

S E L O U N I V E R S I D A D E

L I N G U Í S T I C A



LÉXICO E
IDEOLOGIA NA
EUROPA OCIDENTAL



ALDO BIZZOCCHI



ANNA BOLLME

UNIP
EDITIVO

FAPESP

BUSCANDO
ENCONTRAR AS
CAUSAS DAS
CONVERGÊNCIAS E
DIVERGÊNCIAS
DAS FORMAÇÕES
LÉXICO-GRAMATICAS
DE IDIOMAS DISTINTOS
(FRANCÊS, INGLÊS,
ALEMÃO, PORTUGUÊS
E ITALIANO), ALDO
BIZZOCCHI, VALENDO-
SE DO MÉTODO
COMPARATIVO,
SUBMETE OS
VOCABULÁRIOS DE
DISCURSOS
CIENTÍFICOS,
FILOSÓFICOS E
JORNALÍSTICOS

CONTEMPORANEOS
A UM TRATAMENTO
QUALITATIVO E
QUANTITATIVO
BASEADO EM
PRINCÍPIOS DA
LINGÜÍSTICA
QUÂNTICA E
POSTULADOS DA
LINGÜÍSTICA DO
DISCURSO, NUMA
INTERFACE COM
DADOS DA
LEXICOLOGIA.
O RESULTADO É UMA
PESQUISA
FUNDAMENTAL PARA
OS ESTUDOS DOS MAIS
DIFERENTES CAMPOS
DA LINGUAGEM

**LÉXICO E IDEOLOGIA NA
EUROPA OCIDENTAL**

ALDO BIZZOCCHI

LÉXICO E IDEOLOGIA
NA EUROPA
OCIDENTAL

B6251

Bizzocchi, Aldo Luiz

Léxico e ideologia na Europa Ocidental / Aldo Bizzocchi –

São Paulo : Annablume, 1997.

271 p.

ISBN 85-85596-97-X

1. Línguas indo-européias – Lexicologia. 2. Línguas indo-européias – Filosofia. 3. Título.

CDD 491

LÉXICO E IDEOLOGIA NA EUROPA OCIDENTAL

Aldo Luiz Bizzocchi

Revisão:

Mara Guasco

CONSELHO EDITORIAL

Eduardo Peñuela Cañizal

Willi Bolle

Norval Baitello Junior

Carlos Gardin

Lucrécia D'Aléssio Ferrara

Plínio de Arruda Sampaio

Maria Odila Leite da Silva Dias

Ivan Bystrina

Salma T. Muchail

Ubiratan D'Ambrósio

Maria de Lourdes Sekeff

Gilberto Mendonça Teles

1ª edição: fevereiro de 1998

© Aldo Luiz Bizzocchi

ANNABLUME editora . comunicação
Rua Ferreira de Araújo, 359 — Pinheiros
05428-000 . São Paulo . SP . Brasil
Tel e Fax. (011) 212.6764
<http://www.annablume.com.br>

AGRADECIMENTOS

Ao Professor Doutor Cidmar Teodoro Pais, pela orientação segura e precisa, pelos inestimáveis conselhos e pelo permanente estímulo.

Aos Professores Norberto Luiz Guarinello, Franklin Leopoldo e Silva, Leonardo Melo e Silva, José Flávio Motta, Maria Regina Kawamura, Lisbeth Kaiserlian Cordani e Zara Issa Abud, pelas preciosas informações fornecidas.

À CAPES e ao CNPq, pelo apoio financeiro à pesquisa e à FAPESP, pelo auxílio à publicação deste livro.

A todos que, direta ou indiretamente, contribuíram para a realização deste trabalho.

*Aos meus pais, pelo incansável
apoio e constante incentivo.*

SUMÁRIO

APRESENTAÇÃO	15
ABREVIATURAS E SÍMBOLOS	19
INTRODUÇÃO	21
CAPÍTULO I: NOÇÕES PRELIMINARES	27
ETIMOLOGIA E LEXICOGÊNESE	29
A EUROPA OCIDENTAL	31
AS LÍNGUAS EUROPÉIAS OCIDENTAIS	35
O ASPECTO IDEOLÓGICO DO LÉXICO	40
CARACTERÍSTICAS DO <i>CORPUS</i> ESTUDADO	45
Modalidade oral x modalidade escrita	45
Análise sincrônica x análise pancrônica	46
Línguas utilizadas na pesquisa	47
Universos de discurso abordados	51
CAPÍTULO II: FUNDAMENTOS TEÓRICOS	53
LÉXICO E VOCABULÁRIO: LEXIAS, VOCÁBULOS, PALAVRAS	55
CONVENÇÕES ADOTADAS NESTE TRABALHO	57
TIPOS DE EMPRÉSTIMOS	58
EMPRÉSTIMOS DE VOCÁBULOS GRECO-LATINOS	59
Processos não vulgarizantes	62
Processos vulgarizantes	64
TIPOLOGIA LEXICOGÊNICA DOS VOCÁBULOS DE UMA LÍNGUA	68

LATIM NORMALIZADO	71
VULGARISMOS	73
Vulgarismos resultantes de herança vernácula	73
Vulgarismos alogenéticos (empréstimos de vulgarismos estrangeiros)	77
Alógenos que retornam à língua de origem	80
Empréstimo de vulgarismos com as respectivas desinências	80
Empréstimos de tradução	81
Vulgarismos resultantes de metamorfismo	100
<i>Metamorfismo espontâneo e metamorfismo deliberado (analógico)</i>	100
Vulgarismos resultantes de combinação sintagmática (composição ou derivação)	110
Co-ocorrência de processos lexicogênicos no mesmo vocábulo	115
Vocábulos cujas formas flexionais pertencem a categorias lexicogênicas diferentes	116
GRECOLATINISMOS	116
Grecolatinismos resultantes de empréstimo direto ou indireto do grego/latim	116
Refecções de vulgarismos	117
Empréstimos de restituição	119
Grecolatinismos resultantes de combinação sintagmática (composição ou derivação)	125
VOCÁBULOS NÃO CLASSIFICÁVEIS NAS CATEGORIAS ANTERIORES	127
Vocábulos derivados de nomes próprios	127
Siglas e seus derivados	128
Vocábulos formados de fractomorfemas	129
Neologismos fonológicos específicos	131
PERCURSO GERATIVO DE CRIAÇÃO VOCABULAR	132
MECANISMOS DE EMPRÉSTIMO	133

CRITÉRIOS ESPECIAIS PARA A CLASSIFICAÇÃO DOS VOCÁBULOS	137
Dos nomes próprios e seus compostos e derivados	137
Classificação dos advérbios de modo derivados de adjetivos	138
Vocábulo derivados de radicais greco-latinos com afixos vernáculos de aparência erudita	140
Vocábulo alótipos	143
Tratamento da vogal pós-tônica	145
Vocábulo de origem grega terminados em <i>-ma</i>	150
Vocábulo classificáveis em mais de uma categoria	152
CAPÍTULO III: APLICAÇÃO DA TEORIA LEXICOGÊNICA À ANÁLISE ESTATÍSTICA CONTRASTIVA DE LÉXICO	155
PROCEDIMENTOS LEXICOLÓGICOS	157
Estabelecimento do <i>corpus</i>	158
A classificação dos vocábulo	164
<i>O Índice Etimológico (IE)</i>	167
PROCEDIMENTOS ESTATÍSTICOS	170
Tamanho e representatividade do <i>corpus</i>	170
A estatística lexical	177
CAPÍTULO IV: RESULTADOS E DISCUSSÃO	185
ANÁLISE ESTATÍSTICA COM BASE NAS PALAVRAS-OCORRÊNCIA	187
Comportamento estatístico do discurso científico	187
Comportamento estatístico do discurso jornalístico	188
Análise comparativa do comportamento lexicogênico em ambos os discursos	189
ANÁLISE ESTATÍSTICA COM BASE NOS VOCÁBULOS	196
Comportamento estatístico do discurso científico	196
Comportamento estatístico do discurso jornalístico	200

Comparação entre os perfis lexicogênicos em ambos os discursos com base em palavras-ocorrência e em vocábulos	203
Natureza e fluxo dos formantes – o Índice Etimológico	205
Esquema de difusão dos empréstimos	209
PALAVRAS DE CARACTERIZAÇÃO E VOCABULÁRIO	
FUNDAMENTAL	212
Palavras de caracterização	212
Vocabulário fundamental	213
CONSIDERAÇÕES ADICIONAIS SOBRE A ANÁLISE DOS RESULTADOS	
CIRCUNSTÂNCIAS HISTÓRICAS E OS RESULTADOS OBTIDOS	225
CONCLUSÃO	233
BIBLIOGRAFIA	237
GLOSSÁRIO	243
ILUSTRAÇÕES	251
TABELAS	251

APRESENTAÇÃO

Léxico e Ideologia na Europa Ocidental, do professor Aldo Bizzocchi, é um texto formalizador de um discurso científico que, como todo discurso, apresenta uma multiplicidade de sentidos possíveis a serem construídos pelo processo da leiturização. Contudo, não se deve ignorar o fato de que tais sentidos devem ter por ancoragem a reconstrução do espaço de significações construído pelo autor. Esse espaço de significações é aquele de um estudo que, situado no campo da filologia contemporânea, tem por sustentáculo o princípio da historicidade, delimitado à investigação do universo lexical, recortado por diferentes idiomas do indo-europeu e do anglo-saxão. Postulando ser a linguagem um fenômeno social, tanto do ponto de vista genético quanto do ponto de vista de sua natureza e/ou das condições em que ela se desenvolve, Bizzocchi a concebe como um processo complexo que torna possível a aquisição e o emprego de uma dada língua. A língua, concebida fora do uso, é compreendida como sistema autônomo e imanente que se explica pela descrição de um conjunto de elementos e de regras combinatórias, nos níveis fonológico, morfológico e sintático. Contudo, na dimensão do uso, a língua compreende um conjunto de traços antro-po-sócio-cultural-ideológicos que a caracterizam em face de outras línguas que lhe são cognatas, razão pela qual nessa dimensão ela se faz idioma, implicando a concepção de unidade lingüística de uma nação em relação às demais. Logo, o termo “idioma” só se refere a uma língua nacional, presupondo a existência de um estado político do qual ela é expressão de uma visão de mundo de um grupo humano. Dentro desse quadro, o autor, privilegiando o sistema lexical de diferentes idiomas de famílias lingüísticas diferenciadas, considera-os como espaço de estruturação e reestruturação de diferentes conhecimentos, de diferentes valores culturais e ideológicos que se aproximam e se

distanciam; logo, de convergências e divergências marcadas no quadro das formalizações léxico-gramaticais de idiomas distintos. Objetivando encontrar as causas dessas convergências e divergências, o autor, valendo-se do método comparativo, submete o material de amostra por ele selecionado — vocabulário de discursos

científicos, filosóficos e jornalísticos contemporâneos — a um tratamento qualitativo e quantitativo, tendo por critério de análise princípios da Lingüística Quântica e postulados da Lingüística do Discurso, numa interface com dados da Lexicologia. Esse quadro teórico possibilita-lhe revelar dados de extrema valia para o estudioso dos mais diferentes campos da linguagem.

Entende-se que, por não ser a lexia, o vocábulo compreende uma construção formal de conteúdos, filtrados de *continuum* de conhecimentos, arquivados na memória semântica a longo prazo dos interlocutores de uma dada comunidade lingüística, para expressar a visão de mundo de um grupo, enquanto verdade por definição. Logo, a lexia em estado de dicionário, isto é, o vocabulário, pode se remeter a alguns dados do conhecimento, organizados a partir de experiências coletivas e adquiridas pelo grupo. Nesse sentido, ao privilegiar pelo étimo vocábulos comuns de sistemas lexicais de idiomas diferenciados, Bizzocchi se defronta com dois aspectos relevantes: um mesmo discurso fundador a perpassar esses diferentes sistemas vocabulares e o modo como esse discurso fundador foi retomado por diferentes povos de diferentes formações antro-po-sócio-cultural-ideológicas, fundadoras de diferentes Estados que, delimitados por espaços político-geográficos distintos, se sustentam no tripé nação-cultura-idioma: uma só nação, uma cultura una, mas não necessariamente uniforme, falando uma só língua. Assim é que a permanência desse patrimônio lingüístico-cultural dá aos homens, organizados em grupos, a consciência da sua identidade nacional.

Na busca desse discurso fundador, o pesquisador se faz autor de um texto historiográfico que lhe permite entender o passado pelo olhar do presente, de modo a projetar seu ponto de vista para além de estudos meramente etimológicos de bases históricas propostos para o tratamento do vocabulário pela Gramática Histórica. Reconstrói, pelo processo de deriva, as bases dos sistemas vocabulares do português, do francês, do italiano, do inglês e do alemão e, presentes neles, encontra traços de formações discursivas e ideológicas do discurso do poder greco-latino, projetando-se e/ou

propagando-se na Europa Ocidental. Descobertas as causas da área de convergência, essas vão sendo apresentadas ao leitor pelo mesclar de conhecimentos científicos, filosóficos e litúrgicos aos “saberes” desses diferentes povos. Tem-se assim, pelo estudo de lexias institucionalizadas pelo uso (vocabulário), extraídas de discursos científico-filosófico-jornalísticos contemporâneos, o modo como cada um desses povos filtrou os diferentes sentidos de um mesmo discurso fundador, apossando-se de conhecimentos por ele veiculados e, ao mesmo tempo, do vocabulário que formalizava tais conhecimentos, adaptando-o ao sistema léxico-fonético de suas línguas maternas. Tal adaptação implicou e implica, necessariamente, o reorganizar do conteúdo sêmio-lingüístico original pelas próprias experiências coletivas de cada um desses povos. Tem-se, assim, a cultura greco-latina a perpassar toda a formação vocabular da Europa Ocidental, fazendo-se ali substrato de diferentes idiomas que a retomam por suas visões específicas de mundo. Assim, a etimologia e a neologia lexical se fazem âncora dos estudos lexicogênicos, propostos por Bizzocchi, a partir de um ponto de vista da lingüística contemporânea.

Ressalta-se que, ao privilegiar o sistema vocabular de diferentes idiomas, considerando-o como espaço formalizador de diferentes elementos lexicais, específicos de cada língua, Bizzocchi demonstra que a produção lexical de um idioma não se dá de maneira caótica, pois a dinâmica do enriquecimento desse universo é susceptível a um controle e a uma taxionomia, quanto à tipologia ou quanto aos processos que possibilitam o aparecimento de novas unidades em seu inventário. Por conseguinte, agora abandonando o eixo da historiografia e da socioetnologia que lhe possibilitaram a descoberta de um mesmo discurso fundador e de suas diferentes retomadas em regiões que delimitam hábitos, culturas e ideologias diferenciadas, o autor se volta para um trabalho eminentemente lingüístico para estabelecer tipologias, valendo-se de princípios da estatística. Quantificando e confrontando esses vocabulários entre si, descobre no sistema da flexão as regularidades apresentadas pelos paradigmas dos novos vocábulos, emprestados do sistema greco-latino e incorporados aos sistemas de cada um dos idiomas estudados. Defronta-se, então, com a regra dos vazios lexicais existentes no sistema lexical do português, do francês, do italiano, do inglês e do alemão decorrentes da importação de conceitos, idéias e/ou as coisas do mundo, o que

leva também à importação de vocábulos com que são designados nesses mesmos idiomas. Analisa as transformações morfofonêmicas, decorrentes do processo de incorporação de um vocábulo a um outro sistema lingüístico e descobre os idiomatismos: formas resultantes de processos analógicos que só se explicam à luz da história, da cultura e da ideologia de cada um dos povos, razão pela qual suas significações não decorrem dos vocábulos que as compõem e nem tampouco das regras sintáticas que os articulam. Por conseguinte, como o próprio étimo da palavra indica, o autor chega ao sistema particular dos processos de derivação de cada um dos idiomas estudados. Ali descobre a assistematização, ou seja, aponta que uma derivação pode aparecer em um dado vocábulo de um dos idiomas estudados e faltar para um vocábulo congêneres em um outro idioma, comprovando que o sistema de derivação não se estrutura, como é o caso da flexão, por paradigmas coesos. Assim, dentro de um mesmo sistema de derivação, há variação para a estruturação vocabular em cada um dos idiomas estudados. Logo, ao comparar a formação de vocábulos semelhantes em idiomas diferentes, Bizzocchi, por um lado, aponta dificuldades para sistematizá-los e, por outro, comprova que tais dificuldades decorrem do fato de ser o universo lexical o espaço onde o velho se redimensiona pelo novo, sendo o novo a revelação de identidades, de valores antro-po-sócio-cultural-ideológicos.

Jeni S. Turazza

ABREVIATURAS E SÍMBOLOS

al. = alemão

cat. = catalão

esp. = espanhol

fr. = francês

gr. = grego

ing. = inglês

it. = italiano

lat. = latim

nor. = norueguês

port. = português

rom. = romeno

sue. = sueco

>, < = evolução fonético-fonológica e/ou gráfica; derivação

→, ← = empréstimo entre línguas

≈ = equivalência entre formas ou vocábulos de duas línguas

[] = transcrição fonética, representando os sons da fala

// = transcrição fonológica, representando os fonemas da língua

* = forma hipotética ou não documentada

INTRODUÇÃO

A Europa Ocidental é talvez o melhor exemplo de povos étnica e lingüisticamente diferentes que, no entanto, devido a uma influência cultural comum — no caso, a influência da Grécia e de Roma, exercida sobretudo através de suas línguas (o grego e o latim) e de sua religião (o cristianismo) —, constituíram uma única grande civilização. A influência unificadora da cultura clássica reflete-se na ideologia desses povos e, conseqüentemente, em suas línguas, especialmente no que concerne ao léxico. A história dessas línguas sustenta-se, assim, numa constante tensão entre a influência greco-latina e o elemento vulgar.

Por outro lado, numerosos estudos têm sido empreendidos acerca do léxico dessas línguas, muitos dos quais visando à explicação dos processos de criação das palavras, à detecção de uma ideologia subjacente ao léxico, e mesmo à quantificação dos fenômenos observados, para isso recorrendo inclusive a análises estatísticas. Procuramos aqui abordar todas essas questões sob um novo enfoque, buscando, para tanto, explicar a dinâmica da criação e renovação do léxico das línguas européias ocidentais, pertencentes às famílias lingüísticas românica e germânica, de um ponto de vista ao mesmo tempo neológico e etimológico, isto é, com base na origem e nos processos de produção das unidades lexicais, a partir da supramencionada tensão *clássico/vulgar*, processos estes cuja tipologia pretendemos estabelecer.

Procuramos, igualmente, fazer uma análise quantitativa de tais processos, mediante a aplicação de métodos estatísticos, para, ao final, comparar o comportamento dessas línguas com base nos perfis dos respectivos léxicos. Tal estudo é particularmente relevante na medida em que tem como objeto o léxico das línguas dos dois grandes grupos étnicos que constituem o núcleo da Civilização Ocidental.

Como se sabe, parte do material léxico das línguas européias de cultura entrou nas mesmas por evolução direta a partir das línguas que as precederam historicamente (como é o caso do latim vulgar em relação às línguas românicas), parte resulta de importação a partir de outras línguas, e parte é fruto de criação intralingüística. A importação lexical pode, por sua vez, ser direta, isto é, a partir da própria língua na qual a palavra foi criada, ou indireta, quando aporta elementos léxicos a partir de línguas nas quais tais elementos também são resultado de importação. Cumpre ressaltar, ainda, que tanto as palavras resultantes de importação quanto aquelas engendradas intralingüisticamente estão igualmente sujeitas ao processo da evolução histórica.

Os elementos léxicos resultantes de importação ou de criação intralingüística passam inicialmente por um estágio de neologicidade, podendo a seguir ver sua freqüência de uso aumentada até o ponto de vir a integrar uma ou mais normas discursivas, ao mesmo tempo que unidades pertencentes às diversas normas da língua podem sofrer uma progressiva queda de freqüência, tendendo a tornar-se arcaísmos. No limite, tais unidades poderão deixar de pertencer à língua.

Nem sempre a criação neológica visa à introdução de signo destinado a dar conta de um novo recorte cultural, resultante do permanente processo de evolução histórica da comunidade social que fala uma determinada língua, mas, em muitos casos, restringe-se a propor um novo significante para um signo já existente. Em razão disso, formas de expressão novas, recém-introduzidas na língua, convivem, num primeiro momento, com formas preexistentes, passando em seguida a concorrer com as mesmas, podendo, eventualmente, suplantá-las.

Nosso princípio de análise baseia-se na tensão permanente entre as línguas clássicas do Ocidente — grego e latim — e as línguas ditas vulgares, na verdade línguas nacionais dos povos da Europa Ocidental, tensão esta que se encontra na base de toda a tradição literária, religiosa, filosófica, intelectual, e também filológica e gramatical do Ocidente. De fato, inúmeras razões fazem que o grego e o latim (aquele principalmente por intermédio deste) estejam em posição hierárquica superior em relação às diversas línguas vulgares, das quais constituem o grande paradigma. Embora o Império Romano do Ocidente tenha desmoronado frente às invasões bárbaras em 476, seu desaparecimento enquanto unidade

política não significou a extinção de sua influência cultural sobre os povos da Europa. As instituições políticas, os fundamentos jurídicos, os cânones literários e artísticos da Roma Antiga, bem como a própria língua latina, sobreviveram até a era moderna. Todas as vezes que a civilização européia procurou reerguer-se, como à época de Carlos Magno, ou durante o Renascimento, sua fonte de inspiração foi a civilização romana, que, por sinal, representava, para a Europa Ocidental, a própria civilização antiga, uma vez que, até o século XVI, o conhecimento que se tinha a respeito da Grécia Antiga se dava por meio da língua latina. Além de representarem o veículo de transmissão do patrimônio cultural da Antigüidade, as línguas clássicas, principalmente o latim, foram, durante muito tempo, consideradas as únicas línguas dignas da literatura, da poesia, da ciência, da filosofia, da religião, etc. Nesse sentido, o epíteto “vulgar” dado aos dialetos populares medievais possuía uma forte carga pejorativa. De fato, tais dialetos não eram vistos como línguas, no sentido em que hoje entendemos tal termo, e só tardiamente, e a muito custo, lograram estes impor-se como línguas literárias e, mais tarde, como idiomas nacionais. De qualquer maneira, vê-se que as diversas línguas vulgares mantiveram entre si uma posição de igualdade, de equivalência, e, ao mesmo tempo, uma posição de submissão cultural à influência greco-latina. Ainda que algumas línguas pudessem, na Idade Média e na Renascença, desfrutar de algum prestígio cultural, como é o caso do provençal da poesia trovadoresca, nenhuma língua rivalizava com o grego ou o latim. Mesmo atualmente, quando essas línguas já não exercem mais uma hegemonia absoluta sobre as demais línguas do Ocidente, sua influência está longe de ser desprezível. Com efeito, ainda hoje, a maioria dos termos técnicos, científicos, ou simplesmente os vocábulos de cultura de qualquer língua ocidental são empréstimos das línguas clássicas ou estão formados com elementos morfológicos dessas línguas e segundo seus modelos.

Entretanto, a enorme influência cultural das línguas clássicas, enquanto repositório da cultura greco-romana e cristã, deixou marcas desse estatuto hierárquico superior, que acima mencionamos, também ao nível estritamente lingüístico. Conforme dissemos, as línguas clássicas representam o grande acervo de elementos léxicos que supre as necessidades de expressão das diversas normas de cultura das línguas européias ocidentais. Assim, quan-

do se introduz numa determinada língua um grecismo ou latinismo, este não é em nenhum momento sentido como um vocábulo estrangeiro (tomemos como exemplo os vocábulos portugueses *organismo* e *temperatura*). Na verdade, o mesmo já é introduzido na língua com todas as adaptações fonológicas, morfológicas e ortográficas que lhe garantem, desde o início, o estatuto de unidade genuinamente pertencente ao léxico da língua vulgar. Já um vocábulo proveniente de uma outra língua qualquer, seja ela europeia ou não, entra no léxico da língua sempre como um estrangeirismo, o que é revelado pela grafia, pela pronúncia e pela morfologia, e somente mais tarde poderá vir a perder esse caráter estrangeiro (por exemplo, ing. *football* → port. *futebol*, fr. *abat-jour* → port. *abajur*). É nesse sentido que podemos afirmar que um empréstimo ao grego ou ao latim, mesmo quando vulgarizado, não tem o mesmo estatuto de um empréstimo a uma língua vulgar. Dado que todas as línguas do Ocidente sofreram — e ainda sofrem — a influência das línguas clássicas, e que, ao mesmo tempo, todas elas intercambiam elementos léxicos clássicos ou vulgares entre si, a opção de uma língua, em face de uma determinada necessidade de criação neológica, entre um vocábulo greco-latino e um vocábulo vulgar estrangeiro, entre a importação e a criação intralingüística, a partir de material léxico autóctone, pode revelar algo a respeito da própria ideologia dessa língua, e, conseqüentemente, do povo que a fala. Eis por que nenhuma análise etimológica do léxico das línguas europeias ocidentais pode ignorar essa tensão importantíssima entre o clássico e o vulgar, entre o nacional e o importado. Se é verdade que há inúmeras línguas no mundo que jamais sofreram algum tipo de influência externa, tais como certas línguas aborígenes da América e da África, em relação às quais nossas considerações de ordem lexicogênica não teriam nenhum valor, as línguas do Ocidente configuram um caso único de línguas diferentes, cujo parentesco é por vezes remoto, formadas no entanto segundo um mesmo modelo, que confere às mesmas um verdadeiro caráter de unidade cultural, histórica e, sobretudo, ideológica, uma vez que fazem parte da mesma civilização.

Em vista disso, qualquer análise do léxico dessas línguas deve estar baseada na natureza *lexicogênica* das palavras, isto é, deve ter como ponto de vista sua origem e as transformações e adaptações fonético-fonológicas, morfológicas e semânticas por elas sofridas, tanto no momento de sua introdução na língua quan-

to ao longo de sua evolução. Tal abordagem do léxico é, conforme dissemos anteriormente, simultaneamente neológica e etimológica, e conduz a uma classificação das unidades lexicais da língua segundo sua tipologia genética. A partir dessa classificação, pode-se estabelecer um perfil lexicogênico de cada língua. Uma análise quantitativa e qualitativa desse perfil permite estudar a dinâmica neológica da língua em questão, bem como identificar suas tendências predominantes, suas preferências, suas coerções, e, conseqüentemente, correlacionar tal comportamento a uma “visão de mundo” subjacente.

Procuramos, assim, construir um modelo teórico que desse conta dos principais fenômenos lexicogênicos conhecidos, ao menos no domínio das línguas ocidentais; para tanto, incorporamos os processos neológicos já descritos em trabalhos clássicos de lexicologia, bem como as categorias classificatórias constantes das obras filológicas mais importantes e de grandes dicionários etimológicos, ao mesmo tempo que redefinimos alguns conceitos e instituímos outros.

Cumpramos notar que a maior parte das pesquisas etimológicas nas línguas românicas e germânicas foram levadas a efeito ainda no século passado, por lingüistas e filólogos da corrente histórico-comparativa. Desde então, o *status quaestionis* evoluiu muito pouco, permanecendo até hoje em vigor os conceitos e categorias classificatórias estabelecidos naquele período. A quase totalidade dos trabalhos de filologia e lingüística diacrônica realizados no presente século consistem basicamente de aplicações dos modelos teóricos desenvolvidos no século XIX a *corpora* específicos, ao passo que trabalhos de revisão da teoria vigente praticamente inexistem, embora, a nosso ver, o arsenal teórico de que hoje dispomos já se tenha mostrado insuficiente, na medida em que não dá conta do aspecto dinâmico dos processos de criação lexical, nem tampouco os analisa do ponto de vista ideológico.

CAPÍTULO I

NOÇÕES PRELIMINARES

ETIMOLOGIA E LEXICOGÊNESE

Entende-se por *etimologia* o estudo da origem das palavras, a busca do *étimo*, palavra considerada como ancestral daquela que constitui o objeto da pesquisa etimológica. Essa definição já é, por si mesma, insuficiente. Dada, por exemplo, a palavra portuguesa *ópera*, seu *étimo* deve ser buscado no it. *opera*, no lat. *opera*, do qual a forma italiana provém por herança, ou, ainda, numa hipotética raiz indo-européia **ə3opos*? Tal indagação não é de forma alguma irrelevante, visto que só poderemos compreender o que seja a lexicogênese na medida em que conseguirmos distingui-la da etimologia propriamente dita.

Longe de ser a “ciência do verdadeiro sentido das palavras” (do grego *ἔτυμος*, “verdadeiro”, e *λόγος*, “ciência”), como pretendia Dionísio de Halicarnasso, criador do termo, a etimologia tornou-se, ao longo do tempo, a disciplina que estuda a história das palavras. Porém, desde cedo, colocaram-se obstáculos de difícil transposição à tarefa do etimólogo: descrever a história de uma palavra é detectar o momento exato de seu aparecimento no léxico da língua, rastreando, em seguida, todo o seu percurso ao longo do tempo, até o presente? Em caso afirmativo, como detectar tal momento? Com efeito, na maioria das vezes, os etimólogos e filólogos só têm acesso ao léxico através da modalidade escrita da língua. Sabe-se, por outro lado, que muitas palavras de introdução antiqüíssima, e mesmo vernáculas, só foram tardiamente registradas em documentos escritos. Além disso, toda língua tem a capacidade de gerar sempre novas unidades léxicas a partir de outras, já existentes, e o número de criações internas a qualquer língua ultrapassa, de muito, o número de unidades primitivas (vernáculas) ou importadas de outras línguas. Assim, a história de uma palavra como *guarda-roupa* se inicia a partir do momento em que duas unidades léxicas autônomas, *guardar* e *roupa* se juntam

para formar o composto, ou antes, nas ancestrais germânicas *wardÇn* e *raubÇ*, ou, ainda, no composto francês *garde-robe*, no qual a palavra portuguesa se inspirou? As coisas se complicam ainda mais quando introduzimos na discussão acerca da história das palavras o problema do significado. Muitas palavras mudam de significado ao longo do tempo, da mesma forma que muda o significante. Assim, conforme procuraremos demonstrar neste trabalho, toda pesquisa que ignore o aspecto semântico de uma palavra correrá o risco de agrupar numa mesma classe objetos muito diferentes. Por essa razão, devemos ter como um princípio fundamental o fato de que o verdadeiro objeto de qualquer estudo etimológico e lexicogênico deve ser o vocábulo, unidade de norma, monossêmico e, muitas vezes, monoisotópico.

Está claro, também, que muito mais importante do que saber-se onde surgiu um determinado elemento léxico é saber-se como, e por quê, surgiu tal elemento. Tal observação é importante para que possamos distinguir entre a etimologia propriamente dita e a lexicogênese. A primeira se ocuparia da origem das palavras, numa perspectiva estritamente diacrônica, e de um ponto de vista exclusivamente fonético. Já a segunda estudaria os processos e mecanismos de criação e renovação lexicais, o percurso através do qual as palavras penetram no léxico, procedendo a uma classificação das unidades léxicas da língua segundo sua natureza genética. Dito de outra maneira, a etimologia pesquisaria os antepassados das palavras, ao passo que a lexicogênese estudaria a gestação, o nascimento, a vida e, eventualmente, a morte dos vocábulos no seio da língua. Desse modo, unidades de mesmo étimo serão diferentemente consideradas conforme os diferentes modos através dos quais chegaram ao léxico. Vê-se assim que, enquanto a etimologia possui vocação eminentemente diacrônica, a lexicogênese é uma disciplina de índole sincrônica, que, entretanto, não hesita em recorrer à diacronia para explicar o porquê dos fenômenos sincronicamente localizados que presidem à dinâmica da criação e renovação lexicais. Juntas, elas poderiam con-

figurar uma espécie de “lexicologia pancrônica” (ver Análise sincrônica x análise poancrônica, p. 46).

A EUROPA OCIDENTAL

Poder-se-ia afirmar que, sob todos os pontos de vista, a Europa se constitui numa unidade. Sendo o menor dos continentes, possui algumas particularidades: todos os seus habitantes nativos pertencem à raça branca, fazem parte da mesma civilização, professam, salvo algumas pequenas minorias religiosas (judeus, muçulmanos), uma única religião, o cristianismo, através de suas várias confissões ou seitas, e, com raras exceções, falam línguas pertencentes ao mesmo tronco lingüístico, o indo-europeu. Todavia, apesar dessa aparente unidade, desde o início a história da Europa se viu marcada por um contraste entre Ocidente e Oriente, que perdura até os dias de hoje.

O continente europeu emergiu tardiamente das brumas da Proto-História. A primeira grande civilização genuinamente européia foi a Civilização Grega, que pela primeira vez na Antigüidade implantou uma cultura do homem, baseada numa visão antropocêntrica do mundo, de que decorreram o aperfeiçoamento — senão a própria instituição — do conceito de Estado, bem como a gênese do pensamento racional, marca indelével da cultura ocidental até hoje. Realizando a separação entre o mundo natural e o sobrenatural, e propondo-se a dominar o primeiro através do uso da razão, bem como incorporando as técnicas e aquisições culturais do Oriente (o alfabeto, a moeda), os gregos fizeram que o centro de gravidade do mundo antigo rapidamente se deslocasse da Ásia Central e Ocidental para a Europa.

A Civilização Grega foi historicamente sucedida por Roma, que não só herdou a cultura helênica, ampliando-a com suas contribuições originais (por exemplo, o direito romano), mas que teve como principal mérito o haver estendido essa cultura através da constituição de um grande império, que uniu povos muito diferentes, dando-lhes uma mesma cidadania e uma mesma língua, e abrindo assim caminho para a edificação de uma verdadeira civilização européia. Esse império universal viria a ser o grande vetor da fé cristã, elemento espiritual que garantiu a unidade cultural

européia mesmo após a ruína política do Estado romano. Encontrava-se assim a Europa dividida entre o universo greco-romano-cristão e o mundo “bárbaro”, que mais tarde viria a ser assimilado por aquela cultura, dando origem à verdadeira Civilização Européia ou Ocidental.

Como se vê, a civilização difundiu-se na Europa a partir de dois grandes centros, a Grécia e Roma, o primeiro mais sintonizado com as culturas asiáticas e o segundo mais voltado para suas conquistas a Oeste. Assim, desde logo, o Império Romano se estruturou com base numa dualidade de línguas e culturas. Na parte oriental do Império se falava grego, ao passo que a língua do Ocidente era o latim. Além disso, as desigualdades econômicas entre as duas regiões conduziram a uma diferenciação progressiva entre ambas que culminou, após a morte do imperador Teodósio, em 395, com a divisão do Estado em Império do Ocidente e Império do Oriente.

O Império Romano do Ocidente, menos desenvolvido e militarmente enfraquecido, caiu em 476 diante dos sucessivos ataques bárbaros, ao passo que seu irmão oriental, o Império Bizantino, sobreviveria ainda por mil anos. Estava definitivamente selada a divisão da Europa. A partir daí, os dois ramos da Civilização Européia iriam seguir caminhos separados, quando não opostos, que apenas esporadicamente voltariam a se cruzar.

Em primeiro lugar, a divergência religiosa entre o Papa, bispo de Roma, e o Patriarca Bizantino, bispo de Constantinopla, em torno da primazia espiritual e temporal sobre a Cristandade levou ao Cisma do Oriente (1054), através do qual nasceu a Igreja Ortodoxa, a partir de então separada da Igreja Católica Apostólica Romana. Enquanto o Ocidente permanecia fiel ao catolicismo, o Oriente abraçava sua nova confissão. Tal fato teve também uma importante repercussão lingüística, na medida em que a língua oficial da ortodoxia era o grego, ao passo que o idioma da Igreja de Roma continuava a ser o latim. Essas duas línguas iriam influenciar sobretudo o léxico das línguas vulgares de suas respectivas esferas de abrangência, e deixariam seus traços inclusive no alfabeto: no Ocidente, as línguas vulgares passaram a ser escritas através do alfabeto latino; no Oriente, prevaleceram os alfabetos grego e cirílico, este derivado daquele. Isso não significa que a Europa Ocidental não tenha sofrido também uma considerável influência da cultura helênica. Mas, sem dúvida, tal fato se deu

indiretamente, através do latim, enquanto no Oriente essa influência foi direta.

Entretanto, as divergências entre Europa Ocidental e Oriental são ainda mais profundas: no Ocidente, desenvolve-se um novo sistema político e econômico, o feudalismo, a partir do século IX; no Oriente, tal sistema chegaria tardiamente, e apenas a algumas regiões. Por outro lado, tendo sido progressivamente substituído nos países ocidentais pelo capitalismo, a partir de fins da Idade Média, o feudalismo sobreviveu no Leste Europeu, sobretudo na Rússia, até o início do presente século. O sistema capitalista, por sua vez, tendo sua origem associada à da própria burguesia emergente por volta dos séculos XII e XIII, só começa a dar timidamente seus primeiros passos nos países do Leste a partir da presente década, com o desmoronamento da União Soviética. Pode-se dizer, assim, que a Europa Oriental passou diretamente do regime feudal para o socialismo, sem nunca ter conhecido verdadeiramente o sistema capitalista. A dicotomia *capitalismo/socialismo*, que dominou toda a história política do século XX, é mais um traço diferenciador das duas regiões.

Também a Revolução Industrial, iniciada na Inglaterra em meados do século XVIII, e difundida nos demais países ocidentais ao longo do século XIX, jamais chegaria aos países do Leste. Com efeito, muitos deles só vieram a iniciar sua industrialização a partir das revoluções socialistas deste século, enquanto outros ainda se encontram sob uma economia essencialmente agrícola.

Do ponto de vista político, são também profundas as diferenças entre o Leste e o Oeste. Enquanto as nações da Europa Ocidental vivenciaram a partir do século XVIII grandes transformações políticas e sociais, com o fim do absolutismo, derrubado por várias revoluções de índole liberal, que culminaram com a implantação de regimes democráticos e, em grande parte, republicanos, as ex-repúblicas comunistas somente agora começam a adotar a democracia como regime de governo.

No aspecto cultural, a Europa Ocidental foi palco de vários movimentos que jamais atingiram, senão muito levemente em alguns casos, o Leste Europeu. É o caso dos estilos artísticos românico e gótico, com suas majestosas catedrais, dos quais foi contemporâneo no Oriente o estilo bizantino. É o caso igualmente do Renascimento, muito mais do que um mero movimento de renovação estética, uma verdadeira revolução intelectual, econô-

mica, política e social, que marcou a retomada da civilização greco-romana, aliada à busca do progresso filosófico e científico. Durante esse período de fundamental importância para a história européia e mundial, a Europa Oriental encontrava-se em grande parte sob o jugo otomano.

Um dos fatores de desenvolvimento da Europa Ocidental que assinalam o contraste entre Leste e Oeste é a unidade política dos Estados ocidentais, consumada ainda em fins da Idade Média, enquanto na Europa Oriental encontramos ainda hoje uma colcha de retalhos de etnias, que ao longo da História foram integradas aos mais diversos reinos e impérios, e que mais tarde deram origem a Estados artificialmente construídos, em virtude de partilhas políticas consecutivas às grandes guerras, como é o caso da União Soviética, da Iugoslávia e da Tchecoslováquia, que recentemente se dissolveram. Em função disso, o Leste Europeu ainda se constitui num cadinho de grandes conflitos étnicos e políticos.

Vê-se assim, claramente, a existência de duas Europas, ambas nascidas da herança greco-romana, ambas participantes da mesma civilização, porém cada uma com suas diferenças históricas e ideológicas específicas. Do ponto de vista étnico, observamos que a chamada Europa Ocidental reúne basicamente os povos latinos (com exceção dos romenos), célticos e germânicos, ao passo que a Europa Oriental congrega eslavos, bálticos, gregos, ilírios e ugro-fineses. Na esfera ocidental, temos um certo equilíbrio entre os grupos latino e germânico em termos de densidade demográfica, distribuição territorial, variedade lingüística, importância histórica e influência cultural, enquanto os celtas, menos numerosos e menos destacados ao longo da História, parecem ter sido quase completamente absorvidos por esses dois grupos. Já na Europa Oriental prevalecem duas grandes culturas: a grega e a eslava. Entretanto, não existe entre elas o equilíbrio verificado entre o mundo latino e o germânico. A Europa Oriental não possui verdadeiramente uma civilização unitária.

Entre as duas divisões da Europa existe uma zona de transição, composta de povos que gravitavam entre o Ocidente, representado pelo Sacro Império Romano-Germânico e depois pelo Império Austro-Húngaro, e o Oriente, materializado pelo Império Russo. São eles os poloneses, húngaros, tchecos, boêmios, croatas, etc. Por essa razão, os eslavos não possuem unidade cultural,

já que parte deles é católica e parte, ortodoxa; parte das línguas eslavas sofreram influência latina e parte, influência grega.

De tudo o que foi dito até agora, percebe-se que a Europa Ocidental é o resultado do cruzamento étnico-cultural entre o mundo romano e o mundo germânico, sob a influência ideológica da tradição judaico-cristã, de que resulta uma civilização que poderíamos chamar de romano-germano-cristã. A Europa Oriental, por outro lado, resulta da combinação da cultura eslava com a cultura grega tardia, contaminada por sua vez de elementos asiáticos, igualmente sob a égide da religião cristã. Pode-se falar assim de uma civilização greco-eslavo-cristã.

AS LÍNGUAS EUROPÉIAS OCIDENTAIS

As Civilizações Grega e Romana surgiram de forma praticamente espontânea, quando populações de origem indo-européia vindas da Ásia Menor e instaladas em seus nichos europeus (Grécia e Itália) deram início ao grande desafio de passar da vida rural à vida urbana, e de organizar o meio social em torno da noção de Estado. Tais civilizações não se inspiraram, portanto, em nenhum modelo civilizatório anterior (especialmente no caso da Grécia). Assim, todo o léxico de cultura do grego e do latim, isto é, todo o conjunto de palavras pertencentes à esfera das atividades intelectuais, se originou sobretudo por criação intralingüística. No caso do latim, parte do léxico de cultura foi criado intralingüisticamente e parte se originou de empréstimos ao grego.

Já a Civilização Ocidental é uma continuação histórica da Civilização Greco-Romana, acrescida de aportes culturais judaico-cristãos e germânicos. Pode-se dizer que a Civilização Ocidental nasceu dos escombros da Civilização Greco-Romana e os poucos valores culturais que manteve foram herdados de Roma através da Igreja. Assim, quando ao final da Alta Idade Média houve um primeiro renascimento cultural, o único modelo de civilização disponível era o greco-romano. Por isso, as primeiras palavras cultas das línguas vulgares européias foram importadas do latim ou do grego, ou cunhadas segundo o modelo das línguas clássicas.

Como se sabe, as línguas românicas ou neolatinas descendem diretamente do latim, numa evolução constante que em ne-

num momento sofreu solução de continuidade. Tal fato suscita a questão de sabermos em que instante e por que razão se estabelece o limite entre a língua latina e suas sucedâneas históricas. Assim, antes de prosseguirmos nossa explanação acerca das línguas vulgares européias ocidentais por oposição às línguas clássicas, é prudente definirmos alguns conceitos, que nos serão extremamente úteis mais adiante, quando passarmos a discutir a questão acima colocada. Para tanto, é particularmente ilustrativo o que diz o lingüista italiano V. Pisani:

“O conceito de ‘língua’ é uma abstração baseada numa série de fatos reais que são os atos lingüísticos isolados, unidades de expressão correspondentes a unidades de intuição dos falantes. Mas estes atos não são imagináveis fora da sociedade, na qual atuam todas as capacidades humanas, e na qual aparecem como comunicações (um reflexo das quais são também os colóquios do indivíduo consigo mesmo, seus pensamentos dialéticos, isto é, de determinadas intuições); comunicações que podem ter lugar somente na medida em que o indivíduo que fala cria a sua expressão servindo-se de certos signos (palavras e nexos) de que o ouvinte recria uma intuição mais ou menos correspondente àquela de seu interlocutor. Para que se desenvolva esta dupla atividade por parte de quem fala e de quem escuta, é preciso que ambos atribuam aos signos usados, os quais não têm por si sós nenhuma relação necessária com seu significado, um valor fundamentalmente igual. (...) Mas, naturalmente, servindo a cada vez para exprimir uma nova intuição, o signo não pode jamais ter um valor idêntico àquele do seu modelo, de onde um duplo motivo de inovação: modificação do valor, mesmo reproduzindo-se substancialmente inalterada a forma fônica do modelo, ou modificação também da forma fônica deste.

“Só com uma certa aproximação se pode, pois, falar de identidade dos signos contidos nos atos dos falantes isolados de uma dada comunidade lingüística. Mas é na admissão desta identidade que se baseia o nosso conceito de ‘língua’, o qual abrange os signos comuns aos atos lingüísticos de uma determinada quantidade de indivíduos, ou, se se preferir, o sistema de isoglossas que reúne tais atos lingüísticos, de modo que ele pode assumir âmbito maior ou menor, temporal e espacialmente, conforme os atos isolados considerados. (...) Dada a natureza do conceito de ‘língua’, é claro (...) que a determinação de certas ‘línguas’ se faz

com base em outros princípios que não estritamente glotológicos: motivos históricos, culturais, tradicionais nos fazem decidir falar como um todo da língua latina do século III ou talvez do século VI a. C. até o século V–VI d. C. e de uma língua italiana do século XIII aos dias de hoje, e não por exemplo de toda uma língua abrangendo a tradição latina e neolatina (...).¹

Tradicionalmente, divide-se a história lingüística do latim em três etapas: latim arcaico, latim clássico e baixo latim ou latim medieval. Paralelamente ao latim clássico e medieval, de caráter culto e literário, falava-se o latim vulgar, do qual derivam as línguas românicas. Esse latim vulgar, na verdade, nunca se constituiu numa língua fixa e estável. Com efeito, era falado ao longo de uma grande extensão territorial, que incluía a Ibéria, a Gália, a Itália, os Alpes, a África do Norte, etc., e, em cada uma dessas regiões, se havia sobreposto a uma língua preexistente, a que chamamos língua de substrato, falada pelas populações autóctones antes da conquista romana, e que, por vezes, deixou traços na fonética, na morfologia e no léxico do latim dessas populações. É, pois, natural que a diversidade dialetal do latim vulgar fosse bastante significativa. Não obstante, enquanto o Império Romano sobreviveu, a comunicação entre as diversas províncias permaneceu suficientemente intensa para impedir que divergências dialetais se aprofundassem a ponto de romper a possibilidade de intercompreensão. Já a partir do definitivo esfacelamento do Império, no século V de nossa era, as comunicações se tornaram mais difíceis, as relações comerciais entre as províncias cessaram, e a cultura literária, outro importantíssimo fator de união entre os falantes do latim, experimentou um período de extrema decadência. Some-se a isso a invasão dos territórios romanizados por povos germânicos que, embora em muitos casos adotando a língua latina — ou melhor, seus dialetos locais —, aportavam a esses dialetos seus hábitos lingüísticos próprios, contribuindo assim para aumentar a diversificação lingüística entre as regiões. As línguas germânicas representavam, desse modo, línguas de substrato em relação ao latim vulgar das províncias invadidas. Inicia-se assim um novo período na história da língua, em que não mais

¹ PISANI, V. *Introduzione alla linguistica indeuropea*. Roma, Edizioni Universitarie, 1939, p. 9 s.

convém falar-se de um latim vulgar, mas sim de um romance, ou, antes, de vários romances.

O período romance principia com a queda do Império Romano do Ocidente e se estende até o momento em que alguns de seus dialetos ascendem à posição de línguas literárias e de cultura, passando os demais dialetos a ser vistos não mais como dialetos do latim ou do romance, mas sim como variedades regionais das línguas literárias utilizadas nos territórios em que são falados.

Percebe-se assim que a história lingüística latina pode ser dividida didaticamente em três grandes períodos: um período latino vulgar, um período romance e um período neolatino, que por sua vez se dividirá em antigo (séculos IX a XI), médio (séculos XII a XV) e moderno (a partir do século XVI).

Uma divisão historiográfica semelhante pode ser aplicada ao domínio das línguas germânicas. É bem verdade que as obras tradicionais de filologia germânica distinguem apenas dois períodos, um chamado de germânico comum, e outro em que já aparecem os testemunhos escritos dos dialetos germânicos, dando origem ao período literário. Contudo, se observarmos mais atentamente a história dos povos germânicos, veremos que muito do que dissemos acerca da tradição latina em suas várias fases se aplica também ao germânico. Em primeiro lugar, o germânico primitivo, também chamado de proto-germânico, embora nunca tenha sido documentado, tendo sido reconstruído por comparação com as demais línguas indo-européias e a partir das evidências de suas línguas sucessoras, apresenta uma fase arcaica, comumente denominada pré-germânico, postulada pelos germanistas para explicar a transição entre o indo-europeu e o germânico propriamente dito, e uma fase de aparentemente maior estabilidade lingüística, em que os principais traços das línguas germânicas já estão consolidados: trata-se do acima citado germânico comum. A exemplo do latim vulgar, o germânico comum provavelmente nunca foi uma língua unitária, mas antes um conjunto de dialetos de origem indo-européia portadores de isoglossas comuns, que os aproximavam mutuamente e, ao mesmo tempo, os distinguiam dos demais dialetos indo-europeus: dialetos itálicos, célticos, gregos, bálticos, etc. Nessa etapa, tais dialetos eram ainda suficientemente próximos, em virtude de as tribos que os falavam se encontrarem circunscritas a um território relativamente pouco extenso. Entretanto, a partir sobretudo do século V, impelidos pela invasão dos

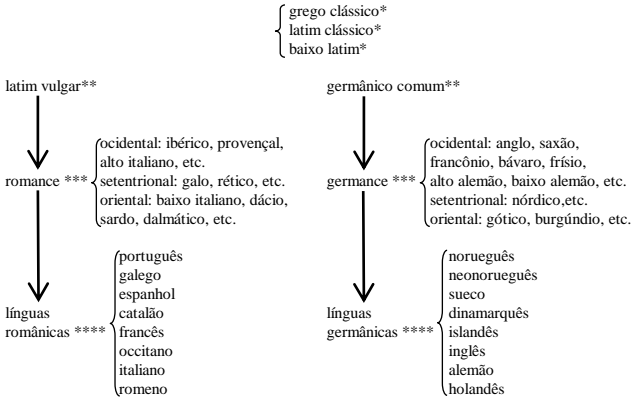
hunos de Átila vindos do Oriente, os germanos dispersaram-se progressivamente, especialmente nas direções oeste e sul, havendo inclusive algumas tribos invadido o território romano e fundido-se aos povos aborígenes, como é o caso dos francos na Gália. A partir desse momento, a diversidade lingüística entre os germanos também começa a intensificar-se, de modo a podermos dizer que tal fase de sua história se compara ao período romance, razão pela qual propomos aqui denominar o germânico dessa época de *germance*, por analogia ao romance.

O período *germance* se inicia, portanto, a partir da grande onda de invasões e incursões realizadas pelos germanos ao Império Romano (especialmente durante os séculos IV e V) e perdura até o início da tradição literária das línguas germânicas.

Em síntese, propomos a divisão da história do germânico em três etapas distintas, a saber: um período germânico comum, um período *germance* e um período neogermânico.

A postulação de um estágio intermediário entre o germânico comum e as línguas germânicas historicamente documentadas, que estamos aqui denominando período *germance*, é, a nosso ver, de suma importância para o estudo da lingüística germânica, uma vez que muitos fenômenos fonéticos, morfológicos e léxicos comumente atribuídos ao germânico comum ocorreram na verdade já no período *germance*, como, por exemplo, a passagem de *ē* germânico a *ā* e o empréstimo de vocábulos gregos e latinos, contemporâneo da intensificação do contato entre romanos e germanos.

Conforme dissemos na Introdução, o grego e o latim possuíam um estatuto hierárquico superior em relação às línguas vulgares, tanto românicas quanto germânicas. Estas se desenvolveram como línguas literárias segundo o modelo greco-latino. Desse modo, o parentesco existente entre o latim clássico e as línguas românicas, por ser indireto, já que, como vimos, as mesmas na verdade provêm do latim vulgar, não torna tais línguas mais propícias ao influxo das línguas clássicas do que o seriam as línguas germânicas. Nesse sentido, o efeito da influência greco-latina é exatamente o mesmo sobre ambas as famílias lingüísticas. Esquemmatizemos, então, o processo evolutivo dessas línguas e a relação existente entre as mesmas e as línguas clássicas da seguinte maneira:



* línguas clássicas → influenciam a norma culta das línguas vulgares

** línguas não documentadas diretamente

*** línguas pouco documentadas (séculos V a IX)

**** línguas plenamente documentadas (a partir dos séculos VIII–IX)

O período romance/germance é de fundamental importância para a gênese da Civilização Ocidental, pois é exatamente nessa etapa histórica que se dá o cruzamento entre as duas culturas básicas que compõem essa civilização: a cultura greco-romano-cristã e a cultura germânica. Durante o período latino/germânico primitivo, a maior parte do acervo léxico dessas línguas era autóctone, com uns poucos empréstimos feitos ao céltico e ao grego (no caso do latim). Na fase romance/germance, o intercâmbio de valores culturais e ideológicos (cristianismo, feudalismo, cavalaria, amor cortês, direito consuetudinário, etc.) implicou o intercâmbio de palavras. Nesse momento, os diversos romances se povoam de expressões germânicas (por exemplo, port. *guardar*, *dançar*, etc.), ao mesmo tempo que os germances recebem contribuições léxicas românicas (por exemplo, ing. *cup*, *dish*, etc.).

O ASPECTO IDEOLÓGICO DO LÉXICO

Conforme dissemos anteriormente, o léxico das línguas europeias ocidentais se sustenta basicamente sobre as dicotomias *autóctone/importado* e *clássico/vulgar*. Isso significa que tais

línguas sofrem influências verticais do grego e do latim, línguas que lhes servem de paradigma, e influências horizontais das demais línguas vulgares européias, bem como de línguas extra-européias. Em princípio, essas influências se exercem de igual maneira em todas as línguas consideradas. No entanto, o léxico de cada uma delas vai apresentar um comportamento diverso, vale dizer, vai criar ou renovar suas unidades segundo processos lexicogênicos diferentes, de que resultam léxicos de configuração bastante desigual. Torna-se evidente, então, que cada língua realiza certas opções e demonstra certas preferências por um ou outro processo de criação lexical. Essas opções e preferências obviamente mudam ao longo do tempo, de modo que o aspecto do léxico de uma língua numa determinada etapa sincrônica de seu desenvolvimento é o resultado de todas as tendências lexicogênicas verificadas nessa língua desde o início de sua história até aquele momento. Tais tendências, historicamente delimitadas, são responsáveis pela feição que o léxico apresenta a cada novo corte feito em sua história. Mas, mais do que isso, determinam de certa forma o porvir desse léxico, na medida em que muitos dos traços que o caracterizam se repetem de modo sistemático, tornando-se com o tempo injunções em nível de norma ou mesmo de sistema. Assim, se num primeiro momento a escolha em favor de um determinado processo lexicogênico, quando da introdução de um novo vocábulo, é totalmente livre e resulta exclusivamente do arbítrio do criador mesmo desse vocábulo, a partir do instante em que tal escolha, inspirada em modelos anteriores (vale dizer, em vocábulos dotados das mesmas características previamente criados), começa a sofrer algum tipo de coerção em termos de norma, dela resultando uma forma fonético-fonológica e/ou morfológica “recomendada”, passam a existir, ainda que informalmente, regras de criação lexical que, de certo modo, vão condicionar as perspectivas futuras do léxico em questão. É nesse sentido que podemos dizer, conforme o fizemos mais acima, que tais escolhas, tais coerções positivas e negativas da norma, dizem respeito à própria ideologia do léxico de uma língua, especialmente no que tange aos elementos léxicos de origem externa, pois é sobretudo em face da influência estrangeira, particularmente de línguas que exercem o papel de paradigmas culturais em relação a outras, que o caráter ideológico do léxico vai sobressair. Esse comportamento lexical ideologicamente orientado, firmando-se historicamente através da

repetição temporal, e socialmente através da difusão e da proliferação das formas que engendra, acaba, como vimos, por cristalizar-se no seio do sistema. Pode-se falar então de uma *ideologia lexical cristalizada*.

Além disso, sendo a escolha dos processos lexicogênicos uma coerção as mais das vezes em nível de norma, é de se esperar que cada um dos conjuntos vocabulários correspondentes aos vários universos de discurso que compõem o sistema léxico da língua apresente um comportamento lexicogênico próprio daquele universo de discurso específico. Desse modo, a ideologia lexical da língua seria uma espécie de média dos comportamentos lexicogênicos ideologicamente orientados dos seus universos de discurso.

Assim sendo, não é o material léxico vulgar puro, isto é, o conjunto das unidades herdadas ou criadas intralingüisticamente, que nos vai revelar a ideologia da língua, mas sim o material de origem estrangeira, seja ela greco-latina ou vulgar. Com efeito, do ponto de vista do material vulgar puro, as línguas européias não se diferenciam. É a escolha feita por cada língua entre conservar um empréstimo em sua forma original ou de algum modo adaptá-lo fonético-fonológica ou morfologicamente a fim de torná-lo mais familiar aos falantes da língua, mais próximo, pois, das palavras genuinamente nacionais, que vai diferenciar as línguas em termos lexicogênicos. Mais ainda, uma vez tendo-se decidido pela nacionalização do empréstimo, qual dentre os vários processos disponíveis para realizá-la será o escolhido é outra característica individualizadora de cada língua.

Como vimos, o grego e, principalmente, o latim, por serem as duas únicas línguas de cultura da Antigüidade e da Alta Idade Média européias, são o grande paradigma em que se inspiraram as línguas vulgares quando ascenderam ao *status* de línguas literárias e de cultura. Diz Einar Haugen a esse respeito:

O latim continuou a ser usado como a língua do governo e da cultura na Europa muito tempo depois da queda do império romano. Não só era usado nos países onde se falava uma forma de latim, mas também nos países onde a Igreja Católica Romana estabeleceu sua forma de cristianismo. Por esse motivo o latim se tornou até certo ponto o modelo do que uma língua deveria ser se ela fosse servir de veículo de cultura nacional e internacional. Tanto o grego como o latim são línguas às quais podemos nos referir como normalizadas,

ou padronizadas, ou em resumo, línguas padrão: elas foram codificadas, ou em outras palavras, foram analisadas na forma escrita e se tornaram um conjunto de regras que pode ser ensinado na escola. (...) A maioria das línguas mais importantes da Europa não foi normalizada até o tempo da Renascença. Isso significa que até o século XVI elas ocupavam só uma posição secundária em seus países, já que o latim era a língua cultural universal.²

Durante toda a Idade Média, os países da Europa Ocidental desconheciam a noção de Estado: o poder encontrava-se pulverizado em inúmeros feudos, cada um dos quais funcionava como entidade política praticamente autônoma, com suas próprias leis e costumes, ditados pelo senhor local. O rei era assim uma espécie de suserano máximo, cujo poder se exercia sobre seus vassalos diretos, os barões do reino, os quais por sua vez dominavam senhores feudais menos poderosos, e assim sucessivamente, formando uma grande pirâmide, em cuja base estavam os servos da gleba. Ora, o fato de o monarca encontrar-se no topo dessa pirâmide não lhe garantia na prática grande poder nem lhe permitia impor ao país praxes sociais, culturais e lingüísticas de caráter nacional.

Quando, na Renascença, a ascensão da burguesia abriu caminho para a revolução filosófica, científica, estética e econômica que redundou na implosão do sistema feudal, o poder do soberano viu-se, por conseqüência, significativamente aumentado. Os reinos medievais, orientados por um critério geopolítico decorrente da própria ideologia renascentista, erigiram-se em impérios segundo o modelo romano. Tal ideologia conduziu não só à reprodução do modelo político de Roma, mas também à imitação dos cânones estéticos da Antigüidade Clássica e, sobretudo, à imitação lingüística do grego e do latim. Os dialetos vulgares medievais que gozavam de maior prestígio, quer por serem veículo de difusão literária, quer pela importância política da região onde eram falados, foram então elevados à condição de línguas literárias e de cultura em escala nacional, calcados no modelo das línguas clássicas, consideradas como as únicas manifestações lingüísticas dignas de ser imitadas. Tal atitude baseava-se numa concepção clássica, também chamada de concepção imperial de língua, segundo a qual

² Apud HILL, A. A., org. *Aspectos da lingüística moderna*. São Paulo, Cultrix/EDUSP, 1974, p. 110-111.

toda língua nasce rude, atinge um estado de perfeição, correspondente ao período de maior esplendor cultural e de apogeu político do Estado em que é falada, após o que se segue um período de decadência, associado igualmente à decadência política, social e cultural desse Estado. Tal concepção, desenvolvida sobretudo a partir dos estudos filológicos e gramaticais empreendidos pelos sábios de Alexandria, no período helenístico (séculos III-I a.C.), considerava o belo como um conceito imutável, negando-lhe, portanto, toda possibilidade de evolução. Assim, era natural que o mundo ocidental, recém-saído de uma era que os próprios europeus de então chamaram de Idade das Trevas, se voltasse para a Civilização Greco-Romana.

Uma vez que todas as línguas literárias européias beberam na mesma fonte greco-latina, e tendo sido esta a principal influência lingüística sofrida pelas mesmas, o caráter ideológico do léxico de cada uma delas vai-se mostrar principalmente na forma particular assumida pelos incontáveis empréstimos de vocábulos gregos ou latinos efetuados a partir de fins da Idade Média, passando pela Renascença e chegando até nossos dias, uma vez que, ainda hoje, a influência do léxico greco-latino nas línguas vulgares está longe de ser desprezível, especialmente em certos universos de discurso, como o da ciência e o do direito, por exemplo.

Mais modernamente, outras línguas de cultura, como o francês e o inglês por exemplo, passaram também a gozar de grande prestígio, seja em virtude da preponderância política e econômica dos países que as falam, seja em função da importância da produção cultural (literária, científica) que possibilitaram. Esse fato faz que as próprias línguas vulgares passem a apresentar uma hierarquia entre si: algumas línguas exercem marcada influência sobre outras, que por sua vez são mais susceptíveis a receber do que a exercer influências. De qualquer modo, toda língua exerce e ao mesmo tempo recebe influências. Cumpre perceber, no entanto, que a intensidade da influência que uma língua exerce ou sofre varia em função do domínio do conhecimento considerado. Assim, o inglês será muito influente no discurso tecnológico, ao passo que o italiano se mostrará hegemônico no discurso da música, e assim por diante. Também nas trocas vocabulares entre línguas vulgares é possível depreender um comportamento léxico ideologicamente muito bem definido.

CARACTERÍSTICAS DO *CORPUS* ESTUDADO

A análise estatística do léxico pressupõe a escolha de um *corpus*, que servirá de amostra sobre a qual serão feitas inferências sobre o próprio léxico a partir do qual tal amostra foi extraída. No presente trabalho, nosso *corpus* consiste de textos pertencentes a dois diferentes universos de discurso, a saber, o discurso científico e o discurso jornalístico, escritos em cinco línguas europeias modernas, três românicas (português, francês e italiano) e duas germânicas (inglês e alemão). As razões da escolha dessas línguas e desses universos de discurso serão apresentadas a seguir. Informações mais detalhadas sobre as características do *corpus*, bem como explicações sobre a metodologia estatística utilizada, serão fornecidas no Capítulo III.

Modalidade oral x modalidade escrita

Embora os processos lexicogênicos por nós estudados digam respeito ao léxico de um modo geral, nossa análise prende-se sobretudo à modalidade escrita da língua. Evidentemente, os mesmos processos neológicos se verificam em todos os universos de discurso, quer naqueles ditos “cultos” (discurso científico, tecnológico, filosófico, jurídico, literário, etc.), nos quais se privilegia a modalidade escrita, quer naqueles ditos “populares”, como o discurso coloquial em geral, a gíria, o calão, etc., que, via de regra, só existem no âmbito da oralidade. Não obstante, são os discursos escritos aqueles que apresentam a maior riqueza em termos lexicogênicos, em face da influência greco-latina e estrangeira mais acentuada, fator, como vimos, fortemente ideológico. Mais ainda, em alguns casos, o aspecto gráfico do vocábulo é de fundamental importância para a própria caracterização dos processos lexicogênicos de que tratamos nesta obra. Constituem também os textos escritos o principal *corpus* do qual se extrai o repertório léxico dos dicionários etimológicos, instrumentos de trabalho indispensáveis à pesquisa lexicogênica, uma vez que é impossível catalogar palavras caídas em desuso que jamais tenham sido registradas.

Durante muito tempo, apenas a modalidade escrita da língua interessou aos estudos lingüísticos. Tal postura remonta ainda aos

primeiros trabalhos de gramática e filologia, encetados na Grécia antiga, e perpassa toda a tradição medieval e renascentista, chegando ao século XIX. Apenas a partir da contribuição de Saussure, e de algumas correntes lingüísticas dela derivadas, é que se inicia o estudo da língua oral (cite-se, por exemplo, a tagmêmica americana de Pike, destinada a descrever línguas ágrafas). É interessante notar que, embora a escrita tenha sido originalmente concebida como um meio auxiliar que permite o registro material da fala, após algum tempo ela assumiu características próprias, sendo hoje nítidas as diferenças gramaticais, vocabulares e estilísticas entre a expressão escrita e a expressão falada de qualquer língua de cultura. Mais ainda, por vezes a escrita determina fenômenos no plano fonético-fonológico. Assim, a palavra alemã *Efeu* (outrora grafada *Epheu*) pronuncia-se atualmente [ˈeːfə] por influência da grafia com *ph*, quando sua pronúncia original era [ˈeːphə], uma vez que provém do médio alto alemão *ebehöu*.³ Da mesma forma, já se escuta no Brasil, entre as camadas populares, a palavra *tóxico* pronunciada como *tóchico*, por força da grafia com *x*.

Análise sincrônica x análise pancrônica

Numa visão estreita, nossa análise poderia ser classificada como sincrônica, uma vez que examina a configuração do léxico num determinado momento de sua história, tal qual testemunhada pelos textos analisados. Entretanto, o léxico de uma determinada etapa sincrônica reflete todos os processos evolutivos por ele sofridos e, ao mesmo tempo, condiciona de certa maneira seu próprio futuro. Os processos lexicogênicos que aqui descrevemos só puderam ser estabelecidos com rigor na perspectiva histórica da língua, de tal modo que a verdadeira compreensão dos fatos ligados à gênese e à renovação lexicais somente pode dar-se mediante uma concepção dinâmica (isto é, dialética) do sistema, em que evolução e funcionamento se articulem de tal maneira a se fundirem num único e mesmo processo. Com efeito, a configuração do léxico de uma língua num dado momento é o resultado de todos

³ Cfr. ROSENTHAL E. T. *A língua alemã*. Desenvolvimento histórico e situação atual. São Paulo, Herder, 1963, p. 284.

os processos lexicogênicos nele verificados ao longo do tempo até o momento considerado. Por outro lado, qualquer criação neológica que venha a ocorrer nessa etapa sincrônica da língua será decerto condicionada pela configuração que o léxico apresenta nessa etapa. Assim, ao introduzir na língua uma nova palavra, o falante, ainda que ingênuo, se pautará, consciente ou inconscientemente, pelos padrões lexicais vigentes na língua, vale dizer, procurará enquadrar seu neologismo nos modelos vocabulares preexistentes, por sinal os únicos disponíveis. Evidentemente, sendo tal processo essencialmente dinâmico, mudanças de comportamento se verificarão ao longo do tempo. Desse modo, a cada nova etapa sincrônica encontraremos um léxico de diferente configuração, que por sua vez determinará uma evolução futura numa direção diferente da até então seguida.

Nesse sentido, é mais prudente tratar nossa análise como *pancrônica*. Numa análise estritamente sincrônica as palavras não se diferenciam aos olhos de um falante ingênuo: considerações de ordem etimológica não têm a menor importância do ponto de vista do usuário cotidiano da língua, assim como palavras introduzidas há uma década têm o mesmo peso de palavras introduzidas há um milênio. Se é verdade que nosso *corpus* retrata a língua de hoje, é igualmente verdade que através dele podemos compreender todo o passado dessa língua, e mesmo prever-lhe o futuro. Nossa análise será então ao mesmo tempo descritiva, histórica e prospectiva. Através dela poderemos conhecer a realidade atual das línguas estudadas, para compará-las e daí extrair informações valiosas. Poderemos também lançar uma luz sobre o passado dessas línguas, a fim de explicar etapas mais ou menos nebulosas de sua evolução e através disso compreender o presente. Poderemos enfim planejar o léxico das línguas com base nesse passado e nesse presente, propugnando a adoção de critérios de escolha e normalização neológica mais racionais, uma vez que baseados no saber científico.

Línguas utilizadas na pesquisa

Conforme dissemos mais acima, nosso *corpus* consiste de textos escritos em cinco línguas européias vulgares, três românicas (português, francês e italiano) e duas germânicas (inglês e ale-

ção). A rigor, teria sido proveitoso se pudéssemos estender nossa análise às demais línguas dessas famílias, como, por exemplo, o espanhol, o catalão, o holandês, as línguas escandinavas, etc. Entretanto, tal pretensão esbarra na dificuldade de obtenção de textos nessas línguas, seja porque os jornais e revistas dos países que as falam, com exceção dos de língua espanhola, não circulam no Brasil, seja porque tais línguas não são veículo de divulgação científica (ver Universos de discurso abordados, p. 51). Some-se a isso a inexistência em nossas bibliotecas de dicionários etimológicos dessas línguas, com os quais forçosamente teríamos de trabalhar. Além disso, a inclusão de um número muito grande de línguas implicaria a manipulação de um *corpus* de dimensões muito superiores àquelas que poderíamos tratar com os recursos de que dispúnhamos. Assim, em virtude dos universos de discurso escolhidos para compor o *corpus*, bem como das dificuldades técnicas envolvidas, tais línguas tiveram de ser descartadas. Optamos então por trabalhar com o português e com mais duas línguas de cada família. A escolha dessas línguas se deu em função da importância das mesmas, decorrente sem dúvida da importância histórica, política, econômica e cultural dos países europeus que as falam, importância essa que gostaríamos aqui de nuançar.

Se observarmos a história da Europa Ocidental, que em grande parte se confunde com a história mundial, sobretudo a partir da Renascença e das Grandes Navegações, veremos que a mesma se desenrola basicamente em torno de quatro países, cuja esfera de influência se estende, por vezes, a todo o planeta: França, Inglaterra, Itália e Alemanha. Com efeito, quase todos os fatos históricos de relevo e repercussão internacional, desde a Idade Média até meados do presente século, se deram dentro do território desses países ou foram de algum modo ditados por seus interesses. Aí ocorreram quase todas as grandes descobertas científicas, aí nasceram os grandes movimentos artísticos, bem como as grandes revoluções sociais, com a implantação de novos sistemas políticos (o absolutismo, o fascismo, a democracia) e econômicos (o feudalismo, o capitalismo). Igualmente, todas as grandes guerras tiveram invariavelmente a participação desses Estados, como, por exemplo, a Guerra dos Cem Anos, a Guerra dos Trinta Anos, as Guerras Napoleônicas, a Guerra Franco-Prussiana, as duas Guerras Mundiais. Esses quatro países foram também sedes de grandes impérios europeus ou coloniais (Sacro Império Romano-

Germânico, Império Napoleônico, Império Colonial Francês, Império Britânico, III Reich).

Desse modo, podemos dizer que os quatro países acima citados constituem uma espécie de núcleo da Civilização Européia, em relação ao qual outros países (países ibéricos, escandinavos, eslavos, balcânicos, bálticos) ocupam uma posição periférica, não só em termos geográficos, mas sobretudo do ponto de vista histórico, político, econômico e cultural. Tal núcleo, entretanto, não se constitui apenas desses quatro países, mas alguns outros Estados menores e regiões geográficas próximas gravitam em torno de seu centro. Podemos adotar então uma definição dessas unidades geográficas com base no critério etnolingüístico, ao invés do critério estritamente político. Desse modo, poderíamos falar de uma “Grande França”, de uma “Grande Inglaterra” (ou Grã-Bretanha), e assim por diante, reunindo todos os territórios e populações de mesma língua. Nossa definição de França incluiria então a Bélgica e a Suíça francófonas; a Grã-Bretanha congregaria, além da Inglaterra, da Escócia e do País de Gales, toda a Irlanda; o que chamamos de Alemanha incluiria a Áustria e parte da Suíça; outra parte da Suíça pertenceria em termos étnicos à Itália.

A esses quatro grandes blocos etnolingüísticos poderíamos somar ainda certas regiões pertencentes à sua esfera de ação direta e que tiveram sua história intimamente ligada à história e aos interesses políticos dos mesmos. Assim, a Alemanha tem nos Países Baixos, Polônia, Tchecoslováquia e Hungria sua área de influência. A Itália possui penetração cultural em parte dos Balcãs e do Mediterrâneo.

Decorrente da importância histórica dos países em questão é a preponderância econômica e militar dos mesmos na atualidade. De fato, das sete grandes potências econômicas do Primeiro Mundo quatro são países europeus, exatamente os quatro que formam aquilo que convencionamos aqui chamar de núcleo europeu. São também as maiores potências bélicas da Europa Ocidental e, em escala mundial, perdem somente para os Estados Unidos e a ex-União Soviética.

Já assinalamos acima a importância desse núcleo europeu como berço de grandes movimentos artísticos (o Gótico, o Neoclassicismo, o Romantismo, o Impressionismo, por exemplo, nasceram na França; o Humanismo, o Renascimento e o Barroco são produto da Itália). Os países desse núcleo logo assumiram

posição de destaque ou de liderança em algum dos diversos domínios da arte (por exemplo, a França na literatura, a Inglaterra no teatro, a Itália nas artes plásticas, a Alemanha na música).

Também nos domínios filosófico e científico tais países sempre detiveram a primazia. As primeiras universidades aí surgiram (Bolonha, Paris, Oxford). Quase todos os grandes inventos e descobertas dos tempos modernos, bem como as principais correntes de pensamento (a escolástica, o mecanicismo, o iluminismo, o positivismo, o marxismo, o estruturalismo) e teorias científicas (a mecânica clássica, a relatividade, a física quântica, o evolucionismo), aí nasceram, criados por sábios desses países. Mais do que isso, algumas ciências, como a química, a sociologia e a lingüística, foram aí criadas. A rigor, a própria ciência moderna teve sua gênese no âmbito desses países, através dos trabalhos de pensadores como Galileu, Newton, Descartes, Bruno, etc.

No aspecto religioso, a Itália é a sede da mais importante igreja do Ocidente, a Igreja Católica, instituição cuja influência e atuação foram e continuam sendo determinantes sobre o destino espiritual e político do Ocidente. Por outro lado, a Reforma Protestante, principal movimento religioso de oposição ao Catolicismo, e responsável pela grande polarização religiosa do mundo ocidental, nasceu na Alemanha, estendendo-se aos demais países anglo-saxões, enquanto os países latinos permaneciam católicos. Já a França e a Inglaterra deram berço às duas mais importantes igrejas nacionais, isto é, subordinadas ao Estado: a igreja galicana, de inspiração católica, na França; e a igreja anglicana, de inspiração protestante, na Inglaterra.

Mas é possível afirmar a importância das línguas por nós escolhidas também do ponto de vista estritamente lingüístico. Para tanto, levamos em conta:

a) a influência exercida por cada uma dessas línguas sobre as demais;

b) o interesse despertado por seu aprendizado, por parte de estrangeiros;

c) o pioneirismo dessas línguas na criação de neologismos posteriormente compartilhados pelas demais;

d) o número de falantes e a difusão internacional dessas línguas (embora, a rigor, este não seja um fator determinante).

Sobre este último ponto, é interessante observar que o espanhol, apesar de muito falado e muito difundido, não é considerado

língua de divulgação científica, a não ser em escala regional (América Latina), talvez até em função da pouca relevância da produção científica dos países hispânicos. Já o português, embora muito falado, é pouco difundido, e, igualmente, só se presta à divulgação científica em escala regional. Essas duas línguas começam a se destacar no cenário internacional devido ao fato de os países que as falam, especialmente os países latino-americanos, representarem atualmente um grande mercado de investimento do capital internacional. Entretanto, sob os pontos de vista histórico, político e cultural, tais línguas continuam a ocupar uma posição de segunda plana em relação às línguas de cultura tradicionais do Ocidente.

Universos de discurso abordados

Conforme mencionamos anteriormente, o *corpus* de nossa pesquisa consiste de textos representativos de dois diferentes universos de discurso: o discurso da ciência, ou, mais simplesmente, discurso científico, e o discurso jornalístico. Como pretendemos, dentre outras coisas, detectar eventuais variações no comportamento lexicogênico das línguas estudadas em função do universo de discurso considerado, procuramos escolher discursos que, do ponto de vista da influência lexical greco-latina e estrangeira, apresentem comportamentos tão distanciados entre si quanto possível.

Numa abordagem apriorística, já que baseada em observações não sistemáticas, nos parece que o discurso científico, por ser um discurso eminentemente “culto”, ligado ao universo acadêmico e a uma tradição secular, é um dos discursos em que mais se verifica a influência do léxico das línguas clássicas. Além disso, trata-se de um discurso especializado, fortemente normativizado e dotado da propriedade da monoisotopia semântica.

O discurso coloquial, ao contrário, parece ser o que menos acusa a influência greco-latina, dado inclusive seu caráter exclusivamente oral. Todavia, justamente pelo fato de tal discurso somente existir no plano da oralidade, sua utilização em nossa pesquisa se vê prejudicada. Mesmo as poucas obras literárias redigidas num estilo mais coloquial são, ainda assim, elaborações de cunho literário, que procuram imitar de modo artificial a fala

popular, não representando, dessa forma, o legítimo discurso coloquial.

Assim sendo, o discurso jornalístico seria, dentre os discursos disponíveis na modalidade escrita, um daqueles que mais se aproximam — se não o que mais se aproxima — do discurso coloquial, uma vez que se dirige a um público bastante variado, na sua maioria formado de pessoas de escolaridade média ou baixa, além de ser um discurso não técnico, não especializado e não hermético, voltado a leitores leigos. Evidentemente, tais considerações apóiam-se, como dissemos, numa análise não sistemática e foram feitas *a priori*. A confirmação ou não desses pressupostos é justamente uma de nossas tarefas.

A escolha do discurso jornalístico se deu não somente por sua suposta proximidade em relação ao discurso coloquial, mas também por refletir de forma bastante fiel a língua da atualidade, na medida em que, na imprensa, o tempo e o cuidado na revisão dos textos é menor, inviabilizando ou desaconselhando muitas vezes a utilização de recursos estilísticos de inspiração literária que poderiam mascarar os usos lingüísticos mais populares. A esse respeito, cumpre salientar ainda que nos jornais diários o tempo de revisão é ainda menor do que nos jornais e revistas semanais.

CAPÍTULO II

FUNDAMENTOS TEÓRICOS

LÉXICO E VOCABULÁRIO: LEXIAS, VOCÁBULOS, PALAVRAS

Àquilo a que vulgarmente chamamos de *palavra* correspondem no jargão das ciências da linguagem três diferentes termos, representativos de três conceitos diferentes de unidades léxicas: a *lexia*, o *vocábulo* e a *palavra-ocorrência*, ou mais simplesmente *palavra*. Partindo da tricotomia *sistema/norma/fala*, proposta por Coseriu⁴, vários autores definem a *lexia* como unidade léxica em nível de sistema, isto é, de língua, e o *vocábulo* como unidade léxica em nível de norma, reservando o termo *palavra* para cada uma das ocorrências concretas das lexias/vocábulos no texto.

Pode-se dizer então que a *lexia* corresponde ao verbete dos dicionários de língua, que apresenta significados diversos, porém relacionados entre si, já que possuem uma base semântica (um núcleo sêmico) comum. A cada um dos diferentes significados da *lexia* dá-se o nome de *acepção*. Quando uma *lexia* é empregada no âmbito de um texto pertencente a determinado universo de discurso (por exemplo, o discurso jurídico, o discurso da medicina, o discurso da engenharia civil, e assim por diante), ocorre que, de todas as acepções elencadas pelo dicionário, apenas uma é a que está efetivamente em jogo; as demais acepções são “filtradas” naquele contexto. A *lexia* transforma-se assim em *vocábulo*, unidade de norma, isto é, de universo de discurso. Cada vez que um determinado *vocábulo* é empregado num texto dado, temos uma *palavra*. Ao conjunto das lexias de uma língua dá-se o nome de *léxico*; o conjunto dos vocábulos de um universo de discurso é o que chamamos *vocabulário*.

⁴ COSERIU, E. *Teoría del lenguaje y lingüística general*. Madrid, Gredos, 1967.

A distinção acima feita entre *lexia* e *vocábulo*, e, conseqüentemente, entre *léxico* e *vocabulário*, é particularmente importante quando se trata de analisar e classificar as unidades léxicas de uma língua de um ponto de vista lexicogênico, como veremos a seguir. Com efeito, quando uma *lexia*, unidade de língua, é atualizada em discurso, sofre necessariamente uma redução de seu significado, em função da natureza do universo de discurso, bem como da situação temática dada.⁵ Isso significa que a *lexia* polissêmica (isto é, portadora de vários significados), unidade disponível no sistema para atualização em diversas normas discursivas, torna-se monossêmica (portadora de um único significado) ao ser efetivamente atualizada num universo de discurso definido. Devido ao aspecto dinâmico do léxico, a cada nova atualização em discurso, uma *lexia* pode provocar um novo “efeito de sentido”, resultante de uma nova combinatória sintagmática, o qual, se recuperado pelo sistema, faz que a *lexia* se veja acrescida de um novo significado, vale dizer, de uma nova acepção. Surge, desse modo, um novo *vocábulo*. Cada um dos *vocábulos* que constituem uma determinada *lexia* possui uma mesma forma de expressão fonológica, uma vez que o que distingue os diferentes *vocábulos* de uma *lexia* é o seu conteúdo. Assim, uma classificação das unidades lexicais de uma língua baseada estritamente no aspecto etimológico destas ficará circunscrita ao plano do significante; por conseguinte, não levará em conta senão a *lexia*. Ocorre, todavia, que cada *vocábulo* de uma mesma *lexia* é gerado em condições diversas dos demais *vocábulos* dessa mesma *lexia*, segundo um processo neológico que lhe é específico. Não se pode, pois, afirmar que o *vocábulo* *casa*, na acepção de “moradia”, tenha a mesma natureza lexicogênica do *vocábulo* *casa* em sua acepção matemática. Eis por que o objeto de análise da lexicogênese deve ser o *vocábulo* e não a *lexia*.

OBSERVAÇÃO: Embora tenhamos definido a *palavra* como unidade de texto, nos permitiremos, no presente trabalho, empregar igualmente tal termo para neutralizar semanticamente os termos *lexia* e *vocábulo*, toda vez que a distinção entre os mesmos não seja pertinente ou relevante.

⁵ Cfr. PAIS, C. T. *Ensaio semiótico-lingüísticos*. Petrópolis, Vozes, 1977, p. 69.

CONVENÇÕES ADOTADAS NESTE TRABALHO

Nesta obra, empregamos freqüentemente a expressão *línguas clássicas* para designar o grego clássico e o latim literário (clássico e medieval), bem como designamos por *línguas vulgares* às línguas românicas (português, espanhol, francês, italiano, etc.) e germânicas (inglês, alemão, sueco, norueguês, etc.), tanto em sua fase arcaica quanto moderna. As designações *clássico* e *vulgar* não têm, portanto, para nós, a conotação eufórica/disfórica que se lhes costuma atribuir: são simplesmente termos de caráter operacional que utilizamos para representar os dois grandes pólos da tensão na qual se baseiam todos os processos lexicogênicos das línguas européias ocidentais.

Embora já tenhamos definido na Lista de Abreviaturas e Símbolos os sinais gráficos que utilizamos neste trabalho, achamos conveniente explicá-los mais detalhadamente aqui, na medida em que o significado de alguns deles foge um pouco do usualmente convencionado.

O sinal > pode indicar a transmissão de um vocábulo de uma língua historicamente delimitada para outra língua, sua sucedânea histórica, através de herança vernácula (por exemplo, lat. *caballu* > port. *cavalo*), bem como pode indicar a própria evolução fonético-fonológica de um vocábulo no âmbito de uma determinada língua (por exemplo, port. arcaico *nembrar* > port. moderno *lembrar*); pode ainda indicar simples evolução da grafia de um vocábulo, como em port. *castello* > *castelo*.

O sinal → indica o empréstimo de um vocábulo, de um morfema, de um significante ou de um significado de uma língua para outra (por exemplo, lat. *templu* → port. *templo*, ing. *football* → port. *futebol*).

O sinal ≈ indica a equivalência semântica entre vocábulos de duas línguas (por exemplo, port. *árvore* ≈ ing. *tree*), sendo necessário ressaltar aqui que a equivalência semântica não deve ser entendida como identidade semântica, isto é, como a existência de vocábulos de diferentes línguas com o mesmo significado; na verdade, a equivalência semântica é antes uma *equivalência de sentido*, nos termos de Hjelmlev. Assim, o port. *árvore* e o ing. *tree*, por exemplo, não têm, obviamente, o mesmo significado, até porque os recortes culturais que lhes subjazem são diferentes. Entretanto, pode-se dizer que *árvore* e *tree* têm o mesmo sentido

de conteúdo, pois recortam a mesma substância semântica, o mesmo *continuum* amorfo dos dados da experiência; codificam, enfim, a mesma informação potencial. É nesse sentido que dizemos que *árvore* e *tree* são semanticamente equivalentes.

TIPOS DE EMPRÉSTIMOS

A primeira coisa a notar acerca do neologismo alogenético (palavra importada de outra língua) é que, em geral, ao introduzir em seu léxico uma nova lexia, de procedência estrangeira, uma língua não importa, juntamente com o significante, todos os significados que aquela lexia possui na língua de origem. De fato, dado que todo neologismo surge dentro de um contexto lingüístico definido, no âmbito de um determinado universo de discurso, para dar conta de um fato cultural específico, tal neologismo tende, no momento de sua implantação, a ser, salvo raras exceções — neologismos poéticos ou publicitários, por exemplo —, monossêmico, o mesmo se verificando também para o neologismo alogenético. Assim, quando, por exemplo, se decide adotar em português o anglicismo *spread* no domínio do discurso econômico, filtram-se todas as outras acepções que essa palavra apresenta na língua inglesa. Evidentemente, o significante desse empréstimo poderá ser reutilizado, com novos significados, referentes ou não a novos universos de discurso, significados estes que poderão ser, igualmente, transpostos do inglês para o português, ou, mesmo, criados já no âmbito da língua portuguesa. Exemplo deste último caso é a palavra *bonbonnière*, que em francês designa o recipiente em que se guardam bombons, e em português, além de admitir essa acepção, significa também “estabelecimento que vende chocolates”. Em decorrência do que foi dito, podemos dizer que uma língua não importa uma lexia, unidade de léxico, polissêmica, mas sim um vocábulo, unidade de norma, de universo de discurso, monossêmica, a qual sofre, ao ser transplantada de uma língua para outra, uma redução de significado, e, ainda, uma adaptação a uma nova rede léxico-conceptual.

Do ponto de vista das relações *significante/significado* existentes no interior do signo, o empréstimo de um vocábulo de uma língua A para uma língua B pode ser de três tipos, a saber:⁶

a) *Empréstimo total*, quando a língua B toma emprestado o vocábulo como um todo, isto é, significante e significado, ressalvadas as adaptações fonético-fonológicas, morfológicas e ortográficas obrigatórias, além da redução de significado acima referida;

$$\begin{array}{ccc} \text{Língua A} & & \text{Língua B} \\ \frac{C}{E} & \longrightarrow & \frac{C}{E} \quad \text{onde } C = \text{conteúdo} \\ & & E = \text{expressão} \end{array}$$

b) *Empréstimo de significante*, quando a língua B toma emprestado um significante de A, atribuindo-lhe um novo significado, criado em B;

$$\begin{array}{ccc} \text{Língua A} & & \text{Língua B} \\ \frac{C}{E} & \longrightarrow & \frac{C'}{E} \end{array}$$

c) *Empréstimo de significado*, quando a língua B toma emprestado um significado de A, substituindo o respectivo significante por um outro, já existente em B, ou criado especialmente nessa língua para constituir, com aquele significado, o novo vocábulo.

$$\begin{array}{ccc} \text{Língua A} & & \text{Língua B} \\ \frac{C}{E} & \longrightarrow & \frac{C}{E'} \end{array}$$

⁶ Cfr. BARBOSA, M. A. *Léxico, produção e criatividade: processos do neologismo*. São Paulo, Global, 1981, p. 293-294.

EMPRÉSTIMOS DE VOCÁBULOS GRECO-LATINOS

Quando uma língua vulgar decide tomar emprestado ao grego ou latim um vocábulo, este sofre necessariamente um processo de modificação estrutural, a fim de compatibilizá-lo ao novo sistema lingüístico em que deve enquadrar-se. Em primeiro lugar, devemos lembrar que quase nunca o significado que tal vocábulo deverá ter na língua receptora do empréstimo coincide exatamente com o significado original grego ou latino, até porque a própria rede semântico-conceptual em que esse significado se insere é diferente para cada língua. Muitas vezes, em especial na norma técnico-científica, um empréstimo greco-latino é adotado para dar conta de um conceito totalmente novo, próprio daquele universo de discurso. Tem-se aí puramente um empréstimo do significante greco-latino, ao qual se atribui um novo significado. Entretanto, também o significante greco-latino pode sofrer alterações, as quais podem ser espontâneas ou deliberadas, obrigatórias ou facultativas, em função da própria estrutura fonético-fonológica da língua receptora. De qualquer maneira, convém lembrar que, quando do empréstimo, necessariamente ocorre a adaptação da pronúncia do significante ao sistema fonético-fonológico da língua receptora, bem como a adaptação da grafia do vocábulo ao sistema ortográfico dessa língua. Cabe ressaltar aqui que quase nunca é possível preservar ao mesmo tempo a grafia e a pronúncia greco-latinas. Mesmo quando o sistema ortográfico de uma língua vulgar permite conservar, total ou quase totalmente, a grafia greco-latina, como ocorre, por exemplo, em francês ou inglês, isso invariavelmente se dá em detrimento da pronúncia grega ou latina, pois, evidentemente, os mesmos grafemas possuem valores diferentes nos diferentes sistemas. Da mesma forma, se a língua vulgar optar pela conservação da pronúncia original, deverá fazê-lo em prejuízo da grafia, já que sons idênticos costumam ser representados graficamente de forma diversa em cada língua. Na grande maioria dos empréstimos greco-latinos por línguas vulgares européias, a grafia apresenta nítida precedência sobre a pronúncia. Tal fato se deve a duas razões principais: em primeiro lugar, tais empréstimos possuem quase sempre caráter erudito, e por isso entram nessas línguas por via escrita, e não oral; em segundo lugar, devemos lembrar que o próprio conhecimento que se tem das línguas clássicas

se baseia em textos escritos nessas línguas, não se sabendo ao certo qual era a exata pronúncia das palavras gregas ou latinas. Mesmo durante a Idade Média, quando o latim ainda era falado, seu uso era até certo ponto meramente convencional, já que estava restrito a situações e discursos formais, enquanto nas situações e discursos informais se utilizava normalmente a língua vulgar. Em consequência, todos os falantes do latim eram, na verdade, bilíngües. Desse modo, enquanto língua de cultura da Idade Média, o latim era muito mais uma língua escrita do que falada. Se ainda não era propriamente uma língua morta, já não podia tampouco ser considerada uma língua “viva”, entendida como língua natural de um povo, de uma comunidade etnicamente definida. Por essa razão, o latim medieval acabava por admitir uma pronúncia diferente em cada país, em cada região, pronúncia esta modelada em grande parte pelo sistema fonético-fonológico da língua vulgar dessa região. Havia então, *grosso modo*, um latim “com sotaque francês”, um latim “com sotaque italiano”, e assim por diante.

Com relação aos empréstimos, convém lembrar, por outro lado, que, além das adaptações e/ou modificações que uma palavra pode sofrer no momento de sua introdução no léxico da língua, esta sofre também o processo da evolução histórica, que poderá provocar mudanças tanto no significante como no significado. Entretanto, antes de prosseguirmos, é preciso fazer uma clara distinção entre a adaptação fonético-fonológica pura e simples, que não é senão uma adaptação da pronúncia do vocábulo à grafia latina (por exemplo, o *th* de *theatrum* pronunciado como fricativa linguodental em inglês, visto ser esta a pronúncia usual de tal grafema nessa língua), e a mutação fonético-fonológica, resultante esta quer da evolução histórica natural da língua, quer da alteração intencional do significante do vocábulo, quando de sua introdução na língua, por analogia às palavras vernáculas da mesma, fenômenos a que damos o nome de *metamorfismo*. Ao contrário da adaptação fonético-fonológica, o metamorfismo confere ao empréstimo um aspecto formal vulgar. Veremos, a seguir, os tipos possíveis de processos de adaptação e/ou mutação, ao nível do significante, que um empréstimo greco-latino pode sofrer.

Processos não vulgarizantes

a) Empréstimo com conservação da grafia original, com ou sem enquadramento morfológico (isto é, substituição da desinência grega ou latina pela vernácula), porém com adaptação fonético-fonológica a essa grafia. Exemplos:

port. *campus* (plural *campi*) /'kâpus/, /'kâpi/ — sem enquadramento morfológico (conservação das desinências latinas *-us/-i*) e adaptação fonológica.

ing. *theatre* /'θiɪtɹ/, fr. *juste* /'ʒyst/ — com enquadramento morfológico e adaptação fonológica.

b) Empréstimo com adaptação ortográfica obrigatória, sem mutação fonológica. Exemplos:

lat. *fabricare* → fr. *fabriquer*

lat. *physica* → port. *fisica*, al. *Physik*

OBSERVAÇÃO: A diferença de grafia entre o latim e as línguas vulgares pode ser resultado quer de coerções de natureza fonológica (por exemplo, a substituição de *c* por *qu* antes de *e* em francês para assegurar a conservação da pronúncia oclusiva velar), quer de reforma ortográfica (por exemplo, o port. *física*, que outrora já se escreveu à moda latina, com *ph* e *y*).

c) Empréstimo do grego ou latim, com posterior evolução fonética, porém com conservação da grafia original. Exemplos:

lat. *versu* → al. *Vers* /'vers/ > *Vers* /'fers/

lat. *gradu* → ing. *grade* /'gra:d/ > *grade* /'grejd/

d) Empréstimo com adaptação fonético-fonológica e posterior reforma ortográfica não motivada por essa adaptação, e, portanto, sem alteração do valor fonológico dos grafemas implicados. Neste caso, não há vulgarização, pois nenhum dos dois processos é vulgarizante. Exemplo:

Processos vulgarizantes

a) Empréstimo do grego ou latim, com posterior evolução fonética, acompanhada de adaptação ortográfica, simultânea ou posterior, visando a dar conta dessa evolução fonética. Exemplos:

lat. *productu* → port. *producto* /pro'duktu/ > *producto* /pro'dutu/ > *produto* /pro'dutu/

lat. *latinu* → al. *latin* /la'ti:n/ > *latin* /la'tejn/ > *latein* /la'tejn/ > *Latein* /la'tajn/

OBSERVAÇÃO: Nota-se neste caso uma fase intermediária, em que já houve evolução fonética, mas a grafia permanece latina. Nesta fase, o vocábulo encontra-se na situação descrita em Processos não vulgarizantes, p. 62, alínea c. Trata-se de um estágio de transição, em que o significante do vocábulo ainda mantém o caráter erudito, exceto quanto à pronúncia (vide na p 58, alínea a). Na verdade, é no momento em que uma reforma ortográfica reconhece a evolução fonética que fica consagrada a vulgarização do vocábulo. De fato, podem conviver na língua formas resultantes tanto do processo descrito na p. 62, alínea a, quanto do processo descrito na alínea c; numa abordagem estritamente sincrônica, tais formas não se diferenciam. Por exemplo, comparem-se o ing. *grade*, “grau” e *matrix*, “matriz” (pronunciado /'mejtriks/). Em ambos os casos, há conservação da grafia e alteração da pronúncia. Entretanto, *grade*, como vimos, sofreu mutação fonética, ao passo que *matrix* já entrou na língua com a grafia e pronúncia atuais. Neste último caso, houve apenas adaptação fonológica, visto que à época da introdução de *matrix* o grafema *a* já se pronunciava /ej/ em inglês. Note-se, entretanto, que para que se configure a vulgarização, é preciso que a mudança da grafia do vocábulo seja decorrente da evolução fonética, e, mais ainda, que haja uma alteração do valor fonológico dos grafemas implicados nessa mudança. Assim, tomando o exemplo do al. *Latein*, temos que, no momento em que se adotou o grafema *ei* para representar a seqüência fonológica /ej/, simultaneamente o grafema *i*, que até então tivera essa função, sofreu uma redistribuição.

buição, perdendo tal valor. Desse modo, os grafemas *ei* e *i* em nenhum momento tiveram valor fonológico equivalente. Muito diferente é a situação descrita na p. 62, alínea d, em que houve adaptação fonético-fonológica e posterior reforma ortográfica, porém tanto o grafema substituto quanto o grafema substituído possuíam ambos o mesmo valor fonológico, e o mantiveram após a reforma.

b) Empréstimo com adaptação fonético-fonológica à grafia greco-latina e posterior reforma ortográfica, com alteração do valor fonológico dos grafemas implicados. Exemplo:

lat. *natione* → nor. *nation* /na'ju:n/ > *nasjon* /na'ju:n/

OBSERVAÇÃO: Note-se que a mudança de grafia do norueguês é conseqüência de uma adaptação da pronúncia. Ocorre que, num determinado momento, todo grafema *ti* com valor fonológico de /f/ foi sistematicamente substituído por *sj*, uma vez que, até então, ambos os grafemas possuíam o mesmo valor fonológico em norueguês. Conseqüentemente, após essa reforma ortográfica, o grafema *ti* deixou de possuir o valor /f/ em norueguês. Nesse sentido, houve uma simplificação ortográfica, em decorrência da qual os grafemas *ti* e *sj* tiveram seu valor fonológico e sua distribuição redefinidos.

c) Empréstimo com mutação fonética arbitrária e correspondente adaptação ortográfica (= metamorfismo). Exemplos:

lat. *jubilu* /'ju:bilu/, → it. *giubilo* /'dʒubilo/

lat. *textu* /'tekstu/ → it. *testo* /'testo/

d) Empréstimo de latinismo a partir de língua vulgar estrangeira, com adaptação da grafia à pronúncia do vocábulo na língua vulgar de origem. Exemplos:

lat. *machina* → fr. *machine* /ma'ʃin/ → al. *Maschine* /ma'ʃi:ne/

lat. *tribuna* → fr. *tribune* /tri'byn/ → al. *Tribüne* /tri'by:ne/

lat. *statua* → fr. *statue* /sta'ty/ → sue. *staty* /sta'ty:/

lat. *media* → ing. *media* /'mi:diə/ → port. *mídia* /'midia/

OBSERVAÇÃO: Note-se que as formas francesas e a forma inglesa não são vulgares, pois apresentam apenas adaptação fonético-fonológica à grafia latina, enquadrando-se portanto no processo descrito na p. 62, alínea a, ao passo que as formas alemãs, a forma sueca e a portuguesa apresentam mutação fonético-fonológica e alteração gráfica em relação ao latim.

e) Empréstimo de latinismo a partir de língua vulgar estrangeira, com adaptação da pronúncia à grafia do vocábulo na língua vulgar de origem. Exemplo:

lat. *servitium* /ser'vi:sium/ → fr. *service* /ser'vis/ → rom. *serviciu* /ser'viʃu/

OBSERVAÇÃO: Aqui também a forma francesa não é vulgar, pois somente apresenta adaptação da grafia à pronúncia latina, substituindo o *ti* latino por *c*, com conservação da pronúncia /s/. Tal forma enquadra-se assim no processo descrito na p. 62, alínea b. Já a forma romena apresenta tanto mutação fonético-fonológica quanto alteração gráfica em relação ao latim.

f) Refecção gráfica de vocábulo vulgar, segundo o modelo latino, porém com conservação do fonetismo vulgar. Exemplos:

lat. *nidu* > fr. *ni* (séc. XI) > *nid* (séc. XVI)

lat. *salute* > fr. *salu* (séc. XI) > *salut* (séc. XIII)

OBSERVAÇÃO: Neste caso, não temos propriamente um empréstimo do significante latino, mas sim uma refecção do significante vulgar por imitação do latim. Trata-se de um falso empréstimo, e por essa razão, o vocábulo conserva seu estatuto vulgar (vide Refecções de vulgarismos, p. 122).

OBSERVAÇÕES:

1) Como vimos, um vocábulo tomado de empréstimo ao grego ou ao latim permanece erudito se ocorre alteração ou adapta-

ção da pronúncia com conservação da grafia, ou conservação da pronúncia com alteração da grafia (por força de adaptação ou reforma ortográfica), ou, ainda, se ocorre adaptação da pronúncia e posterior reforma ortográfica, desde que tal reforma não seja motivada pela adaptação da pronúncia. Por outro lado, um vocábulo é vulgarizado se sofre mutação ou adaptação fonética em relação ao latim, com adaptação gráfica correspondente, concomitante ou posterior à mutação ou adaptação. Note-se que, para que fique configurada a vulgarização, é preciso que o vocábulo objeto de mutação ou adaptação fonética tenha sofrido uma reforma ortográfica especificamente para dar conta desse evento. Desse modo, é perfeitamente possível que, num determinado momento, um vocábulo sofra uma mutação fonética resultante de sua evolução histórica, sem que haja alteração da grafia, e que, posteriormente, tal vocábulo venha a ser alvo de uma reforma ortográfica, sem que esta seja motivada pela mutação fonética. Da mesma forma, é perfeitamente possível que um vocábulo que sofreu adaptação aos padrões fonético-fonológicos da língua passe por uma reforma ortográfica não motivada por essa adaptação. Em ambos os casos, estamos diante de dois eventos independentes, nenhum dos quais, isoladamente, é vulgarizante. Assim, é preciso deixar bem claro que o metamorfismo só se dá quando a mudança fonética e a alteração gráfica apresentam conexão entre si, mesmo que não ocorram simultaneamente.

- 2) Um vocábulo tomado de empréstimo ao grego ou ao latim pode conservar a grafia e a pronúncia originais, sofrer alteração na pronúncia, na grafia ou em ambas. Entretanto, nem todas as fases dessa evolução precisam ter sido documentadas. Assim, se num momento qualquer da história de uma dada língua, vocábulos escritos com *ph* passaram a grafar-se com *f*, em virtude de reforma ortográfica, há por outro lado vocábulos que entraram na língua posteriormente e que portanto se grafam *ab initio* com *f*. O mesmo vale para adaptações fonológicas. Cabe ressaltar que há palavras que diferem do latim tanto fonética quanto graficamente, e no entan-

to não são vulgarizadas (por exemplo, esp. *nación*, al. *Zensur*), pelas razões expostas acima.

TIPOLOGIA LEXICOGÊNICA DOS VOCÁBULOS DE UMA LÍNGUA

Do ponto de vista de sua lexicogênese, isto é, dos processos através dos quais uma língua cria e renova seu léxico, os vocábulos de uma língua podem ser divididos, numa primeira abordagem, em dois grandes grupos:

a) *Vocábulos vernáculos*, isto é, aqueles que constituem o material léxico de base da língua no momento em que esta se individualiza historicamente, e que, numa perspectiva diacrônica, correspondem aos elementos existentes na língua desde sempre, visto que foram herdados diretamente do estágio diacrônico anterior, ao qual os lingüistas histórico-comparativos chamam de *língua-mãe*. Fica claro já de início que a própria conceituação do que seja vernáculo numa língua depende da fixação da “data de nascimento” dessa língua, isto é, do momento histórico a partir do qual a língua-mãe dá origem à língua-filha. Sabemos que esse momento tem existência puramente teórica, visto que do ponto de vista diacrônico não há senão continuidade, ou, antes, mudança constante. Por conseguinte, o estabelecimento dessa “data de nascimento” é um ato arbitrário, porém metodologicamente justificado. Para as línguas européias ocidentais, adota-se como *terminus a quo* o instante de seu primeiro registro escrito, o que se dá por volta dos sécs. VIII–IX. Assim, para as línguas românicas, são vernáculos os elementos léxicos provenientes diretamente do latim vulgar e do romance, por evolução fonética regular. Para as línguas germânicas, são vernáculos as unidades provindas diretamente dos dialetos pré-literários germânicos (anglo, saxão, alemânico, juto, etc.), e que, conforme foi visto em O aspecto ideológico do léxico, na p. 40, agrupamos sob a designação genérica de *germanice*.

b) *Neologismos*, isto é, vocábulos engendrados em época posterior à do início histórico da língua, e que, segundo Guilbert⁷,

⁷ GUILBERT, L. *La créativité lexicale*. Paris, Larousse, 1975.

podem ser fonológicos, sintagmáticos, semânticos e alogenéticos.

Por outro lado, no caso específico das línguas européias ocidentais, a influência das línguas clássicas — grego e latim, especialmente o latim medieval e eclesiástico — é enorme, sobretudo na assim chamada “norma culta” (universos de discurso filosófico, científico, jurídico, etc.), permanecendo até hoje essas duas línguas a grande fonte de elementos léxicos (radicais, afixos, etc.) onde se abastecem os discursos cultos das línguas européias modernas. Conseqüentemente, poderíamos também agrupar os vocábulos dessas línguas em relação às línguas clássicas em *vocábulos eruditos*, aqueles tomados de empréstimo ao grego ou latim, ou resultantes da combinação sintagmática de elementos greco-latinos, e *vocábulos vulgares*, correspondendo basicamente aos vocábulos vernáculos, seus compostos e derivados. Entretanto, muitos vocábulos apresentam hibridismo de elementos eruditos e vulgares, razão pela qual são chamados de *vocábulos semi-eruditos*. Resultam em geral da combinação sintagmática de elementos eruditos e vulgares (por exemplo, composição a partir de um radical grego ou latino e outro vernáculo, ou derivação de radical erudito com afixo vulgar), da substituição de algum(ns) dos morfemas do vocábulo greco-latino original por morfema equivalente vernáculo, ou ainda de empréstimos greco-latinos que sofreram mutação fonético-fonológica natural (resultante da evolução histórica da palavra) ou induzida, mutação esta a que, como vimos em Empréstimos de vocábulos greco-latinos, na p. 60, demos o nome de *metamorfismo*. Contudo, há ainda vocábulos cujos morfemas são todos vulgares, e que no entanto são semi-eruditos, porque são calcados em vocábulos greco-latinos, isto é, são verdadeiras “traduções” desses vocábulos.

Tendo em vista que tanto os vocábulos semi-eruditos quanto os vocábulos vulgares possuem um caráter de “vernaculidade”, ao menos do ponto de vista do significante, e tendo em vista também que o caráter mais classicizante ou mais vulgarizante dos vocábulos de uma língua pode revelar, em última instância, a própria ideologia subjacente ao léxico dessa língua, que preside à escolha das unidades que devem compor esse mesmo léxico, optamos por dividir o léxico das línguas européias ocidentais em duas macrocategorias, a saber, os *grecolatinismos* (vocábulos dos quais a totalidade dos morfemas, salvo as desinências, são greco-latinos, sem nenhum metamorfismo) e *vulgarismos* (vocábulos dos quais pelo

menos um dos morfemas constituintes, à exceção das desinências, é vernáculo ou metamorfizado). Da própria definição que acabamos de dar de *grecolatinismo* e *vulgarismo*, resulta que tais termos não têm para nós exatamente o mesmo significado que usualmente se lhes atribui, particularmente em trabalhos de filologia e etimologia. Cumpre ressaltar, ainda, que tais termos dizem respeito exclusivamente ao aspecto formal, lexicogênico, dos vocábulos, e não ao seu aspecto pragmático, não importando, pois, se um determinado vocábulo, formalmente classificado como *grecolatinismo* ou *vulgarismo*, é efetivamente de uso culto ou popular, do ponto de vista diacrônico e diafásico. Não há, portanto, correlação necessária entre o estatuto lexicogênico — *greco-latino* ou *vulgar* — de um vocábulo e o(s) universo(s) de discurso preferencial(is) de sua ocorrência, até porque muitos termos técnicos, pertencentes a normas discursivas especializadas, dentre os quais muitos compostos de elementos greco-latinos, podem sofrer um processo de banalização lexical, isto é, podem ser incorporados a normas não técnicas, inclusive a norma coloquial. Citem-se, por exemplo, termos como *neurose*, *paranóia*, *psicopata*, dentre outros, extraídos do discurso psiquiátrico, e que hoje são de uso corrente no discurso banal. Por outro lado, muitos significantes de vocábulos vulgares pertencentes ao discurso banal podem sofrer um processo de especialização, isto é, podem ser adotados por algum tecnoleto (jargão característico de uma determinada atividade profissional), como suporte de um novo significado, específico desse tecnoleto, transformando-se assim em termo técnico. É o caso de vocábulos como *árvore*, *eixo*, *raiz*, etc., no discurso da matemática.

Assim, para evitar uma eventual ambigüidade terminológica, convencionamos que os termos *erudito*, *semi-erudito* e *vulgar* terão para nós valor estritamente lexicogênico, de tal modo que o termo *erudito* se referirá aos grecolatinismos, e os termos *semi-erudito* e *vulgar* dirão respeito aos vulgarismos, tais como definimos essas macrocategorias lexicogênicas mais acima. Por outro lado, utilizaremos os termos *culto* e *popular* para indicar o tipo de universo de discurso preferencial de um vocábulo. Assim, serão ditos cultos aqueles vocábulos pertencentes a normas especializadas, como, por exemplo, a norma técnico-científica, jurídica, religiosa, etc., qualquer que seja seu estatuto lexicogênico (*erudito*, *semi-erudito* ou *vulgar*), ao passo que serão ditos populares

aqueles vocábulos pertencentes ao discurso banal, qualquer que seja seu estatuto lexicogênico.

Mais adiante, apresentaremos uma tipologia dos grecolatinismos e dos vulgarismos, a qual, embora possa ser insuficiente, nos balizará por ora em nosso labor classificatório. É de notar que, dada a complexidade dos fenômenos que intervêm na produção das unidades léxicas, nossa classificação, por partir de uma perspectiva lexicogênica, procura, especialmente no tocante aos neologismos, abranger e, ao mesmo tempo, superar a teoria do neologismo de Guilbert, acima mencionada.

LATIM NORMALIZADO

Nossa classificação dos empréstimos greco-latinos em eruditos e semi-eruditos (vulgarizados) baseia-se no tipo de percurso através do qual o vocábulo emprestado chega à língua vulgar e do qual podem ou não resultar divergências formais ao nível do significante (gráficas e/ou fonético-fonológicas) entre o vocábulo grego ou latino que dá origem ao empréstimo e o vocábulo da língua vulgar resultante desse empréstimo. Portanto, para decidirmos sobre o caráter erudito ou vulgarizado de um vocábulo, temos de compará-lo ao vocábulo grego ou latino que lhe dá origem, a fim de podermos distinguir entre a simples adaptação fonético-fonológica e/ou ortográfica (processos não vulgarizantes) e a mutação ou evolução fonético-fonológica, com correspondente adaptação gráfica (processos vulgarizantes). Assim sendo, é imprescindível para a análise lexicogênica dos empréstimos greco-latinos das línguas vulgares escolhermos um padrão fonológico e ortográfico para as línguas clássicas, que servirá de parâmetro dessa análise.

Devemos lembrar aqui que a maior parte dos vocábulos de origem grega que chega às línguas vulgares o faz através do latim, que, por razões históricas e culturais conhecidas, efetuou em larga escala importações de termos gregos. Entretanto, sobretudo a partir da Renascença, começa a crescer nas línguas vulgares o número de vocábulos e elementos de formação vocabular (radicais, afixos) importados diretamente do grego, principalmente no âmbito do discurso científico. Assim, quando um vocábulo grego

existir também em latim, a grafia que nos servirá de parâmetro será a grafia latina desse vocábulo. Quando um vocábulo grego, que não exista em latim, é emprestado diretamente a uma língua vulgar, a grafia que adotaremos como parâmetro para tal vocábulo — e que, por sinal, as línguas vulgares adotam em seus empréstimos — será a grafia latinizada, isto é, transliterada para o alfabeto latino, seguindo as mesmas convenções ortográficas dos vocábulos gregos emprestados ao latim (por exemplo, o vocábulo grego *ὀφθαλμός* será transcrito em latim *ophthalmos*). Deste modo, nossa comparação se dará sempre em relação à forma latina dos vocábulos. Ao padrão ortográfico e fonológico latino por nós adotado como parâmetro de análise, chamaremos *latim normalizado*. Tal padrão baseia-se na pronúncia do latim medieval, que é a forma que serviu de paradigma para as línguas vulgares românicas e germânicas, admitindo para alguns grafemas⁸ mais de uma pronúncia, quando tais variantes fonológicas forem tradicionais. A tabela a seguir mostra a correspondência entre os grafemas latinos e seu valor fonológico normalizado.

GRAFIA	PRONÚNCIA	GRAFIA	PRONÚNCIA
<i>c</i> (+ <i>e, i, y</i>)	[tʃ], [ts], [s]	<i>x</i>	[ks], [gz]
<i>g</i> (+ <i>e, i, y</i>)	[dʒ], [ʒ]	<i>thi</i> (+ vogal)	[t]
<i>ph</i>	[f]	<i>titi</i> (+ vogal)	[tt], [t]
<i>th</i>	[t]	<i>sti</i> (+ vogal)	[sti]
<i>ch</i>	[k]	<i>xiti</i> (+ vogal)	[ksti]
<i>rh</i>	[r]	<i>j</i>	[j]
<i>ti</i> (+ vogal)	[tsi], [si]	<i>v</i>	[v]
<i>z</i>	[ts], [dz], [z]	<i>y</i>	[i]
<i>h</i>	∅	<i>ae</i>	[e], [e]
<i>s</i> intervocálico	s, z	<i>oe</i>	[e]

OBSERVAÇÃO: Os demais grafemas latinos mantêm a pronúncia usual.

⁸ Em nossa terminologia, utilizaremos o termo *grafema* para designar uma seqüência de letras qualquer que, por ocorrer no sistema ortográfico da língua de forma freqüente, sistemática e regular, pode constituir um traço característico dessa língua. Assim, em latim, a seqüência *-tia-* é mais freqüente do que *-toa-*, por exemplo. Note-se que o grafema não corresponde necessariamente a um único fonema. Menos ainda, a noção de grafema, tal como definida por nós, coincide com a de *letra*.

Outro aspecto importante a ser levado em conta no momento da comparação entre a forma original de um vocábulo greco-latino e a forma assumida na língua vulgar pelo empréstimo correspondente diz respeito às equivalências ortográficas. Esse aspecto é relevante na medida em que a uma grafia divergente do grego/latim pode corresponder uma mutação fonético-fonológica, o que configura a vulgarização do vocábulo. Outras vezes, a divergência gráfica se deve tão-somente a uma adaptação da grafia greco-latina ao sistema ortográfico da língua vulgar. Além das conhecidas transformações gráficas do tipo *ph > f*, *y > i*, etc., mormente devidas à inexistência daqueles grafemas latinos na ortografia da língua vulgar, existem também grafemas tradicionais, em geral surgidos na Idade Média, que não substituem propriamente os grafemas latinos, mas, antes, são formas gráficas alternativas aos mesmos, podendo comutar com eles em todos os contextos, especialmente quando falte o sinal tipográfico correspondente. Frequentes nas línguas germânicas, damos aqui os grafemas alternativos utilizados na ortografia da língua alemã, e sua correspondência com a grafia latina:

lat. *ae* ≈ al. *ä*

lat. *oe* ≈ al. *ö*

lat. *ss* ≈ al. *ß*

Assim, sempre que encontrarmos tais sinais gráficos em vocábulos alemães de origem greco-latina, devemos convertê-los nos grafemas latinos correspondentes, antes de realizarmos a análise lexicogênica desses vocábulos.

VULGARISMOS

Vulgarismos resultantes de herança vernácula

Como já vimos, os vulgarismos resultantes de herança vernácula são aqueles vocábulos herdados diretamente do sistema lingüístico anterior à consolidação de uma língua vulgar como língua literária, e, portanto, como língua de cultura autônoma, não mais como simples dialeto. A fixação de uma fronteira temporal que separe a língua-mãe da língua-filha, na terminologia glotoló-

gica tradicional, é tarefa das mais árduas e complicadas; dela depende entretanto a própria definição do que seja vernáculo ou empréstimo. Por um lado, até aproximadamente o séc. VIII (um pouco antes, em algumas regiões, ou pouco depois, em outras), havia nos falantes dos diversos romances um sentimento de que não falavam outra coisa que não o próprio latim, sob forma um pouco diferente. Nessa época, o intercâmbio de elementos lingüísticos entre o latim literário e seus dialetos romances era intenso, em ambas as direções, num processo constante de alimentação e realimentação lexical. Citemos como exemplo, o sufixo latino *-aticum*, que evoluiu no galo-romance para *-atge*, sendo em seguida reintroduzido em latim na forma *-agium*.⁹ Note-se que tais contribuições recíprocas entre latim e romances não eram vistas como empréstimos, já que, conforme dissemos, prevalecia um sentimento de unidade lingüística. Por outro lado, em meados do séc. IX, surge o primeiro documento escrito em língua francesa, os famosos Juramentos de Estrasburgo, em que a distinção entre a “latina lingua” e a “romana lingua” está patente. A partir daí — sobretudo a partir da Reforma Carolíngia —, já se pode falar em “língua francesa”, e os próprios falantes dessa língua já a sentem como tal. Mesmo os demais dialetos falados no recém-nascido Reino de França já não são mais dialetos do latim, mas sim variedades do mesmo idioma que fora guindado à condição de língua literária. Contribuições latinas a essa língua são agora efetivamente resultado de empréstimos, as mais das vezes eruditos, de modo que a individualização lingüística coincide, não por acaso, com a individualização política do Estado francês. Cabe observar que o italiano só tardiamente se individualizou como língua literária (séc. XIII), e tal fato se dá justamente quando aparecem os primeiros poemas escritos em *volgare*, numa tentativa bem sucedida de provar que tal língua, construída sobre a base do dialeto florentino, era perfeitamente adequada à expressão dos mais nobres sentimentos e dos conceitos mais sofisticados.

Assim, sendo-nos forçoso, para nossa tarefa de classificação, realizar um recorte no *continuum* que é a língua, parece-nos que a escolha mais razoável de uma fronteira temporal entre o latim e as línguas românicas ou entre o germânico e as línguas germânicas,

⁹ Cfr. MAURER JR., T. H. *A unidade da România ocidental*. São Paulo, FFCL-USP, 1951, p. 111-112.

deva recair na data do primeiro registro escrito, ressaltando, contudo, que se trata apenas de um corte metodológico, sendo óbvio que a transição de um estágio de consciência lingüística para outro não se dá instantaneamente, mas constitui um longo processo, de modo que tal fronteira, por nós adotada, é necessariamente arbitrária e fluida.

No caso das línguas românicas, é vernáculo tudo quanto já existia no latim vulgar e continuou a existir, sem solução de continuidade, nos diversos romances regionais (galo-romance, ítalo-romance, ibero-romance, etc.), de modo que, quando esses romances ganharam o estatuto de línguas literárias autônomas, distintas do latim, e, eventualmente, línguas de comunicação comum aos membros de uma mesma comunidade étnico-cultural (a designação “línguas nacionais” seria, sem dúvida, anacrônica para esse período), esse acervo de material léxico (e também gramatical, evidentemente) passou a constituir o léxico de base dessas novas línguas.

Será vernáculo nas línguas germânicas aquele material léxico e gramatical que já existia nos diversos dialetos integrantes da família germânica (gótico, anglo, saxão, juto, franco, lombardo, alemânico, bávaro, burgúndio, nórdico, etc.), e que, conforme já dissemos, convenciamos denominar *germances*, por analogia aos dialetos romances. Poderíamos, assim, falar de um anglo-germance, de um goto-germance, e assim por diante.

A característica básica de um vulgarismo vernáculo é a regularidade das mutações fonéticas (metaplasmos) que sofreu ao longo de seu percurso histórico. Há, sem dúvida, uns poucos casos de vocábulos vernáculos que apresentam evolução fonética irregular, mas trata-se de casos excepcionais, que, no mais das vezes, podem ser explicados historicamente. Entretanto, deve-se evitar absolutizar a idéia, tão cara aos lingüistas histórico-comparativos e aos neogramáticos, da infalibilidade das leis fonéticas, visto que o próprio progresso da ciência lingüística provou o seu não fundamento. Convém ressaltar, a esse respeito, que os chamados metaplasmos fonéticos têm existência e vigência relativa, isto é, restrita a um lugar e a um tempo determinados. Aplicam-se, além disso, a um certo número de elementos fonéticos, em dados contextos, e, portanto, é mister, antes de tudo, situar espaço-temporalmente tais processos evolutivos, para reconhecer, inclusive, a existência de um “metaplasmo zero”, isto é, a ausência de

mutação fonética. A palavra italiana *porta*, por exemplo, não parece apresentar nenhuma diferença em termos de qualidade fonética em relação ao latim *porta*, que lhe deu origem. Tal vocábulo não é menos vernáculo, por causa disso.

É interessante notar que as línguas européias ocidentais apresentam muitos elementos léxicos vernáculos que podem parecer à primeira vista estranhos às suas respectivas línguas-mães. Assim, há nas línguas românicas inúmeros vocábulos de origem germânica, e que seriam erroneamente classificados como empréstimos, visto que tais vocábulos não passaram diretamente de uma língua germânica para uma língua românica, mas, sim, passaram de um dialeto germance para o latim vulgar ou dialeto romance, e daí para a língua românica que é sua continuação histórica. O mesmo se deve dizer de palavras vernáculos das línguas germânicas que são, entretanto, de origem latina. Um exemplo tornará isso mais claro. Temos em francês a palavra *garder*, que provêm do franco *wardÇn*. Todavia, tal palavra não foi do franco diretamente ao francês, até porque essas duas línguas não coexistiram na mesma época histórica. O que verdadeiramente ocorreu foi o empréstimo do franco *wardÇn* ao latim vulgar, na forma *guardare*. Esta forma, por sua vez, é que produziu o fr. *garder*, it. *guardare*, port. e esp. *guardar*. Do mesmo modo, o al. *Kopf*, ing. *cup*, sue. *kopp*, etc., por exemplo, vêm do lat. *cuppa*. Contudo, tal empréstimo não foi efetuado do latim pelo alemão, inglês, etc. Na verdade, o lat. *cuppa* passou, ainda à época dos primeiros contatos entre os romanos e os bárbaros germânicos (sécs. I–III d.C.), ao gótico e demais germances, sendo por esses transmitido, por herança, às línguas germânicas modernas. Do que foi dito resulta que tanto o fr. *garder* quanto o al. *Kopf* são perfeitamente vernáculos. Não obstante, muitos trabalhos de pesquisa etimológica não dão conta de tal fato. O próprio latim clássico apresenta vocábulos vernáculos de origem alogenética, como por exemplo, *ambactus*, *cambiare*, *carrus*, *carruca*, de procedência céltica, porém ingressos no latim em época pré-histórica, quando coabitavam a Península Itálica povos falantes de dialetos paleolatinos (como o falisco, o prinstino e o próprio latim) e povos falantes de dialetos célticos, como os volscos e gauleses, dentre outros.

Vulgarismos alogenéticos (empréstimos de vulgarismos estrangeiros)

Um vulgarismo alogenético é um vocábulo resultante do empréstimo de um vulgarismo estrangeiro. Esse empréstimo se dá sempre em etapas. Segundo Guilbert¹⁰, num primeiro momento, o termo estrangeiro é introduzido num determinado ato de fala em referência a um significado próprio à língua estrangeira. É o que ele chama de *xenismo*, isto é, termo que permanece efetivamente estrangeiro. Nessa categoria entram inicialmente todos os nomes próprios de pessoas, nomes geográficos de rios, de cidades, bem como todas as palavras da língua que exprimem realidades que não possuem correspondente na língua do falante ou que são voluntariamente integrados por ele à sua elocução justamente para ressaltar seu caráter estrangeiro. O empréstimo propriamente dito se inicia a partir do momento em que se introduz na sociedade o objeto ou conceito designado inicialmente na língua estrangeira, de tal modo que a comunidade lingüística acolha ao mesmo tempo o termo que o designa, ou então quando ela recorre ao termo estrangeiro em referência a um significado já denominado em sua própria língua. A essa nova situação, em que o termo ainda conserva seu caráter estrangeiro, mas já se incorporou aos hábitos lingüísticos dos falantes, Guilbert dá o nome de *peregrinismo*. Segundo ele, o empréstimo é a fase posterior, em que, dada a generalização de seu uso, o termo se integra de tal modo ao léxico, a ponto de não ser mais sequer percebido como termo estrangeiro.

Retomando os conceitos de *xenismo*, *peregrinismo* e *empréstimo*, acima definidos, diremos que o *xenismo* se caracteriza como um elemento de baixa freqüência e distribuição irregular, e que não se integrou ainda ao léxico da língua em que ocorre. Muitos *xenismos* podem mesmo ocorrer uma única vez, num único discurso, ou ser reempregados algumas vezes, durante certo tempo, e depois ser abandonados. Isso se deve ao fato de que muitas palavras estrangeiras são objeto de modismos passageiros. Podemos citar como exemplos atuais de *xenismos* em português do Brasil as palavras inglesas *soft*, *light*, *clean*, dentre outras.

¹⁰ GUILBERT, L. *La créativité lexicale*. Paris, Larousse, 1975, p. 92-93.

Quando uma palavra estrangeira, havendo ingressado na língua como xenismo, passa a ser reiterada em novos discursos, aumentando de frequência e assumindo distribuição regular, ela acaba por integrar-se ao léxico da língua, embora ainda conserve seu caráter estrangeiro, o que é atestado pela pronúncia, em geral estranha ao sistema fonético-fonológico da língua, e, principalmente, pela grafia discordante das regras ortográficas da língua, além de ser convencionalmente grafada em caracteres itálicos. Temos aí o peregrinismo. Tal é o caso de palavras como *premier* (= *primeiro ministro*), *habitué*, *surf*, *whisky*, etc.

Quando, finalmente, um peregrinismo perde totalmente seu caráter estrangeiro, através da adaptação fonético-fonológica, ortográfica e, principalmente, morfossintática ao padrão da língua que o abriga, está definitivamente consumado o empréstimo. Esse é o caso de *futebol*, *redingote* e *abacaxi*, por exemplo.

Em virtude de o xenismo não pertencer ao léxico da língua, entendemos não deva — na verdade, não possa — o mesmo ser classificado lexicogenicamente. Já o peregrinismo e o empréstimo estrangeiro nacionalizado fazem efetivamente parte do léxico da língua, razão pela qual são ambos classificáveis como vulgarismos aloenéticos.

Convém notar que existem duas diferenças básicas entre o empréstimo de um vocábulo grego ou latino (grecismo ou latinismo) e o empréstimo de um vocábulo vulgar estrangeiro (estrangeirismo):

a) Embora palavras e expressões latinas não vernaculizadas (por exemplo, *statu quo*, *ipso facto*, *ibidem*, etc.) possam ser introduzidas na língua como xenismos ou peregrinismos, um grecismo ou latinismo, em geral, sofre sempre um processo de adaptação fonético-fonológica, morfossintática e ortográfica, que o integra perfeitamente, desde sua introdução na língua, no léxico da mesma. Já, em relação ao vulgarismo estrangeiro, isso quase nunca acontece de imediato, mas, na verdade, o estrangeirismo deve passar por todas as etapas de sua adoção.

b) Um significante grego ou latino pode não ser importado juntamente com seu(s) significado(s), mas pode adquirir, já de início, um novo significado na língua que importa o significante. Ao contrário, o estrangeirismo penetra na língua sempre como um vocábulo completo, com significante e significado estrangeiros, embora, por se tratar justamente de um vocábulo, unidade lexical

de norma discursiva, por definição monossêmico e monoisotópico, esse significado — que, em relação ao vocábulo, também se denomina *acepção* — seja, conseqüentemente, único, até porque, conforme vimos em tipos de empréstimos, p. 58-59, o empréstimo de um estrangeirismo pressupõe a adoção de uma das acepções da lexia original, com a concomitante filtragem das demais.

Quanto mais antigo é o empréstimo de um estrangeirismo, maior a probabilidade de que este já se tenha nacionalizado. Assim, a maior parte dos galicismos da língua inglesa penetraram nesta durante o período de dominação francesa sobre a Inglaterra, o que se deu na Baixa Idade Média, entre os anos de 1066 e 1399. Por isso mesmo, tais galicismos já não são mais, há vários séculos, sentidos pelos falantes do inglês como palavras estrangeiras. Além de termos como *change, channel, city, country, cover, danger, money, point, royal, valley, value, vine*, etc., também elementos gramaticais, tais como os sufixos *-ty* e *-ous*, depreendidos de galicismos como *pity* e *nervous*, passaram a fazer parte do acervo gramatical do inglês, entrando posteriormente em novos derivados, muitos deles autóctones, como, por exemplo, *safety* e *flatterous*. Interessante é notar que, no caso específico do inglês, é tal a quantidade de galicismos, que mesmo empréstimos mais recentes e ainda não totalmente nacionalizados, tais como *fiancé, fête, naïve, amateur*, dentre outros, nem por isso são grafados em caracteres itálicos, o que demonstra a grande receptividade que os vocábulos franceses continuaram tendo no léxico da língua inglesa.

Por outro lado, há línguas extremamente mais resistentes à adoção do estrangeirismo, como é o caso do italiano e do espanhol, por exemplo. Nestas, só recentemente parece estar-se subvertendo essa tendência, por força da influência das novas tecnologias e das telecomunicações, o que tem feito tornarem-se mais comuns vocábulos como *software, hardware, computer, charter*, além de outros, já internacionalizados, como *hotel, choc*, etc. Nessas línguas, a grande barreira à penetração de estrangeirismos é de índole fonológica: com efeito, numa língua como o italiano, em que a maioria absoluta das palavras termina por vogal, termos estrangeiros, de terminação consonantal, representam, sem dúvida, elementos de difícil assimilação pelo sistema da língua. Esses estrangeirismos tenderão, com certeza, a permanecer sempre como peregrinismos.

Alógenos que retornam à língua de origem

Tomemos o seguinte exemplo: o vocábulo francês antigo *desport*, “recreação”, deverbal regressivo de um antigo verbo *desporter*, “distrair”, passou ao inglês nas formas *disport* e *sport*, “esporte”, e daí para o italiano, o alemão, e, inclusive, o próprio francês. Hoje, o vocábulo *sport*, embora originário do francês, existe no léxico dessa língua como alógeno, ao lado do vernáculo *déport*, que significa “recusa, demora”. Da mesma forma, o francês antigo *test*, “vaso de prova”, que subsiste no francês moderno na forma *têt*, “crânio”, foi, ainda na Idade Média, emprestado ao inglês com seu significado original. Depois de ter caído em desuso em francês nessa acepção, tal vocábulo retorna à sua língua de origem por empréstimo do inglês. Portanto, *test*, vocábulo vernáculo em francês medieval, é atualmente vulgarismo alogênico nessa língua.

Empréstimo de vulgarismos com as respectivas desinências

Muitas vezes, quando uma língua importa um vocábulo de outra língua, o faz de forma integral, isto é, importa não apenas o radical do vocábulo, mas também suas desinências. Ocorre que tais desinências não têm na língua importadora função gramatical, como têm na língua de origem do vocábulo. Assim, tais morfemas acabam por integrar-se ao radical do vocábulo importado. Um exemplo tornará isso mais claro. Para tanto, tomemos o adjetivo francês *privé*, o qual nada mais é do que o particípio passado do verbo *priver*. Assim, a terminação *-é* desse vocábulo constitui uma desinência verbal, que se agrega ao radical verbal *priv-*. Tal vocábulo foi emprestado ao inglês na forma *privy*. Ocorre que em inglês a terminação *-y* não é desinência de particípio, assim como *priv-* não é radical verbal, uma vez que não existe em inglês um suposto verbo **to prive*. Desse modo, o radical do adjetivo inglês *privy* contém o elemento gramatical *-y*, originário da desinência francesa.

Esse processo de empréstimo de vocábulos juntamente com as desinências aconteceu em larga escala sobretudo com verbos franceses importados pelo alemão durante a Baixa Idade Média e

Renascença. Dessa forma, temos o al. *marschieren*, a partir do fr. *marcher*, o al. *garantieren*, a partir do fr. *garantir*, etc. Nota-se, assim, que as terminações francesas *-er* e *-ir*, desinências de infinitivo, deram origem em alemão a um sufixo verbal *-ier-*, uma vez que a desinência alemã de infinitivo é *-en*. Tal sufixo tornou-se posteriormente produtivo em alemão, sendo muitíssimo usado junto a radicais verbais latinos, como, por exemplo, em *applaudieren*, *demonstrieren*, *denunzieren*, *formulieren*, *informieren*, *promovieren*, *prozessieren*, etc.

Também o inglês algumas vezes importou verbos franceses com as respectivas desinências de infinitivo: fr. *dominer* → ing. *to domineer*, fr. *flatter* → ing. *to flatter*, fr. *rendre* → ing. *to render*, etc.

Empréstimos de tradução

Partindo das definições de *empréstimo total*, *empréstimo de significante* e *empréstimo de significado* que demos na p. 59, podemos definir empréstimo de tradução como um empréstimo de significado em que o significante é vulgar, isto é, total ou parcialmente constituído de morfemas vernáculos ou metamorfizados. Por tratar-se de um tipo de empréstimo, a tradução do significante é sempre intencional, motivada pelo vocábulo original. Isto significa que o criador do empréstimo de tradução conhece a língua cujo vocábulo deve ser traduzido, e, além disso, opta conscientemente pela tradução do mesmo. Faz-se necessária essa observação, visto que há vocábulos que, à primeira vista, parecem traduções de vocábulos estrangeiros, mas que em verdade não o são, pois falta-lhes justamente o caráter de intencionalidade. Trata-se apenas de soluções análogas, porém independentes. Por exemplo, o ing. *go out* ou o al. *ausgehen* poderiam parecer traduções do lat. *exire* (*ex* ≈ *out* ≈ *aus*, *ire* ≈ *go* ≈ *gehen*). Entretanto, *go out* e *ausgehen* foram cunhados de forma totalmente independente do latim, e certamente em época anterior ao contato entre romanos e germanos. Da mesma forma, o ing. *fishy* tem a mesma composição semântica do lat. *piscosus*. Todavia, *fishy* não foi motivado por *piscosus*, mas, sim, por *fish* e pelo sufixo *-y*, depreensível a partir de uma série de pares existentes em inglês, tais como *rain/rainy*, *dirt/dirty*, *salt/salty*. Chamaremos tais casos de *falsas traduções*.

Entre as falsas traduções, podemos citar, por exemplo, inúmeras designações de utensílios, instrumentos, profissões, etc., em que a escolha do significante é motivada pelo próprio significado do vocábulo, sendo natural, assim, que o significado de tais vocábulos apresente a mesma forma semântica em diversas línguas. É o caso, por exemplo, do português *saca-rolhas* em face do francês *tire-bouchons*, ou do português *leiteiro* em face do inglês *milkman*. Note-se, entretanto, que, exatamente por serem motivados pelo seu próprio significado, muitos vocábulos semanticamente equivalentes apresentam formas semânticas diferentes, por exemplo, port. *motorista* (motivado por *motor*) x fr. *chauffeur* (motivado por *chauffer*, “aquecer”), port. *padeiro* (motivado por *pão*) x ing. *baker* (motivado por *bake*, “assar”). Isso demonstra que casos como o de *leiteiro* em face de *milkman* não constituem realmente traduções, mas sim resultados coincidentes.

Os empréstimos de tradução podem ser *sintagmáticos* ou *semânticos*, *diretos* ou *indiretos*, *obrigatórios* ou *facultativos*, e, ainda, podem partir de um grecolatinismo ou de um vulgarismo.

Empréstimo de tradução sintagmático

O empréstimo de tradução sintagmático é um neologismo sintagmático (composto ou derivado) vulgar criado para traduzir um vocábulo estrangeiro, o qual, em geral, também é composto ou derivado. Normalmente, substituem-se um a um os morfemas do vocábulo original por seus correspondentes na língua receptora do empréstimo. Temos assim o *empréstimo de tradução total*. Mesmo que alguns desses morfemas (ou todos) não tenham exatamente o mesmo significado nas duas línguas, isso não descaracteriza a tradução se há o caráter de motivação em relação ao vocábulo original. Assim, por exemplo, comparando-se o lat. *describere* e o al. *beschreiben*, a correspondência entre *scribere* e *schreiben* é patente. Poder-se-ia, porém, argüir que os prefixos *de-* e *be-* não significam a mesma coisa em latim e alemão, mas, ainda assim, a motivação do modelo de *describere* na criação de *beschreiben* resta evidente, o que é comprovado por dados históricos e filológicos referentes à gênese desse vocábulo em alemão.

Pode ocorrer, entretanto, que nem todos os morfemas de um vocábulo emprestado sejam traduzidos, mas que alguns sejam

mantidos em sua forma original. Nesse caso, o empréstimo de tradução será dito *parcial*. Por exemplo, no fr. *promouvoir*, tradução do lat. *promovere*, traduz-se *-movere* por *-mouvoir*, mas mantém-se o prefixo *pro-* inalterado (a forma vernácula seria *pour-*). Da mesma forma, temos o fr. *naturel* em face do lat. *naturale*, ou o port. *relação* em face do lat. *relatione*. Quando a tradução é parcial, pode ocorrer que o vocábulo tradutor não tenha exatamente o mesmo significado do vocábulo traduzido, embora deva manter com este uma intersecção semântica: é o caso do port. *computação*, tradução do lat. *computatione*, que, no discurso da informática, significa algo um tanto diferente do que significa originalmente em latim. Ocorre que o termo português emprestou seu significado informático do inglês *computation* (mais tarde substituído por *computing*). Pode-se dizer, nesse caso, que o vocábulo português é empréstimo de tradução de um grecolatinismo inglês, e não latino.

Vale a pena chamar a atenção para o fato de que, nos trabalhos tradicionais de filologia e etimologia, apenas aquilo que denominamos *empréstimo de tradução total* é considerado como empréstimo de tradução. Desse modo, vocábulos como *contemplação* ou *televisão* ou são ambos classificados como eruditos, com total desconsideração do sufixo vernáculo, ou, ainda, é classificado o primeiro como erudito, por existir em latim, e o segundo como empréstimo do inglês, sendo então equiparado a vocábulos de natureza lexicogênica completamente diferente, como *nocaute* (do ing. *knockout*), por exemplo.

Empréstimo de tradução semântico

O empréstimo de tradução semântico é um neologismo semântico vulgar criado para traduzir um vocábulo grego ou latino. Consiste, portanto, em atribuir-se a um significante vulgar já existente na língua um novo significado, emprestado a um vocábulo grego, latino, ou a um grecolatinismo de outra língua vulgar. A escolha do significante tradutor em geral não é arbitrária, mas parte de uma relação semântica entre os vocábulos tradutor e traduzido. Um exemplo tornará isso mais claro. Tomemos o lat. *angulus*, que significa “canto de uma casa, esquina” e também “ângulo (conceito geométrico)”. O alemão possuía uma palavra

vernáculo correspondente, *Ecke*, apenas para a primeira acepção. Quando a língua alemã, já estabelecida como idioma de cultura, necessitou de um termo para dar conta da segunda acepção, de caráter erudito, tomou ao latim essa acepção, atribuindo-a ao significante vernáculo *Ecke*, em razão da acepção comum de *angulus* e *Ecke*, “canto, esquina”.

Contudo, é mister assinalar que nem todo empréstimo de significado configura tradução. Com efeito, quando uma língua vulgar realiza empréstimo de significado de vocábulo de outra língua vulgar, na qual esse vocábulo é neologismo semântico — que poderíamos considerar como empréstimo de significante intralingüístico —, não temos empréstimo de tradução, mas sim *empréstimo de sentido*.¹¹ Assim, se em inglês *tree* significa inicialmente “árvore”, e, posteriormente, ganha, no discurso tecnológico, o significado adicional de “eixo rotativo que transmite energia” (cfr. *árvore de transmissão*), e se ao vernáculo português *árvore*, “árvore”, se acrescenta a acepção tecnológica tomada de empréstimo ao inglês, a palavra portuguesa *árvore* continuará vernáculo em qualquer das duas acepções, pois houve apenas empréstimo do sentido inglês. Observe-se que este último caso é diferente do al. *Ecke* em relação ao lat. *angulus*, pois as duas acepções de *angulus* já existiam em latim quando do empréstimo da segunda pelo alemão, ao passo que a criação da acepção tecnológica de *tree/árvore* é praticamente simultânea em ambas as línguas.¹²

Nos diagramas a seguir, procuraremos exemplificar graficamente os mecanismos de empréstimos de tradução semântico e de

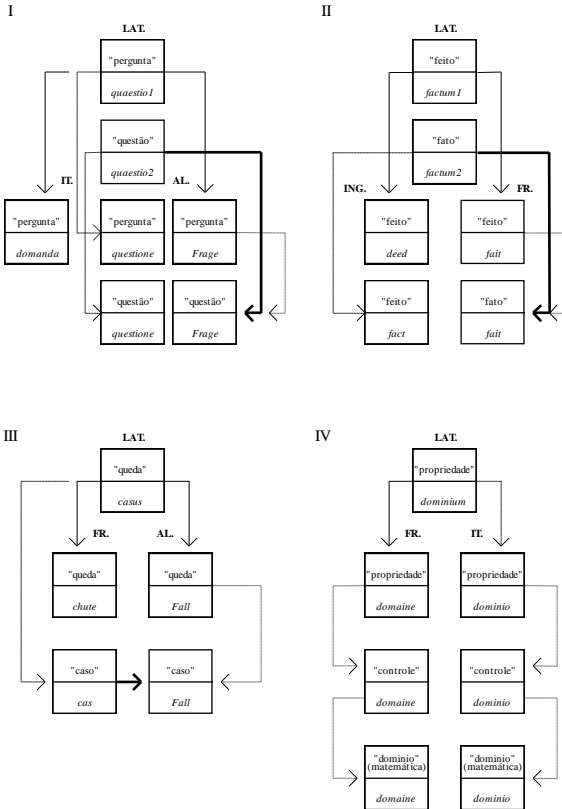
¹¹ A respeito da designação *empréstimo de sentido*, cfr. MAURER JR., T. H. *A unida-de da România ocidental*. São Paulo, FFCL-USP, 1951, p. 76. Embora consideremos tal designação terminologicamente inadequada, por ser ambígua e pouco precisa, optamos por mantê-la, por razões tradicionais e para não confundir o fenômeno por ela designado com o que chamamos de *empréstimo de significado*.

¹² Poder-se-ia acusar, com justa razão, tal critério de classificação de arbitrário. Entretanto, o ponto de vista de nossa análise lexicogênica parte da relação entre as línguas clássicas e as línguas vulgares. Assim, quando estas últimas vão buscar no léxico daquelas os formantes de seus neologismos eruditos, encontram um léxico “fechado”, até porque o grego clássico e o latim são línguas mortas. Portanto, se um dia a acepção geométrica de *angulus* — que, aliás, foi importada do grego — constituiu neologismo semântico em latim, tal fato estava totalmente perdido de vista quando as línguas vulgares européias foram utilizadas pela primeira vez em algum tratado de geometria, por exemplo.

empréstimo de sentido. Para tanto, adotaremos as seguintes convenções:

- herança vernácula do latim ou equivalência entre o vocábulo latino e o vulgar
- → empréstimo direto do latim
-→ neologia semântica (ressemantização)
- empréstimo de tradução

Convencionaremos, também, que a língua vulgar situada à esquerda no diagrama realiza os processos lexicogênicos de empréstimo ao latim ou neologia semântica, que aqui chamamos também de ressemantização, em primeiro lugar, razão pela qual o eixo cronológico em nossa representação gráfica vai sempre da esquerda para a direita.



No diagrama I, temos o lexema latino *quaestio*, ao qual correspondem (pelo menos) dois vocábulos, com as acepções “pergunta” e “questão, problema”, respectivamente. O italiano e o alemão possuíam vocábulos vernáculos, respectivamente *domanda* e *Frage*, correspondentes à acepção latina “pergunta”. Posteriormente, o italiano emprestou ao latim a palavra *questione*, com ambas as acepções, isto é, “pergunta” e “questão”, ao passo que o alemão ressemantizou o significante de *Frage* para dar conta da acepção “questão”.

No diagrama II, temos o lexema latino *factum*, com as acepções “feito, façanha” e “fato, ocorrência”. Correspondente à primeira acepção, temos o fr. *fait* (herança de *factum*) e o ing. *deed*. Para a segunda acepção, de caráter mais erudito, o inglês empresta o vocábulo latino, na forma *fact*, enquanto o francês realiza neologia semântica de *fait*, configurando, assim, o empréstimo de tradução semântico.

No diagrama III, temos uma situação semelhante à anterior, porém agora o latim *casus* possui apenas uma acepção, “queda”¹³, à qual correspondem os vernáculos fr. *chute* e al. *Fall*. Quando as línguas vulgares, dentre elas o francês, sentem a necessidade de dar conta de uma nova acepção, “caso, situação”, emprestam o significante *casus* ao latim (por exemplo, fr. *cas*). Já o alemão ressemantiza *Fall*. Neste caso, portanto, o empréstimo de tradução se dá em relação aos grecolatinismos fr. *cas*, it. *caso*, ing. *case*, etc., e não propriamente em relação ao latim.

No diagrama IV, temos o lat. *dominium*, “propriedade, moradia do *dominus*”. Em francês, temos um vernáculo correspondente, *domaine*, e, em italiano, um empréstimo do latim, *dominio*. A seguir, uma dessas duas línguas, que convencionamos, por uma questão meramente prática, tenha sido o francês, ressemantiza o vocábulo com a acepção “controle, poder, esfera de ação”. Esse neologismo semântico é imediatamente adotado pelas outras línguas (em nosso exemplo, o italiano), configurando o empréstimo de sentido do vocábulo francês. Em seguida, o fr. *domaine* ganha nova acepção, no discurso da matemática, qual seja, “conjunto dos valores que podem ser assumidos pelas variáveis independentes de uma função”. Ocorre então novo empréstimo de sentido pelo italiano. Neste caso, houve dois empréstimos de sentido e nenhum empréstimo de tradução, razão pela qual se torna irrelevante saber em qual das duas línguas se deu a neologia semântica original e qual efetuou o empréstimo de sentido correspondente.

Se observarmos os diagramas acima, notaremos que sempre que temos, na língua à esquerda do diagrama, um empréstimo ao latim, e, na língua à direita, um neologismo semântico vulgar

¹³ Na verdade, *casus* possui outras acepções, inclusive como termo de gramática, que estamos aqui desconsiderando, por não interessarem ao processo de tradução semântica objeto de nosso exemplo.

correspondente, temos, na língua à direita, também, um empréstimo de tradução semântico, o que corresponde exatamente à nossa definição desse processo lexicogênico. Já, quando temos neologia semântica de ambos os lados, temos o empréstimo de sentido.

Empréstimos de tradução diretos e indiretos

Até o momento discutimos os chamados empréstimos de tradução diretos, ou seja, aqueles que provêm diretamente de vocábulo estrangeiro, por substituição do significante original por outro, indígena. Entretanto, há vocábulos que podem constituir-se em empréstimos de tradução por via indireta. Se um vocábulo, dito primitivo, pertencente a uma determinada língua A possui um derivado nessa língua, e se a esse vocábulo primitivo corresponde numa língua B um vocábulo igualmente primitivo, que, por sua vez, admite um derivado vulgar, este derivado pode ser um empréstimo de tradução indireto do derivado de A, se satisfizer a uma das seguintes condições:¹⁴

a) O vocábulo primitivo de B é tradução do vocábulo primitivo de A. Exemplos:

lat. <u>condemnare</u>	→	it. <u>condannare</u>	(tradução direta)
<u>condemnatio</u>		<u>condanna</u>	(tradução indireta)
lat. <u>accipere</u>	→	al. <u>annehmen</u>	(tradução direta)
<u>acceptio</u>		<u>Annahme</u>	(tradução indireta)
lat. <u>angulus</u>	→	al. <u>Ecke</u>	(tradução direta)
<u>angularis</u>		<u>eckig</u>	(tradução indireta)

b) O vocábulo primitivo de B é empréstimo, vernaculizado ou não, do vocábulo primitivo de A. Exemplos:

lat. <u>natura</u>	→	fr. <u>nature</u>	(empréstimo não vernaculizado)
<u>naturalis</u>		<u>naturel</u>	(tradução indireta)

¹⁴ Os exemplos por nós apresentados referem-se a traduções a partir do latim; contudo, essas condições são válidas para quaisquer línguas.

lat. <u>persona</u> <u>personalis</u>	→	al. <u>Person</u> (empréstimo não vernaculizado) <u>persönlich</u> (tradução indireta)
lat. <u>regula</u> <u>regularis</u>	→	al. <u>Regel</u> (empréstimo vernaculizado) <u>regelmäßig</u> (tradução indireta)
lat. <u>particula</u> <u>particularis</u>	→	it. <u>particola</u> (empréstimo vernaculizado) <u>particolare</u> (tradução indireta)

OBSERVAÇÃO: Se o vocábulo primitivo de B é vernáculo, o derivado não é necessariamente tradução. Exemplos:

lat. <u>aurum</u> <u>aureus</u>	→	ing. <u>gold</u> <u>golden</u>
lat. <u>piscis</u> <u>piscosus</u>	→	ing. <u>fish</u> <u>fishy</u>

Empréstimos de tradução obrigatórios e facultativos

De modo geral, o empréstimo de tradução é um procedimento facultativo — e portanto voluntário — de substituição de morfemas de um vocábulo importado por correspondentes vernáculos. Por seu caráter de não obrigatoriedade, esse tipo de tradução, quando ocorre, revela, por parte de quem a realiza, um sentimento lingüístico “nativista”, se nos é lícito empregar esse termo. Essa modalidade de tradução costuma afetar tanto os radicais quanto os afixos do vocábulo, e ocorre quer pela proximidade fonética entre o morfema tradutor e o morfema traduzido (lat. *comprobare* → port. *comprovar*¹⁵, ou lat. *litteratura* → it. *letteratura*, por influência do vernáculo *lettera*), quer pela necessidade de enquadrar o empréstimo nos paradigmas de flexão da língua (lat. *praevidere* → port. *prever*, enquadrável no paradigma de conjugação de *ver*), quer ainda pelo simples sentimento nativista a que acima nos referimos, e que conduz à substituição de morfemas por outros, sem nenhum parentesco genético ou semelhança fonética, apenas

¹⁵ Compare-se com port. *comprobatório* ← lat. *comprobatorius*.

por haver entre eles equivalência semântica (lat. *expressio* → al. *Ausdruck*).

Entretanto, em muitos casos, a tradução de um empréstimo torna-se obrigatória, pela inconveniência de se introduzir na língua importadora um elemento gramatical que lhe seria estranho. Por exemplo, o sufixo deverbal latino *-atione* é sistematicamente traduzido, em português pelo correspondente vernáculo *-ação*, ao passo que o francês, que já possuía o sufixo vernáculo *-aison*, introduz, posteriormente, o sufixo erudito correspondente *-ation*. Assim, a tradução portuguesa do lat. *comparatione* por *comparação* é obrigatória, enquanto a tradução francesa *comparaison* é facultativa (compare-se com as formas francesas *nation*, *transformation*, *éducation*, etc., em que não há tradução). A tradução obrigatória afeta em geral apenas os morfemas gramaticais (gramemas), especialmente os sufixos.

Empréstimos de tradução de grecolatinismos

O empréstimo de tradução de grecolatinismo é aquele que se origina de vocábulo grego ou latino, ou de vocábulo proveniente de qualquer outra língua vulgar, na qual dito vocábulo seja exclusivamente formado de elementos greco-latinos. Assim, o port. *contemplação* traduz o vocábulo latino *contemplatione*, ao passo que *televisão* é tradução do grecolatinismo inglês *television*, composto do grego *tele-* e do latim *-visione*. Por sinal, boa parte dos empréstimos de tradução que encontramos nas línguas germânicas, especialmente na norma técnico-científica, provêm de grecolatinismos criados em outras línguas vulgares (por exemplo, o al. *Wasserstoff*, tradução do fr. *hydrogène*, criado a partir do gr. *hydor* e *genos*).

É interessante notar que boa parte dos próprios vocábulos latinos de caráter culto são traduções do grego. Isso se deve ao grande prestígio da língua e da cultura gregas em todo o mundo antigo, particularmente em Roma, principal beneficiária e continuadora da civilização helênica. Termos técnicos, científicos, filosóficos, religiosos, termos de gramática e retórica, termos médicos, enfim, elementos vocabulares dos mais diversos domínios do conhecimento, desenvolvidos pelos gregos, foram trans-

portados em massa para o léxico do latim, alguns em sua forma original, outros na forma de traduções. Vejamos alguns exemplos:

GREGO

ἄμορφος

ἄτομος

διάλογος

ἔντομον

μεταμόρφωσις

μεταφορά

μονογία

πρόλογος

σιμμετρία

σιμπάθεια

συμφωνία

συνείδεσις

τρίγωνον

ὑπόθεσις

LATIM

*informis**individuus**colloquium**insectum**transformatio**translatio**soliloquium**praeafatio**commensuratio**compassio**consonantia**conscientia**triangulum**suppositio*

A introdução de empréstimos de tradução de vocábulos gregos e latinos nas línguas vulgares europeias — assim como também a introdução de grecolatinismos puros — foi particularmente freqüente durante a Baixa Idade Média (sécs. XI–XV), quando as novas línguas de cultura que então emergiam precisaram dar conta de uma realidade que até então só havia encontrado expressão nas línguas clássicas. Diz a esse respeito Malmberg:

Quando os povos germânicos, os “bárbaros” tiveram que adquirir dos povos do sul e da Igreja conceitos que não tinham, pois o mundo ideológico da religião cristã lhes era alheio, tomaram emprestadas todas as palavras correspondentes, com sua expressão e conteúdo, mas adaptando-as primeiro a seus próprios sistemas fonéticos e morfológicos. (...) Mas formaram também novas palavras com material léxico próprio seguindo modelos estrangeiros: do latim *expressio* o alemão *Ausdruck*; do latim *impressio* o alemão *Eindruck*; do latim *conscientia* o sueco *samvete*. É o que chamamos empréstimos de tradução.¹⁶

¹⁶ MALMBERG, B. *A língua e o homem*. Introdução aos problemas gerais da lingüística. Rio de Janeiro, Nórdica/São Paulo, Duas Cidades, 1976, p. 79.

Deve-se notar, todavia, que, embora o principal veículo da influência greco-latina tenha sido inegavelmente a Igreja e sua língua oficial, o latim, o influxo da cultura greco-romana sobre a Europa Ocidental medieval transcendeu de muito o campo exclusivamente religioso, conforme atesta o próprio Malmberg:

O Ocidente é, em muitos aspectos, uma unidade, sob o ponto de vista cultural e da estrutura social. A educação ocidental construiu-se sobre bases grego-latinas. Todo o nosso léxico científico, e, em geral, filosófico e 'espiritual', vem das línguas clássicas ou está formado sobre modelos clássicos. As línguas escritas do Ocidente foram elaboradas em grande parte por quem conhecia o latim e tinha o latim como modelo (*Idem, ibidem*).

A necessidade de tornar as línguas vulgares instrumentos de expressão intelectual e literária apoiava-se num conflito constante entre latinidade e vernaculidade. Com efeito, durante toda a Idade Média, e ainda na Renascença, a maior parte dos textos escritos utilizavam exclusivamente a língua latina. Esta era a língua dos padres, dos filósofos, dos professores, dos magistrados, dos médicos... O primeiro desafio das línguas vulgares era o de afirmar-se como sistemas lingüísticos capazes de fazer frente a essa língua, que desfrutava de prestígio e credibilidade milenares. Assim, a introdução na língua vulgar de conceitos abstratos, de caráter erudito, provenientes do latim, implicava a necessidade da escolha de um significante para tais conceitos, escolha esta que era o próprio cerne do conflito acima referido. Adotar o próprio termo latino, com algumas adaptações fonéticas e gramaticais, ou traduzi-lo, era a questão que se propunham os gramáticos e intelectuais da época. Se, por um lado, a tradução de termos latinos dava maior transparência aos mesmos, por outro lado, nem sempre era possível encontrar exata correspondência entre os elementos a serem traduzidos e seus supostos equivalentes vernáculos. Ao mesmo tempo, a necessidade de conferir à língua vulgar a mesma confiabilidade e precisão do latim, além da influência do espírito do Humanismo, aconselhavam a opção por um léxico latinizado.

Cumprе ressaltar que a tendência à tradução de vocábulos latinos não deve ser tomada como um traço absoluto de uma determinada língua, mas, antes, como um fenômeno lingüístico relativo a uma época e a um contexto sócio-político-cultural.

Assim sendo, observa-se no desenvolvimento léxico das línguas européias um período de predominância da influência latina, seguido de uma fase de predominância francesa, e, posteriormente, inglesa. Não obstante, essa tendência à tradução assumida por certas línguas permanece de modo geral ainda hoje, quer de forma residual, quer porque já se incorporou à estrutura léxica da língua.

Vejamos, a seguir, alguns exemplos de traduções de vocábulos latinos, nas diversas línguas européias:

LATIM

concurrere
consecrare
emotio
producere
superponere

LATIM

cognomen
conducere
obtinere
populatio
supponere

LATIM

abstractus
combinatio
doctor
equinus
infinitus
mundialis
peninsula
perfectus
providere
realis
reducere
subtrahere
superficies
verisimilis

PORTUGUÊS

concorrer
consagrar
emoção
produzir
sobrepor

ESPAÑHOL

cognombre
conllevar
obtener
población
suponer

FRANCÊS

abstrait
combinaison
docteur
chevalin
infini
mondial
presqu'île
parfait
pourvoir
réel
reduire
soustraire
surface
vraisemblable

LATIM

aquaticus
condemnare
confirmare
consequentia
conscientia
crucifixus
districtus
extraordinarius
multiplicare
mundialis
naturalitas
notabilis
planeta
projectare
reducere
supervivere

LATIM

continere
duplex
felinus, felideus
femininus
fraternitas
humanitas
masculus, masculinus
maternus
multiplex
participare
paternus
patria
perdonare
praecursor
praedicere
praefatio
praevidere
progressus
subire
subjacere

ITALIANO

acquatico
condannare
confermare
consequenza
consapevolezza
crocifisso
distretto
straordinario
moltiplicare
mondiale
naturalezza
notevole
pianeta
progettare
ridurre
sopravvivere

INGLÊS

withhold
twofold
catlike
womanly
brotherhood
mankind, manhood
manly
motherly
manifold
partake
fatherly
fatherland
forgive
forerunner
foretell
foreword
foresee
headway
undergo
underlie

LATIM

aequatio
circumdare
circumstantia
collaborare
communis
conscientia
distantia
elaborare
exceptio
excludere
extraordinarius
includere
influentia, influxus
inscribere
introducere
multiplex
participare
peninsula
percipere
percurrere
perdonare
persequi
possibilis
praesidens
proprietas
providere
repetere
satisfacere
scientia
superficies
superfluus
transferre
transitio
triangulum
unitas

LATIM

collaborare

ALEMÃO

Gleichung
umgeben
Umstand
mitarbeiten
gemein, gemeinsam
Gewissen
Abstand
ausarbeiten
Ausnahme
ausschließen
außerordentlich
einschließen
Einfluß
einschreiben
einleiten, einführen
vielfach
teilnehmen
Halbinsel
vernehmen
durchlaufen
vergeben
verfolgen
möglich
Vorsitzer
Eigenschaft
versehen
wiederholen
genug tun
Wissenschaft
Überfläche
überflüßig
übertragen
Übergang
Dreieck
Einheit

SUECO

samarbeta

<i>conscientia</i>	<i>samvete</i>
<i>descriptio</i>	<i>beskrivelse</i>
<i>elaborare</i>	<i>utarbeta</i>
<i>excludere</i>	<i>utesluta</i>
<i>expressio</i>	<i>uttryck</i>
<i>extraordinarius</i>	<i>utomordentlig</i>
<i>impressio</i>	<i>intryck</i>
<i>influens</i>	<i>inflytelse</i>
<i>peninsula</i>	<i>halvö</i>
<i>percipere</i>	<i>förnimma</i>
<i>persequi</i>	<i>förfölja</i>
<i>praecedere</i>	<i>föregå</i>
<i>praeferre</i>	<i>föredra</i>
<i>progressus</i>	<i>framsteg</i>

É interessante notar que também no domínio da gramática ocorrem traduções, como, por exemplo, a preposição alemã *ausgenommen*, a partir do latim *excepto*, “exceto”.

As traduções obrigatórias, que, como vimos, resultam em geral da inconveniência — ou, mesmo, impossibilidade, por contrariar a estrutura fonológica da língua — de introduzir elementos morfológicos estranhos ao sistema, atingem, por isso mesmo, principalmente os afixos. Um exemplo disso é a inexistência, em francês, de proparoxítonos, razão pela qual sufixos como *-abile*, *-ariu*, *-tudine*, são sistematicamente substituídos, nessa língua, pelos vulgares *-able*, *-aire*, *-tude*, etc. Mesmo quando uma determinada língua dispõe de um afixo erudito, pode ocorrer que o afixo vulgar correspondente seja de uso mais freqüente, justamente por ser sentido como mais familiar pelos falantes. Damos, a seguir, uma breve lista dos principais afixos greco-latinos que costumam sofrer tradução nas diversas línguas vulgares:

OBSERVAÇÃO: O asterisco indica tradução facultativa; o travessão indica a inexistência de afixo vulgar.

PREFIXOS

LATIM	PORTUGUÊS	FRANCÊS	ITALIANO	INGLÊS	ALEMÃO
<i>sub-</i>	–	<i>sous</i> -*	<i>so-, sotto-</i>	<i>under</i> -*	<i>unter-</i>
<i>super-</i>	<i>sobre</i> -*	<i>sur</i> -*	<i>sopra</i> -*	<i>over</i> -*	<i>über-</i>
<i>contra-</i>	–	<i>contre-</i>	<i>contro</i> -*	<i>counter-</i>	<i>gegen-, wider-</i>
<i>anti-</i>	–	–	–	–	<i>gegen, wider-</i>
<i>archi-</i>	–	–	<i>arci</i> -*	–	<i>erz-</i>
<i>per-</i>	–	<i>par</i> -*	–	–	<i>ver-</i>
<i>pro-</i>	<i>por</i> -*	<i>pour</i> -*	–	–	<i>fort-</i>
<i>prae-</i>	–	–	–	<i>fore</i> -*	<i>vor-</i>
<i>post-</i>	–	–	–	–	<i>nach-</i>
<i>com-, con-, co-</i>	–	–	–	–	<i>mit-, zusammen-</i>
<i>circum-</i>	–	<i>circon-</i>	<i>circon-</i>	–	<i>um-</i>
<i>inter-</i>	<i>entre</i> -*	<i>entre</i> -*	<i>tra</i> -*	<i>enter</i> -*	<i>unter</i> -*
<i>auto-</i>	–	–	–	<i>self-</i>	<i>selbst-</i>
<i>in-(neg.)</i>	–	–	–	<i>un</i> -*	<i>un-</i>
<i>in-, intro-</i>	–	–	–	–	<i>ein</i> -*, <i>hinein</i> -*
<i>ex-</i>	–	–	<i>s</i> -*	–	<i>aus</i> -*
<i>extra-</i>	–	–	<i>stra-</i>	–	<i>außer-</i>
<i>de-</i>	–	–	<i>di</i> -*	–	<i>ab</i> -*
<i>re-</i>	–	–	<i>ri</i> *, <i>ra</i> -*	–	<i>er-, wieder-</i>

SUFIXOS

LATIM	PORTUGUÊS	FRANCÊS	ITALIANO	INGLÊS	ALEMÃO
-abilis	-ável	-able	-evole*	-able	-abel, -bar
-alis	—	-el*	—	—	-ell*
-arius, -aris	—	-aire, -ier	—	—	-är*
-orius	—	-oire	—	—	-orisch
-ivus	—	-if	—	—	—
-icus	—	—	—	—	-isch
-ator	-ador	-ateur, -eur	—	-er*	-er*
-anus	—	-ain, -en	—	—	-(an)isch
-tas	-dade, -tade	-té	-tà	-ty	-tät
-tudo	-tude	-tude	—	-tude	—
-atio	-ação	-aison*	—	—	—
-ia	—	—	—	—	-ei*
-ismus	—	—	-esimo*	—	—

Empréstimos de tradução de vulgarismos estrangeiros

O empréstimo de tradução de um vulgarismo estrangeiro possui as mesmas características do empréstimo de tradução de grecolatinismo, porém origina-se de um neologismo sintagmático vulgar — isto é, do qual pelo menos um dos morfemas é vulgar — pertencente a uma outra língua vulgar. Conseqüentemente, o empréstimo de tradução de vulgarismo só apresenta a modalidade sintagmática (Conforme vimos em Empréstimo de tradução semântico, na p. 83, quando a um neologismo semântico de uma determinada língua corresponde em outra língua outro neologismo semântico, não temos empréstimo de tradução, mas sim empréstimo de sentido). Além disso, só se pode falar de tradução de vulgarismo quando há efetivamente correspondência entre os morfemas do vocábulo original e do vocábulo traduzido. Caso contrário, tem-se simplesmente o empréstimo do significado, concomitante ao empréstimo do fato cultural que é seu referente, com a criação de um novo signo, para dar conta desse novo recorte cultural. Assim, o port. *computador* e o fr. *ordinateur* são traduções do ing.

computer, mas, sem dúvida, o port. *escavadeira* não é tradução do ing. *bulldozer*. Além disso, não é demais recordar que toda tradução é motivada, razão pela qual um vocábulo como port. *sapateiro* não é tradução do ing. *shoemaker* ou vice-versa, mas são apenas construções semelhantes, motivadas pelo mesmo fato cultural.

Vejam, então, alguns exemplos de empréstimos de tradução de vulgarismos:

FRANCÊS	ITALIANO	INGLÊS	ALEMÃO
<i>aéroport</i>	<i>aeroporto</i>	<i>airport</i>	<i>Flughafen</i>
<i>chef-d'œuvre</i>	<i>capolavoro</i>	<i>masterwork</i>	<i>Meisterwerk</i>
<i>chou-fleur</i>	<i>cavolfiore</i>	<i>cauliflower</i>	<i>Blumenkohl</i>
<i>développement</i>	<i>sviluppo</i>	–	<i>Entwicklung</i>
<i>entreprendre</i>	<i>intraprendere</i>	<i>undertake</i>	<i>unternehmen</i>
<i>fonctionnement</i>	<i>funzionamento</i>	<i>functioning</i>	<i>Funktionierung</i>
<i>gentilhomme</i>	<i>gentiluomo</i>	<i>gentleman</i>	–
<i>outré-mer</i>	<i>oltremare</i>	<i>overseas</i>	<i>Übersee</i>
<i>processement</i>	<i>processamento</i>	<i>processing</i>	<i>Prozessierung</i>
<i>gratte-ciel</i>	<i>grattacielo</i>	<i>skyscraper</i>	<i>Wolkenkratzer</i>
<i>récepteur</i>	<i>ricevitore</i>	<i>receiver</i>	<i>Empfänger</i>
<i>sous-développé</i>	<i>sottosviluppato</i>	<i>under-developed</i>	<i>unterentwickelt</i>
<i>sous-marin</i>	<i>sottomarino</i>	<i>undersea</i>	<i>unterseeisch</i>
<i>sous-sol</i>	<i>sottosuolo</i>	<i>underground</i>	<i>Untergrund</i>
<i>souterrain</i>	<i>sotterraneo</i>	<i>underground</i>	<i>unterirdisch</i>
<i>supermarché</i>	<i>supermercato</i>	<i>supermarket</i>	<i>Supermarkt</i>
<i>superpuissance</i>	<i>superpotenza</i>	<i>superpower</i>	<i>Supermacht</i>

Quando várias línguas vulgares traduzem do grego ou do latim um determinado vocábulo, é de qualquer modo irrelevante para a classificação lexicogênica desses empréstimos saber qual língua efetuou o empréstimo de tradução primeiro, e qual depois. Em outras palavras, sendo o vocábulo original pertencente ao léxico de línguas historicamente anteriores às línguas vulgares, e que exercem justamente a função de *línguas-parâmetro* em nosso sistema de classificação, todas as línguas vulgares estão, portanto, em posição idêntica em relação a essas línguas: são todas línguas importadoras em relação ao grego ou ao latim. Ao contrário,

quando uma língua vulgar traduz um vocábulo de outra língua vulgar, é evidente que a datação dos vocábulos correspondentes das diversas línguas se torna indispensável para sua classificação. Por exemplo, se tomarmos o ing. *computer* e o fr. *ordinateur*, só poderemos saber qual dos dois é o vocábulo original e qual a tradução, se soubermos a data da primeira ocorrência de cada um desses vocábulos em suas respectivas línguas. Obviamente, tendo o computador sido inventado nos Estados Unidos no início do século XX, a ocorrência mais antiga é a do ing. *computer*, sendo portanto o vocábulo francês uma tradução do inglês.

Vulgarismos resultantes de metamorfismo

Conforme vimos em Empréstimos de vocábulos greco-latinos, na p. 60, chama-se *metamorfismo* ao processo de mutação fonético-fonológica, acompanhado de adaptação gráfica correspondente, sofrido por um empréstimo de origem grega ou latina, ou por um grecolatinismo emprestado de uma língua vulgar estrangeira, no âmbito da língua vulgar receptora do empréstimo. Chama-se igualmente *metamorfismo* ao vocábulo resultante desse processo. Convém mais uma vez distinguir a mutação fonético-fonológica, seja ela espontânea (resultante de evolução fonética natural) ou deliberada, da simples adaptação fonético-fonológica, que, por sinal, todas as línguas realizam, ao importarem um vocábulo estrangeiro. O único caso em que o metamorfismo se origina de uma adaptação e não de uma mutação fonético-fonológica é o que descrevemos em Processos vulgarizantes, p. 64, alínea b.

O metamorfismo pode, como dissemos, ser espontâneo ou deliberado (analógico). Pode ser também obrigatório ou facultativo.

Metamorfismo espontâneo e metamorfismo deliberado (analógico)

a) *Metamorfismo espontâneo*: resulta da evolução fonética natural que um empréstimo greco-latino pode sofrer ao longo do tempo. Em alguns casos, coincide com aquilo que os gramáticos chamam de corrupção ou corruptela. Atinge em geral os emprés-

timos mais antigos (por exemplo, os vocábulos gregos ou latinos introduzidos numa língua vulgar nos primeiros séculos de sua história), e, por seu caráter natural, decorre do próprio uso da língua pelos falantes comuns, especialmente pelos menos letrados. Além disso, costuma ocorrer em etapas (ver Etapas do processo de metamorfismo espontâneo, p. 102). Metamorfismos desse tipo são:

lat. *capitulu* → port. *cabido*

lat. *titulu* → fr. *title* > *titre*

lat. *melancholia* → it. *malinconia*

b) *Metamorfismo deliberado (analógico)*: consiste na mutação fonético-fonológica e gráfica intencional de um empréstimo greco-latino, no momento de sua introdução, por analogia aos vocábulos vernáculos da língua. Trata-se de atribuir aos significantes dos novos vocábulos o mesmo padrão fonológico já vigente na língua. Assim, na medida em que uma determinada seqüência fonológica, característica do grego ou do latim, ainda que fonologicamente possível na língua vulgar (e nem sempre o é), presente na mesma uma freqüência de uso muito baixa, há grande probabilidade de que um vocábulo greco-latino que contenha tal seqüência a tenha substituída por outra, mais freqüente na língua, e que geralmente é a forma vernácula resultante daquela seqüência mediante os metaplasmos da evolução fonética regular. O metamorfismo deliberado atinge, em geral, os empréstimos mais recentes (especialmente a partir do final da Idade Média), principalmente os vocábulos cultos (termos técnicos, científicos, etc.), e é, por isso mesmo, realizado em grande parte por eruditos, conhecedores das línguas clássicas. Devido a seu caráter intencional, esse tipo de metamorfismo revela, por parte de seu criador, um sentimento nativista semelhante ao que norteia o empréstimo de tradução. São exemplos de metamorfismos deliberados:

lat. *doctore* → port. *doutor*

lat. *functione* → fr. *fonction*

lat. *sollicitu* → it. *solleccito*

lat. *directu* → it. *diretto*

Metamorfismo obrigatório e metamorfismo facultativo

Na medida em que a estrutura fonológica da língua vulgar não seja compatível com a das línguas clássicas, vê-se o falante introdutor do empréstimo greco-latino constrangido a introduzir o vocábulo em questão com sua estrutura fonológica original, o que, se aceito pelos demais falantes, conduzirá à modificação do próprio sistema fonológico da língua¹⁷, ou a modificar o vocábulo para adaptá-lo a esse padrão, procedimento este que, se realizado de maneira sistemática numa determinada língua, irá configurar o metamorfismo obrigatório. Exemplo disso é o *e* protético das línguas ibéricas diante de *s* seguido de consoante (lat. *statua* → port. *estátua*). Entretanto, quase todas as línguas européias ocidentais metamorfizam empréstimos greco-latinos, sem que tal procedimento seja imposto por coerções da estrutura fonológica dessas línguas. Temos assim o metamorfismo facultativo. Em certas línguas, como o italiano, por exemplo, um determinado tipo de metamorfismo, embora facultativo, é realizado de forma sistemática. Nesse caso, diremos tratar-se de um metamorfismo recomendado.

Etapas do processo de metamorfismo espontâneo

Conforme vimos em Empréstimos de vocábulos greco-latinos, na p. 60, o vocábulo metamórfico difere do original greco-latino tanto fonológica quanto graficamente. Portanto, somente a mutação fonológica ou somente a alteração de grafia, isoladamente, não configuram o metamorfismo. Enquanto o metamorfismo deliberado já é introduzido na língua como tal, o metamorfismo espontâneo em geral surge na língua inicialmente como grecolatinismo, o qual, com o passar do tempo e o efeito do uso lingüístico continuado, sofre mutações fonético-fonológicas, que a grafia, em geral, não costuma acompanhar de imediato, visto que a modalidade escrita da linguagem é sempre mais conservadora e, portanto, de evolução mais lenta. Observa-se, assim, uma etapa intermediária no processo de vulgarização de um vocábulo, em que o mesmo, embora ainda mantenha seu *status* greco-latino do ponto de

¹⁷ Tal fato ocorreu, por exemplo, em francês, que não possui palavras vernáculas iniciadas por *s* subverteu o sistema fonológico primitivo do francês, que, a partir de então, passou a admitir tal combinação de fonemas.

vista gráfico, já começa a dar mostra de que caminha no sentido de perdê-lo, a menos, é claro, que tal tendência seja refreada, quer pelo ensino normativo de uma pronúncia “cultu”, que suplante a inovação fonética, quer pelo não acompanhamento dessa inovação pela ortografia (como ocorre na língua inglesa).

Tomemos o exemplo da palavra alemã *Litanei* (← lat. *litania*):

ETAPA	GRAFIA	PRONÚNCIA	DATA	PROCESSO SOFRIDO
I	<i>litani</i>	[lita'ni:]	séc. XI	
II	<i>litani</i>	[lita'ne]	séc. XII	evolução fonética
III	<i>litanei</i>	[lita'ne]	séc. XIV	reforma ortográfica
IV	<i>Litanei</i>	[lita'na]	séc. XVI	evolução fonética

Introduzida na língua alemã como latinismo (etapa I), sofreu evolução fonética sem alteração gráfica (etapa II). Portanto, até o séc. XII, esse vocábulo manteve sua forma erudita. Quando, por volta do séc. XIII, começa-se a grafar o ditongo [eɪ] como *ei*, é que se começa a configurar efetivamente o metamorfismo do vocábulo. Assim, no séc. XIV, quando a grafia em *ei* já está totalmente consagrada, o vocábulo em questão já pode ser considerado um perfeito metamorfismo. Note-se, entretanto, que a passagem do estágio erudito ao estágio vulgar (metamórfico) se dá lentamente, não sendo possível estabelecer uma data precisa da ocorrência dessa passagem, até porque as grafias *litani* e *litanei* convivem durante algum tempo, numa demonstração de hesitação entre conservar a grafia latina, em detrimento da pronúncia, ou modificar a grafia, em prejuízo da latinidade do vocábulo. Note-se, ainda, que uma nova evolução fonética ocorre posteriormente (etapa IV), sem que a grafia atual do vocábulo dê conta da mesma.

Em certas línguas, como o francês e o inglês, por exemplo, a mudança ortográfica é extremamente lenta, razão pela qual muitas palavras que já sofreram evolução fonética ainda conservam grafia greco-latina. Um exemplo disso é o inglês *Christ*, cuja pronúncia já passou pelas etapas ['kri:st], ['kri:st] e ['kri:st], sem alteração da grafia. Tais línguas, por isso mesmo, acabam por assumir um aspecto mais latinizado.

Embora todo metamorfismo difira do vocábulo greco-latino que lhe dá origem tanto gráfica como foneticamente, nem todo vocábulo de uma língua vulgar que apresente divergência do

grego ou latim em ambos os aspectos é um metamorfismo. Com efeito, muitas vezes ocorre de um vocábulo greco-latino já entrar na língua vulgar com sua grafia adaptada ao sistema ortográfico dessa língua. Posteriormente, pode haver evolução fonética do vocábulo, o que o fará diferir do original greco-latino, sem que isso, no entanto, configure metamorfismo. Vejamos o exemplo abaixo:

lat. *natione* → esp. *nación* [na'tsjon] — adaptação ortográfica
nación [na'θjon] — evolução fonética
nación [na'sjon] — evolução fonética

Também não constitui metamorfismo a adaptação fonético-fonológica de um vocábulo greco-latino, com posterior reforma ortográfica. Assim, por exemplo o al. *Prozeß* (← lat. *processu*) surge primeiro como *process* [pʁo'tsɛs], portanto, com simples adaptação do *c* à pronúncia que essa letra tem em alemão; uma posterior reforma da ortografia alemã substituiu todo *c* com valor de [ts] por *z*. Em nenhum momento houve evolução fonética dessa palavra.

Igualmente, não constitui metamorfismo o vocábulo que já seja introduzido em língua vulgar com grafia e pronúncia divergentes do original greco-latino, mas que se enquadre no paradigma de vocábulos correspondentes às situações descritas acima. Assim, o esp. *abducción* [aβðuk'sjon] já entrou na língua nessa forma, porém se enquadra no paradigma de *nación*. O mesmo vale para o al. *Holzän* em relação ao paradigma de *Prozeß*. Já um vocábulo como o it. *giubilo* não se enquadra em nenhum desses dois paradigmas: com efeito, não há em italiano nenhuma palavra cuja grafia tenha passado de *j* a *gi*, bem como também não há nenhuma palavra cuja pronúncia inicial [j] tenha evoluído para [dʒ], pois todas as palavras italianas com essas características são vernáculos. Portanto, *giubilo* é uma forma latina vulgarizada por analogia com palavras vernáculos: é um legítimo metamorfismo.

Convém ainda lembrar que, quando uma língua vulgar toma emprestado um vocábulo ao latim, substituindo algum dos morfemas desse vocábulo por um morfema correspondente vulgar, o qual por sua vez procede daquele morfema latino por metamorfismo, não temos aí um empréstimo de tradução, mas, na verdade, um novo metamorfismo. Seja, por exemplo, o lat. *librarius*, que o

francês toma emprestado na forma *libraire*, isto é, metamorfozando o sufixo. Igualmente, existe em francês uma infinidade de outros vocábulos de origem latina com tal sufixo: *apothicaire*, *funéraire*, *littéraire*, *primaire*, etc. Se nos fosse possível dizer ao certo qual dentre todos os vocábulos franceses terminados por *-aire* surgiu primeiro, poderíamos afirmar que tal vocábulo sofreu metamorfismo do sufixo em relação ao latim, e os demais vocábulos do paradigma em questão constituiriam então traduções do latim mediante tal sufixo vulgarizado. Entretanto, não é possível determinar com exatidão qual desses vocábulos foi o primeiro a surgir, até porque muitos deles devem ter surgido mais ou menos na mesma época. Eis porque é mais sensato considerar todos esses vocábulos como metamorfismos em relação aos respectivos vocábulos latinos. Em igual situação se encontram os vocábulos alemães do paradigma do acima citado *Litanei*, como é o caso de *Polizei*, por exemplo.

Radicais que sofrem metafonía

Como vimos, a disparidade gráfica e fonética de um empréstimo greco-latino em relação ao vocábulo original é condição necessária porém não suficiente para caracterizar o metamorfismo. Assim, muitos empréstimos greco-latinos apresentam irregularidades que os distinguem da forma greco-latina original. Entretanto, se tais irregularidades se devem ao enquadramento do vocábulo num paradigma de flexão gramatical próprio da língua, mais uma vez não se tem metamorfismo. Por exemplo, em português, a toda palavra terminada em *-al* corresponde um plural em *-ais*. Conseqüentemente, *normal* forma o plural *normais*, com perda do *l* do radical latino, o que não constitui vulgarização, mas sim enquadramento da palavra no mesmo paradigma de vernáculos como *mortal/mortais*, por exemplo. Da mesma forma temos o fr. *normal/normaux*, port. *progredir/progride*, etc.

Metamorfismo por hiperurbanismo

O hiperurbanismo é um fenômeno que consiste da pronúncia incorreta (do ponto de vista gramatical) de uma palavra, por ex-

cesso de zelo do falante, o qual, por incultura, toma por errada uma forma que na verdade é correta. Assim, é comum ouvir-se em português *rúbrica* por *rubrica*, *púdico* por *pudico*, *siclano* por *sicrano*, ou ainda *telha de aranha* por *teia de aranha*, *demás* por *demais*, etc. Pois bem: não só os falantes incultos, mas por vezes também os eruditos cometem deslizes desse tipo. Quando tais deslizes ocorrem no exato momento em que um vocábulo de origem grega ou latina está sendo introduzido numa determinada língua, o que temos é um desvio da forma original que, se produzir a dupla divergência (gráfica e fonético-fonológica) a que anteriormente nos referimos, configurará o metamorfismo. Desse modo, observamos exemplos como o it. *corinzio*, que é grafado com *-zi-* por analogia a inúmeras palavras italianas que ao latim *-ti-* respondem com *-zi-* (por exemplo, lat. *natione* → it. *nazione*). Entretanto, a forma original latina é *corinthius*, com *-thi-* portanto, a qual não poderia ser transliterada para *-zi-*, mas somente para *-ti-* (cfr. port. *coríntio*, por exemplo). Um outro exemplo, também extraído do italiano, é o verbo *analizzare*, empréstimo do fr. *analyser*, por sua vez formado por derivação a partir do substantivo *analyse* (← gr. *analysis*). O italiano identificou a terminação *-yser* daquele vocábulo ao sufixo francês homófono *-iser* (← lat. *-izare* ← gr. *-izein*), o qual possui em italiano uma forma semi-erudita *-izzare*.

Principais metamorfismos

Vejamos, a seguir, os principais processos metamórficos que as línguas européias ocidentais costumam apresentar.

METAMORFISMO

EXEMPLO

PORTUGUÊS

<i>-e- → -i-</i>	lat. <i>professio, processio</i> → port. <i>profissão, procissão</i>
<i>-ct-, -pt- → -it-</i>	lat. <i>respectus, conceptus</i> → port. <i>respeito, conceito</i>
<i>-ct-, -pt- → -t-</i>	lat. <i>productus, optimus</i> → port.

	<i>produto, ótimo</i>
<i>-ct-</i> → <i>-ut-</i>	lat. <i>doctor</i> → port. <i>doutor</i>
<i>-kd-</i> , <i>-gd-</i> → <i>-d-</i>	gr. <i>anekdotos</i> , <i>amygdale</i> → port. <i>anedota</i> , <i>amígdala</i>
<i>-cti-</i> , <i>-pti-</i> → <i>-ç-</i>	lat. <i>productio</i> , <i>adoptio</i> → port. <i>produção</i> , <i>adoção</i>
<i>-nct-</i> , <i>-mpt-</i> → <i>-nt-</i>	lat. <i>junctus</i> , <i>assumptus</i> → port. <i>junto</i> , <i>assunto</i>
<i>-ncti-</i> , <i>-mpti-</i> → <i>-nç-</i>	lat. <i>functio</i> , <i>redemptio</i> → port. <i>função</i> , <i>redenção</i>
<i>hy-</i> → <i>j-</i>	lat. <i>hyacinthus</i> → port. <i>jacinto</i>
<i>pl-</i> , <i>bl-</i> → <i>pr-</i> , <i>br-</i>	lat. <i>plumbum</i> , <i>obligare</i> → port. <i>prumo</i> , <i>obrigar</i>
<i>-mn-</i> → <i>-n-</i>	lat. <i>omnipotens</i> → port. <i>onipotente</i>
<i>e</i> protético diante de <i>s</i> seguido de consoante	lat. <i>stimulus</i> → port. <i>estímulo</i>
sonorização de surdas intervocálicas	lat. <i>vicarius</i> , <i>capitulum</i> → port. <i>vigário</i> , <i>cabido</i>
síncope de <i>d</i> intervocálico	lat. <i>Paradisus</i> , <i>concludere</i> → port. <i>Paraíso</i> , <i>concluir</i>
síncope de <i>l</i> intervocálico	lat. <i>articulus</i> , <i>tabula</i> , <i>populus</i> → port. <i>artigo</i> , <i>tábua</i> , <i>povo</i>

ESPAÑHOL

<i>-x-</i> → <i>-j-</i>	lat. <i>exemplum</i> → esp. <i>ejemplo</i>
<i>-ct-</i> → <i>-t-</i>	lat. <i>fructus</i> , <i>respectus</i> → esp. <i>fruto</i> , <i>respeto</i>
<i>-nct-</i> , <i>-mpt-</i> → <i>-nt-</i>	lat. <i>junctus</i> , <i>assumptus</i> → esp. <i>junto</i> , <i>asunto</i>
<i>-ncti-</i> , <i>-mpti-</i> → <i>-nci-</i>	lat. <i>functio</i> , <i>redemptio</i> → esp. <i>función</i> , <i>redención</i>
<i>hi-</i> , <i>hy-</i> → <i>j-</i>	gr. <i>hierarchia</i> , lat. <i>hyacinthus</i> → esp. <i>jerarquía</i> , <i>jacinto</i>

-m- seguido de
consoante → -n-

lat. *circumferentia* → esp.
circunferencia

e protético diante de s
seguido de consoante

lat. *statua, scena* → esp. *estatua,*
escena

FRANCÊS

-un- → -on-

lat. *functione* → fr. *fonction*

síncope do s pré-
consonantal

lat. *hospitale* → fr. *hôpital*

síncope da vogal pós-tônica

lat. *articulus* → fr. *article*

síncope da vogal pós-tônica
e rotacismo

lat. *titulus, capitulum, apostolus* →
fr. *titre, chapitre, apôtre*

ITALIANO

-u- pós-tônico → -o- e
respectivos derivados

lat. *regula, regularis* → it. *regola,*
regolare

-i- → -e-

lat. *sollicitus* → it. *sollecito*

-i- pós-tônico → -a- e
respectivos derivados

lat. *chronica, syndicus* → it.
cronaca, sindaco

-ct-, -pt-, -bt- → -tt-

lat. *directus, aptus, obturare* → it.
diretto, atto, otturare

-chth-, -phth- → -tt-

gr. *ichthyologia, diphthongos* → it.
ittiologia, dittongo

-kd- → -dd-

gr. *anekdotos* → it. *aneddoto*

-cti-, -pti- → -zi-

lat. *productio, adoptio* → it.
produzione, adozione

-nct-, -mpt- → -nt-

lat. *junctus, assumptus* → it.
giunto, assunto

-ncti-, -mpti- → -nzi-

lat. *functio, exemptio* → it.
funzione, esenzione

-ps- → -ss-

lat. *ellipsis* → it. *ellisse*

-x- → -ss-, -s-

lat. *complexus, auxilium, externus*
→ it. *complesso, ausilio, esterno*

-bs-, -ds- → -ss-, -s-	lat. <i>abstractus, adstringere, obsoletus</i> → it. <i>astratto, astringere, ossoleto</i>
-xc- → -cc-	lat. <i>excellens</i> → it. <i>eccellente</i>
-mn- → -nn-	lat. <i>hymnus</i> → it. <i>inno</i>
-dm- → -mm-	lat. <i>admirare</i> → it. <i>ammirare</i>
-dv- → -vv-	lat. <i>adverbium</i> → it. <i>avverbio</i>
-dj- → -gg(i)-	lat. <i>adjunctus, adjectivum</i> → it. <i>aggiunto, aggettivo</i>
j- → g(i)-	lat. <i>jubilum</i> → it. <i>giubilo</i>
hi- → g-	lat. <i>hieroglyphicus</i> → it. <i>geroglífico</i>
hy- → gi-	lat. <i>hyacinthus</i> → <i>giacinto</i>
-ms-, -ns- → -s-	lat. <i>circumstantia, instans</i> → it. <i>circostanza, istante</i>
-pl- → -pi-	lat. <i>exemplum, templum</i> → it. <i>esempio, tempio</i>
-m- seguido de consoante → -n-	lat. <i>triumphus</i> → it. <i>trionfo</i>
-chi- → -ci-	gr. <i>schistos</i> → it. <i>scisto</i>
-bl-, -bj- → -bbl-, -bbi-	lat. <i>publicus, objectus</i> → it. <i>pubblico, obbietto</i>
geminção consonantal	lat. <i>tolerare</i> → it. <i>tollerare</i>

INGLÊS

-pium → -ple	lat. <i>principium, participium</i> → ing. <i>principle, participle</i>
	<i>participle</i>

ALEMÃO

metafonia da vogal pós-tônica	lat. <i>particula, metaphora</i> → al. <i>Partikel, Metapher</i>
-------------------------------	--

SUECO

metafonia/síncope da vogal pós-tônica lat. *titulus* → sue. *titel* (plural *titlar*)

Nota-se, pelo quadro acima, que o metamorfismo é um fenômeno muito mais comum nas línguas românicas do que nas germânicas. Isso se deve ao fato de, naquelas línguas, o próprio vernáculo, por analogia com o qual os empréstimos greco-latinos são vulgarizados, proceder do latim, havendo, portanto, um sentimento de parentesco muito nítido entre os vocábulos vernáculos e os empréstimos. Já as línguas germânicas emprestam elementos do grego ou do latim sempre em sua forma original, salvo poucas alterações na vogal pós-tônica, justamente por estarem geneticamente mais distantes do latim. Assim, pode-se dizer que o metamorfismo de empréstimos greco-latinos é tanto mais freqüente quanto mais próxima do latim está uma língua vulgar. Por essa razão, é o italiano, dentre as línguas românicas, a que mais metamorfiza, visto ser a língua que menos diferenças fonético-fonológicas apresenta em relação ao latim. Inversamente, o francês é a língua românica foneticamente mais distante do latim; por conseguinte, é também a que menos metamorfiza.

Vulgarismos resultantes de combinação sintagmática (composição ou derivação)

Compostos de radicais vulgares ou derivados de radical vulgar

Um dos mecanismos mais freqüentes de criação vocabular é a chamada neologia sintagmática, a qual compreende basicamente os processos de composição e derivação, esta última também chamada de afixação. É oportuno destacar aqui que, malgrado alguns autores prefiram classificar a prefixação como uma forma de composição, por entenderem que certos prefixos (por exemplo, *contra*, *extra*, *sobre*, etc.) são também vocábulos autônomos, optamos por considerar tanto a prefixação quanto a sufixação como formas de derivação, reservando o termo *composição* ao processo de união sintagmática de dois vocábulos lexicais, isto é,

portadores de morfema lexical (lexema, segundo a terminologia de Pottier). Com efeito, nem todos os prefixos são vocábulos autônomos, e, além disso, são todos vocábulos gramaticais, de modo que não há grande diferença funcional entre a prefixação e a sufixação enquanto processos lexicogênicos.

Já vimos que os empréstimos de tradução sintagmáticos são essencialmente compostos e derivados vulgares criados para traduzir um vocábulo estrangeiro. Entretanto, a maior parte dos compostos e derivados de uma língua são de tipo autogênico¹⁸, isto é, criações intrínsecas da língua, sem nenhuma motivação externa, quer de significante, quer de significado. Essas criações podem resultar da combinação de elementos exclusivamente greco-latinos, da combinação de elementos exclusivamente vulgares, ou ainda podem resultar da combinação de ambas as espécies de elementos. No primeiro caso, que estudaremos mais adiante (Grecolatinismos resultantes de combinação sintagmática, p. 125), o vocábulo resultante será um grecolatinismo. Nos demais casos, o vocábulo resultante será um vulgarismo. No presente Item, nos ocuparemos dos compostos e derivados com radicais vulgares, tratando dos compostos e derivados vulgares com radical greco-latino no Item seguinte. Essa divisão se faz necessária, visto que, no primeiro caso, o vocábulo (primitivo) que dá origem ao composto ou derivado é vulgar, e, no segundo caso, greco-latino. Nosso método de classificação está, portanto, centrado primordialmente no estatuto lexicogênico do elemento lexical da palavra, e, apenas secundariamente, no elemento gramatical. A razão disso é que, dado um radical vulgar, o vocábulo derivado que o contém será vulgar, qualquer que seja a natureza dos afixos, quer sejam estes eruditos ou vulgares. Já, dado um radical greco-latino, o vocábulo derivado resultante só será greco-latino se também os afixos o forem; basta que um dos afixos seja vulgar para que todo o vocábulo também o seja.

É interessante notar que a derivação — sobretudo a sufixação — é o principal processo neológico sintagmático das línguas românicas, ao passo que nas línguas germânicas predomina a composição. De fato, nessas línguas, existem morfemas e palavras lexicais que acabaram por assumir uma função meramente grama-

¹⁸ Cfr. Mecanismos de empréstimo, p. 133.

tical nas composições. Assim, comportam-se como se fossem legítimos sufixos. Por exemplo, um elemento como *-man* no ing. *milkman* desempenha o mesmo papel que o sufixo *-eiro* no port. *leiteiro*. Tal elemento apenas não se tornou um puro sufixo porque existe também como palavra autônoma (por exemplo, ing. *He is a good man*, “Ele é um bom homem.”). Por vezes, um antigo elemento independente pode transformar-se em sufixo quando não seja mais reconhecido como elemento lexical. Isto se dá em geral quando a palavra autônoma correspondente cai em desuso, permanecendo apenas nos compostos. É o que ocorre com o sufixo alemão *-lich*, de *herzlich*, por exemplo, que já foi um adjetivo autônomo, significando “igual, semelhante” (cfr. ing. *like*), e que se perdeu completamente nessa função.

Por ser a composição o mecanismo predominante de formação sintagmática de palavras nas línguas germânicas, apresentam estas uma gama de possibilidades combinatórias bastante superior à das línguas românicas. Com efeito, na derivação, dispõe-se de um número ilimitado de radicais, mas um número limitado, e relativamente baixo, de afixos. Já na composição, qualquer palavra lexical pode ser combinada a qualquer outra. Decorre daí que a muitas palavras compostas germânicas correspondem nas línguas românicas palavras complexas, isto é, formadas de mais de uma palavra simples (compare-se, por exemplo, o al. *Kindergarten* com o port. *jardim de infância*).

São compostos vulgares em português vocábulos como *guarda-chuva*, *cata-vento*, *parapeito*, *corrimão*, *barbirruivo*, *cabisbaixo*, etc., todos contendo apenas radicais vulgares. Derivados de radicais vulgares em português temos: *sapateiro*, *redondeza*, *formigamento*, *empobrecer*, *imperdível*, etc.

Compostos de radicais greco-latino e vulgar ou derivados de radical greco-latino com afixo vulgar

Este tipo específico de combinação sintagmática entre elementos greco-latinos e vulgares, de que resulta um vocábulo lexicogenicamente híbrido, é particularmente interessante por mostrar como, numa língua vulgar, um grecolatinismo pode dar origem a vulgarismos, o que revela, em última instância, a perfeita integração desse grecolatinismo ao léxico da língua. Um exemplo

desse tipo é a palavra *meter* em inglês. Além de existir de forma independente, com o significado de “medidor”, também se depreende de compostos eruditos como *thermometer*, *dynamometer*, *photometer*, dentre outros. A existência desses compostos eruditos em inglês autorizou a criação de termos como *speedometer*, por exemplo. Mais freqüentes ainda são as derivações (por exemplo, *metering*), especialmente deverbais de verbos latinos e denominais de adjetivos latinos, como o al. *Planung*, a partir de *planen*, ou o ing. *politeness*, a partir de *polite*.

São compostos híbridos, isto é, formados de um radical grego ou latino e outro vulgar: port. *auriverde*, *rubro-negro*, *semi-aberto*; fr. *bureaucratie*, *demi-ton*; ing. *biochemistry*, *self-confidence*, etc. Derivados de radical greco-latino com afixo vulgar são: port. *desequilíbrio*, *recusador*, *sobre-humano*; fr. *actanciel*, *réactionnaire*, *révolutionnaire*; ing. *computer*, *typing*, etc. Entram aqui também as derivações regressivas a partir de verbos greco-latinos, pois esse processo de derivação em latim jamais se aplicava aos verbos.¹⁹

Compostos e derivados “anômalos” a partir de elementos greco- latinos

Muitas vezes, podemos ter um vulgarismo em que todos os morfemas formantes são greco-latinos. Isso ocorre quando da combinação de elementos greco-latinos resulta um vocábulo cuja estrutura fonológica é incompatível com a das línguas clássicas. Neste caso, o caráter vulgar do vocábulo não reside em seus elementos mórficos, mas sim no processo morfossintático segundo o qual estes são combinados.

Tomemos como exemplo os vocábulos espanhóis *inmenso* e *disminuir*. Embora os morfemas *in-*, *mens-*, *dis-* e *minu-* sejam todos latinos, os encontros consonantais *nm* e *sm* não são admissíveis na prefixação em latim. Com efeito, as formas latinas correspondentes são *immensus* e *diminuere*. Desse modo, não se pode dizer que o esp. *inmenso* provenha diretamente do lat. *immensus*,

¹⁹ Cfr. a esse respeito MAURER JR., T. H. *A unidade da România ocidental*. São Paulo, FFCL-USP, 1951, p. 88.

mas, antes, que foi construído dentro da própria língua espanhola a partir dos elementos latinos *in-* e *mensus*, segundo um processo de combinação sintagmática totalmente estranho à língua latina, o mesmo valendo para *disminuir*. Na verdade, tais vocábulos podem ser explicados como resultado de uma corrupção popular dos vocábulos latinos de que provêm por analogia com outros, tais como *distracción* e *infeliz*, por exemplo.

Tomemos agora o composto inglês *dosemeter*. Formado a partir dos grecismos *dose* e *meter* por simples justaposição, esse termo não obedece à regra de composição greco-latina, que somente admitiria as formas *dosometer* e *dosimeter* (esta última forma, por sinal, coexiste em inglês com *dosemeter*). Nota-se, assim, que, não obstante os formantes dessa palavra serem gregos, sua combinação se deu conforme um processo próprio às línguas germânicas, a justaposição de morfemas, sem a utilização de vogais de ligação.

Ainda em inglês, temos, ao lado de *normality*, o vocábulo *normalcy*. Este último vocábulo foi engendrado a partir do adjetivo erudito *normal* com o sufixo igualmente erudito *-cy*, tal qual abstraído de formas como *efficacy* (← lat. *efficacia*), ou *accuracy* (← lat. hipotético **accuratia*, de *accuratus*), ou ainda *privacy* (← lat. hipotético **privatia*, de *privatus*), por exemplo. Entretanto, o vocábulo inglês deveria ter como base uma forma latina hipotética **normalitia* (segundo o modelo de *justitia*, *malitia*, etc.), a qual produziria em inglês **normalicy* e não *normalcy*. Esta última forma pressupõe um lat. **normaltia*, que não tem qualquer justificativa morfológica. Desse modo, a combinação entre *normal* e *-cy* se dá de forma totalmente irregular, incompatível com o padrão morfológico dos vocábulos eruditos.

Muitos termos técnicos formados a partir de radicais greco-latinos apresentam combinação irregular desses radicais. É o caso do port. *leucemia*, formado a partir dos radicais gregos *λευκός* e *αἷμα*, mais o sufixo *-ία*. Ocorre que a combinação desses morfemas em grego teria dado **λευχαιμία*, que transliterado para a grafia latina produziria **leuchaemia*. Ora, o *ch* latino, assim como o *P* grego, resultam em português em *c* diante de *a*, *o*, *u* e consoantes, e *qu* diante de *e* e *i*. Assim, a forma portuguesa mais adequada, embora a menos usada, é *leuquemia* e não *leucemia*.

Muitas vezes, derivados greco-latinos anômalos se originam por analogia a derivados greco-latinos perfeitos. É o caso do ing.

investor, derivado do verbo *to invest* (← lat. *investire*), por analogia com o par *to direct* x *director* (← lat. *director*), por exemplo. Entretanto, a forma mais condizente com a estrutura morfológica do latim seria **investitor*.

O efeito da analogia com vocábulos de natureza etimológica diferente pode conduzir muitas vezes a um erro na análise dos elementos mórficos do vocábulo em questão. Assim, em português, o vocábulo *médium* (← lat. *medium*) produz o derivado *mediúnico*, com *n* ao invés de *m*, provavelmente pelo fato de que, em nossa língua, a vocábulos terminados em *m* correspondem derivados em *n*: *fim* x *final*, *som* x *sonoro*, *jardim* x *jardineiro*, etc. Ocorre que, nesses casos, a etimologia dos vocábulos em questão (lat. *finis*, *sonus*; fr. *jardin*) justifica o uso de *n* na derivação, o que não acontece com *médium*. Note-se que, além de etimologicamente equivocado, o vocábulo *mediúnico* ainda inclui em seu radical a desinência latina *-um*, como se a mesma pertencesse ao radical da palavra.

Co-ocorrência de processos lexicogênicos no mesmo vocábulo

Um mesmo vocábulo vulgar pode resultar de mais de um processo de vulgarização. Assim, o port. *redução*, por exemplo, provém do lat. *reductione* por tradução do sufixo latino *-tione* pelo vernáculo *-ção*, bem como por síncope do *c* latino do grupo *-ct-*, o que configura um metamorfismo. Da mesma forma, o vocábulo correspondente it. *riduzione* vem do lat. *reductione* por tradução do prefixo latino *re-* pelo vernáculo *ri-* e, além disso, por metamorfismo do grupo consonantal *-ct-* para *-z-* (síncope do *c*). A incidência de dois processos de vulgarização no mesmo vocábulo pode ser simultânea, no instante mesmo da criação deste, ou sucessiva. Neste caso, um vocábulo greco-latino pode sofrer empréstimo de tradução parcial, podendo, mais tarde, aqueles morfemas que não foram traduzidos ser alvo de metamorfismo. Esses dois fenômenos podem também ocorrer na ordem inversa.

Por outro lado, pode ocorrer simultaneamente num dado vocábulo a vulgarização de um morfema erudito e a substituição de um morfema vulgar por seu correspondente erudito. Seja, por exemplo, o ing. *extravaganza*, empréstimo do it. *stravaganza*. O inglês procedeu à substituição do prefixo vernáculo italiano *stra-*

pelo correspondente latino *extra-* por um processo que denominamos *restituição*, conforme veremos mais adiante (Empréstimos de restituição, p. 119). Entretanto, o elemento italiano *-vaganza*, formado a partir do verbo erudito *vagare* mediante o sufixo igualmente erudito *-anza*, constitui por sua vez metamorfismo em inglês. Ocorre que, em inglês, *-vaganza* difere do lat. hipotético **-vaganzia* tanto pela grafia quanto pela pronúncia, o que configura o metamorfismo.

Igualmente, seja o it. *rivoluzionario*, empréstimo do fr. *révolutionnaire*. Por um lado, o italiano substituiu o prefixo erudito francês *ré-* pelo vernáculo *ri-*; por outro lado, substituiu o sufixo vulgar francês *-aire* pelo erudito *-ario*.

Vocábulos cujas formas flexionais pertencem a categorias lexicogênicas diferentes

Alguns adjetivos e verbos apresentam formas flexionais pertencentes a categorias lexicogênicas diferentes, podendo mesmo algumas ser eruditas enquanto outras são vulgares. Seja o adjetivo francês *protecteur*, empréstimo de tradução do lat. *protector*: tal adjetivo forma um feminino *protectrice*, que está perfeitamente conforme ao latim. Por outro lado, o verbo italiano *distinguere*, perfeito latinismo, possui um particípio passado *distinto*, metamorfismo do lat. *distinctus*.

GRECOLATINISMOS

Grecolatinismos resultantes de empréstimo direto ou indireto do grego/latim

Constitui grecolatinismo todo vocábulo formado exclusivamente de morfemas de origem grega ou latina, não metamorfizados, excetuadas as desinências. Um grecolatinismo pode resultar do empréstimo de uma palavra efetivamente pertencente ao léxico do grego ou do latim, mas pode também ser criada, no âmbito da língua vulgar, através da combinação sintagmática de morfemas greco-latinos. Esta última possibilidade será discutida mais adian-

te, em Para que um vocábulo se caracterize como grecolatinismo, basta que seu significante seja greco-latino. Assim, o grecolatinismo pode resultar de um empréstimo total ao grego ou ao latim, ou de um empréstimo de significante em relação a essas línguas. Um grecolatinismo, uma vez introduzido na língua vulgar, com um determinado significado, pode, assim como todas as outras palavras da língua, adquirir novos significados, mediante o processo da neologia semântica. Desse modo, inicialmente, um grecolatinismo como o port. *núcleo* não significa, como em latim, “caroço, semente”, mas tal signo foi tomado apenas em seu significado abstrato, o de “essência, âmago, cerne”. Posteriormente, tal vocábulo adquire novos significados, alguns especializados (em biologia, em química, em física, em lingüística, etc.), inclusive o significado original de “semente”.

É muito comum que, quando uma língua vulgar necessite de uma designação para um novo conceito, busque essa designação nas línguas clássicas, ou porque tal conceito já exista nas mesmas, ou porque se trate de conceito referente a alguma norma discursiva culta ou especializada. Assim, é também comum que, quando uma primeira língua vulgar opte pelo empréstimo ao grego ou ao latim, essa tendência seja seguida pelas demais línguas vulgares, as quais, importando daquela primeira língua o conceito, importam também a expressão desse conceito. Portanto, quando uma língua vulgar importa um vocábulo diretamente do latim, dá-se um *empréstimo greco-latino direto*, ao passo que, quando uma língua vulgar importa um vocábulo greco-latino de outra língua vulgar, tem-se um *empréstimo greco-latino indireto*. Um exemplo disso é o fr. *direction*, importado diretamente do latim *directione*; já o ing. *direction* é um empréstimo do grecolatinismo francês.

Refecções de vulgarismos

Muitas vezes, um vocábulo vulgar ou vulgarizado, de origem latina, pode ser substituído por seu equivalente erudito, num processo que se chama *refecção*. Na verdade, a substituição do significante vulgar pelo erudito não se dá instantaneamente, mas durante algum tempo a forma vulgar e a erudita convivem na língua como alótopos. O grecolatinismo, surgindo inicialmente como neologismo, proposto como sinônimo do vulgarismo, uma

vez aceito pela comunidade lingüística, tende a ter uma freqüência de uso cada vez maior, ao passo que o vulgarismo começa então a diminuir de freqüência, até tornar-se um arcaísmo e, eventualmente, desaparecer. Algo do gênero parece estar ocorrendo agora mesmo em português, em relação à palavra *mosteiro*. Este vernáculo vem sofrendo a concorrência cada vez mais intensa do greco-latinismo *monastério*, provavelmente por influência do ing. *monastery*, visto que a tradução do termo inglês por *monastério* ao invés de *mosteiro* tem sido freqüente em especial nas legendas e dublagens de filmes de cinema de língua inglesa.

Entretanto, o apogeu da tendência à refeção sobre o modelo greco-latino parece ter ocorrido durante a Renascença, quando uma verdadeira torrente de grecismos e latinismos invadiu as línguas européias ocidentais. Facilitada pela moda literária e, principalmente, pela ideologia da época, a adoção desses grecolatinismos banuiu progressivamente os vulgarismos correspondentes, embora em alguns casos tenha havido alguns refluxos posteriores²⁰, ou mesmo a distinção de significado entre as formas erudita e vulgar, o que conduziu à permanência de ambas na língua.

Por vezes, a adoção de uma grafia latinizada pode induzir a uma falsa refeção. Isso ocorre quando a grafia de uma palavra reassume a forma latina, porém no plano fonético-fonológico nada ocorre que possa indicar tratar-se de um novo vocábulo. Essas falsas refeções foram particularmente freqüentes em francês, em que palavras como as vernáculos *ni* e *salu* tiveram, em fins da Idade Média, sua grafia alterada para *nid* e *salut*, por inspiração do latim *nidu* e *salute*, respectivamente. Entretanto, não se trata aí de empréstimos ao latim, senão de simples reforma ortográfica. Note-se que, nesse caso, não houve concorrência na língua entre as formas antigas e as novas, e os falantes iletrados, que obviamente jamais tomaram conhecimento dessa mudança de grafia, em nenhum momento surpreenderam o aparecimento de duas novas unidades no léxico da língua francesa. Por conseguinte, os vocábulos *nid* e *salut* permanecem vernáculos, em que pese sua grafia latinizada.

²⁰ Vejam-se, como exemplo, as formas portuguesas terminadas em *-ável*, *-ível*, que sofreram, no Renascimento, a concorrência de formas em *-ábil*, *-ível*, as quais, não tendo tido grande aceitação fora da norma literária, acabaram por cair em desuso, em benefício das formas anteriores, vulgares.

Alguns exemplos de vocábulos refeitos sobre o latim são: port. *chor* > *frol* > *flor*, *coa* > *cauda*, *diago* > *diácono*, *dino* > *digno*, *eivigar* > *edificar*, *esmar* > *estimar*, *fremoso* > *formoso*, *inhorar* > *ignorar*, *seenço* > *silêncio*, *sordo* > *surdo*, *zeo* > *zelo*, etc.; fr. *aerdre* > *adhérer*, *esmer* > *estimer*, *rançon* > *rédemption*, *soutil* > *subtil*, *treü* > *tribut*, etc.; it. *cerchio* > *circolo*, *rovina* > *ruina*, *romore* > *rumore*, etc.

Empréstimos de restituição

A partir das definições de *empréstimo total*, *empréstimo de significante* e *empréstimo de significado* que demos em Tipos de empréstimos, p. 58-59, definiremos o empréstimo de restituição como um empréstimo de significado de um vulgarismo estrangeiro, em que o significante é greco-latino, isto é, totalmente constituído de morfemas de origem grega ou latina não metamorfizados. Pode-se dizer, assim, que o empréstimo de restituição é o processo inverso do empréstimo de tradução, pois, neste, parte-se de um vocábulo greco-latino e chega-se a um vocábulo vulgar, ao passo que, na restituição, parte-se de um vulgarismo e chega-se a um grecolatinismo. O empréstimo de restituição, a exemplo do empréstimo de tradução, pode ser de dois tipos, a saber:

a) *Empréstimo de restituição sintagmático*: ocorre quando uma língua vulgar realiza neologismo sintagmático (composto ou derivado) com morfemas greco-latinos e significado emprestado de vocábulo vulgar estrangeiro. Exemplo:

ing. *feed back* → fr. *réalimentation*

b) *Empréstimo de restituição semântico*: ocorre quando uma língua vulgar realiza empréstimo de significante de vocábulo grego ou latino, com empréstimo de significado de vocábulo vulgar estrangeiro, o qual é neologismo semântico (= empréstimo de significante intralingüístico) na língua de origem. Exemplo:

fr. *entier*, “número inteiro (conceito matemático)” → ing. *integer*

Da mesma forma que o empréstimo de tradução, também o empréstimo de restituição pode ser facultativo ou obrigatório. Será obrigatório especialmente quando a um morfema vulgar estrangeiro (por exemplo, um sufixo) corresponder na língua importadora

exclusivamente um morfema greco-latino (por exemplo, ao sufixo vulgar francês *-aire* correspondem em português os sufixos *-ar* e *-ário*, ambos eruditos).

É interessante notar que, nas línguas românicas, o processo de empréstimo por restituição é particularmente freqüente na norma técnico-científica. Como se sabe, as línguas clássicas, especialmente o grego, apresentam uma grande facilidade para a formação de compostos, o que também ocorre em relação às línguas germânicas. Já as línguas românicas privilegiam a criação sintagmática por derivação, até devido à própria estrutura morfológica dessas línguas.²¹ Desse modo, as línguas germânicas ganham em coesão e economia, pois a uma locução nominal românica opõem uma palavra composta ou mesmo uma locução nominal de menor extensão (compare-se o port. *atraso devido à chuva* com o ing. *rain delay*, por exemplo). Assim sendo, quando uma língua germânica produz um neologismo técnico-científico (neônimo) por composição, este resulta ser, devido à estrutura peculiar da língua, um termo compacto e sintético. Por força das próprias exigências de normalização terminológica, as línguas românicas devem traduzir tal termo por outro, igualmente sintético. É justamente nas línguas clássicas que elas encontram o material léxico e os recursos morfosintáticos necessários à formação desse termo sintético. Assim, a um termo técnico inglês como *English-speaker*, o português contrapõe *anglofone*, mais sintético do que *falante de língua inglesa*. Vê-se aí por que grande parte dos termos técnico-científicos românicos são empréstimos de restituição de termos técnico-científicos ingleses ou alemães.

Vejamos, a seguir, alguns exemplos de empréstimos de restituição, nas diversas línguas européias.

ing. *bedding* (termo de geologia) → fr. *stratification*

ing. *football* → port. *balípedo* (desusado)

fr. *noyau* (termo de lingüística) → port. *núcleo*

fr. *actantiel* → esp. *actancial*

fr. *fonctif* → ing. *functive*

al. *Weltanschauung* → it. *cosmovisione*

²¹ Cfr. REY, A. *La terminologie: noms et notions*. Paris, P.U.F., 1979, p. 71.

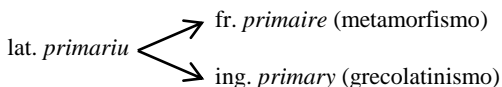
al. *Überich* → port. *superego*

it. *umanesimo* → port. *humanismo*

Relações entre metamorfismo e empréstimo de restituição

Se, dadas duas línguas vulgares, para uma determinada acepção, uma delas apresenta um grecolatinismo e a outra um vulgarismo correspondente, podemos ter duas situações, que descrevemos abaixo.

1) Se o significante existe em grego ou latim, teremos, por exemplo:



independentemente de qual vocábulo seja o mais antigo, o francês ou o inglês, pois ambos são empréstimos do latim.

2) Se o significante não existe em grego ou latim, podemos ter, por exemplo:

2a) fr. *révolution* + *-aire* = *révolutionnaire* (derivado vulgar)
→ ing. *revolutionary* (empréstimo de restituição)

(O vocábulo francês é o mais antigo)

2b) ing. *evolution* + *-ary* = *evolutionary* (derivado grecolatino) → fr. *évolutionnaire* (metamorfismo)

(O vocábulo inglês é o mais antigo)

Portanto, neste caso, a classificação lexicogênica dos vocábulos francês e inglês depende de datação, para se saber qual é o original e qual o empréstimo.

Relações entre empréstimo de tradução e empréstimo de restituição

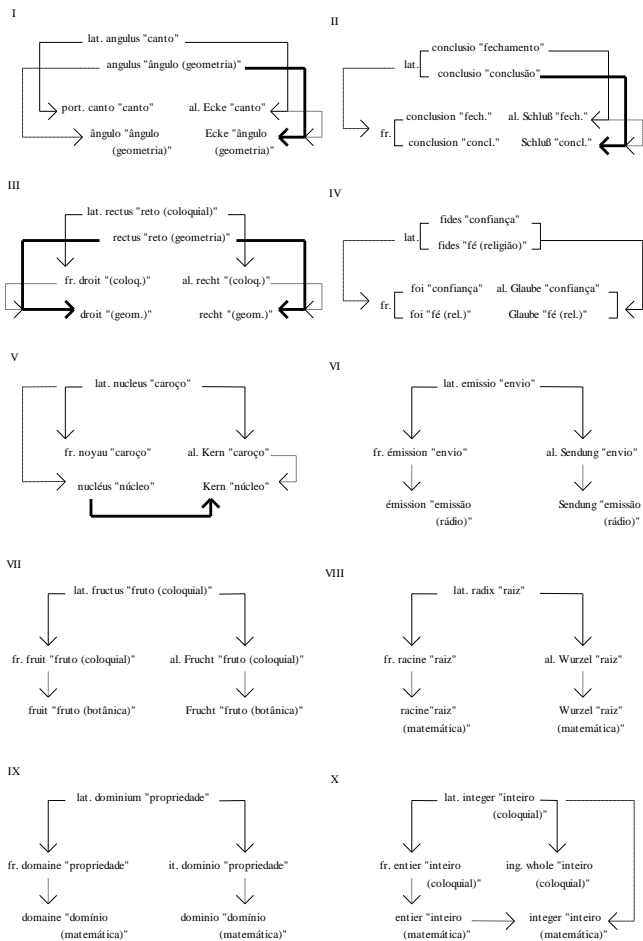
Da mesma forma que no item anterior, também aqui podemos ter duas situações. Sejam os vocábulos fr. *formel* e ing. *for-*

mal. Como ambos provêm do lat. *formalis*, temos que o inglês é empréstimo direto do latim, ao passo que o francês é empréstimo de tradução, independentemente de qual dos dois vocábulos seja o mais antigo. Se tomarmos agora os vocábulos fr. *informationnel* e ing. *informational*, criações intralingüísticas dessas línguas, visto que não existe em latim um suposto vocábulo **informationalis*, teremos novamente duas possibilidades: se o ing. *informational* surgiu primeiro, então o fr. *informationnel* é tradução do mesmo; se, ao contrário, foi a forma francesa a que surgiu primeiro, então a inglesa é restituição daquela. Assim, também neste caso a classificação lexicogênica depende da data de introdução dos vocábulos.

Empréstimo de tradução, empréstimo de sentido e empréstimo de restituição: configuração e tipos possíveis

Tanto o empréstimo de tradução quanto o empréstimo de sentido e o empréstimo de restituição caracterizam-se por ser empréstimos de significado, com a conseqüente substituição do significante. Vejamos então, graficamente, alguns exemplos de diferentes tipos possíveis de empréstimos de tradução, de sentido e de restituição, suas relações e sua configuração esquemática, mais uma vez adotando as seguintes convenções:

- herança vernácula do latim ou equivalência semântica entre o vocábulo latino e o vulgar
- → empréstimo direto do latim
-→ neologia semântica (ressementização)
- empréstimo de tradução
- empréstimo de restituição



O exemplo I representa um empréstimo de tradução do alemão em relação a uma acepção já existente em latim, no caso, a acepção geométrica de *angulus* (ver p. 86). O português, na mesma situação, optou pelo empréstimo total do vocábulo latino.

No exemplo II, o francês empresta ao latim o significante *conclusio* (fr. *conclusion*) juntamente com seus dois significados.

Para o primeiro deles, “fechamento, término”, o francês e o alemão já possuíam vocábulos vernáculos correspondentes, *clôture* e *Schluss*, respectivamente. A segunda acepção, “conclusão (de um raciocínio lógico, por exemplo)”, é inédita em ambas as línguas. O reemprego de *Schluss*, em alemão, configura a tradução.

No exemplo III, o francês e o alemão possuem vocábulos vernáculos correspondentes à primeira acepção do lat. *rectus*, “reto (na norma coloquial)”, respectivamente fr. *droit* e al. *recht*. Neste caso, ambas as línguas efetuam tradução em relação à segunda acepção de *rectus*, “ângulo reto (no discurso geométrico)”. Conseqüentemente, há neologia semântica em ambas as línguas.

No exemplo IV, temos um vocábulo latino, *fides*, que significava originariamente “crença, confiança”, tendo posteriormente adquirido, sob a influência da religião cristã, o significado adicional de “fé cristã”. As línguas românicas recebem por herança vernácula ambas as acepções. Os dialetos germânicos, que já tinham vocábulos próprios para designar a primeira acepção, emprestam o sentido religioso de *fides* ainda nos primeiros séculos da era cristã. Desse modo, transmitem por via vernácula para as línguas germânicas que os sucedem historicamente esses vocábulos com ambas as acepções. É o caso do al. *Glaube*, que, a exemplo do fr. *foi*, é vernáculo e herdou tanto a primeira quanto a segunda acepção do lat. *fides*. Trata-se de um caso interessante de empréstimo de sentido ocorrido entre o latim e o germance em época pré-histórica do ponto de vista das línguas germânicas.

O exemplo V, já visto anteriormente, apresenta um empréstimo de tradução alemão em relação a um grecolatinismo francês. Neste caso, o francês empresta o significante do latim, dando-lhe novo significado. O alemão, por sua vez, ressemantiza um vernáculo correspondente.

Em VI, temos um grecolatinismo francês, *émission*, correspondente a um vulgarismo alemão, *Sendung*, ambos com o significado de “envio, lançamento”. Nessa acepção, o fr. *émission* comuta com vulgarismos como *envoi*, por exemplo. Mais tarde, tanto *émission* quanto *Sendung* assumem um novo significado, “emissão (radiofônica)”. Tem-se, portanto, neologia semântica em ambas as línguas. Uma delas criou o neologismo, e a outra realizou empréstimo de sentido.

Os exemplos VII e VIII são análogos. A partir de um vernáculo pertencente à norma coloquial, criam-se neologismos semân-

ticos referentes a normas técnico-científicas (“fruto” no discurso da botânica e “raiz” em matemática). Novamente temos simples empréstimo de sentido. Vale notar, entretanto, que, no exemplo VII, o significado coloquial e o científico de *fruit/Frucht* possuem um núcleo sêmico comum. Já os significados coloquial e científico de *racine/Wurzel* são disjuntos. No primeiro caso, temos parassinonímia e, no segundo, homonímia.

Se no exemplo VI, temos o empréstimo do sentido de um grecolatinismo por um vulgarismo (*émission* → *Sendung*), em IX temos exatamente a situação inversa, isto é, o empréstimo do sentido do vulgarismo francês *domaine* pelo grecolatinismo italiano *dominio*.

Finalmente, em X, temos um empréstimo de restituição: ao vocábulo latino *integer*, “inteiro (coloquial)”, correspondem os vernáculos fr. *entier* e ing. *whole*. Quando o francês confere a *entier* sua acepção aritmética (“número inteiro”), o inglês vai buscar no latim o significante do vocábulo correspondente.

Já havíamos chamado a atenção para o fato de que o empréstimo de tradução se caracteriza graficamente por um empréstimo direto ao latim à esquerda e por um neologismo semântico à direita, em nosso diagrama. Note-se que o empréstimo de restituição apresenta exatamente a configuração inversa, ou seja, neologismo semântico à esquerda, empréstimo ao latim à direita, lembrando que, em nossa convenção, o fenômeno anterior é sempre o da esquerda. Pelo diagrama, podemos notar também que o empréstimo de sentido caracteriza-se sempre por neologia semântica de ambos os lados.

Grecolatinismos resultantes de combinação sintagmática (composição ou derivação)

Além da importação de vocábulos às línguas clássicas, as línguas vulgares também podem constituir grecolatinismos através da combinação sintagmática de elementos léxicos greco-latinos. Essa combinação sintagmática, como se sabe, pode ser de dois tipos: derivação ou composição. No primeiro caso (também chamado de afixação), tem-se a combinação de radicais gregos ou latinos com afixos igualmente gregos ou latinos; no segundo caso, tem-se a combinação de dois vocábulos lexicais (isto é, que con-

tenham pelo menos um lexema) gregos ou latinos, sendo que cada um deles pode, por sua vez, já ser o resultado de combinações anteriores, podendo conter, pois, mais de um radical, ou um radical e um ou mais afixos.

Convém lembrar, contudo, que, por vezes, a derivação pode consistir da combinação de um radical com afixo “zero” (derivação regressiva) ou da simples mudança de categoria gramatical de uma palavra. Assim, em inglês, a partir do grecolatinismo substantivo *experience* formou-se o derivado verbal *to experience*. Este último, por sua vez, passou ao português na forma *experenciar*.

A composição e derivação greco-latinas são particularmente frequentes no discurso técnico-científico, no qual abundam radicais e afixos oriundos do grego e do latim, e cujas próprias normas terminológicas aconselham a formação de novos termos através da combinação sintagmática desses elementos léxicos.

Grecolatinismo sintagmático criado em substituição a outro grecolatinismo

A criação sintagmática de grecolatinismos nem sempre dá origem a vocábulos inéditos, visto que, algumas vezes, um grecolatinismo resultante de derivação está na verdade calcado em outro grecolatinismo. É o caso, por exemplo, do fr. *consacrer*, criado a partir do lat. *consecrare*, tendo o radical latino *secr-* sido substituído pelo alomorfe igualmente latino *sacr-*, por influência do verbo *sacer*, empréstimo do lat. *sacrare*. Nesse sentido, pode-se considerar o verbo *consacrer* como derivado de *sacer* por prefixação.

Algumas vezes, um grecolatinismo sintagmático criado em substituição a outro grecolatinismo pode na verdade ter surgido de forma indireta, como empréstimo de restituição de um vocábulo estrangeiro, que por sua vez é empréstimo de tradução em relação ao latim. Tomemos, por exemplo, o ing. *military*, resultante da combinação do radical latino *milit-* (de *miles*, *militis*) com o sufixo erudito *-ary* (do latim *-arius*). Tal adjetivo corresponde ao lat. *militaris*. Note-se entretanto que o sufixo latino *-aris* foi substituído em inglês pelo correspondente a *-arius*. Dessa forma, embora o ing. *military* remeta ao lat. *militaris*, não constitui um empréstimo direto do vocábulo latino, mas na verdade *military* está calcado sobre o fr. *militaire*, em que o sufixo latino *-aris* foi traduzido

pelo sufixo semi-erudito francês *-aire*, proveniente do lat. *-arius*. Ora, *military* é, assim, um empréstimo de restituição do vocábulo francês, através da substituição do sufixo *-aire* pelo seu cognato inglês *-ary*.

VOCÁBULOS NÃO CLASSIFICÁVEIS NAS CATEGORIAS ANTERIORES

Nem sempre é possível classificar um vocábulo como greco-*latinismo* ou *vulgarismo*. Isto porque há vocábulos que contêm elementos não pertencentes propriamente ao léxico da língua, tais como nomes próprios (principalmente *topônimos* e *antropônimos estrangeiros*), siglas, *onomatopéias*, morfemas fragmentários, etc. Analisaremos, a seguir, cada um desses tipos possíveis de vocábulos.

Vocábulos derivados de nomes próprios

Há muitos vocábulos que derivam de nomes próprios, como sobrenomes de personagens históricos, ou de designativos não vernaculizados de cidades, regiões e países. Com efeito, nomes próprios de pessoas (chamados *antropônimos*) como *Newton*, *Kant*, *Hitler*, etc., e nomes geográficos (*topônimos* ou *corônimos*) como *Besançon*, *Cambridge* ou *Zimbabwe* não podem, a rigor, ser considerados como integrantes do léxico do português, embora ocorram frequentemente em textos de nossa língua. Por essa razão, palavras como *newtoniano*, *kantismo* ou *hitlerista*, assim como *gentílicos* do tipo de *zimbabwês*, por exemplo, apresentam dificuldade de classificação lexicogênica, visto que os únicos elementos realmente classificáveis são os sufixos. Poder-se-ia, talvez, adotar o estatuto lexicogênico do sufixo como critério para classificar tais vocábulos. Assim, o port. *kantiano* seria um greco-*latinismo*, ao passo que o fr. *kantien* seria um *vulgarismo*. Entretanto, tal critério nos parece um tanto artificial, de sorte a considerarmos mais aconselhável classificar vocábulos dessa natureza numa terceira categoria.

Mesmo nomes próprios vernaculizados, como *Platão*, *Lutero*, *Napoleão*, e seus derivados *platônico*, *luterano* e *napoleônico*, por exemplo, não são de classificação menos problemática. O mesmo vale para nomes geográficos como *França*, *Roma*, *Londres*. Embora tais palavras sejam suscetíveis de análise lexicogênica, não figuram em dicionários, quer de língua, quer etimológicos. Já os gentílicos correspondentes *francês*, *romano*, *londrino*, constantes dos dicionários, são unidades perfeitamente incorporadas ao léxico.

Compostos e derivados de nomes próprios são abundantes no discurso técnico-científico. Exemplos disso são termos como *amperímetro* (< Ampère), *andaluzita* (< Andaluzia), *berquélío* (< Berkeley), *devoniano* (< Devon), *férmio* (< Fermi), *kimberlita* (< Kimberly), *macadamizar* (< McAdam), *mendelévio* (< Mendeleiev), *voltagem* (< Volta), *wattímetro* (< Watt), etc.

Embora a gênese e a evolução dos nomes próprios obedeça, sem dúvida, às mesmas leis que regem a gênese e a evolução dos nomes comuns, constituem aqueles o objeto de estudo da onomástica, reservando-se, assim, o termo *lexicogênese* ao estudo destes últimos.

Siglas e seus derivados

Uma sigla não é, a princípio, uma unidade do léxico, mas, antes, a condensação de uma lexia complexa. Assim, por exemplo, *USP* = *Universidade de São Paulo*, *EUA* = *Estados Unidos da América*, *PT* = *Partido dos Trabalhadores*, etc. É evidente que a sigla, por seu poder de condensação informacional, com conseqüente economia de codificação fonológica, acaba, muitas vezes, por transformar-se em palavra, passando inclusive a ser realizada como uma seqüência fonológica coesa, e não como simples seqüência “soletrada”. É o caso de *USP*, por exemplo, que se pronuncia como /'uspi/.²²

De qualquer maneira, do ponto de vista lexicogênico, o objeto analisável não é a sigla, mas sim a locução nominal subjacente a essa sigla, locução, aliás, que só poderá ser analisada após

²² Cfr. BARBOSA, M. A. *Léxico, produção e criatividade: processos do neologismo*. São Paulo, Global, 1981, p. 188.

decomposta em seus constituintes simples (por exemplo, em *Partido dos Trabalhadores*, analisar-se-iam *partido* e *trabalhador*). Ainda que muitas siglas possam ser realizadas fonologicamente como palavras, existe em relação a elas uma consciência lingüística muito clara de que não são, a rigor, palavras, mas sim seqüências de letras iniciais de palavras. Por isso mesmo, as siglas têm sua origem sempre na modalidade escrita da língua. As coisas se complicam quando nos deparamos com um vocábulo composto ou derivado de uma sigla, como por exemplo *abecedário*, *aidético*, *petista*, *ufologia*, etc. Neste caso, já não se tem mais uma sigla, isto é, uma condensação gráfica de uma seqüência de vocábulos, mas sim um verdadeiro vocábulo, que cumpre todas as exigências da estrutura fonológica e morfossintática da língua, e que, no entanto, possui um formante não classificável lexicogenicamente. Há ainda casos de derivados de siglas como *anti-AIDS*, em que o elemento síglico entra em combinação com um prefixo, sem perder seu estatuto gráfico e fonológico de sigla.

Há casos em que a formação da sigla, processo a que damos o nome de siglação, é muito antiga, remontando mesmo ao latim, como em *snob*, vernaculizado para *esnobe*, originalmente grafado *s.nob.* (= *sine nobilitate*, “sem nobreza”, epíteto conferido, na Idade Média, aos plebeus que ascendiam à aristocracia).

Muitas siglas hoje já não são mais sentidas como tais, por se ter perdido a referência à locução nominal ou frase que lhes subjaz. É o caso do ing. *nylon*, que remete à frase “Now you lousy old Nipponese”, pronunciada por uma senhora americana quando em visita à indústria desse material nos Estados Unidos. Da mesma forma, temos o ing. *jeep* (→ port. *jipe*), originalmente *G.P.* (*vehicle for general purposes*), e ainda *radar* (= *radio detection and ranging*).

Vocábulos formados de fractomorfemas

Existem na língua elementos que, de um ponto de vista morfossintático estrito, não são, em princípio, morfemas, pois resultam de um processo de truncamento fonológico dos mesmos, mas que, por entrarem freqüentemente em composições e derivações, acabam por assumir as mesmas propriedades gramaticais dos morfemas. São os chamados *fractomorfemas* (ou *quase-*

morfemas). Esses elementos, muito comuns no discurso coloquial e no vocabulário das telecomunicações e das novas tecnologias, são o resultado da abreviação de vocábulos plenos por motivo de economia enunciativa. Temos, assim, *auto* por *automóvel*, *cine* por *cinematógrafo*, *foto* por *fotografia*, *metrô* por *metropolitano*, etc. Esses elementos podem, às vezes, coincidir com radicais gregos ou latinos, como no caso de *auto* ou *foto*, assim como podem também, e mais freqüentemente, não corresponder a nenhum radical preexistente na língua (caso de *cine* e *metrô*). Os fractomorfemas, além de ser encontrados como vocábulos autônomos, também entram como formantes em composições e derivações, como nos seguintes exemplos: *aquaginástica* (< *aquático* + *ginástica*), *autopeças* (< *automóvel* + *peças*), *Brasilsat* (< *Brasil* + *satélite*), *fotomontagem* (< *fotografia* + *montagem*), *informática* (< *informação* + *automática*), etc. Podemos citar, ainda, *busílis*, *eurotúnel*, *maxidesvalorização*, *minissaia*, *motel*, *pop-art*, *show-mício*, etc. Como se pode ver, há vocábulos resultantes da combinação de fractomorfemas com morfemas plenos, ao passo que há outros formados exclusivamente de fractomorfemas.

No discurso científico, o uso de fractomorfemas é particularmente freqüente na terminologia da química. De fato, muitas designações de compostos químicos são formadas de fragmentos de vocábulos gregos ou latinos. Dentre os principais fractomorfemas de uso corrente em química, citemos:

met- < *methyl* ← gr. *μεθί*+ *ὄλη*

et- < gr. *αιθερ*

prop- < *propiónico* ← gr. *πρό* + *πίων*

but- < lat. *butyrum*

Há exemplos de vocábulos cujo caráter de compostos de fractomorfemas pode, à primeira vista, não ser evidente. É o caso, por exemplo, do port. *reprografia*, formado por *repro-* (redução de *reprodução*) e *-grafia*, que poderia também ser analisado como sendo formado pelos prefixos latinos *re-* e *pro-* e pelo radical grego *-grafia*. Somente o conhecimento da origem desse vocábulo, isto é, da motivação inicial de sua criação, permite classificá-lo

com segurança como fractocomposto e não como derivado greco-latino.

Um caso relativamente freqüente em que ocorre enxerto de fractomorfemas em um vocábulo é o trocadilho, processo através do qual se substitui uma seqüência fonológica de determinado vocábulo por outra, extraída de outro vocábulo, provocando assim um efeito de sentido humorístico: *compositor* (= aquele que compra músicas de outros autores e as publica como suas), *entidade pilantrópica*, etc.

Embora quase todos os fractomorfemas se originem de morfemas greco-latinos ou vulgares, é bastante problemático classificar os vocábulos formados por tais elementos quer como grecolatinismos quer como vulgarismos.

Neologismos fonológicos específicos

A neologia fonológica pode ser de dois tipos, a saber, específica ou complementar. Este último tipo resulta da introdução, no léxico, de um novo significante, em decorrência de um outro tipo de neologia. Entretanto, nos interessa aqui a neologia fonológica específica, aquela que não parte de elementos mórficos preexistentes na língua, mas, sim, cria novos morfemas pela produção de seqüências inéditas de fonemas, desde que permitidas pelo sistema fonológico da língua.

O neologismo fonológico específico pode ser:

a) *ex-nihilo*, isto é, resultante de uma combinação totalmente arbitrária de fonemas (por exemplo, *pirlimpimpim*, *Xangrilá*);

b) onomatopaico, isto é, parcialmente motivado pelo som do referente (por exemplo, *miau*, *tique-taque*).

O primeiro tipo é bastante freqüente no discurso publicitário e na gíria, sendo, ao contrário, bastante raro em outras normas discursivas. Em geral, não dá origem a compostos ou derivados. Já o segundo tipo costuma apresentar derivações, os chamados vocábulos impressivos (*miar*, *ziguezaguear*, *crocante*, etc.).

Os neologismos fonológicos *ex-nihilo* e as onomatopéias puras são, evidentemente, inclassificáveis, quer como grecolatinismos, quer como vulgarismos. Já os vocábulos impressivos podem ser classificados, até porque alguns deles são de origem vernácula

(por exemplo, *cuco* < lat. *cuculus*, *piar* < lat. *pipiare*, *uivar* < lat. *ululare*) ou estrangeira (*pingue-pongue*, do ing. *ping-pong*).

PERCURSO GERATIVO DE CRIAÇÃO VOCABULAR

Há casos em que, para classificar lexicogenicamente um vocábulo, temos de postular a existência de etapas evolutivas não documentadas desse vocábulo, e que, por sinal, jamais existiram efetivamente, mas que são necessárias para explicar a configuração atual do vocábulo. Alguns exemplos tornarão isso mais claro.

1) Tomemos o vocábulo francês *privauté*. Tal vocábulo não pode ser explicado como empréstimo ao latim, visto que não existe nessa língua um suposto vocábulo **privalitas*, nem tampouco como empréstimo a qualquer outra língua vulgar, como, por exemplo, o ing. *privacy* ou o port. *privacidade*. Todavia, esse vocábulo também não pode ser explicado como um neologismo sintagmático vulgar formado pelo adjetivo *privé* e pelo sufixo *-té*, já que a combinação desses dois elementos jamais produziria a forma *privauté*. Assim, a única maneira de explicar *privauté* é admitir que não deriva diretamente de *privé*, mas sim de uma forma hipotética **privat* por analogia com os pares *royal/royauté*, *loyal/loyauté*.

2) Se tomarmos agora o vocábulo francês *onguiculé*, termo de zoologia, veremos que tal vocábulo não se explica como empréstimo ao latim, já que não existe um suposto lat. **unguiculatus*. Igualmente, não se pode cogitar de empréstimo a outra língua vulgar, uma vez que os correspondentes port. *ungüiculado*, it. *unguicolato*, ing. *unguiculate*, etc., são cronologicamente posteriores a *onguiculé*. Também não podemos explicar esse vocábulo como combinação do radical *onguicul-* com o sufixo *-é*, já que o elemento *onguicul-* não preexistia em francês. Desse modo, é forçoso postular a forma latina hipotética **unguiculatus*, da qual o vocábulo francês resulta por metamorfismo do radical (passagem de *u* a *o* diante de nasal) e por tradução do sufixo *-atus* por *-é*. Evidentemente, **unguiculatus* jamais existiu em latim ou em qualquer outra língua; na verdade, tal forma somente teve existência na mente do zoólogo que cunhou *onguiculé* em francês.

3) Seja o vocábulo inglês *chronicle*. Tal vocábulo não pode ser explicado diretamente a partir do latim, uma vez que não existe **chronicula* nessa língua. Tal forma também inexistente nas demais línguas vulgares, e tampouco se explica por derivação a partir do lat. *chronica* com um sufixo *-le*, já que tal sufixo não existe em inglês. Só nos resta então postular um latim hipotético **chronicula* para explicar *chronicle*.

Nota-se, assim, que os processos de vulgarização não incidem apenas sobre vocábulos greco-latinos efetivamente existentes, quer nas línguas clássicas, quer nas línguas vulgares, mas podem afetar também vocábulos hipotéticos. É interessante notar que alguns desses processos de vulgarização estão tão fortemente arraigados em certas línguas, que chegam a intervir no próprio momento da criação de um novo vocábulo, fazendo que este apresente, desde seu nascimento, as mesmas características lexicogênicas de outros vocábulos semi-eruditos já existentes. Pode-se dizer, nesse sentido, que os falantes cultos conhecem, ainda que de forma intuitiva em certos casos, os processos de vulgarização da língua, e os aplicam a suas criações vocabulares, como se fossem verdadeiras regras de conversão fonético-fonológico-ortográfica do grego e do latim para sua língua que devem ser respeitadas.

MECANISMOS DE EMPRÉSTIMO

Os processos lexicogênicos que acabamos de discutir podem ser agrupados em duas categorias básicas, segundo o vocábulo resultante seja constituído apenas de elementos pertencentes à própria língua, sem nenhum empréstimo a outro sistema lingüístico, quer em nível de conteúdo, quer em nível de expressão, ou, ao contrário, contenha algum elemento importado de outro sistema lingüístico. Temos, assim:

- 1) Processos lexicogênicos autogenéticos (autogenia):
 - 1a) herança vernácula;
 - 1b) neologia fonológica;
 - 1c) neologia sintagmática;
 - 1d) neologia semântica.
- 2) Processos lexicogênicos alogenéticos (alogenia):

- 2a) empréstimo sem vernaculização;
 2b) empréstimo com vernaculização fonológica;
 2c) empréstimo com vernaculização sintagmática;
 2d) empréstimo semântico;
 2e) refecção.

Agora que já vimos todos os processos lexicogênicos que atuam nas línguas européias ocidentais, vejamos, no quadro abaixo, alguns exemplos das diferentes possibilidades de empréstimo de vocábulos do grego/latim pelas línguas vulgares e de vocábulos de uma língua vulgar por outra língua vulgar, lembrando, mais uma vez, que o sinal > indica herança vernácula e o sinal → indica empréstimo.

GREGO/LATIM	LÍNGUA A	LÍNGUA B	TIPO DE FENÔMENO
1) <i>directio</i> →	fr. <i>direction</i> →	ing. <i>direction</i>	fr.: grecolatinismo direto ing.: grecolatinismo indireto
2) <i>conclusio</i> →	fr. <i>conclusion</i> →	al. <i>Schluß</i>	fr.: grecolatinismo direto al.: empréstimo de tradução
3) <i>emissio</i> →	fr. <i>émission</i> →	al. <i>Sendung</i>	fr.: grecolatinismo direto al.: empréstimo de sentido
4) <i>insectus</i> →	fr. <i>insecte</i> →	it. <i>insetto</i>	fr.: grecolatinismo direto it.: metamorfismo
5) <i>hydor, phobein</i> →	fr. <i>hydrophobie</i> →	it. <i>idrofobia</i> al. <i>Wasserscheu</i>	fr.: grecolatinismo sintagmático it.: empréstimo de grecolatinismo al.: empréstimo de tradução
6) <i>helix, pteron</i> →	fr. <i>hélicoptère</i> →	it. <i>elicottero</i>	fr.: grecolatinismo sintagmático it.: metamorfismo
7) <i>cambiare</i> >	fr. <i>changer</i> →	ing. <i>change</i>	fr.: vernáculo ing.: empréstimo de vulgarismo
8) <i>arbor</i> > (coloquial)	fr. <i>arbre</i> → (mecânica)	it. <i>albero</i> (mecânica)	fr.: neologismo semântico (vulgarismo) it.: empréstimo de sentido (vulgarismo)
9) <i>dominium</i> > (coloquial)	fr. <i>domaine</i> → (matemática)	it. <i>domínio</i> (matemática)	fr.: neologismo semântico (vulgarismo) it.: empréstimo de sentido (grecolatinismo)

GREGO/LATIM		LÍNGUA A	LÍNGUA B	TIPO DE FENÔMENO
10) <i>integer</i> (coloquial)	>	fr. <i>entier</i> (matemática)	ing. <i>integer</i> (matemática)	fr.: neologismo semântico (vulgarismo) it.: empréstimo de restituição
		→		
11) <i>percipere</i>	→	fr. <i>percevoir</i> →	ing. <i>perceive</i>	fr.: empréstimo de tradução ing.: empréstimo de vulgarismo
12) <i>emittere</i> (coloquial)	→	fr. <i>émettre</i> → (física)	al. <i>senden</i> (física)	fr.: neologismo semântico al.: empréstimo de sentido
13) <i>rectus</i> (geometria)	→	fr. <i>droit</i> (geometria)	ing. <i>straight</i> (geometria)	fr.: empréstimo de tradução ing.: empréstimo de tradução
		→		
14) –		fr. <i>achèvement</i> →	ing. <i>achievement</i>	fr.: vulgarismo sintagmático
–		fr. <i>publicité</i> →	ing. <i>publicity</i>	ing.: empréstimo de vulgarismo
15) –		fr. <i>sous-marin</i> →	ing. <i>undersea</i>	fr.: vulgarismo sintagmático
–		fr. <i>sous-sol</i> →	ing. <i>underground</i>	ing.: empréstimo de tradução
16) –		ing. <i>feed back</i>	fr. <i>réalimentation</i>	ing.: vulgarismo sintagmático fr.: empréstimo de restituição
17) <i>rigidus</i>	>	fr. <i>raide</i>		alótipos vernáculo e refeito (parassinônimos coexistentes)
	→	fr. <i>rigide</i>		
<i>flos</i> >	>	port. <i>chor</i>		alótipos vernáculo e refeito (parassinônimos não coexistentes)
	→	port. <i>flor</i>		
<i>auricula</i>	>	port. <i>orelha</i>		alótipos vernáculo e refeito (significados disjuntos)
	→	port. <i>aurícula</i>		
18) <i>ruina</i>	>	it. <i>rovina</i>		alótipos vernáculo e refeito (sinônimos coexistentes = alomorfes)
	→	it. <i>ruina</i>		

Nos exemplos 1 a 4, temos empréstimo direto, empréstimo indireto, empréstimo de tradução, empréstimo de sentido e empréstimo metamórfico, a partir de vocábulos latinos.

Nos exemplos 5 e 6, temos a criação, dentro de uma língua vulgar, de grecolatinismos sintagmáticos a partir de radicais gregos, e o empréstimo direto, empréstimo de tradução e empréstimo metamórfico em outras línguas vulgares.

Nos exemplos 7 a 10, temos a transmissão por herança de vocábulos do latim a uma língua românica, e o empréstimo direto, empréstimo de sentido a um vulgarismo estrangeiro, empréstimo de sentido a um grecolatinismo estrangeiro e empréstimo de restituição desses vocábulos vernáculos em outras línguas vulgares.

Nos exemplos 11 e 12, temos o empréstimo de tradução de vocábulos latinos por uma língua vulgar, e o empréstimo direto ou empréstimo de sentido em relação a esse vulgarismo por outras línguas vulgares; no exemplo 13, temos o empréstimo de tradução simultâneo em duas línguas vulgares.

Nos exemplos 14 a 16, temos a criação, numa língua vulgar, de um vulgarismo sintagmático, e o empréstimo direto, empréstimo de tradução e empréstimo de restituição do mesmo por outra língua vulgar.

Nos exemplos 17 a 20, temos, numa mesma língua românica, a herança vernácula e o empréstimo direto do mesmo vocábulo latino, o que configura o fenômeno da *alotropia*. Em outro trabalho²³, já tivemos a oportunidade de discutir esse fenômeno de forma mais detalhada. Aqui, nos limitamos a retomar as principais características desses diferentes casos de alotropia, com algumas alterações terminológicas resultantes do próprio avanço de nossa pesquisa nesse campo.

No exemplo 17, temos dois alótipos sincrônicos parassinônimos, isto é, coexistentes na mesma etapa sincrônica da língua, e com significado e distribuição parcialmente coincidentes.

No exemplo 18, temos dois alótipos diacrônicos sinônimos, ou seja, duas formas correspondentes a duas diferentes etapas sincrônicas da língua, cujo significado e distribuição são coincidentes.²⁴

No exemplo 19, temos dois alótipos sincrônicos de significado e distribuição disjuntos, isto é, coexistentes na mesma etapa

²³ BIZZOCCHI, A. L. Aspectos da alotropia nas línguas românicas: esboço de uma tipologia. *Anais da 43ª Reunião Anual da SBPC*, Rio de Janeiro, SBPC, 1991, p. 401-402.

²⁴ Estamos aqui considerando o vocábulo *flor* em sua acepção coloquial, não especializada. É evidente que o vocábulo *flor* em sua acepção científica (na terminologia botânica) apresenta significado e distribuição diferentes tanto de *flor* coloquial quanto de *chor*, que só possuía, aliás, a acepção coloquial.

sincrônica da língua, porém com significados e distribuições totalmente distintos. Trata-se de simples cognatos coetimológicos.

No exemplo 20, temos dois alótropos sincrônicos sinônimos, isto é, coexistentes na mesma etapa sincrônica, e, além disso, de significado e distribuição coincidentes, visto comutarem em todos os contextos. Diferenças, se há, entre tais alótropos, são de natureza estritamente estatística (*rovina* é de uso mais freqüente que *ruina*, especialmente na norma coloquial). São simples variantes alomórficas do mesmo signo.

CRITÉRIOS ESPECIAIS PARA A CLASSIFICAÇÃO DOS VOCÁBULOS

Existem alguns casos específicos de vocábulos cuja classificação exige a adoção de critérios adicionais, que passamos a discutir a seguir.

Dos nomes próprios e seus compostos e derivados

Conforme já vimos em Vocábulos derivados de nomes próprios, na p. 127, os nomes próprios não são, via de regra, classificáveis. Entretanto, os vocábulos compostos e derivados de nomes próprios podem ser tratados de duas maneiras diferentes, segundo a procedência do nome próprio, na seguinte conformidade:

a) Os compostos e derivados de nomes próprios de origem grega ou latina (por exemplo, port. *crísticida*, *platonismo*) podem ser normalmente classificados e enquadrados nos processos lexicogênicos já descritos;

b) Os derivados de nomes próprios vulgares, sejam eles nacionais (como, por exemplo, o port. *camoniano*, de *Camões*) ou estrangeiros (por exemplo, o port. *kantiano*, derivado do nome próprio alemão *Kant*), bem como os gentílicos, em grande parte derivados de nomes geográficos, são de classificação problemática, devido ao fato de que, as mais das vezes, a pesquisa etimológica desses vocábulos é tarefa demasiado complicada para que se possa realizá-la com margem razoável de acerto. O mesmo vale para os compostos de nomes próprios, tais como fr. *ampèremètre*,

port. *voltímetro*, etc. Na verdade, tais vocábulos possuem uma parte não classificável, o que impede que sejam considerados quer como grecolatinismos quer como vulgarismos.

Cabe aqui lembrar que o que distingue o nome próprio do nome comum é o fato de que este último designa uma classe de objetos, ao passo que aquele designa um único objeto ou, se se preferir, uma classe composta de um único objeto. Nesse sentido, *cadeira* é um substantivo comum pois pode aplicar-se a qualquer manufato que possua as características físicas de uma cadeira, enquanto *Napoleão* é um substantivo próprio na medida em que não há dois indivíduos no mundo com essa identidade. Dito em outras palavras, *cadeira*, enquanto signo, representa um sem-número de objetos, ao passo que *Napoleão* é um signo associado a um único objeto. Mesmo que haja outros indivíduos no mundo com esse mesmo nome, serão meros homônimos, isto é, a indivíduos distintos corresponderão signos diferentes, cujo único traço em comum é possuírem o mesmo significante.

Partindo dessa definição, convém observar que alguns vocábulos que se escrevem usualmente com inicial maiúscula não são nomes próprios no sentido exato do termo, na medida em que são na verdade particularizações de nomes comuns. É o caso de vocábulos como port. *Reforma*, *Presidente*, *Renascimento*, etc.; fr. *Dieu*, *État*, *Nord*, etc. Outras vezes, nomes próprios sofrem um processo inverso, generalizando-se semanticamente e, conseqüentemente, dando origem a nomes comuns: port. *bíblia*, *igreja*, *máfia* (a partir de *Bíblia*, *Igreja*, *Máfia*, respectivamente). Há ainda casos de vocábulos que se escrevem com maiúscula por mera convenção ortográfica (por exemplo, ing. *American*, *November*, *Wednesday*, etc.).

Classificação dos advérbios de modo derivados de adjetivos

Os advérbios de modo derivados de adjetivos podem ser de quatro tipos diferentes:

a) advérbios que possuem forma idêntica à do próprio adjetivo (por exemplo, ing. *fast*, al. *möglich*);

b) advérbios que resultam de um adjetivo por metafonía do radical (por exemplo, port. adj. *bom*, adv. *bem*);

c) advérbios que, embora se relacionem semanticamente a um adjetivo, procedem de outra etimologia (por exemplo, ing. adj. *good*, adv. *well*);

d) advérbios que provêm do adjetivo por sufixação. Neste caso, os sufixos usuais são: port. *-mente*, fr. *-ment*, it. *-mente*, ing. *-ly* e *-wise* e al. *-weise*.

Nos três primeiros casos, deve-se proceder à classificação do advérbio como um todo. No caso dos advérbios compreendidos em d), é necessário primeiramente abstrair-se o sufixo, para em seguida realizar-se a classificação, pois, caso contrário, todos os advérbios portadores de sufixos seriam classificados como vulgarismos, já que os sufixos adverbiais são todos vulgares. Como os advérbios pertencentes às demais categorias são sempre vulgares, resultaria daí que todos os advérbios da língua seriam considerados vulgares. Segundo o critério de abstração do sufixo, serão vulgares os advérbios derivados de adjetivos vulgares; os advérbios derivados de adjetivos greco-latinos serão igualmente greco-latinos. Acreditamos ser esse critério de classificação o mais adequado, tendo em vista que a sufixação é o único processo de geração de advérbios produtivo na maioria das línguas européias ocidentais (as demais categorias compreendem apenas formas cristalizadas), sendo o sufixo adverbial um elemento obrigatório na formação do advérbio de modo, o qual, portanto, não deveria, a rigor, ser considerado como um sufixo, mas sim como uma desinência adverbial.

É interessante notar que nem todo advérbio terminado por um dos sufixos relacionados na alínea d) acima é de fato advérbio de modo. Tomando alguns exemplos do português, temos, por um lado a equivalência semântica:

automaticamente = de modo automático

perpendicularmente = de modo perpendicular

mas, por outro lado, temos:

novamente ≠ de modo novo

justamente ≠ de modo justo

Observe-se que os dois últimos exemplos, malgrado sejam derivados de adjetivos, possuem um caráter eminentemente gramatical.

Outro caso digno de nota é o de alguns advérbios de modo do francês e do inglês cujo radical sofre alteração morfológica quando da junção dos sufixos adverbiais fr. *-ment* e ing. *-ly*. Seja, por exemplo, o adjetivo francês *profond*, o qual forma o advérbio *profondément*, portanto com a aposição de um sufixo *-é*, o que nos levaria a pensar num eventual radical *profondé* ao invés de *profond*. Da mesma forma, temos o ing. *automatically* em face de *automatic*, portanto com a aposição de um sufixo *-al*. Deve-se desconsiderar tais sufixos, visto que as formas **profondé* e **automatical* são meras variantes combinatórias de *profond* e *automatic*, respectivamente, e não formas autônomas. Desse modo, as formas efetivamente analisáveis são *profond* e *automatic*.

No francês, encontram-se, ainda, advérbios de modo em que ocorre alteração fonológica do radical, no momento da junção do sufixo *-ment*. São advérbios derivados de participípios presentes latinos ou franceses, tais como *évident* x *évidemment* ou *effrayant* x *effrayamment*. Novamente neste caso, as formas a ser consideradas são *évident* e *effrayant*, e não **évidem-* ou **effrayam-*, pela razão exposta acima.

Vocábulos derivados de radicais greco-latinos com afixos vernáculos de aparência erudita

Alguns afixos pertencentes às línguas românicas, embora perfeitamente vernáculos quanto à origem, não apresentam em algumas dessas línguas nenhuma mutação fonético-fonológica ou gráfica em relação aos afixos latinos de que procedem. É o caso, por exemplo dos sufixos port. e it. *-mento*, cat. e fr. *-ment*, em relação ao lat. *-mentum*; port., esp. e it. *-oso*, em face do lat. *-osus*; port., esp., cat. e it. *-ura*, fr. *-ure*, em vista do lat. *-ura*, etc. Tais sufixos aparecem em palavras vernáculas dessas línguas, tais como port. *cimento*, *costura*; esp. *hermoso*, *armadura*; fr. *serment*, *tournure*, etc. Por outro lado, tais línguas importaram do latim diversas palavras que também possuem esses sufixos: port. *elemento*, *religioso*, *quadratura*; esp. *monumento*, *figura*; fr. *document*, *censure*, etc. Desse modo, passaram a coexistir nessas línguas sufixos eruditos e vulgares alótipos. Sabe-se, por exemplo, que o sufixo português *-mento* é vernáculo em *cimento* e erudito em *emolumento*. Entretanto, tal sufixo é bastante usado nessa

língua para formar derivados, tais como *acabamento*, a partir de *acabar*, e *processamento*, a partir de *processar*. Coloca-se aí uma questão interessante: qual é o estatuto lexicogênico do sufixo *-mento* nos vocábulos acima citados? Note-se que dessa informação depende a própria classificação desses vocábulos, pois, se *-mento* for vernáculo, *acabamento* será derivado de radical vulgar com afixo vulgar, e *processamento* será derivado de radical erudito com afixo vulgar, configurando-se ambos os vocábulos, portanto, como vulgarismos; se, ao contrário, o sufixo em questão for erudito, *acabamento* continuará sendo um vulgarismo, mas *processamento* será derivado de radical erudito com afixo erudito, consubstanciando um grecolatinismo. Coloca-se o mesmo problema em relação a prefixos vernáculos que coincidem formalmente com seus alótipos eruditos: *re-*, *de-*, *con-*, etc.

Uma consulta a obras de lingüística românica nos mostra que os afixos realmente produtivos no processo de criação lexical por derivação no romance são, via de regra, os afixos vernáculos. Isso se deve ao fato de que, à época em que se deram tais derivações, somente estavam disponíveis para tal fim os afixos integrantes de palavras que já pertencessem ao léxico romance. Ora, até esse momento, praticamente todas as palavras eram de origem vernácula, pois, tratando-se de época pré-literária, a maioria dos vocábulos eruditos ainda não haviam sido introduzidos na língua vulgar. Dessa maneira, os afixos que entram em derivações vulgares nessa época são os afixos vernáculos. Os afixos alotrópicos correspondentes de origem erudita somente chegaram à língua vulgar posteriormente, já em época literária, ficando muitos deles restritos a vocábulos diretamente importados do latim. Nesse sentido, podemos dizer que os afixos eruditos não apresentavam até então grande vitalidade nas línguas românicas. A partir da introdução nas línguas vulgares de empréstimos latinos portadores de afixos, muitos desses afixos tornaram-se produtivos nessas línguas, passando a concorrer com seus alótipos vernáculos.

A forma mais segura de sabermos se um determinado afixo, formalmente coincidente com seu alótipo latino, é de fato erudito ou vernáculo é a comparação entre as formas desse afixo nas diversas línguas românicas, uma vez que em algumas dessas línguas o afixo vulgar pode ter sofrido normalmente os processos de mutação fonético-fonológica. Se tal ocorre, a distinção entre o afixo vernáculo e o erudito correspondente é clara. Assim, em

SUFIXOS

LATIM	PORTUGUÊS	ESPAÑHOL	CATALÃO	FRANCÊS	ITALIANO
-mentu	-mento	*-miento	-ment	-ment	-mento
-osu	-oso	-oso	-ós	*-eux	-oso
-ura	-ura	-ura	-ura	-ure	-ura
-ivu	-ivo	-ivo	*-iu	*-if	-ivo
-ante	-ante	-ante	-ant	-ant	-ante
-ente	-ente	*-iente	-ent	-ent	-ente
-ale	-al	-al	-al	*-el	-ale

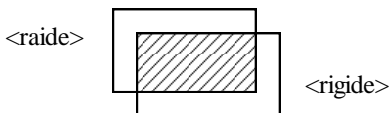
É interessante notarmos que, por força da flexão, certos sufixos vernáculos apresentam tanto uma forma coincidente com a latina quanto uma forma própria, evoluída foneticamente em relação ao latim. Assim, por exemplo, o sufixo vernáculo fr. *-if* possui um feminino *-ive*, coincidente com o latim. Do mesmo modo, o sufixo vernáculo cat. *-(i)ó* (< lat. *-ione*: por exemplo, lat. *ratione* > cat. *raó*; lat. *natione* → cat. *nació*) faz o plural *-(i)ons* (cfr. cat. *nacions* ← lat. *nationes*). De qualquer maneira, tais sufixos são sempre vulgares, e, conseqüentemente, os vocábulos que os portam são vulgarismos.

Vocábulos alótropos

Em Mecanismos de empréstimo, p. 133, tivemos a oportunidade de apresentar alguns exemplos de vocábulos alótropos, os quais se dividem nas seguintes categorias:

a) alótropos sincrônicos parassinônimos (significado e distribuição parcialmente coincidentes);

Ex.: fr. *raide/rigide*



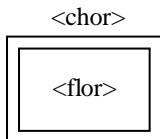
b) alótropos sincrônicos de significado e distribuição disjuntos;

Ex.: port. *orelha/aurícula*



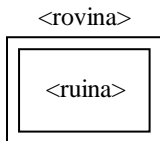
c) alótipos diacrônicos sinônimos (significado e distribuição coincidentes);

Ex.: port. *chor/lor*



d) alótipos sincrônicos sinônimos (significado e distribuição coincidentes).

Ex.: it. *rovina/ruina*



Em todos os casos acima citados, temos uma forma vulgar, de origem vernácula, e uma forma erudita, resultante de empréstimo ao latim. Nos dois primeiros casos, temos alótipos com significado e distribuição diferentes, ao passo que nos dois últimos, temos alótipos com significado e distribuição coincidentes. Convencionamos, assim, com base na definição de refecção de vulgarismo que demos em *Refecções de vulgarismos*, p. 117, que um vocábulo greco-latino alótipo de um vulgarismo somente seja tratado como refecção daquele se houver entre ambos coincidência de significado e distribuição (exemplos *c* e *d*); caso contrário, o grecolatinismo será simples empréstimo erudito, ou empréstimo de restituição, se for o caso.

Tratamento da vogal pós-tônica

Sabe-se, com base em leis fonéticas bem conhecidas, que a pressão articulatória no aparelho fonador diminui ao longo da cadeia falada. Assim, essa pressão (que equivale à quantidade de ar acumulada nos pulmões) é maior no início de cada palavra e menor no fim. Por essa razão, são os sons finais de cada palavra, sobretudo os sons pós-tônicos, aqueles mais sujeitos a alterações fonéticas, especialmente a metafoia (mudança de qualidade) e a redução (encurtamento da duração), cuja conseqüência última pode ser o desaparecimento completo desses sons. É por essa razão que muitos proparoxítonos se haviam tornado paroxítonos em latim vulgar. É por isso também que em francês, por exemplo, a desinência *-a* do latim transformou-se em *-e*, e as desinências latinas *-e* e *-o* foram suprimidas. Dessa forma, a vogal pós-tônica de palavras proparoxítonas sofreu em muitas línguas metafoia ou síncope (por exemplo, lat. *bibitus* > port. *bêbado*; lat. *opera* > port. *obra*). Tal fenômeno ocorre igualmente nas línguas germânicas, atingindo tanto palavras vernáculas quanto empréstimos.

Menos freqüente, mas igualmente importante, é o desenvolvimento de uma vogal pós-tônica entre duas consoantes em palavras onde anteriormente não existia, a fim de facilitar a pronúncia, fenômeno que recebe o nome de *anaptixe* — ou *suarabácti* —, e que configura um tipo particular de epêntese (acrêscimo de fonemas no interior da palavra). Pode, ainda, ocorrer inversão de posição entre a vogal pós-tônica e uma consoante adjacente, metaplasmo fonético que chamamos de metátese. Tais fenômenos, a princípio característicos de palavras vernáculas, estenderam-se eventualmente a empréstimos estrangeiros, inclusive empréstimos greco-latinos. Como a vogal pós-tônica pertence, via de regra, ao radical da palavra, os processos em questão, quando aplicados a empréstimos greco-latinos, podem constituir metamorfismo, razão pela qual é importante levá-los em conta para uma correta classificação lexicogênica dos vocábulos por eles afetados.

Dividiremos nosso estudo do tratamento da vogal pós-tônica em quatro fenômenos específicos, a saber, *síncope*, *anaptixe*, *metafoia* e *metátese*, e discutiremos em que circunstâncias esses fenômenos devem ou não ser considerados como metamorfismos.

a) *Síncope*. A vogal pós-tônica costuma sofrer síncope quando é seguida de outra vogal (em geral, a vogal temática da palavra²⁵), ou quando é seguida de consoante líquida (*l* ou *r*). Exemplos:

lat. *tendentia* → it. *tendenza*

lat. *justitia* → port. *justiça*

lat. *articulum* → fr. *article*

De modo geral, a síncope constitui um processo vulgarizante. Entretanto, entendemos que a mesma não deva ser considerada como um metamorfismo quando a vogal pós-tônica é imediatamente seguida de outra vogal, pois, nesse caso, a vogal pós-tônica, que se reduz foneticamente a uma semivogal, encontra-se muito pouco individualizada, havendo inclusive uma acentuada tendência à fusão de ambas as vogais (considere-se, a título de exemplo, o it. *studio*, que tem como plural *studi*, outrora grafado *studii*). Desse modo, não apenas a vogal final, mas o ditongo formado por ela e pela vogal pós-tônica, é, como um todo, sentido como uma desinência. Já nos vocábulos proparoxítonos do latim vulgar, a vogal pós-tônica seguida diretamente pela vogal final do vocábulo perde seu valor silábico ou desaparece, passando tais vocábulos a ter valor de paroxítonos. Além disso, em virtude da inexistência de uma consoante entre a vogal pós-tônica e a vogal final, quando esta última cai, a vogal pós-tônica acaba por ficar em posição final absoluta na palavra. Ocorre que as línguas que eliminam sistematicamente a vogal final possuem, por isso mesmo, como caracterís-

²⁵ A rigor, as gramáticas tradicionais costumam distinguir entre *vogal temática*, indicativa de um paradigma de declinação, e *desinência*, morfema responsável pela flexão da palavra. Entretanto, em diversas línguas indo-européias, inclusive em latim, a vogal temática dos substantivos e adjetivos funde-se, em muitos casos, à desinência da palavra, devido ao fenômeno da assimilação (por exemplo, lat. *lup-i*, a partir de **lup-o-i*, o que leva o singular *lupus*, outrora *lup-o-s*, a ser decomposto como *lup-us*). Conseqüentemente, o morfema resultante dessa fusão acumula as funções temática e gramatical, indicando, ao mesmo tempo, a classe gramatical, a função sintática e o paradigma de declinação da palavra (Veja-se, a esse respeito, SAUSSURE, F. *Curso de lingüística geral*. São Paulo, Cultrix, 1969, p. 180). Mais ainda, nas línguas européias ocidentais modernas, a perda do sistema de casos levou ao desaparecimento da maioria das antigas desinências, passando a vogal temática a assumir tal papel. Em nosso trabalho, estamos considerando como desinência todo morfema final de palavra que não pertença ao radical nem constitua afixo. Dessa forma, na palavra *leiteiro*, *leit-* é o radical, *-eir-* é sufixo e *-o* é desinência.

tica fonológica ter suas palavras terminadas por consoante. Desse modo, a manutenção da vogal pós-tônica torna-se insustentável nesse caso. Tome-se, por exemplo, o lat. *tendentia*. Tal vocábulo, ao ingressar na língua alemã, teve de sofrer um processo de enquadramento morfológico que implicou a substituição da desinência latina *-a* por uma desinência “zero” em alemão. Em decorrência disso, a forma resultante alemã teria sido **Tendenzi*; entretanto, o padrão fonológico do alemão não admitiria a terminação em *-i*, razão pela qual essa vogal também teve de ser suprimida, resultando a forma *Tendenz*. O mesmo não ocorreu, por exemplo, com o lat. *symbolus*, que produziu em alemão a forma *Symbol*. Aqui, a presença de um *l* entre a vogal pós-tônica e a vogal desinencial impediu que a primeira fosse afetada pela queda da desinência *-us*. Pode-se dizer, desse modo, que a síncope da vogal pós-tônica seguida de outra vogal é quase uma coerção do sistema fonológico de certas línguas, constituindo-se em fenômeno obrigatório, comparável ao processo de metafonia do radical que descrevemos em Radicais que sofrem metafonia, p. 105. Em vista dessas considerações, em nossos dois primeiros exemplos de síncope da vogal pós-tônica (it. *tendenza*, port. *justiça*) não há vulgarização, mas apenas no terceiro (fr. *article*).

OBSERVAÇÃO: Igualmente, não configura vulgarização a queda da vogal pré-tônica de verbos derivados de substantivos ou adjetivos que tenham sofrido síncope da vogal pós-tônica, nas condições acima estabelecidas. Desse modo, o substantivo francês *divorce* (← lat. *divortium*) não apresenta vulgarização, o mesmo se dando com o verbo *divorcer* (← lat. *divortiare*), por exemplo.

b) *Anaptixe*. Em grego e latim, esse fenômeno ocorre em geral no caso nominativo de substantivos e adjetivos cuja desinência é zero. Assim, dado o radical latino *ministr-*, que forma as flexões *ministerium* (acusativo), *ministro* (ablativo), *ministri* (genitivo), etc., temos o nominativo *minister*, com *e* anaptíctico. Tal fenômeno se dá por analogia com palavras de origem indo-européia que apresentam apofonia de grau: lat. *pater* (grau normal) x *patris* (grau zero).

Em empréstimos greco-latinos, a anaptixe ocorre mais frequentemente nas línguas germânicas, onde a supressão da desi-

nência requer, em contrapartida, a inserção de uma vogal de apoio, por razões exclusivamente fonéticas. Assim, ao importar o lat. *theatrum*, o alemão suprime a desinência *-um*; conseqüentemente, uma vez que o sistema fonético-fonológico do alemão não admite palavras terminadas em *tr*, faz-se necessária a intercalação de uma vogal de apoio entre *t* e *r*, formando sílaba, de que resulta o vocábulo alemão *Theater*. Uma vez que a anaptixe de uma vogal pós-tônica já era processo corrente em grego e latim, e sua ocorrência é em geral obrigatória do ponto de vista fonético-fonológico, já que a ausência de uma vogal de apoio tornaria o vocábulo impronunciável, entendemos que tal fenômeno não configura metamorfismo.

c) *Metafonia*. Frequentemente, a vogal pós-tônica não chega a desaparecer, mas sofre modificação de sua qualidade. Vejamos alguns exemplos:

lat. *angulus* → it. *angolo*

lat. *articulum* → al. *Artikel*

lat. *metaphora* → al. *Metapher*

lat. *philologus* → cat. *filòleg*

lat. *nucleus* → cat. *nucli*

É interessante notar que, algumas vezes, temos a metafonia pura, como no caso do it. *angolo*, em que o *u* pós-tônico do lat. *angulus* passa a *o* por analogia com palavras vernáculas, enquanto, outras vezes, a metafonia na verdade resulta de síncope e posterior anaptixe. É o caso do sufixo alemão *-abel*, originário do lat. *-abilis*. Aqui, houve primeiramente a queda do *i* pós-tônico latino: *-abile* > *-able*; a seguir, a queda da desinência, o que levaria a uma forma hipotética **-abl*, a qual, por ser foneticamente inadmissível, exigiu a inserção de uma vogal de ligação, resultando na forma *-abel*. As formas flexionadas correspondentes desse sufixo são *-abler*, *-able*, *-ables*, etc., isto é, sem vogal de ligação. Em todos os casos, a metafonia da vogal pós-tônica constitui vulgarização.

d) *Metátese*. Este fenômeno é particularmente comum em francês, por razões prosódicas. Temos, assim:

lat. *tertiarius* → fr. *tertiaire*

lat. *illusorius* → fr. *illusoire*

Ocorre que, nessa língua, dada a inexistência da acentuação proparoxítona, a metátese da vogal pós-tônica é usada como um recurso para impedir o deslocamento do acento tônico para a sílaba seguinte. A transformação dos sufixos latinos *-arius* e *-orius* em *-aire* e *-oire* em francês seguiu o modelo de palavras vernáculas, como *paire*, do lat. *paria*, e *miroir*, do lat. *miratorium*, por exemplo. A metátese, nesse caso, além de ser motivada pela imitação de formas vernáculas, é um procedimento foneticamente não obrigatório, pois formas como **tertiarie*, **illusorie*, com deslocamento do acento tônico, ou **tertiare*, **illusore*, com queda da vogal pós-tônica, seriam perfeitamente possíveis em francês. Assim, a metátese constitui sem dúvida alguma um tipo de metamorfismo.

Até agora, analisamos os fenômenos que podem afetar a vogal pós-tônica de um vocábulo isoladamente. Há casos, entretanto, em que tais fenômenos co-ocorrem num mesmo vocábulo. Seja o al. *Bibel*, empréstimo do lat. *biblia*. Tivemos aí em primeiro lugar a síncope da vogal pós-tônica seguida de outra vogal, bem como a queda da vogal desinencial, processos que, conforme vimos acima, não são vulgarizantes. A seguir, tivemos a epêntese de uma vogal de ligação, processo igualmente não vulgarizante. Todavia, comparando-se o vocábulo alemão ao latino, observa-se que aquele difere deste o suficiente para não mais possuir o estatuto de grecolatinismo.

Seja agora o ing. *proper*, proveniente do fr. *propre*, que por sua vez é empréstimo do lat. *proprius*. Em francês, houve apenas a queda da vogal pós-tônica seguida de outra vogal, portanto o vocábulo francês é um perfeito grecolatinismo. Já em inglês, houve a metátese do grupo *re* para *er*, ou, se se preferir, houve a apócope do *e* final, com subsequente anaptixe de um *e* de ligação. Assim, em inglês, somou-se ao metaplasmo já herdado do francês um novo metaplasmo, redundando num vocábulo cujo radical difere do latino por mais de um caráter. Partindo da própria definição de metamorfismo, que consiste de mutação fonético-fonológica do radical, acompanhada da adaptação ortográfica correspondente (ver Empréstimos de vocábulos greco-latinos, p. 60), vê-se que qualquer radical pertencente a uma língua vulgar que apresente diferenciação fonético-fonológica conexa a uma diferenciação gráfica em relação ao radical grego ou latino de que provém por empréstimo constitui metamorfismo. Estamos aqui

aceitando como únicas exceções a esta regra a síncope da vogal pós-tônica seguida de outra vogal e a anaptixe de vogal pós-tônica de ligação, pelas razões já expostas anteriormente. No entanto, o acúmulo desses processos num mesmo vocábulo dá origem, na verdade, a um terceiro processo, de que resulta uma diferenciação entre os radicais erudito e vulgar suficientemente grande para que não se tenha mais aí um vocábulo genuinamente erudito. Pode-se dizer, assim, que a ocorrência isolada de cada um dos metaplasmos acima estudados não consubstancia a vulgarização, mas a co-ocorrência de dois ou mais deles no mesmo vocábulo constitui metamorfismo.

Vocábulos de origem grega terminados em *-ma*

Existem nas línguas vulgares inúmeros vocábulos eruditos resultantes de empréstimo de substantivos deverbiais gregos terminados pelo sufixo *-ma*, tais como port. *esquema*, *panorama*, *problema*, *programa*, *tema*, etc. Enquanto em português, espanhol e italiano tal sufixo possui a mesma forma do grego, nas demais línguas tal sufixo geralmente se apresenta nas formas fr. *-me* (*thème*), ing. *-me* (*programme*) ou *-m* (*problem*), al. *-m* (*Problem*). Nota-se assim que em tais línguas o *a* final desses vocábulos sofre o tratamento morfológico normal de uma desinência *-a*, ou seja, transforma-se em *-e* ou é eliminado. Ocorre porém que tal fonema não corresponde realmente a uma desinência, mas na verdade faz parte do sufixo, já que em grego as palavras com essa terminação pertencem à terceira declinação, cujo tema é zero, e não à primeira, de tema em *a*. Desse modo, pode-se dizer que o *a* final do vocábulo português *problema* não é o mesmo *a* de *mesa*, por exemplo. Na realidade, em *mesa*, o *a* é uma vogal temática nominal, ao passo que o *a* de *problema* é parte integrante do radical da palavra, que possui conseqüentemente vogal temática zero.

Em decorrência disso, deveríamos concluir que a metafonia ou a queda do *a* final dos vocábulos em questão em francês, inglês e alemão, por atingirem o radical do vocábulo, constituem metamorfismo. Por conseguinte, tais vocábulos deveriam ser considerados vulgarismos nessas línguas. Entretanto, devemos lembrar, em primeiro lugar, que no próprio grego, assim como em latim, palavras da terceira declinação terminadas em *a* freqüentemente se

confundem com palavras da primeira declinação, igualmente terminadas em *a*. Note-se, por exemplo, que, em latim, a palavra *schema* admite tanto a flexão *schema/schematis*, conforme a terceira declinação, quanto a flexão *schema/schetae*, conforme a primeira. Desse modo, percebe-se que já nas línguas clássicas o *a* do sufixo *-ma* era por vezes sentido como vogal temática. Em consequência disso, inúmeros compostos e derivados cultos de palavras portadoras de tal sufixo pressupõem uma vogal temática *a*. Assim, temos em português *macrossomatia* e *macrossomia*, *sistemático* e *sistêmico*, etc. Igualmente temos *hematófago*, mas *hemofílico*; *programático*, mas *panorâmico*, e assim por diante.

Em segundo lugar, é comum nas línguas vulgares a mudança de declinação de algumas palavras gregas ou latinas. Dessa forma, tem-se que o fr. *climat* e o ing. *climate* partem de uma forma grega hipotética *κλίματος*, pertencente à segunda declinação, e não da forma real *κλίμα*, pertencente à terceira. Por outro lado, o it. *automa* pressupõe um hipotético vocábulo grego *αὐτόμα*, da terceira declinação, por oposição ao vocábulo real *αὐτόματον*, da segunda declinação.²⁶ Todos esses vocábulos, incontestavelmente eruditos, demonstram que a passagem dos vocábulos de origem grega terminados em *-ma* da terceira para a primeira declinação é procedimento perfeitamente lícito, que não compromete o estatuto erudito desses vocábulos.

Por essa razão, se certas línguas tratam o *a* do sufixo grego *-ma* como desinência, submetendo-o, em consequência, a um enquadramento nos paradigmas de flexão próprios dessas línguas, o fazem porque já na tradição clássica tal procedimento era aceitável. Assim, vocábulos como fr. *thème*, ing. *programme*, al. *Problem*, etc., devem ser tratados normalmente como grecolatinismos.

²⁶ O vocábulo it. *automa* pode entretanto ser considerado também como um metamorfismo por hiperurbanismo (cf. *Metamorfismo por hiperurbanismo*, p. 104), uma vez que o gr. *αὐτόματος*, “que age por iniciativa própria”, remete a um radical indoeuropeu **mētis* (> lat. *mens*), que significa “pensar”. Nesse sentido, a terminação *-ματος* nada teria a ver com o sufixo deverbal *-μα*.

Vocábulos classificáveis em mais de uma categoria

Pode ocorrer de um vocábulo ser de classificação problemática, em virtude de ser passível de inclusão em mais de uma categoria lexicogênica. Isso se dá quando dois processos lexicogênicos diferentes, aplicados ao mesmo vocábulo de origem, conduzem a uma mesma forma resultante, e, além disso, os dados fonético-fonológicos, ortográficos, bem como as informações a respeito da história desse vocábulo não permitem elucidar o problema. Seja, por exemplo, o fr. *inondation*, empréstimo do lat. *inundatione*. Sabemos que se trata de um vulgarismo, uma vez que apresenta divergência fonológica e gráfica em relação ao latim. Entretanto, por um lado, podemos supor que o vocábulo francês proceda do latino por empréstimo de tradução do radical erudito *und-* pelo vulgar *ond-*, destacável do substantivo vernáculo francês *onde*. Por outro lado, tal vocábulo pode ter resultado de um metamorfismo, semelhante ao que produziu em francês *fonction* a partir do lat. *functione*. Note-se que ambas as explicações conduzem ao mesmo resultado, sendo impossível dizer ao certo qual dos dois processos foi o que realmente ocorreu. Todavia, neste caso dispomos de algumas indicações que podem nos auxiliar a classificar tal vocábulo. Se tomarmos o it. *inondazione*, e o compararmos ao fr. *inondation*, poderemos de início pensar que se trata em ambas as línguas do mesmo processo lexicogênico. No entanto, conforme vimos na p. 106, a respeito dos principais metamorfismos de cada língua, o italiano apenas transforma *u* em *o* em posição tônica ou pós-tônica. Por outro lado, a passagem de *un* para *on* em francês é um metamorfismo bastante freqüente nessa língua. Do ponto de vista do empréstimo de tradução, pode-se dizer que no francês tal processo atinge mais freqüentemente os afixos, raramente os radicais. Em italiano, ao contrário, tal processo incide igualmente sobre radicais e prefixos; apenas os sufixos não costumam ser alvo freqüente de traduções. Dessa maneira, somos levados a acreditar que, enquanto o fr. *inondation* resultou de metamorfismo, seguindo o mesmo modelo de *fonction*, o it. *inondazione* resultou de empréstimo de tradução, a partir do vocábulo vernáculo *onda*.

Cumpramos ressaltar que, embora no exemplo dado tenhamos conseguido apresentar uma hipótese plausível, nem sempre a classificação de um vocábulo de natureza lexicogênica dúbia terá

o mesmo sucesso. Com efeito, podem ocorrer casos em que nem mesmo informações externas, como por exemplo o comportamento estatístico dos processos lexicogênicos em cada língua, podem dirimir satisfatoriamente a dúvida na classificação. É o que ocorreria se, por acaso, tanto o francês quanto o italiano admitissem o mesmo tipo de metamorfismo, e se o empréstimo de tradução tivesse nessas duas línguas a mesma distribuição. É o que ocorre também com o vocábulo italiano *fattore* (← lat. *factor*), que tanto pode provir de um metamorfismo (passagem de *-ct-* a *-tt-*), quanto pode originar-se de uma tradução do radical latino *fact-* pelo vernáculo *fatt-*, abstraído de *fatto*, por exemplo. Em italiano, ambos os processos são bastante frequentes.

CAPÍTULO III

APLICAÇÃO DA TEORIA LEXICOGÊNICA À ANÁLISE ESTATÍSTICA CONTRASTIVA DE LÉXICO

PROCEDIMENTOS LEXICOLÓGICOS

Conforme expusemos na Introdução, nosso trabalho consiste da elaboração de uma teoria que explique a dinâmica da criação e renovação lexicais do ponto de vista dos processos de geração das palavras (lexicogênese), e, conseqüentemente, da proposição de um método de análise e classificação das unidades lexicais de discurso (vocábulos) com base nessa teoria, bem como de uma análise estatística contrastiva do léxico de cinco línguas européias, utilizando o referido método. Nosso objetivo é, portanto, estudar o comportamento do léxico dessas línguas mediante procedimentos estatísticos, para, finalmente, realizar um confronto entre as mesmas, tendo por base as semelhanças e diferenças no comportamento lexicogênico dos respectivos léxicos. Tendo em vista ser impossível estudar o léxico de uma língua diretamente, uma vez que o mesmo se configura como um conjunto aberto, com número indefinido e variável de elementos, que contém elementos de alta e baixa freqüência, bem como elementos virtuais, isto é, ainda não efetivamente criados, porém criáveis e utilizáveis a qualquer momento, nosso estudo teve de ser realizado a partir de inferências estatísticas sobre uma amostragem, ou seja, sobre um conjunto finito, fechado e pouco extenso de unidades léxicas efetivamente existentes. Nosso trabalho consistiu, portanto, em selecionar textos escritos nas cinco línguas, coletar todas as ocorrências efetivas dos vocábulos desses textos que preenchem certas condições por nós estabelecidas (ver Escolha das unidades léxicas a serem analisadas e organização do banco de dados, p. 160), tabulá-los, classificá-los lexicogenicamente, indexá-los, para, em seguida, com base nessa indexação, procedermos às operações estatísticas de contagem, determinação de taxas de freqüência, etc., e, finalmente, elaboração de tabelas e gráficos, nos quais

baseamos a análise e a interpretação dos resultados de nossa investigação.

Estabelecimento do *corpus*

Escolha dos textos

Realizamos nossa análise a partir de dois diferentes universos de discurso, a saber, o discurso científico e o discurso jornalístico, a fim de verificar a ocorrência de variações no comportamento léxico de cada língua em função do universo de discurso considerado. Pretendemos também reconhecer alguns dos diferentes processos neológicos empregados por cada língua, bem como identificar, a partir da análise estatística, suas tendências predominantes, para sabermos, inclusive, em que medida esta ou aquela tendência, diacronicamente delineada, determina o comportamento neológico atual e até que ponto tal comportamento já não se incorporou à própria ideologia da língua e, conseqüentemente, da cultura que lhe subjaz.

Em ambos os universos de discurso, houve a preocupação de utilizar textos que efetivamente retratassem a língua em seu estado atual. Assim, a atualidade e a representatividade dos textos foram os requisitos básicos que balizaram nossa escolha. No caso do discurso jornalístico, utilizamos edições recentes dos jornais e revistas de maior circulação em seus respectivos países. Quanto ao discurso científico, entendemos ser representativos livros e artigos produzidos por pesquisadores e pensadores eminentes, cujo trabalho e cuja linguagem, aí incluído o vocabulário pessoal de cada autor, bem como os neologismos por ele criados, influenciam de um modo geral a expressão lingüística de toda a comunidade científica e intelectual. No que tange à questão da atualidade, decidimos trabalhar com textos publicados nos últimos vinte anos. Entendemos que, em termos do discurso científico, trabalhos publicados há mais tempo poderiam não mais refletir a realidade da língua atual, especialmente tendo em vista a própria dinâmica da ciência, cuja velocidade o discurso científico deve acompanhar. Por outro lado, se adotássemos um intervalo de tempo menor (digamos, os últimos cinco ou dez anos), correríamos o risco de

não encontrar nessa amostra nenhum trabalho realmente representativo, já que, pelas características peculiares através das quais se dá a transmissão da informação nos meios científicos, um determinado trabalho costuma levar em média um tempo superior a esse para obter repercussão e tornar-se, portanto, obra de referência para outros autores.

O discurso científico foi subdividido em oito áreas do conhecimento, a saber:

- 1) ciências naturais:
 - a) matemática;
 - b) física;
 - c) química;
 - d) biologia.
- 2) filosofia e ciências humanas e sociais:
 - a) filosofia;
 - b) história;
 - c) sociologia;
 - d) lingüística e semiótica.

Fomos forçados a não incluir livros científicos na amostra, em virtude da dificuldade de encontrar livros de autores de renome, publicados pela primeira vez há menos de vinte anos, sobretudo na área das ciências naturais. De fato, a partir do final da Segunda Guerra Mundial, tem havido uma tendência crescente no sentido de os grandes cientistas publicarem os resultados de suas pesquisas na forma de artigos ao invés de livros. Desse modo, a maior parte dos livros de publicação recente são compilações de artigos publicados originariamente em revistas, em geral há mais de vinte anos. Em virtude disso, decidimos trabalhar exclusivamente com revistas científicas consideradas de grande prestígio dentro da comunidade lingüística em cuja língua são publicadas, incluindo anais e atas de sociedades científicas renomadas. Essa decisão traz inclusive uma vantagem metodológica: enquanto um livro é, em geral, obra de um único autor, uma revista contém trabalhos de diversos autores, minimizando-se com isso eventuais interferências de caráter estilístico no comportamento do vocabulário aí depreendido. A escolha de revistas de grande prestígio se deve ao fato de que é nessas revistas que pesquisadores conceituados costumam publicar seus trabalhos.

Cumprir observar que algumas das principais publicações científicas brasileiras, dentre elas as duas únicas publicações representativas na área de matemática, não puderam ser utilizadas por serem editadas exclusivamente em inglês.

Metodologia para a coleta de material

Uma vez escolhidos os títulos das publicações científicas, os volumes ou fascículos a ser utilizados foram selecionados, dentre aqueles publicados nos últimos vinte anos, segundo um método estatístico, através da aplicação de uma tabela de números aleatórios. Feito isto, as páginas dos fascículos a ser utilizadas na amostra foram obtidas através do mesmo método.

A seleção das páginas das publicações jornalísticas sofreram tratamento semelhante, feita a ressalva de que foram desprezadas mensagens publicitárias e outros textos não pertinentes ao discurso jornalístico, os quais no entanto também costumam ocorrer em jornais e revistas. No caso específico dos jornais, em que várias matérias diferentes figuram numa mesma página, também a matéria foi escolhida aleatoriamente.

Uma vez selecionado todo o material a ser utilizado, foram coletadas todas as ocorrências das unidades léxicas que interessam a nosso estudo, que, conforme veremos no Item seguinte, são os vocábulos lexicais. Procedemos, assim, à coleta de oito amostras de exatamente 250 palavras-ocorrência para cada língua no discurso científico, correspondendo cada amostra a uma das oito áreas do conhecimento em que tal universo de discurso foi subdividido (ver p. 158). Para o discurso científico em língua portuguesa, foram utilizadas sete amostras ao invés de oito, em virtude da inexistência de textos de matemática em português, conforme explicamos acima. No discurso jornalístico, coletamos quatro amostras de 500 palavras-ocorrência por língua, cada amostra correspondendo a um jornal ou revista. Foram coletadas, portanto, 2.000 palavras-ocorrência para cada língua e para cada universo de discurso, à exceção do discurso científico em português, que contou 1.750 palavras-ocorrência, totalizando assim 19.750 palavras-ocorrência. Em razão de termos optado por trabalhar com amostras de igual tamanho, é evidente que nem todo o material coletado foi utilizado integralmente, mas, na verdade, a coleta se

deu até que o número total de palavras estipulado para cada amostra fosse atingido.

Escolha das unidades léxicas a serem analisadas e organização do banco de dados

Como já dissemos, nosso objeto de estudo é a unidade de norma, isto é, o vocábulo. Nossa análise restringiu-se aos vocábulos lexicais, isto é, aqueles cuja estrutura morfológica comporta pelo menos um morfema lexical (*lexema*, segundo a terminologia de Pottier), tendo sido desconsiderados os vocábulos gramaticais, ou seja, aqueles formados exclusivamente de morfemas gramaticais (*gramemas*, segundo Pottier).²⁷ A razão de tal recorte metodológico é que os vocábulos gramaticais, por sua própria natureza, não estão sujeitos, via de regra, aos processos de criação e renovação lexical que constituem o escopo de nossa análise lexicogênica, fundamentalmente baseada na dicotomia *erudito/vulgar*. Com efeito, a maior parte dos vocábulos gramaticais de qualquer língua são vernáculos ou resultantes de neologia sintagmática (composição ou derivação). Raramente se observam nesse domínio outros processos lexicogênicos, tais como o empréstimo, a tradução, o metamorfismo, etc., razão pela qual é praticamente impossível estabelecer no âmbito do léxico gramatical oposições entre grecolatinismos e vulgarismos, entre elementos autóctones e alógenos, e menos ainda detectar a transformação de unidades de uma categoria em outra. Além disso, o conjunto dos vocábulos gramaticais de uma língua constitui um inventário fechado, pouco extenso, cujas unidades permanecem praticamente as mesmas ao longo do tempo. Enquanto no subconjunto das lexias lexicais do universo léxico de uma língua, a cada instante, pode ser criada uma nova unidade, unidades de baixa freqüência podem sofrer um aumento

²⁷ É bem verdade que muitos vocábulos gramaticais se originam de vocábulos lexicais (citem-se, a esse título, não só locuções preposicionais e conjuncionais, tais como port. *em face de*, *por meio de*, *junto a*, *de acordo com*, etc., como também preposições e conjunções como port. *segundo*, *conforme*, etc., pronomes como fr. *personne*, advérbios como esp. *mañana*, al. *morgen*, etc.). Por outro lado, existem também vocábulos lexicais formados exclusivamente de gramemas, como por exemplo o port. *ensimesmado*.

na frequência de seu uso, unidades de alta frequência podem decair de frequência e mesmo desaparecer, e, além disso, lexias correntes podem assumir novos significados, dando origem a novos vocábulos, no subconjunto das lexias gramaticais do léxico reina uma grande estabilidade nesse aspecto, até porque os signos gramaticais possuem um caráter estritamente interoceptivo, isto é, indicam funções e relações que se estabelecem entre os próprios elementos da língua, ao passo que os signos lexicais são exteroceptivos, ou seja, remetem a fatos da realidade exterior à linguagem, realidade esta que está em constante mutação.

Assim, procedemos à coleta de todas as ocorrências de vocábulos lexicais dos textos do *corpus*, quais sejam:

- a) substantivos;
- b) adjetivos;
- c) verbos, à exceção dos verbos auxiliares, devido ao seu caráter eminentemente gramatical;
- d) advérbios derivados de adjetivos.

Os vocábulos assim coletados constituíram o banco de dados sobre o qual realizamos nosso estudo. Para organizar esse banco de dados, utilizamos processos informáticos, através do uso de microcomputador, e com o auxílio de um *software* gerenciador de banco de dados. Os vocábulos foram organizados nesse banco em listas por amostra, correspondendo a cada língua doze amostras, oito científicas e quatro jornalísticas (onze amostras no caso do português), o que totaliza 59 listas. Em cada lista, os vocábulos foram ordenados segundo a classe gramatical a que pertencem (substantivo, adjetivo, verbo ou advérbio), bem como alfabeticamente no interior de cada classe. Os vocábulos encontram-se, assim, prontos para a próxima etapa do trabalho, a classificação. Para realizar essa etapa, é necessário que a cada vocábulo corresponda uma ficha lexicológica. No presente caso, uma vez que nosso banco de dados é informatizado, não utilizamos fichas reais, confeccionadas em cartão, como tradicionalmente se costuma fazer, mas, ao invés disso, arquivamos na própria memória do computador todas as informações correspondentes a cada vocábulo. Pudemos, com isso, substituir um volumoso fichário por um simples disquete de microcomputador. Mesmo assim, por questões

práticas, continuaremos a chamar cada um dos registros de nosso arquivo de *ficha lexicológica*.

As fichas lexicológicas

As fichas lexicológicas que utilizamos em nosso trabalho são compostas dos seguintes campos:

VOCÁBULO: Este campo corresponde à transcrição do vocábulo coletado no texto. Nos casos em que encontramos no texto o mesmo significante com significados diferentes (diversas acepções de uma mesma lexia ou lexias homônimas), utilizamos os índices 1, 2, etc., após cada um dos vocábulos coincidentes. Os substantivos foram registrados no singular, exceto quando são pluralícios, isto é, não admitem o singular (por exemplo, port. *férias*). Os adjetivos foram registrados no masculino singular, em francês e italiano; em alemão, na forma-base, ou forma predicativa, isto é, sem desinências; em inglês, não existindo a flexão dos adjetivos, foram estes registrados na sua forma invariável. Os adjetivos pluralícios (por exemplo, port. *vários, diversos*, etc.) foram registrados no plural. Os verbos foram registrados no infinitivo. No caso de verbos defectivos, que não possuam o infinitivo (por exemplo, os verbos modais ingleses *can, may, must*, etc.), registramos a forma-base, que corresponde ao presente do indicativo. Os advérbios, por serem invariáveis, foram registrados em sua forma única.

SIGNIFICADO: Neste campo registramos uma definição sucinta do vocábulo ou um vocábulo equivalente em português, a fim de podermos identificar rapidamente de que acepção da lexia se trata.

CLASSE GRAMATICAL: Indica a classe gramatical a que pertence o vocábulo (Substantivo, Adjetivo, Verbo ou Advérbio). Tais classes são respectivamente indicadas pelas siglas SUB, ADJ, VER e ADV.

ÍNDICE LEXICOGÊNICO (IL): Trata-se do índice de classificação lexicogênica. Sua composição e sua função serão explicadas a seguir, em O índice lexicogênico, p. 163.

ÍNDICE ETIMOLÓGICO (IE): Trata-se do índice de classificação etimológica. Sua composição e função serão explicadas em O índice etimológico, p. 166.

NÚMERO DE OCORRÊNCIAS: Neste campo registramos o número de ocorrências efetivas do vocábulo em questão na amostra examinada.

OBSERVAÇÕES: Reservamos este campo para eventuais informações adicionais porventura necessárias, tais como a citação do trecho do texto em que o vocábulo figura, indicações quanto à etimologia ou à data de introdução do vocábulo na língua, eventuais dúvidas na classificação, etc.

A classificação dos vocábulos

Para o estabelecimento de um perfil lexicogênico do léxico de cada língua estudada, cumpre classificar cada vocábulo individualmente, e, a seguir, indexá-lo, isto é, atribuir-lhe um código alfanumérico, referente a essa classificação. Tal código, que funciona como uma etiqueta própria de cada vocábulo, é composto de dois índices, um índice lexicogênico e um índice etimológico. Através desses índices é possível realizar dois tipos de estatística, conforme veremos a seguir.

O índice lexicogênico (IL)

No Capítulo 2, apresentamos uma tipologia lexicogênica dos vocábulos de uma língua, que consta de uma série de categorias, correspondentes aos diversos processos de geração vocabular, ou processos lexicogênicos. O índice lexicogênico, ou simplesmente IL, como passaremos a designá-lo, baseia-se diretamente nessas categorias, e é através dele que podemos quantificar cada processo lexicogênico no interior do léxico de uma determinada língua. Este índice compõe-se de um algarismo (1 ou 2) e de uma letra. O algarismo 1 indica que o vocábulo em questão é um vulgarismo, enquanto o algarismo 2 representa os grecolatinismos. Dentro de cada uma dessas duas macrocategorias, as letras representam as

categorias lexicogênicas em que as mesmas se dividem. Teremos, então, a seguinte correspondência:

Macrocategoria 1 – Vulgarismos

1a – Vulgarismos resultantes de herança vernácula (ver p. 73)

1b – Vulgarismos alogenéticos (ver p. 77)

1c – Empréstimos de tradução de grecolatinismos (ver p. 90)

1d – Empréstimos de tradução de vulgarismos estrangeiros (ver p. 98)

1e – Vulgarismos resultantes de metamorfismo (ver p. 100)

1f – Vulgarismos sintagmáticos compostos de radicais vulgares ou derivados de radical vulgar (ver p. 110)

1g – Vulgarismos sintagmáticos compostos de radical greco-latino e vulgar ou derivados de radical greco-latino com afixo vulgar (ver p. 112)

1h – Vulgarismos sintagmáticos resultantes de composição ou derivação “anômala” de elementos greco-latinos (ver p. 113)

Macrocategoria 2 – Grecolatinismos

2a – Grecolatinismos resultantes de empréstimo direto ou indireto do grego ou latim (ver p. 116)

2b – Grecolatinismos resultantes de refecção de vulgarismos (ver p. 117)

2c – Empréstimos de restituição (ver p. 119)

2d – Grecolatinismos resultantes de composição ou derivação (ver p. 125)

Os vocábulos não classificáveis nem como vulgarismos nem como grecolatinismos, dos quais tratamos em Vocábulos não classificáveis nas categorias anteriores, p. 127 (nomes próprios e seus derivados, siglas, etc.), foram rotulados com a sigla NC (= não classificável). Desse modo, tais vocábulos, embora não classificados, fizeram parte de nosso *corpus* e foram computados para fins estatísticos.

De modo geral, os rótulos definidos acima para o IL coincidem com as categorias lexicogênicas apresentadas no Capítulo II. Todavia, quanto às categorias correspondentes aos empréstimos de tradução (1c e 1d, respectivamente), não fizemos distinção entre traduções sintagmáticas e semânticas, ou entre a tradução total e a tradução parcial. Tampouco levamos em consideração se

se trata de empréstimo de tradução direto ou indireto, obrigatório ou facultativo. Na verdade, tais distinções são de grande importância teórica para a própria compreensão do fenômeno da tradução, mas na prática revelam-se pouco produtivas. Com efeito, empréstimos de tradução totais ou parciais, sintagmáticos ou semânticos, diretos ou indiretos, configuram, todos, casos de tradução.²⁸ São todos, portanto, vulgarismos do mesmo tipo. A nosso ver, um vocábulo que possua todos os seus morfemas vulgares não é mais vulgar do que outro que possua apenas um. O que nossa estatística visa quantificar, no presente caso, é a porcentagem de vocábulos resultantes de empréstimo de tradução existentes no léxico de uma determinada língua, não importando o número de morfemas do vocábulo que tenham sido traduzidos, ou o modo como se tenha dado a tradução. Por outro lado, mantivemos a distinção entre o empréstimo de tradução de grecolatinismo e o empréstimo de tradução de vulgarismo estrangeiro. Isso porque, no primeiro caso, estamos diante da vulgarização de um vocábulo greco-latino, isto é, da transformação de um grecolatinismo em vulgarismo, o que, conforme já vimos anteriormente (cfr. Introdução), confere ao processo um forte cunho ideológico, ao passo que, no segundo caso, temos simplesmente a transformação de um vulgarismo em outro. Tal fato não aumenta nem diminui o caráter erudizante ou vulgarizante da língua em questão, não havendo, portanto, um traço ideológico tão marcante a presidir este último processo. Poderíamos mesmo dizer que a vernaculização de estrangeirismos — quer por adaptação fonológico-ortográfica, quer por tradução —, fruto da reação a tudo que é estranho ao sistema lingüístico, é algo natural em qualquer língua, enquanto que a vulgarização de grecolatinismos, por não ser a regra geral na maioria das línguas, é algo que, quando ocorre, revela aspectos interessantes da ideologia que preside a lexicogênese dessa língua. Por essa mesma razão, mantivemos a distinção lexicogênica entre os vulgarismos sintagmáticos derivados de vocábulos greco-latinos e os derivados de vocábulos vulgares.

²⁸ Na verdade, nem sempre é possível distinguir ao certo entre o empréstimo de tradução parcial e o total, como se pode ver pelo exemplo a seguir. Seja o it. *acquoso*, empréstimo do lat. *aquosus*. Sabe-se que o radical it. *acqu-*, extraído do vocábulo vernáculo *acqua*, traduz o radical lat. *aqu-*, mas é difícil dizer se o sufixo it. *-oso* é, no presente caso, erudito ou vernáculo.

Quanto aos metamorfismos, também optamos, por razões práticas, por agrupá-los numa única classe, já que as possíveis subdivisões dessa categoria (metamorfismo espontâneo, deliberado, obrigatório, facultativo) não são pertinentes do ponto de vista estatístico. Isso não impede, evidentemente, que possamos observar em cada língua a predominância deste ou daquele tipo de fenômeno, e que possamos inclusive interpretar tal observação. Apenas julgamos que a consideração de todos esses aspectos no exame de cada vocábulo acarretaria uma série de dificuldades técnicas, complicando desnecessariamente a análise, sem que tal esforço se revelasse produtivo ou vantajoso do ponto de vista estritamente científico.

O índice etimológico (IE)

Além de quantificar cada categoria lexicogênica, é interessante identificarmos a origem dos formantes, bem como observarmos a migração de elementos léxicos de uma macrocategoria para outra, uma vez que o aspecto central de nosso estudo é a constatação de que o léxico não é algo imóvel, mas está em constante movimento; do ponto de vista da tensão *erudito/vulgar*, vê-se que essas duas categorias não são estanques. Se o fossem, nosso estudo não teria maior interesse. Entretanto, a inexistência de uma fronteira rígida entre o que é erudito ou vulgar na língua, a possibilidade de converter unidades de uma categoria em outra e vice-versa, é o que vai, em última análise, determinar a ocorrência de uma escolha a cada criação vocabular, escolha essa revestida de um forte caráter ideológico, conforme já assinalamos. Desse modo, ao estudarmos a dinâmica desse constante movimento do léxico e das forças que, presidindo cada ato de criação ou transformação lexical, regem tal movimento, estamos na verdade detectando, ou pelo menos rastreando, a ideologia desse léxico e, conseqüentemente, a ideologia da língua, que nada mais é do que a ideologia da comunidade humana que a fala. A fim de podermos observar essa dinâmica, decidimos adotar um segundo tipo de controle, o índice etimológico²⁹, ou simplesmente IE, que consis-

²⁹ A rigor, a designação *índice etimológico* não é muito adequada, uma vez que o que procuramos detectar não é pura e simplesmente o étimo do vocábulo, mas sim o seu

te de siglas representando as línguas das quais procedem os formantes das palavras, na seguinte conformidade:

GL = formante ou vocábulo de origem grega, latina, ou a resultante da combinação de ambos;

PO = formante ou vocábulo de origem vernácula portuguesa;

FR = formante ou vocábulo de origem vernácula francesa;

IT = formante ou vocábulo de origem vernácula italiana;

IN = formante ou vocábulo de origem vernácula inglesa;

AL = formante ou vocábulo de origem vernácula alemã;

OU = formante ou vocábulo de outras origens.

Para indicarmos a passagem de um vocábulo de uma categoria a outra, ou a combinação sintagmática de elementos de diferentes procedências, utilizamos o sinal >. Alguns exemplos tornarão mais claro como tal sistema de classificação etimológica funciona. Para tanto, tomemos o exemplo da língua francesa. Os vocábulos vernáculos dessa língua receberão um IL = 1a e um IE = FR. Os vocábulos franceses resultantes de composição ou derivação em que o(s) radical(ais) seja(m) vernáculo(s) (IL = 1f) receberão igualmente um IE = FR. Um vocábulo francês que tenha sido tomado de empréstimo diretamente ao grego ou latim, ou indiretamente, através de outra língua em que tal grecolatinismo já existisse, e que não tenha sofrido nenhum processo de vulgarização, será classificado como 2a GL; um vocábulo resultante de composição ou derivação a partir de formantes exclusivamente greco-latinos receberá o rótulo 2d GL, quer tenha o mesmo sido engendrado em francês, quer em outra língua e posteriormente importado pelo francês (empréstimo indireto).

Suponhamos agora um vocábulo francês resultante do empréstimo de um vulgarismo italiano, por exemplo. O mesmo terá um IL = 1b, indicando que se trata de um estrangeirismo, e um IE = IT, indicando que provém do italiano. Se tal estrangeirismo passar depois a outra língua, por exemplo o inglês, será igualmen-

percurso lexicogênico, da língua de origem à língua de destino, relacionando esse percurso aos processos identificados pelo índice lexicogênico (IL). Ocorre que, para tanto, somos obrigados justamente a pesquisar a origem dos vocábulos. Por essa razão, e na falta de um termo melhor, optamos por conservar a designação *índice etimológico*, com a ressalva que acabamos de fazer.

te classificado nesta última língua como 1b IT, pois se trata de um empréstimo vulgar, feito ao francês, de um vocábulo italiano. Note-se que o IL pode sofrer alteração quando um vocábulo é transplantado de uma língua para outra — o vocábulo em questão é vernáculo em italiano (IL = 1a) e estrangeirismo em francês e inglês (IL = 1b) —, mas o IE permanece o mesmo em todas as línguas (no caso, IE = IT), pois a etimologia do vocábulo não muda. Assim, no caso dos alógenos que retornam à língua de origem (ver p. 80), ocorre algo curioso, como podemos observar tomando o já mencionado exemplo do fr. *sport*: esse vocábulo, cuja forma arcaica *desport* deriva do verbo vernáculo francês *desporter* (portanto, 1f FR), passou ao inglês (1b FR). Mais tarde, retornou ao francês na forma *sport* como empréstimo do inglês. Atualmente, esse vocábulo possui em francês o rótulo 1b FR, pois constitui-se num estrangeirismo provindo do inglês, cujo étimo é genuinamente francês!

Até o momento, apresentamos exemplos de vocábulos cujo estatuto etimológico não sofreu alteração: vernáculos que permanecem vernáculos, grecolatinismos que permanecem grecolatinismos. Há, contudo, muitos casos — e são os que mais nos interessam — de vocábulos lexicogenicamente híbridos. Dentre eles estão os chamados vocábulos semi-eruditos. De modo geral, resultam da combinação sintagmática de formantes de diferentes procedências, ou da substituição de um ou mais morfemas de um empréstimo estrangeiro por morfemas vernáculos. No caso dos semi-eruditos, podemos ter a combinação de um radical grecolatino com um radical ou afixo vulgar, o metamorfismo ou a tradução, total ou parcial, de um grecolatinismo, e, ainda, a refecção ou a restituição de um vulgarismo. Vejamos como o índice etimológico se comporta nesses casos.

Seja um vocábulo grego ou latino importado pelo francês na forma de um empréstimo de tradução: seu rótulo será 1c GL>FR. Tem-se aqui um vocábulo originalmente greco-latino (GL) que, por tradução, se tornou vulgar (FR). O sinal > indica essa transformação. Se tal vocábulo for emprestado ao inglês, por exemplo, terá nesta língua o rótulo 1b (i.e., empréstimo de vulgarismo francês) GL>FR, já que, como vimos, o IE não sofre alteração nesse caso. O mesmo vale para um metamorfismo francês, que teria o rótulo 1e GL>FR. Um empréstimo de tradução francês de um vulgarismo italiano, por exemplo, teria um rótulo 1d IT>FR.

Um composto de um radical latino e de outro francês ou um derivado de um radical latino com sufixo francês seriam rotulados como 1g GL>FR.

Por outro lado, um antigo vulgarismo francês que tivesse sido refeito sobre o modelo latino, fato este que foi muito comum na Renascença, teria um rótulo 2b FR>GL. Um empréstimo de restituição francês de um vulgarismo alemão seria rotulado como 2c AL>GL.

A sigla OU fica reservada aos formantes e vocábulos oriundos de outras línguas que não as cinco que constituem nosso objeto de estudo, como, por exemplo, o holandês, o provençal, o espanhol, o árabe, as línguas ameríndias, etc.

Como se vê, através do índice etimológico é possível fazer uma estimativa da porcentagem de vocábulos eruditos (grecolatinismos puros), da porcentagem de vocábulos vulgares (vulgarismos puros) e da porcentagem de vocábulos semi-eruditos ou híbridos da língua. Pode-se também traçar um espectro etimológico do léxico de cada língua, observando em que medida a mesma efetua empréstimos de outras línguas e de que línguas ela empresta com maior ou menor frequência.

PROCEDIMENTOS ESTATÍSTICOS

Tamanho e representatividade do *corpus*

Tamanho do corpus

Conforme vimos na p. 158, nosso *corpus* constitui-se de 59 amostras divididas em 12 amostras para cada língua (8 amostras de 250 palavras-ocorrência para o discurso científico e 4 amostras de 500 palavras-ocorrência para o discurso jornalístico), com exceção da língua portuguesa, que conta com 11 amostras (7 amostras de 250 palavras-ocorrência para o discurso científico e 4 amostras de 500 palavras-ocorrência para o discurso jornalístico). A partir do conjunto dessas amostras depreendemos os seguintes totais de vocábulos:

DISCURSO CIENTÍFICO	DISCURSO JORNALÍSTICO
Português: 1.173 vocábulos	Português: 875 vocábulos
Francês: 1.259 vocábulos	Francês: 972 vocábulos
Italiano: 1.290 vocábulos	Italiano: 986 vocábulos
Inglês: 1.284 vocábulos	Inglês: 987 vocábulos
Alemão: 1.306 vocábulos	Alemão: 967 vocábulos

Com isso, passamos a ter cerca de 1.300 vocábulos para cada língua no discurso científico, e cerca de 1.000 vocábulos para cada língua no discurso jornalístico. Como vemos, embora tenha sido coletado exatamente o mesmo número de palavras-ocorrência em ambos os universos de discurso (2.000), o número de vocábulos distintos efetivamente depreendidos foi relativamente menor no discurso jornalístico, o que indica que, em todas as línguas estudadas, esta norma discursiva é mais redundante em termos de vocabulário do que a norma científica. Isso parece inclusive ser de fato uma característica inerente a esse universo de discurso, o que equivale a dizer que, no discurso jornalístico, o mesmo vocábulo ocorre num mesmo texto com maior frequência do que costuma acontecer no discurso científico. Além disso, muitos vocábulos da norma jornalística figuram em diversos textos e em diversas amostras. Isso se dá, também, pelo fato de que, no discurso científico, cada amostra corresponde a uma diferente área do conhecimento, refletindo portanto um jargão e uma temática diferentes, enquanto no discurso jornalístico as amostras correspondem a diferentes jornais e revistas e não a diferentes discursos.

Em termos da riqueza vocabular de um texto, representada por R , a qual se define como a razão entre o número de vocábulos (V) e o número de palavras-ocorrência (N) desse texto, temos os seguintes valores:

	R=V/N (disc. cient.)	R=V/N (disc. jorn.)
Português	0,67	0,43
Francês	0,62	0,48
Italiano	0,64	0,49
Inglês	0,64	0,49
Alemão	0,65	0,48

Com base nas cifras acima, podemos notar que, no discurso jornalístico, o português se revelou ligeiramente mais redundante que as demais línguas.

A frequência média de ocorrência de vocábulos por língua e por universo de discurso é o inverso da riqueza vocabular, ou seja: $\bar{f} = N/V$. Temos então:

	$\bar{f} = N/V$ (disc. cient.)	$\bar{f} = N/V$ (disc. jorn.)
Português	1,49	2,32
Francês	1,61	2,08
Italiano	1,56	2,04
Inglês	1,56	2,04
Alemão	1,53	2,08

Representatividade do corpus

O tamanho do léxico está diretamente ligado ao grau de complexidade e de progresso atingido pela sociedade a cuja língua o mesmo pertence. Assim, as pesquisas no campo da etnolingüística demonstraram que praticamente todas as línguas ágrafas do mundo possuem um léxico cuja extensão gira em torno de 1.500 palavras. Isso parece indicar que o universo cultural de qualquer povo dito “primitivo”, que vive da caça, da pesca, da coleta de alimentos e de formas rudimentares de agricultura e pecuária, é composto de cerca de 1.500 conceitos, dentre os quais estão representações de animais, plantas, artefatos de uso diário, alimentos, partes do corpo humano, divindades, doenças, acidentes geográficos, etc. Numa sociedade desse tipo, não existem mais do que

cinco ou seis atividades diferentes, representadas por funções sociais tais como: caçador, guerreiro, feiticeiro, chefe de tribo, etc. Nessas sociedades, igualmente, verificam-se apenas dois universos de discurso distintos: o discurso sagrado, reservado às práticas rituais de caráter religioso e à medicina — que, nessas culturas, aliás, não se distingue da feitiçaria e, portanto, da religião — e o discurso profano, destinado às práticas do dia-a-dia. Além disso, pode-se dizer que todo o patrimônio de saber da comunidade, e conseqüentemente todo o seu léxico, é partilhado de forma igual por todos os seus membros.

Já o léxico do grego ou o do latim — únicas línguas de cultura da Antigüidade Clássica e da Alta Idade Média européia — contavam no século VI de nossa era cerca de 50.000 palavras. Esta era também a extensão aproximada do léxico das línguas de cultura européias no século XVI. É interessante notar que no mundo grego havia por volta de cinquenta profissões, e na Europa renascentista, cerca de uma centena delas. Na França do século XIX, Comte relatava a existência de cento e trinta profissões, ao passo que, hoje, levantamentos de nosso Ministério do Trabalho dão conta da existência de mais de trinta mil atividades profissionais. Essa assustadora diversificação da atividade profissional, reflexo da crescente divisão do trabalho provocada pelo progresso científico e tecnológico, que se acelerou sobretudo a partir de princípios do presente século, trouxe como conseqüência natural o crescimento do léxico das línguas de cultura moderna, com a proliferação de novas normas de discurso, muitas vezes privativas de um único grupo social ou profissional, constituindo assim o seu jargão específico. Desse modo, estima-se que a extensão do léxico de uma língua de cultura oscile atualmente entre seiscentas e setecentas mil palavras. É interessante notar, portanto, que essa hipertrofia do léxico, se é que assim podemos chamar tal fato, se dá exatamente em função da proliferação das novas técnicas, fenômeno típico das sociedades industriais e pós-industriais. Em tais sociedades, extremamente heterogêneas, o saber acumulado é cada vez maior, mas em contrapartida é compartilhado por cada vez menos indivíduos.

É óbvio, portanto, que qualquer falante de uma língua de cultura moderna, por maior que seja o seu grau de instrução, jamais vai dominar senão uma pequena parcela desse gigantesco léxico, e isso lhe basta para conviver socialmente de forma satisfa-

tória. Em primeiro lugar, é preciso lembrar que tal falante, enquanto ser social, participa de vários grupos sociais ao mesmo tempo, e, por conseguinte, possui competência lingüística em diversas normas discursivas. Tal falante pode ser considerado, desse modo, como lugar de convergência e de conflito dessas normas. Assim, todo falante domina um determinado conjunto de vocábulos, aplicáveis a diferentes universos e situações de discurso, que é o que poderíamos chamar de *vocabulário fundamental* da língua. Grande parte desse vocabulário fundamental constitui a base vocabular do chamado discurso coloquial ou banal. Como se vê, o vocabulário fundamental é o repertório vocabular que todos os falantes de uma língua dominam de forma mais ou menos equivalente e através do qual se intercomunicam.

Do que foi dito decorre que, qualquer que seja a norma discursiva que tomemos, aí encontraremos vocábulos privativos dessa norma ao lado de outros susceptíveis de ser encontrados em outras normas discursivas. O conjunto dos vocábulos pertencentes a todas as normas discursivas do léxico de uma língua constitui aquilo a que chamamos de *dianorma* desse léxico. A dianorma léxica define-se, pois, como o conjunto-intersecção dos conjuntos vocabulários das diversas normas de discurso. Graficamente, temos:

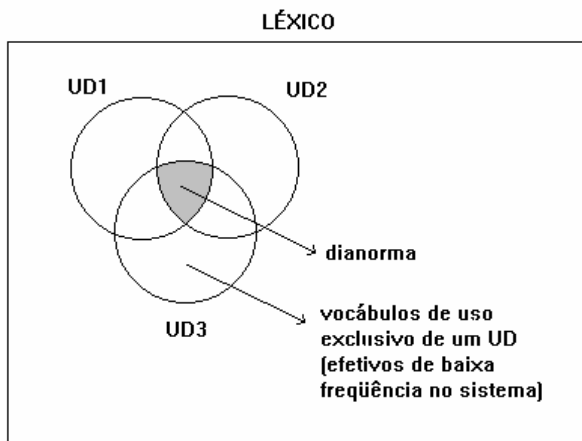


FIGURA 1 – A dianorma léxica

A dianorma léxica reúne, portanto, todos os vocábulos de alta frequência em seus respectivos universos de discurso e que, além disso, são comuns a todos esses universos de discurso. Podemos dizer, assim, que a dianorma léxica é o subconjunto do vocabulário fundamental da língua que contém os elementos léxicos efetivos de alta frequência. No limite, a dianorma pode coincidir com o próprio vocabulário fundamental.

Assim, dado um léxico de grandes dimensões, como o das línguas de cultura modernas, observamos que a maior parte das unidades léxicas a ele pertencentes são exclusivas de um único universo de discurso (UD) e, portanto, de baixa ou baixíssima frequência de uso fora desse universo de discurso (algumas unidades são de baixa frequência mesmo dentro de seu universo de discurso privativo). Paralelamente, qualquer falante, a qualquer momento, pode utilizar em discurso uma nova unidade lexical, que ele próprio engendrou, graças à produtividade de que o léxico de qualquer língua é dotado. Tais criações lexicais (neologismos) podem ser utilizadas novamente, em outras situações discursivas, terminando por ser incorporadas pela língua, ou ser simplesmente abandonadas. Dessa forma, muitos elementos do léxico possuem na verdade existência meramente virtual.

A conclusão a que se chega é que o acervo léxico comum a todos os membros de uma comunidade lingüística é bem menor que o próprio léxico dessa língua. Ora, estima-se que esse acervo comum (a dianorma léxica) varie nas línguas de cultura modernas entre 600 e 1.200 vocábulos. Assim, se compararmos nosso *corpus* de cerca de 2.200 vocábulos para cada língua estudada com a extensão total do léxico da língua em questão, concluiremos evidentemente que nossa amostra não é representativa, pois corresponde aproximadamente a apenas 0,3% do léxico. Entretanto, parece-nos mais sensato confrontar nossa amostra com a dianorma do léxico, em relação à qual a mesma é plenamente representativa. Por outro lado, cumpre também considerar o tamanho do *corpus* em relação ao número de categorias lexicogênicas a serem detectadas no mesmo, a saber, treze. Em face dessas considerações, entendemos que nosso *corpus* pode ser considerado representativo, em termos do rigor científico com que se pretende realizar a análise. De qualquer modo, é preciso também ter presente que a utilização de um *corpus* muito maior, embora fosse desejável, exigiria recursos materiais (apoio financeiro, programas de computador especiais, intercâmbio de informações em nível internacional, etc.) e humanos (grande número de pesquisadores, assessoria técnica especializada, etc.) que iriam muito além dos propósitos e dos limites deste trabalho.

Vale lembrar ainda que o aspecto mais importante desta pesquisa é o aspecto qualitativo, do qual a análise quantitativa é apenas um instrumento. De fato, um *corpus* maior sem dúvida refinaria os valores medidos, mas certamente não produziria um resultado muito diferente do efetivamente obtido. Assim, se o escopo de nosso trabalho é detectar as principais tendências de cada uma das línguas estudadas no que concerne à criação e à renovação lexicais, confrontando tais línguas para observar em que aspectos elas se aproximam ou se distanciam, parece-nos ocioso buscar uma precisão de medida que em nada acrescentaria à nossa análise, e que, de resto, é sempre um ideal inatingível quando se utilizam métodos estatísticos.

A estatística lexical

A estatística lingüística baseia-se em alguns princípios básicos, tais como a definição de uma variável aleatória (digamos, por exemplo, o número de vocábulos de um *corpus* que pertencem a uma determinada categoria) e a delimitação do intervalo numérico dentro do qual a mesma pode variar. Uma vez feito isso, e auferidos os valores efetivos registrados por essa variável, procede-se ao cálculo da média desses valores e do respectivo desvio-padrão. Admitindo-se que o comportamento dessa variável obedeça à lei de Laplace-Gauss, ou *lei normal*, o que normalmente acontece, pode-se estudar a amostra e fazer inferências a respeito da população da qual a mesma foi retirada, através da aplicação de tabelas e gráficos bastante conhecidos.

Para ilustrarmos como isso foi aplicado ao presente caso, tomemos como exemplo a frequência relativa de uso de vocábulos vernáculos (IL=1a) de um dos dois universos de discurso estudados em uma determinada língua. Uma vez que não nos é possível realizar essa medição diretamente sobre o léxico da língua, coletamos diversas amostras de igual tamanho, compostas de textos ou fragmentos de textos representativos do universo de discurso em questão, e realizamos a contagem de todas as ocorrências de vocábulos vernáculos de cada uma dessas amostras, obtendo, assim, o número absoluto dessas ocorrências. Definimos, então, uma variável aleatória x , a saber, o número de vocábulos vernáculos de uma amostra qualquer. Supondo que tenhamos coletado n amostras representativas e que tenhamos obtido um valor de x para cada amostra, observaremos que tais valores, embora possam diferir uns dos outros, concentram-se em torno de um determinado número, que é sua média, representada por \bar{x} . Temos, assim:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

onde x_i é o número de vocábulos vernáculos da amostra i , e n , como já dissemos, o número total de amostras examinadas. A média representa, assim, o valor mais provável da variável aleatória x .

Calculamos também a dispersão dos valores da variável, medida pelo desvio-padrão σ :

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}}$$

Através desse procedimento, poderemos também avaliar quanto os valores de x se desviam da média, o que, de certa forma, nos permite estimar a magnitude do erro existente entre nossa medida, realizada sobre uma amostra da população, e o valor que encontraríamos, se efetuássemos a pesquisa sobre a própria população.

Se nossa variável aleatória x obedece à lei normal, então seu comportamento pode ser descrito graficamente através da curva de Gauss, ou *curva em sino*:

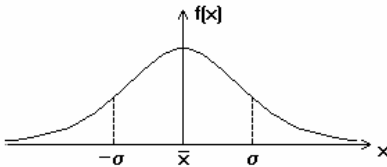


FIGURA 2 – Curva de Gauss

Nessa curva, temos no eixo das abscissas os valores de x , e no eixo das ordenadas a frequência relativa de ocorrência $f(x)$ associada a cada um dos valores possíveis de x . Dados dois valores quaisquer x_1 e x_2 da variável x , podemos conhecer a probabilidade de ocorrência de valores de x situados entre esses limites, integrando a função normal entre esses dois pontos:

$$p(x_1 \leq x \leq x_2) = \int_{x_1}^{x_2} a e^{-\sigma(x-\bar{x})^2} dx$$

Uma vez medido o valor de x para cada amostra, calculamos a média e o desvio-padrão respectivos, e estamos aptos a aplicar a lei normal para descrever o comportamento do *corpus* analisado e

realizar inferências estatísticas sobre a população da qual esse *corpus* foi retirado.

Entretanto, os resultados que se obtêm a partir da análise de um *corpus* somente podem ser transpostos para o conjunto da população se as amostras utilizadas respeitarem a hipótese nula, segundo a qual a distribuição de um fenômeno é absolutamente regular ao longo de toda a população, vale dizer, a probabilidade de obtenção de um determinado resultado é a mesma, qualquer que seja a amostra examinada. É evidente que o vocabulário de qualquer texto sofre influências de caráter estilístico ou temático que fazem que determinados vocábulos apresentem uma frequência sensivelmente diferente da que possuem no léxico da língua. Isso pode eventualmente fazer que determinadas categorias lexicogênicas sejam super ou subestimadas em certas amostras. Conseqüentemente, o desvio da variável x verificado será maior que o desvio teoricamente previsível. Alguns testes podem ser realizados para saber-se em que medida uma interferência estilística ou temática pode tornar nossas inferências inverossímeis ou pouco confiáveis. Dentre eles, está o teste de Pearson, também conhecido como teste do χ^2 , e a comparação entre o desvio-padrão teórico e o desvio-padrão experimental. Esses testes pressupõem o estabelecimento de um limiar de aceitação ou rejeição da hipótese nula. Cumpre ressaltar que a escolha desse limiar é arbitrária e se prende a critérios subjetivos, razão pela qual nunca nos encontraremos diante de uma escolha binária do tipo *aceitar/rejeitar* que implique 100% de certeza. Na verdade, a fronteira entre a aceitação e a rejeição da hipótese nula é fluida, e pode ser traduzida em termos de uma maior ou menor probabilidade de verificação dessa hipótese. Isso significa que nosso *corpus* sem dúvida alguma não retrata fielmente a realidade da língua, mas podemos estimar quanto ele se afasta dessa realidade, e, a partir daí, construir um modelo aproximativo do comportamento lexicogênico da língua. Em outras palavras, podemos dispor de amostras mais ou menos verossímeis, mas nunca de uma amostra totalmente verossímil ou totalmente inverossímil. Uma vez que nossa análise é sobretudo qualitativa, um perfil quantitativo aproximado do léxico é perfeitamente aceitável, desde que o erro aí implicado esteja dentro de limites que não comprometam a veracidade das conclusões da análise.

Como tivemos a oportunidade de ver até agora, nosso estudo tem por objeto a unidade lexical de discurso, isto é, o vocábulo. No entanto, a estatística lexical baseia-se na unidade de texto, a palavra-ocorrência. Desse modo, todos os procedimentos estatísticos até aqui descritos referem-se sempre a palavras e não a vocábulos. Não obstante, também realizamos uma segunda estatística, desta vez com base exclusivamente nos vocábulos. Nesta nova análise, cada vocábulo foi computado uma única vez, encontrando-se, pois, todos eles em pé de igualdade dentro do *corpus*, vale dizer, tendo todos o mesmo peso estatístico. Nosso objetivo agora é abstrair a freqüência relativa de cada vocábulo, procurando estimar apenas o número de vocábulos gerados por cada processo lexicogênico. Este novo perfil do léxico será confrontado com o obtido pelo método anterior, a fim de verificarmos se os vocábulos encontrados no *corpus* apresentam freqüências regulares, ou se existem processos de criação lexical cujos vocábulos, embora muito numerosos, são em sua maioria de baixa freqüência, e assim inversamente.

Esta segunda estatística lexical partiu da contagem do número de vocábulos de cada categoria lexicogênica (representada pelo IL) em cada um dos dois universos de discurso considerados. Como desta vez desejávamos que cada vocábulo fosse computado uma única vez em cada universo de discurso, não dividimos o *corpus* em várias amostras, mas, na verdade, tratamos cada uma das duas grandes seções do mesmo — científica e jornalística — como uma única amostra. Com isso, garantimos que um mesmo vocábulo figurasse em cada universo de discurso uma só vez.

Por tratar-se agora de apenas uma amostra por discurso, não é possível calcular média ou desvio-padrão experimental (o desvio-padrão teórico pode ser calculado, como veremos mais adiante), mas, por outro lado, pode-se demonstrar matematicamente que os valores assim obtidos coincidem com aqueles que obteríamos se dividíssemos o *corpus* em amostras. Senão vejamos.

Se dividíssemos, por exemplo, a seção científica do *corpus* em n amostras de tamanho N , e obtivéssemos o valor da variável x (que em nosso exemplo representa o número de vocábulos vernáculos) para cada amostra, o valor médio \bar{x} dessa variável seria dado por:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

onde x_i é o número de vocábulos vernáculos da amostra i .

A taxa percentual de freqüência de vocábulos vernáculos no discurso científico, que representaremos por t , seria então:

$$t = \frac{\bar{x}}{N}$$

Por outro lado, considerando toda a seção científica do *corpus* como uma única amostra, obteremos um valor de x dado por:

$$x = \sum_{i=1}^n x_i$$

Nossa taxa de freqüência t seria dada agora por:

$$t = \frac{x}{nN}$$

Substituindo a expressão (3) em (4), teremos:

$$t = \frac{\sum_{i=1}^n x_i}{nN} = \frac{\bar{x}}{N}$$

que é igual à expressão (2). Vê-se, assim, que o resultado é o mesmo por ambas as vias, porém este segundo procedimento é mais prático que o primeiro.

Podemos também estimar o desvio-padrão teórico associado ao valor x obtido, que será dado por:

$$\sigma = \sqrt{Npq}$$

onde p é a probabilidade de obtenção de um vocábulo vernáculo numa amostra qualquer, q é a probabilidade complementar, isto é, a probabilidade de obtenção de vocábulo pertencente a qualquer

outra categoria lexicogênica, e N o tamanho da amostra. Uma boa estimativa de p para o discurso científico é a razão entre o número x de vocábulos vernáculos encontrados nessa amostra e o número total de vocábulos da amostra, que é nN , donde se conclui que $p = t$. Como q é complementar de p , temos que $q = 1 - p \Rightarrow q = 1 - t$.

Nosso desvio-padrão pode agora ser reescrito como:

$$\sigma = \sqrt{Nt(1-t)}$$

A contagem do número de vocábulos de cada categoria lexicogênica se deu por classe gramatical, de modo a podermos saber, por exemplo, que percentual de substantivos, de adjetivos, de verbos e de advérbios são de origem vernácula, ou empréstimos vulgares, ou grecolatinismos, e assim por diante. A partir daí, foi possível saber também o percentual de cada categoria lexicogênica no total do *corpus*, bastando para isso somar os totais de vocábulos de cada uma dessas categorias por classe gramatical, e dividir essa soma pelo número total de vocábulos do *corpus*.

Procedemos também a um levantamento da natureza etimológica dos vocábulos, desta vez com base no IE. Através disso, pudemos perceber a procedência dos morfemas constituintes de cada vocábulo e, conseqüentemente, quantificar o volume de morfemas de cada procedência, bem como analisar o caminho trilhado por esses vocábulos e as transformações por eles sofridas ao longo de sua história.

Outro levantamento também efetuado por nós diz respeito às palavras de caracterização e ao vocabulário fundamental do *corpus*. Com efeito, alguns vocábulos ocorreram uma única vez, e em apenas uma das amostras analisadas. Tais vocábulos, chamados de palavras de caracterização, ou *hápax legómena* (isto é, “ditas uma única vez”), assumem especial importância, pois estão diretamente ligados à temática do texto em que figuram. Por outro lado, temos aqueles vocábulos que aparecem em todas as amostras e que constituem, portanto, o conjunto intersecção das mesmas. Trata-se do vocabulário fundamental (VF) dessas amostras. Para o levantamento desse VF, procedemos de duas maneiras: em primeiro lugar, depreendemos o VF resultante da intersecção das amostras científicas, e fizemos o mesmo em relação às amostras jornalísticas. Com isso, obtivemos dois conjuntos intersecção, referentes aos dois universos de discurso. A seguir, cruzamos esses dois

conjuntos, para obter um novo conjunto intersecção, isto é, o VF de todo o *corpus* (vide Figura 3). Nosso objetivo é averiguar quantos e quais vocábulos não são privativos de um só universo de discurso e traçar o perfil lexicogênico desses vocábulos.

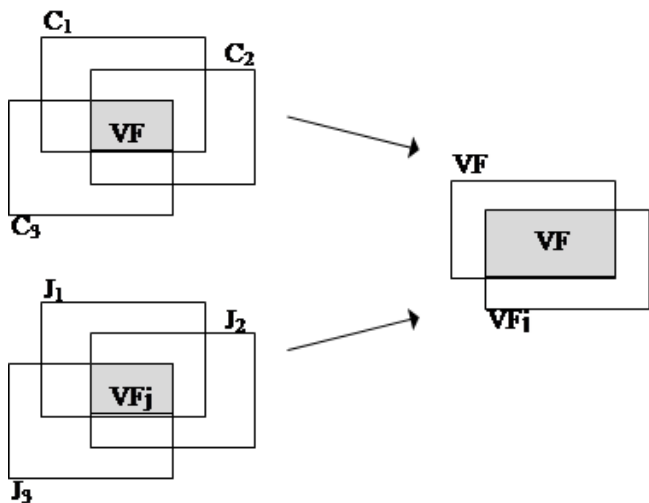


FIGURA 3 – Representação esquemática do procedimento para obtenção do vocabulário fundamental do discurso científico (VFc), do vocabulário fundamental do discurso jornalístico (VFj), e do vocabulário fundamental de todo o *corpus* (VF), onde C_1 , C_2 , C_3 representam amostras científicas e J_1 , J_2 , J_3 representam amostras jornalísticas.

CAPÍTULO IV

RESULTADOS E DISCUSSÃO

ANÁLISE ESTATÍSTICA COM BASE NAS PALAVRAS-OCORRÊNCIA

Comportamento estatístico do discurso científico

Foram construídas tabelas específicas para cada língua estudada, com base nas quais foram feitos os testes estatísticos de confiabilidade das amostras, como o teste de Pearson, para o cálculo do χ^2 , e a comparação entre o desvio-padrão teórico e o desvio-padrão experimental, dentre outros. Os testes apresentaram resultados satisfatórios, o que equivale a dizer que as amostras utilizadas são confiáveis, isto é, satisfazem a hipótese nula. Não reproduziremos aqui tais tabelas, ou os valores obtidos nos referidos testes, por entendermos que os mesmos não são de maior interesse. No entanto, o leitor que desejar aprofundar-se no exame dos dados estatísticos completos poderá recorrer diretamente à Tese de Doutorado que deu origem a este livro, na qual os mesmos são minuciosamente apresentados.

A partir desses dados, chegamos à tabela abaixo, da qual constam as taxas percentuais de frequência de cada categoria lexicogênica e seus respectivos desvios para cada uma das línguas estudadas.

TABELA 1 – Discurso científico

	PORTUGUÊS		FRANCÊS		ITALIANO		INGLÊS		ALEMÃO	
	Méd. %	σ s %	Méd. %	σ s %	Méd. %	σ s %	Méd. %	σ s %	Méd. %	σ s %
1a	18,4	1,3	23,8	1,6	27,6	3,1	16,4	1,4	30,2	1,4
1b	2,3	0,8	0,8	0,6	1,2	0,8	17,1	1,6	1,4	0,6
1c	14,0	1,2	11,4	0,7	12,9	1,4	1,0	0,6	19,7	1,8
1d	1,6	0,5	0,1	0,2	2,6	0,6	1,0	0,6	1,5	1,0
1e	8,6	1,3	4,6	1,2	15,1	1,5	1,0	0,6	0,8	0,4
1f	5,4	0,8	7,9	1,2	5,2	1,5	7,2	1,7	19,8	1,0
1g	4,3	0,7	6,4	1,6	2,6	0,8	2,8	0,9	5,4	0,9
1h	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
TOTAL DE VULG.	55,0	1,5	55,2	2,3	67,6	2,1	46,9	2,0	79,0	2,4
2a	33,9	1,4	33,8	3,0	25,4	1,4	36,4	1,5	15,8	1,3
2b	0,5	0,3	0,2	0,3	0,2	0,2	0,0	0,0	0,0	0,0
2c	1,4	1,0	0,0	0,0	0,4	0,2	2,0	0,8	0,4	0,6
2d	8,6	1,0	10,1	3,3	5,5	1,0	13,8	1,1	3,4	0,6
TOTAL DE GRECO- LAT.	44,6	1,8	44,3	2,1	31,7	2,0	52,4	1,8	19,8	1,1
NC	0,4	0,2	0,5	0,9	0,6	0,9	0,6	0,9	1,1	2,0

Comportamento estatístico do discurso jornalístico

Para o discurso jornalístico, repetimos os mesmos procedimentos estatísticos já explicados na p. 187, referentes ao discurso científico. Assim, nos limitaremos a seguir a apresentar a tabela a seguir, que compara o comportamento lexicogênico das cinco línguas estudadas no que se refere ao discurso jornalístico.

TABELA 2 – Discurso jornalístico

	PORTUGUÊS		FRANCÊS		ITALIANO		INGLÊS		ALEMÃO	
	Méd. %	σ s %	Méd. %	σ s %	Méd. %	σ s %	Méd. %	σ s %	Méd. %	σ s %
1a	29,6	0,4	35,0	0,7	34,0	1,1	26,6	0,5	49,4	1,2
1b	5,4	0,5	3,9	0,4	4,9	0,8	25,9	1,1	4,9	0,9
1c	11,3	0,5	7,3	0,2	13,6	0,9	0,8	0,2	12,4	1,0
1d	3,6	0,6	0,8	0,4	3,3	0,6	1,3	0,2	1,8	0,8
1e	6,2	0,4	3,2	0,6	9,1	1,0	0,5	0,4	0,4	0,3
1f	6,2	0,1	11,6	0,6	7,3	0,5	11,1	1,0	18,1	1,2
1g	1,7	0,2	4,2	0,4	2,5	0,6	2,9	0,4	3,1	0,3
1h	0,0	0,0	0,0	0,0	0,0	0,0	0,2	0,1	0,0	0,0
TOTAL DE VULG.	64,1	1,5	66,2	0,1	74,8	1,2	69,5	1,6	90,4	1,7
2a	29,4	1,4	26,5	0,7	19,4	1,4	22,9	1,6	7,1	1,0
2b	0,2	0,1	0,1	0,1	0,1	0,1	0,0	0,0	0,0	0,0
2c	0,9	0,6	0,0	0,0	0,5	0,4	1,5	0,3	0,0	0,0
2d	5,1	0,3	6,7	0,8	4,5	1,1	5,7	0,6	1,8	0,3
TOTAL DE GRECO-LAT.	35,7	1,3	33,4	0,8	24,7	0,8	30,3	1,2	9,0	0,7
NC	0,2	0,1	0,4	0,3	0,5	0,3	0,2	0,2	0,6	0,4

Análise comparativa do comportamento lexicogênico em ambos os discursos

Apresentamos abaixo a tabela resultante da fusão dos dados constantes das Tabelas 1 e 2. Através dela é que passaremos a partir de agora a analisar e a comparar o comportamento lexicogênico das cinco línguas estudadas nos dois universos de discurso considerados. Nessa tabela, os dados exprimem porcentagens. Houvemos por bem suprimir os respectivos desvios, por questão de simplicidade, e uma vez que os mesmos são quase todos inferiores a 1% e, quando não, raramente ultrapassaram 2%.

TABELA 3 – Comportamento lexicogênico em ambos os discursos

	PORTUGUÊS		FRANCÊS		ITALIANO		INGLÊS		ALEMÃO	
	CIENT.	JORN.	CIENT.	JORN.	CIENT.	JORN.	CIENT.	JORN.	CIENT.	JORN.
1a	18,4	29,6	23,8	35,0	27,6	34,0	16,4	26,6	30,2	49,4
1b	2,3	5,4	0,8	3,9	1,2	4,9	17,1	25,9	1,4	4,9
1c	14,0	11,3	11,4	7,3	12,9	13,6	1,0	0,8	19,7	12,4
1d	1,6	3,6	0,1	0,8	2,6	3,3	1,0	1,3	1,5	1,8
1e	8,6	6,2	4,6	3,2	15,1	9,1	1,0	0,5	0,8	0,4
1f	5,4	6,2	7,9	11,6	5,2	7,3	7,2	11,1	19,8	18,1
1g	4,3	1,7	6,4	4,2	2,6	2,5	2,8	2,9	5,4	3,1
1h	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,2	0,0	0,0
TOTAL DE VULG.	55,0	64,1	55,2	66,2	67,6	74,8	46,9	69,5	79,0	90,4
2a	33,9	29,4	33,8	26,5	25,4	19,4	36,4	22,9	15,8	7,1
2b	0,5	0,2	0,2	0,1	0,2	0,1	0,0	0,0	0,0	0,0
2c	1,4	0,9	0,0	0,0	0,4	0,5	2,0	1,5	0,4	0,0
2d	8,6	5,1	10,1	6,7	5,5	4,5	13,8	5,7	3,4	1,8
TOTAL DE GRECO-LAT.	44,6	35,7	44,3	33,4	31,7	24,7	52,4	30,3	19,8	9,0
NC	0,4	0,2	0,5	0,4	0,6	0,5	0,6	0,2	1,1	0,6

Diagnóstico da macrocategoria Vulgarismos

a) *Categoria 1a*: Observa-se que, de um modo geral, a taxa de vocábulos vernáculos gira em torno de 20% no discurso científico (DC) e em torno de 30% no discurso jornalístico (DJ) em todas as línguas estudadas, exceto em alemão, em que essas taxas foram sensivelmente superiores (respectivamente 30% e 50%).

b) *Categoria 1b*: Em todas as línguas, os estrangeirismos são mais frequentes no DJ. Nesta categoria, o italiano e o alemão apresentam exatamente o mesmo comportamento. Chama especial atenção a língua inglesa, cujas taxas de estrangeirismos (por volta de 20%) são em muito superiores às das demais línguas, de modo que esse processo lexicogênico parece ser, dentre os vulgarismos, o mais produtivo nessa língua. Por outro lado, é o francês a língua que menos importa vocábulos de outras línguas vulgares. Isso se deve obviamente ao fato de, ao longo da história medieval e moderna, ter sido sempre o francês a língua vulgar de cultura por excelência, tendo-se mantido, também, durante todo esse período, como a língua internacional da Civilização Européia. Por essa razão, o francês sempre exerceu sobre as outras línguas influência maior do que a que sofreu delas. Tal fato repercute nitidamente no léxico dessas línguas, como aqui podemos ver. Mais adiante, quando estudarmos o *corpus* do ponto de vista do Índice Etimológico (IE), estaremos aptos a esmiuçar melhor o fluxo dos emprés-

timos vulgares entre as línguas estudadas, e poderemos tratar qualitativamente tais influências.

c) *Categoria 1c*: As traduções de grecolatinismos são pouquíssimo freqüentes em inglês, correspondendo a aproximadamente 1% do léxico em ambos os discursos. Já no alemão científico, obtivemos a taxa mais alta de traduções dentre todas as línguas. No DC, a taxa 1c oscilou entre 11% e 14% nas línguas românicas, e no DJ ficou em torno de 12%, exceto em francês e inglês, cujas taxas foram inferiores.

d) *Categoria 1d*: Quanto às traduções de vulgarismos estrangeiros, cumpre notar que as taxas de freqüência se apresentam baixas em todas as línguas (menos de 4%), o que demonstra ser tal processo de um modo geral pouco produtivo. Todavia, cumpre notar que o francês é a língua que menos importa vocábulos vulgares estrangeiros, sejam eles diretos (cfr. acima a alínea b) ou traduzidos.

e) *Categoria 1e*: Observa-se, em todas as línguas, que o metamorfismo é mais utilizado no DC do que no DJ. É também mais freqüente nas línguas românicas do que nas germânicas. O italiano se destaca nesta categoria, apresentando em ambos os discursos taxas bastante superiores às das demais línguas. Dentre os processos vulgarizantes, o metamorfismo é, depois da herança vernácula, o mais produtivo nessa língua.

f) *Categoria 1f*: Os compostos e derivados vulgares apresentam freqüências semelhantes em português e italiano, de um lado, e em francês e inglês, de outro. Em alemão, tais vocábulos são sensivelmente mais freqüentes que nas outras línguas.

g) *Categoria 1g*: Os compostos e derivados semi-eruditos são pouco freqüentes em todas as línguas, apresentando taxas nunca superiores a 7%.

h) *Categoria 1h*: Os compostos e derivados anômalos de elementos eruditos somente ocorreram no inglês jornalístico, ainda assim a uma taxa baixíssima (0,2%). Isso nos permite inferir que tal processo lexicogênico é absolutamente excepcional e assistemático.

Diagnóstico da macrocategoria Grecolatinismos

a) *Categoria 2a*: No DC, a taxa de empréstimos de vocábulos greco-latinos girou em torno de 35% em português, francês e inglês. Em italiano e alemão, essa taxa foi mais baixa (cerca de 25% em italiano e cerca de 16% em alemão). A mesma tendência se verifica também em relação ao DJ, sendo que, nesse discurso, as taxas de 2a são um pouco inferiores às do DC em todas as línguas, o que mostra uma maior influência do acervo léxico greco-latino em universos de discurso mais cultos. No entanto, percebe-se que também universos de discurso mais populares, tais como o jornalístico, apresentam influência considerável das línguas clássicas. Na verdade, boa parte dos vocábulos dessa categoria entraram na norma jornalística e mesmo na coloquial provindas da norma científica, configurando o fenômeno da banalização vocabular, de que tratamos no Capítulo II.

b) *Categoria 2b*: As refecções de vocábulos vernáculos segundo o modelo latino só ocorreram nas línguas românicas, embora nas línguas germânicas também haja vocábulos vernáculos de origem latina. Vocábulos refeitos foram ligeiramente mais frequentes em português, mas em nenhum caso ultrapassaram 0,5%.

c) *Categoria 2c*: No DC, o empréstimo de restituição ocorreu com a mesma frequência no italiano e no alemão (0,4%). Em ambos os discursos, as taxas mais altas foram registradas no inglês. Em francês, não houve nenhum caso de empréstimo de restituição. Isso se deve ao fato de que, tendo sido durante muito tempo o francês, como já dissemos a propósito dos estrangeirismos (categoria 1b, acima), a língua de cultura européia por excelência, essa língua sempre emprestou muito mais do que tomou emprestado. Por essa razão, as taxas do francês são baixas em todas as categorias cujo vocábulo de origem é um vulgarismo estrangeiro (categorias 1b, 1d e 2c). No caso específico do empréstimo de restituição, é interessante notar que todos os vocábulos classificáveis nessa categoria encontrados em nosso *corpus* são restituições de vocábulos estrangeiros cujo radical é erudito e um ou mais sufixos são vulgares. Estudando-se os principais sufixos produtivos das línguas abordadas (ver Empréstimos de tradução de grecolatinismos, p. 90), verifica-se que são na maioria vulgares em francês e alemão e eruditos em inglês e italiano. Vê-se assim

que, sendo o francês uma língua essencialmente exportadora de vocábulos e sendo as demais línguas essencialmente importadoras, tem-se com mais freqüência a transformação de um vocábulo semi-erudito em erudito do que o inverso. Tendo o alemão a mesma característica do francês no que respeita aos sufixos, entendemos porque no DJ a taxa 2c do alemão também foi nula.

d) *Categoria 2d*: No DC, os compostos e derivados greco-latinos são menos freqüentes em italiano e alemão. Nesta última língua, tal categoria de vocábulos apresenta taxa inferior também no DJ. Na norma científica, a taxa mais alta pertence ao inglês, ao passo que na jornalística se destaca o francês.

Diagnóstico dos vocábulos não classificáveis nas macrocategorias anteriores

Pudemos observar que os vocábulos não classificáveis quer como grecolatinismos quer como vulgarismos (derivados de nomes próprios, siglas, fractocompostos, vocábulos onomatopáicos, etc.) são muito pouco comuns em todas as línguas, não representando mais do que 1% do léxico. São também ligeiramente mais freqüentes no DC que no DJ. Tal fato se deve, sem dúvida, à presença nos textos científicos de vocábulos derivados de nomes próprios, tais como *newtoniano*, por exemplo, de fractocompostos em química (por exemplo, *hidroximeleína*), de derivados de letras em matemática (por exemplo, *k-ésimo*), etc.

Análise geral do comportamento lexicogênico das línguas estudadas

Os dados referentes ao número total de vulgarismos e de grecolatinismos das cinco línguas estudadas são a seguir apresentados na forma de gráficos, referentes aos dois universos de discurso abordados, a partir dos quais é possível tirar conclusões a respeito do comportamento lexicogênico dessas línguas.

GRÁFICO 1 – Discurso científico

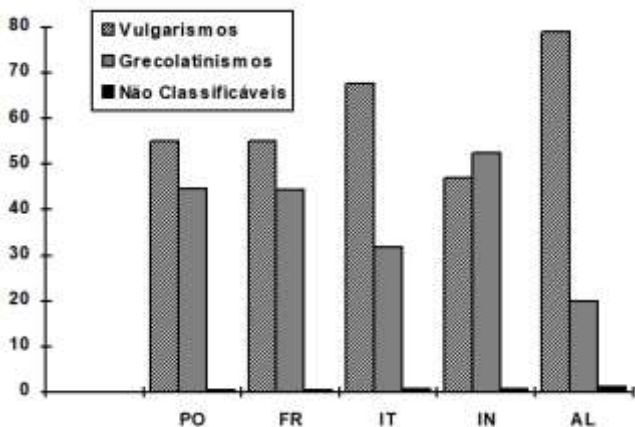
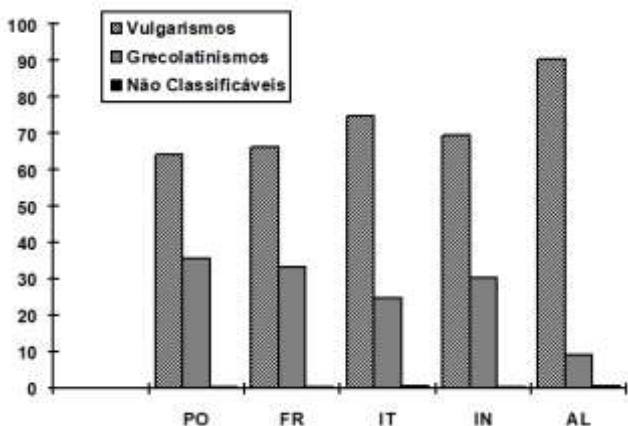


GRÁFICO 2 – Discurso jornalístico



Podemos observar, em primeiro lugar, que o francês e o português apresentam um comportamento semelhante em ambos os universos de discurso. No DC, o inglês é a única língua em que os grecolatinismos superam os vulgarismos. De um modo geral, nessas três línguas a taxa de vulgarismos no DC gira em torno de

50%. Nesse discurso, observa-se também o caráter vulgarizante mais forte do italiano e do alemão, com taxas de vulgarismos situadas nos patamares de 70% e 80%, respectivamente.

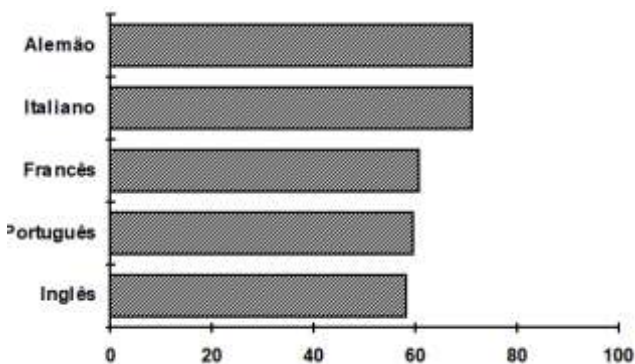
No DJ, as línguas que apresentam as maiores taxas de vulgaridade são o alemão (90% de vulgarismos contra 9% de grecolatinismos) e, a seguir, o italiano, com 75% de vulgarismos contra 25% de grecolatinismos. Nas demais línguas, as taxas situam-se na proporção de cerca de 66% de vulgarismos para cerca de 33% de grecolatinismos.

Calculando-se a taxa média de vulgarismos de cada língua, isto é, somando-se as taxas de vulgarismos dos dois universos de discurso e dividindo-se este resultado por 2, obtém-se um *Índice de Vulgaridade* (IV) de cada língua. Os valores obtidos foram:

LÍNGUA	IV
Português	59,6
Francês	60,7
Italiano	71,2
Inglês	58,2
Alemão	84,7

Plotando-se a seguir os índices assim obtidos num eixo orientado cujos valores vão de 0 a 100, é possível perceber o grau de vulgaridade/grecolatinidade de cada uma das línguas estudadas.

GRÁFICO 3 – Índice de Vulgaridade



Vê-se assim claramente que o italiano e o alemão se destacam como línguas em que a tendência à vulgarização é mais acentuada. O significado dessa constatação será discutido mais adiante, na p. 210.

ANÁLISE ESTATÍSTICA COM BASE NOS VOCÁBULOS

Apresentamos a seguir as tabelas contendo os dados referentes ao levantamento por vocábulos. Por questões práticas, os valores constantes das tabelas já estão convertidos em porcentagens. Além disso, os vocábulos aparecem distribuídos por classe gramatical, o que nos permitirá verificar eventuais variações no comportamento lexicogênico das línguas segundo a classe gramatical do vocábulo.

Comportamento estatístico do discurso científico

Todas as amostras científicas apresentaram aproximadamente 50% de substantivos, 30% de adjetivos, 16% de verbos e 4% de advérbios. Essa distribuição parece ser, portanto, uma característica normal desse universo de discurso. Outro fato a ser notado é a

inexistência em todas as línguas (e em ambos os universos de discurso) de verbos ou advérbios pertencentes à categoria NC.

Vejamos então como as diversas categorias lexicogênicas se distribuem por classes gramaticais.

Português científico

TABELA 4 – Português científico

	SUBSTANTIVO	ADJETIVO	VERBO	ADVÉRBIO	TOTAL
1a	6,7	3,3	4,2	0,4	14,6
1b	2,6	0,3	0,3	0,0	3,2
1c	11,2	1,1	1,5	0,3	14,1
1d	1,5	0,3	0,1	0,0	1,9
1e	3,8	3,0	1,1	0,5	8,4
1f	4,9	1,2	1,1	0,1	7,3
1g	3,7	1,3	0,8	0,0	5,8
1h	0,0	0,0	0,0	0,0	0,0
TOTAL DE VULG.	34,4	10,5	9,1	1,3	55,3
2a	10,3	11,6	6,6	2,0	30,5
2b	0,4	0,3	0,1	0,1	0,9
2c	0,1	0,8	0,1	0,4	1,4
2d	3,5	5,7	1,7	0,5	11,4
TOTAL DE GRECO-LAT.	14,3	18,4	8,5	3,0	44,2
NC	0,1	0,4	0,0	0,0	0,5
TOTAL	48,8	29,3	17,6	4,3	100,0

Na distribuição por vocábulos, o que nos chama particularmente a atenção na norma científica do português é o fato de a taxa de adjetivos greco-latinos ser superior à de substantivos, tanto em números absolutos quanto relativos. Também digno de nota é o fato de as traduções serem muito mais freqüentes nos substantivos do que nas demais classes gramaticais.

Francês científico

TABELA 5 – Francês científico

	SUBSTANTIVO	ADJETIVO	VERBO	ADVÉRBIO	TOTAL
1a	9,4	5,0	5,6	0,6	20,6
1b	0,9	0,0	0,2	0,0	1,1
1c	3,5	5,0	2,5	0,6	11,6
1d	0,1	0,1	0,0	0,0	0,2
1e	2,6	0,9	0,7	0,2	4,4
1f	6,8	1,4	2,0	0,1	10,3
1g	2,8	3,8	1,0	0,3	7,9
1h	0,0	0,0	0,0	0,0	0,0
TOTAL DE VULG.	26,1	16,2	12,0	1,8	56,1
2a	17,6	7,4	4,4	1,6	31,0
2b	0,0	0,1	0,1	0,1	0,3
2c	0,0	0,0	0,0	0,0	0,0
2d	4,7	4,8	2,1	0,2	11,8
TOTAL DE GRECO-LAT.	22,3	12,3	6,6	1,9	43,1
NC	0,2	0,6	0,0	0,0	0,8
TOTAL	48,6	29,1	18,6	3,7	100,0

Cumprer notará aqui que o francês científico não apresenta adjetivos ou advérbios de importação estrangeira. Tampouco apresenta verbos ou advérbios traduzidos de estrangeirismos ou não classificáveis.

Italiano científico

TABELA 6 – Italiano científico

	SUBSTANTIVO	ADJETIVO	VERBO	ADVÉRBIO	TOTAL
1a	8,1	5,0	6,0	0,8	19,9
1b	1,1	0,3	0,2	0,0	1,6
1c	5,1	2,3	3,6	0,2	11,2
1d	1,6	1,2	0,4	0,1	3,3
1e	6,8	3,6	1,3	0,9	12,6
1f	3,3	1,9	1,2	0,0	6,4
1g	1,5	1,2	0,8	0,0	3,5
1h	0,0	0,0	0,0	0,0	0,0
TOTAL DE VULG.	27,5	15,5	13,5	2,0	58,5
2a	14,2	12,0	2,5	2,3	31,0
2b	0,2	0,2	0,0	0,1	0,5
2c	0,1	0,4	0,1	0,0	0,6
2d	3,3	4,1	1,0	0,1	8,5
TOTAL DE GRECO-LAT.	17,8	16,7	3,6	2,5	40,6
NC	0,4	0,5	0,0	0,0	0,9
TOTAL	45,7	32,7	17,1	4,5	100

De um modo geral, a distribuição das categorias lexicogênicas por classe gramatical não apresenta nenhum desvio significativo em relação à distribuição na amostra como um todo.

Inglês científico

TABELA 7 – Inglês científico

	SUBSTANTIVO	ADJETIVO	VERBO	ADVÉRPIO	TOTAL
1a	4,4	3,8	4,4	1,1	13,7
1b	9,2	2,9	4,5	0,6	17,2
1c	0,7	0,3	0,2	0,0	1,2
1d	0,5	0,2	0,3	0,0	1,0
1e	0,5	0,2	0,1	0,0	0,8
1f	3,7	3,1	1,6	0,4	8,8
1g	1,3	1,7	0,0	0,4	3,4
1h	0,0	0,0	0,0	0,0	0,0
TOTAL DE VULG.	20,3	12,2	11,1	2,5	46,1
2a	18,5	8,2	4,4	2,3	33,4
2b	0,0	0,0	0,0	0,0	0,0
2c	0,2	0,9	0,2	0,4	1,7
2d	7,0	5,9	4,4	0,5	17,8
TOTAL DE GRECO-LAT.	25,7	15,0	9,0	3,2	52,9
NC	0,6	0,4	0,0	0,0	1,0
TOTAL	46,6	27,6	20,1	5,7	100,0

Em algumas categorias (1a, 2c), o número proporcional de verbos foi superior ao de substantivos.

Alemão científico

TABELA 8 – Alemão científico

	SUBSTANTIVO	ADJETIVO	VERBO	ADVÉRPIO	TOTAL
1a	9,1	5,2	5,9	1,6	21,8
1b	1,1	0,7	0,2	0,0	2,0
1c	7,8	6,8	2,7	2,1	19,4
1d	1,6	0,2	0,3	0,0	2,1
1e	0,6	0,2	0,1	0,0	0,9
1f	10,2	5,4	5,3	1,4	22,3
1g	2,0	3,4	2,2	0,2	7,8
1h	0,0	0,0	0,0	0,0	0,0
TOTAL DE VULG.	32,4	21,9	16,7	5,3	76,3
2a	12,0	3,0	0,0	1,1	16,1
2b	0,0	0,0	0,0	0,0	0,0
2c	0,2	0,2	0,0	0,2	0,6
2d	4,1	1,0	0,0	0,1	5,2
TOTAL DE GRECO-LAT.	16,3	4,2	0,0	1,4	21,9
NC	0,5	1,3	0,0	0,0	1,8
TOTAL	49,2	27,4	16,7	6,7	100

Há um número proporcionalmente maior de verbos do que de adjetivos nas categorias 1a, 1d e 1f. Note-se ainda que a categoria 2a é muito mais freqüente na classe dos substantivos do que nas demais.

Outro fato que merece destaque é a inexistência de verbos eruditos na amostra estudada. Isso se explica pela razão de que

praticamente todo verbo alemão de radical greco-latino sofre o acréscimo do sufixo vulgar *-ieren*. Dentre as pouquíssimas exceções a essa regra podemos citar os verbos *formen* e *kurven*.

Comportamento estatístico do discurso jornalístico

Das amostras jornalísticas, a portuguesa, a francesa e a alemã apresentaram aproximadamente as mesmas proporções de vocábulos por classe gramatical, a saber: cerca de 55% de substantivos, 20% de adjetivos, 23% de verbos e 2% de advérbios. Já o italiano apresentou proporcionalmente menos substantivos (cerca de 52%) e mais adjetivos (cerca de 26%), ao passo que no inglês o número de substantivos cai para cerca de 48% e o de verbos chega a 27%. Essas diferenças de comportamento podem estar ligadas a fatores estilísticos/temáticos inerentes aos textos utilizados no *corpus* que por razões aleatórias sobressaíram em duas das cinco línguas, ou podem ser indicadoras de uma característica estilística de cunho normativo do discurso jornalístico dessas línguas. Essa tese pode ser reforçada pela observação do comportamento da classe dos advérbios, na qual as taxas muito baixas fazem ressaltar desvios relativamente mais significativos. Nessa classe, podemos verificar que o inglês jornalístico se utiliza de advérbios cerca de quatro vezes mais do que o português.

Português jornalístico

TABELA 9 – Português jornalístico

	SUBSTANTIVO	ADJETIVO	VERBO	ADVÉRBIO	TOTAL
1a	12,3	4,7	9,0	0,3	26,3
1b	5,1	0,2	0,5	0,0	5,8
1c	8,4	1,1	1,7	0,0	11,2
1d	2,7	0,5	0,7	0,0	3,9
1e	3,7	0,9	1,1	0,1	5,8
1f	5,4	1,0	0,8	0,0	7,2
1g	2,5	0,5	0,7	0,0	3,7
1h	0,0	0,0	0,0	0,0	0,0
TOTAL DE VULG.	40,1	8,9	14,5	0,4	63,9
2a	13,1	8,0	7,0	0,5	28,6
2b	0,1	0,1	0,0	0,0	0,2
2c	0,3	0,7	0,1	0,0	1,1
2d	3,0	1,7	1,3	0,0	6,0
TOTAL DE GRECO-LAT.	16,5	10,5	8,4	0,5	35,9
NC	0,1	0,1	0,0	0,0	0,2
TOTAL	56,7	19,5	22,9	0,9	100,0

Nota-se, com base no exame dos dados acima, que as categorias 1c, 1f e 1g afetam sensivelmente mais os substantivos do que as demais classes gramaticais. As refeições (2b) e os processos não classificáveis (NC) só ocorreram em nomes (substantivos e adjetivos).

Francês jornalístico

TABELA 10 – Francês jornalístico

	SUBSTANTIVO	ADJETIVO	VERBO	ADVÉRBO	TOTAL
1a	18,2	5,9	10,7	0,4	35,2
1b	3,9	0,6	0,3	0,0	4,8
1c	4,1	1,3	1,9	0,4	7,7
1d	0,5	0,3	0,2	0,0	1,0
1e	2,2	0,4	0,4	0,2	3,2
1f	8,7	3,0	3,3	0,3	15,3
1g	2,0	2,3	0,5	0,2	5,0
1h	0,0	0,0	0,0	0,0	0,0
TOTAL DE VULG.	39,6	13,8	17,3	1,5	72,2
2a	12,1	4,4	4,0	0,7	21,2
2b	0,0	0,0	0,0	0,1	0,1
2c	0,0	0,0	0,0	0,0	0,0
2d	2,8	2,0	1,1	0,0	5,9
TOTAL DE GRECO-LAT.	14,9	6,4	5,1	0,8	27,2
NC	0,3	0,2	0,1	0,0	0,6
TOTAL	54,8	20,4	22,5	2,3	100,0

Entre os grecolatinismos, é de notar que a categoria das refeições (2b) é praticamente inexistente em francês jornalístico. Curiosamente, o único vocábulo resultante de refeição encontrado foi um advérbio.

Italiano jornalístico

TABELA 11 – Italiano jornalístico

	SUBSTANTIVO	ADJETIVO	VERBO	ADVÉRBO	TOTAL
1a	14,0	6,4	8,2	0,2	28,8
1b	3,9	0,5	0,4	0,1	4,9
1c	5,6	2,6	3,4	0,1	11,7
1d	2,2	0,8	0,3	0,1	3,4
1e	4,5	2,3	0,9	0,4	8,1
1f	5,4	1,9	1,9	0,1	9,3
1g	0,7	1,1	0,7	0,1	2,6
1h	0,0	0,0	0,0	0,0	0,0
TOTAL DE VULG.	36,3	15,6	15,8	1,1	68,8
2a	12,6	7,6	2,6	0,7	23,5
2b	0,1	0,0	0,0	0,0	0,1
2c	0,3	0,3	0,1	0,0	0,7
2d	3,1	2,0	0,8	0,2	6,1
TOTAL DE GRECO-LAT.	16,1	9,9	3,5	0,9	30,4
NC	0,3	0,5	0,0	0,0	0,8
TOTAL	52,7	26,0	19,3	2,0	100,0

Digno de nota é o fato de que o italiano jornalístico importa proporcionalmente mais substantivos estrangeiros do que outras classes de vocábulos.

Inglês jornalístico

TABELA 12 – Inglês jornalístico

	SUBSTANTIVO	ADJETIVO	VERBO	ADVÉRBO	TOTAL
1a	9,1	4,4	8,6	0,9	23,0
1b	13,7	3,2	6,7	0,6	24,2
1c	0,4	0,6	0,2	0,0	1,2
1d	0,9	0,2	0,3	0,0	1,4
1e	0,3	0,1	0,1	0,1	0,6
1f	5,9	4,6	4,0	0,5	15,0
1g	1,8	1,8	0,1	0,2	3,9
1h	0,1	0,0	0,1	0,0	0,2
TOTAL DE VULG.	32,2	14,9	20,1	2,3	69,5
2a	13,5	3,2	4,4	1,4	22,5
2b	0,0	0,0	0,0	0,0	0,0
2c	0,3	0,7	0,0	0,2	1,2
2d	1,6	2,1	2,7	0,2	6,6
TOTAL DE GRECO-LAT.	15,4	6,0	7,1	1,8	30,3
NC	0,2	0,0	0,0	0,0	0,2
TOTAL	47,8	20,9	27,2	4,1	100,0

A exemplo do que ocorre em italiano, também em inglês os estrangeirismos são na sua maioria substantivos. Os empréstimos de tradução de grecolatinismos são mais freqüentes entre os adjetivos. Os dois vocábulos não classificáveis encontrados são substantivos.

Alemão jornalístico

TABELA 13 – Alemão jornalístico

	SUBSTANTIVO	ADJETIVO	VERBO	ADVÉRBO	TOTAL
1a	20,9	6,6	11,6	1,8	40,9
1b	4,2	0,3	0,2	0,0	4,7
1c	7,1	2,2	2,6	0,9	12,8
1d	1,8	0,4	0,4	0,0	2,6
1e	0,5	0,0	0,0	0,0	0,5
1f	12,1	5,9	6,3	0,9	25,2
1g	1,8	1,2	0,6	0,1	3,7
1h	0,0	0,0	0,0	0,0	0,0
TOTAL DE VULG.	48,4	16,6	21,7	3,7	90,4
2a	6,2	0,8	0,0	0,1	7,1
2b	0,0	0,0	0,0	0,0	0,0
2c	0,0	0,0	0,0	0,0	0,0
2d	1,2	0,6	0,0	0,0	1,8
TOTAL DE GRECO-LAT.	7,4	1,4	0,0	0,1	8,9
NC	0,3	0,4	0,0	0,0	0,7
TOTAL	56,1	18,4	21,7	3,8	100,0

Os metamorfismos pertencem todos à classe dos substantivos. Conforme já observamos antes, verbos eruditos em alemão são raríssimos, não tendo ocorrido nenhuma vez em nossa amostra. Também merece destaque o fato de que, dentre os grecolatinismos, somente os processos 2a (empréstimo de vocábulo grecolatino) e 2d (composição ou derivação com morfemas grecolatinos) são produtivos em alemão.

Comparação entre os perfis lexicogênicos em ambos os discursos com base em palavras-ocorrência e em vocábulos

TABELA 14 – Discurso científico

	PORTUGUÊS		FRANCÊS		ITALIANO		INGLÊS		ALEMÃO	
	PALS.	VOCS.	PALS.	VOCS.	PALS.	VOCS.	PALS.	VOCS.	PALS.	VOCS.
1a	18,4	14,6	23,8	20,6	27,6	19,9	16,4	13,7	30,2	21,8
1b	2,3	3,2	0,8	1,1	1,2	1,6	17,1	17,2	1,4	2,0
1c	14,0	14,1	11,4	11,6	12,9	11,2	1,0	1,2	19,7	19,4
1d	1,6	1,9	0,1	0,2	2,6	3,3	1,0	1,0	1,5	2,1
1e	8,6	8,4	4,6	4,4	15,1	12,6	1,0	0,8	0,8	0,9
1f	5,4	7,3	7,9	10,3	5,2	6,4	7,2	8,8	19,8	22,3
1g	4,3	5,8	6,4	7,9	2,6	3,5	2,8	3,4	5,4	7,8
1h	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
TOTAL DE VULG.	55,0	55,3	55,2	56,1	67,6	58,5	46,9	46,1	79,0	76,3
2a	33,9	30,5	33,8	31,0	25,4	31,0	36,4	33,4	15,8	16,1
2b	0,5	0,9	0,2	0,3	0,2	0,5	0,0	0,0	0,0	0,0
2c	1,4	1,4	0,0	0,0	0,4	0,6	2,0	1,7	0,4	0,6
2d	8,6	11,4	10,1	11,8	5,5	8,5	13,8	17,8	3,4	5,2
TOTAL DE GRECO-LAT.	44,6	44,2	44,3	43,1	31,7	40,6	52,4	52,9	19,8	21,9
NC	0,4	0,5	0,5	0,8	0,6	0,9	0,6	1,0	1,1	1,8

Observando-se os números acima, constata-se que em todas as línguas, à exceção do italiano, as taxas de frequência das categorias lexicogênicas por palavra-ocorrência e por vocábulo são praticamente as mesmas, o que demonstra que, nessas línguas, a frequência com que cada categoria ocorreu nos textos utilizados no *corpus* é semelhante à frequência que apresenta na língua.

Já em italiano, nota-se claramente que alguns processos vulgarizantes (1a e 1e, por exemplo) são mais frequentes em nível de fala do que em nível de língua. Igualmente, alguns processos latinizantes (2a e 2d) são menos frequentes na fala do que o são na língua. Vê-se assim que, em italiano, embora a proporção de grecolatinismos e de vulgarismos entre os vocábulos seja semelhante à do francês, inglês e português, no plano das palavras-ocorrência o italiano se mostra notavelmente mais vulgar que essas línguas. Isso significa que, de modo geral, a maior parte dos vocábulos vulgares italianos são de alta frequência, ao passo que grande parte dos grecolatinismos são vocábulos de média e baixa frequência. Uma frequência de palavras superior à respectiva frequência de vocábulos parece ocorrer também em alemão no tocante à categoria 1a (vocábulos vernáculos), mas no entanto tal fato não chega a repercutir de forma sensível no total de vulgarismos da língua, pois o incremento dessa categoria parece ser contrabalançado por pequenos decrementos em outras categorias vulgares.

TABELA 15 – Discurso jornalístico

	PORTUGUÊS		FRANCÊS		ITALIANO		INGLÊS		ALEMÃO	
	PALS.	VOCS.	PALS.	VOCS.	PALS.	VOCS.	PALS.	VOCS.	PALS.	VOCS.
1a	29,6	26,3	35,0	35,2	34,0	28,8	26,6	23,0	49,4	40,9
1b	5,4	5,8	3,9	4,8	4,9	4,9	25,9	24,2	4,9	4,7
1c	11,3	11,2	7,3	7,7	13,6	11,7	0,8	1,2	12,4	12,8
1d	3,6	3,9	0,8	1,0	3,3	3,4	1,3	1,4	1,8	2,6
1e	6,2	5,8	3,2	3,2	9,1	8,1	0,5	0,6	0,4	0,5
1f	6,2	7,2	11,6	15,3	7,3	9,3	11,1	15,0	18,1	25,2
1g	1,7	3,7	4,2	5,0	2,5	2,6	2,9	3,9	3,1	3,7
1h	0,0	0,0	0,0	0,0	0,0	0,0	0,2	0,2	0,0	0,0
TOTAL DE VULG.	64,1	63,9	66,2	72,2	74,8	68,8	69,5	69,5	90,4	90,4
2a	29,4	28,6	26,5	21,2	19,4	23,5	22,9	22,5	7,1	7,1
2b	0,2	0,2	0,1	0,1	0,1	0,1	0,0	0,0	0,0	0,0
2c	0,9	1,1	0,0	0,0	0,5	0,7	1,5	1,2	0,0	0,0
2d	5,1	6,0	6,7	5,9	4,5	6,1	5,7	6,6	1,8	1,8
TOTAL DE GRECO-LAT.	35,7	35,9	33,4	27,2	24,7	30,4	30,3	30,3	9,0	8,9
NC	0,2	0,2	0,4	0,6	0,5	0,8	0,2	0,2	0,6	0,7

Observa-se, a partir dos dados acima, que em italiano e em alemão a frequência de palavras vernáculas é bem superior à dos respectivos vocábulos, o que revela ser tal categoria composta predominantemente de vocábulos de alta frequência nessas línguas.

Quanto à categoria 1f (compostos e derivados vulgares), percebe-se que em todas as línguas a porcentagem de vocábulos é superior à de ocorrências, demonstrando que tal categoria é formada principalmente de vocábulos de baixa frequência. Nesse aspecto, é notável a semelhança entre a configuração do francês e a do inglês.

No que se refere à categoria 2a (empréstimos greco-latinos), o francês e o italiano apresentam comportamentos opostos: em francês, a taxa de palavras dessa categoria é superior à dos vocábulos correspondentes, ao passo que em italiano se dá o inverso. Vê-se assim que, embora os vocábulos eruditos (pertencentes, na sua maioria, à categoria 2a) apresentem taxas bastante próximas em ambas as línguas, são empregados com mais frequência em francês do que em italiano.

Natureza e fluxo dos formantes – o Índice Etimológico

Analisemos agora o comportamento do léxico das línguas estudadas do ponto de vista da natureza etimológica dos morfemas formantes (radicais e afixos) utilizados na formação de seus vocábulos. Através do Índice Etimológico (IE), é possível não só quantificar os elementos de cada procedência, mas também, de certa forma, mapear seu fluxo, já que tal índice permite identificar as transformações morfológicas sofridas pelos vocábulos no curso de sua evolução, tais como vulgarizações fonológicas (metamorfismos) e substituições de certos morfemas por outros (traduções, restituições, refecções). Destarte, podemos descrever a composição morfo-etimológica do léxico dessas línguas.

No quadro a seguir, apresentamos os valores, já convertidos em porcentagens, das taxas de frequência de cada um dos diferentes tipos de IE encontrados no *corpus*.

PTC: português científico
 PTJ: português jornalístico
 PTC: francês científico
 PTJ: francês jornalístico
 PTC: italiano científico

PTJ: italiano jornalístico
 PTC: inglês científico
 PTJ: inglês jornalístico
 PTC: alemão científico
 PTJ: alemão jornalístico

A partir dos valores exibidos na tabela acima, é possível averiguar, por exemplo, que porcentagem de vocábulos de uma determinada língua possuem origem greco-latina (Índices Etimológicos GL, GL>PO, GL>FR, GL>IT, etc.), quantos são eruditos (GL, PO>GL, etc.), quantos são semi-eruditos (PO>GL, FR>GL, etc.), quantos provêm do francês (FR, GL>FR, FR>IN, etc.), e assim por diante. Esses valores secundários foram organizados na tabela abaixo.

TABELA 17 – IE: valores secundários

NATUREZA ETIMOLÓGICA	PTC	PTJ	FRC	FRJ	ITC	ITJ	INC	INJ	ALC	ALJ
vocábulos eruditos	44,8	36,7	43,2	27,4	41,0	30,4	53,5	31,3	21,9	9,2
vocábulos semi-eruditos	28,9	21,9	24,1	15,9	29,3	24,6	10,3	10,0	29,6	17,7
vocábulos de origem greco-latina	72,8	58,4	67,2	43,2	69,9	54,9	63,7	41,2	51,4	26,9
relatinizações	0,9	0,2	0,1	0,1	0,4	0,1	0,1	0,1	0,1	—
vocábulos de origem portuguesa	21,5	31,8	—	—	—	—	—	—	—	—
vocábulos de origem francesa	2,0	4,2	30,6	48,8	1,7	3,9	15,2	22,7	1,7	4,6
vocábulos de origem italiana	0,7	2,0	0,7	2,1	26,0	37,3	0,8	0,8	0,6	0,8
vocábulos de origem inglesa	0,6	0,6	0,2	0,6	0,5	1,0	18,3	31,9	0,3	0,4
vocábulos de origem alemã	0,3	0,1	—	0,1	0,2	0,2	—	0,3	44,1	64,8
vocábulos de outras origens	0,9	2,4	0,5	4,3	0,5	1,6	0,8	2,8	0,1	1,7
vocábulos vulgares mistos	1,3	4,1	0,3	2,1	1,5	1,9	3,1	3,6	1,3	4,0

Com base nos dados contidos nas duas tabelas acima, verifica-se que a taxa de vocábulos tomados de empréstimo ao grego/latim ou formados exclusivamente de morfemas greco-latinos (IE=GL) é sempre maior no discurso científico (DC). De todas as línguas estudadas, o inglês é a que apresenta as mais altas taxas de vocábulos deste tipo no DC, ao passo que no discurso jornalístico

(DJ) a maior incidência desses vocábulos se deu no português. O alemão foi, em ambos os discursos, a língua a apresentar as menores taxas.

Dentre os vocábulos semi-eruditos, observa-se que a taxa do inglês no DC é muito inferior à das demais línguas. O francês também apresenta uma taxa um pouco mais baixa em comparação ao português, italiano e alemão, o que parece indicar que essas duas línguas de fato vulgarizam menos os empréstimos greco-latinos que realizam.

No DJ, observa-se aproximadamente a mesma tendência, com a diferença de que agora o alemão também exibe uma taxa algo inferior, comparável à do francês.

Em ambos os discursos, a única língua a fornecer vocábulos semi-eruditos a todas as demais é o francês. A língua que mais se beneficia de tais empréstimos é o inglês, cujos índices ultrapassam 5%, enquanto nas outras línguas se encontram índices nunca superiores a 1%. Os demais vocábulos semi-eruditos restringem-se praticamente às suas próprias línguas de origem.

No DC, os vocábulos resultantes de empréstimos de origem greco-latina, sejam eles eruditos ou semi-eruditos, giram em torno de 70% nas línguas românicas. Em inglês, esse valor cai para cerca de 60% e, em alemão, para cerca de 50%. No DJ, nota-se praticamente as mesmas taxas no francês e no inglês. O índice mais alto foi obtido em português, e o mais baixo em alemão.

As relatinizações (refecções, restituições) foram ligeiramente mais freqüentes em português, especialmente no DC.

Dentre os vocábulos de origem vulgar que não apresentam hibridismo etimológico, isto é, não resultam da combinação de morfemas de diferentes procedências (dentre os quais estão os vocábulos vernáculos, os compostos e derivados formados de material vernáculo e os empréstimos de vocábulos que na sua língua de origem se enquadram nessas modalidades), observamos que os de origem portuguesa somente ocorrem em português, o que indica a pouca influência dessa língua sobre as demais, resultante sem dúvida do caráter periférico de Portugal no contexto da Civilização Européia. Os de origem italiana, inglesa e alemã são pouco freqüentes fora de suas respectivas línguas. Já os empréstimos de vocábulos franceses são ligeiramente mais freqüentes do que os das demais procedências em todas as línguas estudadas. Além disso, apresentam altas taxas de freqüência particularmente

em inglês. Vocábulos de outras procedências que não as cinco línguas estudadas são pouco freqüentes nas mesmas.

Podemos observar também que a língua alemã é a que menos exporta vocábulos para outras línguas, especialmente no que concerne ao DC (em três das cinco línguas o número de vocábulos alemães no DC foi zero).

Se somarmos todos os empréstimos de estrangeirismos não híbridos de cada língua, obteremos os seguintes resultados:

Português (DC): 24,7	Português (DJ): 37,0
Francês (DC): 31,7	Francês (DJ): 53,8
Italiano (DC): 27,4	Italiano (DJ): 42,1
Inglês (DC): 32,0	Inglês (DJ): 54,9
Alemão (DC): 45,5	Alemão (DJ): 68,3

Como se pode ver, o alemão é, em ambos os discursos, a língua mais receptiva aos estrangeirismos. Já o português apresentou o menor número de empréstimos desse tipo. Outro fato digno de nota é a semelhança de comportamento entre o francês e o inglês, tanto no DC quanto no DJ.

Quanto aos vocábulos vulgares híbridos, isto é, formados de morfemas vulgares de diversas origens (compostos e derivados híbridos, traduções de vulgarismos, empréstimos de vulgarismos híbridos estrangeiros), temos que, no DC, o inglês é a língua que os apresenta em maior quantidade, enquanto o francês é a língua em que os mesmos menos ocorrem. No DJ, temos taxas por volta de 4% em português, inglês e alemão. Em francês e em italiano, essas taxas giram em torno de 2%, o que significa que nestas duas línguas a maioria dos compostos e derivados que figuram no DJ utilizam formantes vernáculos.

Esquema de difusão dos empréstimos

Com base nas informações estatísticas obtidas em nossa pesquisa, bem como nos conhecimentos existentes sobre a história das línguas estudadas, pudemos chegar à elaboração de um mapa (Figura 4, abaixo) que retrata, de modo geral, o percurso realizado pelos vocábulos que, criados numa determinada língua européia

ocidental mediante neologia intralingüística ou importação às línguas clássicas, se difundem pelas demais línguas da Europa, na forma de empréstimos diretos, indiretos ou traduzidos. Evidentemente, assinalamos apenas os percursos principais, isto é, aqueles por onde fluem a maior parte dos empréstimos. Assim, por exemplo, a inexistência no mapa de uma flecha indicando a migração de vocábulos de Portugal para a Inglaterra não significa que não haja no inglês vocábulos de origem portuguesa: significa apenas que a contribuição do português para o léxico do inglês é pequena e, conseqüentemente, pouco importante.

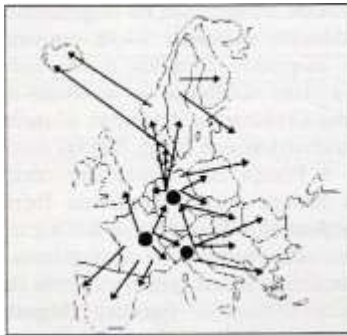


FIGURA 4 – Esquema de difusão dos neologismos – centros difusores de neologismos e áreas de influência (séculos XII a XIX)

De um modo geral, durante toda a Baixa Idade Média e a Idade Moderna, chegando até o século XIX, a língua que detém o pioneirismo na criação de neologismos de difusão internacional é o francês. Com efeito, tendo sido a França durante todo esse período a grande potência política, econômica e cultural européia, é natural que a língua francesa tivesse assumido uma posição de língua-piloto em relação às demais línguas européias ocidentais. Assim, é essa língua a primeira a importar vocábulos gregos ou latinos para dar conta de conceitos ligados ao universo de cultura, vocábulos esses que posteriormente passam por empréstimo às demais línguas. Por essa razão, os vocábulos greco-latinos são geralmente resultado de empréstimo direto do grego ou latim em francês, ao passo que nas outras línguas resultam de empréstimo indireto, através do francês. Na qualidade de língua-piloto, o

francês também detém a primazia na criação de neologismos vulgares por composição ou derivação. Os neologismos criados em francês, sejam eles eruditos ou vulgares, acabam em grande parte sendo exportados para as outras línguas, quer em sua forma original (greco-latina ou francesa), quer através de traduções, metamorfismos e restituições. É interessante notar que, embora o francês exerça uma espécie de hegemonia lingüística sobre toda a Europa Ocidental, outras línguas, como o alemão e o italiano, exercem uma certa hegemonia regional. Dessa maneira se estabelece uma complexa rede de transmissão de empréstimos léxicos que permite a difusão vocabular — e conseqüentemente a difusão dos conceitos culturais e da visão de mundo subjacentes a esses vocábulos — ao longo de toda a área abrangida pela Civilização Ocidental, aí incluídos também os países ultramarinos que falam línguas européias.

Assim, a França constitui-se em centro difusor de empréstimos léxicos para a Península Ibérica (Portugal, Espanha, Catalunha), a Inglaterra, a Itália e a Alemanha. A Itália, por sua vez, transmite os neologismos que cria e os que recebe de outros países, principalmente da França, para a Alemanha e o Sudeste Europeu (Iugoslávia, Grécia, Romênia, Península Balcânica). A Alemanha difunde seus neologismos, em sua maior parte recebidos do francês, diretamente ou na forma de traduções (tanto de grecolatinismos quanto de vulgarismos), para a Holanda, a Escandinávia (Suécia, Noruega, Dinamarca) e para o Leste Europeu (Polônia, Tchecoslováquia, Hungria, Rússia). Finalmente, a Suécia, que, assim como os demais países escandinavos, traduz boa parte dos empréstimos provenientes do alemão, retransmite tais empréstimos para a Finlândia, Rússia e países bálticos, enquanto a Noruega e a Dinamarca enviam suas inovações léxicas para a Islândia.

Vê-se, assim, que os quatro países integrantes daquilo a que chamamos na Introdução de *núcleo europeu* são, de fato, os principais responsáveis pela criação e difusão da cultura européia por todo o mundo ocidental, cultura essa que se materializa sobretudo nos elementos do léxico.

PALAVRAS DE CARACTERIZAÇÃO E VOCABULÁRIO FUNDAMENTAL

Palavras de caracterização

Conforme já dissemos em A estatística lexical, p. 177, chamam-se *palavras de caracterização* àquelas palavras que ocorrem uma única vez num determinado *corpus*. Em nosso caso, sua distribuição foi a seguinte:

DISCURSO CIENTÍFICO	DISCURSO JORNALÍSTICO
Português: 816 palavras ≈ 46,6%	Português: 492 palavras ≈ 24,6%
Francês: 388 palavras ≈ 19,4%	Francês: 489 palavras ≈ 24,5%
Italiano: 741 palavras ≈ 37,1%	Italiano: 448 palavras ≈ 22,4%
Inglês: 773 palavras ≈ 38,7%	Inglês: 419 palavras ≈ 21,0%
Alemão: 940 palavras ≈ 47,0%	Alemão: 567 palavras ≈ 28,4%

No quadro acima, as quantidades absolutas se seguem dos respectivos valores percentuais em relação ao tamanho das amostras (2.000 palavras-ocorrência). Convém lembrar que a amostra correspondente ao português científico contou 1.750 palavras-ocorrência. Conforme vemos, a taxa de palavras de caracterização girou em torno de 42% no DC e em torno de 24% no DJ. O francês apresentou, no DC, um índice bastante inferior à média das demais línguas. Tal fato pode, no entanto, ser fruto de mera casualidade, em função de as amostras terem sido aleatoriamente escolhidas. Cremos, assim, não tratar-se essa discrepância de uma característica inerente à língua francesa, mas, antes, de uma peculiaridade da amostra considerada.

No DJ, o comportamento de todas as línguas mostrou-se regular, embora o alemão tenha apresentado uma taxa um pouco superior à média.

Nota-se que a taxa média de palavras de caracterização do DC é o dobro da que se encontra no DJ. Isso é evidentemente resultado da maior especificidade dos textos científicos, bem

como da diversidade de áreas do conhecimento abrangidas no *corpus* utilizado.

Vocabulário fundamental

Se, por um lado, as palavras de caracterização ocorrem uma única vez no *corpus*, por outro lado, pertencem ao vocabulário fundamental desse *corpus* os vocábulos que figuram em todas as amostras utilizadas. Conforme já dissemos anteriormente, levantamos inicialmente os vocabulários fundamentais de cada uma das duas seções do *corpus* (científica e jornalística), e, a seguir, cruzamos os conjuntos vocabulares assim obtidos para chegarmos ao vocabulário fundamental de todo o *corpus*. Nosso objetivo foi o de observar quantos vocábulos dentre os quase 2.000 coletados são de uso comum a todos os discursos considerados, bem como verificar a natureza lexicogênica dos mesmos.

Em relação ao vocabulário fundamental por universo de discurso, obtivemos os seguintes valores:

DISCURSO CIENTÍFICO	DISCURSO JORNALÍSTICO
Português: 27 vocábulos	Português: 102 vocábulos
Francês: 35 vocábulos	Francês: 55 vocábulos
Italiano: 29 vocábulos	Italiano: 81 vocábulos
Inglês: 39 vocábulos	Inglês: 89 vocábulos
Alemão: 36 vocábulos	Alemão: 98 vocábulos

Vê-se a partir do exame do quadro acima que o número de vocábulos comuns a todas as amostras científicas oscilou entre 27 e 39 (média 33). Já no DJ, excetuado o francês, que apresentou comportamento anômalo, a tendência foi o vocabulário fundamental girar em torno de 90 vocábulos. Dada a especificidade do discurso científico e a variedade de áreas do conhecimento abrangidas, é natural que o vocabulário fundamental dessa norma discursiva seja bem inferior ao do discurso jornalístico.

Apresentamos, a seguir, a relação por língua dos vocábulos integrantes do vocabulário fundamental resultante do cruzamento de ambos os discursos.

a) vocabulário fundamental do português:

VOCÁBULO	CLASSE	IL	IE
ponto	SUB	1a	PO
problema	SUB	2a	GL
processo	SUB	2a	GL
relação	SUB	1c	GL>PO
diversos	ADJ	2a	GL
importante	ADJ	2a	GL
outro	ADJ	1a	PO
dever	VER	1a	PO
estar	VER	1a	PO
fazer	VER	1a	PO
haver	VER	1a	PO
poder	VER	1a	PO
ser	VER	1a	PO
ter	VER	1a	PO

TOTAL: 14 vocábulos

b) vocabulário fundamental do francês:

VOCÁBULO	CLASSE	IL	IE
point	SUB	1a	FR
situation	SUB	2d	GL
autre	ADJ	1a	FR
grand	ADJ	1a	FR
même	ADJ	1a	FR
seul	ADJ	1a	FR
avoir	VER	1a	FR

devoir	VER	1a	FR
être	VER	1a	FR
faire	VER	1a	FR
falloir	VER	1a	FR
pouvoir	VER	1a	FR
y avoir	VER	1a	FR

TOTAL: 13 vocábulos

c) vocabulário fundamental do italiano:

VOCÁBULO	CLASSE	IL	IE
caso	SUB	1c	GL>IT
fatto	SUB	1C	GL>IT
punto	SUB	1a	IT
rapporto	SUB	1a	IT
altro	ADJ	1a	IT
stesso	ADJ	1a	IT
avere	VER	1a	IT
dare	VER	1a	IT
dovere	VER	1a	IT
essere	VER	1a	IT
fare	VER	1a	IT
potere	VER	1a	IT
stare	VER	1a	IT
bene	ADV	1a	IT
solo	ADV	1a	IT

TOTAL: 15 vocábulos

d) vocabulário fundamental do inglês:

VOCÁBULO	CLASSE	IL	IE
-----------------	---------------	-----------	-----------

case	SUB	2a	GL
fact	SUB	2a	GL
problem	SUB	2a	GL
process	SUB	2a	GL
time	SUB	1a	IN
way	SUB	1a	IN
important	ADJ	2a	GL
only	ADJ	1f	IN
other	ADJ	1a	IN
possible	ADJ	1b	GL>FR
be	VER	1a	IN
become	VER	1d	FR>IN
can	VER	1a	IN
consider	VER	2a	GL
have	VER	1a	IN
make	VER	1a	IN
may	VER	1a	IN
there be	VER	1a	IN

TOTAL: 18 vocábulos

e) vocabulário fundamental do alemão:

VOCÁBULO	CLASSE	IL	IE
Beispiel	SUB	1c	GL>AL
Fall	SUB	1c	GL>AL
Punkt	SUB	2a	GL
Weise	SUB	1a	AL
ander	ADJ	1a	AL
ganz	ADJ	1a	AL
groß	ADJ	1a	AL

hoch	ADJ	1a	AL
neu	ADJ	1a	AL
selbe	ADJ	1a	AL
geben	VER	1a	AL
haben	VER	1a	AL
können	VER	1a	AL
lassen	VER	1a	AL
machen	VER	1a	AL
müssen	VER	1a	AL
nehmen	VER	1a	AL
sein	VER	1a	AL
stehen	VER	1a	AL
wissen	VER	1a	AL

TOTAL: 20 vocábulos

Como se pode notar, o número de vocábulos comuns a todas as amostras do *corpus* é bastante pequeno (jamais superior a 20). Como o menor conjunto de vocábulos distintos encontrado no *corpus* foi o correspondente ao português jornalístico (875 vocábulos), e como o número máximo de vocábulos encontrado numa mesma língua, aí incluídos os vocábulos que ocorreram em ambos os discursos, foi 2.276, correspondente à soma do número de vocábulos do DC e do DJ em italiano, pode-se concluir que o número de vocábulos distintos de uma determinada língua do *corpus* está entre esses dois valores; por conseguinte, a razão entre o tamanho do vocabulário fundamental e o tamanho provável do vocabulário total do *corpus* situa-se entre 1% e 2%.

Analisando as listas de vocábulos acima, percebemos que dentre os substantivos, houve alguns que se repetiram em várias línguas. São eles:

PORTUGUÊS	FRANCÊS	ITALIANO	INGLÊS	ALEMÃO
ponto	point	punto		Punkt
problema			problem	
processo			process	
relação		rapporto		
		caso	case	Fall
		fatto	fact	
			way	Weise

O conceito representado em português pelo significante *ponto* apareceu no vocabulário fundamental de quatro das cinco línguas, seguido de *caso*, que ocorreu em três línguas. Pode-se observar que os substantivos acima são todos abstratos e de certa forma devem sua universalidade ao fato de representarem uma espécie de hiperônimos de toda uma gama de conceitos abstratos, sendo portanto capazes de substituir ou de subsumir tais conceitos, ou melhor, as lexias que os recobrem.

Repetindo o mesmo tipo de análise para os adjetivos, temos:

PORTUGUÊS	FRANCÊS	ITALIANO	INGLÊS	ALEMÃO
importante			important	
outro	autre	altro	other	ander
	grand			groß
	même	stesso		selbe
	seul		only	

Os substantivos correspondentes ao port. *outro* aparecem em todas as línguas, e os correspondentes ao port. *mesmo* ocorrem em três línguas.

Praticamente os únicos verbos comuns a todas as amostras são os verbos de ligação e os verbos modais. Com exceção do inglês, em todas as línguas os verbos encontrados são vernáculos.

Somente o italiano registrou advérbios, por sinal apenas dois. Nas demais línguas inexistiram advérbios de ocorrência geral.

Do ponto de vista lexicogênico, o português apresentou 9 vernáculos, 1 empréstimo de tradução e 4 empréstimos de vocábulos greco-latinos, produzindo uma taxa de vulgaridade de 71%.

O francês apresentou 12 vernáculos e 1 grecolatinismo — taxa de vulgaridade de 92%.

No italiano tivemos 13 vernáculos e 2 empréstimos de tradução, perfazendo uma taxa de vulgaridade de 100%.

O inglês apresentou 9 vernáculos, 1 estrangeirismo, 1 empréstimo de tradução de vulgarismo, 1 derivado vulgar e 6 grecolatinismos — taxa de vulgaridade de 67%.

Finalmente, no alemão obtivemos 17 vernáculos, 2 empréstimos de tradução e 1 grecolatinismo, com uma taxa de vulgaridade de 95%.

CONSIDERAÇÕES ADICIONAIS SOBRE A ANÁLISE DOS RESULTADOS

O exame dos vocábulos encontrados no *corpus* permitiu que fizéssemos algumas constatações, que passamos agora a relatar.

a) Convém ressaltar que as amostras referentes à língua portuguesa são de origem brasileira; se houvessem sido extraídas de textos de origem lusitana, provavelmente a taxa de metamorfismos teria sido ligeiramente inferior, em virtude de, no sistema ortográfico em vigor em Portugal, se conservarem as consoantes mudas *c* e *p* (como, por exemplo, em *directo* e *adoptar*), o que coloca os vocábulos portadores dessas consoantes num estágio de transição entre a latinidade e a vulgaridade, uma vez que a pronúncia já se transformou mas a grafia ainda conserva seu caráter erudito. De qualquer maneira, tal fenômeno repercutiria muito pouco em termos estatísticos no perfil lexicogênico da língua, visto que são poucos os vocábulos afetados por essa divergência ortográfica e, além disso, a maior parte deles são também alvo de tradução.

A propósito dessa questão, cumpre observar que em todos os vocábulos de origem greco-latina encontrados no *corpus* que apresentam simultaneamente tradução de morfemas e metamorfismo, a tradução ocorre primeiro (por exemplo, lat. *productio* → port. *produção*); quando tem lugar o metamorfismo (*produção* > *produção*), o vocábulo já é vulgar. Mesmo quando a tradução e

o metamorfismo são realizados simultaneamente no próprio ato da introdução do vocábulo na língua, este decorre daquela, como por exemplo no caso do al. *natürlich* (← lat. *naturalis*), em que a passagem de *u* a *ü* se deve à presença do *i* do sufixo *-lich*, tradução do sufixo latino *-alis*, por força do fenômeno do *Umlaut*.

b) Embora não tenhamos estudado neste trabalho as línguas escandinavas, podemos deduzir que, em termos de índices de vulgaridade, as mesmas se situam muito próximas ao alemão, pois a maior parte dos empréstimos de origem greco-latina dessas línguas aí chegaram vindas do alemão (p. 204). Assim, o que era grecolatinismo em alemão passou ao sueco, norueguês e dinamarquês sem sofrer nenhum processo vulgarizante (por exemplo, gr. *systema* → al. *System* → sue. *system*). Por outro lado, a maioria dos vocábulos alemães que resultam de tradução de vocábulos greco-latinos foram retraduzidos nas línguas escandinavas (lat. *collaborare* → al. *mitarbeiten* → sue. *samarbeta*). Em uns poucos casos, as línguas escandinavas respondem com um grecolatinismo a uma tradução alemã, como no caso do sue. *television* (a partir do ingl. *television*) em face do al. *Fernsehen*, o que sugere talvez taxas de vulgarização ligeiramente inferiores às do alemão. cremos, entretanto, que tal fato seria muito pouco perceptível em termos estatísticos, podendo, mesmo, vir a ser mascarado pelo erro implicado na medida. Evidentemente, essas considerações possuem caráter meramente especulativo, uma vez que não nos foi possível testar experimentalmente tal hipótese, em virtude da dificuldade de trabalhar com essas línguas, conforme explicamos na Introdução.

Quanto ao espanhol, acreditamos que o mesmo apresente um perfil ligeiramente menos vulgar que o português, visto que traduz menos os empréstimos latinos (por exemplo, o sufixo latino *-tione*, que em português é sistematicamente traduzido por *-ção*, conserva-se em espanhol na forma *-ción*), assim como também metamorfiza menos (por exemplo, os grupos consonantais *-ct-* e *-pt-*, que no português costumam evoluir para *-t-*, *-it-* ou *-ut-*, são normalmente preservados em espanhol).

c) Pudemos observar em ambos os discursos analisados que as taxas de ocorrência de vocábulos vulgares resultantes de composição ou derivação (categoria 1f) são sensivelmente mais elevadas em alemão. Pudemos perceber também que muitos desses vocábulos foram criados para corresponder a vocábulos greco-

latinos ou grecolatinismos pertencentes às demais línguas. Embora tais vocábulos não possam rigorosamente ter sido classificados como empréstimos de tradução, uma vez que não existe correspondência entre os morfemas do vocábulo greco-latino e os morfemas do vocábulo alemão, o fato é que também em sua gênese se encontra o intuito de substituir material greco-latino por material vernáculo. Assim, é lícito dizer que, num sentido amplo, o processo da tradução de empréstimos eruditos em alemão é ainda mais abrangente do que o mostram as cifras obtidas somente em relação à categoria 2c.

d) Se compararmos um texto francês com sua tradução para o inglês, encontraremos em ambos os textos um grande número de vocábulos greco-latinos. Em geral, tais vocábulos apresentarão em ambas as línguas uma grafia bastante fiel ao modelo greco-latino (a chamada grafia etimológica). De um modo geral, as únicas divergências ortográficas entre o grego/latim e essas línguas serão:

LATIM	FRANCÊS	INGLÊS
-ae- (<i>cohaesio</i>)	-é- (<i>cohésion</i>)	-e- (<i>cohesion</i>)
-oe- (<i>oeconomia</i>)	-é- (<i>économie</i>)	-e- (<i>economy</i>)
-i- (<i>theoria</i>)		-y (<i>theory</i>)

Não só em termos de grecolatinismos, mas também em termos de vulgarismos encontraremos muitas coincidências entre ambas as línguas, já que, como se sabe, a maior parte dos empréstimos léxicos do inglês provieram do francês, e é inclusive por intermédio dessa língua que a maioria dos vocábulos gregos ou latinos chegaram ao inglês.

Por outro lado, se compararmos um texto italiano e seu correspondente alemão, veremos que as duas línguas divergem bastante do modelo greco-latino. Essas divergências são de duas ordens: ortográficas e morfológicas.

Do ponto de vista ortográfico, temos:

GREGO/LATIM	ITALIANO	ALEMÃO
-ti- (<i>natio</i>)	-zi- (<i>nazione</i>)	
-c- (<i>concursum, centrum</i>)		-k-/-z- (<i>Konkurs, Zentrum</i>)
-ph- (<i>*telephonus</i>)	-f- (<i>telefono</i>)	-f- (<i>Telefon</i>)

-th- (<i>theoria</i>)	-t- (<i>teoria</i>)	
-ch- (<i>technicus</i>)	-c- (<i>tecnico</i>)	
-rh- (<i>rhythmus</i>)	-r- (<i>ritmo</i>)	
-bh- (<i>abhorrere</i>)	-b- (<i>aborrere</i>)	
-dh- (<i>adhaesio</i>)	-d- (<i>adesione</i>)	
-h- (<i>habitus</i>)	-Ø- (<i>abito</i>)	
-j- (<i>projectio</i>)	-i- (<i>proiezione</i>)	
-y- (<i>physica</i>)	-i- (<i>física</i>)	
-ae- (<i>cohaesio</i>)	-e- (<i>coesione</i>)	-ä- (<i>Kohäsion</i>)
-oe- (<i>oeconomia</i>)	-e- (<i>economia</i>)	-ö- (<i>Ökonomie</i>)
-ss- (<i>processus</i>)		-ß- (<i>Prozeß</i>)

Do ponto de vista morfológico, temos no italiano basicamente os diversos tipos de metamorfismo (ver p. 105) e as traduções de radicais e afixos. Em alemão, temos sobretudo as traduções (ver p. 89).

Fenômeno semelhante ocorre entre as línguas ibéricas (português, espanhol, catalão) e as escandinavas (norueguês, sueco, dinamarquês). Divergências ortográficas em relação ao latim são:

GREGO/LATIM	PORTUGUÊS	NORUEGUÊS
-ti- (<i>*nationalis</i>)	-ci- (<i>nacional</i>)	-sj- (<i>nasjonal</i>)
-c- (<i>concursum, centrum</i>)		-k/-s- (<i>konkurs, sentrum</i>)
-ph- (<i>*telephonus</i>)	-f- (<i>telefone</i>)	-f- (<i>telefon</i>)
-th- (<i>theoria</i>)	-t- (<i>teoria</i>)	-t- (<i>teori</i>)
-ch- (<i>technicus, orchestra</i>)	-c-/-qu- (<i>técnico, orquestra</i>)	-k- (<i>teknisk, orkester</i>)
-rh- (<i>rhythmus</i>)	-r- (<i>ritmo</i>)	-r- (<i>rytm</i>)
-bh- (<i>abhorrere</i>)	-b- (<i>aborrecer</i>)	
-dh- (<i>adhaesio</i>)	-d- (<i>adesão</i>)	
-h- (<i>cohaerens</i>)	-Ø- (<i>coerente</i>)	
-x- (<i>mixtus</i>)	-s- (<i>misto</i>)	
-y- (<i>physica</i>)	-i- (<i>física</i>)	
-ae- (<i>cohaesio</i>)	-e- (<i>coesão</i>)	-æ- (<i>kohæsjon</i>)

-oe- (<i>oeconomia</i>)	-e- (<i>economia</i>)	-ø- (<i>økonomi</i>)
consoantes dobradas (exceto <i>rr</i> e <i>ss</i>)	consoantes sim- ples	

As principais divergências morfológicas que as línguas ibéricas apresentam em relação ao latim são os metamorfismos e as traduções. Nas línguas escandinavas se destacam basicamente as traduções.

Com relação a esse ponto, é interessante notar também que o italiano e as línguas ibéricas divergem do francês e das línguas germânicas principalmente no tocante às desinências nominais e às terminações de palavra. De fato, enquanto o francês e as línguas germânicas terminam as palavras basicamente por consoante, semivogal ou *e* (que é mudo em francês e inglês), o italiano, o espanhol e o português apresentam muitas terminações vocálicas (*a*, *e*, *o*; em espanhol e italiano *i*, em espanhol também *u*). Por outro lado, o alemão e as línguas escandinavas divergem do inglês e das línguas românicas especialmente do ponto de vista léxico, pois, como vimos, nessas línguas, a taxa de grecolatinismos é bem inferior à das demais línguas. Pode-se dizer, em resumo, que o francês se distingue das restantes línguas românicas por apresentar terminações consonantais, que são típicas das línguas germânicas, ao passo que o inglês difere das demais línguas germânicas por possuir um léxico bastante latinizado, fenômeno por sua vez peculiar às línguas românicas.

e) Ao examinarmos os metamorfismos e os empréstimos de tradução nas diversas línguas, constatamos que, de modo geral, cada uma dessas línguas aplica tais processos de forma diversa. Assim, podemos dizer que, via de regra, dado um vocábulo latino formado por um radical e por um ou mais afixos (prefixos e/ou sufixos), o inglês tende a conservar a todos esses elementos na forma original latina. Já o francês costuma traduzir os sufixos, ao passo que o italiano em geral traduz os prefixos e metamorfiza o radical. O alemão, por sua vez, apresenta uma tendência à tradução total (radical e afixos). Quanto ao português, observamos um comportamento intermediário entre o francês e o inglês, pois nessa língua as traduções de sufixos são mais freqüentes que em inglês, porém menos que em francês, enquanto o metamorfismo de radicais é mais comum do que no francês.

A explicação do diferente comportamento dos afixos nas diversas línguas é o fato de o francês e o alemão possuírem mais sufixos vulgares do que as demais línguas, ao passo que no italiano e sobretudo no alemão sobressaem os prefixos vulgares, conforme se pode constatar na p. 93.

f) A partir da análise da composição etimológica dos vocábulos das diversas línguas (p. 206), pudemos constatar que o francês é a língua que até hoje mais exportou vocábulos para outras línguas européias, razão pela qual as taxas de galicismos dessas línguas são sempre superiores às de anglicismos, italianismos, germanismos, etc. Vimos também que essa influência francesa se fez sentir de modo particularmente intenso no inglês, do qual uma parcela significativa do léxico (cerca de 20%) é formada de elementos de origem francesa. Ao contrário, o volume de vocábulos ingleses no francês não é superior ao que encontramos nas demais línguas, nem tampouco é superior à quantidade de vocábulos italianos, alemães, etc., existentes no próprio francês. Assim, a nosso ver, a suposta invasão de anglicismos no francês, denunciada e combatida por certos gramáticos puristas, está longe de representar uma ameaça à identidade lingüística desse idioma. Consideramos, portanto, deplorável, por equivocada, a onda de xenofobia e de nacionalismo extremado que de tempos em tempos se abate sobre a língua francesa, e que acaba por ensejar criações lexicais bizarras, como por exemplo o termo *logiciel* (tentativa de tradução do ing. *software*), provavelmente um composto dos fractomorfe-mas *logi-* (de *logique*) e *-ciel* (de *artificiel?*), vocábulo que, por sua estrutura morfológica pouco condizente com a tradição do léxico de cultura do francês, acaba por ser mais alienígena que o próprio estrangeirismo que procura substituir. Essas traduções desnecessárias e pouco criteriosas do francês acabam por distanciar essa língua das demais, como ocorreu, por exemplo, no caso da palavra *ordinateur*, tradução mal-ajambrada do ing. *computer* (já que a principal função do computador não é “ordenar”) e que, além de romper o parentesco etimológico com termos correlatos como *computation*, *computationnel*, etc., destoa da forma adotada por outras línguas (ing., it. *computer*; al. *Computer*; port., esp., cat. *computador*, etc.).

g) Analisando os diversos tipos de grecolatinismos e de vulgarismos existentes nas línguas estudadas, verificamos que há na

verdade vários graus de grecolatinidade ou de vulgaridade, conforme podemos perceber nos exemplos abaixo:

- O al. *Prozeß* é um grecolatinismo (IL=2a), porém o caráter latino desse vocábulo está bastante mascarado pela grafia;

- O ing. *photographer* é um vulgarismo híbrido, pois possui radical erudito e sufixo vernáculo. Não obstante, dois dos três morfemas do vocábulo são gregos. Em termos da extensão desses morfemas, medida em número de fonemas, poderíamos dizer que cerca de 80% do vocábulo é erudito.

- O it. *prevedere* e o al. *versehen* são ambos empréstimos de tradução do lat. *praevidere*, mas a forma italiana está muito mais próxima da latina do que a alemã.

- O al. *Ebene* é empréstimo de tradução em relação ao lat. *planus*, assim como o it. *stato* o é em relação ao lat. *status*, mas o vocábulo italiano praticamente coincide com o latino e poderia mesmo parecer um grecolatinismo. Igualmente, o it. *porta*, que é vernáculo, coincide com o lat. *porta*, do qual se origina.

CIRCUNSTÂNCIAS HISTÓRICAS E OS RESULTADOS OBTIDOS

Comparando o léxico das cinco línguas estudadas, do ponto de vista dos processos lexicogênicos aí verificados, chegamos a algumas constatações interessantes, e até certo ponto surpreendentes, a respeito do comportamento dessas línguas. Pudemos perceber claramente duas diferentes tendências historicamente delineadas: de um lado, uma tendência que poderíamos chamar, com certo abuso terminológico, de *classicizante*, isto é, bastante receptiva às formas gregas e latinas, do ponto de vista fonético-fonológico, morfológico e ortográfico; de outro lado, uma tendência mais *vulgarizante*, em que predominam os elementos vernáculos e em que os empréstimos greco-latinos sofrem de maneira mais sistemática processos de que resultam vocábulos semi-eruditos, isto é, vocábulos de origem greco-latina formados, total ou parcialmente, de morfemas vulgares, ou modificados fonológica e graficamente por imitação das formas vernáculas.

Do ponto de vista dessa dualidade de comportamento, podemos observar uma grande semelhança entre as línguas portu-

guesa, francesa e inglesa, de um lado, e entre o italiano e o alemão, de outro. Se pudéssemos ter incluído em nossa pesquisa mais línguas românicas e germânicas, teríamos decerto podido constatar a vulgaridade como característica das línguas desta última família, à exceção do inglês, bem como comprovar de modo mais cabal a grecolatinidade das línguas românicas, com exceção do italiano. De qualquer maneira, lançando um olhar mais atento sobre as línguas dos quatro países que formam o chamado núcleo europeu, percebemos uma clara analogia entre as duas famílias abordadas. De um lado, encontramos diversos pontos em comum entre o léxico do francês e do inglês. Além de coincidências em termos de dados estatísticos, pudemos notar tendências lexicogênicas bastante semelhantes entre essas duas línguas. De outro lado, encontramos entre o italiano e o alemão tendências igualmente convergentes, e ao mesmo tempo opostas à do francês e do inglês. Podemos dizer, em suma, que o francês e o inglês são línguas mais classicizantes, ou melhor, menos vulgarizantes, que o italiano e o alemão. As línguas ibéricas parecem seguir a tendência do francês e do inglês, ao passo que nas línguas escandinavas encontraremos um comportamento paralelo ao do italiano e do alemão.

Diversas causas históricas devem ser invocadas para explicar tais fatos. A tendência latinizante do francês delinea-se já na Baixa Idade Média (séculos XI a XV), quando grande número de grecismos e latinismos invade a língua, enriquecendo-a. Uma rápida consulta a um dicionário etimológico francês permite verificar que a maior parte dos empréstimos greco-latinos dessa língua, a partir dos quais posteriormente derivarão os outros, se deu no período mencionado. Contudo, o processo de grecolatinização do francês é muito mais extenso e complexo, sendo o resultado de um trabalho árduo e contínuo de escritores e gramáticos cultores da tradição greco-latina. A esse respeito, diz Malmberg:

A cultura clássica teve como porta-voz o grego e, posteriormente, o latim, muito influenciado por aquele. Quando a nova Europa se levantou das ruínas do mundo antigo, sua cultura foi, em princípio, latina. Paris, a capital do estado francês surgido depois da conquista alemã, tornou-se, então, o centro do mundo ilustrado, exercendo através de uma universidade uma grande atração sobre a elite intelectual européia que ali era educada no pensamento científico. (...) Neste clima de puro intelectualismo, foi-se formando, pouco a pouco, a língua francesa, até chegar a constituir a expressão sólida de

um pensamento claro. Isso não aconteceu de um momento para outro: demorou muito tempo e exigiu um paciente trabalho até que o pobre idioma dos Juramentos de Estrasburgo de 842, ou da lenda de Alexius do século XI, adquirisse a forma que encontramos em Pascal ou em Racine. Os gramáticos e puristas possibilitam a evolução, graças a sua familiaridade com a tradição clássica e ao conhecimento do pensamento filosófico. A gramática de Vaugelas, — as famosas “Remarques sur la langue française” de 1647 — codificou definitivamente o francês clássico, que ainda serve, no essencial, de norma para uso lingüístico. (...) o francês se formou e desenvolveu através de um povo que, sem dúvida, teve uma parte maior do que qualquer outro na criação da tradição cultural do Ocidente. Não é, pois, por casualidade, nem constitui também apenas uma consequência do poder político, o fato desta língua ter sido, mais do que qualquer outra, a voz do Ocidente. (...) O caminho seguido pela cultura européia começa na antigüidade clássica e passa pelo cristianismo, a Idade Média, e o Renascimento, antes de chegar à Idade Moderna. Em cada fase, esse caminho esteve intimamente unido a uma forma de expressão lingüística de que não pode separar-se. Pois bem, o grego, o latim (clássico e medieval) e o francês são as línguas fundamentais desta história do pensamento.³⁰

Vê-se assim que, por uma conjunção de fatores, o francês se tornou o herdeiro natural da tradição clássica. Entretanto, a nosso ver esse fato se deve não apenas a razões históricas e culturais, mas também lingüísticas. Conforme dissemos anteriormente, em Principais metamorfismos, p. 106, a vulgarização muitas vezes decorre da proximidade entre as formas latina e vulgar, caso em que a segunda naturalmente substitui a primeira. Ora, sendo o francês a língua românica formalmente mais diferenciada do latim, parece-nos natural que apresente poucas vulgarizações, preservando quase sempre intacto o aspecto clássico das palavras. Vejamos, a seguir, alguns exemplos, citados por Auerbach: o lat. *vigilare*, que existia em francês na forma popular *veiller*, foi introduzido uma segunda vez, e deu origem ao substantivo erudito *vigilance*; o mesmo vale para o lat. *fragilis*, forma popular *frêle*, forma erudita *fragile*; lat. *fides*, adjetivo latino *fidelis*, forma popular do substantivo *foi*, do adjetivo em francês antigo *fëoil*, forma erudita do adjetivo *fidèle*, do qual um substantivo *fidélité*; lat. *directum*,

³⁰ MALMBERG, B. *A língua e o homem*. Introdução aos problemas gerais da lingüística. Rio de Janeiro, Nórdica/São Paulo, Duas Cidades, 1976, p. 84 s.

forma popular *droit*, forma erudita *direct*; lat. *gradus*, forma popular (*de*)*gré*, forma erudita *grade*, etc. Muitas das palavras que entraram no francês por ocasião dessa segunda latinização passaram rapidamente ao uso corrente, como, por exemplo: *agriculture*, *captif* (forma popular *chétif*), *concilier*, *diriger*, *docile*, *éducation*, *effectif*, *énumérer*, *explication*, *fabrique* (fr. pop. *forge*), *facile*, *fréquent*, *gratuit*, *hésiter*, *imiter*, *invalide*, *légal* (fr. pop. *loyal*), *muniton*, *mobile* (fr. pop. *meuble*), *naviguer* (fr. pop. *nager*), *opérer*, *penser* (fr. pop. *peser*), *pacifique*, *quitte* e *inquiet* (fr. pop. *coi*), *rédemption* (fr. pop. *rançon*), *rigide* (fr. pop. *raide*), *singulier* (fr. pop. *sanglier*), *social*, *solide*, *espèce* (fr. pop. *épice*), *tempérer* (fr. pop. *tremper*), *vitre* (fr. pop. *verre*), etc.³¹

O caráter igualmente clássico do inglês se explica pelo francês. Com efeito, durante toda a Baixa Idade Média esteve a Inglaterra politicamente sujeita a dinastias francesas (normandos e plantagenetas). Durante esse tempo foi o francês a língua oficial do país. Isso se reflete no grande número de empréstimos vulgares franceses no inglês. Também a maioria dos grecolatinismos do inglês aí penetraram por via francesa. Assim, é natural que haja muitas semelhanças entre os léxicos dessas duas línguas. Mesmo após a desvinculação política da Inglaterra em relação à França, a língua francesa continuou e continua a ser o principal modelo em que se inspira o inglês. Na verdade, grande parte dos radicais e afixos com que o inglês constrói seus neologismos são franceses. A galicidade — e através dela a latinidade — tornou-se portanto marca indelével da língua inglesa. Diz Kurath:

A conquista normanda da Inglaterra no século XI estabelece um ponto decisivo da história do povo inglês e de sua língua. O francês se tornou, por três séculos, a língua de todos os assuntos de âmbito nacional, quer políticos, quer sociais, quer culturais. A língua literária do saxão ocidental foi logo abandonada, mas as variedades regionais do inglês continuaram em uso nos lares, no culto religioso cristão e nas transações de negócios locais da província. No início do século XIII, surgiram as línguas literárias regionais baseadas na fala oral, primeiro na região do sudoeste (o *Brut* de Layamon, vidas de santos), depois do norte e por último na área de Londres, a sede principal da aristocracia francesa. Na época de Chaucer, a língua da

³¹ Cfr. AUERBACH, E. *Introduzione alla filologia romanza*. Tradução de Maria Rosa Massei. Turim, Einaudi, 1974, p. 115-116.

área focal de Londres e Westminster ficou estabelecida como a língua literária padrão da Inglaterra. A volta do inglês como língua nacional reflete a ascensão da classe média ao poder em assuntos nacionais e culturais.

Durante cerca de trezentos anos de bilingüismo socialmente gradativo, grande número de palavras francesas foi incorporado ao vocabulário inglês, não apenas termos de cunho político, social ou cultural — *country, city, real, royal, value, money, society* — mas também termos corriqueiros como *change, move, use, cover, very*. Durante este período muitas palavras de origem latina foram também adotadas, especialmente em traduções de trabalhos eruditos, tais como tratados sobre Direito, Teologia, Filosofia e Medicina. Boa parte do vocabulário usado nestes campos de estudo data desse período, como bem o demonstra o *Middle English Dictionary* publicado atualmente pela Universidade de Michigan.³²

Se, conforme dissemos, o francês deve em grande parte seu caráter clássico ao distanciamento formal do latim, o elevado grau de vulgarização do italiano pode ser explicado pela razão exatamente oposta. De fato, o italiano é, das línguas românicas, a mais próxima do latim em termos fonéticos. Por essa razão inclusive, o idioma popular só foi tardiamente reconhecido como língua literária, por obra de gênios como Dante e Petrarca, dentre outros.³³ Na verdade, acreditavam os italianos dos séculos XI e XII ainda falar latim, apenas sob forma um pouco diversa. Esse fato, somado à supremacia religiosa de Roma, fez da Itália o centro espiritual do Ocidente e contribuiu particularmente para forjar nos falantes um sentimento de que a península ainda se constituía na capital histórica e cultural da România. Essa crença no mito de que os italianos, mais do que qualquer outro povo latino, eram os legítimos herdeiros da tradição e da cultura romanas foi retomada diversas vezes ao longo da história, a mais recente delas durante o regime fascista, influenciando a evolução da língua a ponto de tornar a vernaculização um traço compulsório (a maior parte dos metamorfismos do italiano são fonologicamente obrigatórios ou recomendados e não facultativos). Por essa razão, a criação lexical no italiano, malgrado tenha sofrido a mesma influência greco-latina

³² Apud HILL, A. A., org. *Aspectos da lingüística moderna*. São Paulo, Cultrix/EDUSP, 1974, p. 75-76.

³³ Cfr., por exemplo, DANTE ALIGHIERI, *De vulgari eloquentia*, 1303-1304.

também verificada nas demais línguas, possui um caráter que poderíamos chamar de *endocêntrico*.

Comportamento ideológico semelhante detectamos no alemão. Cumpre então buscar na história as causas desse endocentrismo lingüístico. A exemplo da Itália, desde muito cedo sentiu-se a Alemanha a capital do mundo germânico. Se não se têm até hoje informações precisas sobre o berço dos povos germânicos, acreditando alguns autores que houvessem os germanos partido da Escandinávia, enquanto outros situam seu *habitat* primitivo na Europa Oriental, o fato é que foi do território alemão que as diversas tribos germânicas partiram para a conquista de Roma. Não por acaso, a região da Europa onde hoje se situa a Alemanha era chamada pelos romanos de Germânia, isto é, pátria dos germanos. Em decorrência disso, em muitas línguas *Germânia* e *germano* são sinônimos de Alemanha e alemão. Mais ainda, a Alemanha conquistou esse predomínio no âmbito dos povos germânicos por estar a Escandinávia em posição geográfica — e conseqüentemente política e cultural — periférica, ao passo que a Inglaterra gravitava na órbita latina.

Sem dúvida, a fraca romanização da Germânia, situada em sua maior parte fora do *limes* imperial, e ao mesmo tempo sua proximidade a essa fronteira, conspirou em muito para que desde cedo surgisse entre as tribos teutônicas de além Reno um sentimento de oposição ao mundo romano. De fato, o alemão foi a primeira das línguas germânicas modernas a produzir documentos escritos, ainda no século VIII, cem anos, portanto, antes do primeiro registro do francês. Ocorre que, enquanto os falantes do romance imaginavam ainda falar alguma forma de latim, os germanos da Alemanha tinham uma clara consciência do não parentesco entre sua língua e o latim. Por essa época, os escandinavos ainda se encontravam fora do alcance ideológico de Roma: nem a cultura greco-latina nem a religião cristã haviam ainda chegado ao extremo norte da Europa. Por essa razão, é natural que nenhum povo germânico tenha vivenciado de forma tão intensa o conflito entre sua identidade étnica e cultural e sua integração, através da cristianização, ao universo ideológico romano.

Não obstante, o nacionalismo alemão desenvolveu-se paralela e conjuntamente ao nacionalismo italiano. Tratava-se de uma disputa entre o imperador germânico e o papa em torno dos poderes espiritual e temporal, disputa esta de caráter eminentemente

político, uma vez que, enquanto Roma liderava a religião no Império, a Alemanha detinha o domínio territorial da Itália. Não é, pois, por uma questão de consciência patriótica, mas principalmente por aversão ao espírito da escolástica, que a maior parte dos autores religiosos adotam a língua alemã. Essa reação à escolástica e à sua expressão lingüística, o latim, viria a acirrar-se por ocasião da Reforma. Contudo, a história da língua alemã é uma contínua tensão entre latinidade e germanidade. Sucedem-se a fases mais receptivas à importação estrangeira, especialmente latina, outras mais refratárias e puristas. O correr do tempo parece ter consagrado esta última tendência. Sem dúvida, o alemão sofreu muito mais do que o inglês ou as línguas românicas o impacto da influência greco-latina, e por isso mesmo apresentou reação tão marcante. Cumpre ressaltar no entanto que o grande número de empréstimos de tradução e de neologismos sintagmáticos criados em substituição a latinismos não significou a rejeição da cultura clássica, mas antes sua perfeita assimilação: por trás desse aparente manto de germanidade o que há é uma organização de conceitos e uma visão de mundo greco-romano-cristã, especialmente no que concerne ao universo da cultura e da ciência. Os conceitos gregos e latinos ganharam no alemão, graças à sua tradução, uma naturalidade comparável à que possuem nas línguas românicas, ao passo que no inglês, a mais latinizada das línguas germânicas, boa parte dos vocábulos gregos e latinos são sentidos, na fala popular, como estrangeirismos, uma vez que têm seu uso restrito à norma culta, permanecendo ainda, de certa forma, mal integrados ao léxico.

Não obstante, esse caráter vernaculizante do alemão contribuiu muito para a consolidação da própria ideologia alemã: a crença nas origens e o apego à mitologia e às tradições dos antigos germanos, que teve seu auge no século XIX com o Romantismo (movimento estético e ideológico que, por sinal, marca o apogeu da cultura alemã), através da obra de gênios como Goethe, Wagner, Nietzsche, entre outros, e que desaguarda, já no século XX, no nacionalismo e na xenofobia exacerbados do nazismo.³⁴ Sabe-se que os períodos de nacionalismo exacerbado caracterizam-se sempre pelo combate à influência estrangeira no idioma. Dentre esses períodos, podemos citar as guerras contra Napoleão, a revo-

³⁴ Cfr. a esse respeito BUTLER, R. d'O. *Raíces ideológicas del nacional-socialismo*. México, Fondo de Cultura Económica, 1943.

lução de 1848, a guerra franco-prussiana de 1870, a Primeira Guerra Mundial e o regime nazista.

Atualmente, a língua parece seguir outra direção, mas o presente é o produto inevitável do passado, e as marcas ideológicas adquiridas ao longo do tempo permanecem. A partir do grande desenvolvimento técnico e científico dos séculos XIX e XX, muitos estrangeirismos penetraram na língua alemã, alguns na forma original, outros na forma de traduções e outros ainda apresentam ambas as formas. Vê-se assim que as diversas tendências da língua estão sujeitas a muitos fluxos e refluxos, e que a escolha desta ou daquela forma, deste ou daquele processo, é basicamente fruto de uma decisão política.

CONCLUSÃO



Neste trabalho, nos propusemos estabelecer um modelo de análise do léxico de línguas européias ocidentais que levasse em conta o aspecto lexicogênico, sobretudo do ponto de vista da tensão entre léxico erudito e léxico vulgar. A partir desse modelo, realizamos uma análise estatística do comportamento de três línguas românicas e duas germânicas, com base em dois diferentes universos de discurso, e, a seguir, levamos a cabo um estudo comparativo dessas línguas a partir dos dados estatísticos obtidos. Como tivemos a oportunidade de verificar, no francês a evolução fonética conduziu a uma clara separação entre os estratos erudito e vulgar; no inglês, por via da influência francesa, se verifica igual distinção. Nas demais línguas românicas, e em especial no italiano, essas duas ordens de elementos se integraram melhor. Evidentemente, a ausência de parentesco direto entre o latim e as demais línguas germânicas teria conduzido o léxico dessas línguas à mesma oposição *erudito/vulgar*, não fosse o fato de as mesmas terem optado pela nacionalização da maioria dos aportes eruditos, o que acabou por propiciar, também neste caso, uma ampla integração entre o elemento greco-latino e o vernáculo.

Pode-se concluir, portanto, pela existência de duas tendências ideológicas básicas na lexicogênese das línguas européias ocidentais: de um lado, uma maior receptividade à influência das línguas clássicas, de que resulta um léxico mais claramente híbrido e heterogêneo; de outro, uma propensão mais acentuada à vernaculização, que conduz a um léxico onde os elementos provenientes das diversas origens se misturam com mais facilidade e convivem mais harmonicamente. Pudemos detectar, dentro das naturais e evidentes limitações de nossa pesquisa, o predomínio da primeira tendência nas línguas românicas, à exceção do italiano, e da segunda nas línguas germânicas, excetuado o inglês. Assim, se restringirmos nossa análise àquilo que definimos na Introdução como sendo o núcleo histórico da Civilização Européia, veremos

que a divisão entre essas duas tendências não coincide com a fronteira das duas famílias lingüísticas, mas na verdade, do ponto de vista da ideologia subjacente aos processos de criação e renovação lexicais, devemos alinhar França e Inglaterra de um lado, e Itália e Alemanha de outro. Desse modo, as duas famílias de línguas passam a formar uma imagem especular uma da outra.

BIBLIOGRAFIA

- ALVES, I. M. *Neologismo: criação lexical*. São Paulo, Ática, 1990.
- AUERBACH, E. *Introduzione alla filologia romanza*. Tradução de Maria Rosa Massei. Turim, Einaudi, 1974.
- BACHELARD, G. *La formation de l'esprit scientifique*. Paris, J. Vrin, 1972.
- BALDINGER, K. *Teoría semántica*. Hacia una semántica moderna. Madrid, Alcalá, 1970.
- BARBOSA, M. A. *Léxico, produção e criatividade: processos do neologismo*. São Paulo, Global, 1981.
- _____. Acrograma e sigla: estatuto semântico-sintático e tratamento na obra lexicográfica. *Anais da 45ª Reunião Anual da SBPC*. Recife, UFPE, 1993, p. 477.
- BARTHES, R. *Elementos de semiologia*. São Paulo, Cultrix, 1970.
- BENVENISTE, E. *Problèmes du langage*. Paris, Gallimard, 1966.
- _____. *Problèmes de linguistique générale*. Paris, Gallimard, 1974.
- BENVENISTE, E.; KRISTEVA, J.; MILNER, J. C. *Langue, discours, société*. Paris, Du Seuil, 1975.
- BIZZOCCHI, A. L. Aspectos da alotropia nas línguas românicas: esboço de uma tipologia. *Anais da 43ª Reunião Anual da SBPC*, Rio de Janeiro, SBPC, 1991, p. 401-402.
- _____. Aspectos semióticos do empréstimo vocabular interlingüístico. *Revista Brasileira de Lingüística*, v. 8, nº 1. São Paulo, SBPL/Editora Plêiade, 1995, p. 31-42.

- BORBA, F. da S. *Introdução aos estudos lingüísticos*. São Paulo, Nacional, 1979.
- BOURCIEZ, E. *Eléments de linguistique romane*. Paris, Klincksieck, 1967.
- BUTLER, R. d'O. *Raíces ideológicas del nacional-socialismo*. México, Fondo de Cultura Económica, 1943.
- CÂMARA JR., J. M. *Para o estudo da fonêmica portuguesa*. Rio de Janeiro, Padrão, 1977.
- CARONTINI, J. & PERAYA, D. *Le projet sémiotique*. Paris, Delarge, 1975.
- CHARAUDEAU, P. *Langage et discours: éléments de sémiolinguistique (théorie et pratique)*. Paris, Hachette, 1983.
- COSERIU, E. *Teoría del lenguaje y lingüística general*. Madrid, Gredos, 1967.
- D'OVIDIO, F. & MEYER-LÜBKE, W. *Grammatica storica della lingua e dei dialetti italiani*. Milão, Ulrico Hoepli, 1932.
- DUBUC, R. *Manuel pratique de terminologie*. Montreal, Linguattech/Bibliothèque Nationale du Québec, 1985.
- ÉMORINE, J. L'indexation du vocabulaire littéraire de la langue portugaise. Toulouse, s.ed., 1969. /Xerocopiado/
- ERBEN, J. Neologismen im Spannungsfeld von System und Norm. *Logos Semantikos: Studia Linguistica in Honorem Eugenio Coseriu 1921-1981*, v. 5: „Geschichte und Architektur der Sprachen“. Berlim, De Gruyter; 1981.
- FERREIRA, M. A. S. de C. *Estrutura e formação de palavras: teoria e prática*. São Paulo, Atual, 1988.
- FOGARASI, M. *Nuovo manuale di storia della lingua italiana*. Florença, Le Monnier, 1990.
- GALISSON, R. *Inventaire thématique et syntagmatique du Français Fondamental*. Paris, Hachette, 1971.
- _____ *Recherches de lexicologie descriptive: la banalisation lexicale*. Paris, Nathan, 1978.
- GENOUVRIER, E. & PEYTARD, J. *Lingüística e ensino do português*. Coimbra, Almedina, 1974.
- GILI Y GAYA, S. *Elementos de fonética general*. Madrid, Gredos, s.d..

- GRANDGENT, C. H. *Introducción al latín vulgar*. Tradução de Francisco B. de Moll. Madrid, Publicaciones de la Revista de Filología Española, 1952.
- GREIMAS, A.-J. *Sémantique structurale. Recherche de méthode*. Paris, Larousse, 1966.
- _____ *Du sens: essais sémiotiques*. Paris, Seuil, 1970.
- _____ *Semiótica do discurso científico*. Da modalidade. São Paulo, DIFEL/SBPL, 1976.
- _____ *Sémiotique et sciences sociales*. Paris, Seuil, 1976.
- GREIMAS, A.-J. *Du sens II: essais sémiotiques*. Paris, Seuil, 1983.
- GUILBERT, L. *La créativité lexicale*. Paris, Larousse, 1975.
- HAENSCH, G. et al. *La lexicografía*. De la lingüística teórica a la lexicografía práctica. Madrid, Gredos, 1982.
- HARTMANN, R. Limites da lexicologia contrastiva. *Problemas da lexicologia e lexicografia*. Porto, Civilização, 1979.
- HERBERG, D. Neologismen: Lexikologisch und lexikographisch betrachtet Sprachpflege. *Zeitschrift für Gutes Deutsch*, v. 37(8), 1988, p. 109-112.
- HILL, A. A., org. *Aspectos da lingüística moderna*. São Paulo, Cultrix/EDUSP, 1974.
- HJELMSLEV, L. *Essais linguistiques*. Paris, Minuit, 1971.
- _____ *Prolegômenos a uma teoria da linguagem*. São Paulo, Perspectiva, 1975.
- JAEGGLI, O. *Studies in Romance linguistics*. Dordrecht, Foris, 1986.
- KARSTEN, T. E.; MOSSE, F.; MEILLET, A. *Anciens germains: introduction à l'étude des langues et des civilisations germaniques*. Paris, Payot, 1931.
- KRAHE, H. *Lingüística germánica*. Tradução de Maria Teresa Zurdo. Madrid, Ediciones Cátedra, 1977.
- KRISTEVA, J. *História da linguagem*. Lisboa, Edições 70, 1969.
- LAUSBERG, H. *Lingüística románica*. Tradução de J. Pérez Riesco e E. Pascual Rodríguez. Madrid, Gredos, 1970.

- LINO, M. T. R. da F. Observatório do português contemporâneo. *Actas do Colóquio de Lexicologia e Lexicografia*. Lisboa, Universidade Nova, 1990, p. 28-33.
- LOPES, E. *Fundamentos da lingüística contemporânea*. São Paulo, Cultrix, 1976.
- LYONS, J. *Eléments de sémantique*. Paris, Larousse, 1978.
- _____. *Introdução à lingüística teórica*. São Paulo, Cultrix, 1979.
- MALMBERG, B. *As novas tendências da lingüística*. São Paulo, EDUSP, 1971.
- _____. *La phonétique*. Paris, P.U.F., 1971.
- _____. *A língua e o homem*. Introdução aos problemas gerais da lingüística. Rio de Janeiro, Nórdica/São Paulo, Duas Cidades, 1976.
- MARTINET, A. *La lingüística sincrónica*. Madrid, Gredos, 1968.
- _____. *Elementos de lingüística geral*. Lisboa, Sá da Costa, 1970.
- MAURER, F. & STROH, F. *Deutsche Wortgeschichte*. 3 v. Berlin, De Gruyter, 1935.
- MAURER JR., T. H. *A unidade da România ocidental*. São Paulo, FFCL-USP, 1951.
- _____. *Gramática do Latim Vulgar*. Rio de Janeiro, Acadêmica, 1959.
- MEILLET, A. *Caractères généraux des langues germaniques*. Paris, Hachette, 1917.
- MEILLET, A. *Linguistique historique et linguistique générale*. 2 v. Paris, Klincksieck, 1936.
- MEYER-LÜBKE, W. *Einführung in das Studium der romanischen Sprachwissenschaft*. Heidelberg, Carl Winters, 1920.
- _____. *Grammaire des langues romanes*. Tomos I a IV. Tradução de A. Doutrepoint e G. Doutrepoint. Paris, G.E. Stechert & Co., 1923.
- MIAZZI, M. L. F. *Introdução à lingüística românica: história e métodos*. São Paulo, Cultrix, 1976.
- MULLER, Ch. *Initiation à la statistique linguistique*. Paris, Larousse, 1968.

- MÜLLER, W. Conceito de estrangeirismo e dicionário de estrangeirismos. *Problemas da lexicologia e lexicografia*. Tradução de Mário Vilela. Porto, Civilização, 1979.
- NOETHER, G. E. *Introdução à estatística*. Uma abordagem não paramétrica. Rio de Janeiro, Guanabara Dois, 1983.
- PAIS, C. T. *Ensaio semiótico-lingüísticos*. Petrópolis, Vozes, 1977.
- _____. *Introdução à fonologia*. São Paulo, Global, 1981.
- PARRET, H. *Semiotics and pragmatics*. An evaluative comparison of conceptual frameworks. Amsterdam/Philadelphia, John Benjamins Pub. Co., 1983.
- PAULUS, J. *A função simbólica e a linguagem*. Rio de Janeiro, Eldorado/São Paulo, EDUSP, 1975.
- PÉCHEUX, M. *et al.* Analyse du discours. Langue et idéologies. *Langages*, nº 37. Paris, Didier/Larousse, 1975.
- PISANI, V. *Introduzione alla linguistica indeuropea*. Roma, Edizioni Universitarie, 1939.
- PONZIO, A. *Producción lingüística e ideología social*. México, Siglo XXI, 1976.
- POTTIER, B. *Linguistique générale*. Théorie et description. Paris, Klincksieck, 1974.
- PRIETO, L. J. *Etudes de linguistique et de sémiotique générales*. Paris/Genebra, Droz, 1975.
- REY, A. *La terminologie: noms et notions*. Paris, P.U.F., 1979.
- ROBIN, R. *História e lingüística*. São Paulo, Cultrix, 1977.
- ROSENTHAL, E. T. *A língua alemã*. Desenvolvimento histórico e situação atual. São Paulo, Herder, 1963.
- SAUSSURE, F. *Curso de lingüística geral*. São Paulo, Cultrix, 1969.
- STREITBERG, W. *Urgermanische Grammatik. Einführung in das vergleichende Studium der altgermanischen Dialekte*. Heidelberg, s.ed., 1943.

GLOSSÁRIO

No presente glossário, apresentamos alguns termos técnicos empregados neste trabalho, acompanhados da definição que adotamos para os mesmos. Houvemos por bem incluir aqui apenas as acepções de cada termo que efetivamente utilizamos nesta obra, excluindo as demais acepções que tais termos eventualmente possuam na linguagem científica. Além disso, relacionamos abaixo os termos cuja definição por nós adotada difira total ou parcialmente da definição correntemente aceita, ou termos cuja definição varie de um autor para outro, situação em que apresentamos a definição adotada na presente obra. Incluímos também neste glossário aqueles termos que, por serem de nossa criação, não constam dos vocabulários e dicionários especializados.

No corpo da definição de um verbete, pode haver a remissão a outro verbe de do glossário. Essa remissão se dá através de um asterisco colocado após o termo ao qual se quer remeter, ou através da palavra “Veja”, seguida do termo ao qual se deve remeter.

Alogenia, *s.f.* Processo lexicogênico* cujo vocábulo resultante contenha algum elemento importado de outro sistema* lingüístico.

Alogênico, *adj.* Relativo à alogenia*.

Autogenia, *s.f.* Processo lexicogênico* cujo vocábulo resultante é constituído apenas de elementos pertencentes à própria língua, sem nenhum empréstimo a outro sistema* lingüístico, quer em nível de conteúdo, quer em nível de expressão.

Autogênico, *adj.* Relativo à autogenia*.

Diacrônico, *adj.* Diz-se do ponto de vista que leva em consideração a língua em sua evolução histórica.

Discurso, *s.m.* Veja *Norma*.

Empréstimo de restituição, *s.m.* Empréstimo* de significado de um vulgarismo estrangeiro, em que o significante é greco-latino, isto é, totalmente constituído de morfemas de origem grega ou latina não metamorfizados.

Empréstimo de significado, *s.m.* Tipo de empréstimo em que uma língua toma emprestado a outra um significado, substituindo o respectivo significante por um outro, já existente na língua que efetua o empréstimo, ou criado especialmente nessa língua para constituir, com aquele significado, um novo vocábulo.

Empréstimo de significante, *s.m.* Tipo de empréstimo em que uma língua toma emprestado a outra um significante, atribuindo-lhe um novo significado.

Empréstimo total, *s.m.* Tipo de empréstimo em que uma língua toma emprestado a outra um vocábulo como um todo, isto é, significante e significado.

Empréstimo de tradução, *s.m.* Empréstimo* de significado em que o significante é vulgar, isto é, total ou parcialmente constituído de morfemas vernáculos ou metamorfizados.

Etimologia, *s.f.* Estudo científico da origem das palavras, que consiste da busca do *étimo*, palavra considerada como ancestral daquela que constitui o objeto da pesquisa etimológica.

Fractomorfema, *s.m.* Elemento resultante do truncamento do significante de um morfema, em geral por motivo de economia enunciativa.

Germance, *s.m.* Por analogia com o conceito de romance*, definimos germance como o conjunto de dialetos resultantes da fragmentação lingüística e da evolução histórica do germânico* comum. O germance corresponde ao período compreendido entre o início da fragmentação do germânico comum, que se acredita ter ocorrido por volta do séc. V, e a aparição dos primeiros documentos escritos em língua vulgar (séculos VIII–IX). O termo *germance* foi por nós proposto por

razões teóricas e operacionais, pois, na terminologia tradicional, emprega-se o termo *germânico* ora para designar o germânico comum ora para designar os diversos dialetos que descendem desse germânico comum, e que soem ser agrupados em três ramos principais: o germânico ocidental, o germânico setentrional e o germânico oriental. Resulta daí que o termo *germânico* possui significado ambíguo, ao passo que, no domínio da lingüística românica, existe uma clara distinção teórica entre o latim* vulgar e os diversos romances regionais que dele descendem. Desse modo, reservamos a designação *germânico* para indicar o estágio histórico em que ainda havia uma língua comum a todos os povos germânicos, designando por *germance* os diversos dialetos em que se fragmenta o germânico. Temos assim três grupos básicos de germances: os germances ocidentais (anglo, saxão, franco, juto, frísio, etc.), os germances setentrionais (antigo nórdico) e os germances orientais (gótico, burgúndio, etc.). Alguns autores preferem agrupar o que chamamos de germanico setentrional e germance oriental num único ramo.

Germânico, *s.m.* Chama-se germânico (ou germânico comum) a uma língua indo-européia da qual não se possui testemunho direto, e que foi reconstruída a partir da comparação das línguas que dela descendem. Há sérias dúvidas de que tal língua tenha sido algum dia uma língua perfeitamente unitária. Na verdade, a tese mais provável é a de que se tratasse de um conjunto de dialetos indo-europeus que possuíam uma série de isoglossas em comum, falados por tribos geograficamente bastante próximas, que conseqüentemente mantinham estreito relacionamento entre si. Assim, acredita-se que as variedades dialetais dessa espécie lingüística eram pouco numerosas e as diferenças entre essas variedades muito pouco acentuadas, de modo que se pode, sem maior erro, falar de uma língua germânica comum, por oposição aos diversos dialetos que a sucedem historicamente, por nós chamados de *germances**, e que, por força da expansão territorial dos povos germânicos, já acumulam diferenças suficientes para não mais poder ser agrupados numa única língua.

Gramema, *s.m.* Morfema gramatical. Exemplos de gramemas são os afixos (prefixos, infixos, sufixos) e as desinências nominais e verbais.

Grecismo, *s.m.* Vocábulo tomado de empréstimo ao grego, ou engendrado a partir de elementos léxicos gregos.

Grecolatinismo, *s.m.* Vocábulo do qual todos os morfemas, salvo as desinências, são greco-latinos, sem nenhum metamorfismo.

Latim vulgar, *s.m.* Denomina-se por latim vulgar à língua falada pelas classes populares do Império Romano e que diferia em muitos aspectos da língua literária de Roma. Tal língua nunca foi perfeitamente unitária, já que era falada em grande extensão territorial, e, além disso, havia sido sobreposta a diversos substratos lingüísticos, que deixaram inúmeros resíduos na pronúncia, na morfologia, no léxico, etc. Entretanto, enquanto o Império se manteve unido, a comunicação ininterrupta entre as várias províncias impediu uma maior diversidade lingüística, de modo que o latim vulgar pode ser considerado como uma língua. A fragmentação política do Império, no séc. V, com o conseqüente colapso das comunicações entre as províncias, provocou uma progressiva diferenciação lingüística, que destruiu a unidade do latim vulgar. Inicia-se assim o período romance*.

Latinismo, *s.m.* Vocábulo tomado de empréstimo ao latim, ou engendrado a partir de elementos léxicos latinos.

Lexema, *s.m.* Morfema lexical, isto é, aquele que representa um conceito, por oposição ao gramema*, que representa uma noção gramatical. Por exemplo, na palavra *meninhas*, o morfema *menin-* é um lexema, enquanto *-inh-*, *-a-* e *-s* são gramemas.

Lexia, *s.f.* Unidade léxica ao nível do sistema* lingüístico.

Léxico, *s.m.* Conjunto de lexias* de uma língua.

Lexicogênese, *s.f.* 1) Conjunto de processos e mecanismos de criação e renovação das unidades léxicas de uma língua. 2) Ramo da lingüística que se ocupa do estudo de tais processos, focalizando o percurso através do qual as palavras pene-

tram no léxico, e procedendo a uma classificação das unidades léxicas da língua segundo sua natureza genética.

Lexicogênico, *adj.* Relativo à lexicogênese*.

Metamórfico, *adj.* Relativo ao metamorfismo*.

Metamorfismo, *s.m.* 1) Processo de mutação fonético-fonológica, acompanhado de adaptação gráfica correspondente, sofrido por um empréstimo de origem grega ou latina, ou por um grecolatinismo emprestado de uma língua vulgar estrangeira, no âmbito da língua vulgar receptora do empréstimo. 2) Vocábulo resultante desse processo.

Monoisotópico, *adj.* Que ocorre numa única norma discursiva.

Monossêmico, *adj.* Que possui um único significado.

Nacionalização, *s.f.* Processo através do qual um xenismo* passa progressivamente a peregrinismo* e finalmente perde totalmente seu caráter de vocábulo estrangeiro. A rigor, o termo *nacionalização* deveria ser preferível ao usualmente empregado *vernaculização*, pois este último termo indica na realidade a transformação de um vocábulo estrangeiro em vocábulo vernáculo, fenômeno que, partindo da própria definição de vocábulo vernáculo, é impossível. Na presente obra, utilizamos ambos os termos como sinônimos.

Neologia, *s.f.* Processo de introdução de uma nova unidade no léxico de uma língua.

Neologismo, *s.m.* Unidade léxica recém-introduzida na língua, de baixa freqüência de uso e ainda não dicionarizada.

Norma, *s.f.* Conjunto de práticas lingüísticas comuns aos membros de um determinado grupo social; conjunto de atos de fala ou de textos escritos que possuem características sociolingüísticas comuns (por exemplo, a norma jurídica, a norma culta, a norma coloquial, etc.); o mesmo que discurso e universo de discurso.

Palavra, *s.f.* 1) Unidade léxica em nível de fala ou de texto; o mesmo que palavra-ocorrência*. 2) Termo usado nesta obra para neutralizar semanticamente os termos *lexia* e *vocábulo*,

toda vez que a distinção entre os mesmos não seja pertinente ou relevante.

Palavra-ocorrência, *s.f.* Cada uma das ocorrências concretas de um vocábulo* num determinado ato de fala ou texto escrito.

Pancrônico, *adj.* Diz-se do ponto de vista que leva em consideração a língua tanto em sua evolução histórica quanto em seu funcionamento. Enquanto o ponto de vista sincrônico* se preocupa apenas com o funcionamento da língua num determinado momento de sua história e o ponto de vista diacrônico* observa apenas sua evolução histórica, o ponto de vista pancrônico parte da premissa de que o funcionamento e a evolução da língua se articulam dialeticamente: a língua só funciona na medida em que evolui, e só evolui porque funciona.

Peregrinismo, *s.m.* Unidade léxica de origem estrangeira, efetivamente integrada ao léxico de uma língua, mas que ainda conserva seu caráter estrangeiro na grafia e/ou na pronúncia.

Polissêmico, *adj.* Que possui vários significados.

Quase-morfema, *s.m.* Veja *Fractomorfema*.

Refecção, *s.f.* Substituição um vocábulo vulgar ou vulgarizado, de origem latina, por seu equivalente erudito.

Ressemantização, *s.f.* Atribuição de um novo significado a um vocábulo. O mesmo que neologia semântica.

Romance, *s.m.* Conjunto de dialetos resultantes da fragmentação lingüística e da evolução histórica do latim* vulgar, mediante a influência do superstrato germânico. O romance situa-se historicamente entre o fim do Império Romano do Ocidente (séc. V) e o aparecimento dos primeiros textos escritos em língua vulgar (sécs. IX–X). Tradicionalmente, divide-se em romance ocidental, representado pelos dialetos ibero-romance, galo-romance, reto-romance, etc., e romance oriental, representado pelos dialetos ítalo-romance, dálmato-romance, daco-romance, etc.

Sincrônico, *adj.* Diz-se do ponto de vista que leva em consideração a língua em um determinado momento de sua evolução histórica, tomada portanto como um sistema estável.

Sistema lingüístico, *s.m.* Na concepção de Coseriu, sistema lingüístico é a língua tomada num determinado momento de sua evolução histórica.

Universo de discurso, *s.m.* Veja *Norma*.

Vernaculização, *s.f.* Veja *Nacionalização*.

Vernáculo, *s.m.* Vocábulo* herdado por uma língua diretamente da língua que a precedeu historicamente (língua-mãe).

Vocábulo, *s.m.* Unidade léxica em nível de norma*.

Vocábulo culto, *s.m.* Vocábulo* pertencente a uma norma* de discurso especializada (norma técnico-científica, jurídica, religiosa, etc.), qualquer que seja seu estatuto lexicogênico* (erudito, semi-erudito ou vulgar).

Vocábulo erudito, *s.m.* Veja *Grecolatinismo*.

Vocábulo popular, *s.m.* Vocábulo* pertencente ao discurso banal, qualquer que seja seu estatuto lexicogênico* (erudito, semi-erudito ou vulgar).

Vocábulo semi-erudito, *s.m.* Vocábulo* que apresenta hibridismo de elementos eruditos e vulgares. Resulta em geral da composição a partir de um radical grego ou latino e outro vernáculo, da derivação de radical erudito com afixo vulgar, da substituição de um ou mais morfemas do vocábulo greco-latino original por morfema equivalente vernáculo (empréstimo* de tradução), ou ainda de metamorfismo*.

Vocábulo vulgar, *s.m.* Vocábulo* vernáculo ou derivado de vocábulo vernáculo, ou, ainda, vocábulo composto de radicais vernáculos.

Vulgarismo, *s.m.* Vocábulo* do qual pelo menos um dos morfemas constituintes, à exceção das desinências, é vernáculo ou resultante de metamorfismo.

Vulgarismo alogénético, *s.m.* Vocábulo* resultante do empréstimo de um vulgarismo estrangeiro.

Xenismo, *s.m.* Vocábulo* estrangeiro, não pertencente à língua, que ocorre num determinado ato de fala ou texto escrito.

ILUSTRAÇÕES

FIGURA 1 – A dianorma léxica	175
FIGURA 2 – Curva de Gauss	178
FIGURA 3 – Representação esquemática do procedimento para obtenção do vocabulário fundamental	184
FIGURA 4 – Esquema de difusão dos neologismos – centros difusores de neologismos e áreas de influência	210
GRÁFICO 1 – Discurso científico	194
GRÁFICO 2 – Discurso jornalístico	194
GRÁFICO 3 – Índice de Vulgaridade.....	196

LISTA DE TABELAS

TABELA 1 – Discurso científico.....	188
TABELA 2 – Discurso jornalístico	189
TABELA 3 – Comportamento lexicogênico em ambos os discursos.....	190
TABELA 4 – Português científico	197
TABELA 5 – Francês científico.....	198
TABELA 6 – Italiano científico.....	198
TABELA 7 – Inglês científico	199
TABELA 8 – Alemão científico	199
TABELA 9 – Português jornalístico	200
TABELA 10 – Francês jornalístico.....	201
TABELA 11 – Italiano jornalístico.....	201
TABELA 12 – Inglês jornalístico	202


TABELA 13 – Alemão jornalístico.....	202
TABELA 14 – Discurso científico.....	203
TABELA 15 – Discurso jornalístico.....	204
TABELA 16 – O Índice Etimológico	206
TABELA 17 – IE: valores secundários.....	207

SOBRE O AUTOR

Aldo Bizzocchi é bacharel em Lingüística e doutor em Semiótica e Lingüística Geral pela Universidade de São Paulo. Vice-coordenador do Curso de Mestrado em Comunicação e Ciências da Linguagem da Universidade Paulista e professor do Curso de Mestrado em Comunicação Social da Fundação Cásper Líbero.

SOBRE A OBRA

Este texto foi apresentado como Tese de Doutorado ao Curso de Pós-Graduação em Lingüística da Faculdade de Filosofia, Letras e Ciências Humanas da USP em dezembro de 1994, onde obteve a nota máxima (10,0 com distinção e louvor). Participaram da banca examinadora Cidmar Teodoro Pais (orientador), Jeni Silva Turazza, Antonieta Laface, Guiomar Fanganiello Calçada e Alcebíades Fernandes Júnior.



A Europa Ocidental é talvez o melhor exemplo de povos étnica e lingüisticamente diferentes que, no entanto, devido a uma influência cultural comum - da Grécia e Roma antigas - constituíram uma única grande civilização. A influência unificadora da cultura clássica reflete-se na ideologia desses povos e, conseqüentemente, em suas línguas.

Em *Léxico e Ideologia na Europa Ocidental*, Aldo Bizzocchi analisa a história dessas línguas sob um novo enfoque, buscando explicar a dinâmica da criação e renovação do léxico das línguas européias ocidentais de um ponto de vista ao mesmo tempo neológico e etimológico, isto é, com base na origem e nos processos de produção das unidades lexicais a partir da tensão entre o *clássico* (de influência greco-latina) e o *vulgar*.