

## ESTRUTURAS RETÓRICAS E ESTRATÉGIAS DE TRADUÇÃO AUTOMÁTICA

Giovanna Martinez Ursulino (PIBIC/CNPq/FA/Uem), Juliano Desiderato Antonio (Orientador), e-mail: giovannam.ursulino@gmail.com.

Universidade Estadual de Maringá / Centro de Ciências Humanas, Letras e Artes /Maringá, PR.

### Linguística - Teoria e Análise Linguística

**Palavras-chave:** Estrutura Retórica, Tradução, Estratégias de Tradução.

#### Resumo:

O presente trabalho tem como objetivo investigar as estruturas retóricas e as estratégias de tradução utilizadas pelo tradutor automático do Google. Tendo como base a *Rhetorical Structure Theory*, analisaram-se as estruturas retóricas que emergem da combinação entre as partes dos textos, partindo da análise do texto fonte em língua inglesa, para, em seguida, observar-se a versão traduzida pelo Google. Os pesquisadores segmentaram os textos em EDUs (unidades discursivas elementares) e, então, dedicaram-se a observar as mudanças ocorridas, as diferenças na estrutura retórica e as possíveis causas para isso, verificando se tais mudanças ocorrem devido a estratégias de tradução empregadas pelo Google ou por falhas no ato tradutório. O intuito é verificar a qualidade dessas traduções e oferecer dados para futuras pesquisas na área voltada para o aperfeiçoamento dos serviços de tradução automática tendo em vista sua importância no mundo de hoje. Os resultados desta pesquisa indicam que o tradutor automático apresenta traduções que tendem a alterar a estrutura retórica do texto ou utiliza estratégias de tradução que diferenciam o texto fonte do texto traduzido. Em alguns casos, essas mudanças são causadas pela polissemia de uma conjunção na língua fonte, em outros verifica-se que o tradutor realiza mudanças de classe gramatical e, conseqüentemente, gera uma mudança de segmentação.

#### Introdução

Com o crescente uso da Língua Inglesa e a enorme demanda por textos traduzidos, um recurso que tem sido vastamente utilizado são os tradutores automáticos. Apesar da melhoria na qualidade das traduções feitas pelo tradutor automático, algumas falhas ainda podem ser observadas nesses textos traduzidos. Um tradutor automático frequentemente utilizado é o Google Tradutor, que funciona a partir da memória de termos e de expressões pré-determinadas.

Este trabalho propõe-se a investigar as estratégias de tradução utilizadas pelo tradutor automático e as estruturas retóricas dos textos fonte e traduzido. A análise, feita no nível da estrutura textual, baseia-se na *Rhetorical Structure Theory* (RST), uma teoria descritiva que estuda a organização dos textos, descrevendo as relações que se estabelecem entre suas partes (MANN; THOMPSON, 1988). A teoria propõe

que, além do conteúdo explícito, existem proposições implícitas que surgem da combinação entre orações e porções maiores do texto chamadas proposições relacionais (TABOADA, 2009).

A investigação das diferentes estratégias utilizadas pelo tradutor automático do Google se dará por meio de uma análise quantitativa e qualitativa, em que será investigada a frequência de concordância entre a estrutura retórica do texto fonte e da tradução realizada automaticamente, levando-se em conta os seguintes parâmetros: quantidade de EDUs; tipo de relações retóricas estabelecidas em cada nó do diagrama; tipo de relação (multinuclear ou núcleo-satélite); motivos de discrepância.

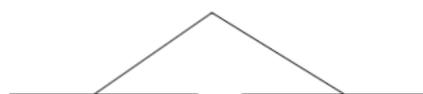
## Materiais e métodos

A RST surgiu de um trabalho de investigação em geração automática de textos, seu principal objetivo é descrever textos. Com isso, ela identifica estruturas hierárquicas dos textos, descrevendo as relações existentes entre elas (MANN; THOMPSON, 1988). Atualmente, há uma lista com 30 relações no website da teoria (MANN; THOMPSON, 2010), essa lista, entretanto, não é um rol fechado e outros autores podem propor novas relações. De acordo com a teoria, as relações organizam-se em:

- Núcleo-Satélite: a porção textual que funciona como satélite traz informações para subsidiar o conteúdo do núcleo. A figura 1 representa uma relação núcleo-satélite, em que o arco vai do satélite em direção ao núcleo.
- Multinuclear: cada porção de texto tem o mesmo estatuto. Na figura 2, que representa uma relação multinuclear, pode-se observar que cada porção textual funciona como um núcleo distinto.

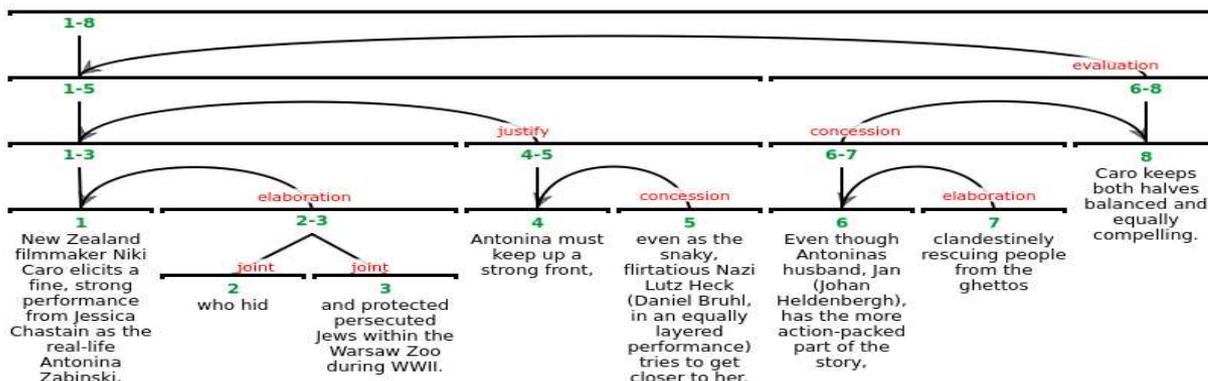


**Figura 1.** Esquema de relação núcleo-satélite  
Fonte: Mann e Thompson (1988, p. 247)



**Figura 2.** Esquema de relação multinuclear  
Fonte: Mann e Thompson (1988, p. 247)

As análises são representadas por diagramas em forma arbórea, conforme figura 3.



**Figura 3.** Diagrama da estrutura retórica de um texto do corpus do trabalho.

Fonte: autor.

As etapas estabelecidas no intuito de ajudar na investigação foram as seguintes: a) compreender a teoria de acordo com os textos de Mann e Thompson (1988) e Taboada (2009); b) realizar a segmentação dos textos em EDUs<sup>1</sup>; c) descrever as relações estabelecidas em cada parte do texto tanto para os textos em inglês quanto para os textos traduzidos; d) elaborar os diagramas de cada texto, com base na estrutura retórica observada em cada um; e) realizar um trabalho de análise quantitativa e qualitativa das traduções, por meio da comparação com o texto fonte; f) descrever as estratégias utilizadas pelo tradutor automático; g) verificar as possíveis causas para as diferenças apresentadas.

## Resultados e Discussão

A investigação proposta nesta pesquisa deu-se a partir da análise de 13 resenhas de filmes retiradas do site do iTunes, divididos pelos subtítulos *review* e sobre o filme. Para a análise, foi necessário traduzir esses textos no Google tradutor e, em seguida, verificar as segmentações em EDUs e as relações retóricas dos textos fonte e compará-los com os textos traduzidos. Com isso, foram observadas algumas mudanças de segmentação e, conseqüentemente, nos diferentes tipos de relações retóricas estabelecidas. Apresentam-se, a seguir, alguns dos erros e estratégias observados durante a pesquisa. Alguns dos erros observados foram os seguintes:

1. O tradutor automático toma a segunda oração de um complexo de orações paratáticas como argumento do verbo da primeira: “[...] / *and learns to look beyond the Beasts hideous exterior /and realizes the kind heart of the true Prince within.*”, “[...] / e aprende a olhar para além do exterior hediondo da Fera e perceber o coração bondoso do verdadeiro Príncipe dentro.”
2. O tradutor automático traduz conectivo equivocadamente: “*As events begin to spin out of control,*”, “Como os eventos começam a sair do controle”
3. O tradutor automático traduz substantivo como verbo: “*The relationship between barista/aspiring actress Mia (Emma Stone) and traditionalist jazz pianist Sebastian (Ryan Gosling) plays out in song and dance in a story that soars in flights of fantasy*”, “A relação entre barista / aspirante a atriz Mia (Emma Stone) e pianista de jazz tradicionalista Sebastian (Ryan Gosling) toca música e dança”

Algumas das estratégias observadas foram as seguintes:

1. No original, utiliza-se adjetivo, ao passo que o tradutor automático utiliza uma oração reduzida de gerúndio: “*Unsure what to believe*”, “Não tendo certeza do que acreditar”.
2. No original, utiliza-se verbo na construção comparativa e o tradutor automático não utiliza: “*The music and choreography are glorious, / and so is the evocation of Los Angeles as a lovers playground,*”, “a música e a coreografia são gloriosas, assim como a evocação de Los Angeles como um playground para os amantes”

<sup>1</sup> *Elementary Discourse Units* são blocos mínimos de construção de uma árvore discursiva.

Geralmente correspondem a orações, com exceção de orações completivas e de orações restritivas (CARLSON; MARCU, 2001).

3. No original, utiliza-se oração explicativa e o tradutor automático utiliza um adjetivo: “*who was obsessed with the idea of discovering an ancient civilization buried deep in the Amazon.*”, “obcecado com a ideia de descobrir uma antiga civilização enterrada nas profundezas da Amazônia.”.

Os erros e as estratégias apresentados acima aconteceram em 11 dos 26 textos utilizados. Além disso, em alguns dos casos esses erros e/ou estratégias ocorreram em mais de um texto analisado.

## Conclusões

Este trabalho teve como objetivo investigar as estruturas retóricas e as estratégias de tradução utilizadas pelo tradutor automático do Google. Na análise, observou-se que o tradutor automático é capaz de utilizar estratégias durante o ato tradutório, conforme mencionado anteriormente, no entanto observou-se que, em alguns casos ele ainda apresenta erros. Com isso, espera-se que este trabalho sirva como ponto de partida para pesquisas futuras, que possam sugerir melhorias nos serviços de tradução automática de forma a tornar seus trabalhos mais confiáveis.

## Agradecimentos

Ao CNPq, pelo subsídio financeiro que tornou possível a realização deste projeto, ao Prof. Dr. Juliano Desiderato Antonio, pela oportunidade e pela assistência durante a pesquisa, à Universidade Estadual de Maringá pelo apoio aos projetos de iniciação científica e a todos os envolvidos com este projeto de pesquisa.

## Referências

CARLSON, L.; MARCU, D. **Discourse Tagging Reference Manual**. Los Angeles: University of Southern California, 2001.

MANN, W. C.; TABOADA, M. **RST Web Site**. Disponível em: <<http://www.sfu.ca/rst>>.

MANN, W. C.; THOMPSON, S. A. Rhetorical Structure Theory: toward a functional theory of text organization. **Text**, v. 8, n. 3, p. 243–281, 1988.

TABOADA, M. Implicit and explicit coherence relations. In: RENKEMA, J. (Ed.). **Discourse, of course**. Amsterdam: John Benjamins, 2009. p. 125–138.