, desapareceu o lobo de repente e as chamas nas lanças eram extinguidas.
Sistemas que traduziram corretamente: Nenhum	
Comentários: O got tem dois complementos: beyond e to. Assim: “Quando foram além / passaram por Stamford e foram para o seu destino....”	
<b>5. The second half has got off to a good start, with slightly higher orders for October.</b>	
Acepção correta: get off – sair-se com – saiu-se com	
FreeTranslation	A segunda metade desceu a um bom começa, com ordens levemente mais altas durante outubro.
Systran	A segunda metade tem fora a um começo bom, com ordens ligeiramente mais elevadas para outubro.
TranslatorPro	O segundo que meio tem fora para um começo bom, com ordens ligeiramente mais altas durante outubro.
Sistemas que traduziram corretamente: Nenhum	
Comentários:	
<b>GETTING – VVG</b>	
<b>1. What a pipe dream, we thought, as many children were getting no education at all.</b>	
Acepção correta: receber – recebendo	
FreeTranslation	O que um sonho de cano, nós pensamos, tantas crianças não recebiam nenhuma educação absolutamente.
Systran	Que sonho da tubulação, nós pensamos, tantas como crianças não começavam nenhuma instrução em tudo.
TranslatorPro	O que um sonho de tubo, nós pensamos, como muitas crianças não estavam adquirindo nenhuma educação nada.
Sistemas que traduziram corretamente: FreeTranslation	
Comentários:	
<b>2. My parents wrote to say that they were still getting some quite nasty tip-and-run raids.</b>	
Acepção correta: participar – participando / fazer – fazendo	
FreeTranslation	Meus pais escreveram dizer que eles ainda recebiam alguma ponta bastante suja-e-ataques repentinos de corrida.
Systran	Meus pais escreveram para dizer que começavam ainda algum completamente nasty derrub-e-funcionam invasões.
TranslatorPro	Meus pais escreveram para dizer que eles ainda estavam adquirindo alguns invasões gorjeta-e-corridas bastante sórdidas.
Sistemas que traduziram corretamente: Nenhum	
Comentários:	
<b>3. He then seduced the daughter of a respectable tradesman and is supposed to have deserted her after getting her pregnant.</b>	
Acepção correta: deixar – deixá-la grávida (engravidar)	
FreeTranslation	Ele então seduziu a filha de um tradesman respeitável e é suposto ter desertado-a depois de receber seu grávido.
Systran	Seduziu a filha de um tradesman respeitável e é suposto então para deserted a após ter começado a grávida.

TranslatorPro	Ele seduziu a filha de um negociante respeitável então e é suposto que tem a abandonado depois da adquirir grávida.
Sistemas que traduziram corretamente: Nenhum	
Comentários:	
4. “Not the prince,” Anne corrected, <b>getting</b> her own back.	
Acepção correta: pegar (de volta) – pegando (de volta)	
FreeTranslation	“Não o príncipe,” Anne corrigiu, recebendo o próprias costas.
Systran	"não o príncipe," Anne corrigido, começando seus próprios para trás.
TranslatorPro	"Não o príncipe", a Anne corrigiu, enquanto adquirindo o próprio atrás dela.
Sistemas que traduziram corretamente: Nenhum	
Comentários:	
5. After breaking the ice last month he was desperately unlucky in an amateur riders' race having to be switched in the final furlong after <b>getting</b> off on the rails.	
Acepção correta: sair – ter saído	
FreeTranslation	Depois que quebrar o gelo último mês ele estava desesperadamente desafortunado numa corrida dos cavaleiros de amador ter que estar trocado no furlong final depois de descer nas barras.
Systran	Após ter quebrado o gelo último mês era desesperadamente unlucky em uma raça dos cavaleiros amadores que tem que ser comutada no furlong final após começar fora nos trilhos.
TranslatorPro	Quebrando o gelo mês passado ele buscou desesperadamente azarado na raça de uns cavaleiros amadores que tem que ser trocado no furlong final depois de descer nas grades.
Sistemas que traduziram corretamente: Nenhum	
Comentários:	

De modo geral, a análise realizada pela tradutora humana permitiu identificar os possíveis sentidos (isto é, traduções) de cada verbo, considerando-se a amostra de sentenças selecionada. Por exemplo, para as sentenças da Tabela 2, os sentidos são indicados na Tabela 5.

Tabela 5 – Sentidos dos verbos nas sentenças selecionadas da Tabela 2

Sentença	Etiqueta do verbo	Sentido
<b>Be</b> so good as to enter.	VBO	ser
<b>Be</b> ready to stake any which show signs of flopping.	VBO	estar
I <b>am</b> 18 years old.	VBM	ter
And he warned that if there <b>was</b> not an honourable settlement, that would explode the whole region.	VBDZ	haver
Most nuns <b>come</b> from the Tibetan countryside and may join the nunnery in their early teens or even when they are as young as nine or 10.	VVO	vir
Well he <b>come</b> into garage the other day and told me his wife had kicked him out!	VVD	entrar
“This city has suddenly <b>come</b> alive,” said her husband, an off-duty border guard.	VVN	renascer
Mr Gonzalez has also <b>come</b> in for criticism from within his own party.	VVN	receber
It's best to be alone when the noises <b>get</b> this loud.	VVO	ficar
“You <b>get</b> three clear benefits out of this,” he explained.	VVO	adquirir
Policemen struggled briefly with the women as they tried to <b>get</b> into the embassy.	VVI	entrar
I was quite sick on the train before we <b>got</b> to Euston, and have not liked long train journeys since.	VVD	chegar
Let those who <b>give</b> us such advice sign it themselves.	VVO	dar
He arrived late yesterday afternoon, it seems, and he has offered to <b>give</b> a little recital to anyone on the island who wishes to attend.	VVI	fazer
The queen broached the matter to me and ostensibly I am now <b>giving</b> thought to my answer — what would you say to going in my place?	VVG	considerar
These systems work well if they <b>know</b> the parameters of Soviet radar very well.	VVO	saber
But I suppose you're right: the public isn't going to notice and those who <b>know</b> me.	VVO	conhecer

No caso do verbo *to come*, por exemplo, como se pode observar na Tabela 5, há várias traduções diferentes nas sentenças da amostra, incluindo “vir”, “entrar”, “renascer” e “receber”.

Vale notar que foram consideradas acepções possíveis aquelas relativas ao uso do verbo em *phrasal verbs* (por exemplo, “*get up*”) ou em construções e expressões nas quais ele não pode ser traduzido independentemente das demais palavras da expressão, uma vez que, isoladamente, não possui sentido algum. Por exemplo, nas construções com o verbo *to be* juntamente com *there*, denotando o verbo “haver” (*there is, there are, there was, etc.*), e na expressão “*I am 15 years old*”. Por outro lado, não são consideradas acepções possíveis aquelas provenientes de expressões de sentido figurado ou do uso dos verbos na função de auxiliares, como *to be, to do, to go (going to...)* e *to have*. No caso do uso de verbos como auxiliares, concentrou-se a análise na verificação da capacidade do sistema de diferenciar o uso do verbo como principal ou auxiliar (somente nos casos em que é o verbo principal, verificou-se sua habilidade em traduzi-lo adequadamente), e não na sua capacidade de identificação das acepções específicas decorrentes da junção do auxiliar com algum verbo principal.

Além das diversas acepções de cada verbo na amostra, a análise permitiu verificar se os tradutores eram capazes de identificar tais acepções nas diferentes sentenças. No caso do verbo *to get*, pode-se verificar, na Tabela 4, que a maioria das suas acepções não foi traduzida corretamente pelos sistemas. Na maior parte dos casos (exceto nas *phrasal verbs*) cada sistema utiliza uma das possíveis traduções do verbo em todas as sentenças. Normalmente, o sistema FreeTranslation traduz o verbo com a acepção “receber”, o Systran, com a acepção “começar”, e o TranslationPro, com a acepção “adquirir”. Com isso, em algumas sentenças a tradução do sistema coincidiu com a mais adequada, contudo, como este verbo é altamente ambíguo, na maioria dos casos, a tradução foi inadequada. Mesmo *phrasal verbs* comuns contendo o verbo *to get*, como *get back* (voltar) e *get into* (entrar), não foram traduzidos corretamente por todos os sistemas.

Em um exame superficial a partir dessa análise, foi possível verificar que alguns verbos apresentavam pouca ou nenhuma ambiguidade nas sentenças selecionadas. Para os verbos ambíguos, em alguns casos, todos os sistemas geraram traduções corretas, enquanto em outros, nenhum ou apenas alguns tradutores levaram a essas traduções. A compilação da análise, visando chegar a resultados conclusivos do experimento, bem como comentários sobre alguns verbos, são apresentados a seguir.

## 4.5 Compilação e resultados da análise

Conforme mencionado, o objetivo da análise das sentenças selecionadas e das suas respectivas traduções, realizadas pelos sistemas escolhidos, era verificar a ocorrência de ambigüidade lexical de sentido no conjunto de verbos em foco e também qual era o tratamento dado pelos sistemas de tradução para esse fenômeno, de modo a identificar o subconjunto dos verbos de fato problemáticos, entre os 15 pré-selecionados, sob o aspecto da ambigüidade lexical de sentido na TA. Para tanto, as análises das traduções do conjunto de sentenças de cada verbo nas tabelas descritas na Seção 4.5 foram minuciosamente examinadas. Os seguintes critérios foram estabelecidos para a identificação dos verbos considerados “problemáticos”.

- 1) Verbos com mais de uma possível acepção identificada pela tradutora humana no conjunto de sentenças da amostra do *corp*us BNC foram considerados potencialmente problemáticos. Apesar de constarem diversas acepções para todos os verbos em dicionários bilíngües (conforme mencionado na Seção 4.1), nem todas essas possíveis acepções foram verificadas na amostra. Isso pode ter ocorrido devido à quantidade relativamente pequena de sentenças da amostra ou ao uso pouco freqüente do verbo com certas acepções.
- 2) Para cada ocorrência do verbo, considerou-se que sua ambigüidade não era tratada adequadamente pelos sistemas nas sentenças para as quais pelo menos dois sistemas não haviam sido capazes de traduzir corretamente o verbo, de acordo com a acepção identificada pela tradutora humana ou sugerida como “aceitável”.
- 3) Sentenças originais cuja acepção do verbo em foco não pôde ser identificada pela tradutora humana foram desconsideradas da análise. Algumas sentenças em inglês apresentavam problemas na sua construção, provavelmente provenientes de erros de interpretação dos textos falados, inviabilizando a identificação da acepção do verbo. Ainda, em algumas

- sentenças, a identificação não foi possível em função da vagueza dessas sentenças, observadas isoladamente do seu contexto.
- 4) Sentenças originais cujo verbo em foco era utilizado com sentido figurado foram desconsideradas da análise
  - 5) Um verbo foi considerado de fato problemático, sob o aspecto da ambigüidade lexical de sentido, caso essa ambigüidade não recebesse tratamento adequado na sua tradução na maioria das sentenças da amostra contendo tal verbo.

Considerando-se esses parâmetros, os números da Tabela 6 foram levantados a partir da análise das traduções das sentenças dos 15 verbos.

Tabela 6 – Compilação da análise da ambigüidade dos verbos

Verbo	Nº de sentenças da amostra	Nº de sentenças desconsideradas <sup>9</sup>	Nº de acepções na amostra <sup>10</sup>	Nº de sentenças problemáticas <sup>11</sup>
<i>to be</i>	51	1	8	7 (14%)
<i>to have</i>	39	1	3	3 (7.9%)
<i>to do</i>	31	2	4	5 (17.2%)
<i>to say</i>	30	0	1	1 (3.3%)
<i>to go</i>	37	0	22	23 (62.2%)
<i>to get</i>	30	0	18	29 (96.7%)
<i>to make</i>	31	0	10	19 (61.3%)
<i>to see</i>	41	4	1	0
<i>to know</i>	30	0	2	3 (10%)
<i>to take</i>	40	0	19	33 (82.5%)
<i>to think</i>	30	0	3	9 (30%)
<i>to come</i>	42	4	15	19 (50%)
<i>to give</i>	33	0	11	11 (33.3%)
<i>to look</i>	30	1	8	15 (50%)
<i>to use</i>	36	0	3	7 (19.4%)

Com base nos critérios estabelecidos, em uma primeira análise, foram considerados verbos problemáticos os seguintes: *to go*, *to get*, *to make*, *to take*, *to come* e *to look*. Entretanto, em alguns casos, apesar do verbo ter sido considerado problemático em uma parcela menor das sentenças, foi possível verificar que a influência das traduções utilizando acepções inadequadas na qualidade da tradução da sentença era bastante negativa. Por essa razão, foi realizada uma análise mais aprofundada, considerando, além dos critérios citados, os efeitos das escolhas incorretas na tradução das sentenças. Para guiar a discussão sobre os resultados dessa análise, as tabelas a seguir mostram as acepções identificadas na amostra do corpúsculo para cada verbo, juntamente com suas respectivas frequências (isto é, com a quantidade de sentenças em que cada acepção ocorre) e com o número de sentenças problemáticas de cada acepção (isto é, o número de sentenças nas quais pelo menos dois sistemas não identificaram a acepção correta). As observações relevantes sobre cada verbo são relatadas após cada tabela.

Tabela 7 – Detalhamento da análise do verbo *to be*

<i>to be</i>	Acepção	Frequência da acepção	Nº Sentenças Problemáticas
	ser	29	1
	estar	10	4
	ter (idade)	1	0

<sup>9</sup> Sentenças nas quais não foi possível para a tradutora humana identificar a acepção adequada ou cuja acepção derivava de uma expressão de sentido figurado.

<sup>10</sup> Excluindo-se, conforme mencionado, as diferentes acepções provenientes do uso do verbo como auxiliar (todas as ocorrências são consideradas como uma única acepção – auxiliar).

<sup>11</sup> O percentual de sentenças problemáticas indicado aqui considera somente o total de sentenças efetivamente analisado, isto é: Nº de sentenças da amostra - Nº de sentenças desconsideradas.

	haver ( <i>there are</i> )	4	0
	auxiliar (ex.: <i>have been fighting</i> )	3	0
	expressão “dever”, “esperar-se que” (ex.: <i>I am to believe...</i> )	3	2

A maioria dos problemas com o verbo *to be* (Tabela 7) consistem de traduções do verbo como “estar” em vez de “ser”. O contrário, ou seja, a tradução do verbo como “ser”, em vez de “estar” ocorre somente uma vez. Os demais problemas ocorrem quando do uso do verbo em expressões, as quais não são reconhecidas pelos sistemas de tradução. Na maior parte dos casos, entretanto, as traduções incorretas não impedem a compreensão da sentença ou modificam significativamente o seu significado. Além disso, somente 14% das sentenças apresentaram esses problemas. Por essas razões, esse verbo não foi considerado problemático.

Tabela 8 – Detalhamento da análise do verbo *to have*

<i>to have</i>	Acepção	Frequência da acepção	Nº Sentenças Problemáticas
	ter	26	1
	auxiliar (ex.: <i>have done</i> )	11	1
	expressão “tomar o chá da tarde” ( <i>have afternoon tea</i> )	1	1

Pode-se perceber que o uso do verbo *to have* (Tabela 8) ocorre predominantemente com a acepção “ter” e que praticamente não há problemas de tradução considerando-se essa acepção. Da mesma maneira, na grande maioria dos casos, os sistemas conseguem identificar o uso do verbo como auxiliar. Na única expressão encontrada na amostra de sentenças, o verbo *to have* é traduzido literalmente como “ter” ou considerado auxiliar. Com isso, apenas 7.9 % das sentenças apresentaram problemas de tradução e o verbo não foi considerado problemático.

Tabela 9 – Detalhamento da análise do verbo *to do*

<i>to do</i>	Acepção	Frequência da acepção	Nº Sentenças Problemáticas
	fazer	14	1
	ir	1	1
	auxiliar (ex.: <i>don't tell me</i> )	13	2
	expressão “ter a ver” ( <i>have to do</i> )	1	1

Os sistemas de tradução apresentam problemas de tradução no caso do verbo *to do* (Tabela 9) na identificação da sua utilização como auxiliar e no seu uso em acepções pouco comuns ou expressões, totalizando apenas 17.2% das sentenças. Esse verbo não foi considerado problemático.

Tabela 10 – Detalhamento da análise do verbo *to say*

<i>to say</i>	Acepção	Frequência da acepção	Nº Sentenças Problemáticas
	dizer	30	1

O único problema de tradução do verbo *to say* (Tabela 10) ocorreu no uso desse verbo no gerúndio (3.3 % das sentenças), isto é, *saying*, que foi traduzido como “provérbio”, “ditado” ou “declaração”. Por isso, ele não foi considerado problemático.

Tabela 11 – Detalhamento da análise do verbo *to go*

<i>to go</i>	Acepção	Frequência da acepção	Nº Sentenças Problemáticas
	ir	9	1
	permanecer	1	1
	dirigir-se	1	1
	avançar	1	1



ir-se	1	1
dar (ex.: <i>go wrong</i> )	1	1
passar ( <i>go by</i> )	2	1
continuar / prosseguir ( <i>go on</i> )	2	2
participar ( <i>go into</i> )	1	1
entrar / dar entrada ( <i>go into</i> )	2	0
sair ( <i>go out</i> )	2	1
passar por / atravessar ( <i>go through</i> )	1	0
cruzar ( <i>go accross</i> )	1	1
ocorrer ( <i>go on</i> )	1	1
voltar ( <i>go back</i> )	1	0
auxiliar do futuro imediato (ex.: <i>going to fail</i> )	5	5
expressão “ser aceito” ( <i>go through</i> )	1	1
expressão “sair como o planejado” ( <i>go to plan</i> )	1	1
expressão “deixar de ser” ( <i>go from being</i> )	1	1
expressão “chegar ao ponto” ( <i>go on the state</i> )	1	1
expressão “interessar-se por” ( <i>go in for</i> )	1	1

Os problemas de tradução do verbo *to go* (Tabela 11) foram bastante variados, concentrando-se especialmente no uso desse verbo como auxiliar do futuro imediato, traduzido corretamente por apenas um sistema. Os demais, como mostra como o exemplo da Tabela 12, realizaram a tradução como futuro simples ou não identificaram o uso do verbo como auxiliar.

Tabela 12 – Exemplo de problema com o uso do verbo *to go* no futuro imediato

Simpler systems don't attempt as much, but are not <b>going</b> to fail so badly," said an analyst with the Centre for Defence Information, Mr Steve Kosiak.	
Acepção correta: não vão falhar	
FreeTranslation	Sistemas mais simples não tentam tanto, mas não fracassarão tão mal," disse um analista com o Centro para Informação de Defesa, Sr Kosiak de Steve.
Systran	Uns sistemas mais simples não tentam tanto quanto, mas não estão indo falhar assim mal, "disse um analista com o centro para a informação da defesa, Sr. Steve Kosiak.
TranslatorPro	Sistemas mais simples não tentam como muito, mas não vai falhar tão mal", disse um analista com o Centro para Informação de Defesa, Mr Steve Kosiak.
Sistemas que traduziram corretamente: TranslatorPro	

Outros problemas ocorreram tanto com o uso do verbo isoladamente quanto em *phrasal verbs* ou expressões, conforme mostram os exemplos na Tabela 13. Considerando-se a quantidade de sentenças que apresentaram problemas na tradução do verbo *to go* (62.2%), ele foi considerado problemático.

Tabela 13 – Exemplos de problemas com o uso do verbo *to go* em *phrasal verbs* ou expressões

The war may well just <b>go</b> on and on.	
Acepção correta: continuar	
FreeTranslation	A guerra bem somente pode ir em e em.
Systran	A guerra pode jorrar apenas vai sobre e sobre.
TranslatorPro	A guerra pode ir bem há pouco sem parar.
Sistemas que traduziram corretamente: Nenhum	
Swapo <b>went</b> into the election with a symbol which was only weeks old: a man with a raised fist.	
Acepção correta: participar - participou	
FreeTranslation	O Swapo entrou na eleição com um símbolo que era só semanas velho: homem com um punho levantado.
Systran	Swapo entrou na eleição com um símbolo que fosse somente semanas velhas: um homem com um punho levantado.
TranslatorPro	Swapo entrou na eleição com um símbolo que era só semanas velho: um homem com um punho elevado.
Sistemas que traduziram corretamente: Nenhum	
We have <b>gone</b> from being the bread basket of Central America to being its basket case; an international beggar with the largest per capita foreign debt in Latin America.	
Acepção correta: deixar de ser – deixamos de ser	

FreeTranslation	Fomos de ser a cesta de pão de América Central a ser seu caso de cesta; um mendigo internacional com o maior dívida estrangeira per capita na américa latina.
Systran	Nós fomos de ser a cesta do pão de América Central a ser sua caixa da cesta
TranslatorPro	Nós fomos de ser a cesta de pão de América Central a ser seu caso de cesta; um mendigo internacional com a dívida externa per capita maior na América Latina.
Sistemas que traduziram corretamente: Nenhum	

Tabela 14 – Detalhamento da análise do verbo *to get*

	Acepção	Frequência da acepção	Nº Sentenças Problemáticas
	<i>to get</i>	ficar / tornar-se	5
receber		4	4
conseguir		2	2
começar		1	1
obter		1	1
fazer		1	1
colocar-se		1	1
ter		1	1
chegar		1	1
pegar		1	1
participar		1	1
entrar ( <i>get into, get through</i> )		2	2
sair ( <i>get off</i> )		2	2
ir (ex.: <i>get outside, get beyond</i> )		2	2
afastar ( <i>get away</i> )		1	1
voltar ( <i>get back</i> )		1	1
deixar (ex.: <i>get her pregnant</i> )		1	1
expressão “ter algo feito” (ex.: <i>get this underbrush cleaned</i> )		2	2

O verbo *to get* (Tabela 14) é altamente ambíguo e apresentava acepções bastante diversas nas sentenças da amostra. Contudo, conforme mencionado, os três sistemas empregam a mesma acepção para traduzi-lo na maioria das sentenças (“receber”, “começar” e “adquirir”, respectivamente), como pode ser verificado pelos exemplos já ilustrados na Tabela 5. Mesmo no caso de *phrasal verbs* comuns, os sistemas não foram capazes de identificar a acepção correta. Isso resultou na tradução incorreta do verbo em 96.7 % das sentenças e o ele foi, portanto, considerado problemático.

Tabela 15 – Detalhamento da análise do verbo *to make*

	Acepção	Frequência da acepção	Nº Sentenças Problemáticas
	<i>to make</i>	fazer	11
tornar		7	7
ter		2	2
deixar		4	4
dar		1	1
pegar		1	1
ir		1	1
compensar ( <i>make up</i> )		1	0
representar / compor ( <i>make up</i> )		2	1
expressão “alertar” ( <i>make aware</i> )		1	1

O verbo *to make* foi traduzido por todos os sistemas, na maioria das vezes, como sua acepção mais comum, “fazer”, o que resultou em traduções incorretas do verbo em 61.3% das sentenças, conforme os exemplos da Tabela 16. Por essa razão, ele foi considerado problemático.

Tabela 16 – Exemplos de problemas com o uso do verbo *to make*

This way you <b>make</b> the song peculiarly Jewish.	
Acepção correta: tornar – torna	
FreeTranslation	Este meio você faz a canção peculiarmente Judeu.
Systran	Esta maneira você faz a canção peculiar jewish.
TranslatorPro	Deste modo você faz a canção peculiarmente judeu.
Sistemas que traduziram corretamente: Nenhum	
“They won't like this but I don't think they will <b>make</b> a violent reaction”, he said.	
Acepção correta: ter – ter	
FreeTranslation	“Eles não gostarão deste mas eu não penso que eles farão uma reação violenta”, disse.
Systran	"não gostarão deste mas eu não penso que farão uma reação violenta", disse.
TranslatorPro	"Eles não gostarão isto mas eu não penso que eles farão uma reação violenta", ele disse.
Sistemas que traduziram corretamente: Nenhum	

Tabela 17 – Detalhamento da análise do verbo *to see*

<i>to see</i>	Acepção	Frequência da acepção	Nº Sentenças Problemáticas
	ver	41	4

As quatro sentenças que apresentaram problemas com a tradução do verbo *to see* (Tabela 17) usavam esse verbo na acepção de “serrar”, cuja forma do presente simples no inglês (*saw*) coincide com a forma do passado simples do verbo *see*. Os sistemas de tradução, porém, não identificaram tal acepção, traduzindo o verbo na acepção “ver”. Apesar de objetivo ter sido selecionar apenas sentenças com flexões do verbo *to see*, como a seleção foi aleatória, as quatro sentenças foram incluídas. A diferenciação entre os sentidos, neste caso, constitui um problema de ambigüidade lexical de sentido, mas não na TA, pois a acepção deveria ser identificada na própria língua (no inglês, neste caso), em um processo de lematização, antes da tradução. Por isso, essas sentenças não foram consideradas nos resultados da análise.

Todas as sentenças restantes utilizavam o verbo na sua acepção mais comum, “ver”, e tal acepção foi corretamente identificada pelos tradutores. Por essa razão, o verbo não foi considerado problemático.

Tabela 18 – Detalhamento da análise do verbo *to know*

<i>to know</i>	Acepção	Frequência da acepção	Nº Sentenças Problemáticas
	saber	25	0
	conhecer	5	3

Nas sentenças da amostra, o verbo *to know* (Tabela 18) é utilizado com apenas duas acepções, sendo que, na maioria das vezes, a acepção correta é “saber”. O único problema encontrado foi a tradução de três ocorrências do verbo na acepção “conhecer” como “saber”, o que não representou prejuízo significativo para a tradução, já que o significado dessas acepções é bastante próximo, como mostram os exemplos na Tabela 19. Além disso, esses casos representam apenas 10% das sentenças. Assim, o verbo *to know* não foi considerado problemático.

Tabela 19 – Exemplos de problemas com o uso do verbo *to know*

She <b>knew</b> them all, she was devastated for them and their families, who would be left husbandless and fatherless.	
Acepção correta: conhecer – conhecia	
FreeTranslation	Soube-os todo, foi devastada para eles e as suas famílias, quem seriam deixado husbandless e fatherless.
Systran	Soube-os todos, ela devastated para eles e suas famílias, que seriam deixados husbandless e fatherless.
TranslatorPro	Ela os conheceu tudo, ela foi devastada para eles e as famílias deles/delas que seriam husbandless esquerdo e órfão.
Sistemas que traduziram corretamente: TranslatorPro	
BRAZIL could be within a week of electing Latin America's first worker President, Mr Luis Inacio da Silva,	

commonly <b>known</b> as Lula.	
Acepção correta: conhecer – conhecido	
FreeTranslation	BRASIL pode ser dentro de uma semana de eleger a américa latina primeiro Presidente de trabalhador, Sr Silva de da de Inacio de Luis, Lula comumente sabido.
Systran	BRASIL podia realizar-se dentro de uma semana presidente latin do trabalhador de América elegendo do primeiro, silva do da do Sr. Luis Inacio, sabido geralmente como Lula.
TranslatorPro	BRASIL poderia estar dentro de uma semana de eleger o primeiro trabalhador President de América Latina, Mr Luis Inácio da o Silva, geralmente conhecido como Lula.
Sistemas que traduziram corretamente: TranslatorPro	

Tabela 20 – Detalhamento da análise do verbo *to take*

	Acepção	Frequência da aceção	Nº Sentenças Problemáticas
	<i>to take</i>	ter	2
fazer		2	2
tomar		4	4
levar		7	7
ir		1	1
ficar		1	1
considerar		2	2
dado		1	1
aceitar		1	1
tirar ( <i>take out</i> )		3	2
tirar (fotos)		1	1
acontecer ( <i>take place</i> )		1	0
assumir (responsabilidades) ( <i>take on</i> )		2	2
assumir (controle, comando) ( <i>take over</i> )		4	1
retomar / aprender / dedicar-se a ( <i>take up</i> )		2	2
tomar parte / participar ( <i>take part</i> )		1	1
decolar ( <i>take off</i> )		2	2
afastar / remover / retirar ( <i>take away</i> )	1	1	

O verbo *to take* (Tabela 20) é altamente ambíguo e foi utilizado nas sentenças da amostra com muitas acepções diferentes, as quais não foram corretamente identificadas na maioria das vezes (em 82.5% das sentenças). A não ser no caso de alguns *phrasal verbs*, os três sistemas optaram por uma única tradução em todas as sentenças (“tomar”, “fazer exame” e “levar”, respectivamente), como mostram os exemplos na Tabela 21. Esse verbo foi, portanto, considerado problemático.

Tabela 21 – Exemplos de problemas com o uso do verbo *to take*

If you <b>take</b> a balanced view, variegated plants — apart from the odd horrendous freak — contribute enormously to awell-planted garden in lots of different ways.	
Acepção correta: fazer – fizer	
FreeTranslation	Se toma uma vista equilibrada, variegou plantas — à parte do ímpar horrendo esquisito — contribui enormemente a jardim awell-plantado em lotes de meios diferentes.
Systran	Se você fizer exame de uma vista equilibrada, as plantas variegated -- aparte do freak horrendous impar -- contribuem enormemente ao jardim awell-plantado nos lotes de maneiras diferentes.
TranslatorPro	Se você leva uma visão equilibrada, plantas matizadas - aparte da extravagância horrenda estranha - contribua enormemente para jardim awell-plantado de muitos modos diferentes.
Sistemas que traduziram corretamente: Systran	
I am taking a course of antibiotics and continually <b>take</b> anti-histamine tablets.	
Acepção correta: tomar – tomo	
FreeTranslation	Tomo um curso de antibiótica e continuamente toma tabelas de anti-histamina.
Systran	Eu estou fazendo exame de um curso dos antibióticos e faço exame continuamente de tabletas do anti-histamine.
TranslatorPro	Eu estou levando um curso de antibióticos e continuamente levo tabletes de anti-histamine!
Sistemas que traduziram corretamente: FreeTranslation	

<b>Take</b> these to your stylist, safe in the knowledge that your new look is going to suit you!	
Acepção correta: levar – leve	
FreeTranslation	Tomem estes a seu estilista, cofre no conhecimento que seu novo olha servirá-o!
Systran	Faça exame destes a seu stylist, seguro no conhecimento que seu olhar novo está indo o servir!
TranslatorPro	Leve estes a seu estilista, seguro no conhecimento que seu olhar novo vai o vestir!
Sistemas que traduziram corretamente: TranslatorPro	

Tabela 22 – Detalhamento da análise do verbo *to think*

<i>to think</i>	Acepção	Frequência da acepção	Nº Sentenças Problemáticas
	pensar	21	0
	achar	6	6
	acreditar	3	3

Os sistemas traduziram todas as ocorrências do verbo *to think* (Tabela 22) como “pensar”, não considerando as acepções “achar” ou “acreditar”, também encontradas na amostra, levando a 30% das sentenças com traduções incorretas do verbo, conforme mostram os exemplos na Tabela 23. Contudo, como a diferença entre essas acepções é bastante sutil, o verbo não foi considerado problemático.

Tabela 23 – Exemplos de problemas com o uso do verbo *to think*

Congress was banking on at least 12 seats here, but most observers <b>think</b> they will be lucky to get 10.	
Acepção correta: achar – acham	
FreeTranslation	O congresso depositava em pelo menos 12 assentos aqui, mas a maioria de observadores pensa que eles estará afortunado receber 10.
Systran	O congress estava depositando ao menos em 12 assentos aqui, mas a maioria de observadores pensam que serão afortunados começar 10.
TranslatorPro	Congresso estava contando com pelo menos 12 assentos aqui, mas a maioria dos observadores pensa que eles terão sorte para adquirir 10.
Sistemas que traduziram corretamente: Nenhum	
“This office has been so discredited that I <b>think</b> it has to be disbanded,” the party's new reformist leader, Dr Gregor Gysi, said at the weekend.	
Acepção correta: achar – acho / a creditar – acredito	
FreeTranslation	“Este escritório foi então discredited que eu penso que tem que ser separado,” o novo líder reformista do partido, Gysi de Gregor de Dr, dito no fim de semana.
Systran	"este escritório foi assim que desacreditado que eu penso lhe tem que disbanded," o líder reformista novo do partido, Dr. Gregor Gysi, dito no fim de semana.
TranslatorPro	"Este escritório foi desacreditado assim que eu penso que tem que ser licenciado", o líder de reformista novo da festa, Dr Gregor Gysi, disse ao fim de semana.
Sistemas que traduziram corretamente: Nenhum	

Tabela 24 – Detalhamento da análise do verbo *to come*

<i>to come</i>	Acepção	Frequência da acepção	Nº Sentenças Problemáticas
	vir	21	4
	ir (ex.: <i>I am coming</i> )	1	1
	voltar ( <i>come back</i> )	1	0
	entrar ( <i>come into</i> )	1	1
	sair ( <i>come off</i> )	2	2
	retirar ( <i>come away</i> )	1	1
	surgir ( <i>come up</i> )	2	1
	renascer ( <i>come alive</i> )	1	1
	cair ( <i>come down</i> )	1	1
	aproximar-se ( <i>come close</i> )	1	1
	receber ( <i>come in for</i> )	1	1

	chegar ( <i>come up</i> )	1	1
	aparecer ( <i>come around</i> )	1	1
	avançar ( <i>come forward</i> )	1	1
	expressão “subir ao poder” ( <i>come to power</i> )	2	2

O verbo *to come* possui muitas acepções, principalmente quando utilizado em *phrasal verbs*. Nas sentenças da amostra, o verbo foi utilizado, na maioria das vezes, na sua acepção mais comum “vir”, sendo que em quatro sentenças essa acepção não foi corretamente identificada. Isso ocorreu nos casos em que havia alguma preposição na sentença, com outra função, mas a acepção o verbo era interpretada considerando-se também a preposição, como no exemplo da Tabela 25, com a preposição *up*.

Tabela 25 – Exemplo de problema com o uso do verbo *to come* seguido de preposição

At this moment about 11 people <b>came</b> up to me with two dogs.	
Acepção correta: vir – vieram	
FreeTranslation	Nesta pessoas de aproximadamente 11 de momento surgiu a mim com dois cachorros.
Systran	Neste momento aproximadamente 11 povos veio até mim com dois cães.
TranslatorPro	Neste momento aproximadamente 11 pessoas me propuseram dois cachorros.
Sistemas que traduziram corretamente: Systran	

Nas demais sentenças, incluindo aquelas com o uso de *phrasal verbs*, o verbo foi traduzido de maneira bastante inadequada, considerando-se apenas, em grande parte dos casos, a acepção mais comum (“vir”), como mostram os exemplos na Tabela 26. No total, 50% das sentenças tiveram o verbo *to come* traduzido incorretamente. Com base nessa quantidade e também no nível de discrepância entre as acepções reais e as identificadas pelos sistemas, o verbo foi considerado problemático.

Tabela 26 – Exemplos de problemas com o uso do verbo *to come*

Anyway they were needed beds, so they sent her home, and er, you know er see that stuff that <b>come</b> away, you see there were still some stuff left to <b>come</b> away.	
Acepção correta: retirar – foram retiradas	
FreeTranslation	De qualquer jeito eles foram necessitados camas, então enviaram seu lar, e er, você sabem que er vê aquele material que vem longe, vê havia ainda alguma esquerda de material vir longe.
Systran	Em todo o caso eram camas needed, assim que emitiram seu repouso, e er, você sabe er veja esse material que vem afastado, você vêem que havia ainda algum material à esquerda a vir afastado.
TranslatorPro	De qualquer maneira lhes precisaram de camas, assim eles enviaram a casa dela, e er, você sabe que er vêem aquela matéria-prima que vai, você vê ainda havia um pouco de materiais partiram para ir.
Sistemas que traduziram corretamente: Nenhum	
“This city has suddenly <b>come</b> alive,” said her husband, an off-duty border guard.	
Acepção correta: renascer – renasceu	
FreeTranslation	“Esta cidade repentinamente veio vivo,” disse seu marido, um guarda de fronteira de fora-dever.
Systran	"esta cidade tem vivo de repente vindo," dito seu marido, um protetor off-duty da beira.
TranslatorPro	"Esta cidade veio viva" de repente, disse o marido dela, um guarda de borda de fora-dever.
Sistemas que traduziram corretamente: Nenhum	
“Yes, I’m <b>coming</b> , but I’ve one or two things to attend to first,” she explained.	
Acepção correta: ir – indo	
FreeTranslation	“Sim, venho, mas tenho um ou duas coisas para assis tir a primeiro,” explicou.
Systran	"sim, eu estou vindo, mas eu tenho um ou duas coisas a atender a primeiramente," explicou.
TranslatorPro	"Sim, eu estou vindo, mas eu tenho uma ou duas coisas para prestar atenção a primeiro", ela explicou.
Sistemas que traduziram corretamente: Nenhum	

Tabela 27 – Detalhamento da análise do verbo *to give*

<i>to give</i>	Acepção	Frequência da aceção	Nº Sentenças Problemáticas
	dar	29	0
	fazer	1	1
	desistir ( <i>give up</i> )	2	2
	ceder ( <i>give in</i> )	2	1
	causar, dar origem ( <i>give rise</i> )	1	0
	abandonar / deixar ( <i>give up</i> )	1	1
	revelar / denunciar ( <i>give away</i> )	1	1
	distribuir ( <i>give out</i> )	1	1
	conceder ( <i>give in</i> )	1	0
	expressão “dar a luz” ( <i>give birth</i> )	1	1
expressão “levar em consideração” / “refletir” ( <i>give thought</i> )	3	3	

Na maioria das sentenças da amostra, o verbo *to give* (Tabela 27) foi utilizado na aceção “dar”, corretamente identificada. Contudo, quando utilizado com outros sentidos, em especial, no caso de *phrasal verbs*, tal sentido não foi identificado, o que levou a traduções bastante problemáticas em 33.3% das sentenças, conforme os exemplos da Tabela 28.

Tabela 28 – Exemplos de problemas com o uso do verbo *to give*

But her hands were trembling uncontrollably, <b>giving</b> away to the duchess the measure of her anxiety.	
Acepção correta: revelar – revelando / denunciar – denunciando	
FreeTranslation	Mas passa tremiam incontrolavelmente, distribui ao duchess a medida de sua ansiedade.
Systran	Mas as mãos estava tremendo uncontrollably, dando afastado à duquesa a medida de sua ansiedade.
TranslatorPro	Mas as mãos dela estavam tremendo incontrolavelmente, dando à duquesa a medida da ansiedade dela.
Sistemas que traduziram corretamente: Nenhum	
I, I think we'd have more effect if, if we'd erm <b>gave</b> them <b>in</b> at the shop where we're known by sight.	
Acepção correta: conceder – se tivéssemos concedido	
FreeTranslation	Eu, eu pensamos que nós teríamos mais efeito se, se ríamos erm deram-nos em na loja onde somos sabidos por visão.
Systran	I, I pensam que nós teríamos mais efeito se, se nós o erm os der dentro na loja onde nós somos conhecidos pela vista.
TranslatorPro	Eu, eu penso que nós teríamos mais efeito se, se nós erm os cederam na loja onde nós somos conhecidos através de visão.
Sistemas que traduziram corretamente: Nenhum	

Tendo em vista os efeitos negativos gerados das traduções incorretas do verbo *to give* para a compreensão da sentença, embora o número de sentenças com tais problemas não tenha atingido 50% da amostra, esse verbo foi considerado problemático.

Tabela 29 – Detalhamento da análise do verbo *to look*

<i>to look</i>	Acepção	Frequência da aceção	Nº Sentenças Problemáticas
	olhar	13	1
	ver	1	1
	parecer	7	5
	aparentar	4	4
	procurar ( <i>look for</i> )	2	2
	esperar, ter expectativa ( <i>look forward</i> )	1	0
	examinar ( <i>look into</i> )	1	1
	rever ( <i>look back</i> )	1	1

O verbo *to look* (Tabela 29) é traduzido em praticamente todas as sentenças considerando-se sua aceção mais comum “olhar”. Assim, em 15 das 17 sentenças em que a aceção adequada não era essa,

os sistemas não foram capazes de identificar tal acepção. Como resultado, 50% das sentenças da amostra apresentam problemas na tradução, como os ilustrados na Tabela 30. Por essa razão, o verbo *to look* foi considerado problemático.

Tabela 30 – Exemplos de problemas com o uso do verbo *to look*

They all manage to <b>look</b> much the same.	
Acepção correta: aparentar – aparentar	
FreeTranslation	Eles todo consegue para olhar muito o mesmo.
Systran	Todos controlam olhar muito mesmo.
TranslatorPro	Todos eles conseguem olhar muito o mesmo.
Sistemas que traduziram corretamente: Nenhum	
Generally we agree on how the situation in South Africa <b>looks</b> today.	
Acepção correta: parecer – parece	
FreeTranslation	Geralmente concordamos em como a situação em África do Sul olha hoje.
Systran	Geralmente nós concordamos com como a situação em África do Sul olha hoje.
TranslatorPro	Geralmente nós concordamos em como a situação na África do Sul olha hoje.
Sistemas que traduziram corretamente: Nenhum	

Tabela 31 – Detalhamento da análise do verbo *to use*

<i>to use</i>	Acepção	Frequência da acepção	Nº Sentenças Problemáticas
	usar	29	0
	costumar	5	5
	acostumar-se	2	2

O verbo *to use* foi traduzido em todas as sentenças na sua acepção mais comum “usar”, dando origem a problemas como os ilustrados na Tabela 32. Em função da quantidade reduzida de sentenças com problemas na tradução do verbo (19.4% das sentenças), o verbo *to use* não foi considerado problemático.

Tabela 32 – Exemplos de problemas com o uso do verbo *to use*

It <b>used</b> to be possible to complete three circuits of the Barkhor before being detained.	
Acepção correta: costumar – costumava	
FreeTranslation	Era possível completar três circuitos do Barkhor antes de deter.
Systran	Usou-se ser possível terminar três circuitos do Barkhor antes de ser detido.
TranslatorPro	Era possível completar três circuitos do Barkhor antes de estar detido.
Sistemas que traduziram corretamente: Nenhum	
The news was welcomed yesterday by the End Conscription Campaign, which has been running an increasingly effective protest against national service on the grounds that servicemen are <b>used</b> to oppress the country's majority population.	
Acepção correta: acostumar – (estão) acostumados	
FreeTranslation	A notícia foi recebida ontem pela Campanha de Recrutamento de Fim, que tem corrido um protesto crescentemente eficiente contra serviço nacional nos chãos que militares são usados para oprimir a a população de maioria do país.
Systran	A notícia era a campanha para o fim ontem dada boas-vindas do conscription, que tem funcionado um protesto cada vez mais eficaz de encontro ao serviço nacional nas terras que os recrutas estão usados oppress a população da maioria do país.
TranslatorPro	As notícias foram dadas boas-vindas ontem pela Campanha de Conscrição de Fim que tem corrido um protesto crescentemente efetivo contra serviço nacional nos chãos que são usados membro das forças armadas para oprimir a população de maioria do país.
Sistemas que traduziram corretamente: Nenhum	

Como resultado dessa análise detalhada do problema da ambigüidade lexical de sentido nas traduções dos verbos em todas as sentenças da amostra, os seguintes verbos foram considerados problemáticos: *to go*, *to get*, *to make*, *to take*, *to come*, *to look* e *to give*.



## 5 Considerações Finais

Neste relatório foi descrito um estudo sobre o problema da ambigüidade lexical de sentido na tradução automática do inglês para o português. O estudo focalizou a tradução dos verbos das sentenças, considerando os 15 verbos mais frequentes no corpus BNC: *to be, to have, to do, to say, to go, to get, to make, to see, to know, to take, to think, to come, to give, to look* e *to use*. O objetivo era analisar a existência do fenômeno da ambigüidade na tradução desses verbos em uma amostra de sentenças do corpus BNC, os efeitos da ambigüidade na compreensão da sentença e o comportamento de três sistemas de tradução automática disponíveis atualmente (Systran, FreeTranslation e Globalink L&H Power Translator® Pro) diante dos verbos considerados ambíguos. Com isso, pretendia-se selecionar um subconjunto desses verbos contendo os verbos considerados mais problemáticos para a tradução, sob o aspecto da ambigüidade lexical de sentido.

Para tanto, para cada um dos 15 verbos com mais de uma possível acepção em dicionários bilíngües foram selecionadas sentenças do corpus BNC, considerando-se as várias flexões desse verbo (tempo, pessoa, etc.). Em seguida, essas sentenças foram submetidas aos três tradutores automáticos e as traduções foram então analisadas por uma tradutora humana. Por fim, os resultados dessa análise foram compilados para levantar o subconjunto dos verbos mais problemáticos. Foram considerados problemáticos os verbos nos quais a ambigüidade podia ser verificada em pelo menos 50% das sentenças, assumindo-se que um verbo era ambíguo em uma sentença se pelo menos dois dos sistemas de tradução não eram capazes de identificar a sua acepção correta naquela sentença. Os seguintes verbos apresentavam essa característica: *to go, to get, to make, to take, to come* e *to look*. Além disso, foram considerados problemáticos os verbos cuja ambigüidade, embora verificada em menos de 50% das sentenças, causava problemas substanciais na compreensão das sentenças traduzidas. Somente o verbo *to give* apresentou essa característica. Como resultado dessa compilação, sete verbos foram considerados problemáticos: *to go, to get, to make, to take, to come, to look* e *to give*.

Esses resultados mostram que, conforme afirmam Ide & Véronis (1998), apesar de a TA ser uma área bastante antiga e ter sido a primeira área na qual o problema da ambigüidade de sentido foi identificado, pouquíssimos sistemas de TA incorporam métodos recentemente desenvolvidos, e potencialmente eficientes para a DLS, reconhecidamente um problema de difícil solução nessa área. Nos sistemas estudados, de uso expressivo atualmente, normalmente escolhe-se por uma das possíveis acepções de um verbo, provavelmente a considerada mais comum, e esta acepção é utilizada na tradução da maioria das suas ocorrências, excetuando-se alguns casos do uso do verbo *phrasal verbs* ou em expressões comuns.

A partir do subconjunto de sete verbos resultante do refinamento do conjunto inicial dos 15 verbos mais frequentes do BNC pretende-se realizar uma nova análise para identificar quais as características morfológicas, sintáticas, semânticas e/ou pragmáticas do verbo e da sentença necessárias para evidenciar a acepção correta do verbo. Além disso, pretende-se analisar possíveis regularidades no problema da ambigüidade nesse subconjunto que permitam tratar tal fenômeno de maneira sistemática. Essas análises serão realizadas visando à criação de um modelo computacional para o tratamento do problema da ambigüidade dos verbos (considerando-se, inicialmente, esse subconjunto de verbos) na tradução automática. Por fim, esse modelo deverá ser acoplado a diferentes sistemas de tradução, auxiliando no processo de escolha lexical, nos casos de ambigüidade.

Posteriormente, pretende-se estender o conjunto de verbos a serem contemplados pelo modelo e, possivelmente, considerar a tradução de palavras de outras categorias gramaticais, de modo a oferecer subsídio ao processo tradução de sentenças completas.

## Agradecimentos

À Marcela Franco Fossey, pelo trabalho técnico de análise das acepções dos verbos nas sentenças da amostra do BNC e das suas traduções pelos sistemas considerados.

## Referências Bibliográficas

- Burnard, L. (2000). *Reference Guide for the British National Corpus (World Edition)*, Oxford University Press.
- Garside, R.G.; Leech, G.N.; Sampson, G.R. (eds.) (1987). *The Computational Analysis of English: A Corpus-based Approach*. Longman, London.
- Houaiss, A. (ed.) (1982). *Dicionário Inglês-Português*. Editora Record, Rio de Janeiro.
- Ide, N.; Véronis, J. (1998). Word Sense Disambiguation: The State of the Art. In *Computational Linguistics*, 24(1).
- Jurafsky, D; Martin, J.H. (2000). *Speech and Language Processing: An introduction to Natural Language Processing, Computational Linguistics and Speech Recognition*. Prentice-Hall, New Jersey.
- Marshall, I. (1983). Choice of Grammatical Word-class without Global Syntactic Analysis: Tagging Words in the LOB Corpus, *Computers and the Humanities*, 17: 139-50.
- Oliveira Jr., O.N.; Marchi, A.R.; Martins, M.S.; Martins, R.T. (2000). A Critical Analysis of the Performance of English-Portuguese-English MT Systems. In *Anais do V Encontro para o processamento computacional da Língua Portuguesa Escrita e Falada*. Atibaia.