PADRÕES DE IDENTIFICAÇÃO DA UNIDADE CENTRAL EM TEXTOS ARGUMENTATIVOS

Daniella Paes da Silva (PIBIC/CNPq/UEM), Juliano Antonio Desiderato (Orientador), e-mail: daniellapaes@live.com

Universidade Estadual de Maringá/Centro de Ciências Humanas/Maringá, PR.

Linguística, Letras e Artes
Teoria e Análise Linguística

Palavras-chave: Unidade Central, Texto dissertativo-argumentativo, Gênero resposta argumentativa

Resumo:
Esta pesquisa tem como objetivo investigar critérios que contribuam para a identificação automática da unidade central de textos dissertativo-argumentativos e de textos do gênero resposta argumentativa. Para tanto, observou-se a posição da UC de cada texto do *corpus*, além de outros traços que auxiliaram na busca de um padrão. Essa análise revela-se importante, pois contribui para o desenvolvimento de recursos capazes de avaliar textos automaticamente. A análise tem como base a RST (*Rethorical Structure Theory*), teoria funcionalista de caráter descritivo a qual procura estudar a organização textual por meio das relações implícitas que se estabelecem entre as partes do texto. Conforme esperado, constatou-se que na maioria das respostas argumentativas, as UCs ocorreram logo no início do texto, já nos textos dissertativo-argumentativos, a maior parte das UCs foram identificadas na porção que abrange entre 20% e 30% do texto. Esse resultado justifica-se pelo fato de que, na resposta argumentativa, o produtor deve iniciar respondendo a pergunta presente no comando, enquanto no texto dissertativo-argumentativo, é necessário que, antes da ideia central do texto, apresente-se um fundo o qual introduza o leitor no tema em questão. Também se destacou, entre os critérios de identificação das UCs, o uso de alguns substantivos e adjetivos os quais corroboraram para garantir que o comando do texto dissertativo-argumentativo fosse atendido, e, no caso das respostas argumentativas, observou-se o uso de verbos como *acreditar*, *crer* e *pensar*, consequência, essa, do caráter subjetivo desse gênero.

Introdução

O objetivo desta pesquisa é investigar critérios que corroborem para a identificação automática da unidade central (Iruskieta et al, 2015) de textos dissertativo-argumentativos e de textos do gênero resposta argumentativa. A análise tem como base a RST (*Rethorical Structure Theory*), uma teoria funcionalista de caráter descritivo a qual procura estudar a organização textual por meio das relações implícitas que se estabelecem entre as partes do texto (Mann e Thompson, 1988) . Essa análise revela-se importante, pois contribui para o desenvolvimento de recursos capazes de avaliar textos automaticamente.

Materiais e métodos

Para constituir o *corpus* da pesquisa, foram selecionados 22 textos dissertativo-argumentativos referentes ao Enem 2013, publicados na internet por terem sido bem avaliados. Após a seleção, os textos foram separados por cláusulas e identificou-se, então, a unidade central de cada um deles. Os critérios observados para identificação das UCs refletiram, de certa forma, o comando da produção textual, são eles: a repetição de determinados substantivos e adjetivos que apareceram no comando ou pertenciam ao mesmo campo semântico e a busca pelo trecho no qual o produtor atendia a proposta, ou seja, opinava sobre os “Efeitos da implantação da Lei Seca no Brasil”. O resultado dessa análise foi comparado com os resultados de uma pesquisa anterior (Antonio, 2015) na qual foram analisados cem textos do gênero resposta argumentativa, produzidos por candidatos ao Vestibular de Verão de 2013 da UEM.

Resultados e Discussão

Como os textos dissertativo-argumentativos possuem tamanhos diferentes, observou-se a posição relativa das UCs. Foi constatado que em seis dos 22 textos, a UC se encontra na porção que vai de 20% e 30% da extensão total do texto. Nessa porção, há a maior ocorrência de UCs, seguido por 10%-20%, cinco textos, e 30%-40%, quatro textos. É possível justificar essa constatação pelo fato de que, no texto dissertativo-argumentativo, é comum que se inicie com um fundo, uma breve

contextualização do tema, e, logo em seguida, apresente-se a ideia central a qual deve ser desenvolvida ao longo do texto.

Quadro 1. Posição relativa da UC nos textos dissertativo-argumentativos

POSIÇÃO RELATIVA	N
0%-10%	2
10%-20%	5
20%-30%	6
30%-40%	4

POSIÇÃO RELATIVA	N
40%-50%	1
50%-60%	3
60%-90%	0
90%-100%	1

A resposta argumentativa, por sua vez, exige que o produtor já comece seu texto respondendo à questão levantada no comando. As UCs costumam ocorrer, portanto logo no início do texto, o que se confirmou em 90,7% do *corpus* analisado. (Antonio, 2015)

Nos quadros a seguir, está quantificado o número de substantivos e adjetivos, respectivamente, que ocorreram três ou mais vezes nas UCs dos textos dissertativo-argumentativos. Os termos em negrito são aqueles que estavam presentes no comando e aqueles separados por barras foram considerados sinônimos. Acredita-se que o maior número de ocorrências dessas palavras é resultado da preocupação dos produtores em atender o comando e não fugir do tema.

Quadro 2. *Substantivos com três ou mais ocorrências no corpus*

SUBSTANTIVOS	N
Lei	16
Números/Estatísticas/Dados	10
Acidentes	9
Diminuição/Queda/Redução	7
Efeitos/Resultados/Impacto	7
Implantação	6

SUBSTANTIVOS	N
Brasil/País	5
Motorista(s)	4
Trânsito	4
Álcool	3
Fiscalização	3

Quadro 3. *Adjetivos com três ou mais ocorrências no corpus*

ADJETIVOS	N
Seca	13

ADJETIVOS	N
Positivos	5

Verbos como “pensar” e “acreditar” tiveram grande ocorrência nas unidades centrais das respostas argumentativas, o que se explica pelo caráter subjetivo de tal gênero, já no texto dissertativo-argumentativo, o uso de verbos não se destacou.

Outra característica notada que merece destaque na identificação de UCs é a utilização de advérbios. No caso dos textos dissertativo-argumentativos, os advérbios “ainda” e “consideravelmente” (assim como o

adjetivo “considerável”) se destacaram. Eles foram utilizados para explicar que a Lei Seca não anulou o índice de acidentes de trânsito causados por motoristas alcoolizados, mas garantiu sua queda.

Já no *corpus* do gênero RA, advérbios e expressões adverbiais de vários tipos foram utilizados, como “na verdade”, “muito”, “apenas”, “principalmente”, “geralmente”, entre outros. Mais uma vez, o caráter subjetivo do gênero RA permitiu que os produtores usassem mais advérbios, assim como os textos de apoio que davam mais liberdade aos produtores da resposta argumentativa. (Antonio, 2015)

Conclusões

Após realização da pesquisa, concluiu-se que os principais critérios para identificação da unidade central foram: posição da UC e ocorrência de determinados substantivos, adjetivos, verbos e advérbios que corroboraram para a construção do texto. Espera-se que essa pesquisa possa contribuir para a criação de aplicações de avaliação automática e de manuais de produção textual para alunos de ensino médio.

Agradecimentos

Ao CNPQ pela concessão da bolsa PIBIC
Ao professor Juliano Desiderato Antonio por tornar o projeto possível por meio de tão dedicada orientação

Referências

ANTONIO, J. D. **Detecting central units in argumentative answer genre: signals that influence annotator’s agreement.** Actas del XXXI Congreso de la Sociedad Española para el Procesamiento del Lenguaje Natural. **Anais...**Alicante (Espanha): Sociedad Española para el Procesamiento del Lenguaje Natural, 2015.

IRUSKIETA, M. et al. **The Detection of Central Units in Basque scientific abstracts**. *Actas del XXXI Congreso de la Sociedad Española del Procesamiento del Lenguaje Natural*. Alicante (Espanha): SEPLN, 2015

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.