

Mineração de texto e multimodalidade representacional aplicadas à aprendizagem de conceitos científicos

Ana Paula Metz Costa

Universidade Federal do Rio Grande do Sul
Av. Paulo Gama, s/nº, prédio 12.201
CEP 90046-900 - Porto Alegre/RS - Brasil
55 51 3517-1552
anapaulametz@gmail.com

Eliseo Berni Reategui

Universidade Federal do Rio Grande do Sul
Av. Paulo Gama, s/nº, prédio 12.201
CEP 90046-900 - Porto Alegre/RS - Brasil
55 51 3308-3099
eliseoreategui@gmail.com

ABSTRACT

Although widely studied, the learning and teaching of basic scientific concepts still remains as a challenge due to the difficulties inherent in these processes, especially in the school's context. This paper presents a research project that investigate the implementation of a software based in text mining and graphic organization of information using multimodal representations as a resource in promoting conceptual learning achievements. The study will be conducted with students from final grades of Ensino Fundamental level (primary school), through the application of a *quasi*-experimental design that involves pre-test and post-test comparison with non-equivalent control group. The theoretical framework underpinning this research is based on the Ausubel's Theory of Meaningful Learning.

RESUMO

Apesar de amplamente estudados, a aprendizagem e o ensino de conceitos científicos básicos ainda apresentam-se como um desafio em função das dificuldades inerentes a esses processos, especialmente em ambiente escolar. Este artigo apresenta uma proposta de investigação sobre os efeitos da aplicação de um software baseado em mineração de texto e organização gráfica multirrepresentacional das informações como um recurso na promoção de ganhos na aprendizagem conceitual. O estudo será realizado com alunos de séries finais do Ensino Fundamental através da aplicação de uma metodologia *quasi*-experimental que consiste na comparação de pré-teste e pós-teste utilizando grupo controle não equivalente. O referencial teórico que fundamenta esta pesquisa é baseado na Teoria da Aprendizagem Significativa de Ausubel.

Descritor de Categorias e Assuntos

K.3.1 [Computing Milieux]: Computer Uses in Education – *Computer-assisted instruction (CAI)*.

Termos Gerais

Experimentation.

Palavras Chaves

SOBEK; Mineração de Texto; Conceitos; Aprendizagem Significativa.

1. EXTENDED ABSTRACT

A aprendizagem conceitual caracteriza-se pela inserção em uma longa tradição de pesquisas, ganhando e perdendo centralidade ao longo dos anos e através das perspectivas analíticas e sendo redefinida em função da ênfase em aspectos como o desenvolvimento cognitivo, a construção individual e social do conhecimento, a evolução do conhecimento científico, entre outros. Na área do Ensino de Ciências, uma extensa literatura científica aponta de maneira geral para a existência de dificuldades recorrentes enfrentadas por aprendizes e professores na aprendizagem e no ensino de conceitos científicos fundamentais [14].

De acordo com Pozo [15], a utilização eficaz dos conteúdos e princípios da Ciência pelos aprendizes de suas disciplinas dependeria de uma mudança nos processos e representações através dos quais se reflete e se aprende sobre os fenômenos científicos. Para este autor, os educadores científicos deveriam empenhar-se em promover uma *mudança representacional*, entendida como a tradução de uma representação a formatos e representações mais potentes e explícitos. A explicitação progressiva das concepções alternativas dos alunos, que são em boa medida implícitas na sua estrutura cognitiva, leva à construção de estruturas de informação mais estáveis, e nesse processo é de fundamental importância o domínio dos códigos comunicativos.

Não há como desvincular o processo de representação mental sobre os conhecimentos formalizados das maneiras pelas quais tanto alunos quanto a comunidade científica representam externamente estes conteúdos ou o que se refere a eles. Zompeiro e Laburu [22] apontam para o fato de que a linguagem científica utiliza diversos recursos semióticos que vão além dos códigos textuais e orais, abarcando a linguagem visual, os símbolos matemáticos e as operações experimentais, formatos com os quais nem sempre os aprendizes estão familiarizados. Autores como Yore e Hand [21], Tytler et al [20] e Hubber et al [8] afirmam que o engajamento dos professores na construção de uma fonte variada de recursos representacionais por parte dos alunos é fundamental para compreensão dos conceitos científicos básicos, habilitando-os a realizar aprendizagens futuras consistentes, a comunicar o conhecimento e a ter uma vivência efetiva das práticas científicas.

Estas preocupações se encontram inseridas no contexto das novas estruturas e práticas sociais influenciadas pelo desenvolvimento tecnológico que se deu de forma exponencial a partir da segunda metade do século XX e, particularmente, pela popularização de tecnologias voltadas para a produção, manipulação e comunicação de informação. Através de uma grande variedade de recursos, viabiliza-se a utilização de diversos sistemas de signos – oralidade, escrita, imagens, filmes, fórmulas matemáticas, etc. – de forma integrada para representar todo o tipo de informação, gerando novas formas de narrativa [4].

Este resumo apresenta um projeto de pesquisa cuja problemática envolve, portanto, as possibilidades dos recursos computacionais, que cada vez mais se difundem nos espaços educativos, a aprendizagem conceitual pensada a partir de concepções construtivistas e cognitivistas desse processo e as implicações de um enfoque representacional no ensino e na aprendizagem de Ciências. O objetivo deste estudo é a análise da utilização de um software baseado em mineração de texto e organização gráfica multirrepresentacional das informações como um recurso para a aprendizagem de conceitos científicos.

A pesquisa se ocupará das situações e conteúdos escolares de ensino de Ciências, com um recorte específico nas séries finais do Ensino Fundamental. O referencial teórico escolhido para fundamentar a investigação é baseado na Teoria da Aprendizagem Significativa de David Ausubel [1].

O programa SOBEK [16] tem como principal funcionalidade a busca de termos relevantes em um texto e a apresentação das informações geradas através de um organizador gráfico multirrepresentacional. Derivadas da área de estudo e desenvolvimento da Mineração de Dados, as estratégias de mineração de texto buscam padrões que identificam informações com significado relevante a partir de fontes de dados não estruturadas [5]. Na ferramenta SOBEK, implementa-se um algoritmo definido por Schenker [19] que realiza uma análise estatística dos termos presentes no texto e os seleciona a partir do valor absoluto de sua ocorrência. Este modelo de mineração textual, denominado *n-simple distance*, considera também as relações entre os componentes do texto, ligando cada termo estatisticamente relevante a outras N subsequentes palavras selecionadas.

Os dados extraídos do texto são apresentados em uma estrutura organizadora denominada grafo, uma forma de representar relações entre elementos de um determinado conjunto baseada na Teoria dos Grafos da Matemática. Estes organizadores são formados basicamente por dois componentes: os nós ou vértices, que contêm as informações que representam os elementos relacionados, e as arestas, linhas que representam o relacionamento entre os nós [2]. Após o processamento estatístico inicial, o software realiza uma busca na internet por imagens associadas a cada um dos termos relevantes selecionados e as apresenta no grafo, junto ao nó correspondente (Figura 1). Os grafos gerados pelo SOBEK são editáveis, de maneira que é possível excluir ou inserir nós, alterar as relações entre os nós e modificar as imagens relacionadas a cada termo.

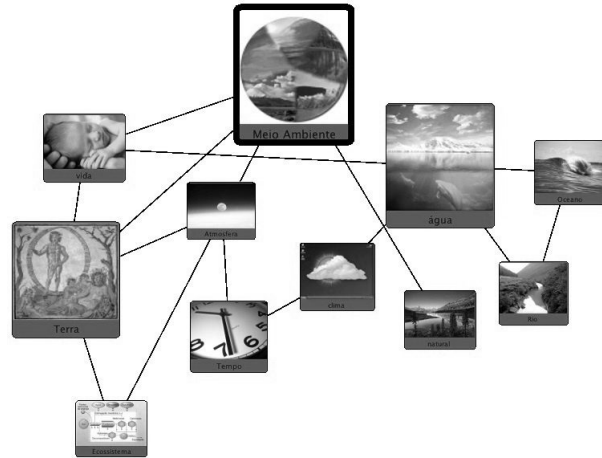


Figura 1. Grafo gerado na ferramenta SOBEK para um texto sobre meio ambiente

O programa foi inicialmente construído para auxiliar professores na avaliação do processo de produção textual coletiva, com o objetivo de minimizar o trabalho de análise da grande quantidade de material que é produzido neste tipo de atividade. A análise de seu uso para este fim revelou que o programa permite uma avaliação qualitativa eficiente dos textos, favorecendo assim o planejamento de práticas pedagógicas em função do desempenho dos aprendizes [10].

A utilização do aplicativo como ferramenta de aprendizagem já foi proposta e analisada em estudos anteriores, como no trabalho de Reategui e colaboradores [17], que averiguou seu uso como apoio à sumarização de textos. Também foram avaliadas as contribuições do trabalho com o SOBEK em atividades de produção textual [9] e de leitura e interpretação de textos [7].

No presente estudo, a utilização do software está alicerçada nos princípios propostos pelo psiquiatra norte-americano David P. Ausubel em sua Teoria da Aprendizagem Significativa. O ponto central desta teoria está no entendimento de que as informações são organizadas no cérebro humano formando uma hierarquia de conceitos. Ao ocupar-se do fenômeno da aprendizagem verbal, que considera preponderante no ambiente escolar, Ausubel trata com detalhes o processo de aprendizagem significativa, caracterizada pela ancoragem das novas as informações de forma não arbitrária e não literal nos subsunçores - as estruturas cognitivas já existentes - dando origem a estruturas diferenciadas [1].

Nessa proposta teórica, são elencadas três condições básicas para a ocorrência da aprendizagem significativa: além da existência e da ativação dos subsunçores apropriados, a disposição do aprendiz para este tipo de aprendizagem e a estruturação lógica do material, de forma que este possa ser relacionado com os conhecimentos prévios de maneira não arbitrária e não literal [1].

Joseph Novak, colaborador muito próximo de Ausubel, desenvolveu e disseminou o uso dos mapas conceituais como uma estratégia facilitadora da aprendizagem significativa. Definidos como diagramas bidimensionais que procuraram mostrar relações hierárquicas entre conceitos de um corpo de conhecimento, os

mapas conceituais são considerados por Novak como representações válidas da estrutura conceitual/proposicional do conhecimento de um indivíduo [12].

A utilização dos mapas conceituais no ensino tem sido amplamente estudada [11, 13]. Os mapas conceituais e os grafos gerados pelo software SOBEK apresentam algumas aproximações tanto em relação à sua estrutura quanto ao processo no qual se engaja o aluno quando utiliza tais ferramentas, de forma que suas implicações no processo de aprendizagem apresentam questões analíticas comuns.

A metodologia definida para esta investigação será a realização de um estudo *quasi*-experimental, com a aplicação de pré-teste e pós-teste com grupo controle não equivalente [3]. O estudo será desenvolvido com duas turmas de 6º ano e duas turmas de 7º ano de uma escola municipal de Educação Básica do município de Esteio, no estado do Rio Grande do Sul. O desempenho dos alunos nos pré-testes e pós-testes serão comparados entre turmas do mesmo ano.

Em uma primeira etapa de coleta de dados, será aplicado um pré-teste sobre um conteúdo a ser estudado para as turmas-controle e para as turmas experimentais, condição essa definida de forma aleatória. Em seguida, serão desenvolvidas sequências didáticas ambientadas no laboratório de informática da escola, planejadas a partir dos princípios da Teoria da Aprendizagem Significativa de Ausubel [1]. Após a realização de uma mesma introdução geral ao tema para todas as turmas, serão aplicadas nas turmas experimentais as intervenções envolvendo atividades de leitura de texto e utilização da ferramenta SOBEK como apoio para o entendimento do material de aprendizagem. Nas turmas-controle, serão desenvolvidas atividades baseadas em leitura de texto (o mesmo utilizado pelas turmas experimentais) e resolução de questionário em editores de texto. Após um intervalo temporal igual serão aplicados os pós-testes para todas as turmas.

Na segunda etapa da coleta de dados, utilizando um novo conteúdo de aprendizagem, serão invertidas as turmas controles e experimentais, de forma que a turma que recebeu a intervenção na primeira etapa será o novo grupo de controle. Serão utilizados pré-testes e pós-testes já validados por estudos anteriores sobre os conteúdos de aprendizagem selecionados.

Os dados coletados nos testes serão submetidos à categorização qualitativa a partir de categorias analíticas definidas em função dos elementos encontrados na literatura científica sobre os conhecimentos prévios e os obstáculos de aprendizagem dos temas estudados. As categorias serão elaboradas de acordo com a proposta de Gray [6] e receberão uma classificação de acordo com a adequação da resposta às convenções aceitas pela comunidade científica e com a capacidade de aplicar o conhecimento a outras situações. A distribuição das respostas e a comparação de desempenho entre as turmas controle e experimentais serão analisadas estatisticamente.

Os conteúdos abordados durante as intervenções deverão fazer parte do currículo oficial da escola e serão escolhidos em função da sequência de trabalho desenvolvida pela professora ao longo do ano letivo e da disponibilidade de pesquisas já publicadas sobre o tema. Com a intenção de diminuir os efeitos de adaptação aos métodos de ensino e ao uso do software, serão desenvolvidas preliminarmente práticas envolvendo a utilização de mapas conceituais e sua produção através do aplicativo Cmap Tools,

além de atividades envolvendo o aplicativo SOBEK planejadas de forma diferente das intervenções realizadas para a coleta de dados.

Espera-se que os resultados produzidos através deste estudo constituam uma contribuição válida para a pesquisa em Ensino de Ciências e para as práticas desenvolvidas por professores na intenção de superar os obstáculos concernentes à aprendizagem de conceitos científicos. A utilização de abordagens pedagógicas baseadas em pressupostos teóricos que afirmem sua eficácia é fundamental na busca por alternativas de práticas condizentes com as atuais demandas educacionais. Dessa forma, buscou-se um suporte sólido para a proposta de utilização do software SOBEK na promoção da aprendizagem conceitual, com a intenção de avaliar e fomentar novos recursos disponíveis para o desenvolvimento das habilidades e saberes necessários ao compartilhamento efetivo da cultura científica.

2. REFERENCIAS

- [1] Ausubel, D. P., Novak, J.D., & Hanesian, H. (1980) *Psicologia educacional* (2a ed., E. Nick, H. B. C. Rodrigues, L. Peotta, M.A. Fontes, & M. G. R. Maron, Trad.). Rio de Janeiro: Interamericana.
- [2] Berry, J. A.; Linoff, G. (1997). *Data Mining Techniques for Marketing, Sales and Customer Support.*: Wiley.
- [3] Campbell, D. T.; Stanley, J. C. (1963) *Experimental and quasi-experimental designs for research*. Boston: Houghton Mifflin.
- [4] Coll, C; Monereo, C. (2010) Educação e aprendizagem no séc XXI: novas ferramentas, novos cenários, novas finalidades. In: *Psicologia da Educação Virtual: Aprender e Ensinar com as Tecnologias da Informação e da Comunicação*. Porto Alegre: Artmed.
- [5] Feldman R.; Sanger, J. (2006) *Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data*. New York, NY, USA: Cambridge University Press.
- [6] Gray, D. E. (2012) *Pesquisa no mundo real*. Porto Alegre: Penso.
- [7] Hessler, J. D. C.; Reategui, E. (2011). Um método para apoio à leitura baseado no uso de uma ferramenta de mineração de texto. *RENTE*, 8(3).
- [8] Hubber, P.; Tytler, R.; Haslam, F. (2010) Teaching and learning about force with a representational focus: Pedagogy and teacher change. *Research in Science Education*, v. 40, n. 1, p. 5-28.
- [9] Klemann, M., Reategui, E., & Lorenzatti, A. (2009). O Emprego da Ferramenta de Mineração de Textos SOBEK como Apoio à Produção Textual. In *Anais do Simpósio Brasileiro de Informática na Educação*, Vol. 1, No. 1.
- [10] MACEDO, Alexandra L. **Rede de conceitos: uma ferramenta para contribuir com a prática pedagógica no acompanhamento da produção textual coletiva**. Porto Alegre, 2010. Tese (doutorado em Informática na Educação) – Universidade Federal do Rio Grande do Sul, Programa de Pós-Graduação em Informática na Educação, Porto Alegre, 2010.
- [11] Mendonça, C. A. S. (2012). O uso do mapa conceitual progressivo como recurso facilitador da aprendizagem significativa em Ciências Naturais e Biologia. Tese

(Programa Internacional De Doctorado Enseñanza De Las Ciencias) – Universidad de Burgos, Burgos

- [12] Moreira, M. A. (1999). Teorias de aprendizagem. Editora Pedagógica e Universitária
- [13] Nesbit, J. C., & Adesope, O. O. (2006). Learning with concept and knowledge maps: A meta-analysis. *Review of educational research*, 76(3), 413-448.
- [14] Özdemir, G., & Clark, D. B. (2007). An overview of conceptual change theories. *Eurasia Journal of Mathematics, Science & Technology Education*, 3(4), 351-361.
- [15] Pozo, J. I. (1999) Más allá del cambio conceptual: el aprendizaje de la ciencia como cambio representacional. En: *Enseñanza de las Ciencias*, V.17, Nº 3, Madrid, Espanha pp. 513-520.
- [16] Reategui, E.; Epstein, D.; Lorenzatti, A.; Klemann, M. (2011) Sobek: a Text Mining Tool for Educational Applications. In: *International Conference on Data Mining*, 2011, Las Vegas, Estados Unidos. *Anais do DMIN '11*. p. 59-64.
- [17] Reategui, E., Klemann, M., & Finco, M. D. (2012) Using a Text Mining Tool to Support Text Summarization. In *Advanced Learning Technologies (ICALT)*, 2012 IEEE 12th International Conference on (pp. 607-609).
- [18] Sasseron, L.H. & Carvalho, A. M. P. (2011) Alfabetização científica: uma revisão bibliográfica. Em *Investigações em Ensino de Ciências*, v. 16, n. 1, p. 59-77.
- [19] Schenker, A. (2003). Graph-theoretic techniques for web content mining (Doctoral dissertation, University of South Florida).
- [20] Tytler, R., Prain, V., & Peterson, S. H. (2006). Picturing evaporation: learning Science literacy through a particle representation. *Teaching Science*, v. 52, n. 1, p. 12-17.
- [21] Yore, L. D., & Hand, B. (2010). Epilogue: Plotting a research agenda for multiple representations, multiple modality, and multimodal representational competency. *Research in Science Education*, 40(1), 93-101.
- [22] Zompero, A. F. & Laburu, C. E. (2011) Significados de fotossíntese apropriados por alunos do ensino fundamental a partir de uma atividade investigativa mediada por multimodos de representação. Em: *Investigações em Ensino de Ciências* V. 16, Nº 2, Porto Alegre, Brasil pp. 179-199.