



UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE LETRAS
DEPARTAMENTO DE LÍNGUAS ESTRANGEIRAS E TRADUÇÃO
LÍNGUAS ESTRANGEIRAS APLICADAS AO MULTILINGUISMO
E À SOCIEDADE DA INFORMAÇÃO

LUCCAS NOGUEIRA MARINHO

**PROJETO GLY: UMA ANÁLISE DA TRADUÇÃO AUTOMÁTICA DE PORTAIS
DE NOTÍCIAS À LUZ DAS INCOMPATIBILIDADES INTERSEMIÓTICAS**

BRASÍLIA - DF
2021

LUCCAS NOGUEIRA MARINHO

**PROJETO GLY: UMA ANÁLISE DA TRADUÇÃO AUTOMÁTICA DE PORTAIS
DE NOTÍCIAS À LUZ DAS INCOMPATIBILIDADES INTERSEMIÓTICAS**

Trabalho de Conclusão de Curso apresentado como requisito parcial à obtenção do título de Bacharel em Línguas Estrangeiras Aplicadas ao Multilinguismo e à Sociedade da Informação (LEA-MSI) da Universidade de Brasília.

Orientador: Prof. Dr. Thiago Blanch Pires, da Universidade de Brasília (UnB).

BRASÍLIA - DF
2021

LUCCAS NOGUEIRA MARINHO

Trabalho de Conclusão de Curso apresentado como requisito parcial à obtenção do título de Bacharel em Línguas Estrangeiras Aplicadas ao Multilinguismo e à Sociedade da Informação (LEA-MSI) da Universidade de Brasília.

Orientador: Prof. Dr. Thiago Blanch Pires, da Universidade de Brasília (UnB).

Aprovado em 28/05/2021

Prof. Dr. Thiago Blanch Pires
Universidade de Brasília
Orientador

Prof. Dra. Fernanda Alencar Pereira
Universidade de Brasília
Avaliadora

Prof. Dr. Fidel Armando Cañas Chávez
Universidade de Brasília
Avaliador

AGRADECIMENTOS

Inicialmente agradeço a Deus por ter me concedido saúde tanto mental quanto físico para suportar e superar esses anos de 2020 e 2021 que foram bastantes difíceis para mim e para a minha família, dedico este trabalho ao meu avô Luiz Alves Marinho, que desde quando entrei na UnB, me dizia que iria na minha formatura, até mesmo quando eu passava o dia com ele na UTI e aos meus avôs Paulo, Margarida e Delzuita que apesar da distância, sempre me apoiaram em todas as minhas escolhas.

Agradeço aos meus pais, Ana e Carlos, que nunca desistiram do meu sonho de me graduar na UnB, apesar das inúmeras vezes, aparecerem obstáculos em meu caminho, me apoiaram de todas as maneiras possíveis, sempre me deram forças para seguir em diante, ensinando-me a nunca desistir mesmo em meio as dificuldades que surgiram em nossas vidas tentando nos afastar do caminho percorrido.

Tenho a grande satisfação de ter sido orientado pelo Professor Thiago que não apenas me auxiliou no TCC, mas também me apresentou ao campo da linguística computacional e da linguística de corpus, quando me deu a oportunidade de fazer parte do Pibic. Agradeço à Professora Fernanda por ter aceitado o convite de fazer parte da banca, pois me marcou com o seu seminário de literaturas africanas, o qual me fez refletir e expandir sobre as literaturas e as várias culturas, principalmente quando mostrou a diferença dos sotaques dos lugares em que ela foi.

Quero agradecer ao professor Fidel que não apenas aceitou o convite de participar da banca, mas também me auxiliou na minha jornada acadêmica, psicológica, com bom humor e seu vasto conhecimento na área da Linguística Aplicada e não menos importante, me ajudou no tema de TCC.

RESUMO

O presente trabalho de conclusão de curso tem a intenção de expandir os gêneros textuais para a identificação de incompatibilidades que são utilizados em documentos multimodais através das incompatibilidades intersemióticas geradas pelo Google Tradutor. Tendo como principal base de apoio os estudos de Liu & O'Halloran (2009), nas análises das relações intersemióticas texto-imagem, e de Vilar et al. (2006) como estrutura na identificação dos desvios léxico-semânticos fornecidos pela plataforma virtual. Os procedimentos utilizados neste presente estudo para adquirir os resultados na coleta de documentos foram os portais de notícias, tendo como idioma central o inglês do *site* de notícias britânicas *The Times*. Sendo requisitado que abranjam ligações semânticas entre texto e imagem, especificamente imagem e legenda, foto e título, figura e corpo de texto. Dessa forma, encontram-se 3 figuras de notícias do inglês e do português traduzido automaticamente pelo site mencionado.

Palavras-chaves: Texturas intersemióticas. Incompatibilidades intersemióticas. Multimodalidade. Tradução automática.

ABSTRACT

This course conclusion work intends to expand the textual genres to identify incompatibilities that are used in multimodal documents through the intersemiotic incompatibilities generated by Google Translate. Having as main support the studies by Liu & O'Halloran (2009), in the analysis of text-image intersemiotic relations, and by Vilar et al. (2006) as a structure in the identification of lexical-semantic deviations provided by the virtual platform. The procedures used in this study to acquire the results in the collection of documents were the news portals, with English as the central language of the British news website The times. Being required to cover semantic links between text and image, specifically image and caption, photo and title, figure and body text. Thus, there are 3 news figures from English and Portuguese automatically translated by the mentioned website.

Keywords: Intersemiotic textures. Intersemiotic incompatibilities. Multimodality. Automatic translation.

LISTA DE FIGURAS

Figura 1 - Categorias de textura intersemióticas de Liu & O'Halloran (2009)	13
Figura 2 - Exemplo de estrutura paralela intersemiótica.....	14
Figura 3 - Categorias de erros propostas por Vilar et al. (2006)	16
Figura 4 - Bear on the hunt for pizza wrecks car.....	20
Figura 5 - Persiga na caça ao carro de destroços de pizza.....	20
Figura 6 - Quaffing Naked Wines may lead to a hangover.....	20
Figura 7 - Beber Vinhos Nus pode causar ressaca.....	20
Figura 8 - “Write off Jimmy at your peril” - Joe Root backs James Anderson for second Test against Pakistan.....	21
Figura 9 - “Cancele Jimmy por sua conta e risco” - Joe Root apoia James Anderson no segundo teste contra o Paquistão.....	21

LISTA DE TABELAS

TABELA 1 - Polissemia Intersemiótica - Relações Lógicas entre língua e imagens (LIU & O'HALLORAN, 2009)	16
TABELA 2 - Palavras Incorretas: Subcategorias.....	17

SUMÁRIO

RESUMO	05
ABSTRACT	06
SUMÁRIO	09
1. INTRODUÇÃO	10
2. FUNDAMENTAÇÃO TEÓRICA	12
3. METODOLOGIA	18
4. ANÁLISE	19
5. CONSIDERAÇÕES FINAIS	23
6. REFERENCIAL TEÓRICO	25

1. INTRODUÇÃO

Depois da segunda guerra mundial, as tecnologias, a medicina, a informação, a comunicação, a ciência, a linguística evoluíram com um crescimento exponencial, tendo o seu grande impulso gerado pela globalização e pela popularização do computador que se tornou uma necessidade funcional econômica de alcance mundial, independentemente da classe social do indivíduo. A dependência humana de uma máquina computacional, como os aparelhos eletrônicos, se tornou ainda mais fortes e frequentes a partir do ano de 2020 por causa da pandemia causada pelo vírus do COVID-19, que se espalhou por toda a população mundial. Devido a isso a utilização de máquinas computacionais se tornaram cada vez mais frequentes como ferramentas de trabalho, com isso surgiram modalidades virtuais para realizar atividades laborais como o *home office*, as aulas remotas, o ensino a distância (EAD) e as conferências *online*. O uso de aplicativos como o uso de tradutores automáticos (TAs) se tornaram cada vez mais frequentes, de modo que os documentos não incluem somente textos, mas imagens e sons, pois têm a intensão de construir significados dessas relações (BATEMAN, 2008). Porém, mesmo que as TAs sejam desenvolvidas para traduzir textos ou sons, não é fácil encontrar um sistema que abarque as relações intersemióticas entre texto e imagem (BATEMAN, 2014).

Em meio a várias abordagens quanto a multimodalidade, Liu e O'Halloran (2009) apresentam a exploração de texturas intersemióticas, que abarcam o uso de diferentes dispositivos coesivos entre texto e imagem para a produção de coerência. Nessas circunstâncias, a utilização de tradutores automáticos em documentos multimodais tais como dispositivos coesivos se encontram, podem resultar em incompatibilidades intersemióticas (PIRES, 2017), isto é, noutra configuração intersemiótica, em sua modificação ou até anulando a construção do significado em relação ao texto de partida e também na abordagem sócio-semiótica interdisciplinar da multimodalidade de (KRESS; VAN LEEUWEN, 1996, 2001) chamados de *modos*, pois utilizam recursos escritos, espaciais, auditivos e visuais para compor mensagens. É com base neste princípio que se faz inescusável a exploração dos resultados gerados de tradutores automáticos em documentos multimodais.

No Brasil ainda não é comum encontrar pesquisas que desenvolvem especificamente os erros gerados pelos tradutores automáticos, no contexto da ciência da informação aplicada na tradução automática, mas existem estudos correlacionados a avaliações nas melhorias destes acontecimentos. (VILAR ET AL., 2006). Dessa mesma forma, o desenvolvimento de pesquisa com relação aos erros de tradução automáticas em documentos multimodais, ou seja,

documentos estabelecidos por Bateman (PIRES, 2017, p. 14) “carregam significados que se utilizam de modos visuais, espaciais e verbais em combinação e cooperação”, fenômenos em relação à catalogação e à avaliação.

Todavia, o uso imediatista da tradução automatizada pode resultar em imprecisão no que se refere a documentos de conteúdos conectando semanticamente texto e imagem, dado que os sistemas não foram desenvolvidos para tal. Os portais de notícias são uma das principais plataformas onde os usuários utilizam ferramentas de tradução automática, pois, esses portais contêm vários gêneros e subgêneros textuais com possibilidades de variações semântica entre uma legenda e uma fotografia. Considerando que o termo “erro de tradução” é constantemente problematizado pela área de Letras especificamente nos estudos de tradução (PIRES, 2017, p. 19). Nesse contexto se faz jus à elaboração do Projeto GLY: uma análise da tradução automática de portais de notícias à luz das incompatibilidades intersemióticas (a sigla GLY é um título criado pelo próprio autor deste texto, que significa *Google By Lucas*), pois as implicações de pesquisa apontam para determinadas variedades de relações semânticas entre texto e imagem encontrados a partir de resultados do Google Tradutor.

Tendo em consideração esses aspectos, as perguntas de pesquisa formuladas são:

- **Como a multimodalidade contribui para a tradução automática em documentos multimodais?**
- **Como se dão as relações texto-imagem em documentos multimodais a partir dos resultados da tradução automática?**

Partindo da estimativa de que há relações de significado entre os elementos linguísticos e visuais possíveis de serem investigados no contexto da qual foi gerada uma tradução automática. O objetivo deste trabalho é identificar relações semânticas entre texto e imagem em documentos multimodais traduzidos automaticamente para a avaliação de tradução automática.

A atual proposta é justificada pela proposta de inovações para o campo da avaliação de tradução automática, ao adicionar o componente da multimodalidade em seu contexto de uso; para o campo da multimodalidade, ao incluir textos multisemióticos gerados por tradução automática; para a expansão do campo inter-e multidisciplinar envolvendo língua e tecnologia. Este estudo não tem propósito de solucionar um problema da tradução automática com o uso da multimodalidade, mas se propõe em expandir a abordagem da multimodalidade na identificação de relações texto-imagem geradas a partir de traduções automáticas.

O trabalho de conclusão de curso contribui para o curso Línguas Estrangeiras Aplicadas ao Multilinguismo e à Sociedade da Informação (LEA - MSI) e para futuras linhas de pesquisas como o “COMPLETT – Corpus Multilíngue para Pesquisas em Línguas Estrangeiras, Tradução e Terminologia”.

A seguir está a fundamentação teórica que é a base central do texto, pois, contém detalhadamente as metodologias usadas na seleção e na coleta das imagens e textos que formam o corpo. Sendo também mencionado o processo das avaliações e catalogações dos erros de TA com o recurso da extensão da captura de tela e edição de imagens, *FireShot* do navegador *Google Chrome*.

2. FUNDAMENTAÇÃO TEÓRICA

A tradução automática se dá por meio de um dispositivo eletrônico com atributos computadorizados. A investigação de programas de softwares que traduzem textos, artigos de um idioma para o outro pertence ao estudo da linguística computacional, mas ainda não é fácil encontrar aplicativos e programas computacionais voltados para a tradução automática multimodal. Este contexto se ampara em Hutchins (2010), pois, desde os anos 80 vem se expandindo a tradução automatizada, devido ao uso do computador e à globalização que se baseiam em regras estatísticas.

De acordo com Kress e Van Leeuwen (2001), é multiplicado os significados que articulam o texto multimodal, descrito em quatro domínios de prática chamados de “estratos”, sendo eles: discurso, *design*, produção e distribuição.

Entende-se como multimodal a compreensão e a representação, interligada e isolada de forma não verbalizada, como gestos, posturas e imagens; o qual define a linguagem como parte de um conjunto multimodal de gestos, posturas, imagens e sons, sendo também em seu uso cultural, social e histórico (JWITT, 2009). Mesmo que os tradutores automatizados sejam aprimorados para traduções de textos e sons, ainda não há um conjunto que aborde as relações intersemióticas resultantes da integração texto-imagem (BATEMAN, 2014).

Devido a uma metodologia que consiste em análises empíricas de documentos multimodais, conhecida como interação e combinação de múltiplos modos dentro de um mesmo artefato, esta pesquisa adiciona o discurso multimodal de Bateman (2008). Para o autor, “modo” é definido como qualquer aspecto visual que apresenta informação, tendo também a sua relevância na metodologia a noção de gênero, que é atribuído como uma vasta possibilidade de

documentos e matérias de artefatos multimodais como essências na concepção de gênero multimodal, tais como gênero de jornais e gênero de websites.

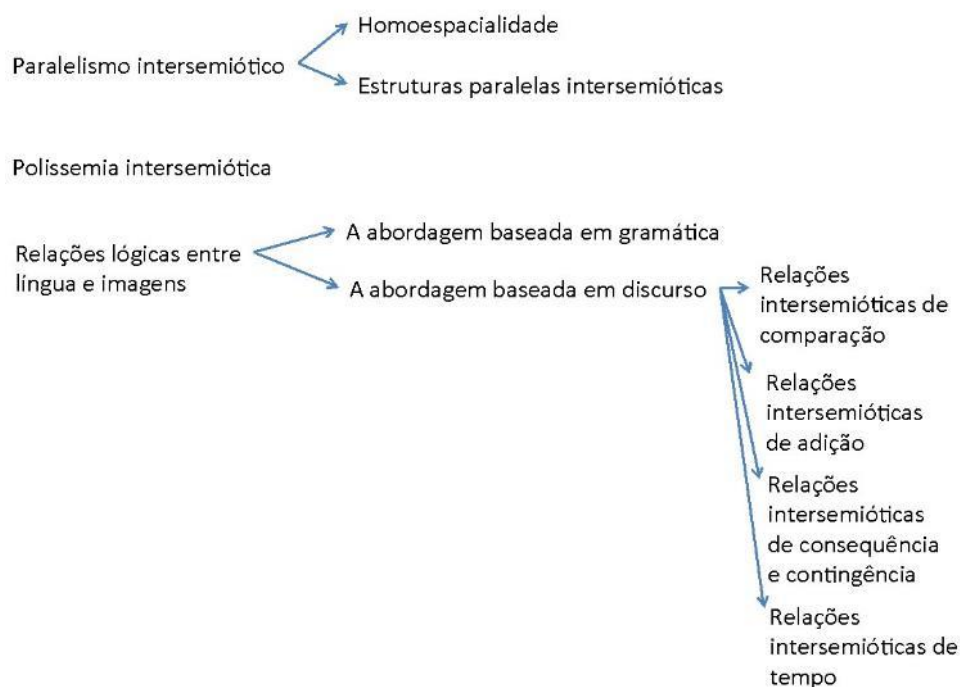
2.1 Incompatibilidade Intersemiótica

A proposta de Liu e O'Halloran (2009), em sua abordagem com modelo de concepção de texturas intersemióticas “que compreende o uso de diferentes dispositivos coesivos entre texto e imagem para a produção de coerência”. Neste contexto, pode-se ocorrer, ao utilizar o tradutor automático, o que chamamos de Incompatibilidade Intersemiótica (PIRES, 2017), ou seja, um novo resultado da tradução incondizente com o texto original de partida.

A única forma pela qual se pode reconhecer essas “mudanças” de significado visual-verbal na tradução é quando há incompatibilidade (mismatch) léxico semântica referindo-se a uma imagem ou parte dela, ou seja, a imagem não representa os componentes semânticos que faltam na tradução. Por essa razão intitulou-se “incompatibilidade intersemiótica” (intersemiotic mismatch) a partir da geração de texto automaticamente produzido. (PIRES, 2017)

O modelo de abordagem da multimodalidade é definido em 3 categorias por Liu e O'Halloran, como pode ser visto na figura 1:

Figura 1 - Categorias de textura intersemióticas de Liu & O'Halloran (2009)



Fonte: Pires (2017)

1. Paralelismo Intersemiótico: é quando os componentes semióticos dividem a mesma forma da relação que interliga imagens e língua.

2. Polissemia Intersemiótica: criação de uma co-contextualização devido ao compartilhamento de significados semelhantes, opostos a significados diferentes.

3. Relações Lógicas entre Língua e Imagens: se trata da abordagem gramatical e discursiva nos significados lógicos dos componentes visuais e verbais.

2.1.A Homoespecialidade

Caracterizado pela língua e imagem no plano de expressão, sendo uma espécie de paralelismo quando o texto é ocupado pelo mesmo espaço da imagem.

2.1.B Estruturas paralelas intersemióticas

Sua principal característica é a “transitividade” gerada pelo compartilhamento da mesma configuração da língua e da imagem.

A figura a baixo mostra uma mulher palestina sendo atacada por um cachorro; que é apresentado na descrição da legenda a mesma ação (cachorro do exército israelense ataca mulher palestina).

Figura 2 – Exemplo de estrutura paralela intersemiótica



Israeli army dog attacks Palestinian woman

Fonte: Liu e O'Halloran (2009)

2.2 Relações lógicas entre língua e imagens

“Buscam tratar das relações lógicas entre língua e imagens, isto é, da análise de significados lógicos entre componentes verbais e visuais baseada em dois tipos de abordagem:(i) gramatical e (ii) discursiva”. (Pires, 2017)

2.2.1. Abordagem baseada em gramática

É a definição de uma descrição primária dos significados lógicos dentre vários recursos semióticos. Já a abordagem baseada no discurso, descrevem a complexidade das relações de sentenças, tornando-se oposto da abordagem gramatical.

A seguir, a tabela apresenta 4 subdivisões da abordagem baseada no discurso.

Tabela 1 – Polissemia Intersemiótica - Relações Lógicas entre língua e imagens

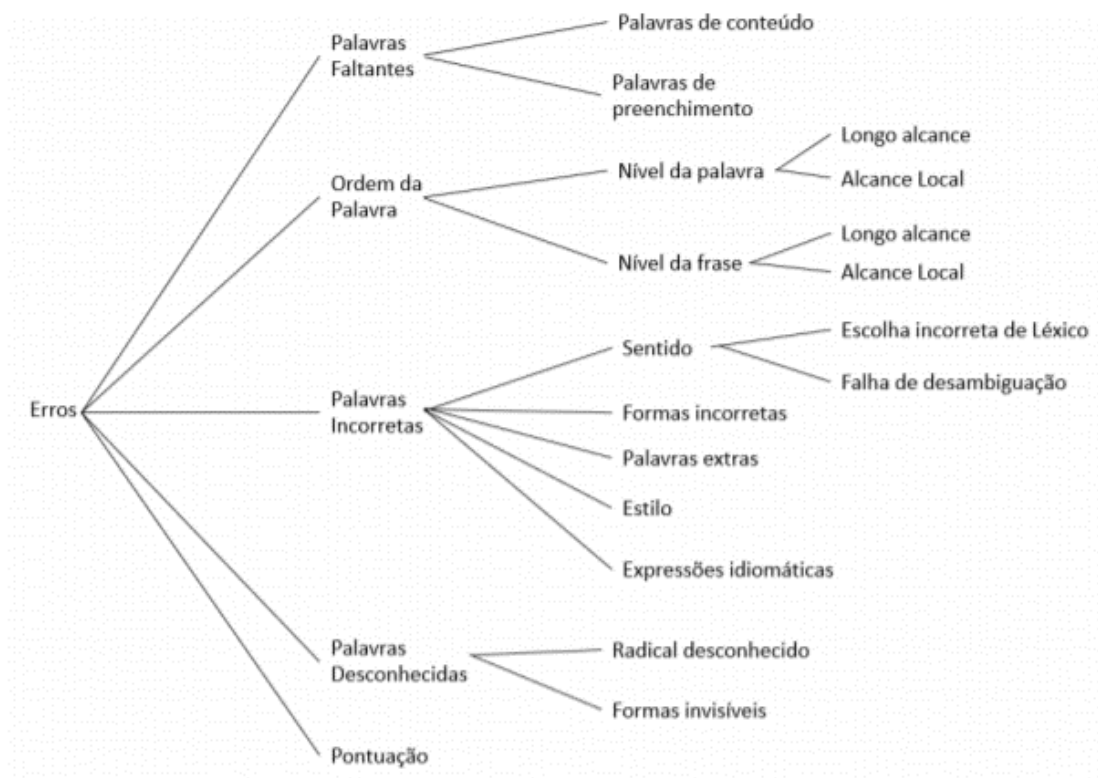
Relação intersemiótica de comparação	Recurso para organização do significado lógico em relação à semelhança entre linguagem e imagens no discurso multimodal. Assim, quando os componentes visuais e linguísticos compartilham significado experimental semelhante, os diferentes modos são sempre reformulações semióticas entre si e suas relações lógicas serão identificadas como comparativas. Além disso, sempre estarão acompanhadas de dispositivos intersemióticos coesivos.
Relação intersemiótica de Adição	Um componente semiótico, ou modo, adiciona novas informações ao outro e, portanto, as duas mensagens são unidas.
Relações intersemióticas de consequência e contingência	Segundo Pires (2017, p. 94), as relações de consequência referem-se à “causais não-modalizadas entre mensagens verbais e visuais onde o efeito foi assegurado”; já as de contingência estão presentes em textos multisemióticos onde a causa tem potencial para determinar uma possibilidade, mas sem um efeito assegurado.
Relações intersemióticas de tempo	Quando diferentes etapas processuais são representadas verbal e visualmente, as relações imagem-texto podem ser descritas com a lógica do tempo. Geralmente apresentados em manuais de procedimentos, por exemplo, manuais técnicos ou o site de procedimentos

Fonte: LIU & O'HALLORAN, 2009

2.3 Erros de tradução automática

Para definir as categorias de incompatibilidade nas classificações de erros de Tradução Automática, Vilar et al. (2006) divide em 5 categorias:

Figura 3 – Categorias de erros propostas por Vilar et al. (2006)



Fonte: Vilar et al. (2006)

1. Palavras Perdidas: ocorre quando se perde alguma palavra no resultado da tradução automática (TA).
2. Ordem da Palavra: é distinguida pelo movimento das palavras individuais ou nos blocos geradas pelas frases, tendo como alcance local ou alcance longo uma necessidade de reagrupar as palavras ou move-las para outro bloco de palavras no contexto local.
3. Palavras Incorretas: mesmo que esteja correta a tradução básica, o sistema não traduz a forma correta da palavra.

Tabela 2 – Palavras Incorretas: Subcategorias

Sentido	Ocorre uma quebra do significado da sentença, podendo apresentar mais duas subdivisões: <ul style="list-style-type: none"> • Escolha de tradução incorreta pela TA; • TA incapaz de desambiguar o significado em dado contexto.
Formas incorretas	A tradução não produz uma forma adequada para a palavra, embora de maneira básica a tradução esteja certa.
Palavras extras	Gera vocábulos sobressalentes que desconfiguram a ideia original.
Estilo	Escolha imprópria de palavras na tradução que, embora mantenham o significado, não podem ser consideradas corretas.
Expressões idiomáticas	TA não reconhece uma expressão idiomática e faz a tradução ao pé da letra, o que acarreta erros, já que essas não poderiam ser transpostas dessa maneira a outro idioma.

Fonte: Vilar et al. (2006)

4. Palavras Desconhecidas: são radicais do idioma inglês de palavras e formas desconhecidas.
5. Pontuação: apenas citado, mas não desenvolvido em Vilar et al. (2006), função atribuída como apoio de citações.
6. Para se entender melhor o que significa portais de notícias utilizados nesse texto, os portais de notícias são *sites online* que publicam informações de quaisquer assuntos de grande valor e de interesse público fornecidas pela imprensa (campo jornalístico), como podemos observar em Ribeiro e Vasco (2014), que justifica em seu livro com a seguinte citação:

Em resumo, nosso plano é, com absoluta franqueza, em nome dos interesses das empresas e instituições públicas, fornecer à imprensa e ao público dos Estados Unidos alerta e informações precisas sobre assuntos de grande valor e interesse público (Cit. in Duarte, 2004, p.36).

3. METODOLOGIA

O presente texto descreve os métodos e instrumentos utilizados para responder às seguintes perguntas de pesquisa:

- I. **Como a multimodalidade contribui para a tradução automática em documentos multimodais?**
- II. **Como se dão as relações texto-imagem em documentos multimodais a partir dos resultados da tradução automática?**

Observando as perguntas acima, essa investigação é caracterizada por método exploratório e qualitativo, pois, nesse estudo foram coletados três pares de documentos multisemióticos, explorando não apenas as incompatibilidades como também as ferramentas utilizadas nas coletas de documentos multisemióticos, tendo seu principal objetivo explorar o achado desse trabalho para uma investigação mais precisa do fenômeno observado nos erros de TAs em documentos multisemióticos dos portais de notícias da *web* (Pires, 2017).

Todos os pares coletados foram retirados do site de notícias *THETIMES*. Essa escolha se justifica na busca em documentos hipertextuais conectados na *World Wide Web*, usados para criar uma unidade semântica, como fotografias com legendas e títulos referentes a imagens em conteúdo jornalístico. (Pires, 2017, p.102). O *site* da *THETIMES* foi escolhido por causa da sua variedade de notícias, possibilitando a exploração de diferentes tipos de notícias, com base em fatos realistas que esse estudo obteve seus achados mais relevantes (da figura 4 até a figura 9).

Com o problema de pesquisa da qualidade da tradução que está contida no problema de informação multimodal em tradução gerada automaticamente, entende-se que o Google Tradutor representa um vasto meio de traduzir automaticamente conteúdo de *websites*, pois oferece um serviço multilíngue e *online* de acesso gratuito. Assim, justifica-se como instrumento para tradução do conteúdo dos *websites* do inglês para o português, pois o critério da seleção da língua estrangeira inglês, se dá pelo motivo da aplicação frequente do idioma inglês nas atividades acadêmicas.

Este trabalho foi feito de forma manual utilizando as ferramentas como *FireShot* para captura de imagens; o *Evernote*, para armazenar e editar as imagens; o Google Tradutor, para a tradução das palavras; e o *Paint*, para a colagem de tela, pois as duas primeiras ferramentas citadas têm um limite de dias para utilizá-las de forma gratuita.

Foram utilizadas como base de análises e estudos as teorias de Liu & O'Halloran (2009), para as relações intersemióticas texto-imagem, e de Vilar et al. (2006), sobre mudanças léxico-semânticas pela plataforma do Google Tradutor.

4. ANÁLISE

Esse capítulo analisa as incompatibilidades intersemióticas em três documentos multimodais produzidos em língua inglesa e suas respectivas traduções geradas pelo Google Tradutor para o idioma português, através do portal de notícias da *THETIMES*.

É importante ressaltar que o presente estudo não leva em consideração um balanceamento de número para se obter uma representação da amostra de análise.

A seguir estão as figuras utilizadas para análise.

Figura 4 - Bear on the hunt for pizza wrecks car.

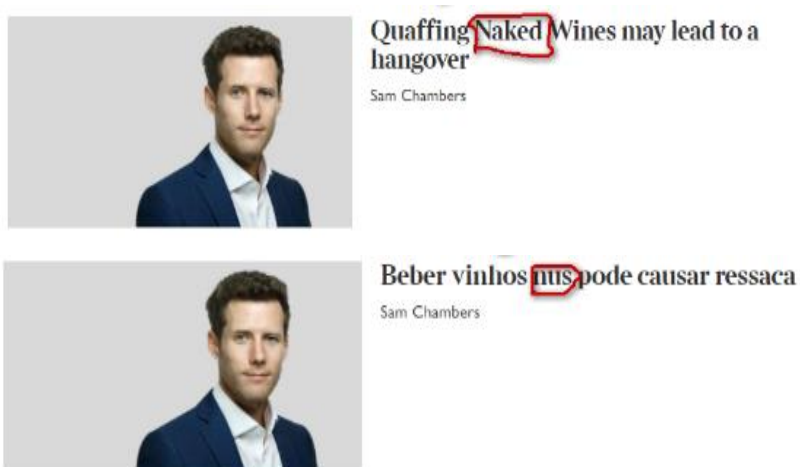
Figura 5 - Persiga na caça ao carro de destroços de pizza.



Portal de notícias: Thetimes.

Figura 6 - Quaffing Naked Wines may lead to a hangover.

Figura 7 - Beber Vinhos Nus pode causar ressaca.



Portal de notícias: Thetimes.

Figura 8 - “Write off Jimmy at your peril” - Joe Root backs James Anderson for second Test against Pakistan.

Figura 9 - “Cancele Jimmy por sua conta e risco” - Joe Root apoia James Anderson no segundo teste contra o Paquistão.

'Write off Jimmy at your peril' - Joe Root backs James Anderson for second Test against Pakistan



'Cancele Jimmy por sua conta e risco' - Joe Root apoia James Anderson no segundo teste contra o Paquistão



Portal de notícias: Thetimes.

A imagem da figura 4 foi capturado da versão original em inglês retirado do *site THETIMES* intitulado “Bear on the hunt for pizza wrecks car”, em português uma possível tradução seria “urso na caça ao carro de destroços de pizza”. A figura 5 mostra a tradução para o português gerada pelo Google Tradutor. Em português a palavra *Bear* significa “urso”. Porém, o problema reside em utilizar a palavra “persiga”, pois descaracteriza o sentido semântico e textual da imagem.

Os erros de tradução geradas pelo Google Tradutor da figura 5 parte da origem léxico-semântico, e, portanto, de sentido de Palavras Incorretas (Vilar et al., 2006) e mais específico na subcategoria Escolha incorreta de Léxico, como se pode perceber através da tradução gerada pela máquina “Persiga na caça ao carro de destroços de pizza”. Baseado no Paralelismo intersemiótico e na subcategoria Estruturas Paralelas intersemióticas (Liu & O’Halloran, 2009), pois a imagem e a língua conectam-se dividindo a semântica entre si, elas se complementam e agregam um valor semântico de transitividade.

Já a figura 6 intitulada “Quaffing Naked Wines may lead to a hangover” traduzida por “Nus” e a figura 8 “Write off Jimmy at your peril” - Joe Root backs James Anderson for second Test against Pakistan” traduzido para “Cancele Jimmy por sua conta e risco”, foram retiradas da versão original em inglês do *site THETIMES*, evidenciam as mudanças semânticas geradas pela máquina de tradução como é apresentado na figura 7 e na figura 9.

Conforme a classificação de erros de tradução proposta por VILAR et al. (2006, p. 698), os erros gerados pelo Google Tradutor estão na categoria das palavras incorretas e na subcategoria Escolha Lexical errada, pois o sistema não encontrou a tradução correta da palavra e essa mesma palavra interfere de forma drástica o sentido semântico da frase, como mostra a figura 7 “Beber Vinhos Nus pode causar ressaca” e a figura 9 “Cancele Jimmy por sua conta e risco” - Joe Root apoia James Anderson no segundo teste contra o Paquistão”. Tendo em vista os aportes de Liu & O’Halloran (2009), essas figuras encontram-se na categoria Relações

Lógicas entre língua e imagens e na subcategoria Relação intersemiótica de Adição, pois o sistema adicionou informações que não contém nas imagens das figuras. Uma possível tradução para a figura 7 seria “Beber Vinho Puro pode causar ressaca” e para a figura 9 “Jimmy eliminado por risco - Joe Root apoia James Anderson para o segundo teste contra o Paquistão”.

Os achados da presente pesquisa, contribuem para os estudos de erros de tradução automática e principalmente na relação intersemiótica entre os significados entre texto e imagem em documentos multimodais.

Por fim, a próxima seção apresenta as considerações finais, respondendo as questões inicialmente propostas na metodologia, além do levantamento de novas hipóteses de pesquisas.

5. CONSIDERAÇÕES FINAIS

Dado o exposto, o presente estudo teve como objetivo mapear incompatibilidades texto-imagem geradas pelo tradutor automático em *websites*. O principal objetivo é a expansão de estudos exploratórios na área da tradução automática relacionando documentos multisemióticos. Os apanhados neste estudo revelam que as relações intersemióticas são bastante interessantes se tratando de documentos multimodais e multisemióticos. A proposta de Liu & O'Halloran (2009) textura semiótica, foi crucialmente um dos instrumentos importantes para a conseqüente percepção do fato e do mapeamento das incompatibilidades.

Todavia, no que se trata de tradução automática, foram perceptíveis alguns padrões de erros com base nas categorias de erros propostas por Vilar et al. (2006). Uma das principais percepções foi a relação de tradução da figura 7 e da figura 9, pois a máquina de tradução gerou outra palavra que não estava no contexto semântico e nem contextual das figuras, ocasionando na perda total de sentido da relação texto imagem.

O presente estudo responde à pergunta de pesquisa: **“Como a multimodalidade contribui para a tradução automática em documentos multimodais?”**. Contribui através da revelação de um padrão das incompatibilidades intersemióticas advindas de palavras incorretas (VILAR et al., 2006). Isso colabora e aprofunda a exploração dos erros de tradução e textura multimodal propostos, sendo satisfatórios para identificar as relações texto-imagem achados na análise.

A segunda pergunta de pesquisa é: **“Como se dão as relações texto-imagem em documentos multimodais a partir dos resultados da tradução automática?”**. Se dão na identificação e catalogação de relações intersemióticas que fogem das funções da máquina e da ampliação do componente intersemiótico na tradução automática.

A limitação de pesquisa encontrada, é a necessidade de um período maior de tempo para pesquisar e coletar em mais *sites* do idioma inglês e não foi encontrado nenhum tipo de plataforma *online* que faça anotações das relações texto-imagem, pois o trabalho foi exercido de forma manual.

Tendo em vista os aspectos observados, o presente estudo propõe as seguintes recomendações para pesquisas futuras:

Expandir a pesquisa incluindo relação texto-imagem através de vídeos que contenham incompatibilidades intersemióticas em determinados gêneros textuais. Outra possibilidade é o desenvolvimento de *softwares* para automatizar as anotações das relações texto-imagem. E por

fim explorar os mesmos estudos contidos nesse projeto em outros idiomas, para realizar testes de usabilidades com o Google Tradutor e outras plataformas que encontrarem.

6. REFERENCIAL TEÓRICO

BATEMAN, J. A. **Multimodality and Genre: A Foundation for the Systematic Analysis of Multimodal Documents**. New York: Palgrave MacMillan, 2008.

BATEMAN, J. A. **Text and Image: A Critical Introduction to the Visual/Verbal Divide**. New York: Routledge, 2014.

HUTCHINS, W. J. **Machine translation: A concise history**. *Journal of Translation Studies*, v. 13, n.1-2, p. 29-70, 2010.

KRESS, G.; VAN LEEUWEN, T. **Reading Images: The Grammar of Visual Design**. 1st. ed. London: Routledge, 1996.

KRESS, G.; VAN LEEUWEN, T. **Multimodal Discourse: The modes and media of contemporary communication**. London: Edward Arnold, 2001.

KRESS, G.; VAN LEEUWEN, T. **Reading Images: The Grammar of Visual Design**. 2nd. ed. London: Routledge, 2006.

LIU, Y.; O'HALLORAN, K. L. **Intersemiotic Texture: analyzing cohesive devices between language and images**. *Social Semiotics*, v. 19, n. 4, p. 367–388, 2009.

PIRES, T. B. **Ampliando olhares sobre a tradução automática online: um estudo exploratório de categorias de erros de máquina de tradução gerados em documentos multimodais**. Tese de doutorado—Brasília: UnB, 2017.

RIBEIRO, VASCO, **O peso do Press Release no processo de produção de notícias**. Porto: U. Porto, 2014.

VILAR, D. et al. **Error analysis of statistical machine translation output**. In: *PROCEEDINGS OF LREC*. Genoa: 2006.