

UTILIZANDO O PROGRAMA *WORDSMITH TOOLS* NA PESQUISA SOBRE APRECIACÃO: UMA SUGESTÃO METODOLÓGICA*

Vander VIANA (Pontifícia Universidade Católica do Rio de Janeiro)

vander.viana@terra.com.br

RESUMO: Estudos baseados em *corpora* requerem o uso do computador. Contudo, quando a análise a ser realizada é de cunho semântico, nem sempre esta ferramenta é empregada. O artigo relata uma proposta metodológica acerca do emprego do programa *WordSmith Tools* (Scott, 1999) em uma pesquisa que utiliza o sistema de Atitude (Martin e Rose, 2003; White, 2000) como aporte teórico para análise de dados. São apresentados alguns resultados obtidos de forma a exemplificar esta proposta. Avalia-se também a metodologia, ressaltando seus pontos positivos e negativos. Os resultados indicam que a análise computacional pode ser utilizada para mapear tais categorias semânticas.

PALAVRAS-CHAVE: metodologia; *WordSmith Tools*; Apreciação; *corpus*; Atitude

ABSTRACT: *Corpus-based studies make use of computers. However, when the analysis to be carried out is of a semantic nature, such tool is not always used. The article consists of a methodological proposal concerning the use of WordSmith Tools (Scott, 1999) in an investigation which relies on the Attitude system (Martin and Rose, 2003; White, 2000) to analyze its data. Some results are presented in order to exemplify this proposal. The methodology is evaluated and its possible advantages and drawbacks are highlighted. The results indicate that computational analysis can be used to map these semantic categories.*

KEYWORDS: *methodology; WordSmith Tools; Appreciation; corpus; Attitude*

0) Introdução

A pesquisa lingüística baseada em *corpus* tem como ponto de partida a linguagem em uso ao invés de abstrações. Tal tipo de investigação utiliza o computador como ferramenta, facilitando o

* Agradeço à Sonia Zyngier pela leitura de uma primeira versão deste artigo e por suas contribuições críticas para o desenvolvimento do mesmo.

VIANA, V. "Utilizando o programa *Wordsmith Tools* na pesquisa sobre apreciação: uma sugestão metodológica". In: *Revista Intercâmbio*, Volume XVI. São Paulo: LAEL/PUC-SP, ISSN 1806-275X-, 2007.

trabalho do pesquisador, uma vez que tarefas laboriosas e repetitivas podem ser realizadas em curto espaço de tempo e com alta precisão. A velocidade e a confiabilidade deste processo permitem que o analista trabalhe com uma quantidade de dados cada vez maior a partir de uma nova perspectiva.

O presente artigo busca explicitar como o subsistema de Apreciação, parte integrante do sistema de Atitude (Martin e Rose, 2003; White, 2000), pode ser mapeado semi-automaticamente com o auxílio do programa computacional *WordSmith Tools* (Scott, 1999). A pergunta de pesquisa deste estudo caracteriza-se por ser essencialmente de cunho metodológico. Tenta-se verificar a possibilidade de utilização do referido programa de análise textual em uma pesquisa lingüística de base semântica a partir da análise de um *corpus* oral.

Na primeira parte deste artigo, discute-se brevemente o sistema de Atitude. Posteriormente, são reportados estudos empíricos anteriores que tenham adotado tal sistema como referencial teórico para investigar como universitários brasileiros se posicionam ideologicamente quanto ao objeto literário. Em um terceiro momento, é descrita a composição do *corpus* de pesquisa. A seguir, discorre-se de forma detalhada a utilização do programa *WordSmith Tools* (ibid) como ferramenta de análise neste estudo. Em uma quinta etapa, são apresentados os resultados obtidos com a metodologia ora proposta. Na penúltima seção, que antecede às conclusões, o procedimento metodológico é avaliado e seus pontos positivos e negativos são levantados.

1) O sistema de Atitude

O arcabouço teórico que embasa o presente trabalho é o sistema de Atitude, já discutido em várias obras como, por exemplo, em Rothery e Stenglin (2000), White (2000), e Martin e Rose (2003). Esse sistema, localizado dentro do sistema maior de Avaliação,

inclui aqueles significados pelos quais textos/falantes anexam um valor ou avaliação intersubjetivo a participantes e processos por referência a respostas emocionais ou sistemas de valores culturalmente determinados¹ (White, 2000:5).

Ele divide-se em três categorias, a saber, Afeto, Julgamento e Apreciação. Ao expressar suas opiniões a respeito de algo, o enunciador pode fazê-lo de três formas que diferem quanto ao ponto de partida.

¹ Tradução livre do seguinte fragmento: "[...] includes those meanings by which texts/speakers attach an intersubjective value or assessment to participants and processes by reference either to emotional responses or to systems of culturally-determined value systems."

VIANA, V. "Utilizando o programa *Wordsmith Tools* na pesquisa sobre apreciação: uma sugestão metodológica". In: *Revista Intercâmbio*, Volume XVI. São Paulo: LAEL/PUC-SP, ISSN 1806-275X-, 2007.

Pode partir de si próprio e de outras pessoas; de instituições socialmente reconhecidas; ou de coisas ou processos. Por conseguinte, ao posicionar-se ideologicamente a respeito de música, o falante pode dizer:

- (a) Eu adoro essa música;
- (b) A música brasileira é péssima;
- (c) Essa música é péssima.

Em (a), tem-se um exemplo de Afeto porque o ponto de partida é o próprio falante e o foco é o sentimento. Em (b), tem-se uma instância de Julgamento, cujo foco é a música brasileira enquanto instituição. O exemplo (c) é característico do subsistema de Apreciação, uma vez que o ponto de partida é um objeto, ou seja, uma música específica.

O subsistema de Apreciação, foco do presente trabalho, envolve a avaliação por parte do enunciador a respeito de um processo ou objeto. Este subsistema encontra-se dividido em outras três categorias mais específicas: Reação, Composição e Valor.

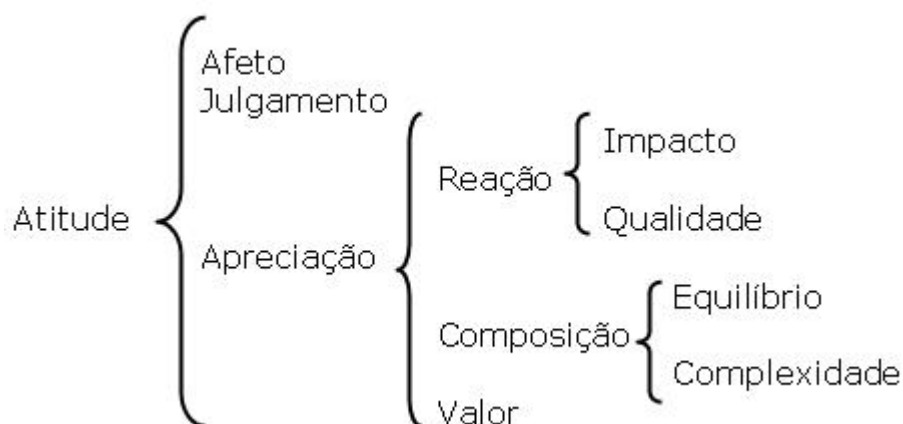


Figura 1: O sistema de Atitude e suas divisões²

Assim sendo, de acordo com o subsistema de Apreciação, quando se quer descrever o impacto emocional que um objeto ou processo tem em uma pessoa, utilizam-se marcadores que podem ser classificados como característicos da subcategoria de Reação: Impacto. Porém, quando o objetivo é avaliar a qualidade de algo, utilizam-se marcadores pertencentes à subcategoria de Reação: Qualidade. A categoria de Composição relaciona-se com a questão da consistência e estrutura. Pode-se descrever um objeto ou processo em termos de sua proporcionalidade (Composição: Equilíbrio) e de seus detalhes (Composição: Complexidade). A categoria de Valor corresponde à avaliação de um objeto ou processo de acordo com convenções sociais.

² Os subsistemas de Afeto e Julgamento também apresentam categorias e subcategorias. As mesmas, contudo, não foram incluídas na Figura 1 já que não fazem parte do escopo deste estudo.

VIANA, V. "Utilizando o programa *Wordsmith Tools* na pesquisa sobre apreciação: uma sugestão metodológica". In: *Revista Intercâmbio*, Volume XVI. São Paulo: LAEL/PUC-SP, ISSN 1806-275X-, 2007.

Todas as categorias e subcategorias podem ser expressas dentro de um contínuo que vai do positivo até o negativo (Rothery e Stenglin, 2000:238-239; Martin e Rose, 2003:63-64).

De forma a ilustrar as diferentes subcategorias analíticas, Martin e Rose (ibid) sugerem alguns adjetivos específicos em língua inglesa que marcam cada uma delas. A Figura 2 contém alguns desses indexadores traduzidos para o português.

Tipo de Apreciação		Positivo	Negativo
Reação	Impacto	notável	tedioso
	Qualidade	bonito	feio
Composição	Equilíbrio	simétrico	desproporcional
	Complexidade	simples	simplista
Valor		inspirador	banal

Figura 2: Exemplos de marcadores apreciativos em português (adaptado de Martin e Rose, ibid)

Segundo Menezes (2002),

um dos grandes problemas em Martin é que seu texto não se constrói a partir de dados empíricos, ou seja, a maioria dos dados são inventados e baseados em adjetivos que, dependendo do contexto, assumem significado[s] diferente[s] daquele[s] por Martin imaginado[s] (p.149)

Por exemplo, a tabela proposta por Martin e Rose (2003) indica que 'bonito' é um indexador da subcategoria Reação: Qualidade com valor positivo. Porém, os dados empíricos nos quais esta pesquisa se baseia mostram que este adjetivo pode ser um marcador de três subcategorias distintas e assumir dois valores diferentes.³

Enquanto Martin e Rose (ibid) sugerem adjetivos como marcadores, Taboada (2004) sugere a utilização da fórmula 'X é Y', em que X é o processo ou objeto a ser avaliado e Y é um adjetivo que marca o tipo de apreciação feita, como em 'O livro (X) é complexo (Y)'. Ao reconhecer esse padrão, torna-se possível a escolha de um ponto de entrada (X ou Y) para se investigar um dado *corpus* de pesquisa através do auxílio computacional. De qualquer forma, ressalta-se que a Apreciação não é unicamente marcada por meio de adjetivos, podendo ser veiculada através de outras classes gramaticais. O foco aqui, no entanto, recai sobre os primeiros.

³ Este assunto é tratado mais detalhadamente na Seção 4 na qual também são apresentados exemplos de forma a clarificar tal idéia.

2) Estudos anteriores: metodologias empregadas

Nessa seção, são revistos alguns estudos a respeito da atitude de universitários em relação à Literatura como área de estudo e/ou ao objeto literário, seguindo a abordagem teórica de Martin e Rose (2003) e de seus colaboradores no que se convencionou chamar de Escola de Sydney. Será dada especial importância aos procedimentos metodológicos adotados em cada uma das pesquisas relatadas de forma que eles possam ser contrastados com a presente metodologia.

Menezes (2002) analisou a resposta dada por 129 alunos de Letras da Universidade Federal do Rio de Janeiro a uma mesma pergunta: "Na sua opinião, os textos literários completam alguma coisa que os outros tipos de texto e a mídia não são capazes de completar?" A análise realizada pela pesquisadora não contou com nenhum auxílio computacional, tendo sido realizada manualmente.

Por sua vez, Barros (2003) estudou o posicionamento de 73 universitários de Letras em relação à literatura no contexto de um evento não-acadêmico, a saber, a Bienal do Livro do Rio de Janeiro. Em relação ao instrumento de pesquisa, o pesquisador optou pela utilização de um questionário. As respostas dadas à pergunta "Qual a sua visão de Literatura?" foram analisadas qualitativamente de forma não-automatizada.

O estudo de Mendes (2002) mapeou a linguagem de posicionamento empregada por alunos da disciplina Língua Inglesa III da Universidade Federal do Rio de Janeiro (UFRJ) em um *corpus* escrito. Nos ensaios que fizeram parte dos dados analisados pela pesquisadora, os alunos deveriam definir 'literatura' e justificar sua importância na área educacional. Em termos metodológicos, a pesquisadora utilizou-se da mesma ferramenta descrita neste trabalho: o programa computacional *WordSmith Tools* (Scott, 1999). Contudo, a análise restringiu-se a uma única palavra: 'literatura'.

Zyngier e Shepherd (2003) também investigaram um *corpus* escrito compilado a partir de 74 ensaios escritos em inglês e produzidos por alunos universitários de Letras no primeiro ano de estudo. O ponto de partida dado para a redação dos ensaios foi a pergunta "What is literature?"⁴. Apesar de afirmarem que as respostas foram digitalizadas e compiladas em um *corpus* de 19.160 palavras, as pesquisadoras não explicitam, na seção de metodologia, como o computador foi utilizado na análise de dados, comentando somente os procedimentos metodológicos utilizados para a sua coleta. Na conclusão do artigo, no entanto, tal questão é retomada e afirma-se que a análise foi realizada somente com as linhas de concordância da palavra '*literature*' (ou

⁴ A pergunta dada como comando motivador por Zyngier e Shepherd (2003) pode ser traduzida como "O que é literatura?" para a língua portuguesa.

VIANA, V. "Utilizando o programa *Wordsmith Tools* na pesquisa sobre apreciação: uma sugestão metodológica". In: *Revista Intercâmbio*, Volume XVI. São Paulo: LAEL/PUC-SP, ISSN 1806-275X-, 2007.

'literatura' em português). Não há, contudo, nenhuma etapa posterior apesar de Zyngier e Shepherd (ibidem) afirmarem que a etiquetagem realizada deva ser aplicada a cada linha de concordância isoladamente, o que não foi feito no estudo.

Mattos e Mendes (2003) discorrem a respeito de um projeto inicial de investigação do posicionamento de alunos de Letras perante o objeto literário. O artigo em questão relata o estudo piloto e a coleta de dados definitiva. O método de coleta de dados adotado pelas pesquisadoras foi o grupo de enfoque. Como o texto é um relato preliminar de uma investigação em curso à época da publicação do artigo, não se detalha a metodologia que seria utilizada para a análise de dados. Argumenta-se unicamente que, após a transcrição dos grupos de enfoque,

serão levantados os itens lexicais mais freqüentes no *corpus*, que fornecerão o ponto de partida para um mapeamento da linguagem dos jovens estudantes de Letras quando emitem opiniões. Para tanto, far-se-á uso da ferramenta *WordList* do programa computacional *WordSmith Tools* (Scott, 1986). Num segundo momento, será realizado um novo Grupo de Enfoque, também com alunos de Letras, que discutirá um tema típico de jovens, a fim de contrastar a linguagem utilizada para falar de poesia e a usada para falar de um assunto geral. A parte final da pesquisa visa à análise de cada uma das categorias de Valoração do sistema de Martin [...] presentes no *corpus* dos dados obtidos, também com o auxílio do referido programa computacional (Mattos e Mendes, 2003:116-117).

Parece, então, que o objetivo maior da pesquisa de Mattos e Mendes (op. cit.) não é de mapear o sistema de Atitude no *corpus* de pesquisa, mas contrastar o léxico presente em duas interações de naturezas distintas.

De forma resumida, a presente proposta metodológica difere-se claramente das empregadas por Menezes (2002) e Barros (2003), uma vez que ambos os estudos não lançaram mão da ferramenta computacional para a análise de dados. Apesar de Mendes (2002) e Zyngier e Shepherd (2003) utilizarem o programa *WordSmith Tools* (Scott, 1999), ambas as pesquisas não o fazem da mesma forma como relatado neste artigo. Além disto, as pesquisadoras analisam *corpora* escritos, que não apresentam a complexidade do trabalho com dados orais, nos quais há a ocorrência de pausas, pausas vocalizadas, auto-reparos e reparos, entre outros fenômenos. A pesquisa de Mattos e Mendes (2003) é a que mais se aproxima da presente proposta por se utilizar do mesmo *corpus* de pesquisa e por empregar a mesma ferramenta computacional; porém, nenhuma metodologia consistente para a análise de dados é sugerida. Tal proposta fugiria ao escopo do

VIANA, V. "Utilizando o programa *Wordsmith Tools* na pesquisa sobre apreciação: uma sugestão metodológica". In: *Revista Intercâmbio*, Volume XVI. São Paulo: LAEL/PUC-SP, ISSN 1806-275X-, 2007.

artigo, uma vez que as pesquisadoras focalizam a questão do uso do grupo de enfoque em uma pesquisa na área de Ciência Empírica da Literatura. Assim sendo, este artigo objetiva suprir esta lacuna.

3) O *corpus* de pesquisa⁵

Inicialmente, cabe definir-se *corpus*. Segundo Sanchez e Cantos (1996:8-9 apud Berber Sardinha, 2004), trata-se de:

Um conjunto de dados lingüísticos (pertencentes ao uso oral ou escrito da língua ou a ambos) sistematizados segundo determinados critérios, suficientemente extensos em amplitude e profundidade, de maneira que sejam representativos da totalidade do uso lingüístico ou de algum de seus âmbitos, dispostos de tal modo que possam ser processados por computador, com a finalidade de propiciar resultados vários e úteis para a descrição e análise (p. 18).

O *corpus*⁶ deste estudo é composto por textos autênticos produzidos por falantes de português como primeira língua. Ele consiste de interações em conversas e, por isso, atende ao propósito de investigação. Por fim, é representativo do uso lingüístico em questão e foi formatado a fim de que pudesse ser analisado pelo computador.

As interações são provenientes de três grupos de enfoque (cf. Barbour e Kitinger, 1999) com a participação, em cada um deles, de cinco ou seis alunos do primeiro período do curso de Letras de uma universidade pública no Rio de Janeiro e de um moderador diferente, mas também aluno de Letras. Como fator motivador para a discussão, os moderadores distribuíram quatro poemas⁷ previamente selecionados e pediram que os participantes os colocassem em ordem de preferência, justificando suas escolhas.

Após a realização dos grupos de enfoque, todas as três gravações foram transcritas e formatadas de maneira que pudessem ser analisadas com o computador.

⁵ O *corpus* de pesquisa é parte integrante do banco de dados do ramo brasileiro do Projeto REDES (*Research and Development in Empirical Studies*). Informações adicionais a respeito do referido grupo de pesquisa podem ser obtidas nos sites internacional (www.redes.lmu.de) e brasileiro (www.lettras.ufrj.br/redes).

⁶ Este mesmo *corpus* já foi utilizado por Mattos e Mendes (2003), Viana e Shepherd (2004) e Viana (2005b). A pesquisa de Viana (2005a) também se utilizou deste mesmo *corpus*, mas a análise não se ateve ao mesmo já que se objetivava contrastar a fala de alunos universitários de Letras em língua portuguesa (L1) e em língua inglesa (L2).

⁷ Os poemas distribuídos aos participantes foram "Soneto de Fidelidade" de Vinícius de Moraes, "Coisas que a vida leva" de Dulce de Melo Monte-Mór, "Retrato" de Cecília Meireles e um poema sem título de Fernando Pessoa.

VIANA, V. "Utilizando o programa *Wordsmith Tools* na pesquisa sobre apreciação: uma sugestão metodológica". In: *Revista Intercâmbio*, Volume XVI. São Paulo: LAEL/PUC-SP, ISSN 1806-275X-, 2007.

O *corpus* pode ser caracterizado pelos seguintes critérios (cf. proposta de Berber Sardinha, 2004:21-22; 26):

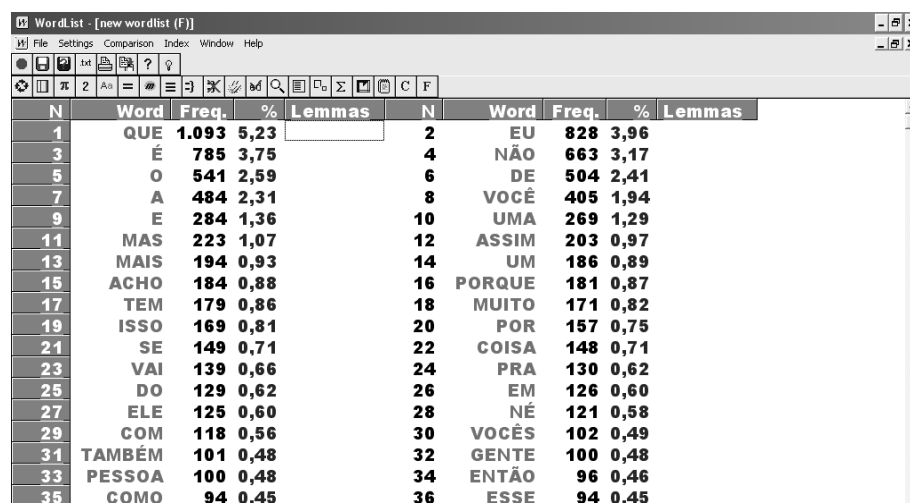
- (a) origem da autoria: todos os textos foram produzidos por falantes nativos em circunstâncias naturais;
- (b) meio: todos os textos são transcrições de interações face a face;
- (c) integralidade: nenhum texto sofreu corte ou fragmentação;
- (d) especificidade: o *corpus* é composto de textos específicos;
- (e) tamanho: o *corpus* é pequeno já que tem menos de 80 mil palavras.

Após a compilação do *corpus* de estudo, foi possível investigar o mesmo com o auxílio do programa computacional *WordSmith Tools* (Scott, 1999), que é abordado na próxima seção.

4) *WordSmith Tools*: sua utilização prática

O programa *WordSmith Tools* (ibid), que pode ser utilizado na análise de textos individuais ou de grandes *corpora*, possui três ferramentas analíticas, a saber, *WordList*, *Concord* e *KeyWords*. As ferramentas mais relevantes para a pesquisa em tela, no entanto, foram o *WordList* e o *Concord*.

A primeira ferramenta utilizada foi o *WordList*, que tem como principal função a criação de listas de palavras em ordem alfabética ou de frequência. A Figura 3 mostra as trinta e seis palavras mais freqüentes no *corpus* de pesquisa.



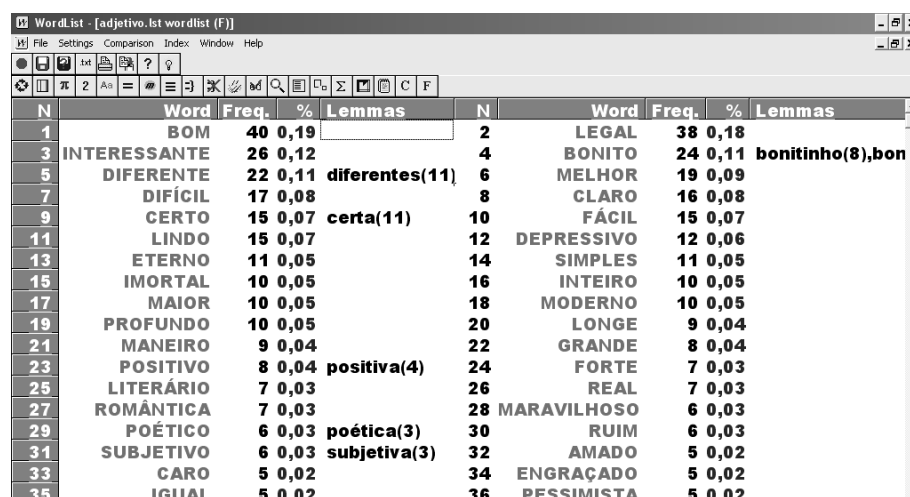
N	Word	Freq.	%	N	Word	Freq.	%
1	QUE	1.093	5,23	2	EU	828	3,96
3	É	785	3,75	4	NÃO	663	3,17
5	O	541	2,59	6	DE	504	2,41
7	A	484	2,31	8	VOCÊ	405	1,94
9	E	284	1,36	10	UMA	269	1,29
11	MAS	223	1,07	12	ASSIM	203	0,97
13	MAIS	194	0,93	14	UM	186	0,89
15	ACHO	184	0,88	16	PORQUE	181	0,87
17	TEM	179	0,86	18	MUITO	171	0,82
19	ISSO	169	0,81	20	POR	157	0,75
21	SE	149	0,71	22	COISA	148	0,71
23	VAI	139	0,66	24	PRA	130	0,62
25	DO	129	0,62	26	EM	126	0,60
27	ELE	125	0,60	28	NÉ	121	0,58
29	COM	118	0,56	30	VOCÊS	102	0,49
31	TAMBÉM	101	0,48	32	GENTE	100	0,48
33	PESSOA	100	0,48	34	ENTÃO	96	0,46
35	COMO	94	0,45	36	ESSE	94	0,45

Figura 3: Lista de palavras mais freqüentes

Todavia, as palavras mais freqüentes não revelam muito a respeito das avaliações feitas pelos sujeitos de pesquisa. Com o objetivo de buscar a linguagem avaliativa – enfocando particularmente o padrão 'X é Y', em que Y é um adjetivo – promoveu-se uma limpeza na lista de palavras de forma a excluir todas as palavras que não pertenciam à classe gramatical de adjetivo. Esta eliminação foi realizada

VIANA, V. "Utilizando o programa *Wordsmith Tools* na pesquisa sobre apreciação: uma sugestão metodológica". In: *Revista Intercâmbio*, Volume XVI. São Paulo: LAEL/PUC-SP, ISSN 1806-275X-, 2007.

manualmente visto que o *corpus* não havia sido etiquetado morfológicamente. Caso tal procedimento tivesse sido realizado em uma etapa anterior, a exclusão poderia ter sido feita de forma automática.⁸ A Figura 4 ilustra o início da lista de adjetivos mais freqüentes.



N	Word	Freq.	%	Lemmas
1	BOM	40	0,19	
3	INTERESSANTE	26	0,12	
5	DIFERENTE	22	0,11	diferentes(11)
7	DIFÍCIL	17	0,08	
9	CERTO	15	0,07	certa(11)
11	LINDO	15	0,07	
13	ETERNO	11	0,05	
15	IMORTAL	10	0,05	
17	MAIOR	10	0,05	
19	PROFUNDO	10	0,05	
21	MANEIRO	9	0,04	
23	POSITIVO	8	0,04	positiva(4)
25	LITERÁRIO	7	0,03	
27	ROMÂNTICA	7	0,03	
29	POÉTICO	6	0,03	poética(3)
31	SUBJETIVO	6	0,03	subjativa(3)
33	CARO	5	0,02	
35	IGUAL	5	0,02	

N	Word	Freq.	%	Lemmas
2	LEGAL	38	0,18	
4	BONITO	24	0,11	bonitinho(8),bon
6	MELHOR	19	0,09	
8	CLARO	16	0,08	
10	FÁCIL	15	0,07	
12	DEPRESSIVO	12	0,06	
14	SIMPLES	11	0,05	
16	INTEIRO	10	0,05	
18	MODERNO	10	0,05	
20	LONGE	9	0,04	
22	GRANDE	8	0,04	
24	FORTE	7	0,03	
26	REAL	7	0,03	
28	MARAVILHOSO	6	0,03	
30	RUIM	6	0,03	
32	AMADO	5	0,02	
34	ENGRAÇADO	5	0,02	
36	PESSIMISTA	5	0,02	

Figura 4: Lista de adjetivos mais freqüentes

Pode-se notar na Figura 4 que os adjetivos já se encontram lematizados⁹ na lista de freqüência. Em outras palavras, todas as instâncias de um mesmo adjetivo (masculino plural, feminino singular/plural, diminutivo e/ou aumentativo) foram reunidas¹⁰ na entrada correspondente ao masculino singular assim como ocorre em dicionários.¹¹ Dessa forma, ao se analisar o adjetivo que está na 9ª posição na lista de freqüência, ou seja, 'certo', percebe-se que há 15 instâncias de 'certo/a' no *corpus* de pesquisa. Como nota-se a existência de 11 instâncias de 'certa' no feminino, conclui-se que só há 4 instâncias de 'certo' no masculino.

Com relação ao adjetivo 'bom', o item mais freqüente na Figura 4 com 40 instâncias, foi necessário descartá-lo da análise pelo fato desse item lexical não aparecer na fórmula 'X é bom', mas por ser usado como pausa vocalizada.

⁸ Ressalta-se, no entanto, que o programa *WordSmith Tools* (Scott, 1999) não realiza tal etiquetagem.

⁹ A lematização é definida por Berber Sardinha (2004:89) como o "agrupamento de duas ou mais formas diferentes em um mesmo item".

¹⁰ Apesar de ser possível realizar a lematização de forma automática com o auxílio do programa computacional em questão, decidiu-se fazer a mesma de forma manual para que possíveis erros fossem evitados.

¹¹ Os únicos casos nos quais dicionários não lematizam palavras são quando elas constituem novos itens lexicais como ocorre, por exemplo, com 'anais' (no plural), 'calcinha' (no diminutivo) e 'portão' (no aumentativo). O mesmo critério foi adotado nesta pesquisa.

VIANA, V. "Utilizando o programa *Wordsmith Tools* na pesquisa sobre apreciação: uma sugestão metodológica". In: *Revista Intercâmbio*, Volume XVI. São Paulo: LAEL/PUC-SP, ISSN 1806-275X-, 2007.

<F5>¹² [...] traduz bem o pensamento do meu relacionamento com essa pessoa e bom... foi exatamente por isso, uma razão totalmente sentimental ahn e vai exatamente contra a proposta do Soneto de Fidelidade né até quando ele fala assim... Pode citar?

Optou-se, então, por trabalhar com os adjetivos mais freqüentes como, por exemplo, 'legal' (38 instâncias), 'interessante' (26 instâncias) e 'bonito' (24 instâncias) que aparecem em 2º, 3º e 4º lugares na lista de freqüência.

A segunda etapa desta pesquisa foi realizada com a ajuda do *Concord*. Essa ferramenta fornece listas de pequenos extratos textuais que precedem e sucedem uma palavra de busca, de forma a explicitar um pouco do cotexto no qual a mesma ocorre. O cotexto é entendido como o ambiente lexical no qual uma palavra é utilizada, ou seja, o seu contexto lingüístico. A título de ilustração, na Figura 5, há uma parte das trinta e oito linhas de concordância obtidas para a palavra 'legal'.¹³

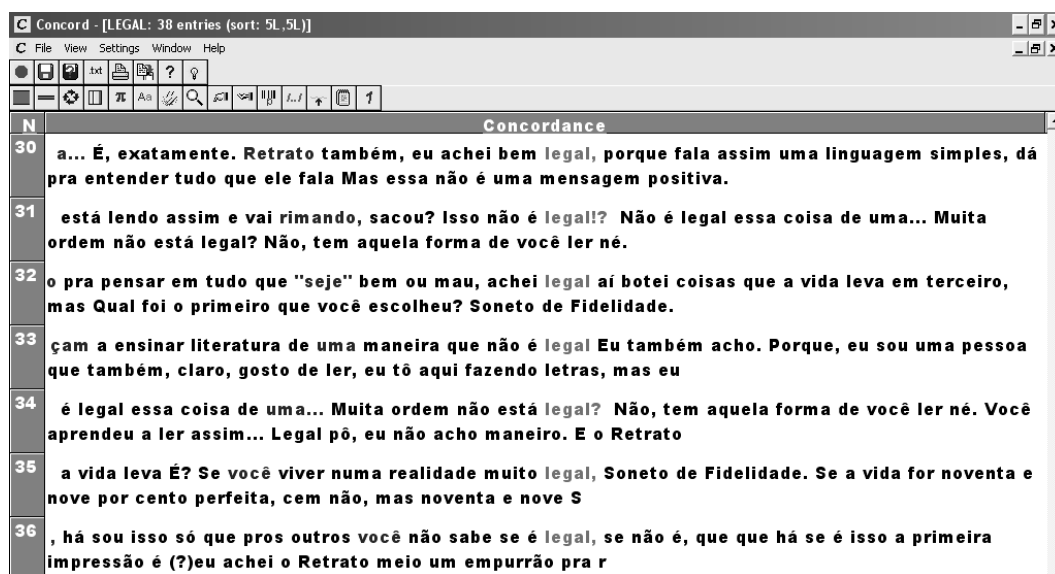


Figura 5: Linhas de concordâncias para a palavra 'legal'

Analisou-se cada linha de concordância individualmente de forma a poder realizar a respectiva classificação. Esta análise foi necessária tendo em vista que um mesmo adjetivo poderia ser indexador de

¹² As letras 'F' e 'M' designam o sexo do participante: feminino e masculino, respectivamente. O número cardinal empregado após a letra foi utilizado para diferenciar os participantes dos grupos de enfoque.

¹³ De todas as instâncias de 'legal' no *corpus*, apenas duas eram instâncias de julgamento como temos na linha de concordância de número 33 na Figura 5. Ao dizer que "a escola faz uma coisa estranha [...] porque eles começam a ensinar literatura de uma maneira que não é legal", a participante, na verdade, emite um julgamento a respeito da escola enquanto uma instituição e não de uma escola em particular.

VIANA, V. "Utilizando o programa *Wordsmith Tools* na pesquisa sobre apreciação: uma sugestão metodológica". In: *Revista Intercâmbio*, Volume XVI. São Paulo: LAEL/PUC-SP, ISSN 1806-275X-, 2007.

múltiplas subcategorias assim como sustentam Rothery e Stenglin (2000:240) e Menezes (2002:149-150).

O adjetivo 'bonito', por exemplo, foi utilizado pelos participantes da pesquisa como um indexador de três diferentes categorias e dois valores distintos. A participante F7 utiliza 'bonito' como um índice da subcategoria de Reação: Qualidade ao apreciar positivamente as imagens criadas quando da leitura de um dos poemas distribuídos no início da discussão.

<F7> eu achei bonitas assim as imagens que vêm à nossa cabeça quando a gente lê.

Já a participante F9 utiliza 'bonito' como um sinônimo para 'interessante'. Dessa forma, o que se tem é que o adjetivo em questão passa a ser um indexador da subcategoria de Reação: Impacto.

<F9> Meu irmão tem dez anos... ele gosta de ler. <F12> Ai, que bonitinho.

Ao analisar a fala da participante F8, pode-se identificar mais um uso diferente para o item lexical 'bonito', ou seja, o de indexador da categoria de Valor. Ao fazer um reparo em sua fala, a participante esclarece o que quis dizer com o adjetivo 'bonito', que, para ela, equivale a 'profundo' neste caso.

<F8> "que não seja imortal posto que é chama / mas que seja infinito enquanto dure" aliás tem escrito várias vezes no meu caderno [...] Porque eu gosto acho muito bonito, profundo.

Quanto ao valor da Apreciação, há também exemplos nos quais 'bonito' é utilizado para apreciar algo de forma negativa:

<F13> Eu acho que quem lê, cara, tem que lê e aquilo tem que chegar muito mais do que na sua cabeça, tem que te emocionar, eu acho que o objetivo do poeta é esse e não escrever uma coisa bonitinha pro cara da academia de letras falar: "é, realmente, isso aqui é bom, hein".

Em sua fala, a participante F13 utiliza a palavra 'bonitinha' para avaliar negativamente a escrita de autores que se mostram demasiadamente preocupados em agradar os membros da Academia Brasileira de Letras.

Tendo em vista a multiplicidade de significados e possibilidades analíticas para cada linha produzida pelo concordanciador, convencionou-se atribuir uma letra correspondente a cada um dos tipos

VIANA, V. "Utilizando o programa *Wordsmith Tools* na pesquisa sobre apreciação: uma sugestão metodológica". In: *Revista Intercâmbio*, Volume XVI. São Paulo: LAEL/PUC-SP, ISSN 1806-275X-, 2007.

de Apreciação existentes no *corpus*. A Figura 6 ilustra os códigos utilizados nesta pesquisa.

Categorias e subcategorias		Valores	Códigos
Reação	Impacto	Positivo	A
		Negativo	B
	Qualidade	Positiva	C
		Negativa	D
Composição	Equilíbrio	Positivo	E
		Negativo	F
	Complexidade	Positiva	G
		Negativa	H
Valor		Positivo	I
		Negativo	J
Não-classificado			K

Figura 6: Códigos adotados na presente pesquisa

Como é possível notar na Figura 6, subcategorias e valores diferentes correspondem a códigos distintos. Assim sendo, a mesma subcategoria de Reação: Impacto pode ser codificada como 'A' ou 'B' a depender do valor da avaliação feita.¹⁴ Criou-se também uma categoria denominada de 'não-classificado' a ser utilizada com as instâncias nas quais a classificação não foi possível porque faltavam elementos textuais para que o posicionamento do falante ficasse claro. Muitas destas instâncias, por exemplo, contêm a palavra 'legal', que é, por sua natureza, indefinida como registra o Dicionário Aurélio: "palavra-ônibus que exprime numerosas idéias apreciativas: ótimo, perfeito, excelente, leal, digno, etc." (Ferreira, 1999:1196).

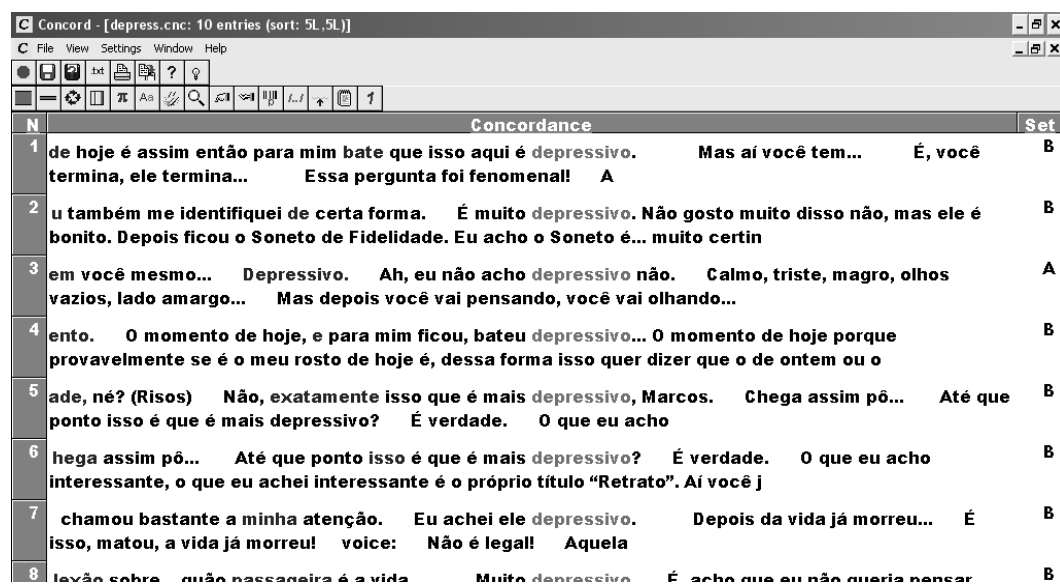
Após a criação dos códigos, atribuiu-se a cada linha de concordância uma letra de acordo com o tipo de Apreciação verbalizada. Na própria tela do *Concord*, o usuário pode selecionar qualquer linha de concordância e clicar em uma determinada letra. A letra digitada, que nesta pesquisa corresponde a um código conforme indicado na Figura 6, fica armazenada na coluna 'set' que sucede a coluna 'concordance'.¹⁵ A

¹⁴ Na versão 3 do programa *WordSmith Tools* (Scott, 1999), pode-se utilizar todo o alfabeto na marcação de linhas de concordância (passo ilustrado na Figura 7) sendo que ele pode ser feito através de letras maiúsculas e/ou minúsculas. Deve-se, contudo, atentar para o fato de que ao ordenar as linhas de concordância de acordo com a letra inserida (passo ilustrado na Figura 8), o programa primeiramente lista todas aquelas que contêm letras maiúsculas em ordem alfabética para só então apresentar as que contêm letras minúsculas, igualmente ordenadas alfabeticamente. Faz-se necessário lembrar que também é possível ordenar as linhas de concordância em ordem alfabética reversa.

¹⁵ De forma a desfazer uma classificação errada, o usuário deve selecionar a linha de concordância específica e clicar na letra correta. Se o objetivo é apagar a classificação anteriormente realizada, deve-se clicar em '0' (zero).

VIANA, V. "Utilizando o programa *Wordsmith Tools* na pesquisa sobre apreciação: uma sugestão metodológica". In: *Revista Intercâmbio*, Volume XVI. São Paulo: LAEL/PUC-SP, ISSN 1806-275X-, 2007.

Figura 7 ilustra uma tela do *Concord* na qual todas as linhas de concordância para o adjetivo 'depressivo' estão devidamente classificadas.



The screenshot shows the Concord software window titled "Concord - [depress.cnc: 10 entries (sort: 5L,5L)]". The interface includes a menu bar (File, View, Settings, Window, Help) and a toolbar with various icons. The main window displays a concordance table with the following columns: N, Concordance, and Set.

N	Concordance	Set
1	de hoje é assim então para mim bate que isso aqui é depressivo. Mas aí você tem... É, você termina, ele termina... Essa pergunta foi fenomenal! A	B
2	u também me identifiquei de certa forma. É muito depressivo. Não gosto muito disso não, mas ele é bonito. Depois ficou o Soneto de Fidelidade. Eu acho o Soneto é... muito certin	B
3	em você mesmo... Depressivo. Ah, eu não acho depressivo não. Calmo, triste, magro, olhos vazios, lado amargo... Mas depois você vai pensando, você vai olhando...	A
4	ento. O momento de hoje, e para mim ficou, bateu depressivo... O momento de hoje porque provavelmente se é o meu rosto de hoje é, dessa forma isso quer dizer que o de ontem ou o	B
5	ade, né? (Risos) Não, exatamente isso que é mais depressivo, Marcos. Chega assim pô... Até que ponto isso é que é mais depressivo? É verdade. O que eu acho	B
6	hega assim pô... Até que ponto isso é que é mais depressivo? É verdade. O que eu acho interessante, o que eu achei interessante é o próprio título "Retrato". Aí você j	B
7	chamou bastante a minha atenção. Eu achei ele depressivo. Depois da vida já morreu... É isso, matou, a vida já morreu! voice: Não é legal! Aquela	B
8	levão sobre, quão passageira é a vida. Muito depressivo. É, acho que eu não queria pensar	B

Figura 7: Classificação das linhas de concordância do item 'depressivo'

Em um estágio posterior ao processo classificatório, as linhas de concordância referentes aos adjetivos analisados foram reunidas em uma listagem única. Tal procedimento pode ser realizado com o comando '*merge*' disponível na ferramenta *Concord*. O único inconveniente deste comando é que ele só permite que dois arquivos sejam reunidos por vez, tornando o processo de fusão demorado quando se tem muitos arquivos.¹⁶ A figura a seguir ilustra a tela resultante após a fusão de todos os arquivos desejados.

¹⁶ Alternativamente, pode-se não utilizar o comando '*merge*', caso o usuário indique todas as palavras de busca ao mesmo tempo antes da geração das linhas de concordância. Neste caso, as palavras devem ser separadas por uma barra ('/').

VIANA, V. "Utilizando o programa *Wordsmith Tools* na pesquisa sobre apreciação: uma sugestão metodológica". In: *Revista Intercâmbio*, Volume XVI. São Paulo: LAEL/PUC-SP, ISSN 1806-275X-, 2007.

N	Concordance	Set
82	difícil de entender, se você tentar analisar, é muito difícil, sabe, tipo assim, parece que é uma coisa profunda, mas eu não vejo nada de profundo, entendeu? Um lúbrico vagar...	H
83	complexo isso. Muito complexo? Sua pergunta foi difícil. Complexo, não sei, depende, sei lá. Vou te falar, fora da faculdade, eu não, não leio mesmo. Só pego, folh	H
84	heria Por que você escolheria esse pedaço? Achei bonito profundo Eu concordo com você esse trecho eu achei interessante, mas eu não escolheria Todo mundo O mesmo..	I
85	u caderno é por que? Porque eu gosto acho muito bonito, profundo. Eu gostei do Coisas que a vida leva mas assim só o fragmentozinho fica meio sem sentido né acho que tem q	I
86	u só li porque meu colega falou assim "pô, é muito legal lê porque é ficção ele tem é... ele cria um mundo totalmente diferente. Aquilo me instigou, eu cheguei, eu nem sabia que o	I
87	com exceção da última parte que eu achei muito legal, mas foi o que achei que tem linguagem menos poética não só do tema, talvez pelo modo como foi abordado não gostei	I
88	cho que a leitura eu acho que tem uma mensagem linda, assim, de esperança eu acho que a leitura mexe muito com o que você acha também, a opinião do leitor. Eu acho	I
89	e milhões de livros não adiantar. E 10 quadrinhas é linda pra ler, começar desde pequena, estimular	I

Figura 8: Mapeamento do *corpus*

A Figura 8 apresenta algumas das instâncias de Composição: Complexidade negativa (H) e Valor positivo (I).

Essa estratégia permitiu, então, que fosse possível ordenar todas as instâncias apreciativas existentes no *corpus* em termos de subcategorias e valores, não obstante o adjetivo utilizado pelos participantes dos grupos de enfoque.

5) Resultados obtidos

Após a etapa de classificação dos adjetivos apreciativos, foi possível quantificar o valor das apreciações e suas categorias em termos percentuais, resultados que são apresentados nesta seção. Objetiva-se aqui demonstrar de que forma o procedimento metodológico anteriormente comentado pode ser útil na realização de uma pesquisa calcada no sistema de Atitude.

Com relação ao valor da Apreciação feita, os alunos participantes da pesquisa preferiram articular suas opiniões de forma positiva mais freqüentemente do que de forma negativa, como indica a Figura 9.

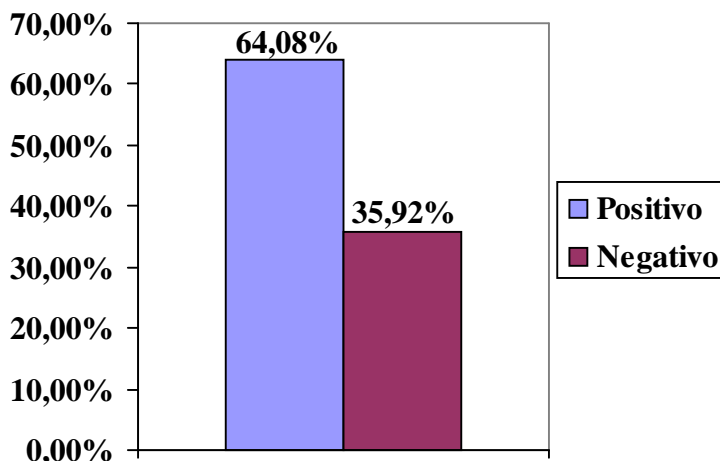


Figura 9: Valor das instâncias apreciativas

Ao todo, 35,92% das instâncias apreciativas são negativas, como pode ser exemplificado pelo comentário realizado pelo participante M2 a respeito do "Soneto de Fidelidade":

<M2> [...] Depois ficou o Soneto de Fidelidade. Eu acho o Soneto é... muito certinho ele é muito sabe qual é? Tem que rimar aqui, fazer o ABBA... Não gosto muito disso – uma coisa padrão. <SMFF>¹⁷ Uma coisa padrão, não gosto muito desse negócio de padrão não, gosto mais de negócio livre...

O participante em questão utiliza o adjetivo 'certinho' para criticar o esquema de rimas empregado por Vinícius de Moraes.

No entanto, grande parte das apreciações é positiva. Um exemplo é a resposta dada pela participante F3 ao comentário anteriormente reproduzido:

<F3> Eu sou o contrário dele. O fato de estar rimando eu acho lindo. Se tiver rima, ah eu eu acho mais bonitinho pela forma do poema e também pelo que tá lá, pelas palavras as figuras...

Apreciações positivas como esta totalizam 64,08% das instâncias apreciativas analisadas no *corpus* de pesquisa.

Com relação às categorias apreciativas, os participantes parecem preferir tecer comentários que se enquadrem em Reação ou em Valor principalmente. A categoria de Composição fica relegada a um segundo plano.

¹⁷ Esta notação serve para indicar que houve uma fala simultânea de uma participante do sexo feminino cuja identificação não foi possível.

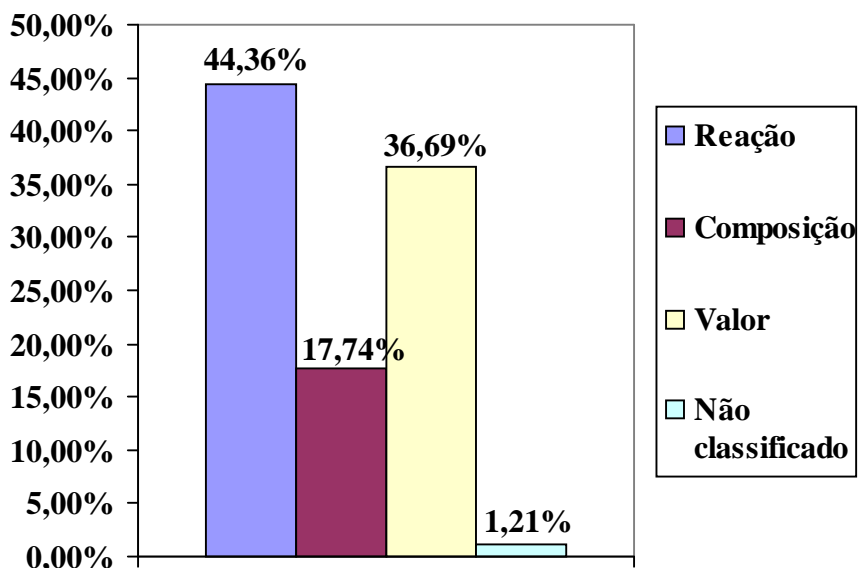


Figura 10: Tipo de Apreciação

Os comentários feitos pelos participantes podem ser classificados como característicos da categoria Reação em 44,36% como no exemplo a seguir:

<F7> Eu achei foi o primeiro que eu escolhi também eu achei interessante porque ele fala assim das coisas da matéria mesmo como é tudo efêmero assim, ele fala que a vida é tipo é... "o sorriso aberto" a juventude, né tudo isso mas como o sentimento, o "ardor sagrado" é uma coisa que ainda tá dentro dele essa chama. O dia que a vida levar isso realmente ele vai tá morto. Acho interessante essa contraposição em que essa chama, o sentimento e que a vida realmente vai levando os anos.

A participante F7 utiliza o adjetivo 'interessante' para qualificar a contraposição existente em um dos poemas distribuídos no início do grupo de enfoque. A opção por este marcador avaliativo revela o impacto que tal contraposição teve na participante.

A segunda categoria mais freqüente nos dados analisados é a de Valor, que soma 36,69% dos casos estudados. A título de exemplificação, cita-se o comentário da participante F6 a respeito do "Soneto de Fidelidade":

<F6> Eu concordo com ela "que não seja imortal posto que é chama / mas que seja infinito enquanto dure" esse que eu escolheria [...] Achei bonito profundo

VIANA, V. "Utilizando o programa *Wordsmith Tools* na pesquisa sobre apreciação: uma sugestão metodológica". In: *Revista Intercâmbio*, Volume XVI. São Paulo: LAEL/PUC-SP, ISSN 1806-275X-, 2007.

Ao utilizar o adjetivo 'profundo', a participante faz uma avaliação do poema a partir das convenções sociais estabelecidas.

Por sua vez, a categoria de Composição totaliza somente 17,74%, sendo a menos freqüente no *corpus* de estudo. Um exemplo de tal tipo pode ser encontrado no comentário de uma participante a respeito de um poema sem título de Fernando Pessoa:

<F8> Porque lendo rapidamente deles foi o que eu achei assim mais melódico que eu gostei mais de ler mais fácil também de você assimilar né foi mais fácil pra mim assimilar porque lendo assim rapidamente é o que chama mais a atenção a linguagem é mais fácil...

O emprego do adjetivo 'fácil' denota a opção da participante de comentar a estrutura do poema em questão, atentando para a facilidade de sua estrutura e da linguagem empregada.

Além das três categorias estabelecidas para a expressão de Apreciação, a Figura 10 também indica a existência de instâncias apreciativas que não puderam ser classificadas. Tal grupo corresponde a 1,21% dos dados analisados e pode ser exemplificado com o seguinte comentário:

<F14> Uma idéia de ser profundo, entendeu.

Tudo o que a participante F14 diz nesta parte do grupo de enfoque é a frase acima transcrita, não havendo nenhuma outra explicação no cotexto que possa permitir a compreensão do real significado da sentença. Não há também nesta parte da interação nenhuma pergunta realizada pelo moderador de forma a clarificar o que a participante deseja expressar. Torna-se difícil, portanto, afirmar em que tipo de marcador apreciativo o adjetivo 'profundo' se constitui neste caso. Assim sendo, optou-se por agrupar tais instâncias em uma categoria de dados não classificados.

De forma resumida, tem-se que os dados analisados apontam que os participantes dos três grupos de enfoque, alunos universitários de Letras, preferem articular suas opiniões apreciativas de forma a indicar como eles reagem aos objetos e processos avaliados e como eles os avaliam dentro de suas próprias sociedades. Estas avaliações são, na maioria dos casos, de valor positivo.

6) Avaliação da metodologia

A metodologia descrita neste artigo apresenta algumas vantagens consideráveis e que atestam a favor da utilização da ferramenta

VIANA, V. "Utilizando o programa *Wordsmith Tools* na pesquisa sobre apreciação: uma sugestão metodológica". In: *Revista Intercâmbio*, Volume XVI. São Paulo: LAEL/PUC-SP, ISSN 1806-275X-, 2007.

computacional em um estudo de cunho semântico. Há, contudo, algumas limitações no emprego da metodologia, que também são explicitadas na presente seção.

Em relação aos pontos positivos, apesar de o *corpus* investigado ser de tamanho reduzido, destaca-se a possibilidade de trabalhar com *corpora* de grandes dimensões com o auxílio computacional. Mesmo sendo rotulado como 'pequeno', uma investigação manual deste *corpus* demandaria muito trabalho e provavelmente geraria erros já que o ser humano não é talhado para a realização de tarefas repetitivas como a busca de determinados itens lexicais, o que é feito em segundos por um computador.

Uma vez que o *corpus* tenha sido formatado de forma que possa ser lido pelo computador, não há diferenças entre a natureza dos dados. Pode-se investigar um *corpus* oral ou escrito. Na primeira categoria, pode-se ter somente o relato de uma única pessoa ou um grupo de enfoque como é o caso aqui. As pausas, pausas vocalizadas, auto-reparos e reparos, fenômenos característicos da interação oral, não prejudicam sobremaneira a análise de dados através do computador.

A opção por um ponto de entrada no *corpus* de estudo, neste caso a classe gramatical de adjetivo, permitiu que o computador pudesse localizar todas estas instâncias nos dados analisados de forma automática e confiável. Acaba-se assim com o problema da segmentação das instâncias apreciativas, algo que era um ponto nevrálgico quando se tentou realizar o mapeamento do *corpus* de forma manual. Não é relevante neste tipo de metodologia onde começa e termina a avaliação. O que importa é detectá-la através do adjetivo utilizado e classificá-la com base em seu contexto.

Apesar de suas vantagens, a escolha de um ponto de entrada também apresenta alguns problemas. Neste tipo de levantamento, só é possível mapear as instâncias apreciativas que são realizadas por meio do ponto de entrada escolhido, deixando-se de fora da análise as possíveis outras instâncias apreciativas. No caso deste estudo, por exemplo, foram mapeados os adjetivos apreciativos, mas deixaram de ser analisadas as outras instâncias de Apreciação que, porventura, não se realizaram por meio de tal categoria gramatical.

Outra limitação é o fato de o programa *WordSmith Tools* (Scott, 1999) somente aceitar uma letra na coluna *set*, o que obriga o pesquisador a utilizar uma série de códigos que correspondam às mais diversas subcategorias que compõem o subsistema de Apreciação. O trabalho de codificação se torna ainda mais difícil se o pesquisador quiser trabalhar, por exemplo, com o sistema de Atitude como um todo.

Pelo mesmo motivo, torna-se difícil a classificação de instâncias que podem ser representativas de duas ou mais possíveis (sub)categorias. Quando não se é possível identificar a (sub)categoria

VIANA, V. "Utilizando o programa *Wordsmith Tools* na pesquisa sobre apreciação: uma sugestão metodológica". In: *Revista Intercâmbio*, Volume XVI. São Paulo: LAEL/PUC-SP, ISSN 1806-275X-, 2007.

apreciativa de um determinado exemplo a partir da análise de seu cotexto, uma solução é a criação de um novo grupo que abarque as instâncias não-classificadas, opção adotada neste estudo.

Por fim, pode-se apontar mais uma desvantagem, a saber, o fato de a ferramenta *Concord* só promover a fusão de dois arquivos por vez. Caso a pesquisa inicial não tenha sido realizada com todos os adjetivos, o pesquisador gastará um tempo considerável até que consiga reunir todos os arquivos de extensão CNC em um único para que possa ter uma visão geral das instâncias que foram classificadas.

7) Conclusão

No início do artigo, propôs-se uma questão de pesquisa de caráter metodológico, a saber, a utilização da ferramenta computacional em uma pesquisa eminentemente semântica a partir da análise de dados orais. Este artigo mostrou em detalhes como o programa *WordSmith Tools* (ibid) pode ser utilizado para extrair, identificar, listar e ordenar as ocorrências do subsistema de Apreciação em um *corpus* específico.

De forma a exemplificar a utilização da metodologia aqui descrita, foram reportados resultados de uma investigação acerca das instâncias apreciativas em um *corpus* oral compilado a partir da realização de três grupos de enfoque (cada um com duração aproximada de 1 hora) nos quais universitários de Letras debateram a questão da Literatura e/ou do objeto literário. Os resultados indicam que estes estudantes preferiram comentar a reação que eles tiveram perante os objetos e/ou processos avaliados, ou seus respectivos valores dentro da cultura na qual eles encontram-se inseridos. Os comentários realizados foram, na maioria dos casos, realizados a partir de uma ótica positiva.

A avaliação da metodologia sugerida ressalta os seus pontos positivos nesta pesquisa, lhe conferindo uma rapidez e confiabilidade com a inserção do computador, quando comparada a uma metodologia manual e baseada no olho humano. Porém, também foram apontadas algumas limitações da ferramenta computacional aqui empregada, que acabam por gerar certo trabalho para o analista, especialmente quando da criação dos códigos a serem utilizados para representar cada uma das categorias e subcategorias.

Apesar de haver desvantagens na utilização da ferramenta computacional neste estudo, nenhuma crítica feita a tal emprego em pesquisas lingüísticas consegue negar a facilidade, rapidez e exatidão que o computador conferiu à exploração de *corpora*. Segundo Tognini-Bonelli (2001:5), o computador "afetou a estrutura metodológica da

VIANA, V. "Utilizando o programa *Wordsmith Tools* na pesquisa sobre apreciação: uma sugestão metodológica". In: *Revista Intercâmbio*, Volume XVI. São Paulo: LAEL/PUC-SP, ISSN 1806-275X-, 2007.

investigação lingüística ao agilizá-la, sistematizá-la e fazê-la aplicável em tempo real a quantidades de dados cada vez maiores"¹⁸.

No caso específico do mapeamento do sistema de Atitude, verifica-se que é possível utilizar o computador de forma a facilitar o trabalho de pesquisa. A ferramenta computacional, contudo, não substitui o pesquisador, já que o trabalho qualitativo de análise e a subsequente classificação das linhas de concordância ainda não podem ser feitos de maneira automática, uma vez que não há formas fixas que identifiquem automaticamente a linguagem atitudinal. Assim sendo, outras pesquisas e novos avanços na área da informática são necessários para que a metodologia aqui apresentada possa ser aperfeiçoada.

REFERÊNCIAS BIBLIOGRÁFICAS

BARBOUR, R.; KITZINGER, J. (Ed.). *Developing focus group research: politics, theory and practice*. London: Sage Publications, 1999.

BARROS, H. dos S. "Verdade ou aparência? A literatura e o aluno além da universidade". In: ZYNGIER, S.; MENDES, M. P.; PINHEIRO, P. (Eds.). *Pontes & transgressões: estudos empíricos de processos culturais*. Rio de Janeiro: Setor de Publicações da Faculdade de Letras da UFRJ, 2003, p. 56-66.

BERBER SARDINHA, A. P. *Lingüística de Corpus*. São Paulo: Manole, 2004.

FERREIRA, A. B. de H. *Novo Aurélio século XXI: o dicionário da língua portuguesa*. 3ª. ed. Rio de Janeiro: Nova Fronteira, 1999.

MARTIN, J. R. "Beyond exchange: Appraisal systems in English". In: HUNSTON, S.; THOMPSON, G. (Eds.). *Evaluation in Text*. Oxford: Oxford University Press, 2000, p. 142-175.

MARTIN, J. R.; ROSE, D. *Working with discourse: meaning beyond the clause*. London: Continuum, 2003.

MATTOS, M. J.; MENDES, M. P. "Uma proposta de projeto sobre grupo de enfoque na Ciência Empírica da Literatura: atitude dos alunos em relação à literatura". In: ZYNGIER, S.; MENDES, M. P.; PINHEIRO, P. (Eds.). *Pontes & transgressões: estudos empíricos de processos culturais*. Rio de Janeiro: Setor de Publicações da Faculdade de Letras da UFRJ, 2003, p. 111-119.

MENDES, M. P. "Questões metodológicas em Lingüística Computacional: a montagem do *corpus* específico". In: ZYNGIER, S.; VALENTE, A. C. (Eds.). *Fatos & ficções: estudos empíricos de literatura*. Rio de Janeiro: Setor de Publicações da Faculdade de Letras da UFRJ, 2002, p. 142-146.

¹⁸ Tradução livre do seguinte fragmento: "has affected the methodological frame of the enquiry by speeding it up, systematizing it, and making it applicable in real time to ever larger amounts of data".

VIANA, V. "Utilizando o programa *Wordsmith Tools* na pesquisa sobre apreciação: uma sugestão metodológica". In: *Revista Intercâmbio*, Volume XVI. São Paulo: LAEL/PUC-SP, ISSN 1806-275X-, 2007.

MENEZES, D. de A. "Conceitos de literatura, textos e mídia". In: ZYNGIER, S.; VALENTE, A. C. (Eds.). *Fatos & ficções: estudos empíricos de literatura*. Rio de Janeiro: Setor de Publicações da Faculdade de Letras da UFRJ, 2002, p. 147-155.

ROTHERY, J.; STENGLIN, M. "Interpreting literature: the role of Appraisal". In: UNSWORTH, L. (Ed.). *Researching language in schools and Functional Linguistic perspectives*. London: Cassell, 2000, p. 222-244.

SCOTT, M. *WordSmith Tools*, versão 3.0. Oxford: Oxford UP, 1999.

TABOADA, M. Re: [AppraisalAnalysis] automating appraisal analyses [mensagem pessoal]. Mensagem recebida por <AppraisalAnalysis@yahoogroups.com> em 09 mar. 2004.

TOGNINI-BONELLI, E. *Corpus Linguistics at work*. Amsterdam/Philadelphia: John Benjamins Company, 2001.

VIANA, V. "Talking about literature in L1 and L2: a study of adjectives". In: ZYNGIER, S.; VIANA, V.; FAUSTO, F. (Eds.). *Venturas & desventuras: coletânea dos trabalhos do V ECEL*. Rio de Janeiro: Setor de Publicações da Faculdade de Letras da UFRJ, 2005a, p. 139-150.

_____. "O papel da literatura na vida de estudantes de Letras: um estudo lingüístico das manifestações de Apreciação". In: SALIÉS, T. M. G.; SHEPHERD, T. M. G. (Org.). *Linguagem: teoria, análise e aplicações*. Rio de Janeiro: Publit Soluções Editoriais, 2005b, p. 225-244.

VIANA, V.; SHEPHERD, T. "Dize-me como falas e eu te direi com quem falas: um estudo sobre a linguagem de Apreciação". *Cadernos do CNLF*, v. VIII, n. 05, p. 74-85, 2004.

WHITE, P. R. R. *An introductory tour through appraisal theory*. 2000. Disponível em: <<http://www.grammatics.com/appraisal>>. Acesso em 16 ago. 2003.

ZYNGIER, S.; SHEPHERD, T. "What is literature, really? A corpus-driven study of students' statements". *Style*, v. 37, n. 1, p. 14-26, 2003.