Análise automática de aspectos relacionados à coerência semântica em resumos acadêmicos

Vinícius Mourão Alves de Souza[1], **Valéria Delisandra Feltrim**[1]

[1]Departamento de Informática – Universidade Estadual de Maringá (UEM)
Av. Colombo, 5.790 – 87020-900 – Maringá – PR – Brazil

{vsouza, valeria.feltrim}@din.uem.br

Abstract. *In this paper we present classifiers responsible for automate the analysis of semantic coherence aspects in academic abstracts. These aspects are based on the schematic structure of the Abstract section and on the semantic similarity among different components that compose the structure. The classifiers were trained and induced by machine learning algorithms based on features automatically extracted from the surface of the text and from the processing of the LSA. Results indicate that all classifiers achieved superior performance compared to baseline measures. Thus, they can be used in environments of aid to writing for emission of suggestions related to coherence.*

Resumo. *Neste artigo são propostos classificadores responsáveis por automatizar a análise de aspectos relacionados à coerência semântica em resumos acadêmicos. Tais aspectos são baseados na estrutura esquemática da seção Resumo e na similaridade semântica entre os componentes que compõem tal estrutura. Os classificadores foram treinados e induzidos por algoritmos de aprendizado de máquina, com base em características extraídas automaticamente da superfície do texto e provenientes do processamento da LSA. Os resultados indicam que todos os classificadores alcançaram desempenho superior às medidas de comparação e que podem ser utilizados, por exemplo, em ambientes de auxílio à escrita para emissão de sugestões relacionadas à coerência.*

1. Introdução

O resumo pode ser considerado uma das seções mais importantes de um trabalho acadêmico, dado que, em conjunto com o título, é utilizado pela comunidade científica como primeiro meio de divulgação de suas pesquisas. Assim como os trabalhos acadêmicos possuem uma estrutura bem definida, em geral enunciada como Introdução – Desenvolvimento – Conclusão, a seção destinada ao resumo também possui um esquema estrutural bem definido e passível de ser modelado. Vários modelos estruturais para resumos têm sido descritos na literatura [Swales 1990][Weissberg e Buker 1990][Aluísio e Oliveira Jr 1996]. Feltrim et al. (2003) propuseram um modelo estrutural específico para resumos de dissertações e teses em Ciência da Computação composto por seis componentes esquemáticos dispostos na seguinte ordem: Contexto, Lacuna, Propósito, Metodologia, Resultado e Conclusão.

A partir desse modelo estrutural e da análise de diferentes aspectos de coerência entre os componentes esquemáticos realizada por Souza e Feltrim (2011), este trabalho

Proceedings of the 8th Brazilian Symposium in Information and Human Language Technology, pages 77–86,
Cuiabá, MT, Brazil, October 24–26, 2011.

propõe a automatização da análise de coerência por meio do desenvolvimento de classificadores baseados em um conjunto de características extraídas automaticamente do texto.

Souza e Feltrim (2011) assumem, de acordo com a definição de Koch e Travaglia (2003) e van Dijk (1983), que a coerência diz respeito a possibilidade de se estabelecer um sentido lógico entre diferentes sentenças de um texto. Desse modo, os autores analisaram a coerência de resumos acadêmicos sob quatro aspectos, denominados de dimensões: (i) Dimensão Título, (ii) Dimensão Propósito, (iii) Dimensão Lacuna-Contexto e (iv) Dimensão Quebra de Linearidade. Essas dimensões consideram o relacionamento semântico entre diferentes componentes do resumo.

A Seção 2 apresenta a anotação e análise do córpus, bem como as dimensões propostas. As características extraídas das sentenças para a indução dos classificadores responsáveis pela análise automática de coerência são apresentadas na Seção 3. A avaliação intrínseca dos classificadores é apresentada na Seção 4 e, por fim, as conclusões são apresentadas na Seção 5.

2. Córpus e Anotação

Com o objetivo de analisar possíveis problemas de coerência que ocorrem em textos acadêmicos escritos em português, foram coletados 385 resumos de monografias de conclusão de curso de Ciência da Computação. Os resumos foram coletados por meio de acesso a bibliotecas digitais públicas e a anotação desse córpus foi dividida em duas partes: (i) anotação dos componentes esquemáticos e (ii) anotação de aspectos de coerência relacionados as dimensões propostas, como descrito a seguir.

2.1. Anotação e Análise da Estrutura Esquemática

A primeira fase consistiu na inserção de marcações no texto que identificam o título do resumo, início e fim de cada sentença, e a sua classificação em um dos componentes esquemáticos previstos no modelo estrutural de Feltrim et al. (2003). Para a anotação dos componentes esquemáticos foi utilizado o AZPort [Feltrim et al. 2006], um classificador *Naive Bayes* que atribui a cada sentença do córpus uma de seis possíveis categorias retóricas: Contexto, Lacuna, Propósito, Metodologia, Resultado e Conclusão. No total, 2.293 sentenças foram automaticamente anotadas e revisadas manualmente para que os erros cometidos pelo AZPort não inteferissem no processo de anotação e análise dos aspectos de coerência. A distribuição das categorias retóricas no córpus pode ser observada na Tabela 1.

Tabela 1. Distribuição das categorias retóricas no córpus

Categoria	Sentenças (N)	Distribuição (%)
Contexto	808	35,23
Lacuna	215	09,38
Propósito	426	18,58
Metodologia	273	11,90
Resultado	451	19,67
Conclusão	120	05,24
Total	2.293	100

2.2. Anotação e Análise de Coerência

Com base na abordagem de Higgins et al. (2004), foram identificadas e anotadas diferentes relações de coerência entre categorias retóricas específicas. Tais relações são baseadas na similaridade semântica entre sentenças de diferentes categorias retóricas e foram chamadas dimensões. As seguintes dimensões, originalmente propostas em Souza e Feltrim (2011), foram analisadas: Título, Propósito, Lacuna-Contexto e Quebra de Linearidade. A análise dos resultados e as conclusões obtidas são apresentadas a seguir.

2.2.1. Dimensão Título

Durante o processo de anotação, verificou-se nos resumos a similaridade semântica de cada sentença com o título do trabalho. Se a sentença é fortemente relacionada com o título, atribuiu-se o valor *alto*. Caso contrário, atribuiu-se o valor *baixo*. Devido a subjetividade da tarefa, foi utilizada uma escala binária para a anotação.

De um total de 2.293 sentenças, foram identificadas 1.243 (54,20%) com alto relacionamento com o título e 1.050 (46,80%) com baixo relacionamento. As sentenças classificadas como Propósito são as mais relacionadas com o título, uma vez que 83,33% dessas sentenças foram classificados como tendo um *alto* relacionamento. Esse valor é considerado elevado se comparado à média de 48,79% do valor *alto* para as outras categorias. A distribuição de acordo com as seis possíveis categorias esquemáticas é apresentada na Tabela 2.

Tabela 2. Relacionamento semântico das sentenças do córpus com o Título

	Sentenças (N)	
Categorias	Alto	Baixo
Contexto	364	444
Lacuna	104	111
Propósito	355	071
Metodologia	139	134
Resultado	220	231
Conclusão	061	059
Total	1.243	1.050

De fato, o título de um texto acadêmico deve apresentar os principais tópicos tratados no trabalho e, do mesmo modo, o resumo deve informar ao leitor sobre esses tópicos, ainda que de forma resumida. Portanto, a falta de relação entre o componente Propósito e o título, pode ser uma evidência de duas situações: (i) o título não é apropriado para o resumo ou (ii) o resumo possui problemas de coerência.

2.2.2. Dimensão Propósito

Para cada resumo do córpus, verificou-se a similaridade entre as sentenças classificadas como Propósito com as demais sentenças do resumo. Se a sentença é fortemente relacionada com o componente Propósito, atribuiu-se o valor *alto*. Caso contrário, atribuiu-se o valor *baixo*. Nos casos em que o resumo não possui sentenças da categoria Propósito ou que a sentença analisada seja classificada como Propósito, atribuiu-se o valor *n/a*.

Excluindo-se um conjunto de 573 sentenças anotadas com o valor *n/a* (426 sentenças de Propósito e 147 sentenças de diferentes categorias e que constituem resumos sem o componente Propósito) tem-se um total de 1.720 sentenças. Desse total, 59,07% possuem alto relacionamento e 40,93% possuem baixo relacionamento com o Propósito. As categorias mais relacionadas com o componente Propósito são as sentenças classificadas como Conclusão, Metodologia e Resultado. A distribuição de acordo com as categorias esquemáticas é apresentada na Tabela 3.

Tabela 3. Relacionamento semântico das sentenças do córpus com o componente Propósito

	Sentenças (N)		
Categorias	Alto	Baixo	N/A
Contexto	378	380	50
Lacuna	129	079	7
Propósito	—	—	426
Metodologia	171	082	20
Resultado	264	135	52
Conclusão	074	028	18
Total	1.016	704	573

Segundo Higgins et al. (2004), o relacionamento entre diferentes componentes da estrutura retórica de um texto determina a sua coerência global. Desse modo, um resumo será difícil de ler e compreender caso determinadas partes não sejam relacionadas. Sendo assim, a partir da análise realizada na Dimensão Propósito, espera-se que o componente Propósito seja relacionado com os componentes Metodologia, Resultado e Conclusão. Caso contrário, pode ser a indicação de um problema de coerência global.

2.2.3. Dimensão Lacuna-Contexto

Nas análises realizadas para a Dimensão Propósito, notou-se que o componente Contexto é, em geral, mais relacionado com o componente Lacuna do que com o Propósito. Assim, espera-se em resumos coerentes que o componente Lacuna seja relacionado com ao menos uma sentença classificada como Contexto. Desse modo, Souza e Feltrim (2011) assumem que a ausência de relação entre esses componentes seja indicação de um problema de coerência.

Para cada resumo do córpus que possua sentenças classificadas como Lacuna e Contexto, verificou-se o relacionamento semântico entre essas categorias. Cada sentença de Lacuna foi anotada com o valor *sim* caso seja relacionada com alguma sentença de Contexto. Caso contrário, atribuiu-se o valor *não*.

Excluindo-se um conjunto de 32 sentenças classificadas como Lacuna e pertencentes a resumos que não possuem sentenças da categoria Contexto, 183 sentenças foram anotadas nessa dimensão. Do total de sentenças, 74,86% foram anotadas com o valor *sim* e 23,14% foram anotadas com o valor *não*. Desse modo, levando em consideração os resultados da anotação dessa dimensão e a análise dos relacionamentos, concluiu-se que a Dimensão Lacuna-Contexto pode indicar possíveis problemas de coerência envolvendo os componentes Lacuna e Contexto.

2.2.4. Dimensão Quebra de Linearidade

Na análise dessa dimensão foi verificada a existência de quebra de linearidade entre sentenças adjacentes do resumo. São possíveis dois valores: *sim*, caso exista alguma dificuldade em se estabelecer um sentido lógico da sentença atual com a sentença anterior e com a próxima sentença; ou *não*, caso a sentença esteja de acordo com o fluxo do texto e o anotador não teve dificuldades em relacioná-la com as sentenças adjacentes.

Do total de 2.293 sentenças analisadas, 153 sentenças foram identificadas com o valor *sim* e 2.140 sentenças com o valor *não*. Das 153 sentenças em que o problema foi identificado, a categoria em que o problema é mais recorrente em termos de proporção é a Resultado, com 41 (9,09%) das sentenças com o valor *sim*. Já a categoria Lacuna é a que menos apresenta problemas, com apenas 7 (3,26%) ocorrências. Esses resultados mostram que no córpus analisado dificilmente existe a ocorrência de uma quebra significativa de linearidade entre sentenças adjacentes, já que os casos identificados correspondem a apenas 7,14% do total de sentenças. Além disso, em boa parte desses casos percebeu-se que tal sentença é justificada em alguma parte do texto, mesmo não sendo uma sentença adjacente. Essa característica dificulta ainda mais o processo de anotação e análise dessa dimensão.

3. Análise Automática de Coerência

O objetivo deste trabalho é desenvolver classificadores capazes de identificar automaticamente possíveis problemas de coerência baseados nas dimensões propostas por Souza e Feltrim (2011), conforme apresentado anteriormente. Foram desenvolvidos e avaliados classificadores para cada uma das dimensões, com exceção da Dimensão Quebra de Linearidade, em que o baixo número de exemplos identificados no processo de anotação impossibilitou o desenvolvimento de um classificador automático para desempenhar a tarefa. Descrevemos nesta seção a etapa de desenvolvimento.

3.1. Desenvolvimento

Para realizar a análise automática das dimensões, desenvolvemos classificadores induzidos por algoritmos de aprendizado de máquina com base em características extraídas da superfície do texto e provenientes do processamento da LSA – Latent Semantic Analysis [Landauer et al. 1998], um método estatístico para a extração e representação de conhecimento aplicado em córpus. A ideia básica da LSA é formar um espaço semântico em que a semelhança entre os termos se dá pela ocorrência em contextos comuns. Pode-se medir a similaridade de conceitos relacionados entre duas palavras ou sentenças, calculando-se o produto co-seno entre os vetores que representam tais palavras ou sentenças. O valor de tal similaridade é limitado entre $[-1, 1]$, sendo -1 o menor valor possível de similaridade e 1 o mais alto grau de similaridade.

Para a indução dos classificadores optou-se pelo algoritmo SMO [Keerthi et al. 2001]. Tal algoritmo é uma implementação da técnica SVM – *Support Vector Machine* [Vapnik 2000], um método de aprendizagem baseado em teorias estatísticas em que cria-se um hiperplano ótimo para a separação das classes. A técnica tem sido utilizada em aplicações de reconhecimento de padrões, tais como categorização de textos [Aizawa 2001] e categorização de *spam* [Drucker et al. 1999]. Também foi

utilizada por Higgins et al. (2004) na tarefa de análise de coerência. Desse modo, acreditamos que o algoritmo SMO seja adequado para a tarefa.

3.2. Extração de Atributos

Extraiu-se um conjunto de 13 atributos de cada sentença do córpus. Todos os atributos foram extraídos de maneira automática para a indução dos classificadores. O conjunto completo de atributos é apresentado a seguir:

1. Categoria retórica da sentença atual;
2. Categoria retórica da sentença anterior;
3. Categoria retórica da próxima sentença;
4. Presença de palavra que pode caracterizar uma anáfora;
5. Posição da sentença no resumo (com base no início do texto);
6. Presença de palavra que pode caracterizar algum tipo de transição;
7. Tamanho da sentença atual (em número de palavras);
8. Tamanho do título do resumo (em número de palavras);
9. Similaridade semântica (LSA) entre a sentença atual e a sentença anterior;
10. Similaridade semântica (LSA) entre a sentença atual e a próxima sentença;
11. Similaridade semântica (LSA) entre a sentença atual e o título do resumo;
12. Similaridade semântica (LSA) entre a sentença atual e a sentença de Propósito; e
13. Valor máximo de similaridade semântica (LSA) entre a sentença de Lacuna e alguma sentença de Contexto.

Os atributos de 1 a 8 são baseados na estrutura retórica e de características superficiais do resumo, enquanto os atributos de 9 a 13 são baseados no processamento da LSA. Os atributos de 1 a 10 são utilizados por todos os classificadores. O atributo 11 é adicionado ao conjunto de atributos do classificador Dimensão Título. O atributo 12 é adicionado ao conjunto de atributos do classificador Dimensão Propósito. Por fim, o atributo 13 é adicionado ao conjunto de atributos do classificador Dimensão Lacuna-Contexto.

4. Avaliação Intrínseca dos Classificadores

A partir da anotação apresentada na Subseção 2.2 e do conjunto de atributos extraídos das sentenças do córpus, foram gerados e avaliados cinco classificadores para as dimensões propostas, sendo um classificador para a Dimensão Título, três classificadores para a Dimensão Propósito, relativos aos componentes Metodologia (M), Resultado (R) e Conclusão (C) e, por fim, um classificador para a Dimensão Lacuna-Contexto.

Foi realizada a etapa de seleção de atributos para cada um dos classificadores utilizando o método *Wrapper* em conjunto com o método de busca *Best-First*. Desse modo, o algoritmo SMO com o *kernel PolyKernel* foi utilizado tanto para a seleção de atributos, quanto para a indução dos classificadores. Utilizou-se o ambiente Weka [Witten e Frank 2005] para as etapas de seleção de atributos, treinamento e teste dos indutores e na avaliação intrínseca. Para o treinamento dos classificadores foi utilizado o método de amostragem *10-fold stratified cross-validation* e alterado o parâmetro *filterType* do algoritmo SMO com o valor "*Standardize training data*", responsável por normalizar os atributos numéricos de maneira que sua média seja zero e o intervalo de variância seja unitária. Na Tabela 4 é possível observar a concordância *Kappa* entre a

anotação humana e cada um dos classificadores treinados com a configuração apresentada anteriormente e com os atributos mais preditivos de acordo com a etapa de seleção de atributos. A Tabela 4 também apresenta a taxa de acerto (ou acurácia) dos classificadores.

Tabela 4. Concordância *Kappa* da anotação humana *vs.* classificadores e taxa de acerto dos classificadores.

	Concordância *Kappa* (K)	Taxa de acerto (%)
Dim. Título	0,871	96,48
Dim. Propósito (M)	0,683	86,17
Dim. Propósito (R)	0,763	89,47
Dim. Propósito (C)	0,748	90,19
Dim. Lacuna-Contexto	0,679	88,52

Pode-se notar na Tabela 4 que todos os classificadores apresentam níveis satisfatórios de concordância *Kappa*, considerando-se a tarefa subjetiva das dimensões. O classificador da Dimensão Título obteve o melhor desempenho, com $K = 0,871$. Mesmo o classificador da Dimensão Lacuna-Contexto, que apresentou o pior resultado entre os classificadores, alcançou um bom nível de concordância com $K = 0,679$. Na Tabela 4 também é possível notar que todos os classificadores alcançaram altas taxas de acerto, entre 86,17% e 96,48%. Entretanto, essa é uma medida que não leva em consideração as classes preditas e, por isso, uma análise mais detalhada dos classificadores é apresentada na Tabela 5, que mostra o desempenho dos classificadores em relação às medidas de avaliação *Precision, Recall, F-Measure* e *Macro-F*. Para efeito de comparação, apresentamos na Tabela 6 os resultados de uma *baseline* simples que sempre atribui a classe prevalente como saída para cada um dos classificadores. Em ambas as tabelas, as classes *Alto* e *Baixo* são referentes aos classificadores da Dimensão Título, Dimensão Propósito (Metodologia), Dimensão Propósito (Resultado) e Dimensão Propósito (Conclusão), enquanto as classes *Sim* e *Não* são referentes ao classificador da Dimensão Lacuna-Contexto.

Tabela 5. Desempenho dos classificadores

	Alto / Sim			Baixo / Não			
	Precision	Recall	F-measure	Precision	Recall	F-measure	Macro-F
Dim. Título	0,975	0,983	0,979	0,912	0,873	0,892	0,936
Dim. Propósito (M)	0,895	0,901	0,898	0,79	0,78	0,785	0,842
Dim. Propósito (R)	0,914	0,928	0,921	0,855	0,83	0,842	0,882
Dim. Propósito (C)	0,921	0,946	0,933	0,846	0,786	0,815	0,874
Dim. Lacuna-Contexto	0,903	0,949	0,925	0,821	0,696	0,753	0,839

Observa-se nas Tabelas 5 e 6, que todos os classificadores superam os valores da *baseline*, em especial o classificador da Dimensão Título, que também apresentou os melhores resultados nas medidas relatadas anteriormente. Nota-se que o valor da medida *F-measure* das classes *alto/sim* é superior ao das classes *baixo/não*. Embora exista um desbalanceamento que possa favorecer as classes *alto/sim*, acredita-se que o comportamento dos classificadores se deve ao fato de ser mais fácil afirmar a existência de um relacionamento alto entre componentes do que um relacionamento baixo. Logo, existe

Tabela 6. *Baseline* dos classificadores

	Alto / Sim			Baixo / Não			
	Precision	Recall	F-measure	Precision	Recall	F-measure	Macro-F
Dim. Título	0,833	1,000	0,908	0,000	0,000	0,000	0,454
Dim. Propósito (M)	0,675	1,000	0,795	0,000	0,000	0,000	0,398
Dim. Propósito (R)	0,661	1,000	0,840	0,000	0,000	0,000	0,420
Dim. Propósito (C)	0,725	1,000	0,840	0,000	0,000	0,000	0,420
Dim. Lacuna-Contexto	0,748	1,000	0,855	0,000	0,000	0,000	0,428

uma ambiguidade maior para identificação de um relacionamento baixo. Entretanto, os classificadores obtiveram um bom desempenho de modo geral, com valores entre $0,839$ e $0,936$ para a medida *Macro-F*, que leva em consideração ambas as classes.

Desse modo, os resultados positivos obtidos na avaliação dos classificadores possibilitam a sua utilização para desempenhar as tarefas propostas pelas dimensões de Souza e Feltrim (2011) de maneira automática.

5. Conclusões

Este trabalho teve como objetivo desenvolver classificadores para a detecção automática de problemas relacionados à coerência semântica em resumos acadêmicos escritos em português a partir da proposta de Souza e Feltrim (2011) de quatro diferentes dimensões. Apresentamos os resultados de cinco classificadores: Dimensão Título, Dimensão Propósito (Metodologia), Dimensão Propósito (Resultado), Dimensão Propósito (Conclusão) e Dimensão Lacuna-Contexto. Todos os classificadores alcançaram desempenho superior às medidas de comparação utilizadas na avaliação intrínseca.

Os resultados positivos permitem a utilização dos classificadores para a emissão de sugestões à usuários de ambientes de auxílio à escrita, como o ambiente SciPo – *Scientific Portuguese* [Feltrim et al. 2006]. Desse modo, os classificadores deverão ser integrados e avaliados como parte do SciPo em um contexto de uso real, para a emissão de sugestões relacionadas à coerência semântica.

Outra questão a ser abordada em trabalhos futuros é a adequação dos classificadores para outras seções de trabalhos acadêmicos, tais como Introdução ou Conclusão. Além disso, devido as dificuldades encontradas para a construção do classificador da Dimensão Quebra de Linearidade, pretende-se, como extensão deste trabalho, usar a Teoria de Centering [Grosz et al. 1995] ou o Modelo de Entidades [Barzilay e Lapata 2008], para a análise do córpus e extração de novos atributos que possam colaborar para a construção de um novo classificador com o papel de identificar a existência de quebra de sentido lógico em sentenças adjacentes. Trabalhos como o de Miltsakaki e Kukich (2004) e Burstein et al. (2010) aplicaram tais modelos em tarefas semalhantes em redações escritas em inglês e alcançaram resultados positivos.

Sendo assim, acreditamos que este trabalho pode beneficiar diretamente ambiente de auxílio à escrita SciPo e também outras ferramentas de Processamento de Linguagem Natural, como sumarizadores automáticos.

Referências

A. Aizawa (2001). Linguistic techniques to improve the performance of automatic text categorization. Em *Proceedings of NLPRS-01, 6th Natural Language Processing Pacific Rim Symposium*, páginas 307–314, Tokyo, Japan.

S. M. Aluísio e O. N. Oliveira Jr (1996). A detailed schematic structure of research paper introductions: An application in support-writing tools. *Procesamiento del Lenguaje Natural*, 19:141–147.

R. Barzilay e M. Lapata (2008). Modeling local coherence: An entity-based approach. *Computational Linguistics*, 34(1):1–34.

J. Burstein, J. Tetreault e S. Andreyev (2010). Using entity-based features to model coherence in student essays. Em *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, páginas 681–684, Stroudsburg, PA, USA. Association for Computational Linguistics.

H. Drucker, D. Wu e V. N. Vapnik (1999). Support vector machines for spam categorization. *IEEE TRANSACTIONS ON NEURAL NETWORKS*, 10(5):1048–1054.

V. D. Feltrim, S. M. Aluísio e M. G. V. Nunes (2003). Analysis of the rhetorical structure of computer science abstracts in portugese. Em Dawn Archer, Paul Rayson, Andrew Wilson e Tony McEnery, editores, *Proceedings of Corpus Linguistics 2003*, volume 16, part 1, special issue de *UCREL Technical Papers*, páginas 212–218.

V. D. Feltrim, S. Teufel, M. G. V. Nunes e S. M. Aluísio (2006). Argumentative zoning applied to criquing novices scientific abstracts. Em James G. Shanahan, Yan Qu e Janyce Wiebe, editores, *Computing Attitude and Affect in Text: Theory and Applications*, páginas 233–246, Dordrecht, The Neherlands. Springer.

B. J. Grosz, S. Weinstein e A. K. Joshi (1995). Centering: a framework for modeling the local coherence of discourse. *Computational Linguistics*, 21(2):203–225.

D. Higgins, J. Burstein, D. Marcu e C. Gentile (2004). Evaluating multiple aspects of coherence in student essays. Em *Human Language Technologies: The 2004 Annual Conference of the North American Chapter of the Association for Computational Linguistics*. Association for Computational Linguistics.

S. S. Keerthi, S. K. Shevade, C. Bhattacharyya e K. R. K. Murthy (2001). Improvements to platt's smo algorithm for svm classifier design. *Neural Computation*, 13(3):637–649.

I. V. Koch e L. C. Travaglia (2003). *A coerência textual*. Editora Contexto, São Paulo.

T. K. Landauer, P. W. Foltz e D. Laham (1998). Introduction to latent semantic analysis. *Discourse Processes*, 25:259–284.

E. Miltsakaki e K. Kukich (2004). Evaluation of text coherence for electronic essay scoring systems. *Natural Language Engineering*, 10(1):25–55.

V. M. A. Souza e V. D. Feltrim (2011). An analysis of textual coherence in academic abstracts written in portuguese (to be publised). Em *Proceedings of the Sixth Corpus Linguistics Conference: CL 2011*, Birmingham, UK.

J. M. Swales (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge applied linguistics series. Cambridge University Press, Cambridge, UK.

T. A. van Dijk (1983). Studies in the pragmatics of discourse. *American Anthropologist*, 85(1):190–192.

V. N. Vapnik (2000). *The nature of statistical learning theory*. Springer Verlag, New York, NY, USA.

R. Weissberg e S. Buker (1990). *Writing up Research: Experimental Research Report Writing for Students of English*. Prentice Hall Regents.

I. H. Witten e E. Frank (2005). *Data Mining: Practical machine learning tools and technique*. Morgan Kaufmann Publisher.