

Identificação de categorias retóricas sentenciais em *abstracts* de Química

Introdução

Diante da volumosa quantidade de informação escrita disponível nas diversas áreas do conhecimento (ORASAN, 2001), inclusive em artigos científicos, em revistas especializadas e em livros, há necessidade de o pesquisador selecionar criteriosamente seus insumos. Isso pode ser feito, por exemplo, com a leitura dos *abstracts* (resumos) de artigos (SWALES, 2009). A esse respeito, Feltrim *et al* (2000) afirmam:

[...] o resumo deve ser redigido com muito cuidado, de forma a ser completo, interessante e informativo, dispensando a consulta ao restante do texto para que o leitor tenha a idéia do que trata o trabalho e, ao mesmo tempo, estimulando o interesse pela leitura do texto completo (p. 4).

Uma vez que a escrita científica apresenta características próprias de seu gênero discursivo (SWALES, 1990), exigindo o uso de determinadas estruturas lexicais, sintáticas e, principalmente, textuais para sua formulação bem-sucedida, é necessário adequar o texto ao gênero para que a pesquisa seja veiculada e possa contribuir aos estudos da área. Segundo Feltrim *et al*: “[...] para que um trabalho seja bem escrito e, conseqüentemente, comunique bem o seu propósito, é preciso atentar a uma série de fatores, tanto do seu conteúdo quanto da sua estrutura”. (2000, p. 1)

O maior problema advindo dessas condições é o de que boa parte da escrita científica, atualmente, é feita em inglês (SWALES, 1990), criando assim impedimentos para autores não-nativos em seu processo de escritura. O aspecto que apresenta mais dificuldades é o textual, pois ele exige um nível sofisticado de organização, possivelmente mais que os aspectos lexicais e sintáticos. Entre essas dificuldades incluem-se “a) o uso de estruturas retóricas transferidas de sua língua materna; b) mau uso de relações lógicas e c) ausência de referências e expressões convencionais pertencentes ao discurso científico” (ALUÍSO *et al*, 2005).

Diversas estratégias são utilizadas para auxiliar cientistas não-nativos a escrever *abstracts* e artigos. Citamos aqui o ambiente AMADEUS – *AMiable Article DEvelopment for User Support* – (OLIVEIRA JR. *et al*, 2006), que baseia-se na anotação e reutilização de material linguístico não-factual como método para a escritura de textos originais na língua alvo. Embora essa estratégia compreenda todo o artigo, nosso *corpus* se limita a *abstracts*, podendo assim gerar material linguístico para esse subgênero textual (SWALES, 2009).

Como os *abstracts* têm papel determinante para despertar o interesse de leitores, devem seguir estruturas textuais convencionais, que englobam os “tipos de informação

incluídos e a ordem em que aparecem” (FELTRIM *et al*, 2000, p. 5). Os autores sugerem (FELTRIM *et al*, 2004) uma rubrica para resumos que contempla tanto a função retórica principal de cada sentença, não obstante a presença de outras, minoritárias, e a ordem em que elas ocorrem. Os tipos de função retórica, listados na ordem em que se espera que ocorram, são: **contexto** (*background*), **lacuna** (*gap*), **propósito** (*purpose*), **método** (*method*), **resultado** (*result*) e **conclusão** (*conclusion*). É esse modelo de rubrica que utilizamos como base para nossa análise.

Objetivo

Este artigo visa, ao analisar a composição de *abstracts* de Química, a verificar a existência de uma rubrica similar à proposta por Feltrim *et al* (2004). Dessa forma poderemos oferecer as bases para que dois objetivos sejam cumpridos: O de auxiliar estudantes da área, que têm inglês como segunda língua, a adequar a produção de seus resumos ao estilo adotado pelo resto da comunidade, e o de construir uma ferramenta computacional que analise resumos automaticamente.

Metodologia

As abordagens utilizadas nas análises neste artigo foram as do estabelecimento de um discurso no gênero científico proposto por Swales (1990) e da postulação de uma rubrica para áreas específicas (FELTRIM *et al*, 2004). A primeira etapa do projeto foi a criação de um *corpus* composto por 25 *abstracts* de artigos científicos de cada uma das revistas, sendo elas a *Analytical Chemistry* (doravante AC) e *Journal of The American Chemical Society* (JACS), disponíveis na internet e que apresentam fator de impacto 5,874 e 9,023, respectivamente. Os 50 *abstracts* foram selecionados aleatoriamente a partir do ano de 2008, não nos restringindo a nacionalidade do autor ou outros parâmetros, pois consideramos que sua publicação já garante padrões lexicais, sintáticos e retóricos desejados. Todas as sentenças de cada *abstract* foram delimitadas para sua categorização. Desconsideramos a informação factual relatada, atendo-nos às evidências lexicais e gramaticais que nos permitiram classificar a função retórica de cada sentença.

<background> **Conventionally**, CñH oxidation reactions are used to install functional groups. </background> <gap> The use of CñH oxidation to transform simple starting materials into highly versatile intermediates, which enable rapid access to a range of complex target structures, is a new area with tremendous potential in synthesis. </gap> <purpose> **Herein we report** a Pd(II)/sulfoxide-catalyzed allylic CñH oxidation to form anti-1,4-dioxan-2-ones from homoallylic oxygenates. </purpose> <result> These versatile building blocks **are rapidly elaborated** to differentiated syn-1,2-diols, stereodefined amino-polyols,

and syn-pyrans, structures ubiquitous in medicinally important complex molecules found in Nature. </result>
<conclusion> We **also demonstrate** that a C-H oxidation approach to the synthesis of these motifs is orthogonal and complementary to other state-of-the-art methods. </conclusion>

Tabela 1. Abstract anotado, grifos nossos.

As evidências lexicais que utilizamos foram: A ocorrência de locuções adverbiais que conotam frequência no presente do indicativo, como *conventionally* (Fig. 1) que indica **contexto**; marcadores discursivos como *however* para a identificação de **lacuna**; vocábulos metadiscursivos, (*herein*), anáforas indicativas de autoria (*this, we*) e verbos como *intend, show, present e report* (ibidem) para o propósito; utilização de preposições marcadoras de processo *by e through*, ou o vocábulo *method* para representar essa categoria retórica; os verbos *obtain, demonstrate, achieve* são comuns na apresentação do **resultado**, bem como o vocábulo *result*; e advérbios (*also*), conjunções (*furthermore*) e locuções adverbiais (*in the future*) para **conclusão**.

Já as evidências gramaticais se resumiram a flexões verbais de tempo e voz, que nos auxiliaram a identificar a categoria retórica de cada sentença. Nos casos de **contexto, lacuna e propósito** o tempo verbal é o presente do indicativo, usado na voz ativa ou passiva (em menor número). Em **método** é o presente, na voz passiva, e pretérito perfeito, na voz ativa e passiva. Em **resultado** é o presente, na voz passiva, e pretérito perfeito, na voz ativa e passiva. Na **conclusão** é o presente, voz ativa e passiva, o pretérito perfeito, voz ativa e passiva e o futuro do presente, voz ativa e passiva.

Resultados

Verificamos que 98% dos *abstracts* contêm pelo menos três das categorias retóricas previstas e uma sentença que apresenta resultado. Apenas 6% deles são iniciados por sentenças que não são **contexto** ou **propósito** – que inicia os *abstracts* na ausência daquela. O tipo de informação, representada pela função retórica de cada sentença, e sua ordenação obedecem *grosso modo* ao padrão estabelecido em outras áreas, portanto. Apontamos aqui ainda a predominância das categorias **resultado e propósito** em relação às outras nas duas revistas. Isso mostra que, apesar de certas diferenças entre as duas publicações, o cerne da divulgação das pesquisas na área é a demonstração (**propósito**) do que foi alcançado (**resultado**).

A principal diferença entre as duas revistas é a presença de sentenças veiculando o **contexto** da área pesquisada e suas **lacunas**. A AC apresenta 24% e 36% a mais dessas sentenças, respectivamente, do que a JACS. Isso pode estar relacionado ao número de palavras dos *abstracts*, que é 12% maior na primeira. Todos os demais parâmetros são parecidos, o que, nesse *corpus*, indica padronização da estrutura retórica dos textos da área.

Revista	Sentenças	Palavras	Contexto	Lacuna	Propósito	Método	Resultado	Conclusão
AC	193	5016	64%	56%	92%	60%	96%	68%
JACS	175	4398	36%	20%	84%	64%	100%	68%

Tabela 2. Levantamento lexical e sentencial dos abstracts.

Conclusão e discussão

Este trabalho mostrou que a rubrica dos resumos de Química analisados segue o padrão de outras áreas de pesquisa, como previsto por Feltrim *et al*, focando-se principalmente nos resultados, veiculados em suas próprias sentenças ou nas de **propósito**. Notamos ainda que, embora a ocorrência de sentenças de **método** seja parecida nas duas revistas, esta função retórica está embutida muitas vezes nas de **resultado**. O número real, assim, das ocorrências deverá ser categorizado de forma intra-sentencial em futuros trabalhos.

Por fim, a análise oferece material linguístico que pode ser aproveitado como modelo para o ambiente AMADEUS, auxiliando não-nativos na escrita de *abstracts*, e também como base inicial para o desenvolvimento de uma ferramenta computacional de auxílio à composição textual, similar ao SciPo-Farmácia (ALUÍSIO *et al*, 2005). Colaborações com outras áreas, especialmente a Ciência da Computação, serão necessárias para a expansão do corpus e o desenvolvimento dessa ferramenta.

Bibliografia

- ALUÍSIO, Sandra Maria. *et al*. Evaluating Scientific Abstracts with a Genre-specific Rubric. In: 12TH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE IN EDUCATION (AIED 2005), 2005, Amsterdã. Proceedings of the 12th International Conference on Artificial Intelligence in Education (AIED 2005). 2005. v. 1, p. 738-740.
- ALUÍSIO, S. M. *et al*. How to Learn the Many Unwritten 'Rules of the Game' of the Academic Discourse: A Hybrid Approach Based on Critiques and Cases to Support Scientific Writing In: IEEE International Conference on Advanced Learning Technologies 2001, 2001, Madison. *IEEE International Conference on Advanced Learning Technologies*. Loalamos CA: IEEE Computer Society, 2001. v.1. p.257 – 260.
- ALUÍSIO, S.M.; OLIVEIRA JR., O.N. (1996). A Detailed Schematic Structure of Research Papers Introductions: An Application in Support-Writing Tools. *Revista de la Sociedad Espanyola para el Procesamiento del Lenguaje Natural*, n. 19, pp. 141-147.
- Analytical Chemistry. Disponível em: <<http://pubs.acs.org/journal/ancham>> Acesso em: 23/11/2011.
- BIBER, D.; CONRAD, S.; REPPERN, R. **Corpus linguistics**: investigating language

- structure and use. Cambridge: Cambridge University Press, 1995.
- BERBER SARDINHA, Tony. **Linguística de Corpus**. Barueri: Manoele, 2004.
- FELTRIM, V.D.; ALUÍSIO, S.M.; NUNES, M.G.V. (2000). **Uma Revisão Bibliográfica sobre a Estruturação de Textos Científicos em Português**. Relatório Técnico, Série Computação, n.120, ICMC-USP, São Carlos.
- FELTRIM, Valéria. *et al.* Applying Argumentative Zoning in an Automatic Critiquer of Academic Writing. In: 17º BRAZILIAN SYMPOSIUM ON ARTIFICIAL INTELIGENCE, 2004, São Luis, Maranhão. Proceedings of the 17o Brazilian Symposium on Artificial Intelligence. 2004. v. 1, p. 214-223.
- MARQUIAFÁVEL, Vanessa Silva. Um processo para a geração de recursos linguísticos aplicáveis em ferramentas de auxílio à escrita científica. 2007. 273 f. (Dissertação) – (Mestrado em Linguística) – Universidade Federal de São Carlos, São Carlos, 2007.
- Journal of the American Chemical Society. Disponível em: <<http://pubs.acs.org/journal/jacsat>> Acesso em: 27/08/2011.
- OLIVEIRA JR, Osvaldo N.; ZUCCOLOTO, Valtencir; ALUÍSIO, Sandra M. Developing strategies to produce better scientific papers: a Recipe for non-native users of English. 2006. Disponível em: < <http://arxiv.org/abs/cs/0611013/>>. Acesso em: 13 fev. 2011.
- ORASAN, C. (2001): Patterns in scientific abstracts. In: Proceedings of Corpus Linguistics, 2001 Conference, Lancaster University, Lancaster, UK, pp. 433 – 443.
- SWALES, John M. **Genre analysis**: English in academic and research settings. Cambridge; Cambridge University Press, 1990.
- SWALES, John M; FEAK, Christine B. **Abstracts and the writing of abstracts**. Michigan; University of Michigan Press, 2009.