

Especificidades de um Discurso

Dulce CARVALHO; Isabel Rego SANTOS; Isabel SARAIVA, Maria de Fátima F. SILVA
(Universidade Aberta)

0. Introdução

Abordar a produção da Professora Maria Emília Ricardo Marques, na década de 90, a partir de um Modelo Lexicométrico¹, tem como finalidade reconstituir um Universo de referências, redes léxico-textuais, redes temáticas veiculadas por formas específicas atestadas no *corpus*².

Tal metodologia implica o registo e tratamento de textos sobre suporte informático, assegurando um inventário rigoroso e exaustivo de todas as ocorrências de todas as formas do *corpus*.

As análises estatísticas deste modelo permitem-nos medir a frequência relativa de cada forma lexical, localizá-la em contexto textual, seleccioná-la e caracterizá-la em função de parâmetros de situação, fornecendo-nos assim resultados objectivos e sistematizados.

Este modelo garante-nos a distância necessária entre os materiais utilizados e o EU analisante, tornando a análise da produção discursiva o mais objectiva possível. É aliás uma metodologia que a Professora Maria Emília Ricardo Marques tem vindo a utilizar ao longo da sua vasta produção científica.

1. Constituição do *corpus*

1.1. Textos seleccionados

Seleccionámos 22 artigos exclusivamente redigidos por Maria Emília Ricardo Marques na década de 90. Alguns estão publicados, outros apenas foram alvo de comunicações em congressos e colóquios.

As datas consideradas são as da escrita, exceptuando-se os artigos publicados. Nestes casos a data é a da publicação.

Ao construir o *corpus*, excluímos quadros, esquemas, tabelas, citações de outros autores, notas de rodapé, bibliografias e referências bibliográficas, contextualizadas ou não, legendas de imagens e respectivas imagens, bem como índices.

Mantivemos títulos e subtítulos, embora não tivessem sido alvo de tratamento estatístico.

Data	Título
1991	<i>A tecnologia educativa numa encruzilhada de caminhos</i>
1991	<i>Bilinguismo: algumas reflexões</i>
1992	<i>Audiovisuais, mediatecas escolares e transformações na aprendizagem - Mediatecas escolares</i>
1992	<i>Formação profissional: aprendizagem de línguas em estações de trabalho</i>
1992	<i>O Rigor de um jogo</i>
1993	<i>A metáfora na divulgação científica: um exemplo «bélico»</i>
1993	<i>Do consumismo à banalização</i>
1993	<i>Para uma política linguística</i>
1994	<i>A inovação tentada</i>
1994	<i>O ensino do português na comunidade: a pragmática da inovação</i>
1994	<i>Percursos de palavra</i>
1994	<i>Tentativa de resposta para um equilíbrio em rotura</i>
1995	<i>Linguagem e Poder</i>
1995	<i>O acto de olhar no discurso de Fernão Lopes</i>
1996	<i>Espaços de um agir comunicacional</i>
1997	<i>Rituais parlamentares constituintes: a palavra feminina</i>
1998	<i>Imagem: Mulher</i>
1999	<i>Cesário Verde e a Cidade: que representações</i>
1999	<i>Encruzilhadas de percursos</i>
1999	<i>Multiculturalismo</i>
1999	<i>Rituais parlamentares constituintes: perspectivas</i>
1999	<i>Texto literário e evolução tecnológica</i>

1.2. Normalização

O estudo lexicométrico obrigou-nos a proceder a uma regularização ortográfica dos textos que designamos normalização.

Ela incidiu não só sobre as maiúsculas sistemáticas, puramente tipográficas, que passaram a minúsculas, mas também sobre as maiúsculas iniciais dos nomes de meses, estações de ano e axiónimos que sofreram um tratamento idêntico.

Os vocábulos destacados, quer através de maiúsculas, quer através de negritos, por vontade expressa da autora, escreveram-se com minúsculas precedidas de asterisco.

As siglas passaram também a minúsculas antecedidas de asterisco.

Apenas mantivemos a maiúscula inicial nos antropónimos e topónimos.

Não fizemos qualquer lematização (não reduzimos nem as formas verbais ao infinitivo nem os femininos ao masculino nem os plurais ao singular) nem desambigüizámos homófonas.

1.3. Partições

Porque pretendemos não só inventariar o vocabulário específico da autora como também verificar se ele se alterou ao longo da década de 90, estabelecemos as seguintes partições, procurando equilibrar a extensão do vocabulário de cada uma das partes.

Parte	Nº de ocorrências	Nº de formas	Nº de hapax	Frequência máxima	Forma
1991-93	19083	3834	1795	1131	<i>de</i>
1994-96	20054	4166	2192	1296	<i>de</i>
1997-99	19795	4510	2501	1132	<i>de</i>

Uma leitura rápida da tabela aponta para uma maior diversidade de formas entre 1997 e 1999. Consta-se que neste triénio a relação entre o número de formas diferentes e o número de ocorrências é superior à verificada nos períodos anteriores. Além disso, o número de formas que aparecem uma única vez (hapax) sofreu um visível aumento nesta parte do *corpus*. Caracterizar-se-á este período por uma maior riqueza lexical? Serão as temáticas abordadas mais variadas?

2. Vocabulário Específico

2.1. Método das Especificidades

Para determinarmos o vocabulário específico da autora e para podermos comparar as alterações por ele sofridas ao longo da década de 90, utilizámos o Método das Especificidades.

Este método estatístico utiliza o modelo Hipergeométrico para fazer um julgamento, em probabilidade³, sobre a especificidade ou não de uma forma numa parte do *corpus*.

De facto, ocorrem formas que se evidenciam, quer pelo seu sobre-emprego – **especificidade positiva (S+)**, quer pelo seu sub-emprego – **especificidade negativa (S-)**. Ocorrem ainda formas que não apresentando qualquer tipo de especificidade, numa determinada parte do *corpus*, são consideradas **banais (b)** nessa mesma parte. Quando a forma é banal em todas as partes do *corpus*, considera-se que a forma faz parte do **vocabulário de base (B)**.

2.2. Resultados

Da aplicação do Método das Especificidades constatou-se, em primeiro lugar, que das 7920 formas diferentes que ocorrem no *corpus*, 1128 fazem parte do Vocabulário de Base usado pela autora.

Analisemos as especificidades positivas mais relevantes⁴.

Especificidades positivas do triénio 1991-93

	1991-93	1994-96	1997-99
<i>metáfora</i>	+E07	-E04	-E03
<i>visual</i>	+E05	b	-E02
<i>projecto</i>	+E05	b	-E04
<i>possíveis</i>	+E05	b	-E02
<i>línguas</i>	+E05	+E02	-E14
<i>lado</i>	+E05	b	-E03
<i>jogo</i>	+E05	b	-E04
<i>concepção</i>	+E05	b	-E03
<i>público</i>	+E04	b	-E02
<i>leitor</i>	+E04	b	-E04
<i>estratégias</i>	+E04	-E02	b
<i>documentos</i>	+E04	b	-E08
<i>bilinguismo</i>	+E04	b	-E08

O vocabulário sobre-empregue neste período abre espaço a um cenário hipotético em que as preocupações da autora parecem girar em torno de um projecto que engloba:

- as *línguas* (note-se o plural) e o *bilinguismo*
- um duplo destinatário: *público* e *leitor*
- as estratégias possíveis: *metáfora*, *visual*, *possíveis*, *jogo*, *estratégias*
- os documentos: *visual*, *projecto*, *concepção*, *documentos*

Em contrapartida, no triénio seguinte, estas formas tornaram-se, na sua maioria, banais com excepção de *línguas* que continua específica positiva, mas com um coeficiente de 02 e *metáfora* e *estratégias* que passaram para específicas negativas.

No último período analisado, com excepção de *estratégias* que passou a ser banal, todas as outras formas se tornaram específicas negativas, portanto deixaram de ter relevância para o emissor.

Especificidades positivas do triénio 1994-96

	1991-93	1994-96	1997-99
<i>poder</i>	-E02	+E06	-E03
<i>discurso</i>	-E02	+E05	-E02
<i>situações</i>	b	+E04	-E05
<i>pedagogia</i>	-E03	+E04	b
<i>papéis</i>	-E02	+E04	b
<i>dificuldades</i>	-E02	+E04	b
<i>último</i>	-E02	+E04	b

Neste período as preocupações da autora deslocaram-se para outra esfera de interesses onde sobressai o vocábulo *poder*: *discurso de um poder instalado, língua do poder, relações de poder, grupos de poder, génese do poder, exercício de um poder, duelo por um poder futuro...* são alguns dos contextos em que ocorre a forma *poder* nesta partição⁵.

O retorno ao texto através das concordâncias fornecidas permitiu-nos verificar que ao *poder* se associam todos os outros vocábulos sobre-empregues: *discurso, pedagogia (da negociação), papéis (socio-políticos, profissionais, sociocomunicativos* bem como *mudança de papéis, papéis "normalizados"* [por regras]).

À semelhança do que aconteceu no período anterior estes vocábulos deixaram de ter o mesmo peso no período seguinte.

Especificidades positivas do triénio 1997-99

	1991-93	1994-96	1997-99
<i>formas</i>	-E10	-E02	+E14
<i>deputados</i>	-E05	-E04	+E11
<i>número</i>	-E04	-E02	+E08
<i>ver</i>	-E03	-E03	+E07
<i>partes</i>	-E03	-E03	+E07
<i>sujeito</i>	-E03	-E02	+E06
<i>vocabulário</i>	-E04	b	+E06
<i>deferência</i>	-E03	-E02	+E06
<i>verbos</i>	-E03	-E02	+E06
<i>corpus</i>	-E09	b	+E05
<i>nível</i>	-E03	-E02	+E05
<i>segmentos</i>	-E03	b	+E05
<i>população</i>	-E04	b	+E05
<i>categoria</i>	-E03	b	+E05
<i>categorias</i>	-E03	b	+E05
<i>eu</i>	b	-E03	+E05
<i>família</i>	b	-E03	+E05
<i>conjunto</i>	b	-E02	+E04
<i>hipótese</i>	b	-E02	+E04
<i>vida</i>	-E04	b	+E04
<i>conhecimento</i>	-E02	-E02	+E04
<i>investigação</i>	-E02	b	+E04
<i>ocorrências</i>	-E03	b	+E04
<i>média</i>	b	-E03	+E04
<i>referido</i>	b	b	+E04
<i>figura</i>	-E02	b	+E04
<i>posição</i>	-E02	b	+E04
<i>profissão</i>	-E02	-E02	+E04

As preocupações da autora no último triênio parecem-nos mais diferenciadas. A panóplia de interesses abre-se.

A análise desta última tabela remete-nos fundamentalmente para dois centros de interesse: um, continua a ser o poder veiculado por *deputados, sujeito, deferência, eu, família, media, figura, posição e profissão*; o outro, é o do *conhecimento, da investigação*.

Vocábulos como *formas, número, partes, vocabulário, corpus, segmentos, população, categoria, categorias, hipóteses, conjunto e ocorrências* remetem-nos para uma área de investigação muito particular, a da Análise do Discurso, em que se evidencia uma metodologia e ferramenta de trabalho, também ela muito específica, a Lexicometria, modelo tão do agrado da Professora Maria Emília e que utilizou ao longo da sua vasta carreira de investigação em busca de um conhecimento cada vez mais profundo, objectivo e sistemático e em que o(s) discurso(s) (pedagógico, político, dos *massmedia*, ...) assume(m) especial importância pelos diferentes tipos de relações que se estabelecem entre o EU e o OUTRO.

Foi também graças a esta metodologia que pudemos apresentar este pequeno trabalho. Com ele pretendemos prestar o nosso tributo à Professora Maria Emília que tanto nos ensinou nesta área explorada minimamente, aqui. Os índices alfanuméricos (alfabéticos e hierárquicos), as formas não plenas, o vocabulário de base, os segmentos repetidos fornecidos pelo programa utilizado, LEXICO 3^o, são materiais ricos, a trabalhar e que muito mais nos diriam sobre as especificidades e as idiossincrasias de um discurso muito especial, o discurso da Professora Maria Emília.

Bem Haja, PROFESSORA!

Notas

¹ Por Lexicometria entende-se um conjunto de métodos que permitem operar, a partir de análises estatísticas, reorganizações formais do vocabulário (conjunto de formas actualizadas no discurso, atestadas num texto ou num *corpus* de textos), de um *corpus* recolhido na sequência textual.

² *Corpus* é um conjunto homogéneo, sincrónico e representativo de textos reunidos com o objectivo de serem comparados.

³ Numa investigação lexicométrica, a primeira fase do trabalho incide sobre as frequências. De facto, a tabela de frequências apresenta a listagem das formas que ocorrem no *corpus* e as respectivas sub-frequências em cada uma das partes que o constituem. No entanto, tais frequências, apesar de representarem dados objectivos, não conduzem a uma análise fiável, uma vez que as partes neste estudo não têm exactamente o mesmo número de ocorrências.

⁴ Os coeficientes de especificidade positiva encontrados variam entre +E14 e +E02. Decidimos analisar apenas as formas com um coeficiente de especificidade ≥ 04 .

⁵ Vejam-se as seguintes concordâncias fornecidas pelo programa LEXICO 3.

por conseguinte, o discurso de um	poder	instalado, que se queria seguro.
daí que a língua do	poder	internacional seja, quer a estritamente
sobretudo, em relações de	poder,	são conhecimentos partilhados
como condicionantes de relações de	poder,	queria destacar o primeiro,
no caso em estudo, por relações de	poder	ou de solidariedade.
há ainda quem oponha grupos de	poder	e grupos de solidariedade,
recorde-se que, na génese do	poder	político e na estruturação do estado actual
quer como exercício de um	poder	instável, reflexo de luta quer entre culturas
quer como duelo por um	poder	futuro, quer como exercício de um

⁶ Programa concebido por André SALEM *et alli* (Université de la Sorbonne Nouvelle – Paris III).

Bibliografia

- Carvalho, D.; Marques, M. E.; Silva, M. F. 1999. «Discurso: práticas lexicométricas». In *Linguística Computacional, Investigação Fundamental e Aplicações*. Lisboa: Ed. Colibri, Associação Portuguesa de Linguística.
- Carvalho, D.; Marques, M. E.; Silva, M. F. 2003. «Mediações de uma notícia». In *Razões e Emoção – Miscelânea de estudos em homenagem a Maria Helena Mira Mateus*. Vol. I. Lisboa: Imprensa Nacional-Casa da Moeda.
- Carvalho, D.; Silva, M. F. (Eds.). 1993. *Lexicometria, séminaire de Maurice Tournier*. Lisboa: Universidade Aberta (Colecção Temas Educacionais).
- Charaudeau, P. 1992. *Grammaire du sens et de l'expression*. Paris: Hachette Education.
- Nascimento, F. P.; Rodrigues, M. C.; Gonçalves, J. B. (Eds.). 1996. *Actas do XI Encontro Nacional da Associação Portuguesa de Linguística*. Vol. 1 – *Corpora*. Lisboa: Colibri-Artes Gráficas, Faculdade de Letras da Universidade de Lisboa.