

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL – UFRGS

FACULDADE DE EDUCAÇÃO

PROGRAMA DE PÓS-GRADUAÇÃO EM EDUCAÇÃO

Miriam Noering Klemann

**Apoio à Produção Textual por meio do Emprego
de uma Ferramenta de Mineração de Textos**

Porto Alegre - RS

Maiο – 2011

Miriam Noering Klemann

**Apoio à Produção Textual por meio do Emprego
de uma Ferramenta de Mineração de Textos**

“Dissertação apresentada ao Programa de Pós
Graduação em Educação da Faculdade de
Educação da Universidade Federal do Rio
Grande do Sul, como requisito para obtenção
do título de Mestre em Educação”

Orientador:

Prof. Dr. Eliseo Berni Reategui

Porto Alegre - RS

Maior – 2011

Miriam Noering Klemann

**Apoio à Produção Textual por meio do Emprego
de uma Ferramenta de Mineração de Textos**

“Dissertação apresentada ao Programa de Pós Graduação em Educação da Faculdade de Educação da Universidade Federal do Rio Grande do Sul, como requisito para obtenção do título de Mestre em Educação”

Aprovada em 10 de agosto de 2011

Prof. Dr. Eliseo Berni Reategui (Orientador)

Profa. Dra. Liliana Maria Passerino (CINTED/UFRGS)

Profa. Dra. Clevis Elena Rapkiewicz (DEBAS/UFRGS)

Profa. Dra. Carla Beatris Valentini (NID/ UCS)

AGRADECIMENTOS

Ao concluir este trabalho gostaria de agradecer...

... ao meu orientador Professor Eliseo Reategui, pela paciência e incentivo, pelas correções e disponibilidade, pela incansável dedicação, pelo seu exemplo e por compartilhar seu conhecimento.

... aos professores das disciplinas que cursei na Pós-Graduação em Educação desta Universidade, pelos valiosos momentos de troca e aprendizado.

... ao Wilson, pelo auxílio e por entender as ausências.

... ao Thomas, por fazer parte desde o início desta trajetória e pela permanente torcida.

... às Professoras Carla e Vera do IEE Ernesto Alves pelo acolhimento e pela confiança.

... a todos os que, de alguma forma ou de outra, colaboraram e ficaram na torcida para que este trabalho pudesse se concretizar. Obrigada pelo apoio e por me mostrar que vale a pena ir além...

RESUMO

Esta pesquisa apresenta a proposta de utilização da ferramenta SOBEK no apoio à produção textual através de um processo interativo. Trata-se de uma abordagem para apoio à construção textual empregando uma ferramenta de mineração de textos, considerando as dificuldades que envolvem a expressão principalmente escrita do aluno. A metodologia idealizada para implementação da pesquisa foi baseada em tarefa. Para validar esta metodologia, foi feita uma análise de todo o processo de produção textual, e não apenas do seu resultado, contemplando os estágios de pré-escrita e escrita. Destacamos que já no estágio de pré-escrita, propôs-se uma série de passos que permitiram ao estudante se apropriar do tema, estruturar suas idéias, planejar e preparar-se para a tarefa, para só então partir para o estágio da escrita do texto. Um experimento foi realizado com uma turma de alunos do segundo ano do Curso Normal de Nível Médio, onde uma atividade de leitura e produção textual foi proposta, sendo decomposta em várias tarefas. Foram feitos diferentes registros da participação dos alunos no experimento, como vídeo com a captura de todas as ações dos alunos nos sistemas utilizados, bem como questionário com questões fechadas e abertas. Na fase de análise dos dados coletados, cada tarefa foi analisada separadamente. Os resultados reportaram que uma das estratégias adotadas pelos alunos na busca por uma compreensão mais aprofundada do texto, foi a releitura (integral ou parcial) incentivada pelos questionamentos decorrentes da interpretação dos grafos apresentados. As evidências identificadas estavam principalmente relacionadas às “idas e vindas” ao texto original. Outra estratégia identificada, relacionada à comparação que o aluno fez do grafo inicial gerado, foi a adição ou remoção de conceitos na base de conceitos e geração de um novo grafo. As observações feitas pelos estudantes indicaram uma fluência na compreensão do texto original e estruturação das idéias com relação à forma como conceitos relevantes no texto se relacionavam, o que auxiliou no planejamento e preparação para a tarefa final - escrita do texto pelo aluno. Foi observado que as estratégias adotadas pelos alunos utilizando a ferramenta de mineração de texto lhes permitiram escrever de maneira coesa, articulada, apresentando aspectos relevantes e adequados sobre o tema lançado. A principal contribuição deste trabalho residiu na elaboração da metodologia para emprego da ferramenta de mineração de texto como apoio à produção textual e demonstração de seu potencial quando aplicada no contexto escolar.

Palavras chave: produção textual, mineração de textos

ABSTRACT

This work proposes the use of the tool SOBEK to support writing in an interactive process. This is an approach to text writing with the support of a text mining tool, considering the difficulties involved in the written expression of students. The methodology conceived for the implementation of this research has been based on tasks. To validate this methodology, the whole text production process has been analysed, and not only its results, including the pre-writing and writing stages. We emphasize that in the pre-writing stage, a sequence of steps has been proposed to enable students to better understand the theme, structure the ideas, plan and feel better prepared for the actual task of writing. An experiment has been carried out with second year high school students, where a reading and writing activity has been proposed, being decomposed in several tasks. The students' participation in the tasks were recorded in different ways, such as video capture of all their actions in the mining system and text editor, and a questionnaire with multiple selection and open questions. In the analysis phase of the data collected, each task was analysed separately. The results reported that one of the strategies adopted by the students in the search for a better understanding of the text was the act of re-reading it (completely or partially), incentivated by the questioning resulting from the interpretation of the graphs presented. The evidence identified has been mainly related to the number of times the students were observed going back and forth to the original text. Another strategy observed, related to the comparison the students made with the original graph generated, has been the adding or deleting of terms from the system's concept base when generating a new graph. The observations made by the students indicated a fluency in the understanding of the original text and in the structuring of their ideas in relation to the ways with which the relevant concepts in the text were related, which has supported the planning of the final task – the actual writing of the text. It has been observed that the strategies adopted by the students using the text mining tool has enabled them to write in a cohesive and articulated way, presenting appropriate and relevant aspects of the theme considered. The main contribution of this work has been the development of the methodology for using the text mining tool as a support for text writing, and the verification of its potential when applied in the educational context.

Keywords: text production, text mining

LISTA DE QUADROS

Quadro 1: Detalhamento da 1ª categoria de observação.....	58
Quadro 2: Detalhamento da 2ª categoria de observação.....	60
Quadro 3: Detalhamento da 3ª categoria de observação.....	61
Quadro 4: Detalhamento da 4ª categoria de observação.....	63
Quadro 5: Detalhamento da 5ª categoria de observação.....	64
Quadro 6: Detalhamento da 6ª categoria de observação.....	68
Quadro 7: Detalhamento da 7ª categoria de observação.....	70

LISTA DE TABELAS

Tabela 1- Dados iniciais Textalyser.....	35
Tabela 2 - Resultados do programa Wordcounter	37
Tabela 3 -Tabela termos gerados e termos utilizados nos textos.....	73

LISTA DE FIGURAS

Figura 1 - Mapa Conceitual elaborado por Dutra	29
Figura 2 - Ferramenta TextAlyser - disposto em http://textalyser.net	34
Figura 3 - Ferramenta Wordcounter - disponível em http://www.wordcounter.com/	36
Figura 4 - Ferramenta Tagcrowd disposto em http://tagcrowd.com/	38
Figura 5 – TagCrowd	39
Figura 6 – Grafo	41
Figura 7 – Tela Ferramenta SOBEK –texto “A coerência textual”	42
Figura 8 - Grafo gerado para o texto sobre o tema " A coerência textual "	43
Figura 9 - foto laboratório IEE Ernesto Alves - acervo pesquisadora	55
Figura 10 - Organograma das Tarefas.....	57
Figura 12 - Grafo gerado a partir do texto "Realismo" - Aluno 3	65
Figura 13 - Grafo gerado a partir do texto "Realismo" - Aluno 5	65
Figura 14 - Grafo gerado a partir do texto "Realismo" - Aluno 6	66
Figura 15 - Grafo gerado a partir do texto "Realismo" - Aluno 7	66
Figura 16 - Tela ferramenta SOBEK - Configurações.....	87
Figura 17 - Tela ferramenta SOBEK Base de Dados - gerado a partir do texto "A coerência textual"	87

LISTA DE SIGLAS

CINTED – Centro Interdisciplinar de Novas Tecnologias na Educação

ENEM - Exame Nacional do Ensino Médio

ETC – Editor de Texto Coletivo

FACED - Faculdade de Educação

MEC – Ministério da Educação e Cultura

NUTED – Núcleo de Tecnologia Digital Aplicada à Educação

OCDE - Organização para a Cooperação e Desenvolvimento Econômico

PCN - Parâmetros Curriculares Nacionais

PEAD - Curso de Pedagogia a Distância

PDE - Plano de Desenvolvimento da Educação

PISA - Programa Internacional de Avaliação de Alunos

*(*Programme for International Student Assessment)*

PPGEDU – Programa de Pós-Graduação em Educação

PPGIE – Programa de Pós-Graduação em Informática na Educação

SAEB - Sistema de Avaliação da Educação Básica

UFRGS – Universidade Federal do Rio Grande do Sul

UNERJ - Centro Universitário de Jaraguá do Sul

SUMÁRIO

SUMÁRIO	6
PRÓLOGO	7
Apresentação da trajetória da pesquisadora	7
1 INTRODUÇÃO	11
2 PRODUÇÃO TEXTUAL	17
3 CONSTRUÇÃO TEXTUAL APOIADA NAS TECNOLOGIAS DIGITAIS	26
3.1 MAPAS CONCEITUAIS	28
3.2 A MINERAÇÃO DE TEXTOS PARA ACOMPANHAMENTO DA ESCRITA COLABORATIVA	31
3.3 O PROCESSO DE LEITURA E ESCRITA APOIADO PELA TECNOLOGIA.....	32
3.3.1 TextAlyser	34
3.3.2 Wordcounter	35
3.3.3 TagCrowd	37
3.4 FERRAMENTA DE MINERAÇÃO DE TEXTOS – SOBEK.....	39
3.4.1 Estudos anteriores utilizando a ferramenta SOBEK	43
4 METODOLOGIA PARA APOIO À PRODUÇÃO TEXTUAL EMPREGANDO A FERRAMENTA SOBEK	46
4.1 METODOLOGIA DA PESQUISA	49
4.2 EXPERIMENTO PRELIMINAR	51
4.3 EXPERIMENTO PRINCIPAL.....	54
4.3.1 Tarefas de pré-escrita	58
4.3.2 Tarefas de escrita	68
5 CONSIDERAÇÕES FINAIS	74
REFERÊNCIAS	77
ANEXOS	84

PRÓLOGO

APRESENTAÇÃO DA TRAJETÓRIA DA PESQUISADORA

Professora de Educação Infantil e Ensino Fundamental – anos iniciais de 1982 (após o término do curso de magistério - segundo grau – Professora do Ensino de 1º Grau de 1ª a 4ª série) até 2008.

Inicialmente atuei em Pomerode - Santa Catarina e a partir de 1990 em Santa Cruz do Sul – RS em colégios da Rede Sinodal de Educação¹. Desde 1991 despertou a minha curiosidade sobre o uso do computador no ambiente escolar, pois havia uma estrutura ótima no colégio para aquela época e com uma história de oito anos na informática educativa².

A partir de 1992 passei a trabalhar como professora de informática neste estabelecimento. Um período rico de experimentações onde, mesmo com pouca teorização a respeito, de forma intuitiva usava a disciplina como laboratório de práticas pedagógicas.

Busquei minha qualificação para o uso e produção de softwares educacionais que pudessem atender a demanda. Essa necessidade, a vontade de compreensão e a necessidade de mudanças levaram-me a perceber o impacto da informática no

¹ A Rede Sinodal de Educação é formada por 58 escolas e tem a função de expressar o seu vínculo com a Igreja Evangélica de Confissão Luterana do Brasil (IECLB). A sua missão é “Como escolas comunitárias evangélicas buscar o desenvolvimento do senso crítico, da criatividade e da integridade da pessoa, na dimensão de sua vocação para servir.” A Rede Sinodal de Educação abrange o setor educacional escolar na IECLB. Está constituída pela Assembléia da Rede Sinodal de Educação, pelo Conselho de Educação e pela Direção-Executiva. Reúne escolas localizadas no Rio Grande do Sul, Santa Catarina, Paraná, São Paulo, Rio de Janeiro e Mato Grosso.

² Em 1983, Colégio Mauá, em atividade pioneira, sediou por três dias, o Seminário sobre Informática – título: “INFORMÁTICA HOJE” o primeiro no interior do Estado, sob a responsabilidade da *Advancing Computer Shop*. Na época com cerca de 500 participantes. Este primeiro seminário foi um sucesso total. Colégio Mauá adquiriu um microcomputador – CP200 - que era acoplado num televisor com gravadores de fita cassete para ser usado no ensino, inicialmente com seis programas: matemática, xadrez, estatística, e jogos educativos. Logo em seguida, aconteceu o primeiro curso de Linguagem Básica para programação de micro-computadores. Nesta época já se destacava a importância do computador na Educação.

ambiente educacional. Durante alguns anos desenvolvi projetos com crianças no programa LOGO³ utilizando inicialmente computadores MSX e a partir de 1993, os PS1 da IBM. A linguagem LOGO foi escolhida para o trabalho com a computação das séries iniciais tendo em vista sua fácil compreensão, e diversas atividades foram desenvolvidas em diferentes níveis utilizando a programação. Foi uma experiência muito gratificante e, apesar das limitações das máquinas e pelo fato de não dispor de muitas alternativas de *software* havia a necessidade de produzir algo que fosse útil e interessante para as séries iniciais. Com ajuda de alguns professores, em 1994, foi possível criar alguns *software* utilizando a linguagem LOGO. Destaco o “*Software* Educacional Mauá” – Letras do Alfabeto⁴. Provando que poderíamos fazer “parcerias” entre as disciplinas.

Em 2003 – me formei em Licenciatura em Computação – Licenciatura Plena⁵, pela Universidade de Santa Cruz do Sul – RS. O primeiro curso da área no estado e segundo do Brasil. Neste período, além de freqüentar as aulas normalmente também participei de vários cursos, fóruns e seminários. Meu trabalho de conclusão de curso foi “Construção de Conceitos sobre Medidas utilizando o Ambiente Megalogo” onde desenvolvi, ao longo de três meses com duas turmas de 2ª série, alguns conceitos matemáticos utilizando computador, noções iniciais sobre

³ LOGO é uma linguagem de programação interpretada, voltada principalmente para crianças, jovens e até adultos. É utilizada com grande sucesso como ferramenta de apoio ao ensino regular e por aprendizes em programação de computadores. Ela implementa, em certos aspectos, a filosofia construtivista, segundo a interpretação de Seymour Papert, co-criador da linguagem junto com Wally Feurzeig. Papert, matemático que trabalhou com Jean Piaget (donde a idéia da filosofia construtivista), é co-fundador do Media Lab no Massachusetts Institute of Technology (MIT). No ambiente LOGO o aluno aprende com seus erros. Aprende vivenciando e se algo está errado em seu raciocínio, isto é claramente percebido e demonstrado na tela, fazendo com que o aluno pense sobre o que poderia estar errado e tente, a partir dos erros vistos, encontrar soluções corretas para os problemas. (*disposto em <http://pt.wikipedia.org/wiki/Logo>)

⁴ Os objetivos deste programa, além de fixar as letras do alfabeto, era associar as letras cursivas com as letras de imprensa, reconhecer as letras da tela e localizá-las no teclado, e também memorizar o traçado das letras maiúsculas e minúsculas.

⁵ O Curso de Licenciatura em Computação da UNISC é um curso que tem como objetivo a preparação de profissionais altamente capacitados para o mercado de trabalho. Prepara tanto para docência em Educação Básica e suas modalidades, como para atuação em área empresarial. São profissionais de espírito crítico, com conhecimento das últimas tendências em computadores e suas tecnologias. Possuem ampla formação teórica, tecnológica, pedagógica e humana valorizando a autonomia, a responsabilidade, a ética e a solidariedade.

ângulos, organização espacial, raciocínio lógico, conceitos básicos sobre MegaLogo⁶ visando aplicações práticas utilizando medidas.

Ainda pela Rede Sinodal, neste mesmo período, foram acontecendo encontros de Professores de Informática (Regionais e Nacionais) com temas variados focando a utilização de ferramentas de Informática no processo Ensino/Aprendizagem, Informática na Interdisciplinaridade, o papel da Informática no processo de aprendizagem, ensino on-line – Uma nova forma de Comunicação.

Em 2005, surgiu o Projeto Movida⁷ que é mantido pela Sociedade Escolar de Santa Cruz, mantenedora do Colégio Mauá, onde sou voluntária. É uma iniciativa na área da Responsabilidade Social. A minha contribuição vem sendo na área de informática – com enfoque na preparação para o mercado de trabalho e também na oficina de geração de renda – *Patchwork* e bordado.

Ainda em 2005 surgiu uma necessidade de compreender as outras inserções da informática no meio educacional o que me estimulou a cursar uma especialização, o que aconteceu no período de abril/ 2005 a dezembro/2006 – no

⁶ O MegaLogo é um ambiente de programação e aprendizagem multimídia, permitindo realizar atividades centradas na aprendizagem, na cooperação e na descoberta. É baseado na linguagem de programação Logo e desenvolvido para o ambiente Microsoft Windows, suportando todos os recursos desse ambiente operacional: som, vídeo, imagens, cores, fontes e impressão, entre outros. A programação é desenvolvida de modo estruturado, como no Logo convencional, mas sem usar o chamado "flip side" e sim uma janela de memória, dividida em duas partes: do lado esquerdo, com o nome dos procedimentos relacionados; do direito, o procedimento propriamente dito, permitindo sua edição, correção ou simples visualização. Quando utilizado pedagogicamente com a base teórica construtivista coopera para o desenvolvimento cognitivo das crianças e adolescentes, auxiliando também no desenvolvimento do raciocínio lógico matemático. (*disposto em <http://pt.wikipedia.org/wiki/MegaLogo>)

⁷ O Projeto MOVIDA desenvolve ações de responsabilidade social junto a uma das comunidades carentes do município de Santa Cruz do Sul. Iniciado em março de 2005 o projeto MOVIDA tem como principal objetivo fortalecer os vínculos familiares e escolares de crianças e adolescentes moradores dos bairros focais e que se encontrem em situação de vulnerabilidade social. A metodologia e implementação do projeto foram baseadas no diagnóstico da realidade sócio-econômica da região, que apontou como principal problema o desemprego; a maior parte dos moradores da região desenvolve atividades na qualidade de safristas junto à indústria de tabaco. O público alvo do Projeto MOVIDA são crianças, adolescentes, jovens e adultos, e sua principal atividade são oficinas de educação complementar, formação para o trabalho e geração de renda. O projeto desenvolve 17 oficinas, que atendem a uma média semanal de mais de 200 pessoas. Nesse processo estão envolvidos 58 voluntários do Colégio Mauá, entre eles alunos, professores, funcionários, pais, direção, além de pessoas da comunidade e de organizações parceiras.

curso: Especialização em Informática na Educação – nível de pós-graduação a distância pela Universidade Federal do Rio Grande do Sul – UFRGS – RS. Conheci pessoas e realidades diferentes de vários estados do Brasil. Na minha monografia, sob o título: “Resgate da Trajetória Tecnológica no Olhar da Construção de Textos”, faço um resgate histórico da trajetória do Colégio Mauá apresentando alguns relatos e propostas de atividades para estimular expressão escrita através da utilização das Tecnologias de Informação, ao longo de 13 anos (período entre 1992 a 2005), acompanhando a própria evolução da tecnologia. Desta forma, apresento um olhar sobre a produção textual envolvendo crianças desde Educação Infantil até Ensino Fundamental – séries iniciais.

Esta experiência me motivou para realização desta pesquisa envolvendo uma ferramenta computacional como apoio a produção textual, tema de interesse deste estudo. Hoje percebo que os alunos iniciam cedo o contato com as tecnologias e propor tarefas diferenciadas é fundamental. A tarefa de escrita, por exemplo, nem sempre é uma tarefa simples para aqueles que escrevem, podendo implicar dificuldades de diferentes naturezas. Ao mesmo tempo, o avanço da tecnologia, exige que o professor esteja continuamente pesquisando, se atualizando para que possa desenvolver e criar atividades desafiadoras para seus alunos utilizando metodologias adequadas.

1 INTRODUÇÃO

Este estudo apresenta uma proposta para o uso de uma ferramenta computacional com o objetivo de auxiliar nas atividades de produção de textos através de um processo interativo. Trata-se de uma nova abordagem para apoio à construção textual empregando uma ferramenta de mineração de textos. Esta abordagem foi fundamentada no ensino baseado em tarefas (ELLIS, 2003) considerando as dificuldades que envolvem principalmente a expressão escrita do aluno.

Desde o início da década de 80, o ensino de Língua Portuguesa na escola tem sido o centro da discussão acerca da necessidade de melhorar a qualidade da educação no País, conforme dados do Ministério da Educação e Cultura (PCN, 2001, livro 02, p. 19). Estes dados apontam também que já no ensino fundamental, no que se refere ao fracasso escolar, a questão da leitura e da escrita tem sido um dos principais eixos de discussão. Vieira (1981) constata, a partir de seus estudos, que existe uma estreita relação entre a falta de leitura e a dificuldade de redação. O autor aponta como principal problema a não identificação das idéias centrais do texto bem como a ausência de clareza e coerência na redação.

Tal dificuldade está relacionada à falta de contato dos alunos com textos escritos, ou seja, à falta de leitura. Não é possível escrever bem sobre o que se desconhece ou sobre aquilo que se conhece pouco. Para Marcuschi (2005) o problema da expressão escrita, está associado à linguagem que o indivíduo usa, pois ele normalmente utiliza o código restrito para falar e, para escrever, necessita do código mais elaborado.

Da Silva, Fagundes e Basso (2007) apresentam alguns índices revelados por avaliações nacionais e internacionais, como o Programa Internacional de Avaliação de Alunos (*Programme for International Student Assessment- PISA*), o Sistema de Avaliação da Educação Básica (SAEB), a Prova Brasil, o Exame Nacional do Ensino Médio (ENEM), os quais têm trazido baixos indicadores com relação ao

desempenho dos alunos na educação básica, além de indicadores sobre o analfabetismo e sobre o ensino superior. Estes autores comentam que a gravidade do problema pode ser observada nos resultados que as escolas brasileiras obtiveram na avaliação PISA, onde têm ocupado as últimas posições. Em 2000, dos 31 países que foram comparados no primeiro ciclo, o Brasil ocupou a 31ª posição em leitura. E em 2001 e 2003, 41 países foram comparados e o Brasil ficou na 37ª posição em leitura. Os autores também apresentam dados sobre o rendimento dos alunos no ENEM em 2006. Numa escala de zero a cem, os participantes obtiveram 52,08 na redação. E ainda, dados do SAEB, também apontam problemas no Ensino Médio (EM), pois em 2001 e 2003, do percentual de estudantes da 3ª série do EM, somente 5,3% estavam no estágio adequado de construção de competências em Língua Portuguesa.

No relatório de avaliação divulgado em dezembro/2010⁸ pela OCDE (Organização para a Cooperação e Desenvolvimento Econômico), que tornou público o resultado do PISA 2009, o Brasil aparece entre os três países que mais evoluíram na educação nesta década. A evolução registrada em 2009 pode ser considerada significativa, pois o País foi o terceiro que mais cresceu, mas ainda é insuficiente. Em leitura, por exemplo, quase 50% dos estudantes ainda alcança apenas o nível mínimo (nível 1, sendo que o nível mais alto é 6 e apenas 20 alunos alcançaram este nível), o que significa que a grande maioria é capaz apenas de encontrar informações explícitas nos textos. Estes dados não devem servir somente para trazer constatações, na realidade são diagnósticos que nos levam a repensar nossa prática pedagógica no sentido de superá-los.

O computador e as tecnologias, que estão muito presentes no nosso cotidiano, podem auxiliar na resolução de alguns problemas apontados pelas pesquisas. O educador deve conhecer o que a tecnologia tem a oferecer e como pode ser explorada em diferentes situações educacionais. Com relação ao computador,

⁸ Disposto em http://www.inep.gov.br/imprensa/noticias/internacional/news10_02.htm, acesso 07/12/2010

existem diferentes aplicações que podem ser exploradas, dependendo do que está sendo estudado ou dos objetivos que o professor pretende atingir (VALENTE, 2002). Para Castells (2003, p.160), “... o que a tecnologia tem de maravilhoso é que as pessoas acabam fazendo com ela algo diferente daquilo para que foram originalmente criadas”.

Jolibert (2006) propõe a produção de textos a partir de situações reais de comunicação. Existem várias etapas pelas quais os alunos passam antes de escreverem, o que a autora chama de “obra-prima”, ou seja, o texto pronto. Para ela aprender a produzir textualmente é, desde o início, aprender a organizar idéias e elaborar um texto que tenha significado real, dotado de sentido, contextualizado e proposto mediante situações significativas.

Neste sentido, cabe ao professor desenvolver atividades de produção textual que estejam mais próximas da realidade do aluno e assim ele possa estabelecer uma relação entre a produção escrita e seu(s) destinatário(s). Não se trata apenas de uma elaboração de novos métodos, mas de uma reestruturação completa da maneira de pensar e organizar as idéias e apresentá-las em forma de texto, pois é a partir da comunicação escrita que organizamos os nossos pensamentos (FRANCO, 2007).

Para Jolibert (2006), escrever é reescrever. Um texto é produzido por camadas, com um ir e vir entre intenções do autor e a necessidade lingüística do texto. Ou seja, produzir é assumir a perspectiva de um determinado tipo de texto, dirigido a certo destinatário, com uma intencionalidade específica (JOLIBERT, 2006).

“Um escritor competente é alguém que planeja o discurso e conseqüentemente o texto em função do seu objetivo e do leitor a que se destina (...) é capaz de olhar para o próprio texto como um objeto e verificar se está confuso, ambíguo, redundante, obscuro ou incompleto. (...) é capaz de revisá-lo e reescrevê-lo até considerá-lo satisfatório para o momento. (...) é um leitor competente, capaz de recorrer, com sucesso, a outros textos quando precisa utilizar fontes escritas para a sua própria produção.” (Parâmetros Curriculares Nacionais – livro 02 – 1997, p. 47 e 48)

Para Hayes (2008), escrever não é apenas apresentar idéias prontas, elaboradas, mas também envolve a criação de novas idéias durante o processo de escrita. Escrever é um processo de descobertas (HAYES, 2008). Nesse sentido as

produções estão em constante relação com a vivência de cada indivíduo. Cabe portanto, a nós professores, viabilizar o acesso do aluno ao universo dos textos, orientá-los na sua produção e interpretação. Isso inclui textos de diferentes disciplinas, textos com conteúdos variados, com os quais o aluno se defronta no cotidiano, mas dos quais nem sempre consegue se apropriar. É preciso oferecer aos alunos inúmeras oportunidades de aprenderem a escrever em condições semelhantes às que caracterizam a escrita fora da escola com a finalidade formar escritores⁹ competentes capazes de produzir textos coerentes, coesos e eficazes. (PCN, 1997).

Na escola, desde os anos iniciais os alunos são preparados para desenvolver atividades escritas. Esta preparação passa por várias etapas e atividades diversas. A conquista da escrita alfabética não garante ao aluno a possibilidade de compreender e produzir textos em linguagem escrita. Essa aprendizagem exige um trabalho pedagógico sistemático.

Porém, nas séries finais do Ensino Médio, embora já tenham tido vários anos de contato com o ensino formal da língua tida como padrão, os alunos ainda apresentam dificuldades seja com relação à gramática, seja com relação ao conteúdo temático em si (ALVES, 2007).

Partindo desse contexto, definiu-se como principal objetivo deste trabalho:

- *Investigar como uma ferramenta de mineração de texto pode ser empregada como apoio à produção textual.*

A ferramenta escolhida para esta tarefa foi a ferramenta de mineração de textos SOBEK, idealizada como recurso de apoio aos professores no acompanhamento de trabalhos de escrita colaborativa (MACEDO et al., 2009). O principal motivo da escolha desta ferramenta foi o fato de ela já ter sido desenvolvida com um olhar educacional. Além disso, o fato da ferramenta ter sido implementada pelo grupo de pesquisa liderado pelo prof. Eliseo Reategui,

⁹ o termo “escritor” está sendo utilizado aqui para referir-se não a escritores profissionais e sim a pessoas capazes de redigir (PCN)

orientador deste trabalho, mostrou-se um grande facilitador para o desenvolvimento dos ajustes necessários para o uso da ferramenta como apoio à produção textual.

Partindo-se do objetivo principal, foram definidos os seguintes objetivos específicos:

- Desenvolver uma metodologia para utilização de um sistema de mineração de textos como ferramenta de apoio à produção textual.
- Propor melhorias no sistema de mineração de textos utilizado focando em sua aplicação como ferramenta educacional para apoio à produção textual.
- Disponibilizar a ferramenta de mineração de textos para outras aplicações e outras instituições.

Nas produções textuais são utilizadas estratégias variadas que envolvem diferentes tipos de texto. Pesquisas sobre assuntos diversos e de interesse dos alunos são explorados desenvolvendo-se nestes habilidades importantes tais como: construir hipóteses, argumentar, concluir, resumir, classificar, avaliar (estabelecer critérios e utilizá-los para selecionar informações importantes), comparar, decidir, organizar. Assim, os alunos desenvolvem habilidades necessárias para se tornarem eficientes leitores e redatores criativos.

O computador e o uso de programas educacionais podem favorecer o desenvolvimento da criatividade, bem como favorecer a construção de conhecimento (VALENTE, 1993) As atividades podem tornar-se mais interessantes e contribuir na manutenção do interesse do aluno pela atividade, dando-lhe a oportunidade de refazer, re-escrever, inserir, deletar, copiar, colar quantas vezes achar necessário e coerente. Copiar e colar fragmentos é atividade normal em qualquer escrita eletrônica e, a possibilidade de colagens, é uma atividade textual que se tornou possível e muito comum na escrita digital (JONSSON, 1997).

Sendo a leitura uma etapa importante no processo de produção escrita, a proposta de utilização da ferramenta de mineração de textos como apoio à produção textual busca possibilitar aos alunos um meio mais rico e dinâmico de exploração

dos textos lidos. Esta possibilidade de exploração dos textos através de uma ferramenta computacional tem como meta promover uma melhor compreensão dos textos e dos temas tratados pelos estudantes em suas produções escritas.

Assim, tendo em vista os objetivos desta pesquisa, o segundo capítulo do trabalho apresenta o arcabouço teórico que deu sustentação a esta proposta, no que diz respeito à teoria sobre produção textual (Hayes e Flower, 1983; Bronckart, 2003; Marcuschi, 2008; Bakhtin, 2000, Koch, 1998, Jolibert, 1994 entre outros) e construção textual apoiada nas tecnologias (Johnson,2001; Marcuschi, 2008; Ebecken, Lopes&Costa, 2003; Hessler e Reategui, 2010; Shanahan, 2008). O terceiro capítulo apresenta diferentes ferramentas de análise e mineração de textos com vistas a discutir sua possível utilização na Educação, incluindo a ferramenta SOBEK, foco deste trabalho.. Já o quarto capítulo apresenta a metodologia desenvolvida nesta pesquisa, cujo objetivo é de demonstrar como uma ferramenta de mineração de textos pode ser empregada como apoio à produção textual. A metodologia foi baseada no conceito de tarefas (Willis, 1996; Skehan, 1998; Ellis, 2003 e Nunan, 1989).

O quinto capítulo detalha os experimentos realizados com a metodologia desenvolvida e apresenta a análise dos resultados obtidos.

Finalmente, o sexto capítulo apresenta considerações finais, algumas implicações desta pesquisa e perspectivas para trabalhos futuros.

2 PRODUÇÃO TEXTUAL

A leitura e a produção textual são comumente tomadas como atividades essenciais no desenvolvimento intelectual e social dos estudantes. Entretanto, é comum a dificuldade destes em se expressar, tanto de forma oral quanto escrita. Por exemplo, Luft (1999, p.43) aponta que “... os jovens não sabem falar, ... não sabem escrever, a nova geração não tem vocabulário...”.

Estudantes em idade escolar e universitários apresentam estratégias imaturas para resolver a tarefa de compor textos, já os escritores experientes são, em geral, aqueles que possuem um repertório de estratégias adequadas para o desenvolvimento da produção de textos (HAYES & FLOWER, 1983).

Para Hayes & Flower (1980), a escrita é uma habilidade cognitiva complexa. Para estes autores, a escrita requer um conjunto de habilidades expressivas que permitem a produção de um texto fluente apoiado, principalmente, num conjunto de competências receptivas (que é o processo de leitura) e um conjunto de habilidades reflexivas (que é a interpretação de texto).

Conforme Bronckart (2003, p.72), “os textos são produtos da atividade humana e, como tais estão articulados às necessidades, aos interesses e às condições de funcionamento das formações sociais no seio das quais são produzidos.” A escrita é um processo histórico que nos acompanha desde a antiguidade, sendo assim, muito antes de existirem e serem usados os computadores, já existiam estudos sobre o desenvolvimento cognitivo e o processo de escrita de crianças, jovens e adultos. Esses estudos sobre a linguagem escrita são válidos até hoje porque nos dão a base para pensar sobre a escrita antes e depois do advento das Tecnologias de Informação e Comunicação (TICs). As possibilidades de escrita e de busca de informações nesta nova era da comunicação nos fazem repensar as metodologias que contribuem para este desenvolvimento, bem como, nos desafiam na busca de ambientes de aprendizagem mais ricos.

Valente (1999) defende que o computador pode enriquecer ambientes de aprendizagem onde o aluno, interagindo com os objetos desse ambiente, tem chance de construir o seu conhecimento. Um ambiente dinâmico e interativo possibilita práticas diferenciadas para que ocorra aprendizagem a partir dos processos de descobertas. Para este autor, quando uma idéia não produz os resultados esperados, novas estratégias são elaboradas fazendo com que novos conhecimentos sejam construídos. Todos os indivíduos são capazes de encontrar respostas para diversas situações, ou relacionando com experiências anteriores ou projetando uma “nova idéia” ou “nova solução”.

Quando o aluno chega à escola/universidade, traz consigo muitas experiências, pois já aprendeu a se expressar, reproduzir histórias e fatos que fazem parte do seu “mundo”. A cada momento, estamos aprendendo algo novo, palavras novas, significados novos. Quando o aluno aprende a escrever, começa a dominar um sistema lingüístico. Aos poucos, compreende a regularidade que se apresenta na escrita e começa a conviver com algumas regras. Uma das regras que vai aparecer bem rapidamente é que não escrevemos como falamos (OLIVEIRA, 2009).

Neste sentido a produção textual se assemelha a um jogo (MARCUSCHI, 2008), pois antes de iniciar temos que conhecer algumas regras, o espaço onde vai acontecer, qual o papel e função dos atores envolvidos. Também envolve algumas decisões e estratégias. Para o autor, o texto deve ser visto como uma seqüência de atos de linguagens (escritos ou falados) e não apenas como uma seqüência de frases de algum modo coesas.

As atividades de produção textual sempre têm um objetivo resultando em diferentes textos, com enfoques e estruturas diferentes. Escrever uma história, por exemplo, é diferente de produzir um artigo para jornal, um poema ou rimas, ou seja, é preciso escrever o texto de acordo com o gênero e tipo textual. Além disso, é necessário preocupar-se também com o sucesso dos objetivos da produção textual, como a interação entre o produtor do texto e o seu leitor. Escrever é diferente de falar. A linguagem escrita existe, conforme Rizzo (1986), porque existe linguagem oral, que é a necessidade social de comunicação entre os indivíduos. Portanto a escrita reflete a linguagem oral podendo ser expressa de forma clara, para si mesmo

e para os outros, pois todo texto tem um autor que fala e escreve (sujeito) e projeta ao executar o texto, a sua intenção (DEMO, 1996).

Bakhtin (2000) afirma que o texto é uma construção de signos criada por um indivíduo para expressar suas opiniões, valores e crenças, resultantes da sua interação com outros indivíduos, inseridos em uma cultura e dentro de um período histórico. Se esse ser social se expressa através do texto que produz, deseja também que alguém o compreenda. O texto não apenas revela significados construídos pelo processo dialógico entre sujeitos, mas que, para compreendê-lo, só será possível no encontro dialógico entre sujeitos.

Para Bakhtin (2000) existem diferentes tipos de textos (orais, escritos, imagens) e também diferentes formas de expressão. Todo texto é composto de materiais heterogêneos que levam em conta um conjunto de signos que por sua vez possuem significados. Comenta que o texto não é um objeto, e quem toma consciência dele, de certa forma, não fica neutro e faz descobertas a partir dele, a partir de sua visão e entendimento. Para ele, “Ver e compreender o autor de uma obra significa ver e compreender outra consciência: a consciência do outro e seu universo, isto é, outro sujeito (um tu).” (p. 338).

Cada pessoa (sujeito) estabelece relações e significados que de certa forma adquirem vida de acordo com sua própria história, suas vivências, conteúdos, valores ideológicos. Ou seja, com base nas circunstâncias que envolvem o sujeito, no contexto onde está inserido.

Diversas atividades podem influenciar a forma como o aluno produz seus textos, para progressivamente aproximá-lo da escrita mais elaborada. Essas atividades podem ser desenvolvidas desde o Ensino Fundamental, passando pelo Ensino Médio até o Ensino Superior, através de produções envolvendo várias atividades distintas. Um dos desafios educacionais é acompanhar a dinâmica dos tempos modernos, por meio da integração com os novos recursos de linguagem, a exemplo dos editores de textos para produções textuais mais elaboradas, considerando também a influência que o texto exerce nos relacionamentos sociais. Koch (1998a, p.128) sustenta:

“É preciso pensar a linguagem humana como lugar de interação, de constituição de identidades, de representação de papéis, de negociação de sentidos. É preciso encarar a linguagem humana não apenas como representação do mundo e do pensamento ou como instrumento de comunicação, mas sim, acima de tudo, como forma de inter-ação social.”

As dificuldades que os alunos enfrentam quando vão produzir um texto são inúmeras. Cabral (2009) comenta que eles não apresentam dificuldades em se expressar na oralidade através da linguagem coloquial, os problemas aparecem quando surge necessidade de produção textual, pois na linguagem oral são utilizados gestos, sinais e expressões, esses recursos não são explorados na modalidade escrita.

Esta autora apresenta quatro elementos para escrever um texto: [1] a repetição (retomada de elementos feita por pronomes ou por palavras e expressões equivalentes ou sinônimas), [2] a progressão (colocação de informações adicionais, complementando a repetição e buscando equilibrar o que já foi dito com o que se vai dizer, garantindo a continuidade do tema), [3] a não-contradição (não devem surgir elementos que contradigam aquilo que já foi citado) e [4] a relação (relacionamentos entre fatos e conceitos apresentados).

Todas essas partes compõem o texto, elas surgem uma após a outra, relacionando-se com o que já foi dito ou com o que se vai dizer e podem ajudar a avaliar o grau de coesão da produção textual.

Grande parte das pessoas passa o maior tempo planejando a escrita. Afirma-se, inclusive, que dois terços do tempo total de escrita é dedicado ao planejamento, e um terço à escrita e revisão (ELLIS, 1995). No seu livro, Ellis cita que Hayes e Flower (1980; 1986) fazem uma divisão no processo de escrita em três estágios: pré-escrita, escrita e reescrita. E é justamente a pré-escrita que envolve toda a leitura, avaliação e pensamento que devem ser realizados antes da escrita poder ser iniciada.

Jolibert (1994) também destaca três operações que estão envolvidas no ato de compor um texto.

1- O planejamento textual, que significa não o fato de “fazer um plano”, mas de levar em conta, na elaboração do texto, o destinatário e o objetivo;

2- A textualização concerne aos processos postos em ação para linearizar um texto (progressão e conservação de informações durante o texto, conexão, segmentação – problemas de anáforas – substitutos, tempos verbais, conectivos, pontuação em suas funções textuais);

3- A revisão dos textos (ou releitura) durante a produção ou depois do texto terminado (JOLIBERT, 1994, p.25).

Isso requer um trabalho com escrita e reescrita de textos, permitindo que o aluno analise sua escrita, refletindo e comparando seu texto com os que existem fora da escola. Nesta perspectiva, a produção escrita estará em constante transformação e é vista como um processo de contínua construção de conhecimentos. Para Jolibert (1999, p. 47), as reescritas correspondem a um aprofundamento do trabalho na elaboração dos textos, as quais podem ser parciais, referindo-se a um nível de análise ou a uma parte do texto.

A proposta de uso de uma ferramenta de apoio à produção textual se encaixa neste sentido. Como fase de pré-escrita, propõe-se uma série de passos que permitem ao estudante se apropriar de um tema e estruturar suas idéias para só então partir para a escrita do texto. Na fase de pré-escrita, é o momento de gerar idéias, planejar e preparar-se para a tarefa. Depois de elaborado o texto, a ferramenta pode ainda trazer subsídios que permitam ao estudante refletir sobre sua construção e fazer novos ajustes (re-escrita), ou seja, uma revisão.

Desta forma, os próprios estudantes poderão desenvolver formas de construção e produção de escrita de textos a partir dos resultados obtidos com o minerador de textos. Poderão avaliar a sua produção textual – coesão e coerência e desenvolvimento a partir dos grafos apresentados pelo programa.

Antes de tudo, aprende-se a fazer textos produzindo diversos textos, em situações de comunicação real (JOLIBERT, 2006). Segundo a autora é necessário que o aluno encontre um sentido para escrever, e isso somente será conseguido mediante o reconhecimento e necessidade para se comunicar, registrar, solicitar informações do seu dia-a-dia. Não se trata de produzir frases soltas, mas sim, de trabalhar em um texto dotado de sentido, contextualizado e proposto mediante situações significativas. Para Jolibert (2006), escrever é produzir textos em função das próprias necessidades.

Um detalhe importante é possuir uma estratégia de produção de textos que

possa se apoiar (1) a capacidade de representação, (2) o desenvolvimento de competências que permitam ao aluno escolher diferentes tipos de textos e identificar principais características lingüísticas, (3) a aptidão para gestionar a produção escrita e ou reescrita de maneira a considerar os diferentes níveis de estrutura de um texto e ainda (4) o desenvolvimento de competências sintáticas, lexicais e ortográficas.

O professor, por sua vez, poderá propor situações reais e assim localizar e identificar o tipo de dificuldade e propor, a partir de exercícios, correções e novas situações para alcançar os objetivos propostos. Demo (1996, p. 24) argumenta:

“[...] compreender o sentido de um texto implica estabelecer relações entre texto e significado, colocar em movimento modos de entender e compreender, indagar possibilidades alternativas de compreensão, perceber e dar sentidos. Esta dinâmica avança ainda mais, quando se trata de saber fazer e refazer um texto, passando-se de leitor a autor. Aparecendo a elaboração própria, torna-se visível o saber pensar e o aprender a aprender [...]”

Estas situações de aprendizagem são atividades diversas, planejadas e organizadas a partir da observação do processo de aprendizagem dos alunos, em função de algum material/texto específico, a partir de propostas dos alunos, de alguma situação ocorrida na interação estabelecida na aula, ou até mesmo (re)planejadas a partir de nossa reflexão sobre a prática.

Johnson (2001) relata que, nos anos que ainda escrevia com caneta e papel, ou usando uma máquina de escrever, elaborava cada frase antes de começar a escrevê-la. Tudo era planejado antes, pois alterar a seqüência das palavras e/ou frases era possível, mas tornava o texto sem muito sentido, difícil de ler. Era preciso ser capaz de reter toda a seqüência de palavras na cabeça, o que significava que a mente tendia naturalmente para uma sintaxe mais simples, mais direta.

Neste sentido, o meio digital trouxe facilidades e novas possibilidades na elaboração dos textos. Podemos observar uma mudança na forma da organização e da seqüência do pensamento que se reflete nitidamente nas frases elaboradas detalhadamente e cheias de significados.

O uso de processadores de textos mudam a nossa maneira de escrever, não só porque estamos utilizando novas ferramentas, mas porque o computador transforma fundamentalmente o modo como concebemos nossas frases. O processo

de pensamento se desenrola paralelamente ao processo de escrever. Percebe-se uma modificação: processos de pensamento e digitação começam a coincidir (JOHNSON, 2001).

No contexto de produção textual, o desafio maior do professor é fazer um planejamento adequado de suas ações através de propostas de tarefas que explorem processos de produção mais estruturados e interativos.

Alguns detalhes importantes devem ser observados ao descrever tarefas, conforme apresentados por Ellis (2003): [1] descrição de forma concisa do objetivo geral da tarefa; [2] informação verbal ou não-verbal fornecida pela tarefa, por exemplo, fotos, um mapa, o texto escrito; [3] a maneira pela qual a informação é apresentada, ou seja, as condições; [4] os procedimentos metodológicos a serem seguidos para executar uma tarefa, por exemplo, grupo, individual, coletivo e o tempo previsto; [5] quanto aos resultados: [a] o produto que resulta da conclusão da tarefa - se "produto aberto", ou seja, permite várias possibilidades, ou "fechado" permitir apenas uma solução conforme um objetivo especificado e [b] o processo lingüístico e cognitivos gerados para execução da tarefa.

No contexto de aprendizagem de línguas, Ellis (2003) apresenta oito princípios que servem de base para criar oportunidades para a aprendizagem e desenvolvimento de competências através da construção colaborativa de conhecimento.

Os princípios propostos por Ellis (2003) são: [1] garantir um nível adequado de dificuldade da tarefa, [2] estabelecer metas claras para cada aula baseada em tarefas, [3] desenvolver uma orientação adequada para a realização pelos estudantes, [4] assegurar que os alunos adotem um papel ativo nas aulas, [5] incentivar os alunos a assumir riscos, [6] garantir que os alunos estão voltados para o sentido da linguagem quando desempenham uma tarefa, [7] atenção à forma, nas fases de pré-tarefa, durante a tarefa e pós-tarefa, ou seja, não perder o foco, e [8] os próprios estudantes serem responsáveis pela avaliação do seu desempenho e progresso.

Willis e Skehan (1996, 1998) já faziam algumas abordagens baseadas em tarefas para o ensino de línguas, como no modelo proposto por Ellis (2003),

composto pelas três fases de implementação de uma aula baseada em tarefas. Para estes autores, a primeira fase - a fase de “pré-tarefa”- diz respeito às atividades realizadas antes da realização da própria tarefa, como uma fase introdutória ou preparação. A segunda fase - a fase de “durante a tarefa” - diz respeito ao desempenho real da tarefa, e pode incluir outros procedimentos. E a terceira fase - a fase de “pós-tarefa” - diz respeito às atividades realizadas após a conclusão da tarefa.

Outra abordagem para definição de tarefas é a proposta de Nunan (1989) sobre o que ele chamou de "habilidades comunicativas" - tarefas que envolvem o uso da linguagem comunicativa em que a atenção do usuário está centrada no significado e não a estrutura lingüística. Para ele, uma tarefa de comunicação é uma parte do trabalho em sala de aula que envolve os alunos em compreender, manipular, produzir ou interagir na língua-alvo, enquanto sua atenção está principalmente focada no significado do que na forma (NUNAN, 1989).

Há um consenso entre os pesquisadores sobre o papel do professor e dos alunos numa abordagem baseada em tarefas. O professor precisa ter mais habilidade do que os educadores com abordagens tradicionais (SKEHAN, 1996), pois o papel dos alunos está intimamente relacionado com o papel do professor e em alguns momentos há troca de papéis entre eles. O professor sabe que não pode deixar de desenvolver a criatividade e a expressão do aluno nas condições atuais da escola, mas precisa utilizar instrumentos adequados, ou seja, práticas que levam à organização do pensamento e ao desenvolvimento da expressão. Para estas práticas, o professor precisa definir de maneira muito clara o que o aluno precisa aprender, a ordem das tarefas a serem realizadas para que os alunos possam atingir os objetivos definidos previamente.

Em relação à utilização da tecnologia como apoio à realização destas tarefas, o simples uso de ferramentas computacionais não é suficiente se não estiver associado a uma metodologia específica, com objetivos precisos.. Para tanto, esta pesquisa utilizou o referencial teórico de tarefa apresentada por Ellis (2003) e adotou os oito princípios que servem de base para criar oportunidades para a

aprendizagem e desenvolvimento de competências, citado anteriormente. Estes são detalhados no capítulo 4.

3 CONSTRUÇÃO TEXTUAL APOIADA NAS TECNOLOGIAS DIGITAIS

Com a entrada dos computadores no mercado, sua popularização e facilidade de uso, aconteceu uma verdadeira revolução na escrita, sobretudo em função das ferramentas de edição de texto, adicionando à escrita maior flexibilidade e eficiência no processo de construção textual. Isso nos remete à possibilidade de, a partir de um texto elaborado, criar muitos outros. Não partindo exatamente de uma idéia apenas. De acordo com Costa Val (2004) pode-se definir texto “*como qualquer produção linguística, falada ou escrita. De qualquer tamanho, que possa fazer sentido numa situação de comunicação humana, isto é, numa situação de interlocução.*”

Alguns fatores, conforme Costa Val (1993), fazem com que um texto seja um texto: a coerência, a coesão e a informatividade. Costa Val (1993) considera a coerência o fator fundamental da textualidade, porque é responsável pelo sentido do texto. Envolve não só fatores lógicos e semânticos, mas também cognitivos na medida em que depende do partilhar de conhecimentos entre os interlocutores. Para esta autora, cada uma das coisas ditas estabelece relações de sentido e significado tanto com os elementos que a antecedem como com os que a sucedem, construindo uma cadeia textual significativa. A coerência, além de dar sentido ao texto, também permite a sua compreensão. Já Koch e Travaglia (1999) corroboram com esta definição, pois conceituam a coerência como algo que se estabelece na interação, numa situação comunicativa entre dois usuários. Seria a possibilidade de estabelecer, no texto, alguma forma de relação, uma unidade de sentido no texto como um todo, não importando a dimensão.

O segundo fator que faz com que um texto seja um texto, conforme Costa Val (1993) é a coesão. Para esta autora a coesão é a manifestação linguística da coerência, que advém da maneira como os conceitos e relações são expressos no texto. É a conexão, ligação harmoniosa entre os elementos de um texto, dando a ele

unidade. Percebe-se tal definição quando lemos um texto e verificamos que as palavras, as frases e os parágrafos estão entrelaçados, um dando continuidade ao outro. Koch (1998a) define coesão de maneira semelhante, enfatizando a ligação, a relação, os nexos que constituem a superfície textual que se manifesta na organização seqüencial do texto. Já para Marcuschi (2008), a coesão se refere aos mecanismos formais de uma língua que permitem estabelecer, entre os elementos lingüísticos do texto, as relações de sentido.

O terceiro fator que faz com que um texto seja um texto, para Costa Val (1993) é a informatividade. A informatividade diz respeito às informações, notícias, tema e conteúdo que ocorrem no texto. Ou seja, o texto tem que apresentar todas as informações necessárias para que seja compreendido. Neste sentido, a compreensão depende também de conhecimentos prévios do leitor. Para Koch (1998a), a informatividade diz respeito à seleção e apresentação das informações que o texto veicula. Quanto menos conhecidas estas informações, mais interessante e envolvente o tornam.

O uso das tecnologias e o domínio dos recursos oferecidos (ferramentas de corte, de realce, de salvamento de arquivos, possibilidade de apagar e reescrever o texto, podendo fazer isso quantas vezes for necessário) facilitam o processo de escrita e, conseqüentemente, de melhor elaboração dos textos. A tecnologia pode interagir de forma sistemática e significativa nas práticas de escrita e leitura no dia-a-dia da sala de aula, fomentando um trabalho de qualidade. A maneira de ler, interpretar e construir textos está sendo alterada e ocupa espaço importante na aprendizagem, porque tem papel decisivo e fundamental na capacidade de reflexão, expressão e criação de novas informações e conhecimentos.

Além disso, a partir do uso dos processadores de textos gera-se um grande volume de material de estudos e pesquisas. Segundo Ebecken, Lopes & Costa (2003), em virtude desse crescimento contínuo do volume, técnicas de extração de conhecimento automáticas tornam-se cada vez mais necessárias. Uma destas técnicas é a mineração de textos, que é o processo de extrair conhecimento a partir de fontes textuais. Por exemplo, tal ferramenta pode ser empregada como apoio à compreensão textual, auxiliando os estudantes na identificação de conceitos

relevantes em um texto (HESSLER e REATEGUI, 2010). Para Shanahan (2008) a leitura e a escrita requerem habilidades cognitivas e, a implementação de ferramentas que possibilitem o desenvolvimento de processos ou estruturas cognitivas envolvidos nas atividades de leitura podem representar uma contribuição importante no processos de letramento.

As atividades deste modo podem se tornar mais ricas. Um fator que pode contribuir para envolver o aluno nas atividades de produção textual é dar-lhe a oportunidade de refazer, re-escrever, inserir, deletar, copiar, colar quantas vezes achar necessário. Desta forma, a primeira escrita é revisada e melhorada, bem como suas sucessivas “reescritas”, até chegar-se à “obra prima” ou versão final (JOLIBERT, 2006).

Para Johnson (2001) o uso de um processador de textos muda a maneira de escrever, de conceber as frases, sendo que o processo de pensamento passa a acontecer no desenrolar do processo de escrever. O autor lembra que, no início, registrava tudo a mão e somente mais tarde efetuava uma composição completa de suas idéias, fazia a transcrição para o computador. Com o passar do tempo, esta rotina também mudou, ocorreu uma "virada", de modo que a caneta e o papel já lhe pareciam uma tecnologia inferior. Então, processos de pensamento e digitação começaram a coincidir. Conforme o autor, sempre é possível acrescentar mais uma parte ao texto, tornando mais fácil o ato de escrever e mudando a própria substância do que estava escrevendo, trazendo também um impacto ao seu pensamento. O mais interessante para ele é a magia do processador de texto e a mudança na maneira de criar.

3.1 Mapas Conceituais

Giombini (2004) enfatiza outra idéia importante na produção textual, que é a simples necessidade do uso das palavras como ferramenta de comunicação do pensamento. A autora cita que as funções que podem ser identificadas na formação

de textos escritos são o resultado intencional da combinação de som, sinal e significado. O mesmo ocorre na elaboração de um diagrama relacionando estes elementos, na medida em que os símbolos têm um papel de expressão, de representação de algo (significado), criando uma comunicação eficaz. Neste contexto, os mapas conceituais podem entrar em jogo, possibilitando ao sujeito visualizar as construções feitas mentalmente através de diagramas gráficos mostrando conceitos e suas relações. A teoria a respeito dos Mapas Conceituais foi desenvolvida na década de 70 pelo pesquisador norte-americano Joseph Novak. Ele define mapa conceitual como uma ferramenta para organizar e representar o conhecimento. Os mapas conceituais têm por objetivo representar graficamente relações significativas entre conceitos na forma de proposições. Uma proposição é constituída de dois ou mais termos conceituais unidos por palavras para formar uma unidade semântica (NOVAK e GOWIN, 1988), ou seja, unidas por frases de ligação nos arcos que unem os conceitos. A seguir um exemplo de um mapa conceitual sobre os mapas conceituais.

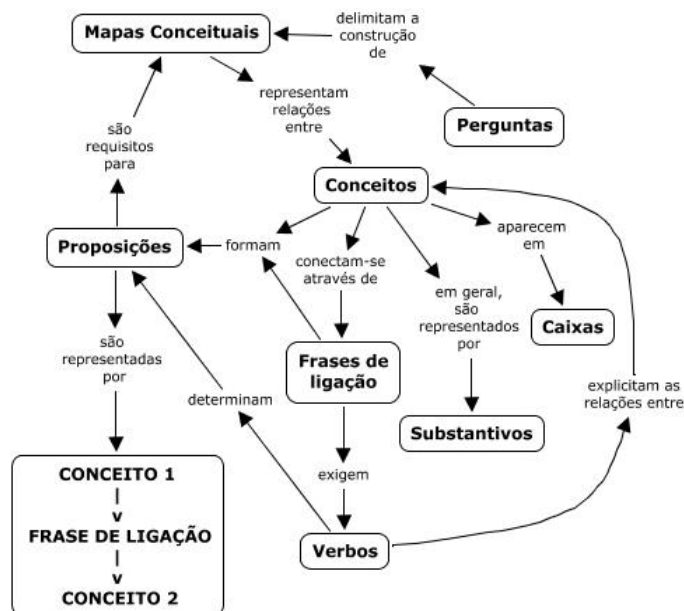


Figura 1 - Mapa Conceitual elaborado por Dutra disponível em <http://mapasconceituais.cap.ufrgs.br/textos/mapasconstrut.pdf>

Vargas¹⁰ apresenta uma das maneiras para elaborar um mapa com os conteúdos conceituais de um texto observando os passos: [1] Após a leitura atenta, listar os conceitos importantes; [2] Agrupar os conteúdos conceituais mais fortemente relacionados; [3] Arranjar, em ordem de importância ou abrangência, os conteúdos conceituais de cada um desses grupos; [4] Escrever cada um desses conteúdos numa folha, dentro de um retângulo (círculo, elipse etc.). É conveniente que os mais abrangentes fiquem em cima e os mais específicos, embaixo; [5] Interligar os retângulos com setas (ou linhas, simplesmente) e escrever próximo a elas uma ou mais palavras de ligação que estabeleçam uma proposição; e finalmente [6] Analisar o mapa para ver em que ele pode ser melhorado: remanejar blocos, estabelecer relações cruzadas, omitir partes menos importantes em prol da clareza, modificar a disposição para facilitar a visualização etc.

Os Mapas Conceituais foram originalmente baseados na teoria da aprendizagem significativa de Ausubel (1982). Para este autor, o conhecimento de um indivíduo influencia seu processo de aprendizagem na medida em que cada aprendiz faz uma filtragem dos conteúdos que têm significado ou não para si próprio e apresentam uma organização cognitiva interna baseada em conhecimentos de caráter conceitual, sendo que a sua complexidade depende das relações que esses conceitos estabelecem entre si. Esse conhecimento anterior resultará num "ponto de ancoragem" onde as novas informações irão encontrar um modo de se integrar e se associar a aquilo que o indivíduo já conhece. Essas relações por sua vez têm um caráter hierárquico, no sentido que as várias idéias se encadeiam de acordo com a relação que se estabelece entre elas. A aprendizagem significativa tem lugar quando as novas idéias vão se relacionando com as idéias já existentes. Entende-se que existe uma relação lógica e explícita entre a nova idéia e alguma(s) outra(s) já existente(s) na estrutura cognitiva do indivíduo.

¹⁰ Disponível em

http://professormarcello.com/index.php?option=com_content&task=view&id=62&Itemid=83.

Professor Marcelo Vargas é professor de física, matemática, ciências e química (nas escolas Dom Bosco, João Paulo I e Rede Estadual de Ensino).

Nesta teoria, a estrutura cognitiva é compreendida, fundamentalmente, como uma rede de conceitos organizados de modo hierárquico de acordo com o grau de abstração e de generalização. Alguns fatores, segundo propõe a teoria ausubeliana, podem contribuir para aprendizagem seja significativa. Estes fatores podem ser divididos em duas classes: [1] fatores cognitivos, que levam em conta [i] a existência de idéia(s) âncora(s) à(s) qual(is) pode se conectar uma nova idéia que se deseja ensinar; [ii] a maneira de o professor apresentar, como a forma com que o aluno estabelece internamente estas relações são importantes para o aprendizado consistente do novo conteúdo; [iii] clareza e a firmeza das idéias que servirão como âncoras determinam o nível e a estabilidade do aprendizado da nova idéia, e [2] fatores afetivo-sociais quando se estabelece uma ligação não-arbitrária e substantiva entre uma nova idéia e uma idéia de esteio ou âncora, internalizar estas relações exige do aluno vontade de fazê-lo, visto que este é um processo ativo.

O exercício de leitura, reflexão e construção de mapas pode também servir de apoio à produção textual. Conforme Giombini (2004), isso requer observar algumas regras de composição, nas quais os conceitos são expressos de forma resumida (ex. palavras, termos) ligados entre si por linhas que narram essas relações.

Apesar do interesse no emprego de mapas conceituais para trabalhar a compreensão de textos, reflexão e produção textual, é instigante a idéia de um processo de construção de textos mais interativo, também fundamentado na análise de conceitos e relacionamentos. A proposta de dissertação aqui apresentada é embasada nesta idéia, sendo detalhada no capítulo 4.

3.2 A mineração de textos para acompanhamento da escrita colaborativa

Nos últimos anos o número de ferramentas de escrita colaborativa tem proliferado e os educadores perceberam o potencial destas ferramentas na aprendizagem (Macedo, 2009). No entanto, apesar das ferramentas computacionais

armazenarem as decisões tomadas por cada aluno na criação de um documento produzido de forma colaborativa, o acompanhamento de cada aluno ainda é uma tarefa muito trabalhosa.

Macedo (2009) apresenta o Editor de Textos Coletivo - ETC, o qual emprega a ferramenta de mineração de textos SOBEK para extrair grafos a partir dos escritos dos alunos. Cabe registrar que a ferramenta ETC foi projetada e desenvolvida na Universidade Federal do Rio Grande do Sul, e pode ser utilizada pelos professores como uma ferramenta de aprendizagem colaborativa em cursos a distância.

O grafos gerados pela ferramenta SOBEK dão aos professores uma visão dos trabalhos dos alunos, mostrando conceitos e relações que parecem ser relevantes. Ao apresentar os resultados, a ferramenta SOBEK fornece pistas sobre os problemas, ou sobre a qualidade de um texto. No estudo de Macedo (2009), o objetivo principal foi o de visualizar a representação gráfica dos conceitos e suas relações em um grafo e, desta forma, apoiar a avaliação qualitativa do material escrito produzido pelos alunos. Embora, conforme a autora, os grafos não podem ser usados para reconstruir o texto original, mas podem dar uma boa noção das principais idéias e conceitos que o aluno julga relevantes.

A ferramenta de mineração de texto pode fornecer pontos positivos e/ou negativos sobre um texto permitindo a identificação de alguns problemas, tais como: a necessidade de explorar mais um determinado tema, a necessidade de produzir um texto mais fluido, e não apenas uma justaposição de termos que não estão bem conectados.

3.3 O processo de leitura e escrita apoiado pela tecnologia

Na tese de doutorado de Machado (2009), a autora faz uma análise cuidadosa de registros documentais de um grupo (curso de Pedagogia a Distância –

PEAD – UFRGS - Pólo de Três Cachoeiras – RS). A autora estudou o processo de aquisição de novas formas de ler e escrever (tanto no espaço do papel quanto da tela do computador). Para ela os conceitos de letramento, autor, leitor, editor são funções que se integram diante do uso do computador e da internet. O estudo teve como objetivo entender as novas práticas que se refletem no letramento dos sujeitos da pesquisa, promovendo mudança nas suas práticas sociais de oralidade, leitura e escrita. Para tanto integraram ao conceito de letramento a compreensão de novos gêneros discursivos e textuais por meio de um estudo dos registros documentais (chats, blogs, fóruns e outros) disponíveis em vários suportes ou espaços em que a escrita se materializa – competências e habilidades específicas, desenvolvidas a partir do início do curso. Neste estudo, letramento (ou a apropriação da escrita e seus usos) é um processo contínuo diante de avanços tecnológicos que têm acelerado as mudanças nos suportes e formas de leitura e escrita.

Segundo Machado (2009), além das mudanças no suporte (do papel para a tela de um computador/celular/palm), a internet e as novas possibilidades de comunicação e interação acabam por produzir novos gêneros textuais pela disseminação de novas modalidades de eventos e práticas sociais de escrita. Para gêneros textuais, a autora adota a definição de Marcuschi (2005), pois para ele este tema não é recente, mas ganha destaque à medida que a sociedade está se tornando mais textualizada, ou seja, a comunicação se expande no plano da escrita. Para Machado (2009) ser letrado é o estado ou condição que adquirem aqueles que sabem empregar os gêneros textuais de forma competente em eventos e práticas sociais, em situações nas quais se faz uso da oralidade, da leitura e da escrita. Mesmo uma pessoa que não sabe escrever, mas sabe ditar uma carta, apresenta certo nível de letramento suficiente para expressar suas idéias. Já a alfabetização digital, para esta autora, envolve a apropriação de novos códigos (iconográficos, por exemplo), novos gestos (uso do teclado e mouse) e novas convenções de leitura e escrita (hipertexto com links).

A seguir, algumas ferramentas que possibilitam a análise de texto são apresentadas.

3.3.1 TextAlyser¹¹

TextAlyser é uma ferramenta de análise de textos online utilizada para destacar grupos de palavras-chave. A utilização desta ferramenta permite ao usuário descobrir rapidamente o assunto principal do texto analisando as palavras e expressões utilizadas. Inicialmente a ferramenta faz uma contagem dos termos utilizados no texto apontando o número total de palavras e apresentando uma série de estatísticas sobre palavras e termos mais frequentes.

Welcome to the online text analysis tool, the detailed statistics of your text, perfect for translators (quoting), for webmasters (ranking) or for normal users, to know the subject of a text. Now with new features as the analysis of words groups, finding out the keyword density, analyse the prominence of word or expressions. Webmasters can analyse the links on their pages. More instructions are about to be written, please send us your feedback !

Textalyser

Enter your text to analyze here :

or analyze a website :

Analysis options :

Minimum characters per word :	<input type="text" value="3"/>
Special word or expression to analyze :	<input type="text"/>
Number of words to be analyzed :	<input type="text" value="10"/>
Ignore numbers :	<input checked="" type="checkbox"/>
Log the query (only for websites) :	<input checked="" type="checkbox"/>
Apply stoplist :	<input type="text" value="English"/>
Apply own stoplist (separe with blanks) :	<input type="text"/>
Make a link analysis :	<input type="checkbox"/>
Exhaustive polyword phrases :	<input type="checkbox"/>

Figura 2 - Ferramenta TextAlyser - disposto em <http://textalyser.net>

O quadro abaixo apresenta algumas destas estatísticas encontradas para um texto sobre coerência textual (texto disponível no Anexo A).

¹¹ Software disponível em <http://textalyser.net>

Total word count :	731
Number of different words :	385
Lexical Density :	52.7%
Readability (Gunning-Fog Index) : (6-easy 20-hard)	21.7
Total number of characters :	6822
Number of characters without spaces :	5631
Average Syllables per Word :	2.45
Sentence count :	34
Average sentence length (words) :	35.24
Max sentence length (words) :	94
Min sentence length (words) :	5
Readability (Alternative) beta : (100-easy 20-hard, optimal 60-70)	-35.9

Tabela 1- dados iniciais Textalyser

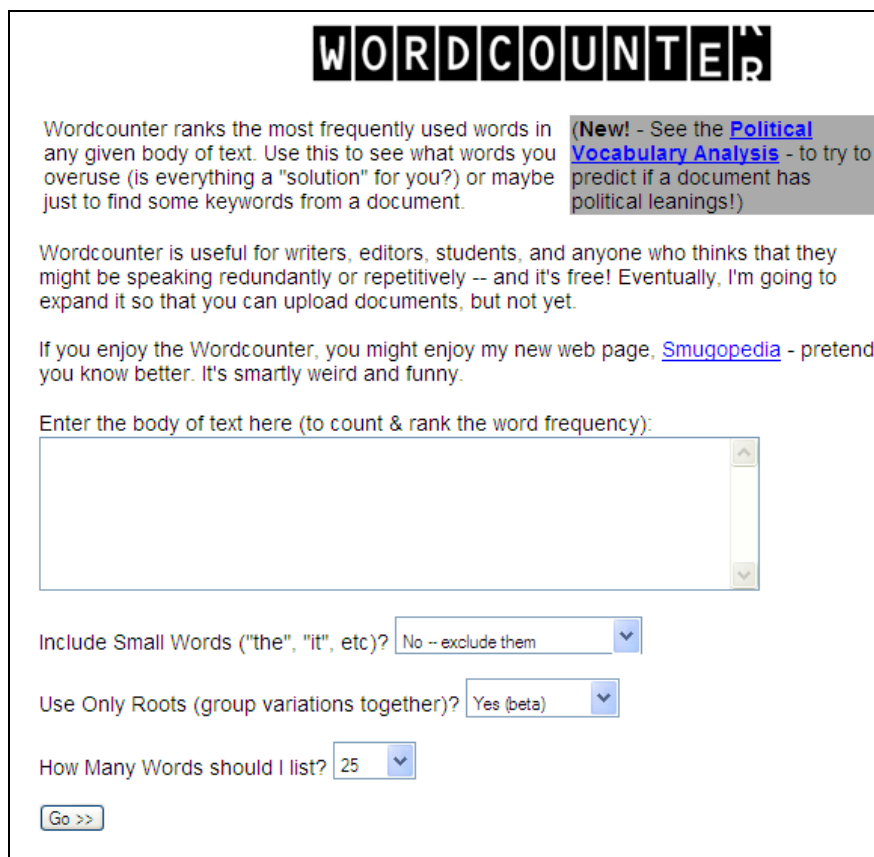
O programa aponta também a frequência com que as palavras mais utilizadas ocorrem no texto, bem como número de palavras, número de sílabas, dentre outros. Além destes fatores, a ferramenta ainda apresenta um índice relativo à “facilidade de leitura” (*readability*), critério obtido a partir do tamanho das frases e estatísticas encontradas. O programa não apresenta nenhuma ferramenta gráfica para visualização das principais informações contidas no texto.

3.3.2 Wordcounter¹²

Wordcounter é uma ferramenta que apresenta a relação das palavras mais utilizadas em um texto. Para os autores de textos é útil, pois mostra as palavras

¹² disponível em <http://www.wordcounter.com/>

repetidas e/ou redundantes numa lista. Nesta lista constam quais palavras foram utilizadas e a frequência com que cada uma aparece no texto.



The image shows the Wordcounter website interface. At the top, the title "WORDCOUNTER" is displayed in a stylized, blocky font. Below the title, there is a main text block explaining the tool's purpose: "Wordcounter ranks the most frequently used words in any given body of text. Use this to see what words you overuse (is everything a 'solution' for you?) or maybe just to find some keywords from a document." To the right of this text, there is a grey box containing a "New!" announcement: "(New! - See the [Political Vocabulary Analysis](#) - to try to predict if a document has political leanings!)" Below the main text, there is a paragraph stating: "Wordcounter is useful for writers, editors, students, and anyone who thinks that they might be speaking redundantly or repetitively -- and it's free! Eventually, I'm going to expand it so that you can upload documents, but not yet." Another paragraph follows: "If you enjoy the Wordcounter, you might enjoy my new web page, [Smugopedia](#) - pretend you know better. It's smartly weird and funny." Below these paragraphs, there is a text input field with the placeholder text "Enter the body of text here (to count & rank the word frequency):". Underneath the input field, there are three dropdown menus: "Include Small Words ('the', 'it', etc)?" with the value "No -- exclude them", "Use Only Roots (group variations together)?" with the value "Yes (beta)", and "How Many Words should I list?" with the value "25". At the bottom left of the form, there is a "Go >>" button.

Figura 3 - Ferramenta Wordcounter - disponível em <http://www.wordcounter.com/>

Novamente, utilizando o texto em anexo para esta análise (texto disponível no Anexo A), e estabelecendo 50 palavras para serem apresentadas na lista, (*How Many Words should I list? 50*) o programa apresenta como resultado:

Word	Frequency
de	60
e	33
que	30
texto	22
da	19
coerência	17
uma	15
na	3
interpretabilidade	3
usuários	3
global	3
considerada	3

Tabela 2 - resultados do programa Wordcounter

O programa Wordcounter também tem como principal objetivo encontrar estatísticas relativas ao uso de palavras e termos no texto, não dispondo de ferramentas gráficas mais complexas para visualização das informações.

3.3.3 TagCrowd¹³

TagCrowd é uma ferramenta online que permite criar nuvens de marcadores (*tagclouds*) de qualquer texto, em diferentes línguas. Para operar a ferramenta, inicialmente estabelece-se qual a frequência mínima que a palavra deve aparecer

¹³ disponível em <http://tagcrowd.com/>

no texto e também o número de palavras que o programa terá que mostrar. Após a definição destes parâmetros, pode-se visualizar os termos do texto em vários formatos e cores diferentes.



The image shows the TagCrowd website interface. At the top left is the logo "TagCrowd" with "BETA" underneath. To the right of the logo is the text "Create your own tag cloud from any text to visualize word frequency." Below the logo is a navigation bar with links: "Start Over", "What is TagCrowd?", "Blog", "Help", and "Contact". On the right side of the navigation bar are social media share icons for Facebook, Twitter, and LinkedIn. The main content area is titled "Choose a text source:" and contains three input options: 1. "Web page URL:" with a text input field and a small example "e.g. http://mysite.com/blog/2007/". 2. "Upload a file:" with a file selection button labeled "Procurar..." and the text "plain text, 6 MB max". 3. "Paste text to be visualized:" with a large text area and the text "plain text, 3 MB max". At the bottom right of the form is a large blue button labeled "Visualize!".

Figura 4 - Ferramenta Tagcrowd disposto em <http://tagcrowd.com/>

A seguir, vemos a nuvem de marcadores criada para o mesmo texto já utilizado anteriormente sobre coerência textual (texto disponível no Anexo A).



Figura 5 – TagCrowd

Diferentemente das ferramentas descritas anteriormente, esta apresenta os principais termos do texto de uma forma mais visual, enfatizando palavras e termos mais frequentes. A ferramenta não busca encontrar relações entre os termos, sendo estes apresentados em ordem alfabética. A próxima seção apresenta a ferramenta SOBEK, a qual busca apresentar visualmente tanto os termos mais relevantes do texto como também seus relacionamentos.

3.4 FERRAMENTA DE MINERAÇÃO DE TEXTOS – SOBEK

Mattison (1999) define mineração de textos como uma aplicação de sistemas de computação que envolve *hardware* e *software* dedicados à análise textual de documentos, com o objetivo de: [1] extrair informação e/ou conhecimento a partir daqueles documentos; [2] descobrir tendências que aqueles documentos possam

expor; [3] obter uma melhor visão sobre as pessoas, lugares e coisas fundamentado no que aqueles documentos podem revelar; [4] classificar, organizar e/ou categorizar os documentos ou a informação/conhecimento que eles contêm; e [5] sumarizar um documento em uma forma mais compacta através da condensação. A ferramenta SOBEK atende a estes objetivos e apresenta graficamente o resultado da análise realizada.

A mineração de textos também pode ser definida como um método de extração de informações relevantes em bases de dados não estruturadas, ou semi-estruturadas (FELDMAN e SANGER, 2006). Trata-se de um campo multidisciplinar que inclui conhecimentos de áreas como Informática, Estatística, Linguística e Ciência Cognitiva. A mineração de textos busca extrair regularidades, padrões ou tendências de textos em linguagem natural, normalmente, para objetivos específicos. Dentre outros, a área tem despertado muito interesse em decorrência da popularidade da Internet, da geração e fácil acesso a vastos repositórios de textos (SHARP, 2001).

A técnica pode ser vista como uma extensão da área de *Data Mining*, cujo foco é na análise de dados estruturados. Também chamada de *Mineração de dados textuais* ou *Descoberta de Conhecimento em Textos*, a mineração de textos permite recuperar informações, extrair dados, resumir documentos, descobrir padrões, dentre outras análises possíveis de se realizar em documentos de texto.

Uma das principais dificuldades encontradas na área é a complexidade no processamento das informações. Tal problema deve-se principalmente à dificuldade de identificação nos relacionamentos entre termos (HEARST, 1999). Por exemplo, os conceitos ‘aprendizagem’ e ‘família’ podem ou não estar relacionados, dependendo do foco de interesse do usuário.

No caso da ferramenta SOBEK, um algoritmo específico definido por Shanker (2003) foi implementado, baseado em análises estatísticas dos textos e representação das informações extraídas em um modelo de grafo chamado de *distância n-simples (n-simple distance)*.

O modelo utiliza um parâmetro n que faz com que sejam considerados n termos subsequentes ao termo analisado. Dessa forma o grafo resultante mostra um termo conectado aos n termos frequentes que aparecem a sua frente. Por exemplo, para $n=2$, o texto “AAA BBB CCC DDD” produziria o grafo da figura 6. Note que não há ligação entre AAA e DDD, sendo a distância do termo “AAA” para “DDD” de 3 (para $n=2$, o valor 3 ultrapassa a máxima distância considerada).

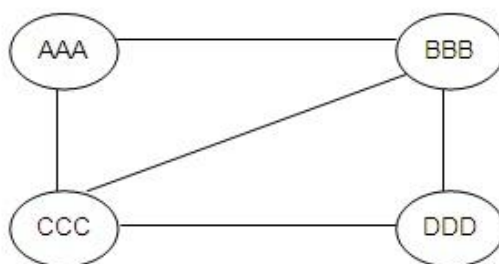


Figura 6 – Grafo

A ferramenta SOBEK foi inicialmente idealizada como recurso de apoio aos professores no acompanhamento de trabalhos de escrita colaborativa (MACEDO et al., 2009). A escrita colaborativa pode ser definida como um processo no qual autores com diferentes habilidades e responsabilidades interagem durante a elaboração de um documento. É um processo que exige gerar idéias, confrontá-las com os outros e entrar muitas vezes em negociações para chegar num consenso (MEDINA e FREITAS, 2004).

A ferramenta opera da seguinte maneira (exemplo construído a partir do mesmo texto utilizado nas análises apresentadas anteriormente, texto disponível no Anexo A).

- (1) Copia-se o texto que se quer trabalhar para a área de entrada de dados da ferramenta. Alternativamente, carrega-se um texto em formato *txt*, *doc* ou *pdf*. A figura 7 mostra a ferramenta com um texto sobre o tema “A coerência textual” (anexo A)

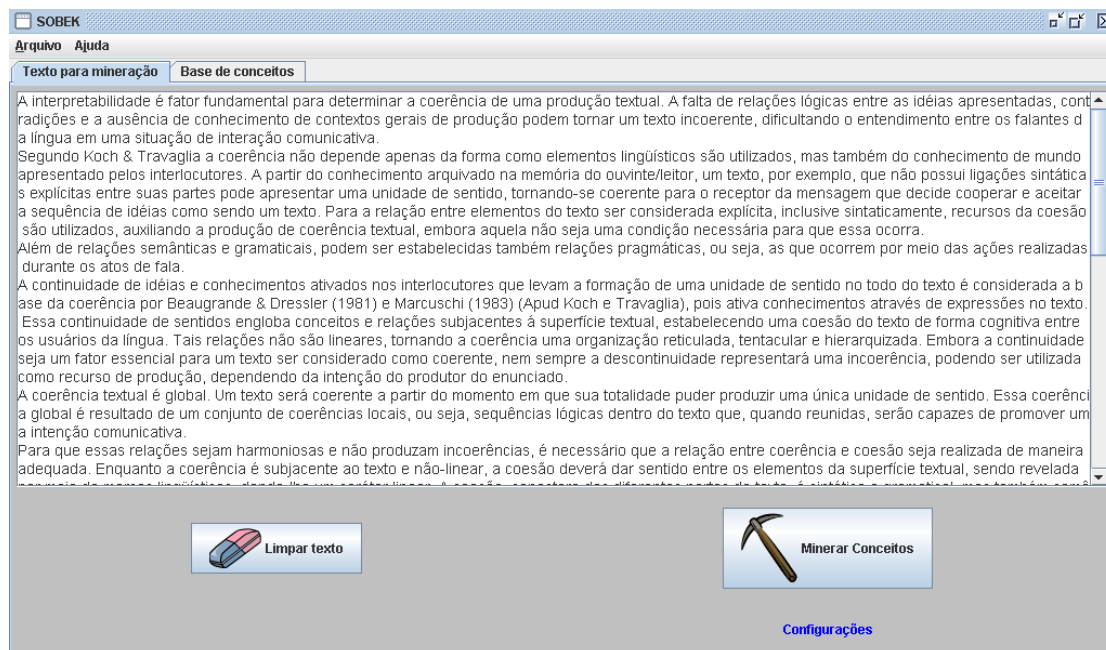


Figura 7 – Tela Ferramenta SOBEK –texto “A coerência textual”

- (2) A partir deste texto, cria-se uma base de conceitos automaticamente através do processo de mineração de textos (botão “Minerar Conceitos”, figura 7). Estes conceitos correspondem a um conjunto de palavras e termos que ocorrem com mais frequência no texto (Anexo B – Figura 16), excluindo-se artigos, preposições, e outras palavras consideradas sem significado relevante (*stop-words*). A procura por palavras registra ocorrências de palavras repetidas através da análise do texto palavra por palavra, o que corresponde à análise léxica do processo de mineração (LORENZATTI, 2007).

- (3) Em seguida, com base nos conceitos extraídos e na lista de *stop-words* utilizada para desconsiderar determinados termos, gera-se o grafo (figura 8).

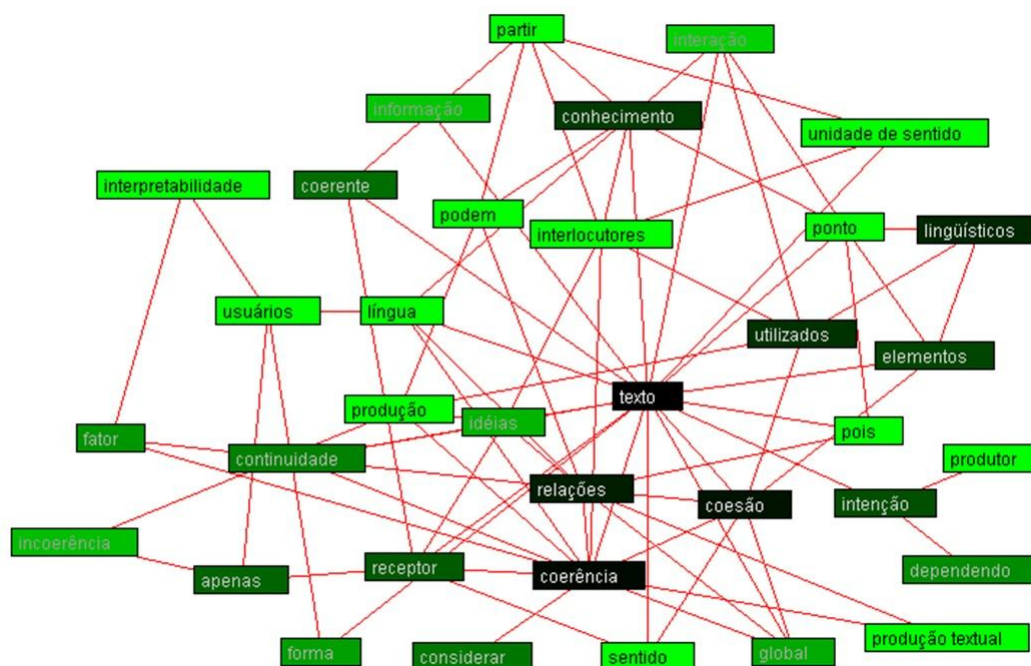


Figura 8 - Grafo gerado para o texto sobre o tema " A coerência textual " com número de mínimo de ocorrências = 3

3.4.1 Estudos anteriores utilizando a ferramenta SOBEK

Alguns estudos, que integram a linha de pesquisa “Ambientes Informatizados e Ensino a Distância” do curso de pós-graduação em Informática na Educação da UFRGS (Universidade Federal do Rio Grande do Sul), estão sendo desenvolvidos no NUTED (Núcleo Tecnologias na Educação).

Um dos estudos apresenta a aplicação de ferramentas para mineração de textos para análise do gênero textual *blog* através das ferramentas SOBEK (LORENZATTI, 2007) e *TagClouds*¹⁴ (BARBOSA et al., 2009). O objetivo foi

¹⁴ Software disponível em <http://www.tagclouds.com/>

constatar a eficiência do uso destas ferramentas e de que forma auxiliam no entendimento do conteúdo dos textos educacionais. O enfoque deste estudo está na mineração de conteúdo, uma vez que a extração de dados deu-se através de dois textos selecionados que abordam as novas tecnologias da educação. Dos resultados analisados, para os autores, a ferramenta SOBEK foi considerada a que tem uma maior capacidade de extrair informações relevantes, uma vez que relaciona conceitos e apresenta resultados mais completos, além de ser de fácil manejo. Outro dado apresentado foi que SOBEK possui características distintas da ferramenta de *Tag Clouds*, sendo que a primeira apresenta uma extração de conceitos encontrados em um texto, acrescidos de relações com outros conceitos encontrados no texto. As relações entre conceitos são obtidas por uma análise da distância entre os termos de um texto. Quanto mais próximos dois termos estiverem em um texto, mais relacionados estarão. Já a segunda ferramenta (*Tag Clouds*), simplesmente realiza uma extração de frequência de termos em um texto.

O segundo estudo, desenvolvido por Azevedo (2009), apresenta uma proposta para a realização de análise das contribuições textuais registradas por alunos em um fórum de discussão. Segundo o autor, esta análise proporciona ao professor a visualização das contribuições relevantes registradas no fórum, observação de como os alunos estão interagindo na discussão, visualização dos que colocaram poucos textos significativos, e estimular aqueles que colocaram muitas contribuições relevantes para interagir com os que colocaram poucas. Para o professor este acompanhamento demanda um tempo bastante considerável.

A abordagem desta proposta envolve mineração de textos utilizando grafos. A ferramenta SOBEK foi utilizada em experimentos anteriores (LORENZATTI, 2007), tendo sido considerada apropriada na obtenção de grafos representativos dos conceitos encontrados em textos. Azevedo (2009), contudo, observa que para permitir ao SOBEK oferecer todos os recursos necessários ao professor, seria necessário aprimorá-lo e criar algumas novas funcionalidades.

Outro estudo dessa linha de pesquisa concentra esforços no desenvolvimento de ferramentas que auxiliem no acompanhamento e análise do grande volume de dados gerados pelas ações dos alunos dentro do ambiente virtual de aprendizagem

(MACEDO et al., 2009). Nesta pesquisa, propõe-se oferecer apoio ao professor a partir do desenvolvimento de um recurso que possibilita o acompanhamento da participação, da contribuição e interação do aluno na produção coletiva em um Editor de Textos Coletivo (ETC). O ETC é definido como uma ferramenta de apoio ao trabalho/aprendizagem em grupo e tem por objetivo oferecer suporte para que todos os usuários, mesmo com diferentes necessidades, possam constituir uma coletividade apoiada na *Web* (MACEDO, 2009). O sistema permite aos usuários salvarem seus textos em diferentes extensões como: .pdf e .txt. e foi implementado a fim de facilitar a visualização dos documentos criados em qualquer plataforma, isto é, em qualquer computador e sistema operacional.

Com relação à ferramenta SOBEK, Macedo (2009) comenta que o recurso é capaz de oferecer ao professor informações acerca da participação e contribuição do aluno no processo de escrita coletiva auxiliando no acompanhamento e análise do grande volume de dados gerados pelas ações dos alunos dentro do ambiente virtual de aprendizagem. Um ambiente de educação a distância comporta um grande número de alunos que interagem e produzem de forma dinâmica e intensa, gerando assim, um alto volume de contribuições oriundas das interações. Os dados provindos das interações refletem na sobrecarga de trabalho do professor que é responsável pelo acompanhamento, comentários e mediações do processo de construção de cada estudante.

Nesta pesquisa, a ferramenta vem atender o problema da dificuldade dos estudantes com relação à produção textual. Neste sentido, a técnica de mineração de texto permite que sejam extraídos e relacionados os principais conceitos tratados em um texto. Com base na extração e relacionamento dos termos, o minerador apresenta um grafo que possibilita a visualização do conteúdo extraído de forma gráfica. Tal facilidade pode ser explorada como forma de apoio à organização de idéias, à compreensão dos principais conceitos de um texto, ao relacionamento destes com conhecimentos já possuídos, e finalmente à construção textual. A proposta apresentada neste documento é baseada nestes preceitos.

4 METODOLOGIA PARA APOIO À PRODUÇÃO TEXTUAL EMPREGANDO A FERRAMENTA SOBEK

O objetivo deste trabalho de pesquisa é investigar como uma ferramenta de mineração de texto pode ser empregada como apoio à produção textual. Esta seção apresenta a metodologia idealizada para implementação da pesquisa.

Para este projeto optou-se pelo uso da ferramenta de mineração de textos SOBEK como instrumento de apoio à produção textual. A escolha por esta ferramenta deu-se por vários motivos:

- A ferramenta foi desenvolvida por um grupo de pesquisa envolvendo professores do PPGEDU e PPGIE (UFRGS), possibilitando que participássemos do seu desenvolvimento e solicitássemos mudanças e ajustes considerados importantes para o projeto;
- A ferramenta pode ser executada em computadores com diferentes sistemas operacionais Linux, Windows ou Mac OS, permitindo que seja utilizada sem maiores restrições;
- A ferramenta é capaz de minerar textos em diversos formatos (txt, doc, pdf) de maneira bastante rápida, possibilitando seu uso efetivo em situações de ensino reais.

A premissa por trás do processo de mineração de textos da ferramenta SOBEK é baseada no fato de que termos importantes aparecem repetidas vezes em um texto. Contudo, tal afirmativa pode não ser sempre verdadeira. Ou seja, conceitos importantes podem não ser frequentes, e/ou conceitos frequentes podem não ser importantes. Tal fragilidade do sistema permite aos estudantes trabalharem “em conjunto” com a ferramenta na identificação dos termos realmente relevantes, e de suas relações.

Deste modo, elaborou-se um processo para construção textual caracterizado pela interação que se estabelece entre o aluno e o sistema, inicialmente na construção e refinamento da base de conceitos, e em um segundo momento na

criação e ajuste do grafo. Estes momentos de reflexão, de análise do texto focalizando na compreensão dos principais conceitos abordados e de seus relacionamentos, permitem ao aluno penetrar no texto de maneira mais profunda, construindo pouco a pouco a rede de relações necessárias para poder estruturar sua própria narrativa.

Um trabalho com escrita e reescrita de textos, permite que o aluno analise sua escrita, refletindo e comparando seu texto com outros disponíveis sobre um determinado tema. Nesta perspectiva, Jolibert (1994a), comenta que as reescritas correspondem a um aprofundamento do trabalho na elaboração dos textos, as quais podem ser parciais, referindo-se a um nível de análise ou a uma parte do texto.

Como apresentado anteriormente, Hayes e Flower (1986) dividem o processo de escrita em três estágios: pré-escrita, escrita e reescrita. Neste sentido, propomos neste trabalho uma forma de emprego da ferramenta de mineração de texto com base nas duas primeiras fases da produção textual – focando principalmente na fase de pré-escrita que envolve toda a leitura, avaliação e pensamento que devem ser realizados antes da escrita ser iniciada.

Assim, estabelecemos algumas tarefas para o trabalho em relação aos alunos nestes dois estágios: pré-escrita e escrita.

A – Tarefas de pré-escrita:

Tarefa 1 - O aluno faz a **leitura** do texto. Esta é uma etapa fundamental no processo de aprendizagem e organização de idéias iniciando assim a fase de compreensão de texto.

Tarefa 2 – O aluno, após realizada a primeira leitura do texto, **insere** o mesmo na ferramenta SOBEK e a utiliza para extrair os conceitos mais relevantes.

Tarefa 3 - O aluno **revisa** os termos extraídos buscando constatar se realmente são os conceitos mais relevantes. Lembrando que os termos irrelevantes podem ser removidos e outros importantes podem ser inseridos no sistema.

Tarefa 4 - O aluno faz a verificação da adequação das relações apresentadas entre os termos. O aluno analisa os relacionamentos – removendo e/ou acrescentando relações: notar que em alguns momentos há a necessidade de voltar ao texto original para releitura, pois podem surgir dúvidas sobre os conceitos e suas relações. Este “ir e vir” faz com que o aluno pouco a pouco se aproprie mais do texto lido inicialmente, observando os detalhes importantes relativos aos conceitos e seus relacionamentos.

B – Tarefas de escrita

Tarefa 5 - O aluno, a partir deste grafo gerado, parte para sua produção textual, buscando novas interpretações para a representação gráfica produzida, agregando novos fatos e idéias relacionados à sua própria realidade. A organização da metodologia proposta em tarefas contempla os estágios de pré-escrita e escrita do processo de produção textual. O estágio de reescrita é visto aqui como uma etapa posterior igualmente importante, cujo principal objetivo é o a revisão e refinamento de um texto inicialmente elaborado. Contudo, optou-se por focar nos dois primeiros estágios por serem estes os principais momentos nos quais a ferramenta de mineração de texto poderia apoiar a produção textual. Neste sentido, a ferramenta pode atuar como um suporte à organização lógica das informações, um processo que relaciona leitura/escrita como etapas de um mesmo processo cognitivo (SHANAHAN, 2008). Leitura e escrita são apontadas pelo autor como aspectos do pensamento que estão essencialmente inter-relacionados. A escrita é um trabalho complexo que requer vários passos ou etapas, nos quais revisamos diferentes aspectos lingüísticos e estruturais. Se as idéias centrais de um tema são apreendidos antes dos conceitos e o aluno é capaz de delinear similaridades e diferenças entre estes conceitos, o aluno desenvolve compreensões mais precisas e integradas

podendo assim reformular novas proposições usando suas próprias palavras (AUSUBEL, 1982). Para este autor os aprendizes processam novas significações que devem estar organizadas progressivamente para que possam ancorar seu conhecimento. Desta forma, a produção escrita estará em constante transformação e vista como um processo de contínua construção de conhecimentos.

4.1 METODOLOGIA DA PESQUISA

Sendo o objetivo do estudo investigar como a utilização de uma ferramenta de mineração de texto pode ser empregada como apoio à produção textual, optamos por realizar nossa pesquisa fundamentada numa metodologia mista (CRESWELL, 2007), com o uso simultâneo de métodos qualitativos e quantitativos de pesquisa. Este tipo de pesquisa busca entender um fenômeno específico em profundidade. Utilizamos métodos quantitativos e qualitativos nas análises dos dados coletados.

Ao invés de estatísticas, regras e outras generalizações, a pesquisa qualitativa trabalha com descrições, comparações e interpretações. As observações podem envolver praticamente todos os sentidos – visão, audição, percepção (FLICK, 2009). Como pesquisadora, minha opção foi manter a posição de observadora e expectadora sem envolvimento com o objeto da observação. Nesse sentido a observação foi não-participante, sem intervenções.

A pesquisa qualitativa também pode ser vista como um processo de reflexão e análise da realidade através da utilização de métodos e técnicas para compreensão detalhada do objeto de estudo em seu contexto histórico e/ou segundo sua estruturação. Esse processo implica em estudos, observações bem como análise de dados e questionário de forma descritiva (OLIVEIRA, 2005). Para tanto, esta pesquisa também foi descritiva, pois se baseou na análise dos registros dos alunos sem manipulá-los, investigando alguns fenômenos, sua relação e conexão com outros (como por exemplo, as idas e vindas ao texto original, leitura, releitura, escrita, reescrita).

Durante a realização desta pesquisa sentimos a necessidade de utilizar outros aspectos que, associados à interpretação qualitativa, pudessem fortalecer o os resultados encontrados. Neste sentido, adotamos métodos quantitativos na fase de estruturação e apresentação de alguns destes resultados.

A análise realizada permitiu a investigação de todo o processo envolvido na realização da produção textual. Foi observada a dinâmica da atividade proposta, avaliando assim os objetivos anteriormente elencados.

A grande maioria dos estudos sobre compreensão de textos procura examinar somente os fatores responsáveis pela compreensão, tais como: [1] fatores lingüísticos (sintáticos, semânticos, léxicos, habilidades de decodificação); e [2] fatores cognitivos (memória de trabalho, capacidade de monitoramento, inferências e capacidade de integrar as informações veiculadas no texto com seu conhecimento prévio) (BRANDÃO e SPINILLO, 2001). No estudo aqui proposto, foram observados diferentes fatores relacionados ao uso da ferramenta de mineração de textos. Para validar este estudo, fizemos uma análise de todo o processo de produção textual, e não apenas do seu resultado.

4.2 EXPERIMENTO PRELIMINAR

Para avaliar o processo de construção textual apoiado pela ferramenta SOBEK, um experimento inicial foi realizado em 2009. Acompanhamos a atividade de construção de texto realizada por um estudante de 20 anos do segundo semestre de Engenharia Mecânica na Universidade de Jaraguá do Sul (UNERJ – SC). Em termos metodológicos, classificamos a investigação realizada como estudo de caso com observação não participante. Foi solicitado ao aluno que fizesse a leitura de um texto na área de contabilidade empresarial e que elaborasse um resumo desta leitura¹⁵.

O estudante iniciou a atividade inserindo o texto fornecido na ferramenta SOBEK. Em seguida, selecionou a opção “Extrair Conceitos” para que a ferramenta selecionasse automaticamente os termos mais frequentes no texto, e que a partir destes se iniciasse a análise dos principais conceitos do texto.

O próximo passo consistiu-se da verificação dos conceitos pertinentes. Foi possível constatar que na lista de conceitos extraídos pela ferramenta, muitas das palavras e termos não contêm significado relevante para a compreensão do tema escolhido, i.e. contabilidade empresarial. Com base nesta verificação, o aluno organizou a lista de conceitos, eliminando os termos considerados pouco pertinentes, e acrescentando termos importantes que apareciam no texto e que não foram identificados pela ferramenta. Palavras escritas no singular e plural com os mesmos conceitos relacionados foram excluídas, bem como termos de ligação entre palavras. Também foram observadas algumas repetições nos relacionamentos e incluídos conceitos não extraídos. Neste processo, o aluno voltou ao texto original algumas vezes para releitura, na medida em que surgiram dúvidas sobre os conceitos e suas relações. O aluno se mostrou muito atento na realização da atividade, e refez duas vezes este procedimento para ter certeza de estar tudo correto. É importante observar que este “ir e vir” fez com que o aluno aprofundasse

¹⁵ O texto sugerido faz parte da apostila sobre Terminologia Contábil (página 5 até 8) do Centro Universitário de Jaraguá do Sul (UNERJ), curso de Engenharia Mecânica, Prof. Paulo Onildo de Matos.

sua compreensão sobre o tema proposto e percebesse detalhes importantes em relação aos conceitos. No passo seguinte, o aluno faz a construção do grafo, selecionando a opção “Gerar Grafo” da ferramenta SOBEK.

Logo após, ao observar o grafo, o aluno imediatamente identificou os conceitos e relacionamentos apresentados como corretos, passando a descrevê-los oralmente numa seqüência bastante coerente, apresentando seus significados e justificando suas posições no grafo. Por exemplo, ao se referir ao conceito *Gastos*, o estudante percebeu a relação entre os termos *bens* e *serviços* adquiridos, bem como *Gastos com mão-de-obra*, e ainda *Gastos na compra* de um imobilizado, como consta no texto original. No passo seguinte, o aluno procedeu à construção do texto a partir do grafo construído. Como os conceitos relacionados ao tema proposto já haviam sido bastante trabalhados em todo o processo de construção do grafo, a produção de texto feita pelo aluno ocorreu de forma muito natural, espontânea e com definições muito claras.

Abaixo, seguem algumas observações do aluno com relação ao processo de construção do texto apoiada pela ferramenta SOBEK

“A possibilidade de extrair de um texto as principais informações e conceitos úteis é muito interessante. Mesmo tendo que fazer ‘uma limpeza’ manualmente nos conceitos relacionados a experiência foi muito válida. O fato de ter que realizar e repetir a leitura de algumas partes do texto para ter a certeza que estava correto ajudou a entender o conteúdo. Desta forma a interpretação e compreensão dos conceitos foi acontecendo de forma rápida. Após observar que o grafo fazia sentido, a produção do texto (resumo) foi fácil de fazer. Bastava olhar o grafo, acompanhar os relacionamentos e transformar a parte gráfica em texto.”

Percebe-se no relato do aluno a total compreensão do método de produção textual proposto, no qual a construção do texto é realizada através de um processo estruturado e interativo. Neste processo, o aluno interage com a ferramenta SOBEK no desenvolvimento de cada uma das etapas de construção textual. O aluno também fez as seguintes considerações com relação a melhorias necessárias na ferramenta:

“O programa deveria sofrer apenas algumas alterações, melhorando a organização das palavras e a definição de linhas de assuntos principais, aqueles ligados diretamente, para

um mais rápido entendimento do assunto. E ainda, a opção para que se possa fazer a alteração dos termos diretamente nos grafos.”

A primeira observação do aluno mostra sua inquietação com relação à funcionalidade da ferramenta de extrair conceitos realmente pertinentes. Efetivamente, no procedimento realizado, muitos termos irrelevantes foram identificados pela ferramenta, sobretudo pelo fato do parâmetro “ocorrência mínima de um termo” ter sido mantido em 2 (Anexo B - Figura 7) – um valor baixo para um texto de três páginas. Consta-se portanto que essa função não está clara para os alunos, e que um processo mais automático de seleção deste parâmetro deveria ser desenvolvido.

Já seu segundo comentário diz respeito não tanto à funcionalidade disponibilizada de alteração da base de conceitos, mas sim à interface do sistema que poderia ser mais adequada para realização das modificações nos grafos. Esta alteração na interface do sistema, dentre outras, está sendo estudada pela equipe de desenvolvimento da ferramenta SOBEK.

Os resultados deste experimento preliminar foram descritos em um artigo completo apresentado no Simpósio Brasileiro de Informática na Educação em 2009 (KLEMANN e REATEGUI, 2009).

4.3 EXPERIMENTO PRINCIPAL

Em 2010 foi realizado o principal experimento desta pesquisa, envolvendo uma turma de alunos do Ensino Médio do Instituto Estadual de Educação Ernesto Alves. A escola, fundada em 21 de março de 1910, com o nome de Colégio Elementar Ernesto Alves, em 1939 passou a formar o Grupo Escolar Ernesto Alves. Em 26 de maio de 1952, após grande empenho da comunidade de Rio Pardo foi criada a Escola Normal Ernesto Alves, passando a formar o Curso de Aplicação.

Os sujeitos envolvidos na pesquisa foram 20 alunos do segundo ano do Curso Normal de Nível Médio, com idade entre 15 e 18 anos. Para a realização deste estudo, foi utilizado o laboratório do IEE Ernesto Alves com 20 computadores para uso dos alunos (figura 10). O trabalho foi realizado individualmente. O tempo de duração deste experimento foram três períodos de aula – iniciando no horário das 7h30min e terminando às 9h45min, ou seja, 2h15min.

Na semana que antecedeu o experimento, a pesquisadora realizou um trabalho junto a esta turma, a pedido da professora da disciplina de Língua Portuguesa (LP). Na ocasião, foi realizada uma oficina sobre Mapas Conceituais e, ao final, a apresentação do programa SOBEK e detalhamento da sua funcionalidade. A atividade também levou 2h15min.

Durante a oficina sobre Mapas Conceituais, foram adotados os passos sugeridos por Vargas¹⁶ (detalhados no capítulo 3.2.1) para elaboração de um mapa com conceitos. O tema para a construção dos Mapas Conceituais foi “Sintaxe”¹⁷. Este tema faz parte do conteúdo programático da disciplina de LP para o 2º Ano do Ensino Médio. Na proposta da professora da disciplina de LP, os alunos deveriam

¹⁶ Disponível em http://professormarcello.com/index.php?option=com_content&task=view&id=62&Itemid=83. Professor Marcelo Vargas é professor de física, matemática, ciências e química (nas escolas Dom Bosco, João Paulo I e Rede Estadual de Ensino).

¹⁷ Sintaxe é a parte da gramática que estuda a disposição das palavras na frase e a das frases no discurso, bem como a relação lógica das frases entre si. Disponível em <http://soportugues.com.br>

apresentar definições, tipos e estrutura relacionado às frases, aos períodos e orações.

Inicialmente foi realizado um mapa de forma colaborativa no quadro da sala e, em seguida, noções básicas e uma prática utilizando o programa CMapTools (v4.09).

Esta atividade vem ao encontro da proposta da pesquisa, pois a construção de mapas também pode servir de apoio à produção textual, assim como os grafos gerados pela ferramenta SOBEK. A partir do diagrama gráfico apresentado pelo programa CMapTools, mostrando conceitos e suas relações, os alunos puderam ter uma melhor compreensão sobre o grafo que a ferramenta SOBEK gera.

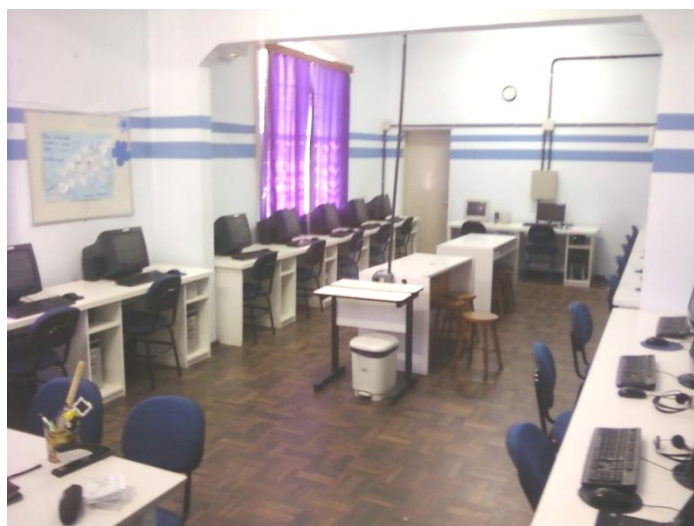


Figura 9 - foto laboratório IEE Ernesto Alves - acervo pesquisadora

Em todos os computadores do laboratório foi instalada a ferramenta SOBEK. Além desta, foram utilizados dois programas: (a) Editor de Textos - *Microsoft Office Word 2007* na etapa de pré-escrita; (b) *Camtasia Studio – TechSmith Corporation – versão 7.1* para geração de filme na etapa de escrita. Este filme foi gerado enquanto os alunos realizavam as tarefas propostas, sendo registradas todas as ações dos mesmos pelo programa nos sistemas utilizados. Após a realização das tarefas, foi aplicado um questionário (Anexo C) com perguntas

abertas. Assim, os participantes ficaram livres para responder com suas próprias palavras às questões formuladas. O objetivo deste questionário foi investigar sobre a importância e a utilização do grafo gerado durante a execução das tarefas descritas no início deste capítulo.

Conforme a professora da disciplina, muitos dos alunos da turma tinham dificuldade na produção textual. Além disso, para os alunos, produzir um texto no computador também era uma tarefa desafiadora. Em entrevista realizada, ela comentou que alguns alunos eram do interior do município e, além de frequentarem as aulas, trabalhavam na lavoura para ajudar na renda familiar. Apesar de existir realidades diferentes e situações difíceis, não existia discriminação na turma. Os alunos eram muito solidários e se ajudavam mutuamente. No seu depoimento ainda comentou, confirmando a afirmação de Luft (1999), que *“estes alunos têm um vocabulário muito restrito (pobre) e falta leitura.”*

Durante a entrevista, a professora da turma ressaltou que os alunos não apresentavam dificuldades em se expressar oralmente numa linguagem coloquial. Os problemas apareciam quando surgia a necessidade de produzir um texto. Para Cabral (2009) esta situação é bem frequente, pois na linguagem oral são utilizados gestos, sinais e expressões, recursos que não são explorados na modalidade escrita. Normalmente os alunos utilizam o código restrito para falar e, para escrever, necessitam do código mais elaborado (Marcuschi, 2005). Portanto, o problema da expressão escrita está associado à linguagem que o indivíduo usa.

Para o desenvolvimento desta pesquisa definimos um plano de trabalho baseado em tarefas (ELLIS, 2003), observando os dois primeiros estágios da produção textual. O organograma da figura 9 resume as tarefas envolvidas neste processo, destacando para cada tarefa, possíveis formas de observação da atividade realizada pelos alunos.

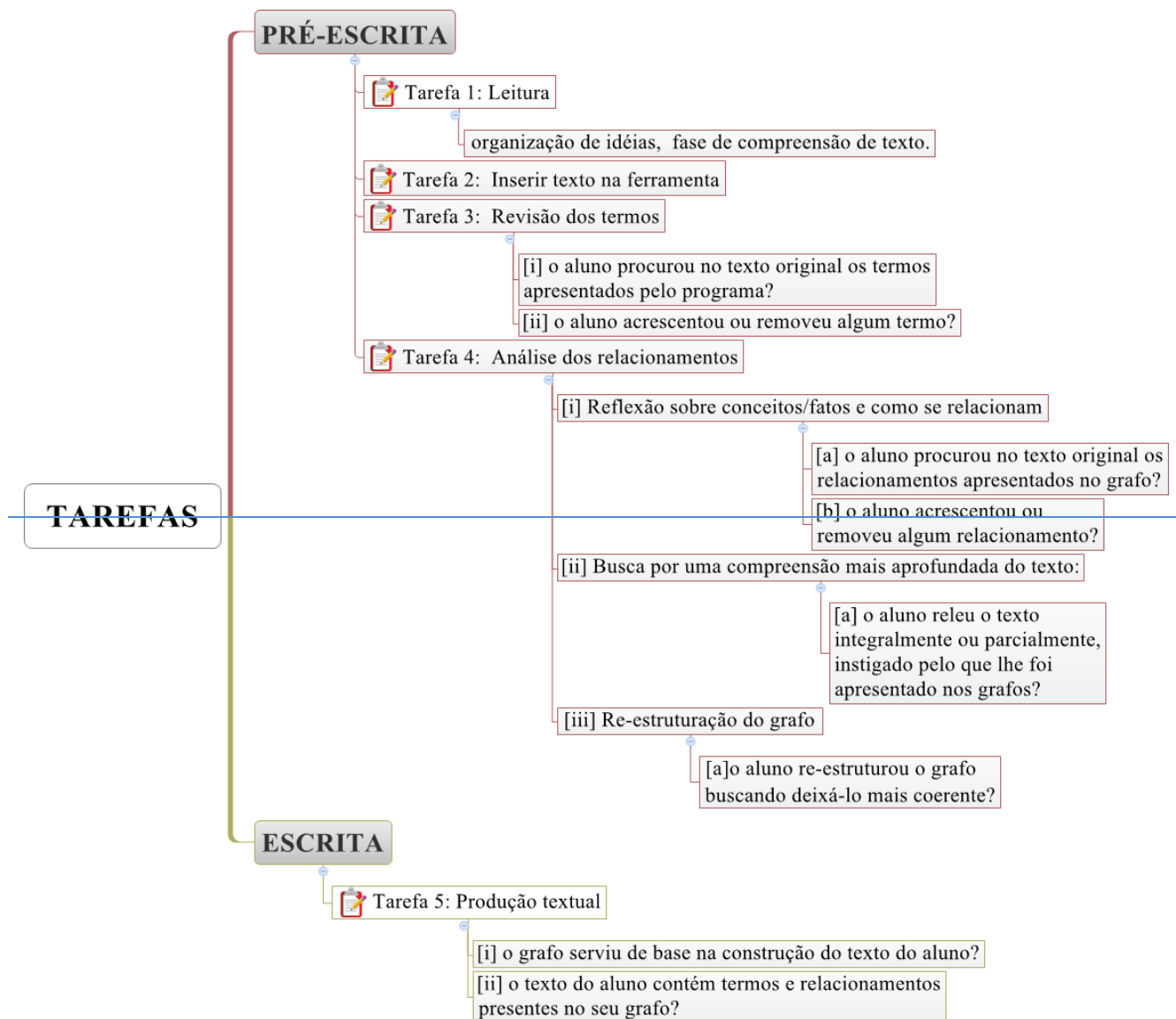


Figura 10 - Organograma das Tarefas

As seções a seguir detalham as tarefas de pré-escrita e de escritas desenvolvidas no experimento.

4.3.1 Tarefas de pré-escrita

Inicialmente os alunos realizaram a **leitura** do texto “Realismo” (anexo D), como estabelecido na **Tarefa 1**. Desta forma, os estudantes iniciaram a fase de compreensão do texto sugerido pela professora da disciplina, pois como salienta Vieira (1981) não é possível escrever sobre o que se desconhece ou sobre aquilo que se conhece pouco.

Na **Tarefa 2**, os alunos inseriram este texto na ferramenta SOBEK, empregando-a como instrumento de apoio à compreensão textual, como proposto por Hessler e Reategui (2010). O objetivo da atividade foi auxiliar os estudantes na identificação de conceitos relevantes extraídos do texto. Estes conceitos correspondem ao conjunto de palavras e termos que ocorreram com mais frequência no texto, a partir do processo de mineração de textos realizado pela ferramenta.

Para as **Tarefas 1 e 2** não foram definidas categorias de observação relacionadas à leitura, pois este foi um procedimento individual no qual não havia muitos elementos observáveis.

Na **Tarefa 3** os alunos revisaram os termos extraídos pela ferramenta buscando constatar se realmente eram os conceitos mais relevantes. Esta fase ainda estava relacionada à compreensão do texto. Duas categorias de observação foram definidas para esta tarefa, sendo descritas nos quadros 1 e 2.

Quadro 1: Detalhamento da 1ª categoria de observação

<i>Categoria 1: Busca por termos apresentados pelo programa no texto original.</i>
<p>Nesta categoria de observação buscamos analisar o momento de reflexão sobre a relevância dos termos extraídos pela ferramenta. A questão abaixo a observação desta etapa do processo de produção textual:</p> <ul style="list-style-type: none"> • <i>o aluno procurou no texto original os termos apresentados pela ferramenta Sobek?</i>

Para responder a questão, buscou-se identificar evidências na utilização da ferramenta Sobek tanto no filme registrado durante o processo de produção textual de cada aluno, bem como nos questionários respondidos pelos participantes.

Observações feitas a partir do filme:

As evidências identificadas no filme sobre esta categoria de observação estavam principalmente relacionadas às “idas e vindas” ao texto original, indicando que os alunos procuraram no texto termos representados no grafo.

Em muitos registros percebe-se que os alunos intercalavam as janelas dos programas, passando da visualização do grafo para a leitura do texto, e depois retornando ao grafo. Tal ação denota a primeira tentativa de uma leitura com mais profundidade do texto original.

Percebe-se também o uso da barra de rolagem na releitura do texto, indicando que os alunos procuraram por mais informações que lhes dessem uma melhor compreensão do material lido. Ao instigar o aluno a explorar o texto original com mais profundidade, a metodologia proposta se alinha à idéia de que a releitura melhora a metacompreensão do aluno, levando-o a um maior esclarecimento sobre o que compreendeu ou não (RAWSON e DUNSLOSKY, 2000).

Análise do questionário:

As perguntas do questionário relacionadas a esta categoria são as de número 1, 2 e 4. Algumas evidências relacionadas a esta categoria são descritas e discutidas abaixo:

“[...] a partir do grafo identifiquei o que se destacava no texto” (Aluno 3)

“[...] percebi que as palavras selecionadas pelo grafo eram importantes, de destaque” (Aluno 4)

“Para encontrar os termos e analisar as relações apresentadas tivemos que ler várias vezes o texto e então com bastante conhecimento conseguimos relacionar o grafo com o texto.”(Aluno 8)

Os relatos acima mostram que a busca pelos termos apresentados pelo programa no texto original, fez com que os alunos aprofundassem sua compreensão sobre o tema proposto e percebessem detalhes importantes em relação aos conceitos. Para Hayes & Flower (1980) é importante desenvolvermos um conjunto de competências receptivas (processo de leitura) e um conjunto de habilidades reflexivas (interpretação de texto), o que naturalmente aconteceu nesta etapa. O comentário do Aluno 8 em relação a “*ler várias vezes o texto*” sugere esta busca por compreensão do texto, ou seja, habilidade reflexiva. Marcuschi (2008) também defende que o trabalho com leitura e compreensão deve ter como meta primordial o desenvolvimento do aluno e suas habilidades. Assim, o estudante desenvolve a capacidade de usar um número maior de recursos da língua e de forma adequada.

Quadro 2: Detalhamento da 2ª categoria de observação

Categoria 2: Acréscimo ou remoção de termos no grafo

Nesta categoria de observação buscou-se identificar evidências que demonstrassem a necessidade dos alunos em fazer ajustes nos grafos, acrescentando ou removendo termos, visando aproximá-los do que compreenderam a partir da leitura do texto. A pergunta que guiou o processo de observação nesta etapa foi:

- *o aluno acrescentou ou removeu algum termo do grafo?*

Para responder a questão, o filme registrado durante o processo de produção textual de cada aluno foi analisado. Nenhuma das questões do questionário estava relacionada a esta categoria de observação.

Observações feitas a partir do filme:

As evidências identificadas no filme sobre esta categoria de observação estavam principalmente relacionadas à comparação que o aluno fez do grafo inicial gerado, com a base de conceitos apresentada pela ferramenta SOBEK. Em alguns casos, os alunos retornaram à base de conceitos, adicionando ou removendo

conceitos manualmente. Em seguida geraram novo grafo. Esta ação denota a tentativa de localizar as idéias principais de forma isolada, reorganizando e estabelecendo algumas relações. Nesta fase, notou-se que os alunos buscaram alternativas e adquiriram uma visão de conjunto sobre o que era tratado no texto, buscando uma compreensão do texto como um todo. Como define Demo (1996), compreender o sentido de um texto implica em estabelecer relações entre texto e significado.

Na **Tarefa 4**, os alunos analisaram os relacionamentos – removendo e/ou acrescentando relações. Em alguns momentos, sentiram necessidade de voltar ao texto original para releitura. Três categorias de observação foram definidas para esta fase, sendo descritas nos quadros 3 e 4.

Quadro 3: Detalhamento da 3ª categoria de observação

<i>Categoria 3: Reflexão sobre conceitos/fatos e como se relacionam</i>
<p>Nesta categoria de observação buscamos analisar os detalhes relativos aos conceitos e relacionamentos apresentados pela ferramenta que o aluno julgou coerentes. As questões que nortearam as observações feitas nesta etapa foram:</p> <ul style="list-style-type: none"> • <i>o aluno procurou no texto original os relacionamentos apresentados no grafo?</i> • <i>o aluno acrescentou ou removeu algum relacionamento?</i> <p>Na tarefa 3, as questões relacionadas à procura ou modificação dos termos relevantes no grafo foram analisadas separadamente. No caso dos relacionamentos, optou-se por fazer estas análises em conjunto, principalmente pela maior dificuldade em identificar ações do aluno focadas nos relacionamentos entre conceitos relevantes.</p> <p>Observações feitas a partir do filme</p> <p>As evidências identificadas no filme relacionadas a esta categoria de observação apontaram para o fato de que 16 alunos, dos 20 participantes, fizeram alterações</p>

nos relacionamentos dos grafos, se reportando ao texto original após efetuarem estas mudanças. Os 4 alunos restantes apenas observaram o grafo apresentado. Desta forma, pode-se concluir que a partir da consulta e modificação dos grafos na busca por relacionamentos que pareciam importantes, os alunos iniciaram a estruturação de suas ideias com relação à forma como conceitos relevantes no texto se relacionavam. Esta é uma etapa importante do planejamento e preparação para a tarefa de escrita do texto. A estruturação das ideias teve como base a compreensão do texto, o que, segundo Brandão e Spinillo (2001), não se resume à capacidade dos alunos em memorizar fatos e relacionamentos, mas também à compreensão de informações que não foram apresentadas explicitamente.

Análise do questionário

As perguntas do questionário relacionadas a esta categoria são as de número 1, 2, 3 e 4. Alguns relatos contendo evidências sobre como os alunos procederam na tentativa de localizar e analisar conceitos e suas relações são apresentados logo abaixo.

No relato do Aluno 1, sobre como procedeu para analisar as relações apresentadas entre os termos, ficou evidente que ele fez uma filtragem dos conteúdos com significado para si próprio:

“ “[...] Julgando o que era necessário ao meu ver, de acordo com meu conhecimento” (Aluno 1)

Esse conhecimento anterior, de acordo com Ausubel (1982), resultou num “ponto de ancoragem”. Assim, novas informações encontraram um modo de se integrar àquilo que o aluno já conhecia, possibilitando estabelecer relações de sentido.

Já o aluno 20 relatou que seu procedimento foi:

“ [...] Julgando o que era necessário a meu ver. E retirei as palavras que julguei desnecessárias... Analisei cada palavra separadamente e depois em relação a outra”.

Esta procura pelas relações entre os termos denota uma busca por

mecanismos de coesão, que permitem estabelecer as relações de sentido do texto, como enfatiza Marcushi (2008).

Quadro 4: Detalhamento da 4ª categoria de observação

Categoria 4: Busca por uma compreensão mais aprofundada do texto

Esta categoria de observação teve como objetivo verificar se o aluno buscou uma melhor compreensão do texto, fazendo sua releitura quando necessário. A pergunta que guiou a observação nesta etapa da pesquisa foi:

- *o aluno releu o texto integralmente ou parcialmente, instigado pelo que lhe foi apresentado nos grafos?*

Observações feitas a partir do filme

O registro obtido pelo filme não permitiu identificar nenhuma evidência relacionada a esta categoria de observação.

Análise do questionário

A pergunta do questionário relacionada a esta categoria é questão de número 4. Sobre a releitura do texto original, buscando localizar termos que apareciam no grafo, **9** alunos responderam que não leram o texto, **3** informaram que leram o texto parcialmente e **8** informaram ter lido o texto integralmente, instigados pelo que lhes foi apresentado nos grafos. Algumas evidências identificadas nas respostas dos alunos:

“Sim, claro, foi muito importante a releitura, pois coisas que às vezes não fazem sentido, na releitura parecem ser muito importantes e fazem sentido com as palavras do grafo.” (Aluno 8)

“Sim porque deu para compreender melhor o texto, ligando uma palavra com a outra.”(Aluno 9)

“Sim, voltava ao texto e achava itens importantes dos quais não tinha visto e por ter que observar os termos colocados, prestava mais atenção ao ler todo texto” (Aluno 17)

Koch e Travaglia (1999) afirmam que o professor deve trabalhar a produção e a compreensão de textos baseado em estratégias. Estas estratégias devem permitir que o aluno, além de compreender e interpretar, possa dar sentido ao texto, estabelecendo alguma forma de relação.

Na busca por uma compreensão mais aprofundada do texto, a estratégia adotada foi a releitura, integral ou parcial, incentivada pelos questionamentos decorrentes da interpretação dos grafos. Desta forma, ficaram evidentes os fatores que fazem com que um texto seja um texto (COSTA VAL, 1993): **a coerência**, “[...]pois coisas que às vezes não fazem sentido, na releitura parecem ser muito importantes” (Aluno 8); **a coesão** “[...] deu para compreender melhor o texto, ligando uma palavra com a outra.”(Aluno 9) e a **informatividade** “[...]voltava ao texto e achava itens importantes dos quais não tinha visto [...]” (Aluno 17).

Quadro 5: Detalhamento da 5ª categoria de observação

Categoria 5: Re-estruturação do grafo

Nesta categoria de observação buscamos analisar de forma mais detalhada os ajustes realizados nos grafos pelos alunos. A pergunta que guiou a observação nesta etapa da pesquisa foi:

- *o aluno re-estruturou o grafo buscando deixá-lo mais coerente?*

Esta ação denota a tentativa do aluno em interpretar o que há no grafo, e ao mesmo tempo ajustá-lo de acordo com o que compreendeu do texto

Observações feitas a partir do filme

A observação do filme permitiu constatar que muitos alunos buscaram re-estruturar seus grafos. Ou seja, interpretaram o que havia no grafo e a partir disso fizeram ajustes de acordo com o que compreenderam do texto. Do total de alunos que participaram do experimento, **16** alunos fizeram ajustes, **4** apenas organizaram o grafo separando os termos para melhor visualização. Os termos selecionados não variaram muito, porém os relacionamentos que cada aluno julgou importante foram

alterados significativamente. Isso influenciou não no tamanho dos grafos gerados, mas nas ligações entre termos apresentadas graficamente, como podemos observar nos grafos a seguir:

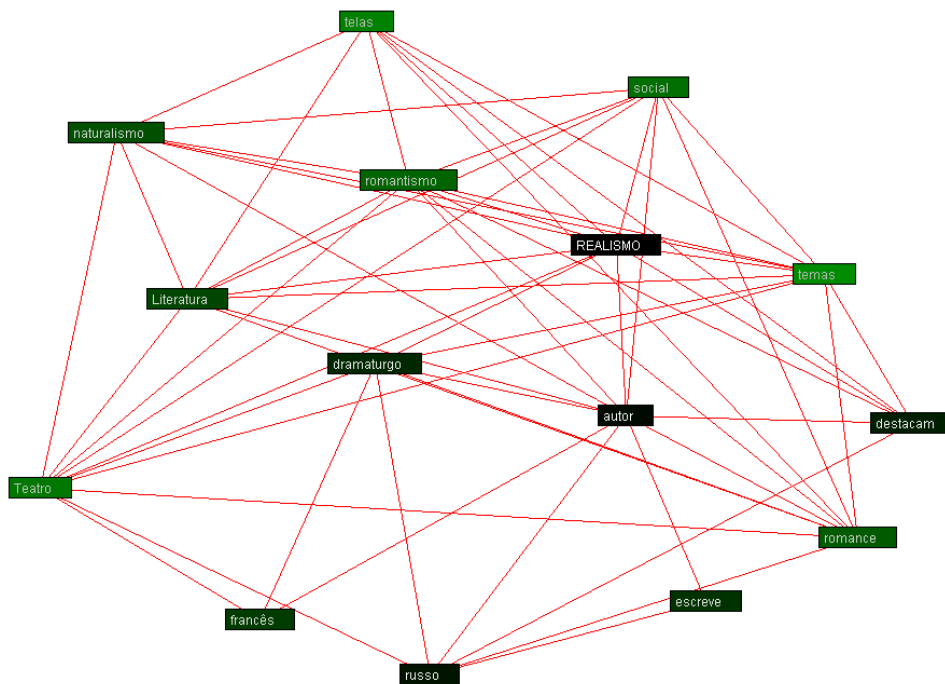


Figura 11 - Grafo gerado a partir do texto "Realismo" - Aluno 3

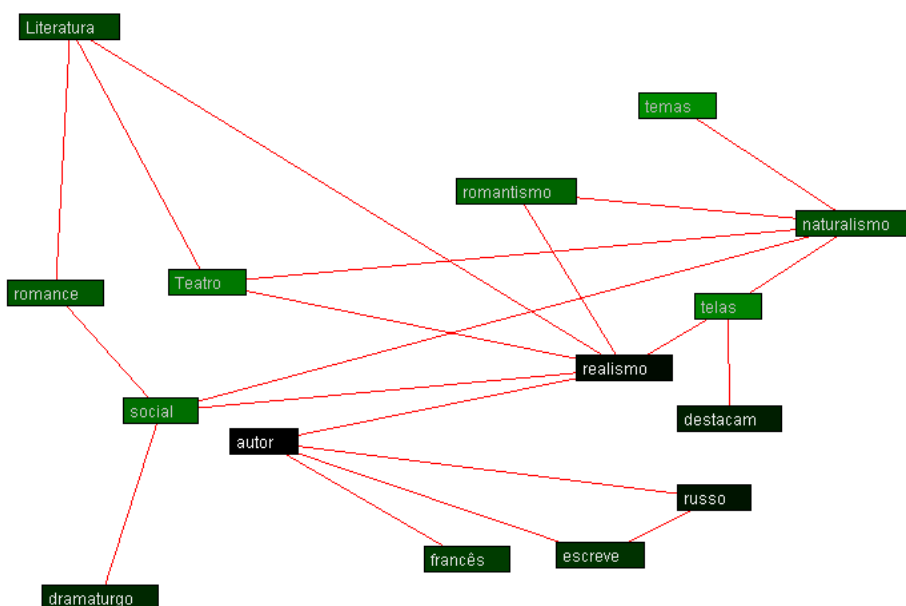


Figura 12 - Grafo gerado a partir do texto "Realismo" - Aluno 5

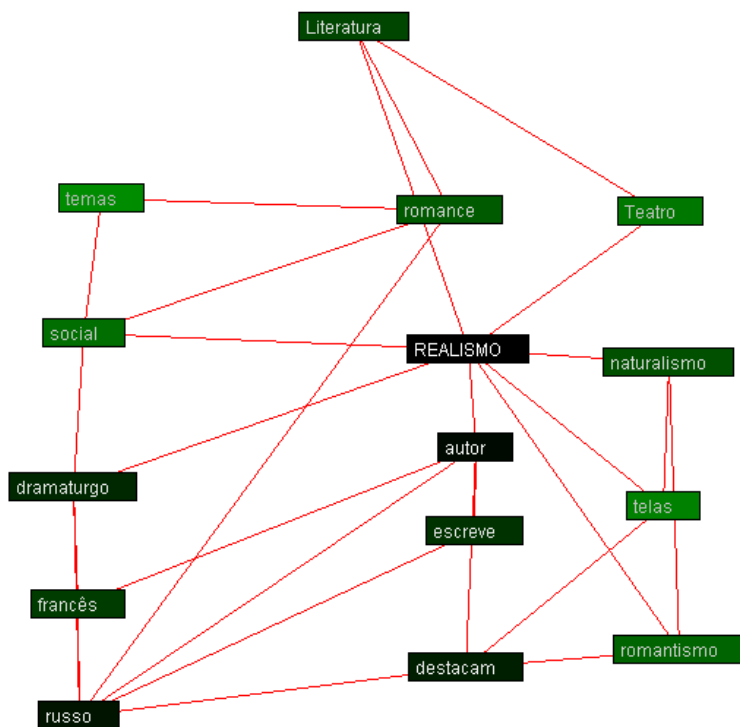


Figura 13 - Grafo gerado a partir do texto "Realismo" - Aluno 6

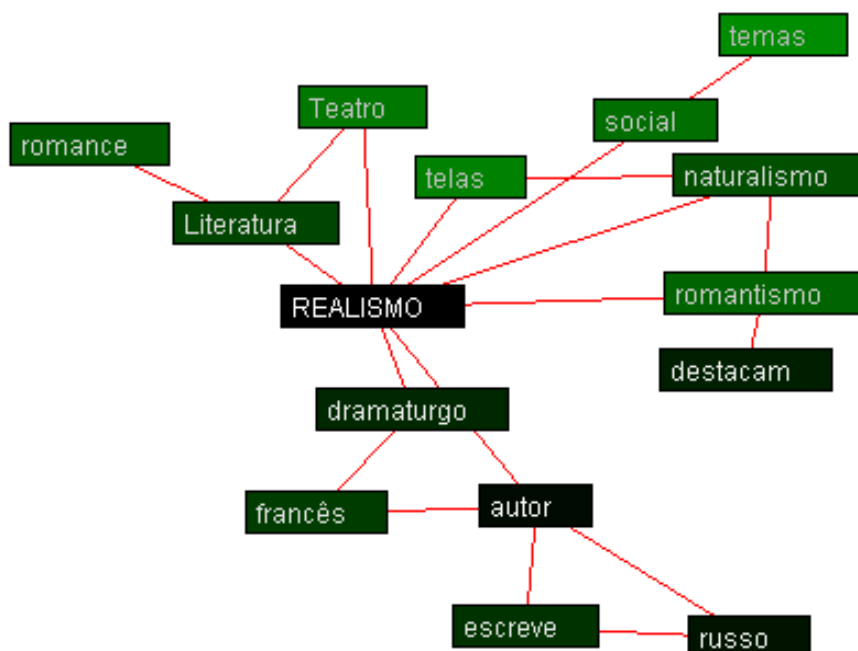


Figura 14 - Grafo gerado a partir do texto "Realismo" - Aluno 7

Nos grafos apresentados como exemplos (figuras 12, 13, 14 e 15), quinze termos foram utilizados. Este mesmo número de termos ocorreu em outros dez grafos que também foram organizados de forma diferente.

Análise do questionário

A pergunta do questionário relacionada a esta categoria é a de número 5. Abaixo são relatados alguns trechos das respostas dos alunos que demonstram que estes tentaram interpretar o que havia no grafo e ao mesmo tempo ajustá-lo de acordo com o que compreenderam do texto:

“Muitas vezes confundiu, mas ao procurar no texto o porque da relação, ajudou a entender.”(Aluno 17)

“De certa forma sim, pois com esse passo fomos conhecendo mais sobre o texto e estabelecendo as relações” (Aluno 16)

“Só organizei melhor.” (Aluno 15)

“Sim, ajudou a organizar as idéias principais” (Aluno 13)

“Sim, ajudou muito, após observar o grafo já tinha mais idéias relacionadas sobre o assunto” (Aluno 18)

“De certa forma sim, pois a medida que íamos reorganizando as palavras íamos compreendendo melhor o texto” (Aluno 20)

Ficou evidente que o Aluno 20 tentou interpretar o texto ao ajustar o grafo de acordo com o que compreendeu. Sabe-se, que a leitura e interpretação textual são fundamentais no desenvolvimento da aprendizagem dos alunos. Trata-se de um caminho muito relevante para a internalização do conhecimento (ANTUNES, 2008) e ainda, no desenvolvimento das habilidades reflexivas (HAYES & FLOWER, 1980) como mencionado anteriormente.

Outra observação interessante foi feita pelo Aluno 16, dizendo que *“[...] com esse passo fomos conhecendo mais sobre o texto e estabelecendo as relações”*. O aluno 13 complementa dizendo que a possibilidade de fazer ajustes nos grafos *“[...] ajudou a organizar as idéias principais”*. Koch (1998b) ressalta que o sentido não está somente no texto, mas é constituído a partir dele, na interação do leitor

com ele. Para esta autora, a leitura/escrita são processos que estão intimamente relacionados e podem levar a aprender a pensar, encadear idéias, argumentar, analisar. Todo o processo de compreensão de texto obedece a regras de interpretação e portanto, a coerência do texto não se estabelece sem levar em conta a interação e as crenças, desejos, querer e preferências (KOCH e TRAVAGLIA 1999). A relação positiva entre as observações feitas pelos alunos indica uma fluência na compreensão do texto original, denotando habilidades reflexivas. O esforço para observar detalhes dos relacionamentos e a re-organização do grafo serviram como um passo importante na organização das idéias, como ressaltou o Aluno 18 “[...] após observar o grafo já tinha mais idéias relacionadas sobre o assunto”. Desta forma, esta “conexão das idéias” contribuiu e facilitou a execução da Tarefa 5 - a produção textual.

4.3.2 Tarefas de escrita

Após terem realizado as tarefas relacionadas à pré-escrita, os alunos passaram para as tarefas de escrita. Na **Tarefa 5** os alunos iniciaram a sua produção textual a partir do grafo gerado, buscando novas interpretações para a representação gráfica produzida, agregando novos fatos e idéias relacionados à sua própria realidade. Nesta fase de escrita, foram definidas duas categorias de observação, apresentadas a seguir.

Quadro 6: Detalhamento da 6ª categoria de observação

Categoria 6: Construção do texto com base no grafo

Nesta categoria de observação buscamos analisar se o aluno utilizou o grafo na produção final do texto. As perguntas que guiaram esta etapa da análise dos dados foram:

- *o grafo serviu de base na construção do texto do aluno?*

- *o texto do aluno contém termos e relacionamentos presentes no seu grafo?*

Observações feitas a partir do filme

A partir da observação do filme, durante a fase de escrita, certas evidências permitiram constatar que a maioria dos alunos observou as relações apresentadas pelo grafo. A busca pela relação, um dos elementos centrais para escrever um texto, citado por Cabral (2009), ficou muito evidente. Para esta autora, num texto coerente, os fatos e conceitos devem estar relacionados. Observou-se que, durante a produção do texto, alguns alunos analisaram o grafo e imediatamente produziram o texto. Outros, porém, observaram o grafo, localizaram os termos novamente no texto original e só depois passaram para a sua produção textual. Em ambos os casos, os termos e as relações apresentados pelo grafo, foram explorados e a maioria incluídos na produção dos textos dos alunos.

Análise do questionário

As perguntas do questionário relacionadas a esta categoria são as de número 5 e 6. Do total dos 20 alunos, **9** afirmaram que o grafo serviu de base para construção do texto, **8** não se basearam no grafo e **3** informaram não terem utilizado o grafo diretamente, mas haviam observado os relacionamentos dos termos. Alguns trechos das respostas dos alunos que mostram como empregaram os grafos em suas produções:

“Utilizei sim, pois quis colocar todos os termos do grafo no meu texto” (Aluno 2)

“Usei várias vezes o grafo; olhava o grafo sempre que precisava para não me perder no texto e não ficar sem sentido.” (Aluno 8)

“Em certos momentos sim, na parte final eu precisei do grafo para saber partes importantes que deveriam aparecer no texto.” (Aluno 9)

“Sim utilizei o grafo (não muito), mas parecia que ele tinha uma ordem para produzir o texto”. (Aluno 11)

“sim, utilizei o grafo porque com as palavras era mais fácil pra produzir o texto”.

(Aluno 19)

Nas respostas dos alunos fica evidente que o grafo gerado pela ferramenta SOBEK colaborou na organização das idéias e influenciou no desenvolvimento das produções textuais. Quando observaram os termos do grafo, os alunos puderam visualizar e esquematizar mentalmente a seqüência dos conceitos e suas relações, e assim formular as suas frases. O fato de alguns alunos retornarem ao texto original reforçou o exercício de leitura e reflexão, pois alguns conceitos poderiam ou não estar relacionados, dependendo do enfoque dado pelo aluno. Para Shanahan (2008), tanto a leitura como a escrita de textos são duas faces do mesmo processo cognitivo. São aspectos do pensamento que estão essencialmente inter-relacionados. Para este autor, a escrita se baseia no conhecimento sobre o tema, na capacidade de organizar informações e na habilidade em encontrar as palavras certas para gerar novas informações. Neste sentido, os grafos serviram de apoio nesta fase de escrita, inicialmente na organização das idéias, bem como na organização das frases, parágrafos e texto como um todo.

Quadro 7: Detalhamento da 7ª categoria de observação

Categoria 7: Presença dos termos e relacionamentos do grafo no texto do aluno

Nesta categoria de observação buscamos localizar e analisar os termos presentes nas produções textuais dos alunos.

O registro obtido pelo filme não permitiu identificar nenhuma evidência relacionada a esta categoria de observação. Quanto ao questionário, nenhuma de suas questões estava relacionada a esta categoria de observação. Contudo, uma análise dos próprios textos mostra que em praticamente todas as produções os estudantes utilizaram muitos dos termos representados no grafo. A tabela 3 apresenta o número de termos do grafo gerado individualmente e o número de termos utilizados em cada uma das produções textuais. Os resultados mostraram que cada aluno utilizou, em média, entre nove e dez termos do grafo na sua produção. Tais resultados ratificam a idéia de que o grafo serviu de apoio aos

estudantes na construção de seus textos.

Abaixo, excertos dos textos produzidos pelos alunos utilizando termos selecionados pelo grafo:

“No Brasil o realismo é representado através do teatro e da literatura, onde cada tela é representada pelo povo, de acordo com seu dia-a-dia.” (Aluno 1)

“No Brasil o realismo marca mais intensamente no teatro e na literatura, usando os temas daquela época, escravidão, burguesia e outros.” (Aluno 2)

“O Realismo na literatura manifesta-se através de conversa [...] através de telas e de teatro, os principais temas era o preconceito racial e [...]” (Aluno 3)

“O Realismo tem o objetivo de mostrar a realidade e o interesse pelos temas sociais. Ele apresenta uma reação no romantismo, muitas vezes o romantismo e naturalismo se confundem.” (Aluno 19)

Nos excertos dos textos produzidos, de uma forma geral, foi possível identificar correlação com o conteúdo abordado no texto original. Ficou evidente que os alunos se basearam nos conceitos e relações apresentados nos grafos para elaboração do seu texto. Os comentários dos alunos 2 e 19, na análise do questionário da categoria 6, confirmam esta hipótese:

“Utilizei sim, pois quis colocar todos os termos do grafo no meu texto” e ainda, “sim, utilizei o grafo porque com as palavras era mais fácil pra produzir o texto”.

Na teoria de aprendizagem significativa desenvolvida por Ausubel (1982), a aprendizagem é um processo em que uma nova informação é vinculada a um aspecto relevante, já existente, da estrutura de conhecimento de um indivíduo. A partir dos termos apresentados nos grafos, os alunos puderam evidenciar significados atribuídos a estes conceitos, bem como as relações entre conceitos, e assim, foram planejando e organizando suas idéias para a sua produção textual.

Planejar o texto significa organizar-se e fazer uma previsão do que irá escrever. Antes de tudo, é preciso saber sobre o que escrever e, para isso, os alunos puderam recolher informações e idéias nos grafos apresentados. Para Cabral (2009), num texto coerente, os fatos e conceitos devem estar relacionados. Essa

relação deve ser suficiente para justificar sua inclusão num mesmo texto. As relações de sentido e significado devem ser estabelecidas tanto com os elementos que antecedem como com os que se sucedem (Costa Val, 1993), pois coerência, além de dar sentido ao texto, também permite a sua compreensão. As observações feitas da forma como os participantes utilizaram os grafos produzidos pela ferramenta SOBEK mostram que os alunos buscaram estabelecer estas relações de sentido muito calcados nas representações gráficas apresentadas pela ferramenta.

ALUNOS	Número de termos do grafo	Número de termos utilizados no texto
Aluno 1	13	5
Aluno 2	22	20
Aluno 3	15	10
Aluno 4	16	10
Aluno 5	15	10
Aluno 6	15	7
Aluno 7	15	13
Aluno 8	15	10
Aluno 9	15	14
Aluno 10	44	6
Aluno 11	15	11
Aluno 12	18	7
Aluno 13	15	11
Aluno 14	15	12
Aluno 15	15	10
Aluno 16	15	9
Aluno 17	15	11
Aluno 18	15	6
Aluno 19	15	13
Aluno 20	17	14

Tabela 3 -Tabela termos gerados e termos utilizados nos textos

5 CONSIDERAÇÕES FINAIS

Este estudo apresentou uma proposta para o uso da ferramenta SOBEK com o objetivo de investigar como esta ferramenta de mineração de texto pode ser empregada como apoio à produção textual. A partir dos objetivos específicos elencados, desenvolvemos uma metodologia para utilização deste sistema de mineração de textos auxiliando nas atividades de produção textual. Esta abordagem foi fundamentada no ensino baseado em tarefa (ELLIS, 2003) considerando as dificuldades que envolvem a expressão escrita do aluno.

Quanto ao uso da ferramenta SOBEK em outras pesquisas na área educacional, Macedo (2009) propôs a utilização da ferramenta para apoiar o trabalho docente na avaliação e acompanhamento de trabalhos de escrita colaborativa. Aqui, nosso foco foi dar suporte ao próprio trabalho dos estudantes, buscando auxiliá-los em tarefas de escrita através do emprego da mineração de textos.

A escolha desta ferramenta aconteceu pelo fato de ela já ter sido desenvolvida com um olhar educacional. Com relação a outras ferramentas de análise de textos apresentadas na seção 3.3, a ferramenta SOBEK se distingue principalmente por apresentar tanto os principais conceitos encontrados em um texto, quanto os relacionamentos entre estes, a partir do grafo gerado.

A idéia de utilização de grafos para apoiar a construção textual é análoga à idéia de emprego de mapas conceituais, abordagem desenvolvida na década de 70 por Joseph Novak (1984) e já mencionada neste trabalho. A ferramenta SOBEK, a construção dos grafos é mais interativa no sentido de que o aluno solicita à ferramenta determinadas ações (extração de termos, identificação de relacionamentos, construção do grafo), e pode intervir ele mesmo neste processo.

Além disso, como mencionado anteriormente, o fato da ferramenta ter sido implementada pelo grupo de pesquisa liderado pelo prof. Eliseo Reategui, orientador deste trabalho, mostrou-se um grande facilitador para o desenvolvimento dos ajustes necessários para o uso da ferramenta. Estes ajustes foram no sentido de prover

melhorias no sistema de mineração de textos utilizado, focando em sua aplicação como ferramenta educacional para apoio à produção textual. Não foi uma questão de desconsiderar outras ferramentas, mas sim propor o uso de uma nova tecnologia no ambiente escolar para o desenvolvimento de tarefas de leitura e principalmente de escrita.

Um experimento inicial foi realizado em 2009. Um estudo de caso no qual um estudante realizou a leitura de um texto na área de contabilidade empresarial, e depois disso empregou a ferramenta num procedimento que envolveu: extração de termos frequentes; criação de uma base de conceitos e relacionamentos a partir dos termos extraídos automaticamente; geração de um grafo correspondente aos termos e relacionamentos estabelecidos; escrita do texto com base no grafo gerado.

Observou-se que o processo de “idas e vindas” do aluno ao texto, para completar cada uma das etapas de utilização da ferramenta, foi construtivo para promover uma melhor compreensão do texto e do tema tratado. A produção do texto de maneira bastante fluida foi atribuída às várias etapas adotadas no uso da ferramenta.

Com relação à ferramenta, alguns problemas foram observados, como a dificuldade de extração de termos efetivamente relevantes, ou a necessidade de construção de uma interface gráfica para manipulação da base de conceitos diretamente nos grafos gerados.

A partir deste experimento, diversas modificações na ferramenta SOBEK foram implementadas para que esta pudesse ser empregada com mais facilidade pelos estudantes. Um estudo com um número de alunos maior foi sendo estruturado para que pudéssemos avaliar de maneira mais ampla os benefícios introduzidos pela ferramenta no processo de produção textual.

Então, em 2010 foi realizado o principal experimento desta pesquisa, envolvendo uma turma de 20 alunos do Ensino Médio. A partir da análise de conteúdo das respostas do questionário, dos textos e vídeos produzidos é possível inferir que o uso da Ferramenta de Mineração de Textos SOBEK é positivo e eficaz na produção textual se aplicado seguindo uma metodologia e planejamento específicos.

Os relatos mostraram que a busca pelos termos apresentados pelo programa no texto original, fez com que os alunos aprofundassem sua compreensão sobre o tema. Desta forma, ao instigar o aluno a explorar o texto original com mais profundidade, o método proposto levou-o a um maior esclarecimento sobre o que compreendeu e o fez perceber detalhes importantes em relação aos conceitos. Nos experimentos realizados, a releitura produziu uma melhora acentuada na compreensão dos textos iniciais. A partir dos termos do grafo, os alunos puderam visualizar e esquematizar mentalmente a seqüência dos conceitos e suas relações, formular as suas frases organizando assim os seus textos.

As principais contribuições deste trabalho foram: [1] a elaboração e avaliação de uma metodologia para emprego da ferramenta SOBEEK como apoio à produção textual. Ressaltamos que o papel do professor é fundamental, porque os professores são responsáveis pelas atividades de aprendizagem e definição das tarefas que serão desenvolvidas em sala de aula, incluindo tarefas de leitura e de escrita. E [2] a concepção de ajustes e novas funcionalidades para a ferramenta, de acordo com necessidades específicas observadas durante o desenvolvimento desta pesquisa.

Pretende-se futuramente disponibilizar a ferramenta de mineração de textos para outras aplicações e outras instituições, buscando compreender como a extração automática de conteúdos e sua representação gráfica pode contribuir em outras atividades educacionais, como por exemplo, em processos de letramento e resolução de problemas.

REFERÊNCIAS

ALVES, M. R. N. R. **Características da escrita de alunos da 1ª série do Ensino Médio**. Revista Eletrônica de Divulgação Científica em Língua Portuguesa, Linguística e Literatura - REVISTA LETRA MAGNA. Ano 04 n.06 -1º Semestre de 2007 ISSN 1807-5193, disponível em:
http://www.letramagna.com/caract_escrita_em.pdf . Acesso em jun. 2010.

ANTUNES, Irandé. **Aula de Português: encontro e interação**. 2. ed. São Paulo: Parábola Editorial, 2008.

AZEVEDO, B. F. T..REATEGUI, E.. BEHAR, P.A. Artigo: **Estudo de Análise Qualitativa em Fórum de Discussão**. Disponível em:
<http://www.cinted.ufrgs.br/renote/dez2009/artigos/8a_brenofabricio.pdf > Acesso em mar. 2010.

AUSUBEL, D. P. **A aprendizagem significativa: a teoria de David Ausubel**. São Paulo: Moraes, 1982.

BAKHTIN, Mikhail. **Estética da Criação Verbal**. Tradução por Maria Ermantina G. Gomes. São Paulo: Martins Fontes, 2000.

BARBOSA, M.L.. SEVERO, C. E. P..REATEGUI, E. Artigo: **Mineração de padrões no gênero textual blog**. Disponível em: <
http://www.cinted.ufrgs.br/renote/dez2009/artigos/8e_marialucia.pdf> Acesso jun. 2010.

BRANDÃO, A. C. P.; SPINILLO, A. G. **Produção e compreensão de textos em uma perspectiva de desenvolvimento**. v. 6.Natal, Estudos de Psicologia, 2001, p. 51-62.

BRONCKART, Jean-Paul. **Atividade de linguagem, textos: por um interacionismo sócio-discursivo**. Tradução por Anna Raquel Machado. São Paulo: EDUC, 2003.

CABRAL, Marina. Artigo **O Texto Escrito. Brasil Escola, 2009**. Disponível em: <<http://www.brasilecola.com/redacao/texto-escrito.htm>>. Acesso em ago. de 2009.

CASTELLS, Manuel. **A Galáxia da Internet: reflexões sobre a internet, os negócios e a sociedade**. Tradução por Maria Luiza X. de A. Borges. Rio de Janeiro. Jorge Zahar Editor, 2003, p.243.

COMPUTER-ASSISTED INSTRUCTION AND WRITING – disponível em <http://www.k8accesscenter.org/training_resources/computeraided_writing.asp> Acesso em maio 2010.

COSTA VAL, Maria da Graça Costa. **Redação e textualidade**. 2 ed. São Paulo: Martins Fontes, 1999.

COSTA VAL, Maria da Graça Costa. **Repensando a textualidade**. In: AZEVEDO, José Carlos (org). *Língua Portuguesa em Debate: conhecimento e ensino*. Petrópolis: Vozes, 2001, p. 34-51.

CRESWELL, John W. **Projeto de pesquisa: métodos qualitativo, quantitativo e misto**. Porto Alegre: Artmed, 2007, 248 p.

DA SILVA, J., FAGUNDEZ, L., BASSO, M. **Metodologia para apoiar o processo de aprendizagem através da produção de objetos de aprendizagem**. En J. Sánchez (Ed.): *Nuevas Ideas en Informática Educativa*, v. 3, Santiago de Chile: LOM Ediciones, 2007, p. 297-320.

DEMO, Pedro. **Educar pela pesquisa**. Campinas: Autores Associados, 1996.

DUTRA, Ítalo M. **Mapas conceituais no acompanhamento dos processos de conceituação. Tese de doutorado** – Universidade Federal do Rio Grande do Sul, Centro de Interdisciplinar de Novas Tecnologias na Educação, Programa de Pós-Graduação em Informática na educação, Porto Alegre, RS, 2006.

EBECKEN, N.F.F.; LOPES, M.C.S & COSTA, M.C.A..- “**Mineração de Textos**”. In: REZENDE, S.O., 2003, “Sistemas Inteligentes – Fundamentos e Aplicações”. Cap. 13. Barueri. SP. Ed. Manole, 2004.

ELLIS, Andrew W. **Leitura, escrita e dislexia: uma análise cognitiva**. Porto Alegre: Artes Médicas, 1995

ELLIS, Rod. **Task-based language learning and teaching**. Oxford: New York. University Press, 2003.

FELDMAN, R., SANGER, J. **Text Mining Handbook**. Cambridge, Inglaterra: Cambridge University Press, 2006.

FLICK, Uwe. **Introdução à Pesquisa Qualitativa (Métodos de Pesquisa)**. Tradução por Joice Elias Costa – 3.ed. – Porto Alegre: Artmed, 2009.

FRANCO, Sérgio. **Educação a Distância**. Disponível em: <<http://www.sinepe-rs.org.br/forum.htm>>. Acesso em fev. 2010.

GIOMBINI, L. **Concept Maps: Theory, Methodology, Technology. Proceedings of the First Int. Conference on Concept Mapping**. Pamplona, Spain, 2004.

HAYES, J.R.. **New Directions in Writing Theory**. In C.MacArthur, S. Graham & J. Fitzgerald (Eds.), *Handbook of Writing Research*. New York: The Guilford Press, 2008, p28-39.

HAYES, J.R. FLOWER, L.S. **Identifying the organization of the writing process**. In L.W. Gregg& E.R. Steinberg (Eds), *Cognitive processes in writing*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc., 1980.

_____. **Writing research and the writer - Identifying the Organization of Writing Processes**. *American Psychologist*, 1986, p. 1106-1113.

HEARST, M. **Untangling Text Data Mining**. Proceedings of ACL'99: the 37th Annual Meeting of the Association for Computational Linguistics, Jun 1999.

HESSLER, Juliana ; REATEGUI, Eliseo . **Um Método para Apoio à Leitura Baseado no uso de uma Ferramenta de Mineração de Texto**. RENOTE. Revista Novas Tecnologias na Educação, v. 8, 2010, p. 1-10.

JOHNSON, Steven. **Cultura da Interface: como o computador transforma nossa maneira de criar e comunicar**. Tradução por Maria Luísa X. de A. Borges; revisão técnica, Paulo Vaz. Rio de Janeiro: Jorge Zahar Ed., 2001.

JOLIBERT, Josette. JACOB, Jeannette e colaboradores. **Além dos muros da escola: a escrita como ponte entre alunos e comunidade**. Tradução por Ana Maria Netto Machado. Porto Alegre: Artmed, 2006.

_____. **Formando crianças leitoras**. Porto Alegre: Artes Médicas, 1994a.

_____. **Formando crianças produtoras de textos**. Tradução por Walkiria M. F. Settineri e Bruno Charles Magne. v. 2. Porto Alegre, Artes Médicas, 1994b.

JONSSON, Ewa. **Electronic Discourse. On Speech and Writing on the Internet**. Luleå University of Technology. Department of Communication and Languages, 1997. Disponível em: < <http://www.ludd.luth.se/users/jonsson/D-essay/ElectronicDiscourse.html> > Acesso em jun. 2010.

KOCH, I. G. V. **A coesão textual**. 10. ed. São Paulo: Contexto, 1998a.

_____. **A coerência textual**. 8. ed. São Paulo: Contexto, 1998b.

KOCH, I.G.V. e TRAVAGLIA, L.C. **Texto e coerência**. São Paulo: Cortez, 1999.

LORENZATTI, A. **SOBEK: uma Ferramenta de Mineração de Textos**. Caxias do Sul: Departamento de Informática/UCS. Trabalho de Conclusão de Curso, 2007.

LUFT, Celso Pedro. **Língua e liberdade**. 7.ed. São Paulo: Ática, 1999.

MACEDO, A. **Rede de Conceitos: Uma ferramenta para contribuir com a prática pedagógica no acompanhamento da produção textual coletiva**. Tese (Doutorado em Informática na Educação) – Universidade Federal do Rio Grande do Sul, Centro de Interdisciplinar de Novas Tecnologias na Educação, programa de Pós-Graduação em Informática na educação, Porto Alegre, RS, 2009.

MACEDO, A., Reategui, E., LORENZATTI, A., BEHAR, P. **Using Text-Mining to Support the Evaluation of Texts Produced Collaboratively**. Education and Technology for a Better World: Selected papers of the 9th World Conference on Computers in Education, Bento Gonçalves, Brazil. Springer, 2009.

MACHADO, Nádie C. F. **Estudo das Trajetórias de Letramento em Curso de Educação a Distância**. Tese (Doutorado em Informática na Educação) – Universidade Federal do Rio Grande do Sul, Centro de Interdisciplinar de Novas Tecnologias na Educação, programa de Pós-Graduação em Informática na educação, Porto Alegre, RS, 2009.

MARCUSCHI, Luiz Antonio. **Produção Textual, Análise de Gêneros e Compreensão**. Coleção: Educação Lingüística – 1.ed. São Paulo: Editora Parábola, 2008.

MARCUSCHI, Luiz Antonio; XAVIER, Antônio Carlos. **Hipertexto e gêneros digitais: novas formas de construção do sentido**. – 2. Ed. Rio de Janeiro, 2005.

MATTISON, Robert. MATTISON, Brigitte Kilger. **Web Warehouse and Knowledge Management**. 1.ed. Mc Graw Hill, New York, 1999.

MEDINA, N. FREITAS, P. J.. Artigo **Desenvolvimento do Pensamento Crítico na Escrita Colaborativa**. Disponível no endereço:
<http://www.cinted.ufrgs.br/renote/nov2004/artigos/r4_pensamento_critico.pdf>
Acesso mar. 2010.

NOVAK, J.D., & Gowin, D.B. **Learning how to learn**. New York, N.Y.: Cambridge University Press, 1984.

NOVAK, J. D.; GOWIN, D. B. **Teoria y practica de la educación**, 1988.

NUNAN, D. **Designing Tasks for the Communicative Classroom**. Cambridge: Cambridge University Press, 1989.

OLIVEIRA, F. L. Artigo: **O Processo de Leitura e Escrita e suas implicações na Aprendizagem dos alunos**. IX Congresso Nacional de Educação – EDUCERE, III Encontro Sul Brasileiro de Psicopedagogia – PUCPR (26 a 29 de outubro de 2009) Disponível em:
<http://www.pucpr.br/eventos/educere/educere2009/anais/pdf/2946_1789.pdf> .
Acesso: jun. 2010.

OLIVEIRA, Maria Marly de. **Como fazer pesquisa Qualitativa**. Recife: Ed. Bagaço, 2005, p. 192.

PCN: Parâmetros Curriculares Nacionais. Secretaria de Educação Fundamental. Brasília: MEC/SEF, 2001 Livro 02. Disponível em
<<http://portal.mec.gov.br/seb/arquivos/pdf/livro02.pdf>>. Acesso em mar. 2010.

_____. **Parâmetros Curriculares Nacionais: Língua Portuguesa**. Ensino de 1^a a 4^a séries Secretaria de Educação Fundamental Brasília, 1997 p 19. Disponível em:
<<http://portal.mec.gov.br/seb/arquivos/pdf/portugues.pdf>> Acesso em mar. 2010.

RIZZO, Gilda. **Fundamentos e Metodologia da Alfabetização – Método Natural**. 4. ed. Rio de Janeiro. Coleção Educação Prática, 1986.

SHANAHAN, Timothy. **Relations among Oral Language, Reading, and Writing Development**. Handbook of Writing Research. In: MacArthur, C., Graham, S. e Fitzgerald, J. (Eds.). New York: The Guilford Press, 2008, p. 171-183.

SCHANKER, A. **Graph-Theoretic Techniques for Web Content Mining**. *PhD thesis* (Tese Doutorado em Ciência da Computação) - University of South Florida, 2003.

SHARP Mark. **Text Mining. Rutgers University, School of Communication, Information and Library Studies, 2001.** Disponível em:
<http://www.scils.rutgers.edu/msharp/textminig.htm>. Acesso em jun. de 2010.

SKEHAN, P. **A Cognitive Approach to Language Learning.** Oxford: Oxford University Press, 1998.

VALENTE, J.A. **A espiral da aprendizagem e as tecnologias da informação e comunicação: repensando conceitos.** Em M.C. Joly (ed.) *Tecnologia no Ensino: implicações para a aprendizagem.* São Paulo: Casa do Psicólogo Editora, 2002, p. 15-37.

VIEIRA, M. C. T. **Levantamento das Dificuldades de Alunos do 1º Ano da Universidade na Compreensão de Materiais Escritos.** Dissertação de Mestrado, PUC-SP, 1981.

WILLIS, J. A. **Framework for Task-based Learning.** London: Longman, 1996.

ANEXOS

ANEXO A

A coerência Textual²⁰.

Ingedore Villaça Koch; Luiz Carlos Travaglia (resumo)

A interpretabilidade é fator fundamental para determinar a coerência de uma produção textual. A falta de relações lógicas entre as idéias apresentadas, contradições e a ausência de conhecimento de contextos gerais de produção podem tornar um texto incoerente, dificultando o entendimento entre os falantes da língua em uma situação de interação comunicativa.

Segundo Koch & Travaglia a coerência não depende apenas da forma como elementos lingüísticos são utilizados, mas também do conhecimento de mundo apresentado pelos interlocutores. A partir do conhecimento arquivado na memória do ouvinte/leitor, um texto, por exemplo, que não possui ligações sintáticas explícitas entre suas partes pode apresentar uma unidade de sentido, tornando-se coerente para o receptor da mensagem que decide cooperar e aceitar a sequência de idéias como sendo um texto. Para a relação entre elementos do texto ser considerada explícita, inclusive sintaticamente, recursos da coesão são utilizados, auxiliando a produção de coerência textual, embora aquela não seja uma condição necessária para que essa ocorra.

Além de relações semânticas e gramaticais, podem ser estabelecidas também relações pragmáticas, ou seja, as que ocorrem por meio das ações realizadas durante os atos de fala.

A continuidade de idéias e conhecimentos ativados nos interlocutores que levam a formação de uma unidade de sentido no todo do texto é considerada a base da coerência por Beaugrande & Dressler (1981) e Marcuschi (1983) (Apud Koch e Travaglia), pois ativa conhecimentos através de expressões no texto. Essa continuidade de sentidos engloba conceitos e relações subjacentes á superfície textual, estabelecendo uma coesão do texto de forma cognitiva entre os usuários da língua. Tais relações não são lineares, tornando a coerência uma organização reticulada, tentacular e hierarquizada. Embora a continuidade seja um fator essencial para um texto ser considerado como coerente, nem sempre a descontinuidade representará uma incoerência, podendo ser utilizada como recurso de produção, dependendo da intenção do produtor do enunciado.

A coerência textual é global. Um texto será coerente a partir do momento em que sua totalidade puder produzir uma única unidade de sentido. Essa coerência global é resultado de um conjunto de coerências locais, ou seja, sequências lógicas dentro do texto que, quando reunidas, serão capazes de promover uma intenção comunicativa.

Para que essas relações sejam harmoniosas e não produzam incoerências, é necessário que a relação entre coerência e coesão seja realizada de maneira adequada. Enquanto a coerência é subjacente ao texto e não-linear, a coesão deverá dar sentido

²⁰ AMORIM,H.Publicado no Recanto das Letras em 05/03/2010. Código do texto: T2122486. Disponível em: <http://recantodasletras.uol.com.br/resenhasdelivros/2122486>

entre os elementos da superfície textual, sendo revelada por meio de marcas lingüísticas, dando-lhe um caráter linear. A coesão, conectora das diferentes partes do texto, é sintática e gramatical, mas também semântica.

A coerência é a responsável para que uma sequência lingüística seja considerada texto, pois é através dela que relações (sintático-gramaticais, semânticas e pragmáticas) poderão ser estabelecidas formando uma unidade global de sentido e dando a ela a textualidade. A coesão auxilia nesse processo, mas não é responsável pelo fenômeno da textualidade, já que não é necessária nem suficiente para isso.

Considerando a produção textual uma atividade interacional dependente não só de aspectos lingüísticos teóricos, mas também de uma situação comunicativa, a incoerência não dependerá apenas de sequências lingüísticas, mas também de seus usuários, intenção e situação comunicativa. Um texto pode ser considerado incoerente em determinado momento, mas fazer sentido em outro, dependendo do contexto em que será aplicado, do conhecimento de mundo e compartilhamento desses conhecimentos entre os falantes e da aceitação ou não pelos usuários, ratificando o conceito de interpretabilidade como princípio da coerência e levando a considerar que não existe um não-texto.

Ainda ao considerar a interpretabilidade como fator fundamental de coerência, não seria possível definir diversos tipos de coerência, mas sim considerar que dependendo da intenção do produtor textual ou da superestrutura do texto que se pretende trabalhar, os mecanismos lingüísticos e cognitivos serão utilizados de forma diferente, a fim de obter o resultado final esperado, alcançando o objetivo determinado previamente pelo enunciador.

Segundo os autores, de acordo com o exposto, fica claro que a construção da coerência depende de vários fatores, sejam eles lingüísticos, cognitivos, culturais, interacionais. Assim, do ponto de vista lingüístico, elementos utilizados no decorrer da produção textual são de indiscutível importância para que a coerência seja estabelecida, pois funcionam como pistas que servirão como ponto de partida para sua realização. Conhecimento de mundo, conhecimento compartilhado e inferências, respectivamente, relacionam-se com a bagagem de conhecimento que cada interlocutor possui, que compartilham e que podem ser ativados a partir de mecanismos lingüísticos utilizados no decorrer da interação; os fatores de contextualização ancoram o texto em determinada situação comunicativa, auxiliando na localização do receptor e na decodificação da mensagem; a situacionalidade permite a adequação do texto à situação sociocomunicativa; a informatividade refere-se ao grau de previsibilidade ou expectabilidade da informação presente no texto. Quanto mais previsível a informação contida no texto, menos informação nova ele proporcionará; a focalização ocorre quando produtor e receptor concentram-se em apenas uma parte de seus conhecimentos, para que não haja problemas de compreensão e incoerência; a intertextualidade é a possibilidade de um texto recuperar ou recorrer ao conhecimento prévio de outros textos por parte dos interlocutores; intencionalidade e aceitabilidade referem-se ao modo como o emissor utiliza-se de produções textuais para alcançar seu objetivo e produzir efeitos esperados, e ao estabelecimento da coerência do receptor que age cooperativamente para dar sentido ao que lhe foi exposto, respectivamente; por fim, consistência e relevância exigem que os enunciados de um texto não sejam contraditórios entre si e que sejam interpretáveis como tratando de um mesmo assunto.

Para finalizar, Koch e Travaglia buscam mostrar pontos principais que possam auxiliar o trabalho de professores de língua Portuguesa, no que diz respeito à coerência, coesão e texto do ponto de vista lingüístico, sem a necessidade de detalhar técnicas de ensino ou se aprofundar em questões teóricas, mas abrindo os olhos dos profissionais quanto as suas metodologias, concepções sobre educação, preparo e segurança diante dos alunos e das possíveis respostas que irão receber durante o processo de interação em sala de aula.

ANEXO B

SOBEK

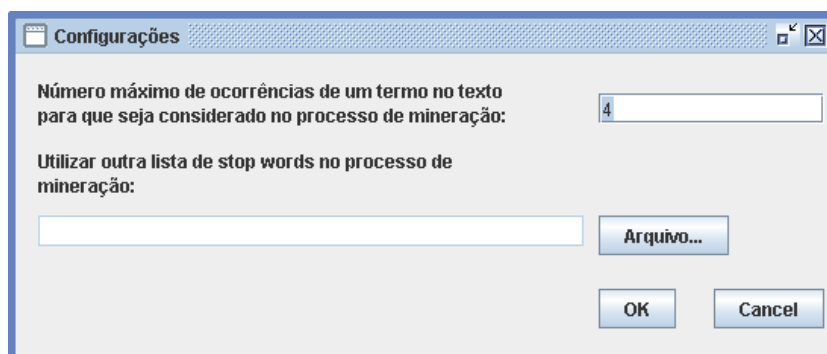


Figura 15 - Tela ferramenta SOBEK - Configurações

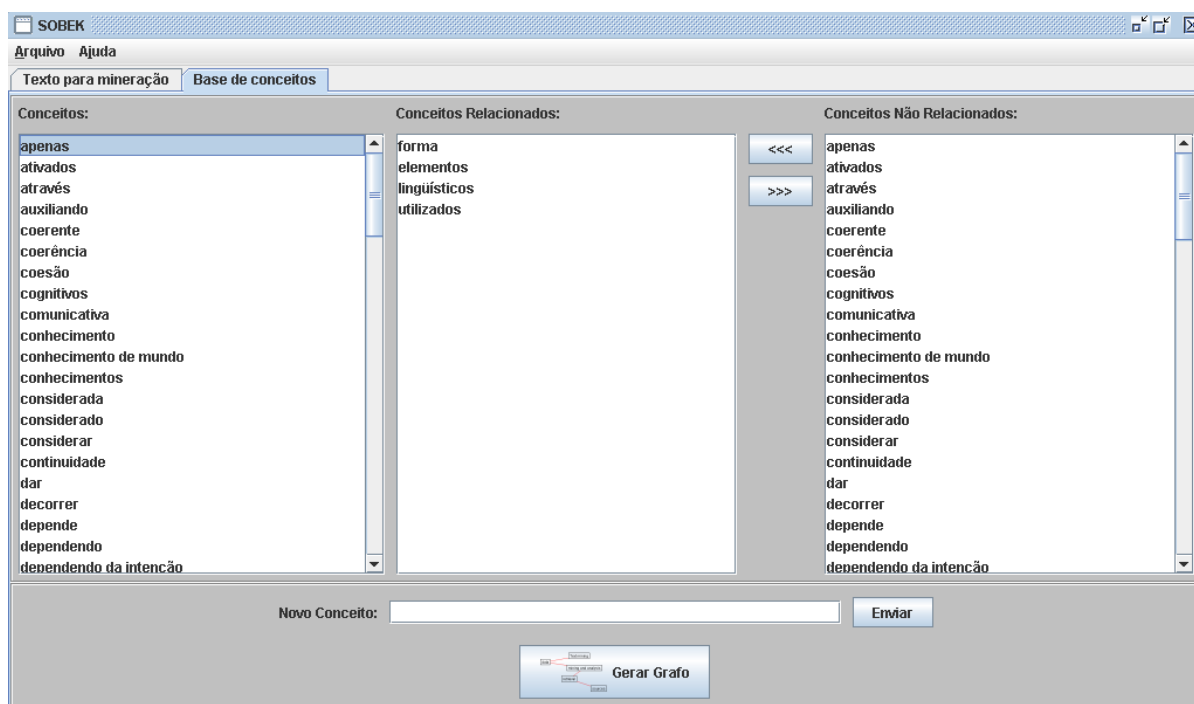


Figura 16 - Tela ferramenta SOBEK Base de Dados - gerado a partir do texto "A coerência textual"

ANEXO C
Questionário

**A utilização da ferramenta de mineração de texto – SOBEK -
como apoio à produção textual.**

Nome: _____

Idade: _____

Sexo: () Masculino () Feminino

Responda:

1. Os grafos lhe auxiliaram a **refletir** sobre o texto, propondo termos e relacionamentos que poderiam ser importantes?
2. A visualização gráfica de termos importantes e relacionamentos lhe ajudaram a melhor **compreender** o texto?
3. Relate como você procedeu para analisar as relações apresentadas entre os termos.
4. A releitura do texto original buscando localizar termos e relacionamentos que apareciam no grafo fez com que você observasse detalhes importantes no texto?

Justifique.

5. A re-organização do grafo acrescentando/retirando termos e alterando os relacionamentos serviu como um passo importante na organização das idéias sobre o que você deveria escrever no seu texto?
6. Você utilizou seu grafo na construção final do texto? Em caso positivo, relate como foi este processo.
7. Sobre a ferramenta SOBEK destaque:
 - a. pontos positivos:
 - b. pontos negativos:

ANEXO D

REALISMO

Movimento artístico que se manifesta na segunda metade do século XIX. Caracteriza-se pela intenção de uma abordagem objetiva da realidade e pelo interesse por temas sociais. O engajamento ideológico faz com que muitas vezes a forma e as situações descritas sejam exageradas para reforçar a denúncia social. O realismo representa uma reação ao subjetivismo do romantismo. Sua radicalização rumo à objetividade sem conteúdo ideológico leva ao naturalismo. Muitas vezes realismo e naturalismo se confundem.

Artes Plásticas - A tendência expressa-se sobretudo na pintura. As obras privilegiam cenas cotidianas de grupos sociais menos favorecidos. O tipo de composição e o uso das cores criam telas pesadas e tristes. O grande expoente é o francês Gustave Courbet (1819-1877). Para ele, a beleza está na verdade. Suas pinturas chocam o público e a crítica, habituados à fantasia romântica. São marcantes suas telas *Os Quebradores de Pedra*, que mostra operários, e *Enterro em Ornans*, que retrata o enterro de uma pessoa do povo. Outros dois nomes importantes que seguem a mesma linha são Honoré Daumier (1808-1879) e Jean-François Millet (1814-1875). Também destaca-se Édouard Manet (1832-1883), ligado ao naturalismo e, mais tarde, ao impressionismo. Sua tela *Olympia* exhibe uma mulher nua que "encara" o espectador.

Literatura - O realismo na Literatura manifesta-se na prosa. A poesia da época vive o parnasianismo. O romance - social, psicológico e de tese - é a principal forma de expressão. Deixa de ser apenas distração e torna-se veículo de crítica a instituições, como a Igreja Católica, e à hipocrisia burguesa. A escravidão, os preconceitos raciais e a sexualidade são os principais temas, tratados com linguagem clara e direta.

Na passagem do romantismo para o realismo misturam-se aspectos das duas tendências. Um dos representantes dessa transição é o escritor e dramaturgo francês Honoré de Balzac (1799-1850), autor do conjunto de romances *Comédia Humana*. Outros autores importantes são os franceses Stendhal (1783-1842), que escreve *O Vermelho e o Negro*, e Prosper Mérimée (1803-1870), autor de *Carmen*, além do russo Nikolay Gogol (1809-1852), autor de *Almas Mortas*.

O marco inicial do realismo na Literatura é o romance *Madame Bovary*, do francês Gustave Flaubert (1821-1880). Outros autores importantes são o russo Fiódor Dostoiévski (1821-1881), cuja obra-prima é *Os Irmãos Karamazov*; o português Eça de Queirós (1845-1900), que escreve *Os Maias*; o russo Leon Tolstói (1828-1910), criador de *Anna Karenina* e *Guerra e Paz*; os ingleses Charles Dickens (1812-1870), autor de *Oliver Twist*, e Thomas Hardy (1840-1928), de *Judas, o Obscuro*.

A tendência desenvolve-se também no conto. Entre os mais importantes autores destacam-se o russo Tchekhov (1860-1904) e o francês Guy de Maupassant (1850-1893).

Teatro - Com o realismo, problemas do cotidiano ocupam os palcos. O herói romântico é substituído por personagens do dia-a-dia e a linguagem torna-se coloquial. O primeiro grande dramaturgo realista é o francês Alexandre Dumas Filho (1824-1895), autor da primeira peça realista, *A Dama das Camélias* (1852), que trata da prostituição.

Fora da França, um dos expoentes é o norueguês Henrik Ibsen (1828-1906). Em *Casa de Bonecas*, por exemplo, trata da situação social da mulher. São importantes também o

dramaturgo e escritor russo Gorki (1868-1936), autor de Ralé e Os Pequenos Burgueses, e o alemão Gerhart Hauptmann (1862-1946), autor de Os Tecelões.

REALISMO NO BRASIL - No Brasil, o realismo marca mais intensamente a literatura e o teatro.

Artes plásticas -Entre os artistas brasileiros, tem maior expressão o realismo burguês, nascido na França. Em vez de trabalhadores, o que se vê nas telas é o cotidiano da burguesia. Dos seguidores dessa linha se destacam Belmiro de Almeida (1858-1935), autor de Arrufos, que retrata a discussão de um casal, e Almeida Júnior (1850-1899), autor de O Descanso do Modelo. Mais tarde, Almeida Júnior aproxima-se de um realismo mais comprometido com as classes populares, como em Caipira Picando Fumo.

Literatura -O realismo manifesta-se na prosa. A poesia da época vive o parnasianismo. O romance é a principal forma de expressão, tornando-se veículo de crítica a instituições e à hipocrisia burguesa. A escravidão, os preconceitos raciais e a sexualidade são os principais temas, tratados com linguagem clara e direta.

O realismo atrai vários escritores, alguns antes ligados ao romantismo. O marco é a publicação de Memórias Póstumas de Brás Cubas, de Machado de Assis, que faz uma análise crítica da sociedade da época. Ligados ao regionalismo destacam-se Manoel de Oliveira Paiva (1861-1892), autor de Dona Guidinha do Poço, e Domingos Olímpio (1860-1906), de Luzia-Homem.

Teatro - Os problemas do cotidiano ocupam os palcos. O herói romântico é substituído por personagens do dia-a-dia e a linguagem passa a ser coloquial.

Entre os principais autores estão romancistas realistas, como Machado de Assis, que escreve Quase Ministro, e alguns românticos, como José de Alencar, com O Demônio Familiar, e Joaquim Manuel de Macedo (1820-1882), com Luxo e Vaidade. Outros nomes de peso são Artur de Azevedo (1855-1908), criador de comédias e operetas como A Capital Federal e O Dote, Quintino Bocaiúva (1836-1912) e França Júnior (1838-1890).

ANEXO E

**Universidade Federal do Rio Grande do Sul
Centro Interdisciplinar de Novas Tecnologias na Educação
Programa de Pós-Graduação em Educação e Informática na Educação
(PPGEDU/PPGIE)**

TERMO DE CONSENTIMENTO INFORMADO

A pesquisadora Miriam Noering Klemann, aluna do curso de Pós-Graduação em Educação e Informática na Educação – nível mestrado - desta Universidade, sob orientação do professor Dr. Eliseo Berni Reategui que realizará formação e investigação através de pesquisa realizada no mês de novembro de 2010. O objetivo desta pesquisa é investigar a utilização de uma ferramenta de mineração de texto (SOBEK) como apoio à produção textual.

Os alunos que aceitarem participar desta pesquisa precisam assinar este consentimento, autorizando a coleta de dados a partir do experimento assim como em outros instrumentos, inclusive entrevistas e questionários.

Os dados desta pesquisa estarão sempre sob sigilo ético, não sendo mencionados os nomes dos participantes em nenhuma apresentação oral ou trabalho escrito que venha a ser publicado. A participação não oferece risco ou prejuízo ao participante. Se no decorrer da pesquisa o participante resolver não mais continuar terá toda a liberdade de fazê-lo, sem que isso lhe acarrete qualquer prejuízo ou constrangimento.

A pesquisadora compromete-se a esclarecer quaisquer dúvidas ou questionamentos que eventualmente os participantes venham a ter no momento da pesquisa ou posteriormente através do email das pesquisadoras.

Após ter sido devidamente informado/a de todos os aspectos desta pesquisa e ter esclarecido todas as minhas dúvidas:

EU _____,

RG sob nº _____, concordo em participar desta pesquisa.

Assinatura do/da participante

Assinatura do (a) responsável

Assinatura da pesquisadora

Miriam Noering Klemann miriamklemann@gmail.com

Rio Pardo, ____ de _____ de 2010