

REITOR
Prof. Dr. Jorge Nagle
VICE-REITOR
Prof. Dr. Paulo Milton Barbosa Landim

CENTRO DE PUBLICAÇÕES CULTURAIS E CIENTÍFICAS
DIRETOR: Prof. Dr. Fernando de Azevedo Corrêa

Conselho Consultivo:
Presidente: Prof. Dr. Fernando de Azevedo Corrêa

Membros: Prof. Dr. Antônio Christofolletti
Prof. Dr. Dilermando Perecin
Prof. Dr. Francisco Habermann
Prof. Dr. Irineu de Moura

Equipe Técnica:
Elga de Souza Pastore
Maria Sylvia C.A. Corrêa
Vanderli P.A. Horikiri

REITOR

Prof. Dr. Jorge Nagle

VICE-REITOR

Prof. Dr. Paulo Milton Barbosa Landim

CENTRO DE PUBLICAÇÕES CULTURAIS E CIENTÍFICAS

DIRETOR: Prof. Dr. Fernando de Azevedo Corrêa

Conselho Consultivo:

Presidente: Prof. Dr. Fernando de Azevedo Corrêa

Membros: Prof. Dr. Antônio Christofolletti

Prof. Dr. Dilermando Perecin

Prof. Dr. Francisco Habermann

Prof. Dr. Irineu de Moura

Equipe Técnica:

Elga de Souza Pastore

Maria Sylvia C.A. Corrêa

Vanderli P.A. Horikiri

UNIVERSIDADE ESTADUAL PAULISTA

ALFA

Revista de Linguística
Lexicografia e Lexicologia

Organizado por Maria Tereza Camargo Biderman

ISSN 0002-5216
ALFAD5

ALFA	São Paulo	V-28 — Suplemento	p. 1-149	1984
------	-----------	-------------------	----------	------

A correspondência e os artigos para publicação deverão ser encaminhados a:
Correspondence and articles for publication should be addressed to:

ALFA: Revista de Lingüística

Instituto de Letras, Ciências Sociais e Educação
Estrada Araraquara-Jaú, Km 1 — C.P. 174
14800 — Araraquara — SP — Brasil

Comissão de Redação

Alceu Dias Lima
Clélia Cândida Abrei Spinardi Jubran
Nildemir Ferreira de Carvalho
Onosor Fonseca
Telmo Correia Arrais

Diretor da Revista

Telmo Correia Arrais

Publicação anual / Annual publication
Solicita-se permuta / Exchange desired

ALFA: Revista de Lingüística. (Universidade Estadual Paulista)
São Paulo, SP — Brasil, 1962-1977, 1980-

Suplemento em: 1984, 28.
1962-1977, 1-23

Publicação interrompida de 1978-1979.
1980-1984, 24-28

ISSN 0002-5216
ALFAD5

Os artigos publicados na ALFA são indexados por:
The articles published in ALFA are indexed by:
Bibliographie Latinoamericaine d'Articles, Clase-Cich-Unam, Indice de Revista, MLA-International Bibliography of Books and Articles on Modern Languages and Literature, Revista de Sumários, Sociological Abstracts, Sumários de Educação.

SUMÁRIO / CONTENTS

APRESENTAÇÃO	V
A CIÊNCIA DA LEXICOGRAFIA. The science of Lexicography. M. T. C. Biderman	1
O DICIONÁRIO PADRÃO DA LÍNGUA. The pattern dictionary of a language. M. T. C. Biderman	27
LÉXICO E DICIONÁRIO. Lexicon and dictionary. J. Rey-Debove; Tradução de C. B. de Moraes	45
SEMÂNTICA E LEXICOGRAFIA. Semantics and Lexicography. J. Casares; Tradução de B. L. Feijóo-Hoyos	71
DEFINIÇÃO LEXICOGRÁFICA EM SEMÂNTICA DESCRITIVA. Lexicographic definition in Descriptive Semantics. U. Weinreich; Tradução de M. C. P. B. Lima	103
A INTEGRAÇÃO DOS NEOLOGISMOS POR EMPRÉSTIMO AO LÉXICO PORTUGUÊS. L'intégration des néologismes par emprunt au lexique portugais. I. M. Alves	119
A PREFIXAÇÃO NEOLÓGICA NO VOCABULÁRIO DA PROPAGANDA CONTEMPORÂNEA. Neological prefixation in the vocabulary of contemporary advertising. R. F. Pereira	127
GLOSSÁRIO M. T. C. Biderman	135
ÍNDICE DE ASSUNTOS	145
SUBJECT INDEX	147
ÍNDICE DE AUTORES / AUTHOR INDEX	149

APRESENTAÇÃO DESTA COLETÂNEA

Desde que comecei a trabalhar em Lexicografia, originalmente na elaboração de um “Dicionário Infantil” (1977) que posteriormente se converteu no “Dicionário Fundamental do Português (ou “Dicionário Escolar da Língua Portuguesa”), venho estudando os variados aspectos teóricos, técnicos e metodológicos da confecção de dicionários. Por causa dessa experiência pareceu-me útil organizar uma coletânea sobre a matéria, em virtude do autodidatismo que impera nos trabalhos lexicográficos produzidos no Brasil; julgo que já chegou o momento de a universidade brasileira ocupar-se de uma tarefa tão básica em qualquer sociedade moderna civilizada, a saber, elaborar os dicionários de que essa sociedade necessita e fazê-lo com competência técnica e científica.

O objetivo desta coletânea de artigos sobre a ciência e a técnica dos dicionários, foi o de montar um pequeno manual que pudesse servir de iniciação e/ou orientação para jovens pesquisadores que possam vir a trabalhar nesta nobre ciência — a Lexicografia. Assim, doravante, poderão contar com um arcabouço teórico publicado em língua portuguesa.

Um volume como este deveria ocupar-se dos problemas lexicográficos fundamentais. Eis por que é preciso tratar também do Léxico e da Lexicologia, pois a ciência dos dicionários é alicerçada na ciência da palavra. Não tenho a pretensão de cobrir todas as questões que este assunto exigiria. Contudo, creio que alguns tópicos básicos estão abordados neste volume.

Como a Lexicografia é uma ciência velha que tem contado com o concurso de especialistas de várias línguas e culturas, procurei reunir um pequeno número de textos representativos. Alguns deles foram escritos por insignes autoridades na matéria, é o caso, por exemplo, dos trabalhos da francesa Josette Rey-Debove e do espanhol Julio Casares.

A estrutura deste volume consta de: a) quatro textos escritos originalmente para ele: dois de minha autoria, um da Prof.^a Ieda Maria Alves e um do Prof. Rony Farto Pereira; b) três traduções, excertos de obras ou artigos, produzidos em outras línguas, e que foram traduzidos por colegas da UNESP. Esses trabalhos podem ser considerados estudos básicos em matéria de Lexicografia. Cada um deles enfoca um ângulo diferente. Por conseguinte, eles se completam mutuamente. Devo acrescentar ainda, antes de passar a considerações de detalhe, que os textos mais substanciais representam a Lexicografia européia e a sua antiga tradição; somente o artigo de Weinreich acrescenta algo do mundo acadêmico americano, cujos lingüistas parecem desinteressados da nossa ciência. Nos comentários que seguem sobre cada artigo, acompanharei a ordenação dos estudos e/ou capítulos.

O meu trabalho “A Ciência da Lexicografia” tenta abordar alguns problemas genéricos desse domínio. Esbocei um amplo painel da problemática, da técnica e dos métodos da Lexicografia, pensando no intelectual que pretenda trabalhar sobre o tesouro léxico da nossa língua materna. Ocupei-me primeiro da história da Lexicografia. O espaço reduzido de um artigo permitiu-me apenas aflorar a matéria. A seguir, forneço informações sobre os principais tipos de dicionários. Na secção seguinte procurei mos-

trar como o computador pode ajudar os lexicógrafos na sua árdua tarefa, compilando, classificando e ordenando dados léxicos e contextuais; dessa forma libera energias dos dicionaristas para o trabalho mais complexo da seleção de dados e elaboração de verbetes.

No segundo artigo examino os vários problemas e métodos relacionados com os dicionários unilingües, especialmente o dicionário-padrão da língua. Trato das etapas e técnicas de confecção de um dicionário como a redação de verbetes. Abordo também um dos problemas centrais na elaboração de um dicionário: a definição léxico-gráfica assunto ao qual se dedicará um espaço maior no artigo de U. Weinreich: “A definição lexicográfica na semântica descritiva”. Também analiso superficialmente a questão da polissemia e da homonímia, da sinonímia e da antonímia, matérias que serão igualmente examinadas em detalhe no artigo de Julio Casares: “Semântica e Lexicografia”.

O artigo de Josette Rey-Debove foi extraído do livro: *Le langage* (Les dictionnaires du savoir moderne. Centre d’Etude et de Promotion de la Lecture. Paris, 1973, pp. 82-108). A autora é uma das principais lexicógrafas da equipe autora dos dicionários *Le Robert: Grand Robert, Petit Robert, Micro Robert*. Foi traduzido pelo Prof. Clóvis Barleta de Moraes.

Inicialmente Rey-Debove trata das relações entre o léxico e a gramática; a seguir, define alguns conceitos operatórios básicos da Lexicologia: a palavra, a unidade léxica, a lexia, o morfema, uma classe aberta (o léxico), uma classe fechada (os morfemas), o repertório léxico de uma língua etc. Mostra como o léxico é o domínio da língua menos especificamente lingüístico por ele se reportar ao universo referencial, físico e cultural, em que se situa o homem. O conjunto de signos que compõem o léxico, conferem a ele um estatuto semiótico, criando um sistema de mundo. Rey-Debove discute também o problema da designação dos referentes dentro do universo físico-cultural, mostrando como se relacionam as palavras e coisas. Outro problema analisado é o dos limites do léxico, já que ele é um conjunto aberto com possibilidades infinitas de expansão. Vários outros tópicos da Lexicologia são aí tratados: os vocabulários especiais, o léxico e a evolução da língua, o significado das palavras como função dependente do contexto, a estruturação do léxico e das significações, os dicionários como nossa única idéia sobre o léxico, o que são os dicionários.

O lexicógrafo Julio Casares é autor de um dos raros dicionários (não alfabético), baseado na estruturação dos conceitos. *Diccionario Ideológico de la Lengua Española* (1942). De sua autoria é o trabalho inserido nesta coletânea “Semântica e Lexicografia” (em três unidades: I, II e III). Foi extraído do livro *Introducción a la Lexicografía Moderna* (Madrid, C. S. I. C., anexo III, 1950; capítulos II, III e IV da segunda parte, pp. 50-100) e traduzido pela Prof.^a Balbina Lorenzo Feijóo Hoyos. Aí se discutem os controvertidos processos de lexicalização em uma língua — no caso, a espanhola — idioma muito próximo do português. Casares trata também do problema da lematização, ou seja, a decisão técnica de escolher como entrada de um dicionário, uma ou outra forma vocabular, o que envolve controvérsias permanentes em meio aos lexicólogos sobre as lexias (palavras) complexas e o como e o quando se dá a categorização lexical de um polinômio vocabular. Esse problema é ampliado por causa da tradição caótica de muitas grafias, particularmente no caso de “locuções vocabulares”. Ensina como se devem ordenar as acepções (significados, valores) de um vocábulo, dada a intrincada malha de significações, criada pela evolução semântica que sofrem as palavras. Advoga as vantagens e as virtudes de um dicionário que tivesse um índice de frequência do uso de cada palavra, ou de cada acepção de uma palavra.

O artigo “Definição lexicográfica em Semântica descritiva”, de Uriel Weinreich, foi tirado de *Problems in Lexicography* (F. Householder & S. Saporta, eds., Bloomington, Indiana University Press, 1967). Foi traduzido pela Prof.^a Maria Cecília P. Barbosa Lima. Depois de opor a descrição semântica à descrição lexicográfica, Weinreich se concentra no problema específico da definição lexicográfica. Analisa os vários aspectos teóricos do ato de definir um termo para um dicionário e os tipos de critérios que podem ser adotados. Uma questão de grande relevância é a do tipo de metalíngua que se deve adotar na redação de um dicionário e a metodologia empregada para definir o “definiendum”, podendo-se empregar várias estratégias: o método analítico, o método sintético, o método denotativo, o método ostensivo ou de mostração, o método implicativo, ou contextual. O método escolhido dependerá da natureza do termo a ser definido: um referente concreto, uma noção abstrata, uma ação ou processo verbal, um instrumento gramatical etc. A sinonímia e a antonímia usadas amplamente nas definições, têm também grande importância lexicográfica. O lexicógrafo, ou a equipe de lexicógrafos que trabalham na confecção de um dicionário, nunca se deve esquecer que as suas definições devem valer para toda a comunidade lingüística a que ele se destina e assim usarem a linguagem comum a todos e não o(s) seu(s) idioletos(s) particular(es).

O léxico é um sistema aberto em expansão. Numa língua viva de civilização é o domínio lingüístico que mais sofre alterações, sobretudo em função de novas criações vocabulares — os neologismos. No português contemporâneo brotam neologismos quase diariamente. Muitos deles serão descartados posteriormente como tantos produtos modernos descartáveis. Outros, porém, se incorporarão definitivamente ao léxico português fazendo-o crescer. Entre os problemas a serem considerados em Lexicologia e em Lexicografia, o da problemática neológica é certamente um dos mais importantes sobretudo nos tempos contemporâneos. Os meios de comunicação de massa (MCM) veiculam uma massa enorme de neologismos, a maior parte deles provenientes das ciências e das técnicas. Os MCM incumbem-se de banalizá-los e integrá-los na língua comum. Os dois últimos artigos desta coletânea são dedicados ao estudo dos neologismos no português contemporâneo. Em ambos, os autores utilizaram MCM de grande penetração no Brasil contemporâneo como fontes para coletar o seu arquivo de formações neológicas.

O trabalho da Prof.^a Ieda Maria Alves “A integração dos neologismos por empréstimo ao léxico português” constitui o penúltimo artigo deste livro. A colega vem trabalhando há alguns anos sobre esse assunto, eterno cavalo-de-batalha para um dicionarista. Aliás, a problemática dos neologismos foi a matéria-prima da tese de doutoramento e também é o tópico da tese de livre-docência da colega. O presente artigo começa por estabelecer uma tipologia dos neologismos. Depois mostra como se tem feito a integração à fonética e à ortografia portuguesas dos mais recentes neologismos incorporados à nossa língua. O caso mais freqüente é o de empréstimos vocabulares feitos a outras línguas, máxime o inglês. O seu corpus baseou-se num amplo levantamento feito em periódicos, jornais e revistas de grande circulação no Brasil contemporâneo (*Veja, Visão, Isto é, Folha de S. Paulo, O Estado de S. Paulo* etc.). Tal levantamento evidenciou a freqüência desses neologismos estrangeiros. São também discutidos os problemas criados no processo de categorização gramatical segundo os padrões léxico-gramaticais da língua portuguesa, ao serem importados e incorporados esse, vocábulos estrangeiros. Os neologismos confirmam uma afirmação que se fez acima (cf. artigo de Rey-Debove) de que o léxico é um sistema aberto e em expansão. Eles só serão consagrados no uso vernáculo quando passarem a fazer parte de um dicionário conceituado da língua. O lexicógrafo constitui, portanto, o juiz que dá o seu aval a

um neologismo, em nome da comunidade lingüística, a fim de integrar esse vocábulo no tesouro léxico do idioma.

No último artigo o Prof. Rony Farto Pereira analisa um corpus de neologismos, recolhido em matéria publicitária das revistas *Veja* e *Isto é*. Constatou que a derivação prefixal é o processo de criação léxica mais utilizado, superior mesmo ao da sufixação. Concluiu que o número de neologismos gerados com recursos primários do português é bem superior aos vocábulos com base em estrangeirismos, pelo menos na linguagem publicitária.

A Lexicografia brasileira digna do nome é muito pobre, com raras exceções: Antenor Nascentes, Antônio Houaiss, Francisco Fernandes, Aurélio Buarque de Holanda Ferreira, Hamílcar de Garcia e Antônio Geraldo da Cunha entre os modernos; entre os antigos só cabe citar o velho Moraes. Muito, senão quase tudo está por fazer no domínio desta ciência, especialmente quando confrontamos a nossa magra produção com a variedade e a qualidade da safra de dicionários em línguas como o francês, o inglês, o espanhol, o italiano, o alemão etc., tanto no passado como no presente. Não vou listar aqui o rol das tarefas que precisam ser executadas em Lexicografia no Brasil. Vou limitar-me a algumas que me parecem mais urgentes.

1) Um *Dicionário padrão da Língua Portuguesa contemporânea*, efetivamente baseado no uso oral e escrito da língua que falamos e escrevemos hoje. Deve ter o rigor técnico de dicionários como os franceses da série *Le Robert*, os ingleses tipo Oxford, Longman, Webster.

2) Um *Dicionário de frequência do português brasileiro contemporâneo* para ser uma fonte de referência permanente sobre os índices de frequência e de uso das palavras, dependendo do registro lingüístico e do gênero literário, nos moldes dos dicionários de frequência de Alphonse Juilland e sua equipe (francês, espanhol, rumeno, italiano e português). Tal dicionário teria uma enorme utilidade didática, sobretudo na elaboração de livros escolares. Assim se evitaria uma atitude comum no ensino da língua materna: a insistência no aprendizado de vocabulário obsoleto, raro, exclusivamente literário ou técnico, típico apenas de um registro, bem como de formas gramaticais raríssimas etc.

3) Um *Dicionário de verbos e regimes* que corrija os erros e os equívocos de Francisco Fernandes e que atualize as suas fontes. Uma equipe de professores do Instituto de Letras, Ciências Sociais e Educação do Campus de Araraquara, UNESP, liderada pelo Prof. Francisco da Silva Borba, já está trabalhando nesta tarefa.

4) Uma série de dicionários técnicos e científicos dos variados domínios da Tecnologia e da Ciência modernas.

Editando o presente livro eu sonhava dar a partida nesta obra enorme que discriminei acima, bem como fornecer subsídios para os grupos que se vierem a constituir, animados do ideal de fazer os dicionários de que as culturas brasileira e portuguesa necessitam.

Maria Tereza Camargo Biderman

A CIÊNCIA DA LEXICOGRAFIA

Maria Tereza Camargo BIDERMAN*

RESUMO: Três tópicos são examinados: 1) História sucinta da Lexicografia de duas línguas latinas (espanhol e francês) e do português. São avaliados os principais dicionários dessas línguas do século XVI ao século XX. 2) Tipologia de obras lexicográficas. São indicados e comentados os principais tipos de dicionário existentes nas línguas latinas e no inglês. 3) O uso do computador na Lexicografia contemporânea. Essa máquina revolucionou a Lexicografia, podendo executar tarefas básicas e enfadonhas como: compilar, classificar e ordenar dados léxicos e contextuais para a confecção de dicionários e depois recuperá-los facilmente e com rapidez.

UNITERMOS: Lexicografia; thesaurus; dicionário, dicionário histórico; dicionário etimológico; dicionário ideológico; dicionário técnico; dicionário científico; dicionário de frequência; banco de dados léxicos; corpus; índices verborum; concordância; dicionário de máquina.

I. HISTÓRIA SUCINTA DA LEXICOGRAFIA DE ALGUMAS LÍNGUAS LATINAS E DO PORTUGUÊS

A antigüidade não produziu obras lexicográficas no sentido que hoje damos a esse termo. Os únicos trabalhos de cunho vagamente lexicográfico daquelas eras são os glossários, sobretudo os produzidos pela escola grega de Alexandria e, entre os latinos, o *Appendix Probi*. Esses precursores do moderno lexicógrafo eram, na verdade, filólogos ou gramáticos, preocupados com a compreensão de textos literários anteriores, ou com a correção de "erros" lingüísticos. Os filólogos alexandrinos, p.ex., buscaram elaborar léxicos e glossários sobre os textos homéricos para a sua melhor compreensão. O gramático romano Varrão (I séc. A.C.) em *De lingua latina* tentou fornecer dados de natureza semântica e etimológica sobre algumas palavras latinas.

Na Idade Média valeria a pena lembrar apenas as *Etimologias* de Santo Isidoro de Sevilha (570-636) e alguns glossários. As *Etimologias* têm escasso ou nulo valor científico e lingüístico, sendo muito fantasistas. Baseiam-se numa concepção mística do mundo e da linguagem que tende a referir a língua e as palavras a um sistema de significação que se reporta a Deus, adulterando-lhes, pois, a natureza. Na verdade, documentam o mundo cultural da Idade Média e sua concepção de universo.

Entre os glossários podemos citar o *Glossário de Reichenau* (séc. VIII D.C.) e o *Glossário de Cassel* (séc. IX D.C.) em terras do império carolíngio. Os dois *Glossários de Reichenau* contêm pouco mais de 2.000 vocábulos. São listas de palavras tiradas da *Vulgata* (versão latina da bíblia) de difícil compreensão para a época do autor, traduzi-

* Departamento de Linguística - Instituto de Letras, Ciências Sociais e Educação - UNESP - 14800 - Araraquara - SP.

das no vernáculo românico da região. *O Glossário de Cassel* (265 palavras) é similar; trata-se de tradução do latim para o vernáculo germânico da região.

Também em terras hispânicas foram elaborados alguns glossários: as *Glosas Emilianenses e Silenses* (séc. X ou XI).

A verdadeira lexicografia, porém, só se vai iniciar nos tempos modernos. Os primeiros dicionários espanhóis foram: o *Universal Vocabulario* de Alonso de Palencia (1490) e os vocabulários *Latino Español* (1492) e *Español Latino* (1495) de Antônio de Nebrija, autor também da primeira gramática espanhola. Aliás, no século XVI na Europa, a lexicografia que se estava formando e desenvolvendo compreendia apenas os dicionários bilingües como esses de Nebrija. Quando o homem renascentista começou a ampliar os seus horizontes culturais abandonando de vez a sua reclusão medieval dentro de sua própria cultura, descobriu a necessidade de aprender línguas, evidentemente as línguas européias mais faladas na época (século XVI). Além da consciência adquirida da distância entre o latim e as línguas vernáculas do seu tempo, o homem renascentista precisava de outros instrumentos de intercâmbio lingüístico num mundo que se abria para um novo diálogo e trocas entre as jovens nações européias. Assim, multiplicam-se os dicionários bilingües na Espanha, na França, na Itália, em Portugal, bem como as gramáticas de cada uma das línguas que se tornaram oficiais para as nações-estado da Europa no século XVI.

Os dicionários seicentistas eram cheios de lacunas e os dicionaristas da época copiavam-se uns aos outros.

A lexicografia monolingüe surge e se desenvolve ao longo do século XVII, aperfeiçoando, aos poucos, as suas técnicas. O *Tesoro de la Lengua Castellana* de Covarrubias é de 1611. Tem muito aspectos positivos até hoje. O dicionário da Academia Espanhola - *Diccionario de Autoridades* — iniciou sua publicação em 1739. Terá sucessivas edições nos séculos XVIII, XIX e XX. A última é de 1983. Foi-se aprimorando através dos séculos. É de tipo seletivo e normativo. Uma sigla que o designa geralmente D.R.A.E. (*Diccionario de la Real Academia Española*).

No século XVII, o “grand siècle” da civilização francesa, vários são os dicionários monolingües do francês de boa qualidade para a época: Richelet (1680), Furetière (1690), o dicionário da Academia Francesa (1694). No século XVIII além da nova versão do Dicionário da Academia (1718) e do excelente *Dictionnaire de Trévoux*, um dos mais importantes feitos lexicográficos da época foi a obra dos enciclopedistas franceses. Com eles nasceu o modelo de enciclopédia que hoje adotamos, ou seja, um repertório geral dos conhecimentos humanos. A despeito de ter ficado aquém das aspirações de Diderot devido à desigualdade entre os seus colaboradores, é um trabalho notável para o seu tempo. No verbete *dicionário* eis a definição escrita por Diderot: “Num dicionário da língua francesa, há principalmente três coisas a considerar: a significação das palavras, o seu uso e o tipo de palavras que devemos incluir neste dicionário. A significação das palavras se estabelece por boas definições; seu uso, por uma excelente sintaxe; seu tipo, enfim, pelo próprio objetivo do dicionário. A esses três objetivos principais, pode-se acrescentar três outros subordinados a esses: a quantidade ou a pronúncia das palavras, a ortografia e a etimologia” (5, p. 102). Essa definição é válida ainda para os nossos dias. No século XIX amplia-se o número de obras lexicográficas francesas e apura-se a sua qualidade. Alguns dicionários da época: Laveaux, Raymond, Landais, Academia (1835), Littré (1872), Larousse (1866-1876), o *Dictionnaire Général* de Hatzfeld e o dicionário medieval de Godefroy. Dentre esses o Littré pode ser considerado uma obra-prima da lexicografia francesa, mesmo para os modernos critérios lexicográficos. Littré dedicou-se monacalmente à confecção do seu dicionário

durante 30 anos. Foi um inovador para o seu tempo; embora o seu exemplário só incluía autores anteriores a 1830 (os clássicos para Littré), constitui um modelo de repertório léxico e de escolha de citações como ilustração das palavras-entrada. O dicionário de Pierre Larousse teve dimensão considerável: 17 volumes. O seu *Grand Dictionnaire Universel du XXème Siècle* mostra uma vocação mais de enciclopedista do que de dicionarista — caso de Littré — de quem se distingüia também por ser menos purista e mais liberal. Esse “dicionário universal” constitui um notável repositório de informações sobre a sua época.

No século XX, na lexicografia francesa, existe uma grande abundância e variedade de dicionários e enciclopédias, fenômeno que se registra em várias das grandes culturas e civilizações contemporâneas. Assim ocorre na italiana, na alemã, na espanhola, na luso-brasileira, na anglo-americana, etc. Atualmente a lexicografia se expande e assume modalidades várias em função do vasto público, das grandes massas sequiosas de informações sobre a sua língua, sobre as línguas estrangeiras e sobre o universo. O dicionário se tornou um objeto de consumo de primeira necessidade.

Quanto à lexicografia francesa contemporânea, convém assinalar os vários dicionários Larousse, Robert, etc. A série dos Larousse vai desde obras elementares como *Mon premier Larousse, Nouveau Larousse des débutants* (1977), *Dictionnaire du vocabulaire essentiel* (1963) até o *Grande Larousse Encyclopédique* de 10 volumes. Como dicionários da língua os *Robert (Grand Robert, Petit Robert, Micro Robert)* são certamente os melhores e modelo exemplar de trabalho lexicográfico. Contêm abundantes sinônimos e excelente exemplificação. Os significados da palavra no interior do verbete são classificados partindo-se do sentido mais antigo. Seu exemplário (moderno e contemporâneo sobretudo) é excelente. O *Grand Robert* procurou ser uma espécie de tesouro do francês contemporâneo e o *Micro Robert* (30.000 verbetes) é um ótimo instrumento de uso escolar e ideal para utilização no ensino do francês como segunda língua.

No domínio das enciclopédias deve-se apontar também algumas obras magistrais: *Encyclopédie Française, Encyclopédie de la Pléiade, Encyclopédie du Savoir Moderne*, etc.

Atualmente o “Institut de la Langue Française” está executando a gigantesca tarefa de compilar o maior arquivo de dados léxicos da língua francesa de todos os tempos, utilizando recursos computacionais e documentais muito modernos, bem como uma grande equipe de lexicógrafos, analistas de sistema, documentalistas e outros. O *Trésor Général des Langues et des Parlers Français* deve ser o maior acervo documental jamais compilado sobre uma língua. Só os fundos documentais sobre o francês dos séculos XIX e XX totalizam cem milhões de palavras (ocorrências). O *Dictionnaire de la Langue Française du 19ème. et du 20ème. Siècle* (nove volumes publicados e mais seis a publicar) constitui certamente o mais monumental trabalho lexicográfico já empreendido e executado sobre uma língua natural. Contém todo tipo de informação lingüística sobre a palavra-entrada, a partir de dados documentais reais que são citados para abonar cada significado, cada uso referido. O verbete inclui construções sintáticas e seus valores semânticos acompanhados de abonações totalmente identificadas: autor, nome da obra, data da publicação, página. Além disso, o dicionário fornece uma seqüência cronológica dessas abonações. Também indica os usos da palavra conforme o registro lingüístico: discurso oral ou escrito, discurso geralmente literário, especificamente literário, literário e raro, antiquado, raro. Informa ainda em que tipo de linguagem especial o termo é usado e qual o seu significado nesse registro, que pode ser: música, liturgia católica, fotografia, etc. Contém também informações sobre a pronúncia, a ortografia, a etimologia e a história da palavra. Finalmente, aponta dados quantitativos sobre

a palavra: a frequência absoluta literária e a frequência relativa literária, dados estatísticos esses que se distribuem em quatro segmentos cronológicos: primeira metade do século XIX; segunda metade do século XIX; primeira metade do século XX; segunda metade do século XX.

A lexicografia portuguesa tem uma história mais pobre que a francesa. O melhor dentre os mais antigos dicionários do português é o bilingüe de Rafael Bluteau — *Vocabulário Portuguez e Latino* em 8 volumes, Coimbra 1712-1721. O dicionário do Padre Bluteau é obra muito rara. Trata-se de um dicionário bilingüe português-latim que contém muita informação e bastante variada sobre essas duas línguas. Foi escrito para um falante do português. Tem características enciclopédicas com numerosos detalhes sobre a realidade e o mundo, evidenciando a vasta cultura do Padre Bluteau. Um dos méritos desse dicionário é o de alistar todos os autores portugueses que compuseram o corpus que forneceu o exemplário das abonações dos verbetes. O dicionarista indica o autor, a(s) obra(s), o local e data da impressão. Não é apenas um dicionário bilingüe cujo objetivo seria fornecer a palavra ou expressão latina que traduzisse um termo português; na verdade, Bluteau elaborou um trabalho misto, pois a parte relativa à língua portuguesa constitui praticamente um dicionário da língua portuguesa.

Um dos mais antigos dicionários unilingües do nosso idioma é o *Elucidário de palavras e frases que em Portugal antigamente se usarão (sic) e que hoje regularmente se ignorão (sic): obra indispensável para entender sem erro os documentos mais raros, e preciosos, que entre nós se conservão (sic)* de Frei Joaquim de Santa Rosa de Viterbo, Lisboa, 1798. Como Frei Viterbo se preocupou com facilitar ao leitor de textos antigos ou arcaizantes, a compreensão dessa linguagem, sua lista de verbetes registra sobretudo vocábulos caídos em desuso, ou com valores semânticos alterados. Por exemplo: *badulaque* = guisado de carne cortado em miúdos; *barro* = lugar pequeno, quinta, casa de campo; *cahimento* = diminuição, falta, queda, desfalecimentos. Raramente indica a classe gramatical da palavra, ou qualquer outra informação gramatical no corpo do verbe. Às vezes o faz como no caso dos advérbios: *cha* adv... *chus* adv... Também há poucos verbos incluídos; a maior parte da nomenclatura é constituída de substantivos. Alguns verbetes mereceram exaustivo tratamento, certamente por identificarem referentes da cultura medieval, já pouco conhecidos ao tempo de Viterbo: *cavallaria*, *cavalleiro*. *Charidade*, p.ex., tem onze entradas diferentes, numeradas com algarismos romanos. Quanto à palavra *cruz*, talvez por causa da condição de religioso do autor, mereceu carinho especial: 35 colunas e meia, o mais amplo texto do dicionário; além disso, inclui 23 desenhos ilustrativos, identificando os diferentes símbolos e tipos de cruz. A rigor o *Elucidário* de Viterbo é mais um glossário do que um dicionário.

Entre os mais abalizados dicionários do passado temos o Moraes, que leva o título de "*Dicionário da Língua Portuguesa* recopilado dos vocabulários impressos até agora e nesta segunda edição novamente emendado e muito acrescentado" por Antônio de Moraes e Silva. A primeira edição de 1789 é obra rara, quase incontrolável. Moraes considerou esta primeira edição como obra do Padre Rafael Bluteau, visto que escreveu no frontespício desse dicionário: "... composto pelo Padre Rafael Bluteau, reformado e acrescentado por Antônio de Moraes Silva, natural do Rio de Janeiro." A segunda edição de 1813, porém, Moraes já atribui a si próprio, antepondo-lhe o seu nome. Essa edição é rara. Felizmente Laudelino Freire fez dela uma reprodução fac-símile em 1922 no Rio de Janeiro. É a versão mais encontrada do velho Moraes em bibliotecas públicas e particulares, embora também seja obra rara. Seguiram-se numerosas edições do Moraes, todas elas, porém, inferiores à edição de 1813 (2.ª ed.).

O dicionário de Morais (2.^a.ed.,1813) constitui um marco na lexicografia de língua portuguesa. É o primeiro dicionário de uso da língua, muito avançado para os padrões lexicográficos da época. Apesar de ter-se baseado na obra do Padre Bluteau, sobretudo na primeira edição, na segunda edição Morais libertou-se de seu modelo, ampliou consideravelmente a obra com respeito ao número de verbetes, incluídos, e mais que isso, apurou o seu trabalho lexicográfico. Omitiu informações de tipo enciclopédico incluídas no Bluteau, revelando consciência de que um dicionário da língua não é uma enciclopédia. No prólogo Morais informa o leitor como executou o seu trabalho, de quais critérios se serviu, repassando problemas como: o corpus usado na abonação dos verbetes, a escolha das entradas, a elaboração do verbete, a ortografia. Depois de render um preito de gratidão a seu protetor, o Senhor de Balsemão, nos tempos do exílio da Inglaterra, conta quão dedicadamente se aplicou à leitura dos clássicos portugueses que enriqueciam a copiosa biblioteca do conde. Lendo e relendo os bons autores durante seis anos, foi apurando o seu domínio da língua materna que, “como muita gente, presumia saber arrazoadamente.”

Veja-se a crítica feita ao Bluteau na passagem que segue:

“Acompanhei este estudo dos livros clássicos com os auxílios de Bluteau, que achei muitas vezes em falta de vocábulos e frases, e mui freqüentemente sobejo em dissertações desapropositadas, e estranhas ao assunto, que fazem avolumar a sua obra.” (7, p. IX).

Continua dizendo que escolheu no Bluteau o que era propriamente português, deixando de lado muitos termos da cultura antiga, seguindo os passos dos “melhores dicionaristas das línguas vivas”. Não omitiu mais termos antigos como gostaria, para não ser acusado de omissão. Por fim, vem esta graciosa publicidade do seu dicionário:

“Do que recolhi de minhas leituras fui suprindo as faltas, e diminuições, que nele achava; e quem tiver lido o Bluteau, e conferir com o seu este meu trabalho, achará que não foi pouco o que ajuntei; e mais pudera acrescentar, se as minhas circunstâncias me não levassem forçado a outras aplicações mais frutuozas. Todavia não venderei ao público por grande o serviço que lhe fiz; basta que conheça, que lhe poupei a despesa de dez volumes raros; que lhe dou o bom que neles há, muito melhorado, por uma décima parte, ou pouco mais do seu custo, com a comodidade de não andar revolvendo tantos tomos; e isto é alguma coisa, enquanto não aparece uma outra melhor.” (7, p. X).

É um excelente dicionário para a sua época. A abonação dos verbetes é geralmente recolhida nos melhores escritores dos séculos XVI e XVII. Assim fazem parte do corpus de autoridades que Morais cita como modelos de boa linguagem: Luis de Camões, Gil Vicente, Damião de Góis, Diogo de Couto, Duarte Nunes Leão, Fernão Mendes Pinto, Francisco de Sá de Miranda, Francisco Rodrigues Lobo, Garcia de Resende, Gomes Eanes Zurara, Frei Heitor Pinto, Jerônimo Corte Real, João de Barros, Padre Antônio Vieira, D. Francisco Manuel de Melo, Padre Manuel Bernardes, etc. Não foram só os escritores literários aqueles de quem Morais recolheu citações para o seu arquivo lexicográfico. Utilizou também autores de obras técnicas e científicas dos seguintes domínios do conhecimento: filosofia, moral, religião, ciências sociais, política, filologia e linguagem, matemática, física, química, astronomia, botânica, geologia, medicina, engenharia, agricultura, artes, história. Preocupou-se ainda com registrar termos de linguagens especiais de uso na linguagem comum. Um dos méritos do seu dicionário é exatamente de indicar o registro linguístico da palavra-entrada.

Assinala, muitas vezes, o fato de o termo ser usado em determinada linguagem especial. Cf. nos exemplos seguintes: *hepático*, *narcótico* (med.), *ácido*, *solução* (quím.), *fluido*, *foco* (fis.), *nebuloso*, *órbita* (astron.), *calmaria*, *quarto*, *zarpar* (náut.),

apóstrofe, hipérbole (ret.), *decisório, usufruto* (dir.), etc. bem como conotações estilísticas de certas palavras típicas de diferentes registros de linguagem falada e/ou escrita: *bajular, barganha* (fam.), *fanico, gana, pespegar* (vulg.), *bandulho, lambada, pança* (chulo), *geringonça, gomarra* (gír.) *gaança, homízio* (ant.), *catapereiro, gelhos* (rúst.).

O Morais teve várias edições nos séculos XIX e XX; 3.^a.ed.-1823; 4.^a.ed.-1831; ...; 7.^a. ed. - 1877. De 1944-1957 a Editora Confluência publicou a 10.^a edição do dicionário de Morais em 12 volumes, versão revista e ampliada por José Pedro Machado. A página de rosto avisa o leitor que se trata de uma edição *corrigida, muito aumentada e atualizada*, pois os editores pretendiam corrigir as muitas inexatidões do Morais em matéria de ortografia (é a do acordo luso-brasileiro de 1945), de etimologia, a forma de definição da palavra e atualizar as informações científicas. Contudo, essa versão desfigurou a obra de Morais. É um outro dicionário, baseado no Morais, que contém várias gralhas, a despeito dos seus méritos.

Um bom dicionário português do século XIX é o de Frei Domingos Vieira: *Grande Dicionário Português* ou *Tesouro da Língua Portuguesa*, 1871-1874. Vieira deixou pronto o plano da obra e um arquivo de verbetes, anotações e abonações. A redação final desse dicionário foi executada por uma equipe que, após a morte de Vieira, aproveitou o trabalho por ele realizado. É um dicionário bastante completo e informativo para o século XIX. Via de regra os significados e usos linguísticos são ilustrados com citações de bons autores. São indicados: o étimo da palavra-entrada, expressões idiomáticas e sintagmas freqüentes em que ocorra esse vocábulo lema. Tome-se como exemplo a palavra *pena*. Há três entradas homônimas: 1.^a) *pena*, s.f. (do lat. *poena*); 2.^a) *pena*, s.f. ant. por *Penha* e 3.^a) *pena*, s.f. (do lat. *penna*). O primeiro verbe, bastante extenso, discorre sobre as várias acepções de *pena*; 'castigo'; 'cuidado, sofrimento'; 'dor, moléstia'; 'dificuldade, trabalho'; sempre acompanhando as explicações de tipo sinonímico com abonações. Seguem-se numerosas expressões em que esta palavra ocorre: 'alma em pena', 'sem pena', 'com duras penas', etc. E ainda os usos especiais dessa palavra enquanto termo forense - 'punição, castigo', acompanhados de sintagmas que o incorporam: *pena capital, pena corporal, pena de talião, pena judicial, pena legal, pena pecuniária*, etc.

Não se vai discutir aqui as outras duas formas homonímicas, pois o propósito dessa exemplificação é apenas dar uma rápida informação sobre o Vieira. Geralmente esse dicionário define bem a palavra-entrada. Às vezes, a extensão do verbe é um pouco exagerada do ponto de vista do léxico em geral: o verbe *paço*, p. ex., cobre três colunas inteiras, incluindo vários sentidos, inúmeras citações e expressões formadas com essa palavra. De uma perspectiva histórica da língua talvez se justificasse; não, porém, para os contemporâneos dos dicionaristas.

O Dicionário de *Aulete, Dicionário Contemporâneo da Língua Portuguesa*; de 1881 só foi planejado e iniciado por Caldas Aulete, pois esse dicionarista morreu antes de concluir sua obra. Foi completado por Santo Valente e colaboradores. Eis a opinião abalizada de Gladstone Chaves de Melo:

"... lugar de destaque cabe ao *Aulete* entre os dicionários portugueses, não só porque ele preenche inteiramente a sua finalidade, como, principalmente, porque traça novos rumos à lexicografia portuguesa, criando o tipo do dicionário moderno. "(6, p. 40). Chaves de Melo considera a segunda edição do *Aulete* "o mais útil e o mais satisfatório de todos os nossos léxicos (sic)". Dizia isso em 1947. Elogia sobretudo a qualidade das definições, embora lamente a carência e os defeitos das informações etimológicas, bem como a insuficiente atualização do léxico técnico e científico. Em 1958 a Edi-

tora Delta publicou uma versão brasileira do *Aulete*, elaborada por Hamilcar de Garcia: *Dicionário Contemporâneo da Língua Portuguesa* (5 volumes). Essa versão tem vários defeitos, embora pretenda ser mais completa que o primeiro *Aulete*. Foram aí incluídos muitos brasileirismos. Além da abonação de autores portugueses, essa edição foi acrescida da abonação de escritores brasileiros. Entretanto, tal documentação é bastante imprecisa e incompleta, indicando-se geralmente apenas o nome do autor citado e, por vezes, de maneira truncada. A revisão gráfica é defeituosa.

A primeira edição do Cândido de Figueiredo (*Novo Dicionário da Língua Portuguesa*) é de 1899. Esse dicionário pretendia ser o repositório mais completo do léxico português de todos os tempos bem como de regionalismos portugueses, brasileirismos, e de territórios onde se falava e fala o português. Cândido de Figueiredo publicou cinco edições do seu famoso dicionário que tem qualidades mas tem também terríveis defeitos. É um dicionário rico sobretudo com respeito ao número de palavras incluídas no seu acervo léxico. Há nele muitas palavras raras. A forma do verbete é muito simples. As definições são curtas, às vezes erradas e tolas, especialmente as relativas a termos técnicos ou de procedência técnico-científica. Muitos significados registrados são obsoletos há muito tempo. Cf. as entradas: *blusa, chingo, cômodo, comunicação, excursão*. Relativamente à nomenclatura que deu entrada neste dicionário veja-se, por ex., a seguinte seqüência de verbetes na letra S: *sueste, suebêrgios, sudro₁, sudro₂, sulvento, sum, sumaca, sumagral, sumagrar, sumagre, sumagreiro, sumalar, sumalário, sumanaís, sumanta, sumatra, sumaúna* etc. Foram também incluídos neologismos de aceitação duvidosa como: *sucessível, sucessoral, sucessorial*, etc. O Volume do léxico arrolado parece constituir uma grande riqueza; na verdade, como o Cândido de Figueiredo não tem a configuração de um *thesaurus*, mas de um dicionário de uso comum da língua, toda essa riqueza vocabular é quase inútil. Em 1949 (?) a Livraria Bertrand publicou uma 14.^a edição (?) do Cândido de Figueiredo com base na 5.^a ed. e prefácio de Júlio Dantas, na época presidente da Academia de Ciências de Lisboa. O acadêmico Júlio Dantas afirmava nesse prefácio que o Cândido de Figueiredo era o melhor dicionário da língua no seu tempo, “o mais opulento, o mais vivo e tecnicamente, o mais perfeito.” Tal afirmação é certamente muito discutível.

Nos tempos contemporâneos é preciso considerar, pelo menos, o mais popular dentre todos os dicionários da língua portuguesa: *Novo Dicionário da Língua Portuguesa* de Aurélio Buarque de Holanda Ferreira, Rio de Janeiro, Nova Fronteira, 1.^a ed., 1975. Baseado numa versão anterior publicada sob o nome de *Pequeno Dicionário Brasileiro da Língua Portuguesa* que tivera sucessivas reedições, Aurélio aumentou substancialmente a sua obra lexicográfica com o auxílio de assistentes e colaboradores. No prefácio o autor nos informa que pretendeu fazer um dicionário de “tipo médio ou inframédio”. Na verdade, o *Aurélio* se aproxima do tipo *thesaurus* no que diz respeito ao número de entradas do dicionário: “bem mais de cem mil verbetes e subverbetes”, segundo o próprio autor. Um dicionário médio seria um dicionário de 40.000 a 50.000 verbetes aproximadamente e um inframédio seria um tipo *Micro Robert* contendo uns 30.000 verbetes. O *Aurélio* chega a lembrar a configuração do Cândido de Figueiredo (5.^a ed.) com relação ao número de entradas e ao tipo de palavras do léxico que ele abriga. Evidentemente distancia-se do Cândido de Figueiredo, relativamente à seriedade e probidade no tratamento de muitos verbetes, como aqueles que descrevem a terminologia técnico-científica. O *Aurélio* acolheu muitas palavras raras, um grande número

de regionalismos, de vocábulos desusados ou obsoletos, de termos exclusivamente literários, um vasto acervo de termos técnicos e científicos, o que inchou demais a sua nomenclatura. Veja-se, por exemplo, uma seqüência qualquer de entradas encontradas na letra A: *aliazar* (var. de *aljazar*), *alicali* (reg.), *alicantina* (espanholismo?), *alico* (var. de *alincorne*), *alidade*, *alismatácea* (bot.), *alissóide* (mat.), *aljazar* (reg.), *almácea* (reg. port.), *almargio*, *alotriógnatos* (zool.). E vai por aí afora: *almonal*, *antomídeos*, *aragonita*, *arapuar*, *ardeiformes*, *ardômetro*, *ardosieira*, *arecal*, *arecineo*, *areiabranquense*, *areia-engolideira*, *areia-gulosa*, *areia-manteiga*, *areia-preta*, *areias-gordas*. Resolvi fazer um teste com a letra R, a qual letra recobre um conjunto médio de palavras do léxico global, considerando-se o número de palavras-entrada de cada letra no dicionário. Num total aproximado de 2.628 palavras (págs. 1189-1225) eu não conhecia 989, ou seja, uns 27% desse total. Claro está que para uma pessoa que tem a competência vocabular que tenho, sendo além disso um especialista em Lexicologia e Lexicografia há vários anos, representa um índice elevado de desconhecimento do léxico total do português. É verdade que nenhum falante por mais competente que seja em matéria vocabular, jamais conseguirá incluir no seu léxico ativo e passivo grandes parcelas do léxico geral da língua. Isso se verifica sobretudo quando o léxico em questão se identifica com o *thesaurus* da língua. Assim concluo que a recolha vocabular feita por mestre Aurélio para povoar o seu dicionário, levou-o a compor uma espécie de *thesaurus* do português, embora não exaustivo, pois não cobriu todas as épocas da história do português (considerando o século XVI como ponto de partida), nem todas as variantes lingüísticas, nem tão pouco as terminologias técnicas e científicas na sua totalidade. Pode-se até especular qual seria o valor quantitativo desse *thesaurus*. Faço uma estimativa de umas 500.000 palavras-tipo (sem considerar os valores polissêmicos), ou cinco vezes o acervo total do *Aurélio*, excluídos os nomes próprios, os quais somariam números muito elevados de tipos vocabulares distintos. Lembro, porém, que Aurélio já incluiu muitos gentilícios e antropônimos, do tipo: *aliancense* (natural de Aliança, PE, Br.), *andaraiense* (natural de Andaraí, BA, Br.), *areia-branquense* (natural de Areia Branca, RN, Br.), *raul-soarense* (natural de Raul Soares, MG, Br.), *reboucense* (natural de Reboças, PR, Br.), *recreense* (natural de Recreio, MG, Br.).

Aurélio afirma que utilizou um vasto corpus para extrair daí as abonações dos verbetes: 770 autores e 1610 obras. Tais fontes de abonação incluem escritores portugueses e brasileiros desde o século XVI até os primeiros anos da década de 1970. Vou citar alguns dentre os mais importantes para que se possa formar um juízo sobre as fontes do *Aurélio*:

1. Ficcionistas:

José de Alencar (bras. séc. XIX)
 Alexandre Herculano (port. séc. XIX)
 José Américo de Almeida (bras. séc. XX)
 Mário de Andrade (bras. séc. XX)
 Ciro dos Anjos (bras. séc. XX)
 Aluizio de Azevedo (bras. séc. XIX)
 Artur Azevedo (bras. séc. XIX)
 Raul Brandão (port. séc. XX)
 Camilo Castelo Branco (port. séc. XIX)
 Euclides da Cunha (bras. séc. XX)
 Osman Lins (bras. séc. XX)

Machado de Assis (bras. séc. XIX)
Fernando Namora (port. séc. XX)
Pedro Nava (bras. séc. XX)
Mário Palmério (bras. séc. XX)
Eça de Queirós (port. séc. XIX)
Graciliano Ramos (bras. séc. XX)
Ricardo Ramos (bras. séc. XX)
José Lins do Rego (bras. séc. XX)
Aquilino Ribeiro (port. séc. XX)
João Guimarães Rosa (bras. séc. XX)
Urbano Tavares Rodrigues (port. séc. XX)

2. Poetas:

Mário de Andrade (bras. séc. XX)
Oswald de Andrade (bras. séc. XX)
Luís de Camões (port. séc. XVI)
Castro Alves (bras. séc. XIX)
Gonçalves Dias (bras. séc. XIX)
Carlos Drummond de Andrade (bras. séc. XX)
Antônio Ferreira (port. séc. XVI)
Alphonsus de Guimarães (bras. séc. XIX)
Jorge de Lima (bras. séc. XX)
Fernando Pessoa (bras. séc. XX)
Antero de Quental (port. séc. XIX)
José Régio (port. séc. XX)
Fagundes Varela (bras. séc. XIX)

3. Ensaístas, historiadores, humanistas:

Alceu Amoroso Lima (bras. séc. XX)
Oscar Araripe (bras. séc. XX)
Frei Amador Arrais (port. séc. XVI)
Rui Barbosa (bras. séc. XIX)
Cleonica Berardinelli (bras. séc. XX)
Sérgio Buarque de Holanda (bras. séc. XX)
Hernani Cidade (port. séc. XX)
D. Francisco Manoel de Melo (port. séc. XVII)
Paulo Mercadante (bras. séc. XX)
João Ribeiro (bras. séc. XX)
João do Rio (bras. séc. XX)
Frei Luis de Sousa (port. séc. XVII)
Padre Antônio Vieira (port. séc. XVII)

Da extensa lista citada extraem-se algumas conclusões. Os escritores compõem um corpus relativamente homogêneo com predomínio do literário e, portanto, não se lhes pode apor a etiqueta “autores dos mais desvairados gêneros” como assevera mestre Aurélio. Além de umas tantas obras na área das ciências humanas (moral, crítica literária, história, sociologia etc.) não fazem parte do corpus publicações no domínio das ciências biológicas e exatas. Por conseguinte, os termos técnicos e científicos não devem ter sido colhidos em obras técnicas e científicas; aliás, esses termos não são abonados. Logo, o *Aurélio* é um dicionário com tendência a constituir um *thesaurus* mas não se fundamenta equitativamente em fontes escritas de todos os domínios do conheci-

mento humano. As abonações são pouco frequentes, a despeito da riqueza das fontes, arroladas no final do dicionário. Cito ao acaso alguns exemplos de verbetes que foram abonados com citações de autores: *aparelhar* (C. Castelo Branco, Camões), *arrancar* (C. Castelo Branco, Machado de Assis, Manoel Bandeira, Júlio Ribeiro), *atenazar* (Alexandre Herculano), *atender* (Ciro dos Anjos), *bicha* (Raul Brandão), *biqueira* (Carlos Drummond de Andrade), *bisaco* (Câmara Cascudo, Osman Lins), *borla* (Machado de Assis), *botar* (José Lins do Rego), *bramar* (Camões), *carecer* (José de Alencar, Machado de Assis), *cataclismo* (João Ribeiro), *coar* (Raul Brandão), *coxia* (Eça de Queirós), *derramar* (Gonçalves Dias), *desempeno* (Euclides da Cunha), *esbaforido* (Eça de Queirós), *esboço* (Eça de Queirós), *esmolar* (Eça de Queirós), *eternizar* (Jorge de Lima), *exarcebar* (Hernani Cidade), *exprimir* (Alceu de Amoroso Lima), *exprobrar* (Rui Barbosa), *ferroar* (José de Alencar). Predominam os chamados modelos de linguagem, os clássicos, como Machado de Assis, Eça de Queirós.

Aurélio registra uma vasta fraseologia (sintagmas lexicalizados ou em vias de lexicalização), dando-lhes entrada na palavra mais importante da expressão: o substantivo em primeiro lugar, seguindo-se em ordem de importância: o verbo, o adjetivo, o pronome, o advérbio. Nesse particular seguiu o critério do *Diccionario de la Real Academia Española*, segundo ele próprio nos informa no prefácio. O autor afirma ainda que a sinonímia nos vocábulos-entrada é abundante, bem como outras informações sobre o campo léxico do verbete em epigrafe: antônimos, parônimos, o que é verdade.

Vou tentar fazer, a seguir, uma rápida crítica dos modelos de verbete do dicionário *Aurélio*. A pedra de toque de um dicionário é a definição da palavra-entrada. Muitas vezes usando os modelos clássicos de definição, Aurélio não foi muito feliz. A melhor definição é aquela que define e/ou descreve a palavra através de uma paráfrase. Tome-se, por exemplo, a definição de *amparo* = “ação ou efeito de amparar”. Essa definição seria boa se ao consultar o verbete *amparar* não encontrássemos lá como definição: “1. Dar, ou servir de amparo a, estear, escorar”. Aurélio incidiu na circularidade, remetendo do substantivo (*amparo*) ao verbo (*amparar*) e vice-versa, o que deixa as coisas no mesmo lugar. Ainda que tenha acrescentado os sinônimos *estear*, *escorar*, de fato não definiu *amparar* com uma paráfrase autônoma, de tal modo que em *amparo* pudesse remeter a *amparar*. Mencionou-se aqui apenas um exemplo, porém, esse fato se repete numerosas vezes. Por outro lado, o recurso de definir por sinônimos, muito utilizado por Aurélio, não é um bom método. Além do exemplo acima citado, cf. *amparar*: 2. proteger, patrocinar, favorecer. A forma do verbete no *Aurélio* é criticável em numerosíssimos casos. São deficientes muitas definições e há abuso do recurso à sinonímia para “definir”. Muitas vezes também é contestável a ordem hierárquica das acepções de uma palavra polissêmica que possua inúmeros significados; nesses casos o dicionário dispõe inadequadamente a seqüência não ficando muito claro o critério de distribuição dos sentidos registrados e, não raro, alguns números abaixo aponta como significação diversa, um sentido que já fora incluído anteriormente (cf. *amparar*, *construção*, *estrela*, *linha*, *montar*, *pensar*). Parece que algumas vezes o critério de distribuição é prioritariamente sintático e não semântico. Aliás, no prefácio, Aurélio refere que deu grande atenção à regência verbal. Por vezes não dá entradas distintas a lexemas que poderiam ser considerados homônimos como: *montar*₁ = “estar ou pôr-se sobre um animal, geralmente cavalo” X *montar*₂ = “colocar em estado de funcionamento”. Também registra termos muito raros, às vezes especificamente literários, sem informar o consulente dessa conotação específica da palavra. Ex.: *aljava*, *aljofre*, *ara*, *bisaco*, *borzeguim*, *brial*, *epístola*, *epitáfio*, *ergástulo*, *ermita*, *escarmento*,

faula, fenecer, fiacre, flamante, flébil, mácula, madeixa, mansuetude, melena, nácar, nefando, noctâmbulo, novel, osedante, obséquias, odorífero, ominoso, onírico, outono, rebuço, recalitrar, recôndito, sáfico, salaz, sege etc.

II. TIPOLOGIA DE OBRAS LEXICOGRAFICAS

1. O tipo mais comum de dicionário é o “dicionário padrão da língua”, ou “dicionário de uso da língua”, de que seriam exemplos vários dicionários da língua portuguesa comentados na última parte do item anterior: o *Morais*, o *Aulete*, o *Cândido de Figueiredo*, o *Aurélio*. Raros são os dicionários que preenchem idealmente o tipo do dicionário padrão da língua. Esse assunto será discutido miudamente no capítulo seguinte: *O dicionário padrão da língua*.

2. O “dicionário ideológico” ou “analógico” organiza os conceitos em campos semânticos, ao invés de ordenar as palavras em ordem alfabética como os dicionários comuns. Essa tradição também é antiga; no século XVII Comenius elaborou o primeiro dicionário desse tipo. O *Thesaurus* de Roget, feito para o inglês, foi publicado pela primeira vez em 1852 pela Longman em Londres.

Um dos melhores dicionários desse tipo é o *Diccionario Ideológico de la Lengua Española* de Julio Casares (Madrid, 1942). Vejamos o que diz o próprio Casares a respeito do dicionário ideológico:

“Os dicionários ordenados com este critério têm duas partes: a primeira é a propriamente ideológica, a segunda é a alfabética, ordenada exatamente como um dicionário semasiológico. Na parte ideológica as palavras se estruturam segundo seu enquadramento em colunas básicas que correspondem à divisão do universo em categorias fundamentais.

Na parte sinótica se encontra o plano geral da classificação; no caso do *Diccionario Ideológico de la Lengua Española* a divisão do universo lexical foi estabelecida em trinta e oito classes, das quais *Deus* compõe uma classe e o *universo*, trinta e sete classes.” (4, p. 439)

Pode-se argüir de arbitrária essa divisão em 38 classes conceptuais, elaborada por Casares. Talvez o progresso dos estudos de semântica e o conhecimento maior do universo poderão proporcionar melhores dados e recursos para a confecção deste tipo de dicionário no futuro.

Dentro deste modelo um dicionário contemporâneo do português muito útil, embora menos refinado, é o *Dicionário Analógico da Língua Portuguesa* (Idéias afins) de Francisco Ferreira dos Santos Azevedo, Brasília, Editora Coordenada, 1974 (?). Estrutura os conceitos e signos léxicos de acordo com o seguinte esquema classificatório: I: Relações abstratas / II: Espaço / III. Matéria. / IV: Intelecto / V: Vontade / VI: Afeições /. Esse modelo de classificação foi certamente emprestado ao *Roget's Thesaurus of English Words and Phrases Newed* prepared by Susan Mc Lloyd. London, Longman, 1982, cujo índice de temas é o seguinte: 1. “Abstract relations” (Relações abstratas) / 2. “Space” (Espaço) / 3. “Matter” (Matéria) / 4. “Intellect: the exercise of the mind” (Intelecto: o exercício da mente) / 5. “Volition” (Volição) / 6. “Emotion, religion and morality” (Emoção, religião e moralidade).

Hallig et Wartburg também elaboraram um dicionário similar para o alemão: R. Hallig & W. von Wartburg. *Begriffssystem als Grundlage für die Lexicographie. Versuch eines Ordnungsschemas*. Berlim, 1953. Em 1963 os autores publicaram uma nova versão reformulada e ampliada. Apresentam aí um esquema de classificação de conceitos.

Em 1981 a Longman de Londres publicou um novo dicionário ideológico para o inglês, cujo editor-chefe foi Tom Mc Arthur: *Longman Lexicon of Contemporary English*, baseado no *Thesaurus* de Roget. Os campos semânticos, ou grandes áreas de significação em que foram organizados os conceitos são os seguintes:

- A. Vida e coisas vivas.
- B. O corpo: suas funções e seu bem-estar.
- C. Os seres humanos e a família.
- D. Construções, casas, a casa, roupas, pertences pessoais, cuidados pessoais.
- E. Alimentação, bebida e agricultura.
- F. Sentimentos, emoções, atitudes e sensações.
- G. Pensamento e comunicação, linguagem e gramática.
- H. Substâncias, materiais, objetos e equipamento.
- I. Arte e artesanato, ciência e tecnologia, indústria e educação.
- J. Números, medidas, dinheiro e comércio.
- K. Entretenimento, esporte e jogos.
- L. Espaço e tempo.
- M. Movimento, localização, viagem e transporte.
- N. Termos gerais e abstratos.

Este tipo de dicionário que estrutura os conceitos em redes de significação pode ser questionável. É possível que o modelo de Casares só se aplique ao espanhol, o modelo do *Lexicon* só se aplique ao inglês e assim por diante. Vale dizer: cada sistema só seria válido para a língua natural para que foi elaborado. Se aceitarmos a teoria de Sapir Whorf e de outros lingüistas sobre o relativismo lingüístico, teremos que admitir que a conceptualização da realidade é típica de cada língua. Isso significaria que cada sistema lingüístico classifica os dados da realidade e da experiência de uma maneira própria, ou seja: o léxico de cada idioma categoriza o mundo e a realidade social e cultural de acordo com o seu próprio modelo classificatório. Para os partidários dos universais lingüísticos tal argumentação seria inaceitável. Talvez no meio termo se situe a verdade. Haveria áreas do conhecimento humano empírico, nomeadas no léxico de cada língua, que seriam exclusivas dessa língua e da cultura que ela expressa. Contudo, no universo cultural em expansão em que hoje vivem os homens, estaria ocorrendo uma convergência dos sistemas classificatórios, expressos por denominações lexicais. E mais ainda: na aldeia global dos meios de comunicação em que está vivendo o homem contemporâneo, intensifica-se a tendência à universalização de conceitos, sobretudo no domínio técnico-científico, fenômeno esse que pode ser bem representado por um organismo internacional para a padronização de termos: o *Infoterm* de Viena.

3. O dicionário histórico constitui uma outra modalidade. Existem vários tipos de dicionários históricos. Há um que se baseia no vocabulário e na língua de determinada época histórica. São exemplos desse tipo dos vários dicionários sobre a Idade Média que possuem algumas línguas européias. Cf. F. Godefroy — *Dictionnaire de l'ancienne langue française et de tous ses dialectes du IXe. au XVe. siècle*, 1881; Boggs, Kasten, Kemiston & Richardson — *Tentative Dictionary of Medieval Spanish* (1946) e outros que descrevem o francês, o inglês, o alemão medievais. Também o *Dictionnaire de la langue française du seizième siècle* de Edmont Huguet (1946) constitui um exemplo de dicionário dedicado e um estágio anterior da língua francesa, nesse caso o século XVI. O *Dictionnaire du français classique* de Dubois, Lagaro & Lerord (1971) trata da língua do século XVII. Esse tipo de dicionário é muito útil na leitura de obras datadas das épocas históricas a que eles se consagram. Assim o dicionário de Godefroy constitui

um instrumento indispensável para ajudar a compreender e interpretar os autores franceses dos séculos IX ao XV.

Outro tipo de dicionário histórico é o pancrônico, muitas vezes rotulado de etimológico. Sendo elaborado a partir da perspectiva da língua contemporânea, ele se ocupa dos estágios anteriores do idioma, remontando à origem das palavras; tenta acompanhar a evolução histórica dos vocábulos, assinalando os diferentes valores semânticos por eles assumidos no decorrer do tempo, indicando *pari passu* as datações de cada um deles.

Um excelente exemplo é representado pelo *Diccionario crítico-etimológico de la lengua castellana* de J. Corominas (1954). Esse dicionário registra os étimos das palavras da língua espanhola contemporânea e procura seguir a evolução dos seus significados e usos no âmbito das línguas e dialetos da Península Ibérica. Assim não é apenas um dicionário histórico do espanhol, mas também de outras línguas e dialetos hispânicos tais como: catalão, português, galego, aragonês, navarro, andaluz etc.

Existem numerosos dicionários enquadrados dentro desta modalidade, tanto de línguas modernas como antigas. Alguns exemplos: A. Meillet & A. Ernout — *Dictionnaire étymologique de la langue latine* (1939); W. Meyer-Lübke — *Romanisches Etymologisches Wörterbuch* (1935); Wartburg & Baldinger — *Französisches Etymologisches Wörterbuch*; J. P. Machado — *Dicionário Etimológico da Língua Portuguesa* (1953).

Um dicionário modelo dentre os históricos é o *Oxford English Dictionary* (OED). Esse dicionário cuja elaboração foi iniciada em 1857, teve sua primeira edição publicada entre 1884-1928. Constitui uma revolução na lexicografia, só superado hoje pelo *Trésor de la langue française*. O OED contém 1.800.000 citações da literatura inglesa desde as primeiras datações de cada palavra. O verbete do OED tem como principal característica a documentação rigorosa das ocorrências da palavra-entrada. As ocorrências das palavras dicionarizadas foram coletadas em 5 milhões de passagens da literatura inglesa desde as suas origens até o começo do século XX. O setor de Lexicografia da Oxford University Press, sob a direção do Dr. Burchfield, trabalha continuamente na atualização deste monumento da língua inglesa. No OED a documentação de cada uma das acepções está ordenada cronologicamente a partir do primeiro registro conhecido. Os autores do OED procuraram fornecer, no mínimo, uma documentação por século; às vezes, várias. Quando a palavra se tornou obsoleta, o OED registra o fato e documenta a última ocorrência datada.

Remeto o leitor ao item relativo à história da lexicografia francesa, feito anteriormente, onde há dados sobre o *Dictionnaire de la langue française du 19ème. et du 20ème. siècle*, que é um dicionário histórico de uma certa forma.

Quanto aos dicionários históricos da língua portuguesa, existem poucas obras a serem comentadas. Sobre o *Dicionário Etimológico da Língua Portuguesa* de Antenor Nascentes (1932), vou transcrever a abalizada opinião de Gladstone Chaves de Melo a seu respeito:

“... está baseado na edição mais antiga do *Romanisches Etymologisches Wörterbuch* de Meyer - Lübke, edição que ficou praticamente inutilizada pela terceira, ultimada em 1935; segundo: porque arrola hipóteses etimológicas, sem indicar quais as imprestáveis, quais as prováveis, quais as inteiramente aceitáveis.” (5, p. 63)

José Pedro Machado publicou em 1951 *Dicionário Etimológico da Língua Portuguesa* em dois volumes. O verbete do dicionário de Machado contém geralmente o seguinte tipo de informação: 1. origem do vocábulo/ 2. discussão do étimo e das variantes quando há controvérsia/ 3. documentação e/ou abonação das formas e sentido

dos valores semânticos anteriores à contemporaneidade/ 4. derivados, para os quais se acrescenta o mesmo tipo de informação; tais derivados aparecem como subentradas da entrada principal, sendo destacados com negritos para facilitar a sua identificação. Por exemplo: em *barro*: **barreiro**, **barrosa**, **barroso**; em *base*: **basear**, **basídio**, **basí**; em *-bata*: **acrobata**, **hidróbata**. Essa forma de agrupar os cognatos, reunindo formas derivadas da raiz portuguesa e formas oriundas da raiz grega de que deriva a portuguesa, dificulta sobremaneira o manuseio do dicionário por uma pessoa não versada em filologia e lingüística histórica. Por essa razão, Machado procurou sanar essa dificuldade incluindo no fim do segundo volume um “índice remissivo”, o qual informa sobre a localização da palavra procurada, indicando a página em que ela se encontra. O dicionário de Machado contém uma série de defeitos dos quais devo assinalar dois, pelo menos: a) inclui muitos vocábulos obsoletos, desusados, raros e regionais, deixando de acrescentar palavras correntes nos tempos modernos (Cf.: *caçabe*, *caçapo*, *cacatu*, *caçagem*, *damões*, *darbar*, *datismo*, *emulgente*, *finicola*, *forda*, *frimário*); b) ao citar autores medievais e clássicos indica uma edição moderna de que se serviu como documentação (p. ex.: Fernão Lopes, *Crônica de D. João I*) sem indicar a época em que o texto foi escrito. Ora, o consulente médio não conhece as datas das obras de Fernão Lopes, nem da maioria dos autores usados como fontes documentais. Acresce que, na introdução do dicionário, não consta uma relação dos autores e edições utilizadas, nem a datação pertinente a ambos.

O mais recente dos etimológicos é o *Dicionário Etimológico Nova Fronteira da Língua Portuguesa* de Antônio Geraldo da Cunha, Rio de Janeiro, 1982. É um bom dicionário no gênero, mas sem grandes pretensões. Pautou-se pelo modelo do dicionário etimológico do francês de Bloch e Wartburg e na versão abreviada do Corominas para o espanhol. No dicionário de A. G. da Cunha o verbete indica o significado da palavra-entrada e registra as variantes históricas do vocábulo com a respectiva datação.

4. Dicionários de tipo especial.

Alguns exemplos de dicionários de tipo especial em língua portuguesa serão referidos a seguir.

Agenor Costa — *Dicionário de Sinônimos e Locuções da Língua Portuguesa* 2.^a edição, 1954, 2 volumes. Francisco Fernandes também é autor de um “dicionário de sinônimos”, mas a sua melhor obra é, sem dúvida, o *Dicionário de Verbos e Regimes*, Porto Alegre, Editora Globo. Teve várias edições e reimpressões. A última que conheço é a 4.^a edição, 13.^a impressão, 1968. A primeira edição foi de 1940. Nessa quarta edição Fernandes nos fornece “mais de 11.000 verbos em suas diversas acepções e regências”. Constitui um excelente trabalho para a época em que foi composto, mas contém defeitos que clamam por uma refacção da sua obra, ou a elaboração de outra similar. Muitos dos verbos incluídos na nomenclatura de Fernandes são desusados, outros, muito raros; julgo até que alguns são invenções suas. Cf. “*desfolegar*, o mesmo que resfolegar”; “*desamigar* o mesmo que inimizar”; “*insalivar*, *oscitar*, o mesmo que bocejar”; “*outar*, o mesmo que joeirar”. Esses dois últimos verbos Fernandes provavelmente os copiou da versão que Machado fez do Morais (cf. item I.). Ora, no Morais-Machado não existe abonação dessas formas e a gente se pergunta se não seriam invenção de Machado, pois elas não existem no velho Morais (ed. de 1813). Quanto a *desfolegar*, os exemplos de Morais são do *Livro de Alveitaria* (séc. XIV) e Gil Vicente (séc. XVI) já obsoletos; deveriam ter sido, portanto, omitidos.

Está sendo elaborado um *Dicionário Gramatical do Português Contemporâneo*, sob a coordenação do Prof. Francisco da Silva Borba (UNESP, Campus de Araraquá-

ra), trabalho esse patrocinado pela FAPESP. A equipe de colaboradores do Prof. Borba são colegas seus do Departamento de Linguística: Cacilda de Oliveira Camargo, Maria Helena de Moura Neves, Sebastião Expedito Ignácio, Odette G. L. A. Souza Campos, José Luiz Fiorin, Antônio Silveira Reis, Djalma Dezotti, Maria Celeste C. Dezotti, Elvira Wanda Vagones Mauro, Clóvis Barleta de Moraes. São os seguintes os objetivos principais do dicionário: a) contribuir para a interpretação correta de uma oração e, conseqüentemente, chegar ao valor mais exato de um texto; b) contribuir para melhorar a atuação do usuário no manejo da língua escrita. São os seguintes seus objetivos suplementares: 1) mostrar como se entrosam, na dimensão pragmática da linguagem, a sintaxe e a semântica; 2) mostrar como a língua é um conjunto de possibilidades: a partir das realizadas, prever as de realização possível; 3) fornecer, pela descrição do sintagma verbal, material de consulta e de cotejo para especialistas que trabalham com fatos sintáticos do português contemporâneo do Brasil; 4) sugerir temas, enfoques e técnicas de análise sintático-semântica. Esse dicionário destina-se a todos aqueles que manejam a forma escrita da língua portuguesa como instrumento de trabalho. As abonações dos verbetes baseiam-se em textos em prosa (mais ou menos 30.000 páginas) dos últimos trinta anos, abrangendo a produção dos vários pontos do País (romances, jornais etc.). Deverá totalizar aproximadamente 7.000 verbetes quando estiver pronto. (Nota: essas informações foram fornecidas pelo Prof. J. L. Fiorin).

Dicionários especializados sobre aspectos particulares da língua portuguesa não são raros. F. da Silva Borba e Z. Jota publicaram cada um deles, dicionários de linguística. Um clássico nessa área é do de J. Mattoso Câmara Jr. — *Dicionário de Linguística e Gramática*, 8.^a ed. Petrópolis, Vozes, 1978. No domínio da linguagem vulgar e da gíria existem vários dicionários publicados recentemente. Cf.: Albino Lapa *Dicionário de calão*, 2.^a ed. Lisboa, Editorial Presença, 1974. Mário Souto Maior, *Dicionário do palavrão*. Recife, Guararapes, 1980. Eduardo Nobre, *O calão — Dicionário da gíria portuguesa*. Lisboa, Casa do Livro 1980. Euclides Carneiro da Silva, *Dicionário da gíria brasileira*. Rio de Janeiro, Bloch, 1973. Ariel Tacla. — *Dicionário dos marginais*. Rio de Janeiro, Forense — Universitária, 1981. Manuel Viotti — *Novo dicionário da gíria brasileira*. 3.^a ed. Rio de Janeiro/São Paulo, Livraria Tupã Editora, s/d.

Finalmente vou considerar os dicionários especialmente dedicados a um domínio do conhecimento, que não a linguagem. São dicionários científicos e/ou técnicos. Nos tempos contemporâneos eles se multiplicam mais e mais, sobretudo por causa da especialização crescente que se verifica em cada ramo da ciência e da técnica.

A editora Longman de Londres está trabalhando num macrojeto de Lexicografia, tendo como base o dicionário Webster, cujos direitos autorais foram por ela adquiridos. No banco de dados léxicos estocados na memória do computador, cada palavra está codificada com um código especial que a distingue: o código em questão pode identificá-la como integrante do léxico comum da língua, ou como termo especializado de uma determinada área do conhecimento. Com esse trabalho classificatório básico, os lexicógrafos da Longman pretendem elaborar muitos dicionários técnicos e científicos, depois da publicação da versão atualizada do Webster. Agiriam da seguinte forma: utilizando o código com que a palavra foi rotulada, pediriam ao computador todo o repertório vocabular daquela área científica ou técnica e já teriam, na saída, a listagem terminológica que comporia o cerne do dicionário, acoplada aos contextos em que as palavras dessa lista tinham ocorrido, dados esses básicos para a redação de cada um dos dicionários técnico-científicos que pretendem publicar.

A título de informação vou alistar, a seguir, uma série de dicionários especializados do tipo técnico-científico que existem em língua portuguesa. Alguns são glossários e não propriamente dicionários. Foram feitos em épocas diferentes, geralmente por especialistas da área, os quais não possuíam formação lexicográfica. Padecem por isso de muitos defeitos.

Blakiston. *Dicionário Médico Ilustrado*. S. Paulo, Organização Andrei Editora S/A, 2.^a ed.; s/d.

Farmacopéia Brasileira. 3.^a ed. S. Paulo, Organização Andrei Editora S/A, 1977.

Pedro A. Pinto. *Dicionário de Termos Médicos*. Rio de Janeiro, Editora Científica, 7.^a ed., 1958 (1.^a ed. 1926 ?).

F. E. Rabello. *Nomenclatura Dermatológica*. Rio de Janeiro, Imprensa Brasileira Ltda., 1974 (1.^a ed. 1950).

Carmelino Scartezzini. *Dicionário Farmacêutico*. Rio de Janeiro, Editora Científica, 1956.

Dicionário de Informática Inglês-Português. 3.^a ed. Revista, aumentada e atualizada. Rio de Janeiro. SUCESU, 1982 (?).

Dicionário Geográfico Brasileiro. 2.^a ed. Porto Alegre, Ed. Globo, 1972. Gilberto Luiz da Cruz. *Livro Verde das Plantas Medicinais e Industriais do Brasil*. (2 vols.). Belo Horizonte, Gráficas Velloso, 1965.

Rodolpho Von Hering. *Dicionário dos Animais do Brasil*. S. Paulo, Editora da Universidade de Brasília, 1968.

Existem outros tipos de dicionário ou repertórios lexicográficos. O dicionário inverso (ou grafêmico) é muito útil para o estudo dos processos de sufixação e da produtividade léxica de determinados sufixos. Por conseguinte, particularmente útil no caso de línguas de tipologia flexional bastante desenvolvida como os idiomas românicos, cuja principal fonte de criação léxica é a derivação sufixal e prefixal. Existe um bom dicionário deste tipo para o francês: A. Juilland - *Dictionnaire Inverse de la Langue Française*. (1965). Um dicionário desse tipo pode ser útil não apenas para lingüistas e especialistas em línguas. Recentemente o alto comando militar da França adquiriu muitas cópias desse "dicionário inverso" do francês de Juilland para utilizá-lo na descodificação de códigos secretos pelos serviços de inteligência. Tenho informação de que existem dois dicionários inversos do português publicados por estrangeiros e dos quais tenho apenas a referência: Dieter Messner. *Dictionnaire Inverse da la Langue Portugaise e Dicionário Inverso Português*, Moscou, 1973.

A cultura luso-brasileira precisa refazer muitas dessas obras para atualizá-las e aprimorá-las lexicograficamente; deve também elaborar outros dicionários relativos a áreas do conhecimento que não possuem nenhum dicionário especializado. O poder público deveria criar um órgão para ocupar-se de tão magna tarefa.

As enciclopédias são obras de referência que buscam reunir o máximo de informação sobre os mais variados domínios do conhecimento para consumo do público em geral, e não de especialistas. Podem ser do tipo alfabético ou por área do conhecimento. Em língua portuguesa existem poucas. A mais antiga é: *Dicionário Enciclopédico Salvat*. Salvat Editores S.A., Barcelona, 1955. A mais recente: *Enciclopédia Mirador Internacional*. S. Paulo/Rio. Enciclopédia Britânica do Brasil Publicações Ltda. 1976. A Editora Delta também publicou uma enciclopédia em português, baseada na Delta-Larousse e organizada por assunto.

III. O USO DO COMPUTADOR NA LEXICOGRAFIA CONTEMPORÂNEA

1. O advento do computador constituiu uma verdadeira revolução dentro da ciência da informática e da lexicografia em particular. A história dos grandes monumentos lexicográficos do passado como o dicionário francês de Littré e o *Oxford English Dictionary* testemunham o penoso sacrifício feito por lexicógrafos heróicos que compilaram, classificaram, ordenaram, organizaram e redigiram um enorme volume de dados léxicos com suas citações documentais, colhidas em centenas de milhares de passagens literárias. Ambas foram obras de várias décadas, a que os dicionaristas dedicaram parte de suas vidas, trabalhando, às vezes, até vinte horas por dia como no caso de Littré. Hoje os dicionaristas não precisam mais dar o seu sangue para elaborar um tesouro lexicográfico porque uma grande parte do trabalho manual, monótono e estafante, pode ser feito pelo computador. Além disso, os contemporâneos contam com a vantagem de produzirem uma obra mais completa e de melhor qualidade, pois economizam sua energia com a parte repetitiva do trabalho, proporcionalmente muito volumosa no conjunto das tarefas; dessa forma poderão utilizar essa energia para a seleção do material compilado pela máquina e para a redação do texto final, que constitui a etapa mais importante de qualquer obra lexicográfica. O computador é particularmente útil na confecção de dicionários históricos, ou de quaisquer dicionários que pretendam documentar a informação fornecida em cada entrada. O testemunho de um lexicógrafo (A. J. Aitken, Edinburgh) que está trabalhando com um monumento da antiga língua escocesa (*Dictionary of the Older Scottish Tongue*, corpus total: 1 bilhão de palavras; corpus selecionado para o dicionário: 200 milhões de palavras) parece-me vir muito a propósito:

“Este artigo pretende desenvolver dois argumentos básicos do ponto de vista daquele que pratica lexicografia histórica. Em primeiro lugar, como a capacidade do computador para contar ocorrências de palavras em extensas passagens de texto pode possibilitar ao planejador de um dicionário histórico um controle muito mais preciso do que sucedeu até hoje, sobre o tamanho e a forma de sua coleção de citações documentais, possibilitando, portanto, o controle da magnitude da tarefa que ele está executando. Em segundo lugar, a manipulação posterior do arquivo do seu dicionário pode beneficiar-se das facilidades oferecidas pelo computador para armazenamento de dados, possibilitando o confronto desse material e a sua seleção; e mais ainda: o acesso flexível aos dados e a sua imediata recuperação bem como a facilidade e a conveniência de inserir revisões e suplementações.” (1, p. 29)

2. O léxico constitui um conjunto aberto em qualquer sistema lingüístico e, por conseguinte, não apenas vastíssimo quando comparado com outros setores e níveis da língua (fonológico, morfológico, sintático) mas também em permanente expansão sobretudo numa língua de civilização. Por essa razão, o quantitativo é uma das propriedades fundamentais do vocabulário. Como a confecção de dicionários é tarefa em que manipulam gigantescos volumes de dados, justifica-se sobremaneira o uso de computadores para a execução de um empreendimento lexicográfico. De fato, o computador é particularmente apropriado e eficaz quando se trabalha com grande quantidade de dados. Inclusive financeiramente o seu uso torna-se recomendável. Vamos dar um exemplo que ilustra as afirmações acima.

Em 1969-1971 B. Richman, J. Carrol & P. Davies elaboraram um trabalho de léxico-estatística — *American Heritage Word Frequency Book* (AHWFB) — a fim de gerar dados léxicos que seriam utilizados posteriormente na confecção de dicionários

pela American Heritage Publishing Company. A partir dos resultados do AHWFB e de outras obras lexicológicas foi criado o *Children's Dictionary* (1979).

O *American Heritage Word Frequency Book* (1971) manipulou um corpus de 5.088.721 palavras, retiradas de 1.075 livros e revistas, dos quais se selecionaram amostras de 500 palavras (tamanho da amostra) em cadeia. Esse grande arsenal de dados léxicos forneceu apenas 86.741 palavras diferentes; chamemo-las de "tipo" (wordtypes). Isso já mostra a fantástica proporção de redundância e repetição nos discursos lingüísticos.

Veja-se o quadro seguinte com alguns dados estatísticos reveladores:

AHWFB	CORPUS TOTAL: 5.088.721		
freqüência acumulada de: <i>the + of + and</i>	happax legomena (tipos que ocorreram uma só vez)	palavras de 2 a 5 ocorrências	palavras de freqüência 1 a 20
653.023	35.079 tipos diferentes	25.358 tipos diferentes	74.686 tipos diferentes

opor aos 86.741 tipos do AHWFB

É impressionante nesta como em qualquer computação das mais variadas línguas, como um número muito pequeno de palavras (instrumentais lingüísticos, palavras-gramaticais) têm uma altíssima freqüência, contra um enorme número de palavras plenas com baixíssima freqüência. No exemplo acima tal fato é típico: três lexemas somam 653.023 ocorrências, enquanto 35.079 (os happax legomena) apenas 35.079 ocorrências. Também é muito pequena, estatisticamente desprezível num corpus de 5.088.721 palavras, o concurso das palavras de 2 a 5 ocorrências com os seus 25.358 tipos diferentes (compare-se com os 86.741 tipos diferentes de todo o corpus!). É por isso que muitos trabalhos de orientação estatística, especialmente os dicionários de freqüência, desprezam as palavras com freqüência abaixo de 5. Quando o corpus é grande, como é o caso deste AHWFB, mesmo as freqüências abaixo de 20 passam a ser irrelevantes. Note-se que elas concorreram com 74.686 tipos diferentes de palavras. Sobram do total de 86.741: 12.055 palavras diferentes, que seriam as palavras dignas de consideração. Esses vocábulos são relevantes tanto para a elaboração de livros pedagógicos como para dicionários de uso geral na língua como o *Children's Dictionary* anteriormente citado. Mais um dado estatístico importante: os 5.000 vocábulos mais freqüentes do AHWFB ocorreram 77 vezes (ou mais) no corpus num total de 4.547.336 ocorrências (de 5.088.721!). Com base nos resultados deste trabalho poderíamos afirmar com Howard H. Keller:

"A disparidade percentual é surpreendente — se um aluno possui um vocabulário de 5.000 palavras, ele conhecerá apenas 5,77% do total de lexemas da língua, mas essas 5.000 palavras constituirão 89% do vocabulário de qualquer texto!" (3, p. 177)

Os autores do AHWFB propuseram um novo parâmetro: FPM (frequency-per-million), ou freqüência-por milhão, para compensar as distorções criadas pela dispersão. (Nota: a dispersão é uma resultante da variedade dos gêneros literários, dos temas tratados, dos níveis de linguagem, dos estilos.). O índice FPM fornece, pois, um fator de reajuste da freqüência à dispersão.

“Os autores acreditam que essa FPM reflete melhor a frequência verdadeira de uma forma que seria encontrada em um corpus de qualquer tamanho, dos pequenos aos infinitamente grandes.” (3, p. 177)

A FPM constitui uma informação importante de um dicionário de frequência. Ela acrescenta uma terceira dimensão à palavra, mostrando o valor ou a importância relativa dessa palavra no conjunto total do léxico, o que falta em um dicionário comum.

Não só os autores do *Children's Dictionary* (Houghton Mifflin, 1979) se serviram dos dados estatístico-vocabulares produzidos pelo AHWFB. O lexicógrafo A. J. Augarde, da Oxford University Press, elaborou o *Oxford Intermediate Dictionary* (12.000 verbetes) usando o AHWFB como referência básica. Selecionou as suas 12.000 palavras-entrada a partir da lista de palavras em ordem decrescente de frequência do AHWFB. Estabeleceu o número de ordem 12.000 como parâmetro e como limiar inferior o vocábulo de ordem 10.000 e como limiar superior, o de número 20.000. Usando sua competência lingüística e seu bom senso, eliminou algumas palavras situadas entre a frequência 10.000 e 12.000 e substituiu-as por outras palavras situadas entre a frequência 12.000 e 20.000, que julgou mais necessárias para um aluno da escola secundária.

No *Frequency Dictionary of Portuguese Words* (Duncan, Ph. D. Dissertation, 1972) o índice de dispersão representa a distribuição irregular de uma palavra nos cinco subcorpus que constituíram o corpus total: 1.º literatura dramática; 2.º romances e contos; 3.º ensaios; 4.º jornais e revistas; 5.º textos técnicos e científicos. Evidentemente a ocorrência de cada vocábulo (excetuados os gramaticais) é diferente conforme o tema tratado e o tipo de linguagem. Um exemplo típico: no corpus do FDPW a palavra *bonito* ocorreu 14 vezes em peças, 6 vezes em romance e contos, 6 vezes em ensaios e apenas 1 vez nos textos jornalísticos, técnicos e científicos. Esse fato mostra que essa palavra é característica de um tipo de linguagem, mas irrelevante em outros gêneros literários e registros lingüísticos. O seu índice de dispersão foi de 45,88%. Uma palavra com dispersão uniforme através de todos os gêneros seria, p. ex., o artigo *o*.

O FDPW propõe um parâmetro semelhante à FPM mencionada acima, isto é, o coeficiente de *uso*. Esse coeficiente reflete o uso de um lexema nas realizações lingüísticas. O coeficiente de uso é função da frequência e da dispersão. A título de exemplo: segundo os dados do FDPW o coeficiente de uso de *bonito* é 19,73 — o do artigo *o* é 23.758,10. A enorme disparidade entre esses dois números evidencia claramente a importância desse parâmetro como propriedade fundamental das palavras dentro do léxico da língua.

3. Evidenciada a dimensão quantitativa do léxico, vejamos quais os materiais de maior utilidade elaborados pelo computador a serem usados numa empresa lexicográfica. São eles: as listas de frequência de palavras, acrescidas de alguns parâmetros de especial interesse (dispersão ou distribuição, uso, FPM) e as concordâncias.

As listas de frequência de palavras ou *índices verborum* são basicamente de dois tipos: a) lista em ordem decrescente de frequência; b) lista das palavras em ordem alfabética com a indicação de seus respectivos parâmetros. Na primeira lista as palavras são ordenadas hierarquicamente, partindo-se da palavra mais frequente até aquelas que ocorreram uma só vez no corpus (os *happax legomena*). Os comentários acima feitos sobre o AHWFB evidenciam a existência de blocos com características típicas dentro dessas listas. Em primeiro lugar o bloco das palavras de alta frequência, geralmente palavras instrumentais e umas poucas palavras lexicais de significação plena (alguns verbos, substantivos e adjetivos). No caso do português as 100 palavras mais frequentes incluem:

4 artigos: a, o, um, uma;

12 adjetivos: algum, aquele, esse, este, grande, mesmo, meu, novo, outra, português, seu, todo;

17 advérbios: agora, ainda, assim, bem, depois, então, hoje, já, mais, menos, muito, não, onde, sempre, só, também, tão;

conjunções: como, mas, nem, ou, quando, se;

numerais: dois;

preposições: a, até, com, de, em, entre, para, per, por, sem, sobre;

pronomes: ela, elas, ele, eles, eu, isso, nos, o que, qual, que (relativo), se (passivo), tudo;

substantivos: ano, casa, coisa, dia, estudo, homem, mulher, pai, parte, tempo, vez, vida;

verbos: chegar, dar, deixar, dever, dizer, encontrar, estar, falar, ficar, haver, haver (impessoal), ir, parecer, passar, poder, querer, saber, ser, ter, ver, vir.

Note-se que essas palavras constituíram 61,98% do corpus total!

As 10 palavras mais freqüentes do FDPW na ordem decrescente de freqüência foram:

1.º o (art.)	4.º que (pn. rel.)	8.º a (prep.)
2.º a (art.)	5.º ele (pn. pes.)	9.º que (conj.)
3.º de (prep.)	6.º ser (v)	10.º em (prep.)
	7.º ela (pn. pes.)	

No outro extremo da lista temos os *happax legomena*, na sua maioria substantivos. Aliás, todos os dicionaristas e estatísticos léxicos sabem que o substantivo é a classe de palavra mais importante. Em todas as línguas estudadas o substantivo ultrapassa de muito numericamente todas as demais categorias sendo, portanto, o tipo de vocábulo mais significativo na composição do léxico. Também do ponto de vista lexicográfico, o substantivo é a pedra de toque na definição de qualquer palavra. As baixas freqüências (2 a 5 ocorrências) são povoadas de substantivos e palavras de significação plena: *aperfeiçoamento, assentimento, absolver, atormentar, afetivo, aliviado*. Há muitos advérbios de modo entre essas palavras: *consideravelmente, positivamente, sensivelmente, sucessivamente* (dados do FDPW). Ora, o advérbio de modo em *-mente* tem um status gramatical curioso. Trata-se de uma palavra de significação lexical por ser formado de adjetivo qualificativo + sufixo *-mente*; e tem um valor gramatical bem distinto de outras palavras por não ter relação de dependência com outras formas léxicas, pois tem colocação praticamente livre na cadeia do discurso.

Os *indices verborum* têm interesse lexicográfico sobretudo quando acompanhados de parâmetros como a dispersão (ou distribuição), o uso, e a FPM. De fato, não basta saber que o verbo *poder* foi utilizado 1.282 vezes num total de 500.000 ocorrências; ou que a preposição *de* foi utilizada 25.313 vezes no mesmo número total de ocorrências. O índice de uso dessas duas palavras aponta claramente a sua importância na língua portuguesa. Para a preposição *de* o coeficiente de uso foi de 21.688,13 e para o verbo *poder*, 1.084,70, relativamente aos dados do corpus do FDPW. Compare-se com o mesmo índice para *bonito* — 19,73 — ou de substantivos como *pensamento* (coeficiente de uso: 54,55; freqüência: 85), *pesquisa* (uso: 3,61; freqüência: 3), *povo* (uso: 85,28; freqüência: 153).

ET ARMÉ DE LA MASSUE, IL SA ÉLANCE FIEREMENT M9984374 RA1144 UNIVERSELLE DU MONDE I IL EST REVETU DE LA PEAU DU LION
ET PARTOUT L' EMPREINTE DE SA GRIFFE, O POUR K2324254 RA1145 QEN DEPIT DE SA FORCE ET DE SA MASSE, CV EST LE PAS OU LION
ET L'ANTILPE D' *HENRI *ROUSSEAU, DEUX TO K7654474 RA1146 , DONT LES GRAVURES FONT UN EFFET GROTESQUE, LE LION
ET LES LOUPS VOIENT *MILON SAISI PAR SA VICTIM M8991159 RA1147 SES DEUX FLANCSCOMME DESTENAILLES INFLEXIBLES, LES LIONS
ET LES A TUÉS DE SA MAIN, IL A VU UN GRAND CH M899115B RA1148 . *MILON *DE *CROTONE, - MILON A JOUË AVEC LES LIONS

Em 1984 o Instituto Nacional de Investigação Científica de Lisboa publicou o *Português Fundamental* (vol. I, Vocabulário), pesquisa sobre o vocabulário básico do português desenvolvida pelo Centro de Linguística da Universidade de Lisboa desde 1970, inicialmente sob a direção do Prof. Lindley Cintra e depois do Prof. João Málaca Casteleiro. Esse projeto tinha por objetivo estabelecer o léxico básico para o ensino do português a estrangeiros. Utilizou-se o modelo de pesquisa do *Francês Fundamental* e do *Espanhol Fundamental*. O corpus que originou o vocabulário de *Português Fundamental* continha 700.000 ocorrências de palavras, coletadas em entrevistas realizadas com 1.400 informantes. A esse acervo de registros do discurso oral, somaram-se os inquéritos escritos, que complementaram os dados colhidos na primeira fase da pesquisa. Os dados lingüísticos foram tratados com um rigoroso modelo de análise de Estatística Léxica, gerando listas de freqüências de palavras. Dessas listas foram selecionadas 2.217 palavras de uso comum, consideradas como o vocabulário básico do português na comunicação oral. Esse trabalho léxico-estatístico constitui uma das tarefas lexicológicas que precisavam ser feitas sobre o nosso idioma. A análise da lista dessas 2.217 palavras mostra que não há coincidência total entre o vocabulário básico do português de Portugal e da variante brasileira. Várias palavras precisariam ser acrescentadas para servir à comunicação falada no Brasil, assim como várias outras poderiam ser eliminadas. Tais diferenças são facilmente explicáveis pela Sociolingüística, em virtude das variações sociais e geográficas da língua nos dois países — Portugal e Brasil.

As concordâncias de textos (KWIC ou Key-Word-In-Context) são listas dos contextos em que ocorreu uma determinada palavra (a palavra-chave) posta em evidência, geralmente à esquerda da saída (print-out) dos dados impressos pelo computador. Por exemplo: *corpo* Carlôs é um rapaz forte: tem ★ ★ corpo ★ ★ de atleta.
O ★ ★ corpo ★ ★ do prefeito está sendo velado na prefeitura.
O ★ ★ corpo ★ ★ da narrativa é constituído por uma trama policial.
No ★ ★ corpo ★ ★ central da casa havia um átrio.
Normalmente antecedem (ou seguem) os dados em código, os quais permitem localizar a passagem, o texto, e a obra de onde foi extraída a citação.

Na elaboração do *Trésor de la langue française* foram produzidas dois tipos de concordância:

1.º) um microcontexto de uma linha (18 a 20 palavras) em que está inserida a palavra-chave:

M40138, "Et, Dérobant L'Éclair à L'Inconnu Sublime," Lier Ce Char d' Un Autre a des Chevaux à Toi? ou M2780325
M40139ET Ligne Directe Ascendanté, La Première Est Celle Qui Lie Le Chef Avec Ceux Qui Descendent de Lui; La D M4870021
M40140 Sur La Peau, Puis, La Douleur Changeait: On Lui Liait La Cheville Avec Un Fil de Fer, On Lui Raidis L4630171
M40141 Veines Qui Tremblent Sur La Main Comme Les Cordes Qui Lient Un Chevreau; — Père * Janet, Qui Est — Ce Qu K6270001

2.º) um macrocontexto de 8 linhas onde ocorreu a palavra-chave:

L679 Maritain Humanisme Intégral 1936
1 Par l'économique. Et c'est bien plutôt la mise de
2 Côté, le rejet dédaigneux de toute idéologie
3 Métaphysique comme expression ou reflet transitoire
4 D'un moment économique, qui, en un sens, et malgré
5 L'insistance avec laquelle le marxisme populaire

6 Exploite ce thème, est une apparence théorique
7 Illusoire, ou, comme les arguments des vieux
8 Sceptiques grecs, un thème drastique destiné à

Thomas M. Paikeday desenvolveu um programa para microcomputador que produz concordâncias de dois tipos:

1.º) microcontexto de 128 caracteres (média de 22 palavras):

The New York Times Everyday Dictionary, TEXT ANALYSIS BY MICROCOMPUTER, Page 2

had the flu, not toxic shock, and that no definitive link *between* Rely and the ailment had ever been established. Kidney arguments of automobile emissions. "There is a positive association *between* the removal of lead from gasoline and the lower blood-lead level swallowing air in a person with that habit. Hold an eraser *between* your teeth *between* meals to help prevent the swallowing mechanism.

2.º) contextos de 14 linhas:

My strength, and then I would go out to put some hurt on the people
a place in orbit where you can go out for a good time on a Saturday
For example, if classified ads go electronic, make sure they're
with only three more states to go, Carter let the movement stumble
contributions, How many politicians go around voting against churches
on auto loans. And so it would go, through mayors supporting the
return to take a vacation. Now go. No whimpering. Not another word
analysis has some merit, I would go further. I really wonder whether
she boarded his helicopter to go to Camp David shortly after war
return for a commitment not to go the plutonium route has never
comeuppance, and now things can go back to normal. There could be no
prisons. The other group would go easy on the criminal and try to
ir, and leave him limp while I go call the police. Would our people
, but that doesn't mean he can go around shooting people. The press
.....
Isaac A. Levi (AP, Mexico City): Elections, Col. Disp., 5 Jul 82

Os programas ALPHA e GLOBAL que geraram essas concordâncias (as do inglês) têm a grande vantagem de não precisarem de um computador de grande porte como aquele utilizado em Nancy, França (as concordâncias do francês acima transcritas) para o *Trésor de la langue française*, e rodarem em um microcomputador, o que coloca ao alcance de mais indivíduos e entidades a possibilidade de gerar concordâncias para trabalhos lexicográficos. De fato, é possível executar essas obras de referência básica em lexicografia como *indices verborum* e concordâncias dispondo de equipamento menos sofisticado que aqueles do *Institut de la Langue Française*. Infelizmente os programas criados por Paikeday para o inglês precisariam ser adaptados para a língua portuguesa, a fim de servirem ao tratamento automático do português e isso não é uma empresa fácil.

As concordâncias de texto são um material riquíssimo para documentar e ilustrar usos semânticos e gramaticais e atestar o que está ocorrendo de fato na língua, quando se trata de um trabalho sobre Lexicografia contemporânea. Um dos mais importantes serviços que uma concordância pode prestar é relativa à sintagmática. Fornecendo-se em bloco seqüencial as ocorrências de uma palavra em seus contextos,

ela fornece ao lingüista, lexicógrafo, ou gramático, o conjunto das combinações e das colocações que a palavra em epígrafe pode ter. Para serem perfeitas as concordâncias exigiriam uma lematização, trabalho de alta sofisticação para um computador, assunto esse que discutiremos mais abaixo.

4. Uma das empresas mais arrojadas no domínio da automação da lexicografia é o “dicionário de máquina”. Um trabalho pioneiro dessa natureza foi o executado por uma equipe de lingüistas e especialistas em computação e matemática do *Istituto di Linguistica Computazionale* de Pisa, equipe essa liderada por G. Ferrari e I. Prodanof. A pesquisa ainda prossegue atualizando-se e aprimorando-se os resultados.

A concepção original do “Dicionário de máquina do Italiano” (DMI) teve como objetivo criar um instrumento básico para lematização automática de textos fornecidos ao computador para análise. Ferrari se baseou num dicionário comum de uso do italiano: N: Zingarelli — *Vocabolario della Lingua Italiana*. A partir desse banco de dados léxicos sobre o italiano contemporâneo, os pesquisadores elaboraram uma lista de formas lematizadas do italiano (forma canônica) de qualquer paradigma ou conjunto das variantes flexionais. Para fazer a máquina operar foi criada uma série de algoritmos, capazes de reconhecer as formas flexionadas e remetê-las à forma lematizada, ou lema.

Para se ter uma idéia da complexidade dessas operações, basta lembrar que o DMI contém 912.618 formas geradas automaticamente (lemas), a partir de 106.152 entradas léxicas. Um dos refinamentos que está sendo elaborado, diz respeito ao difícil problema dos homógrafos no tratamento automático. Os pesquisadores estão tentando reduzir a extensão do dicionário e tornar mais econômico o “look-up”, isto é, a busca das formas em todo o dicionário. Para isso tentam unificar certas classes de homógrafos. Convém lembrar que a homografia constitui um dos mais desafiadores problemas para a lingüística computacional. A propósito: até esta data a distinção entre homógrafos ainda está sendo feita manualmente. Os pesquisadores estão tentando refinar o seu analisador sintático-semântico para poder automatizar esta operação ao menos parcialmente.

O DMI tem operado satisfatoriamente no reconhecimento das formas lingüísticas dos textos processados com o seguinte rendimento porcentual: 86% em textos de jornais, chegando a 96% em textos literários. No caso de textos jornalísticos os maiores embaraços são criados por nomes (geográficos e onomásticos) que, naturalmente, constituem pequena parcela do DMI. E em ambos os casos os homógrafos criam os maiores impasses na análise e identificação automáticas. Por enquanto, os pesquisadores analisam os homógrafos manualmente, a despeito da lentidão dessa tarefa.

O DMI é um arquivo, um tesouro léxico da língua italiana contemporânea, estocado na memória do computador para ser consultado pela máquina quando ela for analisar qualquer texto automaticamente. Por isso a equipe de pesquisadores continua trabalhando na atualização dos dados estocados e no refinamento dos algoritmos que operam a identificação das formas. Esse refinamento diz respeito às tentativas de eliminação de ambigüidades quando o computador propõe mais do que uma lematização. Por isso se trabalha na criação de melhores analisadores sintático-semânticos. O conjunto de operações realizadas pela máquina constitui “uma tentativa de produzir um modelo computacional unívoco do léxico dentro de um modelo de língua natural” no dizer de G. Ferrari. Esse lingüista e sua equipe acreditam que “embora cada realização lingüística se inclua no nível do desempenho, existe um nível de competência para o léxico, que justifica nossa proposta em favor de uma concepção unitária”. (2, p. 8)

Por tudo o que foi dito acima o DMI não é uma simples lista de unidades léxicas; e “deve ser considerado como uma máquina composta de um núcleo de dados e de uma

série de procedimentos que operam em vários níveis lingüísticos. Em outras palavras, cada nível da análise exige a criação de um novo modelo lingüístico a ser associado ao dicionário e a ser executado em conexão com ele. Neste sentido, o dicionário está situado entre a lexicologia, a morfologia, a sintaxe e a semântica. Se essa posição em relação às duas primeiras é relativamente clara, entre as últimas a fronteira não é precisa, sobretudo com relação ao léxico". (2, p. 9)

Várias outras tentativas estão sendo feitas contemporaneamente, mais ou menos similares ao DMI, em outros centros de computação lingüística, mas com objetivos diferentes e concepção original bem distinta. Em Grenoble (grupo GETA) do Centro de Tradução Automática, vários transcodificadores estão em operação (e sendo aprimorados) com vistas à tradução automática das várias línguas da comunidade européia. Na equipe do Prof. Vauquois em Grenoble, está trabalhando nosso colega Paltônio Daun Fraga no "autômato de estados finitos não determinados" para a análise automática do português e sua tradução em inglês.

O uso do computador em Lexicografia está revolucionando essa ciência secular. Atualmente grandes editores de dicionários e obras de referência usam o computador como instrumento básico nos trabalhos de coleta, seleção, armazenamento e recuperação de dados. A rigor, hoje não se pode planejar um dicionário-padrão da língua de porte médio, ou seja, 50.000 verbetes, sem o concurso do computador. Essa máquina pode realizar operações fundamentais na confecção de um dicionário com eficiência muito superior à do homem, a saber: 1) seleção das palavras-entrada para a composição do repertório vocabular do dicionário com base em critérios de Estatística Léxica; 2) controle das definições do dicionário, usando modelos uniformes e um vocabulário básico na formulação dos verbetes; 3) controle rigoroso das referências cruzadas; 4) concordâncias de palavras das quais os lexicógrafos extrairão a documentação dos significados e usos registrados no dicionário. Por maioria de razão, as obras de referência de grande porte não podem dispensar a tecnologia e os recursos científicos contemporâneos, pois os grandes dicionários devem partir de corpora de 30 milhões a 100 milhões de ocorrências de palavras, coligidos em uma gama muito variada de textos: literários, jornalísticos, técnicos, científicos etc. Além disso, existem outros recursos técnicos muito úteis em Lexicografia: a) leitura ótica; b) possibilidade de seleção de textos, correção e inserção de informação diretamente através do vídeo; c) classificação e ordenação automática dos contextos que serão utilizados, contextos esses de fácil recuperação; d) enorme potencial de armazenamento de dados com rápido acesso não-sequencial de um bloco de dados através do uso de discos; e ainda vários meios técnicos de tipo documental que hoje facilitam imensamente a penosa tarefa do lexicógrafo.

Creio que comprovamos suficientemente duas coisas: de um lado, o formidável instrumental que hoje dispomos para a confecção de dicionários e de obras de referência, especialmente o computador; de outro, que a máquina não pode dispensar o concurso do lingüista e do lexicógrafo. Ela chega até um certo ponto e até esse marco trabalha melhor e mais rapidamente que o homem; porém, quando o computador atinge seus limites, o lexicógrafo entra com a sua competência lingüística e com a sua inteligência operadora. Assim, ambos — homem e máquina — podem trabalhar como colaboradores para um produto final de melhor qualidade. Infelizmente, porém, o computador e todos os seus periféricos produzem material numa velocidade muito maior do que uma equipe de lexicógrafos pode digerir, para redigir os verbetes de um dicionário. Provavelmente ainda se passará muito tempo antes que lexicógrafos e analistas de sistemas consigam atingir o estágio ideal em que se criará um analisador automático de uma linguagem natural, capaz de esmiuçar um texto e distinguir como um lingüista compe-

tente as nuances sintático-semânticas do discurso produzido pelos homens. Talvez esse estágio de perfeito entrosamento entre o homem e a máquina não seja jamais atingido. De qualquer forma, a era dos computadores gerou uma autêntica revolução na Lexicografia. Esperemos que a moderna tecnologia enseje a produção de dicionários cada vez mais perfeitos sem que o lexicógrafo se transforme num obscuro mártir da ciência que pratica

BIDERMAN, M.T.C. — The science of Lexicography. *Alfa*, São Paulo, 28(supl.):1-26, 1984.

ABSTRACT: Three aspects are discussed: 1) a short history of Lexicography in Spanish, French and Portuguese. The most important dictionaries of these languages are evaluated, from the XVth to the XXth century. 2) Typology of lexicographical works. The main types of existing dictionaries in Latin languages and English are commented on. 3) Use of computers in modern Lexicography. The computer caused a revolution in Lexicography owing to its capacity to perform basic and tedious tasks such as: compiling, classifying and ordering lexical and contextual data for the organization of dictionaries. In addition, they provide rapid retrieval facilities.

KEY-WORDS: Lexicography; thesaurus; dictionary; historical dictionary; etymological dictionary; conceptual dictionary; technical dictionary; scientific dictionary; frequency dictionary; lexical data bank; corpus; indices verborum; concordance; machine dictionary.

REFERÊNCIAS BIBLIOGRÁFICAS

1. AITKEN, A.J. — Historical dictionaries, word frequency, distributions and the computer. *Cahiers de Lexicologie*, 3(1): 28-47, 1978.
2. FERRARI, G. & PRODANOF, I. — Machine dictionary and lexicon. IN: INTERNATIONAL CONFERENCE ON COMPUTATIONAL LINGUISTICS, Ottawa, 1976 (Comunicação)
3. KELLER, H.H. — The American heritage word frequency book (review). *Language Learning*, 25: 173-178, 1975.
4. MARCOS-MARÍN, F. — *Curso de gramática española*. Madrid, Cincel-Kapelsz, 1980.
5. MATORÉ, G. — *Histoire des dictionnaires français*. Paris, Larousse, 1968.
6. MELO, G.C. de — *Dicionários portugueses*. Rio de Janeiro, S.O.M.E.S., 1947.
7. MORAIS SILVA, A. de — *Dicionário da língua portuguesa*. — Fac-simile da segunda edição, 1813, photographada pela Revista de Língua Portuguesa. Rio de Janeiro, Oficinas de S.A. Littro Typographia Fluminense, 1922. 2v.

BIBLIOGRAFIA CONSULTADA

- BALDINGER, K. — Semasiologia e onomasiologia. *Alfa*, 9: 7-36, 1966.
- BARNHART, C.L. — Plan for a Central Archive for Lexicography in English. *Annals of the New York Academy of Sciences*, 211: 302-319, 1973.
- BIDERMAN, M.T.C. — *Teoria lingüística*. Rio de Janeiro, Livros Técnicos e Científicos, 1978.
- BLUTEAU, R., Pe. — *Vocabulário português e latino*. Lisboa, Colégio das Artes da Cia. de Jesus, 1712-1728. 8v., 2 supl.
- CAPPELLI, A et alii — *Parsing an Italian text with an ATN Parser*. Pisa, Instituto di Lingüistica Computazionale, 1978.
- CASARES, J. — *Introducción a la lexicografía moderna*. Madrid, C.S.I.C., 1950.
- COSERIU, E. — Vers une typologie des champs lexicaux. *Cahiers de Lexicologie*, 2(2): 30-51, 1975.
- DELATTE, L. et alii — Le traitement automatique de la langue française au laboratoire d'analyse statistique des langues anciennes. *Revue des Études Anciennes*, 4:1-55, 1977.
- DICCIONARIO DE LA LENGUA ESPAÑOLA /por/ Real Academia Española. 18.ed. Madrid, Espasa-Calpe, 1956.
- DRETTAS, G. — Les théoriciens allemands du champ. *La Linguistique*, 1(2): 3-22, 1981.
- DUBOIS, J. — *Introduction à la lexicographie: le dictionnaire*. Paris, Larousse, 1971.

- DUNCAN, J. — *Frequency dictionary of Portuguese words*. Stanford, Stanford University, 1972. (Ph. D. Dissertation)
- FERRARI, G. — Dictionnaire automatique et dictionnaire-machine: une hypothèse. In: COMPUTATIONAL AND MATHEMATICAL LINGUISTICS. PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON COMPUTATIONAL LINGUISTICS, Pisa, 1973. Firenze, Leo S. Olschki Ed., 1977. p. 257-262.
- FERREIRA, A.B. de H. — *Novo dicionário da língua portuguesa*. Rio de Janeiro, Nova Fronteira, 1975.
- LARA, L. F. — Méthode en lexicographie: valeur et modalité du dictionnaire de machine. *Cahiers de Lexicologie*, 29(2): 103-28, 1976.
- LONGMAN DICTIONARY OF CONTEMPORARY ENGLISH. London, Longman Group, 1978.
- MC NAUGHT, J. — Terminological data banks: a model for a British linguistics data bank (LDB). *ASLIB Proceedings*, 33(7/8): 297-308, 1981.
- MACHADO, J.P. — *Dicionário etimológico da língua portuguesa*. Lisboa, Ed. Confluência, 1956. 2v.
- MOLINER, M. — *Diccionario de uso del español*. Madrid, Gregos, 1975.
- MORAIS SILVA, A. de — *Grande dicionário da língua portuguesa*. 10. ed. rev. por J.P. Machado. Lisboa, Ed. Confluência, 1949-1957. 12v.
- MURAKAWA, C. de A.A. — *O primeiro dicionário de língua portuguesa de Antonio de Moraes Silva. Estudo crítico da edição de 1813*. Araraquara, Instituto de Letras, Ciências Sociais e Educação, Unesp, 1984. (Dissertação de Mestrado)
- NASCENTE, A. — *Dicionário etimológico da língua portuguesa*. 2a. tiragem da 1.ª ed. Rio de Janeiro, Francisco Alves, s.d.
- PAIKEDAY, T.M. — Language analysis and lexicography by microcomputer. (Comunicação feita no encontro ADS-MLA, 1981).
- PAIKEDAY, T.M. — *The New York Times everyday dictionary*. New York, Times Book, 1982.
- PORTUGUÊS FUNDAMENTAL. Lisboa, Instituto Nacional de Investigação Científica, Centro de Linguística da Universidade de Lisboa, 1984. v.1, Vocabulário e Gramática; T. 1, Vocabulário.
- PRODANOF, I. — A la recherche d'un modèle de derivation en italien. In: COMPUTATIONAL AND MATHEMATICAL LINGUISTIC. PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON COMPUTATIONAL LINGUISTICS, Pisa, 1973. Firenze, Leo S. Olschki Ed., 1977. p. 297-301.
- RAVEN, I. et alii — Lexicography in English. *Annals of the New York Academy of Sciences*, 211, 1973.
- REY-DEBOVE, J. — Le domaine du dictionnaire. *Langages*, (19):3-34, spt., 1970.
- REY-DEBOVE, J. — Lexique et dictionnaire. In: LE LANGAGE: Les dictionnaires du savoir moderne. Paris, Centre d'Étude et de Promotion de la Lecture, 1973. p. 82-108.
- RICHMAN, B. et alii — *The American heritage word frequency book*. New York, Boston American Heritage Publ., Houghton Mifflin, 1971.
- ROBERT, O. — *Dictionnaire alphabétique et analogique de la langue française*. Paris, SNL, Dictionnaire Le Robert, 1972.
- SHERMAN, D. — Retrieving lexicography citations from a Computer Archive of Language Materials. *Annals of the New York Academy of Sciences*, 211:137-142, 1973.
- SHERMAN, D. — Special purpose dictionnaires. *Cahiers de lexicologie*, 32(1): 82-102, 1978.
- TERMINO GRAMME: Bulletin de la Direction de la terminologie. Québec, Office de la Langue Française, 1979.
- THE OXFORD ENGLISH DICTIONARY. Oxford, Clarendon Press, 1933.
- VENEZKY, R.L. — Computer applications in lexicography. *Annals of the New York Academy of Sciences*. 211:287-291, 1973.
- VENEZKY, R.L. — Storage, retrieval and editing of information for a dictionary. *American Documentation*, 19:71-79, 1968.
- VIEIRA, D. Frei — *Grande dicionário português ou thesouro da língua portuguesa*. Porto, Ed. Ernesto Chardon e Bartholomeu H. de Moraes, 1871.
- VITERBO, J. de S.R. de, Frei — *Elucidario de palavras e frases que em Portugal antigamente se usarão (sic) e que hoje regularmente se ignorarão*. Lisboa, 1978-1979.
- WEINREICH, U. — Lexicographic definition in descriptive semantics. In: HOUSEHOLDER, F.W. & SAPORTA, S., eds. — *Problems in lexicography*. Bloomington, The Hague, Mouton, 1967. p. 25-44.
- WEINREICH, U. — Webster's third: a critique of its semantics (review). *International Journal of American Linguistics*, 30:405-409, 1964.
- WICLOW, C.K. — Advanced English vocabulary. *Language Learning*, 24(1): 167-170.
- ZGUSTA, L. — *Manual of lexicography*. The Hague, Mouton, 1971.
- ZINGARELLI, N. — *Vocabolario della lingua italiana*. 10. ed. Bologna, Zanichelli, 1971.

O DICIONÁRIO PADRÃO DA LÍNGUA

Maria Tereza Camargo BIDERMAN*

RESUMO: Os dicionários unilíngües têm vários tamanhos e formatos. O número de verbetes que contém depende do público a que se destina cada dicionário. O dicionário padrão da língua é um dicionário unilíngüe com 50.000 verbetes aproximadamente, incluindo um acervo léxico substancial, sem constituir, porém, um thesaurus que recolhe todas as palavras do léxico. O dicionário padrão é um instrumento cultural muito importante na sociedade contemporânea. Neste artigo são examinados vários problemas relacionados com a confecção de um dicionário padrão e de dicionários unilíngües: a seleção das entradas lexicais, a compilação do corpus (banco de dados), a redação dos verbetes, a definição lexicográfica. Também se analisa a problemática da polissemia e da homonímia na elaboração do dicionário.

UNITERMOS: Dicionário unilíngüe; dicionário padrão; thesaurus; corpus; banco de dados; palavra-entrada; verbe; lematização; forma canônica; índices verborum; concordância; definição lexicográfica; sinônimo; antônimo; referência cruzada.

I. DICIONÁRIOS UNILÍNGÜES E DICIONÁRIO PADRÃO

O dicionário padrão da língua e os dicionários unilíngües são os tipos mais comuns de dicionários. Em nossos dias eles se tornaram um objeto de consumo obrigatório para as nações civilizadas e desenvolvidas. Existem várias modalidades e vários tamanhos de dicionários unilíngües. O tamanho físico é normalmente função da riqueza do repertório léxico nele incluído. Atualmente o planejamento de dicionários da língua se fundamenta nos resultados de quarenta anos de pesquisas em Estatística Léxica. Esses dicionários unilíngües geralmente se distribuem em algumas categorias, dependendo da sua destinação e do tipo de usuário para o qual ele é planejado, a saber:

- 1) o dicionário infantil e/ou básico com 5.000 verbetes aproximadamente;
- 2) o dicionário escolar e/ou médio contendo 10.000 - 12.000 verbetes, podendo totalizar até 30.000 verbetes;
- 3) o dicionário padrão com uma média de 50.000 verbetes, um pouco mais, um pouco menos;
- 4) os "thesauri" que podem incluir 100.000, 200.000, 500.000 verbetes.

Um dicionário padrão é um instrumento para orientar os seus consulentes sobre os significados e os usos das palavras e para que eles possam expressar suas idéias e sentimentos com a maior precisão e propriedade possíveis, utilizando o tesouro léxico que a língua põe à disposição dos falantes do idioma.

O tesouro vocabular se enquadra dentro de um momento histórico da evolução da língua e dentro de uma determinada *norma cultural*. Jean Dubois define assim a norma cultural:

*Departamento de Lingüística - Instituto de Letras, Ciências Sociais e Educação - UNESP - 14800 - Araraquara - SP.

“Essa norma não é definida apenas pela aceitabilidade de todos os termos e de todas as frases contidas no dicionário, mas também por aquela dos enunciados engendrados pelo modelo sociocultural. Os termos não remetem apenas às palavras da língua; eles não são somente objetos da metalingua linguística; eles remetem também a enunciados culturais, a uma visão de mundo. (.....) o dicionário visa tornar-se uma *norma explícita da cultura da comunidade*. A sanção lexicográfica se identifica à sanção pedagógica: aquele que emprega termos não contidos no “tesouro” comum se exclui da comunidade nacional.” (2, p. 99-100)

O dicionário de tipo padrão tende a exercer um papel normativo dentro da comunidade dos falantes. Na sociedade brasileira contemporânea o *Aurélio* vem exercendo essa função há uma década. Se o leitor se reportar à crítica feita ao *Aurélio* no capítulo anterior, verá que o *Aurélio* assumiu o papel de *norma lingüística* em virtude de não existir uma obra de mesmo tipo e melhor elaborada que concorresse com ele, uma vez que toda nação civilizada contemporânea carece desse instrumento cultural. No caso em apreço, até o fato aparentemente secundário de o *Aurélio* constar de um só volume, ao passo que outros dicionários similares do português foram impressos em dois ou mais volumes, é importante. Para o consulente médio é muito mais prático ter sobre a escrivaninha ou na estante um dicionário padrão em um volume. Já não seria tão conveniente consultar dois, três ou quatro volumes. Note-se que o usuário atribui ao dicionário o conhecimento inequívoco da linguagem a fim de dirimir as suas dúvidas. Em outras culturas o papel exercido pelo *Aurélio* no Brasil, é preenchido pelo *Petit Robert* no caso do francês, pelo *Zingarelli*, caso do italiano, pelo *Webster, Longman, Concise Oxford*, caso do mundo de língua inglesa.

O dicionário é um instrumento cultural que remete tanto à língua como à cultura. O lexicógrafo descreve ambas — língua e cultura — como um todo pancrônico, embora se situe numa perspectiva sincrônica. Um dicionário é constituído de entradas léxicas que ora se reportam a um termo da língua, ora a um elemento da cultura. A entrada tem como seu eixo básico a definição da palavra em epígrafe. Essa definição nada mais é que uma perífrase metalingüística da palavra posta como entrada. Essa é a pedra de toque da tarefa lexicográfica, nem sempre executada adequadamente. Nas culturas ocidentais os dicionários têm-se copiado uns aos outros perpetuando erros e inadequações.

Um dicionário padrão (50.000 verbetes) possui um repertório léxico que nenhum usuário jamais utilizará, quer no seu vocabulário ativo, quer no seu vocabulário passivo, mesmo os mais competentes falantes do idioma. Um *thesaurus* (100.000, 200.000 palavras ou mais) inclui necessariamente um grande número de palavras raras, desusadas, obsoletas, outras que são exclusivamente literárias, um grande número de termos técnicos e científicos, de regionalismos e de neologismos.

Alguns exemplos dos tipos de dicionários acima referidos:

1. *Dictionnaire du vocabulaire essentiel* (les 5.000 mots fondamentaux.) George Matore. Paris, Larousse, 1963.

Nouveau Larousse des débutantes. René Lagane et alii. Paris, Larousse, 1977. (16.000 aproximadamente).

2. *Macmillan Dictionary for Children*. New York, Macmillan, 1975.

Children's Dictionary. An American Heritage Dictionary. New York, Houghton Mifflin, 1979.

Esses dois dicionários da língua inglesa destinam-se a escolares de 9 a 15 anos.

Oxford Intermediate Dictionary (J.A. Augarde). Oxford University Press, 1982 (12.000 palavras). Destina-se a escolares de 10 a 15 anos.

Micro Robert (dictionnaire du français primordial). Paris, S.N.L. Le Robert, 1979 (30.000 verbetes).

3) *Dicionário da Língua Portuguesa* (Antônio de Morais Silva). Lisboa, Tipografia Lacerdina, 1813 (40.000 verbetes aproximadamente).

The Concise Oxford Dictionary of Current Words. Oxford University Press (40.000 palavras aproximadamente).

Longman Dictionary of Contemporary English. London, Longman Group Limited, 1978 (55.000 entradas).

4) Na categoria de *thesaurus* podem ser incluídos desde os pequenos “tesouros” léxicos até os magnos “tesouros”, isto é, os de 100.000 até 500.000 verbetes, a saber:

Vocabolario della lingua italiana. (N.Zingarelli). Milão, Stampa OFSA, 1970. (106.152 entradas).

The Oxford English Dictionary. Oxford University Press, 1933. (414.825 palavras).

The American Heritage Dictionary of English Language. New York, American Heritage Publishing Company & Houghton Mifflin Company, 1970 (155.000 entradas).

Webster's New International Dictionary 3rd. ed., 1964. (450.000 verbetes).

Novo Dicionário da Língua Portuguesa (Aurélio Buarque de Holanda Ferreira). Rio de Janeiro, Nova Fronteira, 1975 (mais de 100.000 verbetes).

Grand Larousse de la langue française. (Guibert et alii). (7 volumes).

Dictionnaire de la langue française du 19ème. et du 20.ème siècle Paris, Klincksieck. (15 volumes planejados com aproximadamente 500.000 entradas; 8 já publicados).

O tamanho físico do dicionário, seu tipo de impressão, a existência ou não de ilustração, o seu custo dependerão dos objetivos do dicionário. Além do significado, o dicionário deve responder às dúvidas do consultante sobre a sintaxe das palavras, tais como: regência preposicional, combinações possíveis ou inaceitáveis; deve fornecer ainda informações adicionais sobre a forma do vocábulo-entrada e o paradigma de que ele eventualmente faça parte.

II. A CONFECÇÃO DE UM DICIONÁRIO

Na elaboração de um dicionário padrão e/ou *thesaurus* a equipe de dicionaristas deverá possuir certas qualidades e virtudes e operar segundo procedimentos hoje padronizados. Os lexicógrafos devem conhecer muito bem a sua língua materna e ter uma ampla leitura do seu patrimônio literário e cultural de todas as épocas no caso de idioma de longa tradição cultural como é o caso do português. Devem conhecer igualmente variantes faladas da língua. E devem saber que vão executar uma tarefa científica e cultural que se assemelha muito ao labor dos monges na Idade Média, os quais se aplicavam dedicada e apaixonadamente à cópia de manuscritos e/ou traduções de textos clássicos e científicos de outras línguas, ritualmente, dia após dia, durante toda a sua vida. O dicionarista precisa ser como esse monge. Descrito o perfil do nosso cientista, vejamos como os lexicógrafos devem executar a sua missão.

Um dicionário deve fundamentar-se na recolha de dados léxicos e lingüísticos para a constituição de um corpus representativo da língua escrita e falada. Esse corpus funcionará como fonte de informações sobre o léxico, além de fornecer as abonações dos significados, dos usos e das construções das palavras-entrada do dicionário. Para a elaboração desse arquivo ou banco de dados léxicos, deve-se considerar um período de tempo que represente uma etapa da evolução da língua em que ela possa ser considerada relativamente uniforme. Um corte sincrônico pode recobrir o espaço de uma geração, ou seja, uns trinta anos. O banco de dados precisa incluir todas as variantes escri-

tas da língua e não apenas a linguagem literária. Farão, pois, parte desse arquivo: textos literários, jornalísticos (jornais e revistas), textos técnicos e científicos de todas as áreas do conhecimento humano. Dada a importância dos meios de comunicação de massa no mundo contemporâneo, sobretudo da televisão, é preciso coligir também material da linguagem coloquial dos vários meios de comunicação de massa. O volume de palavras e textos dependerá do tamanho do dicionário que se projeta. Certamente, porém, alguns milhões de ocorrências de palavras (5 milhões no mínimo) distribuídas equitativamente entre as várias modalidades de linguagem e de áreas do conhecimento. O banco de dados será processado no computador e aí armazenado. Dele se extrairão dois produtos básicos para a confecção do dicionário:

1) *indices verborum*: as palavras recolhidas serão ordenadas hierarquicamente em ordem decrescente de frequência (das mais frequentes às menos frequentes). É importante que se obtenha outros parâmetros estatísticos sobre as palavras, além da frequência, tais como a distribuição e/ou dispersão nos vários *subcorpora* que compuseram o banco de dados.

2) Uma concordância das palavras em contexto, ou seja, todas as palavras estarão agrupadas em blocos nos quais a mesma palavra-chave aparecerá em todos os diferentes contextos em que ela ocorreu.

De posse desses dois instrumentos básicos de referência, os dicionaristas poderão iniciar a primeira etapa da confecção do dicionário: a seleção das palavras que constituirão a nomenclatura do dicionário. O índice de frequência baliza os limites de cada tipo de dicionário. Assim, um dicionário infantil e/ou básico só deve conter as 4.000 ou 5.000 palavras mais frequentes do idioma, seja qual for o gênero escrito, seja qual for a variedade de língua. Um dicionário escolar provavelmente incluirá as palavras com frequência igual ou superior a 20 se o corpus for de 5 milhões de palavras, podendo incluir palavras dos “ranking” 10 a 20. Dificilmente incluirá palavras que ocorreram menos de cinco vezes no corpus. Já um dicionário padrão poderá incluir palavras com frequência igual ou superior a 5; em princípio, porém, não deve incluir os “happax legomena”, isto é, as palavras que ocorreram apenas uma vez no corpus. Esses vocábulos são geralmente palavras raras, exclusivamente literárias, obsoletas e deverão ser incluídas apenas num dicionário tipo *thesaurus*.

Selecionadas as 5.000, as 10.000, as 50.000, as 100.000 palavras-entrada que compõem o repertório léxico do dicionário, os lexicógrafos passarão a selecionar, nas concordâncias de textos, os contextos que servirão para abonar os significados, as construções e os usos a serem registrados. Nesse trabalho de seleção deve-se escolher os contextos ideais sob vários aspectos: a) os que melhor explicitam o sentido, uso ou construção que se quer descrever; b) os que efetivamente representam uma boa linguagem; c) os que documentam os diferentes registros lingüísticos, os vários níveis de linguagem. Seria ideal que essa seleção pudesse ser feita em um terminal de computador e que a máquina fosse programada para só registrar, num local de fácil acesso do no banco de dados, os contextos que se pretende usar. Assim quando o lexicógrafo pedisse ao computador os dados contextuais das concordâncias para o seu trabalho de redação dos verbetes não seria submergido por uma avalanche de papel impresso fornecido pela máquina.

III. REDAÇÃO DOS VERBETES

Na etapa de redação dos verbetes o coordenador da equipe de lexicógrafos deve discutir com os dicionaristas as normas a serem seguidas, pôr por escrito esses princípios e fornecer a cada membro da equipe uma cópia das mesmas.

Alguns modelos de verbete são tradicionais e se impõem. Distinguem-se totalmente as palavras de significação plena (substantivo, adjetivo, verbo, alguns advérbios) das palavras instrumentais (artigos, preposições, conjunções, pronomes etc.).

Todo verbete tem um formato típico: após a palavra-entrada na sua forma canônica ou lema, vem a indicação da categorização léxico-gramatical da mesma. Segue-se uma paráfrase do significado, ou das várias acepções de sentido no caso de palavras polissêmicas; ou ainda as construções e/ou usos, no caso de palavras instrumentais. A seguir, vêm outras informações gramaticais sobre a palavra-entrada do tipo: plural irregular, formas verbais irregulares etc. Nas línguas que têm uma complexa morfologia verbal e pronominal, como é o caso do português e de outros idiomas românicos, costuma-se incluir quadros paradigmáticos dos modelos de conjugação verbal e declinação pronominal no início do dicionário e depois, no interior do verbete, se remete a esses modelos. Fornecem-se também informações sobre a pronúncia e especificidades fônicas da palavra (p.ex. *poço*: no sing. o 1.º *o* é vogal fechada enquanto no pl. esse *o* é aberto; *poça*: pronuncia-se o 1.º *o* como vogal aberta). Esse tipo de informação é particularmente importante no caso de um idioma em que a língua escrita se distanciou muito da língua falada como ocorre com o inglês. Enfim, é útil cruzar referências entre si. Assim, por exemplo, aponta-se para vocábulos que compõem o campo léxico da palavra-entrada, fornecendo-se um esquema da sua rede de significações, tanto dos lemas de sentido próximo — os sinônimos — como das palavras de significado oposto — os antônimos.

Alguns itens lexicais são problemáticos com relação à forma. Às vezes o lexicógrafo se pergunta se uma combinação de duas ou mais palavras é um sintagma lexicalizado — uma lexia complexa — ou se é necessário segmentá-la em duas ou mais unidades léxicas. Nesse caso convém aplicar os testes de substituição e de inserção de palavras para decidir sobre o grau de lexicalização da seqüência de itens lexicais.

“1.º) Teste de substituição

Em seqüências como as saudações *bom dia* ou *boa noite* não podemos substituir o primeiro vocábulo por um outro adjetivo mais ou menos sinônimo, dizendo: *ótimo dia* e *ótima noite*, por exemplo. Por conseguinte, tudo leva a crer que *bom dia* e *boa noite* já estejam lexicalizados como lexemas.

2.º) Teste de inserção

Numa seqüência como *dor de cabeça* não diremos *dor* “terrível” *de cabeça*, mas “terrível” *dor de cabeça*, ou *dor de cabeça* “terrível”. Também não se diria *capa* “bonita” *de chuva* mas *capa de chuva* “bonita”. Tampouco se poderia inserir um advérbio em meio aos constituintes de *mercado negro*, resultando: *mercado* “muito” *negro*. Portanto, o teste de inserção demonstra que estão lexicalizados: *dor de cabeça*, *capa de chuva* e *mercado negro*.

Deve-se notar também que, nas lexias complexas já lexicalizadas, não mais ocorrem o artigo, a preposição e outros vocábulos-morfema. Assim em *de cor* já de há muito ocorreu a elipse do artigo, assim em *à medida que*, a elipse da preposição (em). Exemplos similares do inglês: *at hand*, *by heart*, *by hook or by crook*.

Entretanto, em virtude da inconsistência do código escrito, seremos obrigados a chamar de *lexias complexas* vocábulos como *bom dia*, *capa de chuva*, *dor de cabeça*, *mercado negro*, embora as alistemos como unidades categorizadas em língua e devamos colocá-las como legítimas entradas de dicionário. O termo *lexia simples* será, pois, reservado para as unidades que são grafadas como um único segmento.” (1, p. 133).

Também convém consultar o sentimento lingüístico de lexicógrafos, de outros lingüistas e até mesmo de falantes não especialistas. Ainda que se conclua em favor da segmentação da seqüência de palavras, o dicionarista deve incluir os sintagmas com forte tendência à lexicalização como subentradas no fim do verbete. Cf.

Em *cabeça*.
cabeça dura/ cabeça de vento/ andar de cabeça erguida/ levar na cabeça/ perder a cabeça

Em *passar*.
passar a limpo/ passar bem/ passar dos limites/ passar para trás.

Expressões idiomáticas deverão ser incluídas em um dicionário padrão sobretudo se são usadas com freqüência. É evidente que só um dicionário tipo *thesaurus* poderá registrar todos os sintagmas com forte grau de coesão lexical.

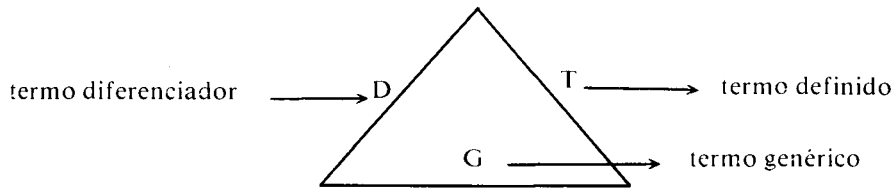
IV. A DEFINIÇÃO LEXICOGRÁFICA

Num dicionário unilingüe qualquer um problema fundamental é a definição da palavra-entrada. A definição de um vocábulo vem a ser uma paráfrase dessa palavra, equivalente a ela semanticamente. Essa paráfrase deve ser redigida em linguagem simples, escoreita e ter sido formulada utilizando-se palavras muito freqüentes na língua, preferivelmente lexemas que façam parte do vocabulário básico. Tal precaução garantiria, em princípio, a compreensão fácil do termo cujo significado o consulente desconhece. A técnica de utilizar um vocábulo básico da língua na elaboração dos verbetes foi utilizada com muita competência lexicográfica pelos autores do *Longman Dictionary of Contemporary English* (1978). Eles estabeleceram uma lista de 2.000 palavras aproximadamente, que foi selecionada com base em numerosas listas de freqüência e trabalhos pedagógicos sobre o inglês. Também foi elaborado um conjunto rigoroso de normas a fim de garantir que apenas os significados mais básicos dessas 2.000 palavras e seus derivados facilmente compreensíveis fossem utilizados pelos redatores do dicionário. Esse repertório de 2.000 palavras foi colocado em um programa de computador. A máquina recebeu instruções adicionais de controle para verificar cada palavra incluída nas definições dos verbetes e rejeitar qualquer vocábulo que não constasse do vocabulário em questão.

A definição de uma palavra-entrada — definição lexicográfica — não se confunde com a definição lógica. Em princípio, a definição lógica só é aplicável à classe dos substantivos. Às vezes, os verbos são definíveis de modo semelhante, mas não as demais classes de palavras.

“A definição lexicográfica e a definição lógica se recobrem até certo ponto, mas apresentam diferenças notáveis. Provavelmente a mais importante delas está em que, enquanto a definição lógica tem de identificar de modo inequívoco o objeto definido (*definiendum*) de tal modo que ele deva, por um lado, contrastar radicalmente com todos os outros objetos susceptíveis de definição, e por outro lado, caracterizar-se de modo positivo e inequívoco como membro da classe mais próxima, a definição lexicográfica enumera os mais importantes traços semânticos da unidade léxica que bastem para distingui-la das outras unidades.” (4, p 53-54).

Na introdução do seu dicionário (*Diccionario de uso del español* p. XV) Maria Moliner propõe um “triângulo definidor” da técnica da definição lexicográfica:



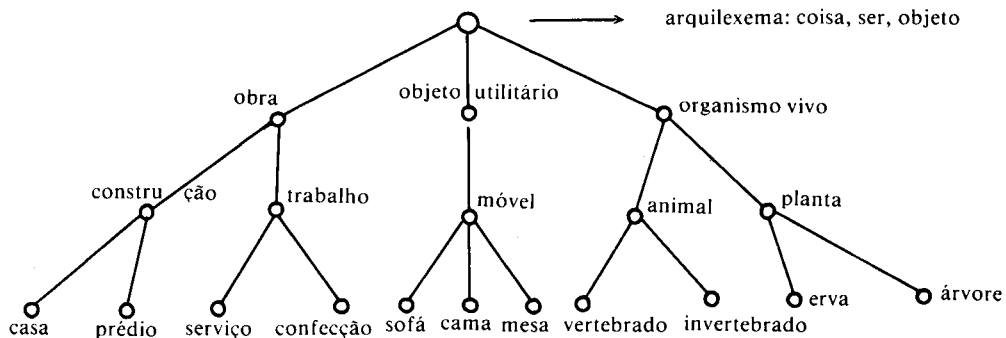
(a figura gráfica é minha)

“Uma definição é uma relação estabelecida entre três termos: termo definido (T), termo genérico (G), que é o conceito de conteúdo mais amplo em que aquele está compreendido, e termo diferenciador (D), que limita a extensão do termo G para que convenha exatamente a T.... As relações entre eles, quanto à sua extensão, são como segue: $G > T$; $G > D$; $D = T$. A maior extensão de G com respeito a T, significa que todos os indivíduos abarcados pelo conceito T estão incluídos no conceito G e este compreende também outros; a maior extensão de G com respeito a D, significa que D é aplicável como predicado somente a uma parte dos indivíduos compreendidos em G; do mesmo modo, a igualdade da extensão entre T e D significa que D é aplicável como predicado a todos os indivíduos abarcados pelo conceito T e, dentro do gênero G, somente a eles.” (3, p. XV) Por exemplo:

cravo:

instrumento musical, de teclado e cordas, diferente do
piano e mais antigo, mas semelhante a ele.

A estruturação das redes léxicas de significação é ascendente: parte-se do genérico, do mais básico para diferenciações progressivas e cada vez mais específicas. Moliner propõe uma estrutura e organização que ela chama de “cone léxico” que teria mais ou menos esta forma:



Nota: o meu desenho apenas imita o da autora adaptando-o grosseiramente ao português. Aliás, não existe simetria no gráfico de M. Moliner como aqui está. No meu desenho cada bolinha representa um lexema.

A figura sugere que se parte dos arquilexemas, ou termos super-ordenados, para os lexemas mais e mais específicos e diferentes.

Na prática lexicográfica alguns modelos básicos de redação de verbete têm sido usados tradicionalmente. Como a natureza dos signos lingüísticos expressos pelas palavras é muito diversa, não se pode trabalhar apenas com um formato de verbete. Os modelos de verbete serão discutidos a seguir.

1.º) Quando o referente designado é um elemento concreto do mundo físico, ou do universo cultural, para que o consulente o conheça ou identifique, o melhor procedimento é a sua descrição, acompanhada de uma imagem do próprio referente. Convém que essa imagem seja colorida para que se possa formular uma idéia aproximada do referente no mundo real. Em muitos casos o *designatum* — cf. exemplos abaixo: *paineira*, *jaguarica* — designa um ser que possui entre suas características distintivas básicas a cor. Pode-se acrescentar os usos, utilidades e destinação do signo referido, o que sempre acrescenta dados complementares sobre o objeto, pois nesses casos temos elementos da cultura ou do mundo físico. Assim os nomes de plantas, flores, frutas, animais, instrumentos, utensílios, etc. têm na definição ostensiva o seu tipo ideal.

paineira = árvore de tamanho grande que dá flores coloridas grandes de cor rosa escuro. Produz a paina, muito usada na confecção doméstica de travesseiros e colchões.

jaguarica = animal semelhante ao gato, de pelo malhado em tons diversos de marrom; é carnívoro e vive nas matas do Brasil.

campina = campo extenso, sem árvores e com vegetação rasteira.

iceberg = grande bloco de gelo flutuante, que se destacou de uma geleira polar.

Vejamos exemplos do universo cultural:

Iemanjá = personagem do folclore religioso brasileiro considerada como uma divindade do mar.

paisagista = pessoa que desenha ou planeja jardins e paisagens que depois serão executados artificialmente.

Na definição de substantivos o dicionarista se serve de conceitos básicos, expressos por palavras de sentido geral como: *árvore*, *animal*, *campo*, *pessoa*, *personagem*, *objeto*, *aparelho*, *instrumento* quando se trata da nomeação de seres concretos e de termos também gerais, mas de significado abstrato como *ação*, *ato*, *estado*, *qualidade*, *fato*, *resultado* quando estiver definindo nomes abstratos.

2.º) No caso de um conceito abstrato, o melhor método é o da definição lógica. Já existem alguns modelos padronizados para esses conceitos na lexicografia ocidental. Por exemplo:

abafamento = estado em que se tem dificuldade de respirar porque o ar parece pesado e se tem a sensação de falta de ar e calor.

abastecimento = ação ou fato de prover/com aquilo que é necessário.

erradicação = ação ou resultado de eliminar pela raiz, isto é, totalmente.

lealdade = qualidade daquele que é sincero e honesto em que se pode confiar.

Os verbos se definem, via de regra, através de outro verbo de significação mais geral como em:

abater = 1. derrubar, pôr abaixo no chão.

2. Derrubar com um golpe mortal, matando.

padronizar = estabelecer um modelo que deve ser imitado e/ou reproduzido.

Os adjetivos são definidos através de uma oração adjetiva.

cotidiano = que é de todos os dias; que é diário.

excitado = que tem atividade mental muito viva e acima do comum.

perecível = que pode estragar-se facilmente (falando de alimentos).

3.º) Outro tipo de definição é aquela em que se inclui o signo seu *genus* e, a seguir, se explicita a *differentia*. Serve freqüentemente para adjetivos que exprimem conceitos primários (primitivos léxicos). Ex.:

azul = uma das cores primárias; a cor do céu, do mar, da água das piscinas.

vermelho = uma das cores do espectro solar; a cor do sangue.

Os primitivos léxicos são de definição lógica muito difícil. Na verdade o lexicógrafo supõe que o falante já possui, no seu repertório léxico, o conhecimento e o domínio do uso desses primitivos. De fato, eles se incluem entre as 2.000 palavras que um indivíduo aprende na sua primeira infância; portanto, quando adolescente ou quando adulto vier a ser um eventual consultante de um dicionário, já deverá ter perfeito domínio dos primitivos léxicos. O lexicógrafo considera, pois, esses primitivos léxicos como axiomáticos.

4.º) As palavras instrumentais como preposições, conjunções, artigos, vários pronomes e advérbios exigem um modelo de definição bem diferente. Preposições e conjunções e outras palavras gramaticais são instrumentos de articulação do discurso; portanto, não só não se referem ao universo físico ou cultural do falante, como também são signos meramente lingüísticos. Assim sendo, a única forma de defini-los é situá-los em um contexto lingüístico, mostrando quais são os usos (às vezes muitos) que a língua faz deles. Sendo algumas dessas palavras aquelas que maior número de ocorrências apresentam em qualquer texto, conseqüentemente possuem uma enorme versatilidade lingüística. Algumas são verdadeiros coringas: o sistema lingüístico lhes atribui usos variadíssimos. É o caso específico de algumas preposições como *de*, *em*, *a* na língua portuguesa. O dicionarista se confunde então com o gramático, procurando alistar os usos mais freqüentes e mais típicos desse instrumento gramatical para formular o seu verbete. Por outro lado, dependendo da finalidade do dicionário e do tipo de usuário a que se dirige o lexicógrafo, esses verbetes sobre palavras gramaticais deverão ser mais ou menos exaustivos.

É freqüente encontrar-se em dicionários do português e de outras línguas, sinônimos para explicar o significado da palavra entrada, ao invés de uma definição. Muitas vezes o consultante vai conferir o sentido dos sinônimos referidos e os verbetes consultados remetem-no de volta à palavra de que partiu, num autêntico círculo vicioso. Além de secular, a técnica de explicar a palavra através de sinônimo é inevitável por vezes, especialmente no caso de adjetivos e verbos. Contudo, sempre que possível, a definição através de uma paráfrase é melhor.

Também se costuma definir o *definiendum* pelo seu contrário.

desarrumar = desfazer o que estava em ordem, arrumado.

desencanto = perda do encanto.

desfavorável = que não é favorável.

desigual = que não é igual em tamanho, peso, quantidade.

V. POLISSEMIA E HOMONÍMIA

5.1. Os signos lingüísticos polissêmicos acarretam várias dificuldades para o lexicógrafo. No passado muitos dicionários operavam da seguinte forma: indicavam primeiro o significado original, etimológico, alistando, a seguir, os significados subsequentes até os valores semânticos contemporâneos ao dicionarista. Um dos raros exemplos de dicionário com esse formato e executado com maestria é o *Oxford English Dictionary* (o grande *Oxford*). Morais na sua segunda edição (a de 1813) tentou operar desse modo. Na prática é extremamente difícil estabelecer a cronologia das significações, pois seria necessário um aparato lexicográfico do porte do *Trésor de la langue française*, isto é, monumental. Evidentemente só uma profunda e detalhada pesquisa na documentação histórica da língua e um gigantesco banco de dados de todos os períodos do idioma autorizaria os lexicógrafos a estabelecer a escala correta dos significados. Por isso os dicionaristas contemporâneos optaram por um modelo mais exequível e intuitivo, porém mais útil para o consulente do dicionário. Ordenam-se os significados com base na sua maior freqüência de uso, ou seja, dos sentidos mais comuns aos menos freqüentes. A gradação dos significados é estabelecida considerando-se ainda: 1.º) o sentido mais geral e não marcado deve figurar em primeiro lugar; 2.º) arrolam-se depois os sentidos mais marcados, mais especializados; 3.º) seguem-se os significados específicos ou técnicos; 4.º) indicam-se depois os valores semânticos nas variedades lingüísticas mais restritas (regionalismos); 5.º) por fim, registram-se os sentidos obsoletos no caso dos dicionários de grande porte. A ordem hierárquica, assim estabelecida, depende do bom senso do lexicógrafo e o dicionarista precisa analisar atentamente o leque de significados da palavra polissêmica, procurando captar os traços semânticos que distinguem os vários sentidos de um vocábulo. Os valores semânticos concretos ou primários antecedem os significados metafóricos. Se a palavra tiver dois ou mais sentidos dominantes será preciso numerá-los de forma distinta e clara: usam-se inicialmente algarismos romanos e depois algarismos arábicos, letras ou outros símbolos na hierarquização dos significados.

Veja-se em exemplos do meu *Dicionário Fundamental do Português* a ordenação dos sentidos:

Cabeça s.f.

1. Parte superior do corpo humano e do corpo dos animais.
2. Extremidade de um objeto (met.)
3. Aquilo que apresenta uma forma arredondada como se fosse uma cabeça. (met.)
4. Inflamação que aparece na pele parecendo uma ponta saliente, avermelhada, dura e dolorida (met.)
5. Cada animal que faz parte de um rebanho.

Fenômeno s.m.

1. Modificação operada nos corpos por ação de agentes físicos ou químicos.
2. Tudo o que é percebido pelos sentidos ou pela consciência.
3. Fato de natureza social ou moral.
4. O que é raro ou surpreendente, que causa admiração.
5. Pessoa que se distingue por um talento extraordinário.

Puxar v.

1. Atrair para perto de si arrastando.
2. Arrastar atrás de si.
3. Fazer força para tirar.

4. Provocar um assunto (ling. col.)

5. Sair semelhante a (ling. col.).

Queimar v.

1. Consumir pelo fogo.

2. Atear fogo a; fazer consumir pelo fogo.

3. Tornar escuro, bronzeado pela ação do sol.

4. Fazer ficar seco e morto.

5. Dar a sensação de calor quase insuportável no contacto com a pele.

6. Destruir pelo gelo ou pela geada.

Um exemplo muito simples para evidenciar uma discriminação muito freqüente no halo de significação que envolve as palavras: a oposição entre um sentido que se pode considerar básico e um valor semântico translato, metafórico:

Fecho s.m.

1. Instrumento ou dispositivo que serve para fechar alguma coisa.

2. Conclusão, encerramento de alguma coisa (met.).

Vamos considerar agora exemplos de outras línguas e de outros dicionários tipo padrão. No *Diccionario de uso del español* (1975) de Maria Moliner conferir a ordenação dos significados nas duas palavras-entrada citadas a seguir:

Realidad (fem.)

1. Cualidad de real; hecho de existir.

2. El mundo real. Lo que existe.

3. Por oposición a lo ilusorio o fantástico, lo efectivo y que tiene valor.

4. Verdad. Lo que ocurre verdaderamente, aparte de las apariencias o de lo que podría imaginarse.

5. Verdad, ingenuidad, sinceridad.

Realizar

1. Convertir una cosa, como “sueños, planes” o “ilusiones” en realidad.

• Converter en cosa real.

• Hacer una cosa designada con um nombre de acción.

2. Vender para convertirlos en dinero las fincas o bienes de cualquier clase que se poseen.

Liquidar.

No *Dictionnaire alphabétique et analogique de la langue française* (1972) de Paul Robert:

Action n.f.

I. 1.° Ce qui fait quelqu'un et par quoi il réalise une intention ou une impulsion.

2.° Fait de produire un effet, manière d'agir.

3.° Absolt. Exercice de la faculté d'agir.

II. Exercice d'un droit en justice.

III. Suite de faits et d'actes constituant le sujet d'une oeuvre dramatique ou narrative.

No *Micro Robert* (1979):

Complet, ète adj.

1.° Auquel ne manque aucun des éléments qui doivent le constituer.

2.° Qui a un ensemble achevé de qualités, de caractères.

3.° (Sens faible: avant ou après le nom). Qui possède tous les caractères de son genre.

4.° Tout à fait réalisé; écoulé.

5.° Avec toutes les parties, tous les éléments qui le composent en fait.

6.° Qui n'a plus de place disponible.

Evidentemente a ordem que se tenta estabelecer é passível de erro, dado o caráter fluido da significação. Tanto no português como em qualquer uma das línguas aqui referidas, o léxico é uma massa nebulosa cujos contornos e limites é praticamente impossível de delimitar com clareza. É inexecutável para um lexicógrafo circunscrever inequivocamente os campos semânticos como um botânico distingue e classifica as espécies de plantas no reino vegetal. Em parte porque os léxicos individuais não coincidem e cada lexicógrafo se reporta ao seu próprio vocabulário na redação dos verbetes. Em parte porque o vocabulário flutua continuamente em virtude da ação da sociedade e dos falantes sobre ele. Como existe uma correlação biunívoca entre o léxico e o mundo em que vivem os homens e como os indivíduos estão perpetuamente transformando o seu universo, esse moto perpétuo se reflete no vocabulário que se altera ininterruptamente.

5.2. Um problema muito diverso mas também complexo para o lexicógrafo é o dos homônimos. No passado a prática mais comum foi distinguir os homônimos dando-lhes entrada diferente no dicionário, caso fossem distintos os étimos de cada palavra. Um exemplo tomado ao *Grande Dicionário Português ou Tesouro da Língua Portuguesa* de Vieira (1871-1884):

Pena, s.f. (do lat. *poena*). Castigo, punição.

.....
Pena, s.f. ant. por *Penha*. Nossa Senhora da Pena.

Penna, s.f. (do lat. *penna*). Tubo natural guarnecido de plumagem que reveste o corpo das aves. // Ainda nos tempos contemporâneos as discriminações homonímicas são estabelecidas com base no étimo, sobretudo latino, que é o mais freqüente, especialmente para as línguas românicas. Assim o *Diccionario de la Lengua Española* da Real Academia Española (ed. 1956), usa a mesma oposição do Vieira para os mesmos vocábulos:

Pena (Del lat. *poena*, y éste del gr.) f. Castigo impuesto por autoridad legitima al que ha cometido un delito o falta.

Pena (Del lat. *penna*) f. Cada una de las plumas mayores del ave, que situadas en las extremidades de las alas o en el arranque de la cola, sirven principalmente para dirigir el vuelo.

No *Diccionario de uso del español* de Maria Moliner (1975) conferir os vocábulos:

1. *Real* (Del lat. “regalis”, deriv. de “rex, regis” v. “rey”). De (del) rey, de {de la} reina, de {de los} reyes o de {de la} monarquia.

2. *Real* (Del b. lat. “realis”, deriv. de “res, rei”, cosa, y éste, posiblemente de “reor”, contar — o la inversa — ...) Se aplica, por oposición a imaginario o inexistente, a las cosas que existen o han existido.

Ainda usando o princípio da discriminação etimológica, vejamos outro tipo de exemplo no dicionário de Nicola Zingarelli, *Vocabolario della lingua italiana* (10a. ed., 1971):

Appunto (1) da *appuntare* (2) s.m. 1. Annotazione scritta, rapida e concisa, fatta per ajuto della memoria.

Appunto (2) o *a punto* lat. *ad p̄unctu(m)* ‘al punto’ avv. proprio

Digitale (1) vc. dotta, lat. *digitāle(m)* ‘relativo al dito’ (*digitus*) agg. Proprio di un dito o delle dita.

Digitale (2) del lat. *digitālis*, per la forma a dito (*lat. digitus*) del fiore. s.f. Planta erbacea.

Digitale (3) ingl. *digital* (relativo al calcolo con elementi numerali (*digit s*, propr. ‘cifre, unità numeriche, dal lat. *digitus* ‘dito (che serve per numerare)’) . agg. Numerico, nei sistemi per il trattamento automatico delle informazioni.

A distinção entre homônimos não precisa ser necessariamente feita através da etimologia. É certo, porém, que a oposição homonímica é resultante, via de regra, da evolução histórica do léxico.

Contudo, sempre que deparamos com vocábulos de forma idêntica mas opostos por seus semas distintivos, devemos atribuir-lhes entradas diferentes no dicionário. Assim, por exemplo, no meu *Dicionário Fundamental do Português* separei *canal*₁ (s.m.) e *canal*₂ (s.m.), *dado* (adj.), *dado*₁ (s.m.) e *dado*₂ (s.m.). Cf.:

*Canal*₁ s.m.

1. Local estreito e comprido para escoar água ou detritos.
2. Passagem construída em um terreno para ligar rios, lagos e rios, mares.
3. Construção semelhante a uma valeta estreita e comprida (v. n.º 1) em que se desvia a água de um rio, ou córrego a fim de fazer irrigação.
4. Qualquer cavidade que dá passagem a gases ou líquidos.

*Canal*₂ s.m.

Qualquer meio físico através do qual podem ser transmitidos sinais de comunicação como, por exemplo, as ondas eletromagnéticas que transmitem sons e imagens para um receptor de televisão.

Dado adj.

1. Que é oferecido como presente.
2. Determinado.
3. Que é amável e simpático com as outras pessoas (ling. fam.).

*Dado*₁ s.m.

Pequeno cubo com que se joga e que apresenta em cada um dos lados pontos que vão do número 1 até o 6.

*Dado*₂ s.m.

1. Elemento ou quantidade conhecida que serve de base para a resolução de um problema.
2. Informação que se põe no computador para ser por ele processado a fim de se obter novas informações (geralmente usado no pl.).

Em casos como os referidos acima no português e no espanhol, é melhor apresentar este formato de entrada separada, ao invés de incluir os homônimos em um único verbete, distinguindo-os depois graficamente através de números. A entrada individual para cada um dos homônimos alertará o consulente para o fato de que se trata de unidades léxicas diferentes da língua.

VI. SINÔNIMOS, ANTÔNIMOS E PARÔNIMOS

No corpo do verbete devem ser indicados os sinônimos quando eles existirem para tornar mais completas as informações sobre o signo lingüístico que constitui a entrada lexical. Será muito útil para o usuário do dicionário que o lexicógrafo forneça os sinônimos da palavra-entrada, distinguindo as nuances dos vários valores semânticos de cada sinônimo. É verdade que não existem duas palavras que se possam considerar totalmente equivalentes, em virtude da riqueza e flexibilidade da língua nos seus diversos matizes de usos afetivos, sociais, técnicos, científicos, gíriáticos, vulgares, etc. Contudo, o conhecimento das relações semânticas entre as palavras possibilitará ao falante ampliar a sua competência vocabular.

O lexicógrafo deve, pois, situar a palavra-entrada dentro da rede de significações de que ela faz parte; assim ele estará explicando simultaneamente o valor dos demais componentes do mesmo campo léxico. A estratégia do contraste tem um enorme efeito

esclarecedor e distintivo para a mente humana. Os vários vocábulos ligados entre si pela significação — sinônimos, antônimos e parônimos — compõem um leque de valores e de empregos diferentes uns dos outros, embora próximos. Esses empregos se originam não apenas no fenômeno da significação em si, mas nas especificações que o uso lingüístico estabelece, correlacionadas com os registros ou níveis de linguagem na maioria das vezes. Assim, por exemplo, os sinônimos de *complexo*: complicado, difícil, confuso. Cf. nos contextos:

- a) A estrutura da matéria é muito *complexa* (registro formal, escrito, científico).
- b) Esta mulher faz as coisas de modo muito *complicado* (registro coloquial).
- c) Este problema da matemática é muito *difícil* para uma criança de dez anos (registro coloquial).
- d) Vinha da sala contígua um ruído *confuso* de vozes, sem que se pudesse distingui-las (registro culto, escrito).

O emprego dos sinônimos grifados nos contextos acima não são, evidentemente, os únicos possíveis. *Complicado* e *difícil* podem aparecer em registro culto e *confuso* em registro coloquial. Trata-se apenas de exemplos para ilustrar a aplicação diferenciada dos sinônimos. De fato, não se usaria nenhum dos sinônimos de *complexo* no enunciado a). Tão pouco se usaria *complexo* na frase d). E ainda que se possa usar *confuso* nos enunciados b) e c), os sentidos serão certamente distintos.

A explicitação da antonímia é também muito útil em um verbete. De fato, há um número expressivo de elementos do léxico que têm estrutura binária, formando pares de contrários. Em decorrência disso, pode-se lançar ainda mais luzes sobre o significado da palavra, situando-a duplamente: 1) no campo semântico de que faz parte (cf. acima); 2) opondo-a ao campo semântico do(s) seu(s) contrário(s). O contraste, feito assim totalmente evidente, atua sobre a mente como se alguém pintasse uma cor qualquer sobre um fundo branco ou preto. Continuando o mesmo exemplo, a rede dos antônimos de *complexo* seria: *simples, fácil, claro*. Colocando-os em contexto, teríamos:

- a) A estrutura da matéria não é *simples*.
- b) Esta mulher faz as coisas de modo muito *simples* (ou *fácil*).
- c) O problema de matemática não era *fácil* para uma criança de dez anos.
- d) Vinha da sala vizinha um som *claro* de vozes, que se podia distinguir bem.

Igualmente os parônimos, ou palavras de forma parecida, ou ainda os cognatos da mesma família de palavras, podem ajudar a compreender melhor o sentido e o valor da palavra-entrada que se consulta. Nesse caso o dicionarista fará referências cruzadas, remetendo o consulente a outros verbetes. Por ex.: em *emigrante* se remete a *imigrante*, em *flagrante* se remete a *fragrante*, em *conserto* se remete a *concerto* e vice-versa. Em *decisão* se remete aos cognatos *decidido, decisivo, indecisão, indeciso*.

A indicação dos registros lingüísticos é muito útil para o consulente do dicionário. Se o falante souber que tal palavra só é usada na linguagem literária, aquela outra apenas no registro coloquial, familiar ou vulgar, ou ainda se o vocábulo é um termo regional ou típico de alguma linguagem especial (direito, botânica, medicina, etc.) essa informação será extremamente valiosa para ele. Se o grau de sofisticação do dicionário for grande ele deverá incluir ainda os termos ou valores desusados e obsoletos. Um dicionário padrão ideal é aquele que dá importância às indicações dos valores semânticos e dos usos lingüísticos característicos de cada nível de linguagem. Quanto mais acurado for o trabalho lexicográfico, tanto mais refinado ele será nesse aspecto.

Enfim, a questão do exemplário de frases e citações que ilustrarão o sentido, a construção ou o uso discriminado pelo dicionarista. Quando o lexicógrafo dispuser de um banco de dados já elaborado em forma de concordância de textos (Cf. item II, A

confeção de um dicionário) o seu trabalho será facilitado, além de ter a chance de ser mais preciso. Nesse caso ele deverá extrair da sua concordância os melhores exemplos, de forma que o contexto citado explicita bem o significado, uso ou construção em epígrafe e também registre o nível de linguagem descrito quando for necessário. Muitas vezes uma passagem muito poética e excepcional literariamente não é adequada como abonação. Pelo contrário, um contexto pouco original mas que ilustre bem a norma lingüística pode ser o mais indicado. Na redação de pequenos dicionários os lexicógrafos geralmente constroem os exemplos de contextos ilustrativos. De fato, como esse tipo de dicionário tem uma finalidade pedagógica, os dicionaristas poderão manipular melhor as informações léxicas se elaborarem eles próprios as frases e contextos que ilustrarão as acepções do verbete. Também nos dicionários maiores como os dicionários padrão, muitas vezes o lexicógrafo precisará construir um exemplo para descrever um significado, uso, construção porque mesmo nos grandes bancos de dados pode não ocorrer um determinado valor lexical.

Convém insistir no fato de que os exemplos devem ilustrar o melhor possível os significados da palavra-entrada e seus usos. Na verdade, a definição, a explicação, a descrição de um conceito só se completam quando postas no contexto do discurso. E assim o dicionário finaliza a sua tarefa de esclarecer, na sua totalidade, um lexema para um falante que o desconhece, ou o conhece imperfeitamente.

O dicionário unilingüe constitui uma das grandes instituições lingüísticas e culturais de uma sociedade, exercendo uma função normativa e prescritiva. Por isso é imprescindível o investimento de grandes recursos humanos, técnicos e econômicos para mantê-lo sempre atualizado e apto a responder às necessidades dos membros da comunidade. Em virtude do afluxo contínuo de termos novos oriundos das aceleradas transformações sociais, culturais, técnicas e científicas no mundo contemporâneo, torna-se indispensável uma freqüente reformulação dos dicionários unilingües. Uma reiterada revisão e atualização sobretudo do dicionário padrão, exige a institucionalização desse processo. Órgãos públicos como o *Institut de la Langue Française* na França ou o *Office de la Langue Française* no Québec (Canadá) parecem preencher idealmente tal função. Podemos lembrar, porém, entidades privadas que institucionalizaram a produção de dicionários unilingües de maneira admirável como a *Oxford University Press* e a empresa *Le Robert*. Todas essas instituições montaram um formidável aparato de equipamentos, recursos documentais e equipes de especialistas, a fim de atualizar continuamente o banco de dados léxicos e lingüísticos e reformular seus dicionários periodicamente.

BIDERMAN, M.T.C. — The pattern dictionary of a language. *Alfa*, São Paulo, 28(supl.):27-43, 1984.

ABSTRACT: Unilingual dictionaries have several sizes and formats. The number of entries in a dictionary depends on the public to which it is addressed. The pattern dictionary of a language is an unilingual dictionary with 50,000 entries approximately, including a substantial lexical repertoire. It is different from a thesaurus which collects all the words of a lexicon. The pattern dictionary is an important cultural instrument in modern society. In this article several problems related to the making of a pattern dictionary and unilingual dictionaries are examined: the selection of lexical entries, compilation of a corpus (data bank), writing of entries, lexicographical definition. Problems like polysemy and homonymy are also analysed.

KEY-WORDS: Unilingual dictionary; pattern dictionary; thesaurus; corpus; data bank; entry-word; entry; lematization; canonical form; indices verborum; concordance; lexicographical definition; synonym; antonym; cross-reference.

REFERÊNCIAS BIBLIOGRÁFICAS

1. BIDERMAN, M.T.C. — *Teoria lingüística*. Rio de Janeiro, Livros Técnicos e Científicos, 1978.
2. DUBOIS, J. — *Introduction à la lexicographie: le dictionnaire*. Paris, Larousse, 1971.
3. MOLINER, M. — *Diccionario de uso del español*. Madrid, Gredos, 1975.
4. ZGUSTA, L. — *Manual of lexicography*. The Hague, Mouton, 1971.

BIBLIOGRAFIA CONSULTADA

- AITKEN, A.J. — Historical dictionaries, word frequency, distributions and the computer. *Cahiers de Lexicologie*, 32(1): 28-47, 1978.
- BALDINGER, K. — Semasiologia e onomasiologia. *Alfa*, 9:7-36, 1966.
- BARNHART, C.L. — Plan for a Central Archive for Lexicography in English. *Annals of the New York Academy of Sciences*, 211: 302-319, 1973.
- BLUTEAU, R., Pe. — *Vocabulário português e latino*. Lisboa, Colégio das Artes da Cia. de Jesus, 1712-1728. 8v., 2 supl.
- CAPPELLI, A. et alii — *Parsing an Italian text with an ATN Parser*. Pisa, Instituto di Linguistica Computazionale, 1978.
- CASARES, J. — *Introducción a la lexicografía moderna*. Madrid, C.S.I.C., 1950.
- COSERIU, E. — Vers une typologie des champs lexicaux. *Cahiers de Lexicologie*, 27(2): 30-51, 1975.
- DELATTE, L. et alii — Le traitement automatique de la langue française au laboratoire d'analyse statistique des langues anciennes. *Revue des Études Anciennes*, 4:1-55, 1977.
- DICCIONARIO DE LA LENGUA ESPAÑOLA / por/ Real Academia Española. 18.ed. Madrid, Espasa-Calpe, 1956.
- DRETTAS, G. — Les théoriciens allemands du champ. *La Linguistique*, 17(2): 3-22, 1981.
- DUNCAN, J. — *Frequency dictionary of Portuguese words*. Stanford, Stanford University, 1972. (Ph. D. Dissertation)
- FERRARI, G. — Dictionnaire automatique et dictionnaire-machine: une hypothèse. In: COMPUTATIONAL AND MATHEMATICAL LINGUISTICS. PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON COMPUTATIONAL LINGUISTICS, Pisa, 1973. Firenze, Leo S. Olschki Ed., 1977, p. 257-262.
- FERRARI, G. & PRODANOF, I. — Machine dictionary and lexicon. In: INTERNATIONAL LINGUISTIC, Ottawa, 1976. (Comunicação)
- FERREIRA, A.B. de H. — *Novo dicionário da língua portuguesa*. Rio de Janeiro, Nova Fronteira, 1975.
- KELLER, H.H. — The American heritage word frequency book (review). *Language Learning*, 25(1): 173-178, 1975.
- LARA, F. — Méthode en lexicographie: valeur et modalité du dictionnaire de machine. *Cahiers de Lexicologie*, 29(2): 103-28, 1976.
- LOGMAN DICTIONARY OF CONTEMPORARY ENGLISH. London, Logman Group, 1978.
- MC NAUGHT, J. — Terminological data banks: a model for a British linguistics data bank (LDB). *ASLIB Proceedings*, 33(7/8): 297-308, 1981.
- MACHADO, J.P. — *Dicionário etimológico da língua portuguesa*. Lisboa, Ed. Confluência, 1956. 2v.
- MARCOS-MARÍN, F. — *Curso de gramática española*. Madrid, Cincel-Kapelsz, 1980.
- MATORÉ, G. — *Histoire des dictionnaires français*. Paris, Larousse, 1968.
- MELO, G.C. de — *Dicionários portugueses*. Rio de Janeiro, S.O.M.E.S., 1947.
- MORAIS SILVA, A. de — *Dicionário da língua portuguesa*. Fac-simile da segunda edição, 1813, photographada pela Revista de Língua Portuguesa. Rio de Janeiro, Oficinas de S.A. Littro Typographia Fluminense, 1922. 2v.
- MORAIS SILVA, A. de — *Grande dicionário da língua portuguesa*. 10. ed. rev. por J.P. Machado. Lisboa, Ed. Confluência, 1949-1957. 12v.
- MURAKAWA, C. de AA. — *O primeiro dicionário de língua portuguesa de Antonio de Moraes Silva. Estudo crítico da edição de 1813*. Araraquara, Instituto de Letras, Ciências Sociais e Educação, Unesp, 1984. (Dissertação de Mestrado)
- NASCENTE, A. — *Dicionário etimológico da língua portuguesa*. 2.ª tiragem da 1.ª ed. Rio de Janeiro, Francisco Alves, s.d.
- PAIKEDAY, T.M. — Language analysis and lexicography by microcomputer. (Comunicação feita no encontro ADS-MLA, 1981).
- PAIKEDAY, T.M. — *The New York Times everyday dictionary*. New York, Times Book, 1982.

- PORTUGUÊS FUNDAMENTAL. Lisboa, Instituto Nacional de Investigação Científica, Centro de Linguística da Universidade de Lisboa, 1984. v.1, Vocabulário e Gramática; T.1, Vocabulário.
- PRODANOF, I — A la recherche d'un modèle de derivation en italien. In: COMPUTATIONAL AND MATHEMATICAL LINGUISTIC. PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON COMPUTATIONAL LINGUISTICS, Pisa, 1973. Firenze, Leo S. Olscki Ed. 1977. p. 297-301.
- RAVEN, I. *et alii* — Lexicography in English. *Annals of the New York Academy of Sciences*, 211, 1973.
- REY-DEBOVE, J. — Le domaine du dictionnaire. *Langages*, (19):3-34, spt. 1970.
- REY-DEBOVE, J. — Lexique et dictionnaire. In: LE LANGAGE: Les dictionnaires du savoir moderne. Paris, Centre d'Étude et de Promotion de la Lecture, 1973. p. 82-108.
- RICHMAN, B. *et alii* — *The American heritage word frequency book*. New York, Boston American Heritage Publ., Houghton Mifflin, 1971.
- ROBERT, O. — *Dictionnaire alphabétique et analogique de la langue française*. Paris, SNL, Dictionnaire Le Robert, 1972.
- SHERMAN, D. — Retrieving lexicography citations from a Computer Archive of Language Materials. *Annals of the New York Academy of Sciences*, 211:137-142, 1973.
- SHERMAN, D. — Special purpose dictionnaires. *Cahiers de Lexicologie*, 32(1): 82-102, 1978.
- TERMINO GRAMME: Bulletin de la Direction de la terminologie. Québec, Office de la Langue Française, 1979.
- THE OXFORD ENGLISH DICTIONARY. Oxford, Clarendon Press, 1933.
- VENEZKY, R.L. — Computer applications in lexicography. *Annals of the New York Academy of Sciences*, 211:287-291, 1973.
- VENEZKY, R.L. — Storage, retrieval and editing of information for a dictionary. *American Documentation*, 19: 71-79, 1968.
- VIEIRA, D. Frei — *Grande dicionario portuguez ou thesouro da lingua portuguesa*. Porto Ed. Ernesto Chardon e Bartholomeu H. de Moraes, 1871.
- VITERBO, J. de S.R. de, Frei — *Elucidario de palavras e frases que em Portugal antigamente se usarão (sic) e que hoje regularmente se ignorarão*. Lisboa, 1778-1799.
- WEINREICH, U. — Lexicographic definition in descriptive semantics. In: HOUSEHOLDER, F.W. & SAPORTA, S., eds. — *Problems in lexicography*. Bloomington, The Hague, Mouton, 1967. p. 25-44.
- WEINREICH, U. — Webster's third: a critique of its semantics (review). *International Journal of American Linguistics*, 30:405-409, 1964.
- WICLOW, C.K. — Advanced English vocabulary. *Language Learning*, 24(1): 167-170.
- ZINGARELLI, N. — *Vocabolario della lingua italiana*. 10. ed. Bologna, Zanichelli, 1971.

LÉXICO E DICIONÁRIO

Por Josette REY-DEBOVE*
Tradução de Clóvis Barleta de MORAIS**

RESUMO: *A autora trata das relações entre léxico e gramática, definindo conceitos básicos da lexicologia: a palavra, a unidade léxica, a lexia, o morfema, classe aberta e classe fechada, o repertório vocabular de uma língua. O léxico é o domínio menos específico da linguagem, reportando-se ao universo referencial. O conjunto de signos que compõem o léxico, conferem a ele um estatuto semiótico, criando um sistema de mundo. A autora discute o problema da designação, mostrando como se relacionam palavras e coisas. Sendo um conjunto aberto, o léxico tem possibilidades infinitas de expansão. Os dicionários constituem nossa única idéia sobre o léxico.*

UNITERMOS: *Léxico; lexicologia; dicionário; unidade léxica; lexia; referência; sistema de mundo; metalingüística; terminologia; designação; taxinomia; polissemia; homonímia; nomenclatura; verbe; definição; dicionário monolingüe; dicionário bilingüe; dicionário histórico.*

SITUAÇÃO DO LÉXICO NA LÍNGUA ***

Gramática e léxico.

Há duas maneiras de aprender uma língua. Uma, natural, por tentativas cada vez mais aperfeiçoadas de comunicação que chegam a conhecimentos memorizados dessa língua (competência natural), como o da criança na família, e, nesse caso, pode-se dominar perfeitamente uma língua sem se ser capaz de descrevê-la. A outra, artificial e metalingüística, pela consulta de dois tipos de obras descritivas conhecidas como indispensáveis e complementares: a gramática e o dicionário.

Essas duas maneiras de aprender uma língua não são quase empregadas exclusivamente: a aprendizagem natural é geralmente acelerada e aperfeiçoada pelo uso da gramática e do dicionário, e a aprendizagem artificial é na maioria das vezes sustentada por verificações duma experiência prática de comunicação. Mas, na realidade, esses dois caminhos de aprendizagem deveriam levar aos mesmos resultados: é por isso que os lingüistas dizem que o objeto dum dicionário e duma gramática é descrever a competência natural do utilizador ideal duma língua.

Aprende-se uma língua estrangeira com um dicionário e uma gramática.

Essa correspondência entre a existência de duas descrições e o funcionamento du-

* Professora da Universidade de Paris. Redatora dos dicionários da série Le Robert.
** Departamento de Lingüística — Instituto de Letras, Ciências Sociais e Educação — UNESP — 14800 — Araraquara — SP.
*** Sempre que foi possível, adaptamos a exemplificação francesa, com algumas exceções; às vezes pusemos a tradução ao lado.
Preferimos *descodificar* por ser uma forma mais vernácula que *decodificar*.

ma língua foi sublinhada pelo emprego que os norte-americanos fazem de *dictionary* e *grammar* para designar os dois componentes duma língua, para o quê nós dizemos *léxico* e *gramática*. O léxico duma língua seria o conjunto das unidades submetidas às regras da gramática dessa língua, sendo a conjunção da gramática e do léxico necessária e suficiente à produção (codificação) ou à compreensão (descodificação) das frases duma língua.

Se a existência das gramáticas e dos dicionários nos sugere, com uma exatidão razoavelmente grande, essa dicotomia da língua, o conteúdo deles nos revela, da mesma maneira, a fragilidade da fronteira traçada entre esses dois domínios complementares. (8, 19)

Gramáticas e dicionários falam frequentemente dos mesmos objetos...

Não é o caso de *um, o, sobre, tudo, exato (-a, -os, -as)* na gramática, quando essas são palavras do dicionário? E não se vêem no dicionário indicações sobre o gênero das palavras, sua concordância e lugar na frase, verbetes como *preposição, advérbio* etc., quando tudo isso constitui objeto das gramáticas? Em qual dos dois livros procurar o sufixo *-agem*, a forma verbal *coubesse*? A imprecisão da situação reflete-se nas variações de conteúdo que se manifestam duma obra a outra, entre as gramáticas e entre os dicionários: uma gramática como *le Bon Usage*, de Grevisse, invade a descrição lexical e uma obra como o *Dictionnaire du français contemporain* (Larousse) estende-se largamente sobre a descrição gramatical. Poderia parecer que se trata dum mesmo objeto encarado de dois pontos de vista diferentes, do conjunto para o elemento na gramática, e do elemento para o conjunto no dicionário. Assim, numa gramática, o capítulo do possessivo enumera as formas *meu, teu, seu, nosso, vosso, seu*, ao passo que em cada forma distribuída no conjunto alfabético do dicionário — *meu, nosso, seu* (sing.), *seu* (pl.), *teu, vosso* — explicita-se que se trata dum possessivo. Mas percebe-se rapidamente que as palavras repertoriadas numa gramática são uma ínfima parte do léxico e que nem todas as regras da gramática são explicitadas no dicionário.

...e os dicionários não recobrem inteiramente seu próprio domínio.

Por outro lado, se se trata de opor regras às unidades a que se aplicam, o dicionário apresenta geralmente graves lacunas; podemos ficar admirados de não encontrarmos nele a lista completa das unidades significativas que são os morfemas presos*, uma parte dos quais aparece nas gramáticas. Contudo os afixos e desinências como *-or, -mento, pre-, -s, -m*, parecem ter um valor que se encontra em *agente, operário* etc. (*-or*, em *lavrador*), *ação(-mento* em *cerceamento*), *antes (pre-*, em *predizer*) e *muitos (-s e -m* em *as crianças cantam*). Ademais, ainda que seja impossível falar uma língua sem empregar nomes próprios, estes não são repertoriados no dicionário de língua, mas somente, quando são afamados, na enciclopédia.

Enfim, seríamos tentados a dizer que é a significação que se opõe ao léxico e à gramática. Mas a significação manifesta-se em toda parte: há tanto uma semântica lexical como uma semântica gramatical. Na frase *O texto cita o ministro*, a significação depende das unidades *o, texto, citar, ministro*; mas depende também da gramática visto que *O ministro cita o texto* já não tem absolutamente o mesmo sentido.

* *Morfema preso*: morfema que não pode ser utilizado sozinho na comunicação porque só existe como parte duma palavra.

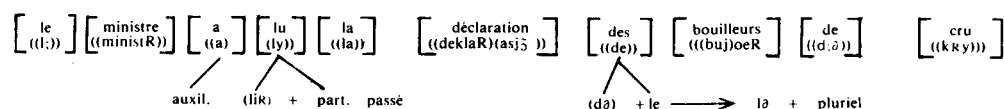
UNIDADES CODIFICADAS SIGNIFICATIVAS.

Trata-se, pois, de considerar que problemas teóricos se põem e como — segundo a maneira de os resolver — se pode traçar uma fronteira aceitável entre gramática e léxico, isto é, definir o léxico.

A gramática utiliza-se de vários tipos de unidades...

Há dois pontos sobre os quais todos estão de acordo: a gramática (sintaxe, morfologia) dá regras para combinar unidades significativas e essas unidades são inferiores à frase que elas constroem. Cada unidade linguística entra na composição duma unidade superior, onde ela funciona como integrante, (1) e essa integração permite produzir um número incalculável de signos com um número restrito de unidades: os morfemas (unidades significativas mínimas) constroem palavras (10) que entram nos constituintes de frase (grupo nominal/grupo verbal).

Pode-se representar do seguinte modo os níveis de integração dos morfemas e das palavras:



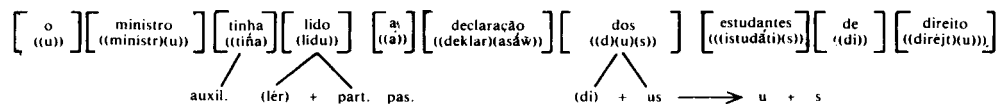
Só as unidades de mesmo nível podem ser estudadas ao mesmo tempo; um morfema que contém um só fonema fica sendo um morfema (*a*) e uma palavra de um só morfema fica sendo uma palavra. É por essa razão que é sempre preferível caracterizar uma unidade no nível superior, por exemplo *ministro* como palavra, e não como morfema (exatamente como há mais exatidão em chamar *quadrado* que *retângulo* a um retângulo de lados iguais).

O morfema.

A separação em morfemas é geralmente possível (*declar-ação*, mas nem sempre: o francês *des*, correspondente aos morfemas *de* + *les*, não pode ser cortado em morfemas, ou elementos concretos separáveis; assim também *men*, que em inglês equivale a *man* + *s*. Os morfemas são livres ou presos: *mar* é dito livre* e *-ação* é preso:** não pode constituir um enunciado significativo por si só. Os morfemas presos constituem o conjunto dos afixos lexicais (prefixos, infixos, sufixos) e das flexões gramaticais (plural, feminino, desinências verbais, casos etc.).

* Na realidade, há o morfema preso de singular, sob a forma zero.

** No original está *-ation*. Em português o sufixo é *-ção*. (Nota do tradutor)



A palavra.

Uma palavra é formada sem dúvida de um só morfema (*mar*) ou de vários (*declaração*).

Define-se como uma forma livre significativa que não pode ser decomposta em outras formas livres significativas menores* e cuja unidade se manifesta por uma coesão interna (também pelo acento de palavra em numerosas línguas). A primeira condição basta para forma como *declaração*, visto que *-ação* não é livre.

A lexia.

A segunda condição é necessária no caso de *bouilleur de cru*, citada acima. As palavras complexas que devem satisfazer a ela são chamadas “lexias” (anteriormente “expressões”, “locuções”).

...e essas unidades estão codificadas na língua.

Todos os integrantes até a lexia são chamados unidades codificadas: o morfema preso é a unidade significativa mínima, e a lexia a unidade significativa máxima. Queremos dizer com isso que elas estão inscritas como unidades no código de nossa memória, a qual deve reproduzi-las tais quais, em bloco, no discurso; não temos a liberdade de mudar, pela mesma razão, a ordem dos elementos nessas unidades, nem mudar um elemento, nem separá-los por inserções (**-temno*, em vez de *-mento*, **declaramento* por *declaração*, **guarda de intenso trânsito*). Ao contrário, o sintagma e a frase não estão codificados. Seus elementos podem ser escolhidos, deslocados e invertidos à vontade, nos limites da gramaticalidade e da semanticidade. Far-se-á, porém, uma exceção importante para os provérbios, que são frases codificadas: *Devagar se vai ao longe*, ou *Quem não tem cão caça com gato* estão inscritas em nossa memória, não as produzimos a partir das palavras que as compõem, do mesmo modo que não produzimos *insuportável* a partir de *in-*, *suport(ar)*, e *-ável*. As unidades codificadas formam conjuntos finitos contáveis (ou mesmo facilmente contáveis), e as unidades não codificadas, sintagma e frase, conjuntos infinitos.

Enfim, é preciso diferenciar a ocorrência particular duma unidade significativa na fala e seu modelo abstrato no sistema da língua (língua/fala) (11). O léxico, quaisquer que sejam as unidades escolhidas, deve fornecer uma lista finita de unidades de língua. Assim, na frase *O ministro cita o texto*, há duas ocorrências na fala duma mesma unidade da língua, *o*. Se certas ocorrências têm um mesmo sentido e uma mesma função, com formas um pouco diferentes, são “variantes”. Tais são, em português, *gatinhar* ou *engatinhar*, *cãibra* ou *câimbra*, que são variantes relativamente livres. Se as varian-

* *Minimum free form* de Bloomfield. — BLOOMFIELD, L. — A set of postulates for the science of language. Psycholinguistics, 1961.

tes são automáticas e estão em distribuição complementar, são “Alomorfes”; assim *São* e *Santo* que não podem empregar-se um pelo outro, sendo obrigatórios *São Paulo* e *Santo Estevão**. Enfim, embora exista uma diferença de sentido entre uma palavra gramaticalmente assinalada como masculino ou feminino (*malvado-malvada, uma-uma*) ou como singular e plural (*jornal-jornais, o-os*), temos o hábito de considerar uma só unidade de língua no masculino (se é o caso não marcado numa língua) e no singular quando o gênero e o número só se realizam na fala. Se o gênero e o número não dependem da fala (*vacance-vacances*), cada forma é uma unidade de língua.

Os problemas levantados pela divisão das unidades significativas variam de uma língua para outra.

TRÊS TIPOS DE LÉXICO.

O conjunto dos morfemas numa língua.

Diversas definições de léxico são válidas, todas diversamente insuficientes. A solução aparentemente mais simples, a dos lingüistas americanos, consiste em escolher como unidade de léxico o morfema, ou unidade significativa mínima**. Mas essa solução apresenta várias dificuldades.

Classes abertas, classes fechadas. Inicialmente, percebe-se que é preciso estabelecer uma distinção entre os morfemas, entre morfemas lexicais ou lexemas, e morfemas gramaticais ou gramemas, o que já é contraditório. Em *a declaração* [adeklarasašw] pode-se distinguir: um artigo definido (*a*), uma raiz verbal (*declar-*), e um sufixo nominal (*-ação*); se se faz variar *ad libitum* cada um dos elementos em sua classe, por exemplo *as lavagens* [azlavãžêjs], constata-se que a escolha é muito grande para substituir *declarar* (quase todos os verbos), que ela é muito menor para substituir *-ação*, e que só se dispõe de *uma* para substituir o artigo definido *a*. Existem, pois, classes mais numerosas ou menos numerosas de morfemas, e certas classes restritas podem ser facilmente descritas pela enumeração de seus elementos.

Constata-se também que quanto mais uma classe é pobre em elementos, tanto mais é fechada, e que quanto mais é rica, tanto mais é aberta. Assim, a classe dos morfemas-nomes, a mais vasta, pode enriquecer-se todos os dias sem prejuízo para a intercompreensão dos usuários mas não a classe dos artigos ou dos pronomes pessoais (imaginemos que perturbação criaria o aparecimento dum pronome X = *eu + tu*, diferente de Y = *eu + ele*).

Enfim, a extrema pobreza e o fechamento numa classe acarretam importantes conseqüências na maneira de identificar os elementos. Basta caracterizar gramaticalmente um elemento desse tipo de classe para que ele seja inteiramente determinado: um artigo definido masculino singular não pode ser senão *o* em português, e não é necessário evocar nem seu sentido nem sua forma fonológica para reencontrá-lo. Isso é impossível, ao contrário, para um elemento dum conjunto rico e aberto; depois de darmos todos os caracteres gramaticais dum verbo: “verbo transitivo direto, regular”, estamos ainda muito longe de ter selecionado *sel(ar)*, por exemplo, e é obrigatório, para uma identificação completa, indicar que seus fonemas são [sel(ár)] e sua significação

* Assim também a preposição *por* é usada sem artigo (*por aqui, por isso*) e *per* com artigo (*pela rua, pelos bosques*) (Nota do tradutor)

** Só os gerativistas definem o léxico em extensão. Ver CHOMSKY, N. — A transformation approach to syntax. In: FODOR & KATZ. — *The structure of language*. New Jersey, Prentice Hall, 1965.

“pôr selo em” (a primeira indicação sozinha poderia fornecer homônimos, *selar* um cavalo, a segunda indicação sozinha, sinônimos, *estampilhar*). A possibilidade de ser um morfema inteiramente definido por caracteres gramaticais, solidários ao fato de que ele pertence a uma classe fechada de poucos elementos, é fundamental e obriga a encarar dois tipos de morfemas em todas as línguas. Adquirimos, pois, o hábito de distinguir os morfemas gramaticais dos morfemas lexicais, sendo os morfemas gramaticais do domínio da gramática. Contudo, a oposição de que se trata aqui, representa os casos mais afastados dum contínuo: as classes são mais, ou menos, ricas e abertas, e não estamos seguros de poder separar os morfemas gramaticais dos outros.

Unidade do código e produção do código. A solução que consiste em definir o léxico como o conjunto dos morfemas, suscita outra dificuldade, a saber, que uma frase não é jamais construída a partir de morfemas, mas das unidades codificadas mais altas, cada vez que se apresenta a escolha, economia bem compreensível. Assim, para produzir a frase “la distribution des journaux se fait rapidement” (“a distribuição dos jornais se faz rapidamente”), dispomos das unidades complexas inteiramente prontas *distribution e rapidement* como escolheríamos *envoi (envio) e vite (depressa)*. Em compensação dispomos de *journal*, em que colocamos o morfema de plural para produzir *journal* (donde o erro freqüente [deʒuRnal])* . Nem todos os morfemas, pois, estão no mesmo plano na produção da frase.

Se se construíssem frases com morfemas, produzir-se-iam palavras. Construir-se-iam, pois, palavras como frases, isto é, unidades não codificadas. Desde então, nada impediria de se dizer *distribu-age*, que tem o mesmo valor que *distribution*, e de inventar mil outras unidades conformes ao sistema, mas não codificadas pelo uso. A gramática gerativa, notadamente, produz, segundo certas regras, palavras até aqui desconhecidas (3). Ora, se é legítimo, para aprender uma língua, estar apto para produzir todas as frases possíveis, vê-se a inadequação de regras que permitem produzir todas as palavras possíveis segundo o sistema, das quais uma grande parte jamais foi empregada. Precisamos concordar plenamente com Bloomfield e Chomsky em que “o léxico é o conjunto completo das irregularidades da linguagem”, mesmo que as irregularidades se manifestem aqui ou ali.

O conjunto das palavras duma língua.

Podemos também retomar a definição tradicional: o léxico é o conjunto das palavras duma língua, o que inclui evidentemente a maior parte dos morfemas (os morfemas livres) e todas as unidades codificadas de vários morfemas (palavras derivadas e compostas, *lexias*). É a imagem do léxico que os dicionários nos dão. O léxico como conjunto de palavras convém bem ao que sabemos da frase, que se constrói com as mais altas das unidades codificadas disponíveis. Divide-se ele em classes funcionais tradicionalmente chamadas categorias gramaticais ou partes do discurso (classes de palavra), e que são definidas por dois caracteres: o contexto, que determina a função na frase, e o paradigma dos morfemas ligados que pode acrescentar-se à unidade (2). Assim, o verbo é definido por sua combinatória relativamente ao nome, ao pronome, ao advérbio etc., e por um tipo de desinências gramaticais que lhe é particular. Existe, pois, uma classificação gramatical do léxico, e dessa classificação estão excluídos os morfemas presos que entram na composição da palavra mas não na da frase. Essa clas-

* Em português só no registro popular é comum a “discordância” *dos jornal* ou *dos jornal*. (Nota do tradutor)

sificação continua sendo aproximadamente a que todos conhecem: nomes, pronomes, adjetivos, verbos, advérbios, artigos, preposições, conjunções, interjeições, com certos melhoramentos segundo os autores (determinantes, dêiticos, substitutos). Ela é válida para numerosíssimas línguas, mas, por exemplo, o chinês quase só distingue substantivos e verbos que servem, na fala, para exprimir outras categorias.

Mas se os morfemas presos, quer sejam lexicais quer gramaticais, são excluídos desse léxico, a gramática é que deve encarregar-se deles e, no caso de não conservar todas as unidades mínimas, o léxico já não se opõe à gramática como um conjunto de unidades a um conjunto de regras.

Essa supressão não basta, aliás, para uniformizar o conjunto descrito, pois que no conjunto das palavras se manifesta a mesma dicotomia que no conjunto dos morfemas: existem palavras gramaticais e palavras lexicais (*quel, duquel e gentil, gentillet*). Reserva-se, em geral, o nome de palavras lexicais às categorias do substantivo, do verbo, do adjetivo e do advérbio (para certos autores, só os advérbios compostos em *-mente* em português, em *-ly* em inglês etc.), reunindo as outras categorias (pronomes, artigos, preposições etc.) as palavras gramaticais. As palavras gramaticais são, pois, registradas no dicionário na qualidade de palavras e na gramática na qualidade de gramaticais (palavras gramaticais, *mots outils* etc.). Enfim a diferença entre morfema preso e morfema livre ou palavra, parece de interesse menor que a distinção entre morfema lexical e palavra lexical, e a distinção entre morfema gramatical e palavra gramatical; se o é livre em português, essa situação não é geral para o artigo (*-ul* romeno é preso como sufixo).

O conjunto das unidades ou o conjunto das palavras de classe aberta numa língua.

É por isso que alguns lingüistas preferem dar prioridade absoluta a essa distinção; segundo eles, a melhor definição do léxico e a mais operatória é então ou o conjunto dos morfemas lexicais (Martinet), ou o conjunto das palavras lexicais somente (Ullmann²⁰, Mathiot). Enquanto unidades significativas codificadas, as palavras formam o conjunto mais vasto (em francês e em numerosas línguas, há mais palavras completas que palavras de um morfema, e mais palavras de um morfema que morfemas presos); e entre as palavras, as palavras lexicais formam de muito longe o subconjunto mais importante. Em francês, as palavras gramaticais quase não ultrapassam uma centena, ainda que componham, só elas, 50% da fala. A frequência está, também, ligada à abertura e ao fechamento dos conjuntos; as palavras mais frequentes estão entre as palavras gramaticais, e as mais raras entre as palavras lexicais, e especialmente os substantivos. A solução do léxico considerada como conjunto aberto seria satisfatória se a distinção entre conjuntos abertos e fechados fosse precisa, e não da ordem do mais ou do menos.

As diversas definições do léxico são exploráveis segundo as necessidades.

Finalmente, escolher-se-á a definição do léxico que pareça mais bem adaptada ao trabalho a efetuar, guardando presentes no espírito as insuficiências de cada solução. Para uma mesma língua, o gramático preferirá sempre falar de morfemas e o lexicólogo (o antropólogo, o sociólogo) de palavras. Além disso, existem tipos de língua que impõem uma definição do léxico de preferência a uma outra, segundo a disposição de suas unidades significativas. O ponto essencial para o lingüista é admitir que a oposição *léxico/gramática* concebida como a oposição *unidades significativas/regras que as combinam*, se ela é suscetível a rigor de dar uma descrição da gramática, não pode bastar para dar uma idéia do léxico.

O ESTATUTO SEMIÓTICO DO LÉXICO

Se o gramático tem o dever de limitar o domínio do léxico em relação ao da gramática, o lexicólogo tem o dever, mais delicado ainda, de reconhecer a zona das trocas entre a língua e o mundo, que se faz essencialmente pelo léxico. Esse segundo objetivo vai contribuir para esclarecer a especificidade do léxico e para avaliar, de outro ponto de vista, as definições precedentemente propostas.

Quaisquer que sejam os elementos escolhidos, o léxico é um conjunto de unidades codificadas significativas, ou signos. Considerado em seu estatuto ótimo, o signo lingüístico é formado por um significante sensível e por um significado abstrato que estão indissociavelmente unidos por um elo "arbitrário" (isto é, convencional, não natural). O signo tem uma função denotativa: serve para evocar os objetos do mundo (aquilo de que se fala), para torná-los presentes à consciência. Não é necessário que esses objetos tenham uma existência real (verificável), basta uma existência sócio-cultural (o centauro, o unicórnio, Mme. Bovary, os Martiens etc.). Essa função denotativa supõe, para o signo, uma mobilidade e uma independência que permitem empregá-lo sozinho com um sentido. Tem o signo lingüístico, além disso, o poder de remeter não somente a um objeto (*Virgílio*), mas a uma classe de objetos (*escritor*). Possui ainda, uma vez que está codificado na língua, um significado (pelo menos um núcleo de sentido) que é ao mesmo tempo determinado e estável, e aquilo a que o signo remete não varia de um contexto para outro; essa condição é indispensável à comunicação entre os usuários duma língua. Finalmente, o signo lingüístico, por sua função na cadeia ordenada da fala, acha-se integrado numa categoria gramatical com implicações semânticas: o substantivo para os objetos, pessoas, ações, fenômenos, qualidades..., o verbo para os processos, o adjetivo e o advérbio para as qualificações dum e doutro.

O quadro apresentado abaixo mostra como se comportam respectivamente as unidades significativas codificadas, palavra e morfema preso, em relação ao modelo do signo lingüístico ótimo precedentemente descrito. Faremos figurar também as unidades não codificadas, sintagma e frase.

	morfema preso		palavra		sintagma	frase
	gram.	lex.	gram.	lex.		
codificação	+	+	+	+	—	—
arbitrariedade do signo	+	+ e -	+	+ e -	+	+
independência material em relação às outras unidades	—	—	+ (?)	+	+	+
independência semântica do contexto (enunciação ou enunciado)	+ e —	+	+ e —	+	+ e —	+ e —
designação	—	—	+ e —	+ e —	+ e —	—
generalidade das coisas designadas			+ e —	+ e —	+ e —	

+ ou— indica que o caráter varia segundo os tipos de unidades. Por exemplo, a arbitrariedade do signo entre as palavras lexicais: + para *gato* e — para *miar*.

A unidade que corresponde melhor à noção de signo lingüístico é a palavra lexical.

A palavra é que representa melhor o signo lingüístico ótimo. Ela é codificada como as unidades inferiores (morfema preso), mas materialmente independente como as unidades superiores (sintagma, frase). Contudo, a palavra lexical aproxima-se melhor ainda do modelo na medida em que, à diferença da palavra gramatical, é semanticamente independente da situação e da enunciação. O sentido das palavras lexicais está quase saturado pelo código, e as circunstâncias de seu emprego quase não a modificam senão para enriquecê-la provisoriamente com uma referência particular (*cão* → *meu cão*). Ao contrário, o sentido de numerosas palavras gramaticais depende da referência determinada por seu emprego. Assim, os dêiticos *este, aqui, agora, ontem, amanhã...* e os pronomes pessoais *eu, tu, ele* ..., que se referem constantemente a objetos diferentes, não têm, portanto, sentido codificado propriamente dito; o aviso “amanhã se farão barbas gratuitamente” não significa nada se não for datado.

A palavra lexical constrói um sistema do mundo.

Enfim, a palavra lexical, graças aos caracteres anteriormente descritos, constitui o instrumento pelo qual as civilizações constroem para si uma visão do mundo;(15) como diz Hegel, “a palavra, só o conceito da qual recebe seu estatuto de indivíduo no universo mental, essa palavra acrescenta sua realidade própria ao conceito; ao mesmo tempo, o conceito encontra na palavra uma fixação e limites”. Certamente tudo é dizível, se se admite com a maioria dos lingüistas a hipótese segundo a qual não existe pensamento independente das palavras que o exprimem e o estruturam; o “indizível” depende apenas da dificuldade de dizer (de linguagem, psicológica etc.). Mas o dizível, notadamente aquilo que designamos pela primeira vez, nem sempre se pode exprimir por uma palavra única: é necessário um grande número de palavras diversamente combinadas. O sintagma desempenha essa função; dizia-se, por volta de 1870, *chemin de fer urbain (estrada de ferro urbana)*, *chemin de fer souterrain (estrada de ferro subterrânea)*, *tramway souterrain (trâmuei subterrâneo)*, para designar o que, dez anos mais tarde, se chamava *métropolitain (metropolitano)*, transformado a seguir em *métro (metro em Portugal, metrô no Brasil)*; falava-se em *arriver dans la Lune (chegar à Lua)*, em *débarquer (desembarcar)*, em *atterrir sur la Lune (aterrizar na Lua)*, muito antes do emprego da palavra *alunir (alunizar)*, aliás mal aceita pelos turistas.

O sintagma-perífrase assume a transição incerta entre os dois sistemas.

O sintagma funciona como uma perífrase que tem valor de palavra, mas que não está codificada (à diferença da lexia). É por isso que há uma infinidade deles, levando-se em consideração o fato de que eles podem ser mais longos ou menos longos (por exemplo: o “*maillot de bain féminin d’une seule pièce dégageant les côtes, le bas du dos et les hanches*” — “maiô de banho feminino duma só peça que deixa livres as costas, a parte de baixo das costas e os quadris”, modelo atual sem nome) e de que cada pessoa é livre para construir sua perífrase como quiser. Ao contrário, a palavra está inteiramente pronta ou completamente construída no código, e o conjunto das palavras é finito.

Pode-se fazer corresponder a cada palavra uma perífrase, isto é, uma expansão; mas nem toda perífrase tem forçosamente uma palavra correspondente, isto é, “con-

densação”. Por isso se pode falar dum léxico finito, definir diferencialmente suas unidades, e tentar descobrir nele estruturas ao mesmo tempo semânticas e sócio-culturais. O sintagma-perífrase, ao contrário, corresponde a uma etapa preclassificatória do mundo, à transição entre dois sistemas de visão do mundo. A formação de terminologias científicas pode dar idéia da criação conjunta dum conjunto de noções e de termos que permitem servir-nos delas, dupla codificação indissociável do conteúdo e da forma.

Aliás, isso é válido do mesmíssimo modo para o imaginário, cujo papel é tão importante nas civilizações. A “mulher com cauda de peixe”, que tem o nome de *sereia*, tem para nós muito mais “realidade” que a “mulher com cabeça de ave”, que não tem nome nenhum. De maneira geral, deve sempre ser demonstrada a inexistência daquilo que não tem nome. Assim, o léxico é o testemunho dos conhecimentos, da mitologia e da ideologia duma civilização; constitui ele o próprio material da etnologia e da história.

As palavras metalingüísticas.

Acrescentar-se-á que não somente o léxico (palavras lexicais) nos dá uma imagem do mundo, mas também que nessa imagem do mundo está incluída uma imagem da língua.

A palavra lexical constrói até um sistema do léxico.

Como falamos duma língua, e justamente de seu léxico, senão com palavras? A palavra *preposição*, por exemplo, serve para falar das palavras *a, de, em*, etc., como a palavra *cão* serve para falar do galgo, do sabujo, do perdigueiro, do buldogue, etc. Essas palavras, ditas metalingüísticas, existem em todas as línguas e figuram sempre no dicionário (visto serem palavras como as outras). Mas constituem objeto da gramática por causa de seu conteúdo, assim como “cão” constitui o objeto de um tratado de zoologia.

A palavra é o melhor representante do léxico.

Se a palavra léxica define melhor os caracteres do léxico que a palavra gramatical, é fácil de ver que, entre as palavras lexicais, o substantivo é o mais representativo.

A ordem quantitativa.

Primeiro quantitativamente, e esse fato não é desprezível: aquele que aprende uma língua, natural ou estrangeira, deve sempre assimilar uma proporção mais elevada de substantivos à medida que progride. Um dicionário de dez volumes contém uma proporção de substantivos muito maior que um dicionário de bolso; constata-se facilmente que as terminologias técnicas e científicas, que fazem à vontade o léxico duma língua, são quase constantemente nominais. É entre os substantivos, e por causa de sua quantidade, que se encontram as palavras mais raras (baixa frequência), as palavras monossêmicas e as palavras mais longas, estando ligados entre si esses três caracteres. Basta citar como exemplo o nome de compostos químicos mencionados nas caixas de remédios.

A designação.

Esse fenômeno está ligado diretamente ao problema da designação. Esquecemos muito facilmente que em francês *nom* (nome) é ao mesmo tempo uma categoria gramatical e um “*appellatif*” (substantivo comum) (o nome duma flor); o inglês, que dispõe de duas palavras diferentes, respectivamente *noun* e *name*, faz-no-lo sentir melhor.

As taxinomias.

Ora, é o substantivo comum que nos permite organizar o mundo construindo classes (de objetos, de fatos, de pessoas, etc.), isto é, escolher quais traços comuns nos fazem encará-las da mesma maneira para opô-las a outra classe concebida do mesmo modo. Pode-se notar, a título de exemplo, que o francês não tem palavra para “animal marinho”, e que as palavras *mammifère* (*mamífero*), *poisson* (*peixe*), *arthropode* (*artropode*) etc., constituem classes no interior das quais só se poderá distinguir indivíduos marinhos, terrestres, etc. O que aconteceu foi que o caráter marinho, que poderia parecer importante, não foi considerado como suficiente para construir uma classe, um conjunto de coerência satisfatória, oponível a outros conjuntos. Inversamente, certas classes suficientemente homogêneas são subdivididas segundo as necessidades da experiência humana. Na maioria das línguas, os animais domésticos têm mais de dois nomes por espécie (*bœuf*, *taureau*, *vache*, *veau*, *génisse*: *boi*, *touro*, *vaca*, *bezerro*, *vite-la*), enquanto que os outros têm geralmente dois ou um só (*grenouille*, *têtard*; *rhinocéros*; *boa* etc.; *rã*, *girino*; *rinoceronte*; *boa*).

Nota-se que em francês as duas denominações mínimas retidas são as da oposição adulto-filhote e não as de macho-fêmea. Nenhum desses fatos de léxico deixa de ter importância, e até frequentemente estes se encontram no nó duma crise ideológica: não devem os franceses admitir que quando se diz “os homens são mortais”, trata-se também das mulheres, mas que em “as mulheres são mortais” os machos estão excluídos?

As substâncias.

Enfim, os substantivos são considerados a justo título, desde a Antigüidade, como portadores de substância e como significantes em si mesmos, à diferença das outras palavras lexicais. Constatam os filósofos que as palavras da linguagem primária (aprendidas pelo contato direto com o mundo) são substantivos (18), observam os psicólogos que a criança, na aprendizagem natural da língua, adquire primeiro os substantivos, depois os verbos e os adjetivos, vindo as palavras gramaticais em último lugar; Chomsky afirma que os caracteres do substantivo lhe são internos e não-contextuais, enquanto que os do verbo são contextuais. O substantivo é a parte do discurso menos gramaticalizada, a mais rebelde à análise distribucional.

O substantivo é também o mais perigoso dos fatores de perturbações.

Contudo, se as palavras lexicais, e especialmente o substantivo, são as mais representativas do léxico tal como o opomos à gramática, é por elas que o léxico tende a escapar do sistema da língua. A palavra lexical, representada principalmente pelo substantivo, une o sistema imanente duma língua a tudo o que não é ele: ao mundo (substância do conteúdo) e às outras línguas, e escapa parcialmente às leis da gramática e até da semiótica lingüística.

A onomatopéia e os ruídos do mundo.

É entre as palavras lexicais e especialmente entre os substantivos que se manifesta a onomatopéia, unidade mal codificada em que falta a arbitrariedade do signo. A onomatopéia é um signo em que a imagem sonora da palavra (significante) e a imagem mental dum ruído (significado) vêm confundir-se. Ademais, todos os ruídos podem ser notados à vontade, quer sejam conformes ao sistema fonológico da língua, quer não.

A sigla e o nome das letras.

Encontram-se também, essencialmente entre os substantivos, formações aberrantes que consistem em lexicalizar a seqüência das iniciais dum sintagma; são as siglas. Ora cada inicial gráfica é lida pelo seu nome, fenômeno inteiramente anormal: “une H.L.M.” [aʒeɛm], “un P.D.G.” [pedeʒe], ora o conjunto das iniciais é lido como uma palavra comum, e aí a despeito do sistema fonológico da língua: “le S.M.I.G.” [smig], “le C.N.I.T.” [knit]. Isso se agrava pelo fato de que a sigla é, como a onomatopéia, produtora de derivados (“smigard”, s.m.; “tictaquer”, v.; etc.).

As palavras e as coisas.

Sendo o substantivo a parte da fala mais diretamente ligada ao mundo exterior, traz ele sempre uma significação complexa, em evolução, ligada a uma função gramatical pobre.* É pelo substantivo que um estado de língua se modifica primeiro, visto que as coisas novas têm necessidade dum nome novo (neologismo).

O caminho do empréstimo.

Ora, essas coisas novas podem vir do estrangeiro, onde elas já têm nome. Consta-se que o empréstimo é não somente um fenômeno essencialmente lexical (oposto a gramatical), mas ainda e principalmente um fenômeno nominal.

O avanço dos nomes próprios.

Ao número esmagador dos nomes comuns vem ajuntar-se o conjunto indeterminado e muito mais numeroso ainda dos nomes próprios. O conjunto dos nomes próprios não é somente mais importante que o dos nomes comuns: ele é dum ordem de grandeza inteiramente diferente. Se essa desproporção não é perturbadora, é porque os nomes próprios têm freqüências relativamente baixas (conseqüência do fato de que eles só designam um objeto singular e não uma classe de objetos). Os nomes próprios fazem parte do léxico dum língua? Fazem, uma vez que não se pode falar uma língua sem usá-los, e não fazem, porque os nomes próprios pertencem a todas as línguas (com algumas acomodações de pronúncia ou de grafia). Estruturas fonológicas estrangeiras penetram, pois, obrigatoriamente em toda língua. Ademais, uma quantidade de nomes próprios, cujo número é indeterminado — lembremo-lo — produzem adjetivos e substantivos comuns (*Nova Iorque*: *nova-iorquino*; *Balzac*: *balzaquiano*; *Mao* [*Tsê-tungue*]: *maoísta*) que vêm integrar-se num conjunto em princípio finito. Muitos mesmo tornam-se substantivos comuns pelo uso de nomes registrados de marcas co-

* No extremo oposto, as palavras gramaticais apresentam o esquema inverso.

merciais: fuma-se um Pall-Mall bebendo um Campari. Os nomes próprios trazem, pois, ao coração do léxico, séries parcialmente estrangeiras e indeterminadas.**

A nominalização infinita.

O poder perturbador do substantivo manifesta-se assim no fenômeno da nominalização. Entende-se por esse fato que tudo, na linguagem, pode ser transformado em substantivo e esse privilégio só pertence ao substantivo; isso vem do fato de que o nome, exprime “o objeto de que se fala” e que se pode falar de tudo. As seqüências ditas “autônimas” são todas substantivos na fala, qualquer que seja a classe delas (o seu *relativamente* não foi ouvido; *não* serve para negar; ele grita: *não se aproxime*). Essa nominalização autonímica pode introduzir qualquer signo exterior à língua considerada, signo estrangeiro (*book* é inglês) ou signo não-lingüístico (= significa “igual”); esses dois exemplos constituem frases portuguesas irrepreensíveis, ao passo que “ele toma seu book” e “um metro = cem centímetros” são híbridas, sendo a primeira frase inaceitável por seu bilingüismo, misturando a segunda a língua natural e a língua matemática.

ONDE ACABA O LÉXICO?

Necessidade duma apreensão sincrônica.

Uma língua não pode ser descrita de improviso em sua duração histórica; não existe língua francesa, especialmente, que seria ao mesmo tempo a de Racine, de Vitor Hugo e de Sartre; pois se Sartre pôde ler Vitor Hugo que pôde ler Racine, o processo inverso é impossível. Ora, uma língua é um fenômeno social definido por trocas completas (recíprocas). O campo das trocas reais ou possíveis é um estado de língua de curta duração; a duração máxima duma sincronia prática é a que pode projetar-se na memória (transformação do tempo histórico vivido em experiência presente), isto é, a duração duma vida humana.

Competência (e incompetência) lexical.

Ninguém conhece perfeitamente uma língua porque ninguém pode conhecer todas as palavras dela. A maioria dos usuários duma língua dominam a gramática, isto é, sabem distinguir uma frase correta duma frase incorreta, e um gramático profissional pode atingir uma competência gramatical ótima.

Não conhecemos jamais todas as palavras de nossa própria língua.

Mas os usuários não dominam jamais o léxico, encontram em todo o decorrer de sua vida palavras desconhecidas, e nenhum lexicólogo ou lexicógrafo pode esperar adquirir uma competência lexical ótima. Deve-se isso, evidentemente, à ordem quantitativa: as regras da gramática são em número restrito, mas não as palavras que elas regem. Além disso, é o léxico que, na língua, muda mais depressa (em francês, a renovação das unidades é da ordem de 10% em 25 anos para cerca de 50.000 palavras). O fato da maioria das pessoas não “compreenderem” um texto é devido às palavras desconhecidas.

** *Silhueta, cavanhaque, gilete*, são exemplos de nomes próprios que se tornaram substantivos comuns. (Nota do tradutor)

das: resulta disso uma espécie de desconforto permanente, de angústia léxica que se transforma freqüentemente em observações desdenhosas sobre o “jargão” dos outros.

O idioleto.

Cada um de nós tem um vocabulário, componente lexical do nosso idioleto; o vocabulário dum indivíduo é único, tanto pela quantidade de palavras conhecidas como pela natureza dessas palavras. É difícil recensear as palavras dum vocabulário. Por um lado, porque nem todas as palavras conhecidas pela pessoa são empregadas efetivamente na fala ou nos textos observados e, por outro lado, porque uma palavra pode ser conhecida ativamente ou passivamente: o vocabulário ativo é o que se tem o costume de empregar; o vocabulário passivo é o que compreendemos quando empregado por outras pessoas, mas que nós mesmos não temos o costume de empregar (assim certas palavras grosseiras muito conhecidas, para tomar um caso típico). Pareceria que, segundo a cultura dos indivíduos, o vocabulário varia para uma língua de civilização entre 3.000 e 40.000 palavras. Ora, os maiores dicionários de língua, que não incluem os nomes próprios (e não históricos), ultrapassam facilmente 100.000 palavras. Contudo, se cada vocabulário é único em quantidade e em qualidade, todos os idioletos têm uma grande parte de palavras em comum, garantia da comunicação e da realidade do léxico.

Cada usuário do português domina pouco mais ou menos todas as palavras gramaticais, e todas as palavras lexicais de primeira necessidade, que têm uma alta freqüência na fala. É entre as palavras de média e baixa freqüência que as diferenças começam a fazer-se sentir (palavras técnicas, científicas, literárias, de gíria, arcaizadas, novas, regionais, etc.). As noções de freqüência e de fechamento de conjuntos estão *grosso modo* em relação, e essa relação justifica também que se faça passar a fronteira do léxico entre as palavras gramaticais e as palavras lexicais.

Léxico comum, léxico total.

Partindo da experiência idioletal, pode-se, pois, abordar o léxico de duas maneiras. O léxico comum dum estado de língua dado é constituído por todas as palavras comuns a todos os usuários (intersecção dos idioletos); o léxico total é constituído por todas as palavras empregadas por todos os usuários (reunião dos idioletos) (9). O léxico comum representa uma língua pobre que funciona bem no conjunto da sociedade, mas que exprime pouco. O léxico total representa uma língua rica que funciona mal no conjunto da sociedade, mas que exprime muito.

Indeterminação do léxico total

Pode-se dizer que uma palavra rara é “menos palavra” que outra, visto como seu valor de troca é menor.

Não estamos seguros de poder reconhecer as palavras como tais.

Uma palavra que funciona para cem pessoas não tem o mesmo estatuto sociolinguístico que uma palavra que funciona para dez milhões; ela apresenta sempre caracteres menos típicos da língua em questão. E uma “palavra” que funciona para uma só pessoa (uma invenção qualquer) já não é uma unidade da língua. O léxico total é um conjunto ideal que se perde na indeterminação e que nenhum dicionário jamais pôde

descrever *in extenso*. Não é a simples quantidade das unidades que torna essa tarefa impossível, mas o fato de que a qualidade linguística das unidades é cada vez mais incerta à medida que elas são mais raras. A isso vêm acrescentar-se as dificuldades práticas (representatividade do *corpus* para as unidades raras, fechamento do *corpus* ao passo que cada dia o léxico se modifica). O léxico total é um “conjunto impreciso” [“ensemble flou”] (Gentilhomme,7).

Escolha de “um” léxico.

De fato, quando se fala de léxico, encaramos um conjunto intermediário entre o léxico total e o léxico comum, podendo um ser estimado em alguns milhares de palavras aproximadamente, e sendo o outro de várias centenas de milhares. Mas entre esses dois limites teóricos, toda fronteira é arbitrária e simplesmente prática: o tamanho variável dos dicionários vem ilustrar essa escolha.

OS SUB-LÉXICOS COMO TIPOS ABSTRATOS.

Os idioletos apresentam diferenças qualitativas que se podem reunir em tipos. Cada pessoa sente que compreende melhor as pessoas de sua região, de sua idade, de seu meio social e profissional que as outras. Encararemos, pois, no interior do léxico total, quatro tipos de sub-léxicos: o das línguas regionais, das línguas sociais, das línguas temáticas e das línguas de gerações.

Esses quatro tipos abstratos não têm nenhuma realidade, se não como componentes de cada idioleto; pode-se imaginar, por exemplo, o vocabulário dum morador de Lille de 60 anos, operário da indústria têxtil, ou o de um parisiense de 20 anos, estudante de farmácia. Os tipos de sub-léxicos a considerar e a importância deles variam com a língua estudada: regiões mais ou menos centralizadas, sociedades com ou sem classes, castas, religiões, etc.

As línguas regionais.

Os falares regionais não devem ser confundidos com os dialetos e as línguas independentes: assim na França o bretão, o provençal, o basco, são línguas diferentes do francês. Mas o francês tal qual se fala em Paris, em Marselha, em Lille ou em Bordeaux apresenta variações léxicas (e fonológicas) que, embora não perturbem gravemente a intercompreensão, são todavia sensíveis quando se passa de uma região a outra. O peixe que se chama *colin* em Paris denomina-se *merlu* em Bordeaux.

As línguas sociais.

As línguas sociais refletem principalmente, em nossas civilizações, a oposição da classe dirigente (que é, em geral, a classe culta) às outras classes. Mas a divisão pode ser diferente. A despeito da democratização e da mistura social, constata-se, na França, que a burguesia tem repugnância em usar certas palavras correntes na classe operária. Mesmo a gíria, que se difundiu em todos os meios, achou-se mais ou menos redistribuída seletivamente segundo as classes. O sistema oficial da educação escolar faz respeitar a norma da classe dirigente.

As línguas das gerações.

As línguas das gerações são línguas lexicalmente diferentes que são faladas num momento dado por pessoas de idades diferentes que coexistem na sociedade. Os moços não falam como as pessoas idosas. Em outras palavras, uma visão rigorosamente sincrônica do léxico traz as marcas do tempo. Em 1972, as pessoas de 60 anos conservavam a maioria dos hábitos de linguagem de seus 20 anos (1932), quase todos os de seus 40 anos (1952), etc. Esses hábitos são mantidos pelo diálogo com pessoas da mesma idade. As trocas entre pessoas de idades diferentes e a renovação constante do léxico se fazem graças ao duplo estatuto, ativo ou passivo, do vocabulário. O pai compreende o filho sem usar as mesmas palavras, e emprega palavras de sua juventude que o filho compreende sem se utilizar delas. As palavras morrem por falta de combatentes: uma palavra que estava na moda em 1932 é cada vez menos empregada à medida que seus utilizadores envelhecem e morrem.

Nas línguas de civilização, a leitura e a cultura vêm corrigir esse esquema; palavras que deveriam cair pouco a pouco no vocabulário passivo dos usuários são às vezes mantidas no vocabulário ativo (acontece-nos falar e sobretudo escrever como Gide), e palavras desde muito tempo fora de uso se mantêm no vocabulário passivo (compreendemos ainda Racine).

As línguas temáticas.

As línguas temáticas agrupam temas de interesse: atividade profissional, ocupações, lazeres. Este subgrupo opõe-se aos outros três na medida em que o tema que o define é exterior ao mesmo tempo ao sistema da língua e à situação de seus usuários. É o domínio do conhecimento, quase inesgotável, que reúne todas as espécies de terminologias de limites incertos e um número considerável de nomes próprios que se acham ligados a elas. As línguas temáticas, por sua relação direta com o mundo, fornecem os elementos do léxico mais numerosos e os mais instáveis, geralmente consignados nas enciclopédias e dicionários especiais.

Entre as línguas temáticas, é preciso contar as da geografia e da história, e não confundir o léxico dessas línguas com o das línguas regionais e das línguas de gerações; a palavra *gabelle* (gabela), por exemplo, é um termo atual que designa um imposto antigo, portanto um termo de história, mas não uma palavra arcaica. A história atualiza o léxico antigo do qual ela se alimenta.

O tema da língua é também gerador duma língua temática, à da lingüística, cujo léxico está consignado nas páginas deste livro*.

SEMÂNTICA LEXICAL.

A semântica lexical representa ao mesmo tempo o essencial da semântica e seu domínio mais confuso, verdadeiro terror dos lingüistas. O estudo do sentido das palavras está com efeito tão tradicionalmente comprometido com os debates filosóficos, psicológicos e literários, que parece difícil abordá-lo em bases científicas.

* A autora se refere à enciclopédia lingüística *Le langage* onde seu trabalho foi publicado.

O sentido das palavras não é nem transcendental nem produzido pelo contexto ...

A primeira reação salutar dos lingüistas consistiu em negar que existam palavras com significado determinado fora da frase contextual que as contém, e que é a única a permitir a manifestação do sentido (o sentido duma palavra é o seu emprego). Se uma palavra é empregada sozinha, é que essa palavra é por si só uma frase, como *Vem!* ou *Garçom!* Se uma só palavra pode funcionar como uma frase, é porque o contexto situacional vem saturar ora o sujeito (“Tu queres vir?”), ora o predicado (“Garçom, traga a conta”). Por isso o turista estrangeiro pode fazer-se compreender com um “pequeno léxico prático” sem conhecer uma só regra de gramática (ao passo que uma competência gramatical sem léxico exclui absolutamente toda comunicação).

Certamente era necessário mostrar que as palavras não têm significado transcendental que viesse de outro lugar senão do discurso: ou do mundo das essências, ou da história, ou da própria forma material delas. Alude-se aqui ao “realismo antigo”, à “prova pela etimologia”^{*} e à “magia das palavras”. Não obstante, se se pode falar da palavra como unidade de língua e não como ocorrência particular na fala, não é preciso dizer que se pode e que se deve falar do significado duma palavra nos mesmos termos. Cada um de nós é capaz de evocar o sentido de *cão*, *morrer*, *feliz*, *nunca...* sem recorrer a uma frase particular.

... é a resultante de contextos já produzidos.

De fato, a maioria dos lingüistas estão de acordo em admitir, num espírito mais sintético, que o significado duma palavra é a somatória e a codificação de todos os significados particulares que se lhe viram tomar recentemente em frases particulares. Isto é, de outro modo, que cada ocorrência duma palavra numa frase está ligada a um significado constante. Quando o lexicólogo fala do sentido duma palavra, não se trata da palavra realizada na fala, mas da palavra abstrata, metalingüística (a palavra que é o nome de todas as suas ocorrências). E a justo título, uma vez que o gramático fala corretamente da interrogação, do subjuntivo, etc., como palavras metalingüísticas das ocorrências de frases interrogativas, de verbos no subjuntivo, etc.

É preciso, pois, pôr de lado duas atitudes extremas: a primeira que consiste em crer que as palavras são entidades com significado determinado do exterior, e que a frase as utiliza tais quais, como cubos dum jogo de construção;^{**} a segunda a de que a palavra só tem sentido numa frase particular, e que tem por corolário a hipótese de que a mesma palavra tem em cada frase um sentido um pouco diferente (6). Esse excesso vem geralmente da confusão entre o elemento e o contexto; nas frases *Este lingüista é inteligente* e *Este lingüista é bonito*, ver-se-iam dois sentidos de lingüista segundo os predicados: o “lingüista-espírito” e o “lingüista-corpo”, o que é absurdo, uma vez que um não existe sem o outro.

^{*} Paulhan denomina assim a interpretação errônea dos letrados que querem explicar o sentido duma palavra por sua etimologia. Por exemplo, a própria palavra *etimologia* significaria *étymos logos*, “sentido autêntico” (o que é falso).

^{**} A hipótese duma denominação primeira e voluntária dos objetos nas origens da linguagem é inteiramente gratuita; no caso raro das palavras definidas antes de serem utilizadas (termos científicos), essas palavras sofrem pelo uso a mesma sorte das da herança comum.

Monossemia, polissemia, homonímia.

Uma palavra pode ser monossêmica (*lingüista*) ou polissêmica (*língua*). Chama-se “sentido” duma palavra polissêmica ao significado de cada conjunto codificado de ocorrências oposto a outros conjuntos codificados de ocorrências da mesma palavra: por exemplo, *língua*, “órgão bucal”, e *língua*, “sistema de expressão comum a um grupo social”.

Esses sentidos diferentes correspondem a distribuições sintático-lexicais diferentes (categorias gramaticais do contexto, ou palavras do contexto) cada vez mais difíceis de descrever à medida que nos afastamos da palavra gramatical para o substantivo (desgramaticalização progressiva das partes do discurso acima evocadas). Aquém da polissemia coloca-se a questão da homonímia; pode-se perguntar, com efeito, se não é necessário considerar duas palavras *língua* homônimas de preferência a uma só palavra *língua* polissêmica.

Graus de codificação.

Para lá da polissemia, abordam-se os efeitos de sentido cuja codificação é fraca ou nula no código da língua corrente (estilística). O conjunto dos agrupamentos semânticos possíveis das ocorrências duma mesma palavra se distribui ao longo duma linha cujos cimos são: palavras diferentes, sentidos diferentes, efeitos de sentido diferentes, sentidos conotativos, sentidos fora do código. O lexicólogo só estuda um sentido por vez e adota a ficção da unicidade monossêmica.

O sentido ou os sentidos só se exprimem por palavras.

Todo significado ou sentido é expresso por uma seqüência de palavras: não há outra maneira de falar sobre isso (*livraria* significa “loja onde se vendem livros”). O significado duma palavra só pode, pois, ser conhecido quando transmitido pelo significante de outras palavras (14). Esse sistema sinonímico circular no interior do léxico é o único que torna possível a descrição da face abstrata do signo.

O lexicólogo deve organizar semanticamente o léxico.

Sendo o léxico essencialmente transmissor de sentidos, trata-se de saber se ele apresenta estruturas que levem em conta ao mesmo tempo o semantismo duma língua (forma do conteúdo) e o sistema conceptual de seus usuários (substância do conteúdo).

Papel da sintaxe e da morfologia.

As classes de palavra, dependentes da sintaxe, têm implicações semânticas que organizam prioritariamente o léxico (substantivo → objeto, fenômeno, estado, ação, qualidade; verbo → processo, etc.). A composição das palavras, que depende da morfologia, na medida em que ela segue as regras do sistema, mostra as estruturas morfo-semânticas do léxico (*triste* → *entristecedor*, *entristecer*, *entristecimento*...). As palavras têm leis de composição que reproduzem modelos semânticos constantes (*raiva* → *raivoso*; *teima* → *teimoso*, etc.) (5). O agrupamento dos compostos em torno duma base dada não deve ser confundido com a tradicional “família de palavras” que reúnia unidades aparentadas somente pela etimologia, isto é, pela história (francês *grève*, *gravier*, *graveleux*; port. *letra*, *literatura*), e não pela expressão e pelo conteúdo.

A abordagem puramente semântica.

A pesquisa das estruturas propriamente semânticas, isto é, daquelas que não se apóiam nem na função nem na forma, é muito mais difícil, porque é fundamental. Parte-se geralmente do significado intuitivo das palavras (competência semântica), que é verificado pela análise distribucional, e que se tenta depois precisar pela aproximação com outras unidades semanticamente aparentadas.

Foram adotados vários métodos: o dos campos lexicais, que são campos temáticos; o das relações lógicas ou conjuntistas (sinônimos, opostos, contrários, complementares, incluíntes ou hiperônimos, incluídos ou hipônimos); o dos marcadores semânticos e dos diferenciadores.

Conjuntos fechados e conjuntos abertos.

Existem microssistemas lexicais fechados (determinados) cuja estrutura é clara porque designam quer um conjunto natural com muito poucos elementos, quer um conjunto convencionalmente bem definido do mundo: os nomes dos dedos da mão, dos dias da semana, dos postos militares, dos signos do zodíaco, etc. Mas outros, aparentemente do mesmo tipo, já são abertos (indeterminados): os nomes dos ossos do esqueleto, dos cursos superiores, dos sinais tipográficos, das bebidas alcoólicas, etc. Ora, se o conjunto é aberto, a descrição semântica diferencial de seus elementos é impossível; *a fortiori* a dos elementos de conjuntos abertos diferentes. Na escala do léxico total, um relacionamento generalizado das palavras torna-se inteiramente aleatório, e ninguém pôde levar seus esforços até aí para nenhuma língua. Se, por outro método, se parte das definições de dicionário, que se admite darem o sentido das palavras, mas que, de fato, nos dizem o que são as coisas, é-se remetido a outro sistema diferencial aberto ainda menos manejável: o do conhecimento e da ideologia.

Tais são os problemas essenciais que um estudo sincrônico, isto é, funcional e estrutural do léxico, coloca. No tempo em que o lexicólogo se contentava em estudar “a vida das palavras” sem nunca tentar descrever uma língua, a lexicologia ia bem. É mais fácil descrever mudanças que avaliar diferenças. Hoje, o lexicólogo é, antes de tudo, requisitado a propor ao gramático uma descrição completa do léxico, especialmente semântica, que venha articular-se com a do gramático. Idealmente, uma língua seria descrita se se pudesse fornecer às máquinas de traduzir o programa gramatical e lexical necessário a todas as traduções corretas e somente a elas. (4).

O DICIONÁRIO DE LÍNGUA

Mas o dicionário não é essa descrição completa do léxico? (16) A maioria dos usuários o crêem, e as publicidades não procuram desenganá-los. O dicionário é um dos objetos culturais mais usuais e mais mal conhecidos.

O dicionário de língua deve inicialmente ser situado entre grande número de obras que apresentam com ele algumas semelhanças. Um dicionário é um texto duplamente estruturado que apresenta: a) uma seqüência vertical de itens, ditos “entradas”, geralmente dispostos em ordem alfabética, seqüência essa chamada “nomenclatura”; b) um programa de informação sobre essas entradas, que forma com elas os verbetes. As entradas são sempre signos lingüísticos, e a informação dada deve aplicar-se, ainda que em pequena parte, ao signo, como o faria, por exemplo, a lista telefônica. Considera-se

que a definição é uma informação sobre o signo (seu significado) e sobre a coisa designada pelo signo (o que essa coisa é).

Os dicionários falam-nos dos signos e das coisas

A dupla estrutura do dicionário faz dele uma obra de consulta e não um texto para ser lido do começo ao fim.

Classificação dos dicionários

Pode-se considerar três tipos de dicionários se se leva em conta a informação sobre os signos ou sobre as coisas: o dicionário linguístico, que só dá informações sobre os signos, com exclusão da definição (dicionário etimológico, por exemplo); a obra enciclopédica, que só dá informações sobre as coisas, incluindo a definição (dicionário técnico de eletricidade, ou então o presente dicionário*), e o dicionário de língua, que dá informações sobre os signos, incluindo a definição. Esses três tipos se dividem em dois grupos: o dicionário geral, que trata de todos os signos numa língua dada ou de todas as coisas numa civilização; e o dicionário especial, que só descreve um setor de uma ou da outra. O dicionário etimológico é um dicionário geral, e o dicionário de sinônimos, um dicionário especial.

O dicionário de língua é um dicionário geral que nos fala do conjunto das palavras numa língua e que dá a definição delas. Sua nomenclatura apresenta, pois, todas as classes de palavra, geralmente com exceção dos nomes próprios, e indica-se a classe de palavra de cada entrada.

A enciclopédia é também um dicionário geral, mas que nos fala do conjunto das coisas numa civilização e que dá a definição delas (seu projeto de ser “universal” é afastado pela língua empregada e pelo sistema cultural que lhe está ligado).

Sua nomenclatura é essencialmente nominal e inclui especificamente nomes próprios e ilustrações com legenda nominal. Não apresenta as classes de palavra, informação aliás inútil, uma vez que só existem substantivos.

Toda nomenclatura é nominal ou nominalizada.

O artigo de dicionário, embora seja escrito em estilo telegráfico (*dentro*: no interior de), deve ser lido como uma longa frase cujo sujeito gramatical é a entrada. Ora, todo predicado** exige um sujeito nominal, e a entrada funciona em todos os casos como um substantivo.

Quando a entrada já é um substantivo, a leitura do artigo é, por exemplo, para *Observatório*: (Um) observatório (é um) estabelecimento científico destinado às observações astronômicas e meteorológicas, ou ainda: *Observatório* significa “estabelecimento científico destinado às observações astronômicas e meteorológicas”. Mas quando a entrada não é um substantivo, só uma leitura é possível, por exemplo, em *dentro*: *Dentro* significa “no interior de”, em que *dentro* se transforma em substantivo uma vez que se trata da palavra *dentro*. É por isso que os fatos de nomenclatura, sozinhos, permitem selecionar qual tipo de informação vai aparecer: sobre as coisas, se são possíveis duas leituras; sobre os signos se só existe uma.

* A Autora refere-se a obra em que seu trabalho foi publicado. (Nota do Tradutor)

** Frase por meio da qual se diz alguma coisa dum objeto.

Realizações híbridas. A oposição *dicionário de língua/enciclopédia* é a de dois modelos abstratos. As obras reais têm geralmente menos coerência (13) e, no limite, já não se pode nem determinar de que elas falam. Assim, a “Enciclopédia” de Diderot apresenta substantivos, verbos, adjetivos, não advérbios, e algumas palavras gramaticais; certos dicionários de medicina dão etimologias, etc. Existe, contudo, um tipo de dicionário heterogêneo bem definível e muito conhecido, o dicionário enciclopédico, que constitui a soma dos dois modelos descritos; o *Webster’s dictionary* e o *Petit Larousse* são excelentes exemplos disso.*

Hierarquia nos signos e nas coisas.

A verdadeira dificuldade, todavia, não é distinguir entre os signos e as coisas, mas apreender um ou o outro desses “conjuntos imprecisos” (“ensembles flous”).

Os dicionários não podem compor livremente sua nomenclatura.

Um dicionário geral da língua teria por finalidade a descrição do léxico total, e uma enciclopédia universal, a descrição de tudo o que existe para uma civilização dada. Sendo indeterminados esses conjuntos, é preciso, para descrevê-los, proceder do mais evidente para o menos evidente, do mais importante para o menos importante. O critério de importância, para a palavra, é a frequência; um dicionário que trata da palavra *balafo* não pode dispensar-se de tratar da palavra *violino*; o critério de importância para a coisa designada pela palavra é a notoriedade: uma enciclopédia que nos fale de Boulez não pode fazer silêncio sobre Mozart (salvo com fins polêmicos, o que pode acontecer).

Reduções estruturadas.

Um dicionário geral permanece geral, qualquer que seja o número das unidades de sua nomenclatura, contanto que ele respeite essa lei. Donde a coexistência de dicionários com nomenclatura de 5000 palavras, de 20.000 palavras, de 50.000 palavras..., devendo a nomenclatura mais vasta teoricamente conter todas aquelas que o são menos (o que não se verifica na prática por causa dos dados incertos sobre as frequências).

Os dicionários são nossa única idéia do léxico.

Os dicionários de língua também se aproximam, do menor ao mais volumoso, da competência léxica ideal** sem jamais atingi-la. Representa o dicionário, por si só, essa competência impossível de ser apreendida, competência essa que nós nos esforçamos por atingir; completa cada vez de modo diferente os milhões de idioletos diversamente compostos. Atualmente, os cursos de lexicologia são de fato cursos sobre os dicionários, objetos sócio-culturais que demarcam a pista dum léxico ao qual ninguém tem acesso diretamente. A situação da gramática é totalmente diferente (17).

* Eles são, na realidade, ao mesmo tempo, enciclopédias (nomes próprios; ilustrações das coisas designadas) e dicionários de língua (todas as palavras, inclusive as palavras gramaticais; informações sobre os signos). (Do *Petit Larousse* há uma adaptação portuguesa — o *Dicionário Prático Ilustrado* — e uma espanhola — *Nuevo pequeno Larousse ilustrado*. nota do tradutor).

** A competência do usuário teórico ideal que conhecesse toda a sua língua.

Definição da unidade.

Quanto ao programa de informação sobre cada unidade, comporta ele pelo menos três elementos: o componente gráfico e fônico, isto é, a menção da palavra com sua grafia, seguida de sua pronúncia; o componente sintático, isto é, a classe de palavra (e, eventualmente, o gênero, o número); o componente semântico, isto é, a definição, ou análise do significado. Essas três informações definem tradicionalmente uma unidade lexical.

O exemplo.

O dicionário de língua dá também exemplos; o exemplo destina-se a mostrar a palavra definida em funcionamento e, em suma, a ministrar provas do que se acaba de afirmar. Nessa perspectiva, poder-se-ia dispensá-lo. Mas acontece que ele é, na verdade, muito mais do que isso. Primeiro, ele assume na maioria das vezes a informação sintática necessária sobre as restrições seletivas (tal verbo tem sempre um sujeito animado, tal adjetivo só se emprega com tal palavra, etc.). Depois, ele serve para atenuar as deficiências das três informações supracitadas: seja, por exemplo, para mostrar, em francês, a elisão do artigo diante de um substantivo, para apresentar uma palavra numa função mais rara não descrita, e principalmente para fazer surgir o sentido que muito vezes foi mal captado na definição. Enfim, serve para corrigir a brutalidade da nomenclatura reintroduzindo todas as unidades codificadas superiores na palavra gráfica tradicional (lexias, locuções, provérbios), e as que estão no caminho da codificação (coocorrências freqüentes: *rectangle blanc*, *image de marque*, *n'être pas sans savoir*, *sévère mais juste...*). Na verdade ele é, pois, essencial. Acreditou-se até, especialmente, que ele poderia substituir uma definição aleatória e mais ou menos subjetiva. Mas não existe nenhuma descrição válida numa língua que consista em mostrá-la: não teríamos, então, nenhuma necessidade dum dicionário, tendo olhos e ouvidos. E se se apresentam todos os exemplos pertinentes e somente os exemplos pertinentes para revelar o significado da palavra em questão, essa escolha dos exemplos implica a mesma reconstrução teórica que a escolha da definição, e exige do leitor um esforço mais dispendioso.

A definição

A definição é uma perífrase que pretende ser sinônima da palavra a ser definida (12,21) e que se acha colocada numa predicação do tipo: “um X é um Y que..., etc.”, isto é, numa identidade entre duas classes de objetos (e não de signos). É só falando das coisas que a definição pode dar-nos o sinônimo do signo. Os lingüistas, todavia, recusam esse método não lingüístico que constitui o verdadeiro interesse do dicionário, embora usando naturalmente definições assim obtidas, em seus trabalhos teóricos. A função gramatical da definição é, por causa da sinonímia, obrigatoriamente a mesma que a da palavra considerada. Não é nunca uma frase.

Estruturas definicionais. Tendo valor de palavra, a definição pode ser tratada pelos mesmos métodos de análise semântica: classificação das definições pela função gramatical (classes de palavra), pela relação morfo-semântica ou puramente semântica com o definido. Assim, a definição de *librairie* por “magasin du libraire” é morfo-

semântica visto que a forma *librair(e)* é conservada; ao passo que a definição por “*magasin ou l’on vend des livres*” é puramente semântica.

Mas, quaisquer que sejam as relações entre o significante e o significado, o fato capital é que a definição tem uma sintaxe que compõe e hierarquiza os elementos de sentido (semas) que ela nos dá. Essa gramática sêmica deve também ser aproximada das leis de composição das palavras, embora seja mais complexa e mais flexível. Ela permite encontrar os modelos lógicos de análise do sentido evocados acima, os quais são universais, especialmente o modelo fundamental da inclusão (todas as livrarias são lojas, ou ainda a classe das livrarias está incluída na classe das lojas).

O dicionário monolíngüe é feito para a tradução...

O dicionário de língua monolíngüe do qual se acaba de tratar, é um dicionário de descodificação, de tradução. Conhece-se o significante da palavra e procura-se o seu significado. A operação inversa de codificação, da versão, é impossível: não se pode encontrar nele o(s) signficante(s) desconhecido(s) dum significado a exprimir. O dicionário *Robert* procurou, em francês, escapar a essa insuficiência pelo emprego de exemplos-perífrases que remetem a uma palavra; assim em *Sanglier*, o exemplo “*Femelle du sanglier. V. Laie*”.

... enquanto que o bilíngüe funciona para a tradução e para a versão.

O dicionário bilíngüe é normalmente duplo, para as necessidades da versão e da tradução, por exemplo francês-inglês e inglês-francês. Mas a situação é mais simples pois se trata, principalmente, nos dois sentidos, de dar o equivalente desconhecido dum palavra conhecida (de signo para signo) e não dum perífrase. O recurso à definição só é em verdade necessário quando aquilo que tem nome numa língua não o tem na outra; enquanto que no dicionário monolíngüe os equivalentes (sinonímia de palavras) são raríssimos e é necessário recorrer à analicidade definicional.

O dicionário histórico é aberrante mas útil.

Tem-se o costume de chamar *dicionário histórico* a um dicionário de língua que informa sobre a história das palavras. Não deve o dicionário histórico ser confundido com as obras antigas ou modernas que tratam dum estado de língua antigo (“*Dicionário do Século XVII*”, por exemplo). O modelo teórico do dicionário histórico, poucas vezes realizado, é uma obra cuja nomenclatura contém todas as palavras dum língua desde as suas origens, isto é, o conjunto das palavras desusadas e das palavras usadas; e para as palavras em uso, contém o conjunto dos sentidos arcaizados e dos sentidos atuais. Essa óptica histórica provoca a informação sobre a etimologia, ou origem da palavra.

O dicionário histórico não descreve de fato nenhuma língua real uma vez que sua nomenclatura acrônica* amontoa palavras de todas as épocas (de vários estados de língua reais) que não funcionaram simultaneamente, e superpõe estruturas lexicais incompatíveis. A vizinhança, na nomenclatura, dum palavra fora de uso e dum pala-

* *Acrônico*: que reúne, sem levar em conta o tempo, elementos situados no tempo.

vra atual poderia ser comparada à vizinhança (não observada) de palavras de línguas diferentes com a duma palavra alemã e duma palavra italiana. E na medida em que todas as palavras antigas são definidas quer por um equivalente atual, quer por uma perífrase atual, o dicionário histórico é, de alguma maneira, um dicionário bilíngüe (e até multilíngüe). Todavia, uma situação lingüística de fato pode justificar sua existência e utilidade: por um lado, o vocabulário passivo das pessoas cultas pode recuar muito longe no tempo; por outro lado, se o dicionário de língua é um dicionário de tradução, é preciso admitir que as palavras desconhecidas devem figurar nele e que as palavras desconhecidas do leitor são muitas vezes palavras antigas; as civilizações da leitura têm uma competência lexical passiva que modifica profundamente a competência normal da comunicação. Finalmente, é sobretudo o mau uso do dicionário histórico que torna discutível a sua utilidade; o leitor, que deveria ater-se somente ao critério do uso atual, vai buscar no passado razões para negar a realidade presente; ele deseja encontrar o sentido duma palavra em sua etimologia, quando a coincidência dos dois é rara; tem toda indulgência para com o arcaísmo, e toda severidade para com o neologismo, embora nem um nem o outro pertença já (ou ainda não) à língua. O dicionário histórico incita-o a confundir a dinâmica do tempo com a degradação dum absoluto, duma "pureza" primitiva. Os puristas são essencialmente conhecedores da história da língua.

REY-DEBOVE, J. — Lexicon and dictionary. Trad. de Clóvis Barleta de Morais. Alfa, São Paulo, 28(supl.):45-69, 1984.

ABSTRACT: Rey-Debove discusses the relationship between lexicon and grammar, defining basic concepts in Lexicology: the word, lexical unit, lexie, morpheme, open class and closed class, and lexical repertoire of a language. Lexicon is the least specific domain of a language, as it reports to referential universe. The set of signs that makes up lexicon gives it a semiotic status by creating a world system (a "système de monde"). The author discusses the problem of designation, demonstrating the relationship between words and things. Being an open system, lexicon has infinite possibilities of expansion. Dictionary is the only manner in which a lexicon can be conceived.

KEY-WORDS: Lexicon; lexicology; dictionary; lexical unit; lexie; reference; world system ("système de monde"); metalanguage; terminology; designation; taxinomy; polysemy; homonymy; nomenclature; entry dictionary; definition; monolingual dictionary; bilingual dictionary; historical dictionary.

REFERÊNCIAS BIBLIOGRÁFICAS

1. BENVENISTE, E. — *Problèmes de linguistique générale*. Paris, Gallimard, 1966.
2. BLOOMFIELD, L. — Form classes and lexicon. *Language*, 38, 1962.
3. BOTHA, R. — *The function of the lexicon in transformational generative grammar*. La Haye, Paris, Mouton, 1968.
4. DELAVENZY, E. — *La machine à traduire*. Paris, P.U.F., 1959 (Collection Que sais-je?)
5. DUBOIS, J. — Recherches lexicologiques: esquisse d'un dictionnaire structural. *Études de Linguistique Appliquée*, (1):43-48, 1962.
6. FIRTH, J.R. — *Papers in linguistics, 1934-1951*. London, Oxford Univ. Press, 1957.
7. GENTILHOMME, Y. — Les ensembles flous en linguistique. *Cahiers de Linguistique Théorique et Appliquée*, 5, 1968.
8. GLEASON, H.A. — The relation of lexicon and grammar. In: HOUSEHOLDER, F.W. & SAPORTA, S., eds. — *Problems in lexicography*. Bloomington, The Hague, Mouton, 1967.
9. HOCKETT, C.F. — *A course in modern linguistics*. New York, Macmillan, 1964.

10. MARTINET, A. — Le mot. *Diogène*, (51) 39-53, juil./sept., 1965.
11. MULLER, C. — Le mot, unité de texte et unité de lexique en statistique lexicologique. *Travaux de Linguistique et Littérature*, 1, 1963.
12. POTTIER, B. — La définition sémantique dans les dictionnaires. *Travaux de Linguistique et de Littérature*, 3(1):33-39, 1965.
13. QUEMADA, B. — *Les dictionnaires du français moderne 1539-1863; étude sur leur histoire, leurs types et leurs méthodes*. Paris, Didier, 1968.
14. QUINE, W. Van O. — *From a logical point of view*. New York, Harper Torchbooks, Harper and Row, 1963.
15. REY, A. — *La lexicologie*. Paris, Klincksieck, 1970.
16. REY, A. — Les dictionnaires: forme et contenu. *Cahiers de lexicologie*, 7(2):65-102, 1965.
17. REY-DEBOVE, J., org. — La lexicographie. *Langages*, (19):1-119, sept., 1970.
18. RUSSEL, B. — *An inquiry into meaning and truth*. Harmondsworth, Penguin Books, 1969.
19. TOGEBY, K. — Grammaire, lexicologie et sémantique. *Cahiers de Lexicologie*, 6(1):3-7, 1965.
20. ULIMANN, S. — *Semantics: an introduction to the science of meaning*. Oxford, Brasil Blackwell, 1962.
21. WEIREICH, U. — La définition lexicographique dans la sémantique descriptive. *Langages*, (19):69-86, sept., 1970.

SEMÂNTICA E LEXICOGRAFIA

Por Julio CASARES*
Tradução de Balbina Lorenzo FEIJÓO-HOYOS**

*RESUMO: A Semântica e a Lexicografia se interpenetram mutuamente porque a Lexicografia não se limita a recolher as palavras do léxico, mas procura descrever a significação dos vocábulos e seus usos. O lexicógrafo também se ocupa de evolução dos sentidos das palavras para estabelecer a escala das acepções de um signo lexical. Casares conceitua acepção e discute o problema da discriminação das acepções e da sua ordenação no caso de palavras polissêmicas. Outra questão delicada para o lexicógrafo é o reconhecimento e a identificação correta dos valores metafóricos. O autor usa como exemplo ilustrativo o verbete lat. *ordo* > esp. *orden* (port. *ordem*), signo polissêmico. Traça gráficos da marcha de significações na semântica evolutiva dessa palavra, do étimo original latino ao espanhol moderno. Casares também trata do problema da lematização, ou seja, a decisão técnica de escolher como entrada de um dicionário, uma ou outra forma vocabular, o que envolve controvérsias permanentes em meio aos lexicólogos sobre as *lexias* (palavras) complexas e como e quando se dá a categorização lexical de um polinômio vocabular. Esse problema é ampliado por causa da tradição caótica de muitas grafias, particularmente no caso de "locuções vocabulares". Advoga as vantagens e as virtudes de um dicionário que tivesse um índice de frequência do uso de cada palavra, ou de cada acepção de um vocábulo.*

UNITERMOS: Semântica; lexicografia; significado; onomasiologia; semasiologia; etimologia; acepção; ordenação das acepções; polissemia; significado figurado; verbete; étimo; frequência do uso; combinação binária; polinômio verbal; tratamento das locuções.

1. Notas diferenciais

Se é verdade o que dizíamos anteriormente¹ sobre a dificuldade de traçar linhas divisórias entre as várias disciplinas que se integram na Lingüística, parece pacífico que essa dificuldade aumenta ao tentarmos circunscrever em campos separados a lexicografia e a semântica. A semântica reivindica como matéria-prima dos seus estudos os dados que a lexicografia recolhe e ordena, e esta, por sua vez, não poderia interpretar nem valorar acertadamente esses dados, se não conhecesse as relações que entre eles vai descobrindo a semântica e as leis que conseguiu formular para explicar os processos evolutivos observados. Temos de tentar, contudo, estabelecer alguma distinção para evitar que o lexicógrafo se perca em elucubrações alheias à sua incumbência específica.

Se considerarmos na palavra, de um lado os elementos fônicos de que consta, os chamados fonemas e, de outro, os elementos formais, que passam ao domínio da gra-

* Lexicógrafo espanhol. Produziu sua obra na década de 40 e 50. É autor do *Diccionario Ideologico de la Lengua Española* e da *Introducción a la Lexicografía Moderna*, de que foram extraídos os capítulos aqui inseridos, publicados, originalmente, no anexo da *Revista de Filología Española* em 1950.

** Departamento de Letras Modernas — Instituto de Letras, História e Psicologia — UNESP — 1980 — Assis — SP.
1 N.T. O autor se reporta à passagem anterior do seu livro de que este artigo é um capítulo.

mática com a denominação de *morfemas*, ainda nos fica, pelo menos, um terceiro componente da palavra, que é o conceptual ou significante, chamado por alguns tratadistas modernos *semantema*. Pois bem, dentro da atual especialização dos estudos linguísticos, há um ramo que se consagra particularmente ao estudo dos semantemas: a semiologia ou *semântica*, que o Dicionário acadêmico² define como segue: “Estudo da significação das palavras”³. Ora, se compararmos esta definição com o conteúdo que o mesmo Dicionário atribui à Lexicografia, enquanto “arte de coleccionar todas as palavras de um idioma e descobrir e fixar o sentido e o emprego de cada uma delas”, veremos que, entre “estudar a significação das palavras” e “descobrir e fixar o seu sentido”, não se percebe senão uma ligeira diferença. Será, pois, conveniente, para torná-la palpável, aprofundar mais os conceitos respectivos.

A primeira distinção que surge, partindo das definições citadas, é a seguinte: a lexicografia é uma “arte”; a semântica, tal como a formulou Bréal, inventor desta denominação, é uma “ciência”. Os problemas que hoje são objeto desta ciência e que já se tinham colocado antes, de outra forma ou agrupados sob outros nomes, foram evoluindo, desde princípios deste século, para uma complexidade cada vez maior. Já não se trata simplesmente de descobrir, cotejar e classificar os fenômenos que se observam nas mudanças de significação, com o fim de buscar as causas que as motivaram e, inferir de tudo isso certas leis; pretende-se, porém, atacar o problema primário, de notório caráter filosófico, referente às relações entre o signo e o significado, entre a palavra e a sua correspondência na mente. E já nesta altura mais propriamente poderíamos dizer “nesta profundidade” — é necessário indagar o que é a palavra e qual é a sua representação correlativa no espírito do indivíduo.

2. A palavra e o “sintagma”.

Para não repetir aqui o que já foi tratado por extenso em outro lugar⁴, bastará recordar que os mais insignes gramáticos não conseguiram jamais pôr-se de acordo sobre uma definição do conceito de *palavra*. As teorias mais modernas sustentam que a palavra é um corte arbitrário praticado na frase, e consideram-na como a última unidade indivisível do sistema expressivo que chamamos linguagem. Assim que fragmentarmos uma oração para dela extrair uma palavra que nela estava inserida, a fim de colocar esse vocábulo na mesa de dissecação, comprovamos que lhe falta vitalidade e que se converteu numa peça anatômica tão inerte como uma víscera separada do corpo. É claro que podemos transplantar essa palavra para outra frase onde voltará a cobrar vida, mas nunca será a mesma vida em cada uma das combinações em que entre. Não se trata só do caso em que um vocábulo tem diversas acepções bem delimitadas. A palavra *direção* se usa com significados diferentes em cada uma das seguintes frases: “quando o carro tomou uma curva, quebrou a *direção*”; “Ocorreu um incêndio na *Direção* da Segurança”; “os pinhões da engrenagem giram em *direção* oposta”; “o chefe mais antigo tomou (para si) o cargo da *direção* da empresa” etc. Exemplos deste teor podem multiplicar-se até a saciedade; mas não ilustram convenientemente a tese. Interessar-nos-ia, pois, procurar um caso em que a palavra transplantada tivesse uma só acepção

2 N.T. Trata-se do *Dicionário de la Lengua Española*, da Real Academia Española, editado pela Espasa-Calpe, S.A. com várias edições (a 19a. é de 1983). Há ainda o *Diccionario Manual e Ilustrado de la Lengua Española*, também da Real Academia Española, que é, ao mesmo tempo, um resumo e um suplemento do anterior. Faz-se necessário lembrar que estes dois Dicionários, sendo o resultado das pesquisas constantes dos académicos espanhóis, gozam de prestígio absoluto. Todas as vezes que aparecer o termo *Diccionario*, sem especificações, o autor refere-se ao *Dicc. de la Lengua Española*.

3 N.T. Note-se que isto se escrevia antes de 1950.

4 V. meu *Nuevo concepto del Dicionario...*, pág. 218.

e, na falta de outro exemplo melhor, vamos tentar a demonstração com o vocábulo *mar*. Comparemos, pois, o valor diferente deste vocábulo, na boca de um mesmo falante, segundo ele diga: “Prefiro o peixe de *mar* ao de rio”; ou “há muito *mar* para se pescar”; “vive num *mar* de rosas”; ou “sua vida é um *mar* de paixões”; ou ainda “de *mar a mar*”. Na primeira oração, “*mar*” é só o meio em que vivem os peixes, um ambiente de água salgada, sem dimensões, contornos ou movimentos; na segunda, *há muito mar*, predomina a idéia de uma massa líquida agitada com dinamismo perigoso; na terceira, *mar de rosas*, não há água nem movimento, nem cor, nem nada que não seja uma noção abstrata de *período de felicidade*; na quarta *mar de paixões*, a idéia preponderante é a de abismo (moral); finalmente, a última, de *mar a mar* só evoca na mente a expressão de *ponta a ponta*. Não se poderia dizer que a palavra *mar* assume, por si só, tão variadas significações, nem cabe pensar que tais significações residam numa ou noutra das palavras que acompanham *mar*. Os últimos três exemplos pouco ou nada têm a ver com o significado de *mar* (massa de águas salgadas), constituem, porém, expressões idiomáticas. A estes conglomerados, onde as palavras se fundem num bloco significante e se influenciam reciprocamente, daremos o nome de “sintagmas”.

Não esqueçamos que este tecnicismo *sintagma*, além do significado exposto, tem outros na lingüística moderna; e são, por certo, tantos e tão contraditórios que quase o tornam imprestável. Diremos somente, a título de exemplo, que quando Von Pirquet, para denominar certos fenômenos semelhantes à anafilaxia, inventou a palavra *alergia*; foi esta tão bem recebida pelos técnicos que até se difundiu na linguagem usual e reclamou um lugar no dicionário de tipo geral. Ao cabo de trinta e cinco anos “*alergia*” chegou a significar tantas coisas que começa a resultar inservível. Tanto é assim que o Dr. Marañón nos dizia recentemente que a entrada dessa palavra no Dicionário acadêmico já não se justificava. Conste, pois, para evitar equívocos, que o termo “*sintagma*” só terá, a seguir, o significado que acabamos de dar-lhe.

3. Flutuação das relações entre signo e significado.

Vemos, pois, que no mecanismo da linguagem não existe uma correlação e-s-t-á-v-e-l e de termos unívocos entre signo e significado, isto é, entre a palavra falada ou a escrita e o conceito que essa palavra evoca. Essa correlação, que se dá no signo natural (a fumaça em relação ao fogo) e no signo convencional (a cor vermelha como símbolo do perigo), não se dá, em troca, no signo verbal, salvo em casos excepcionais, como *binômio*, *hipófise* ou *gasolina*; de onde resulta que a palavra, isolada do sintagma ou da frase, se oferece como um núcleo de p-o-s-s-i-b-i-l-i-d-a-d-e-s s-i-g-n-i-f-i-c-a-n-t-e-s que, por certo, não são sempre as mesmas para todos os membros da respectiva comunidade lingüística. A representação mental que evoca, por exemplo, a palavra *marisma* (*mangue* em português) não será nunca igual para as pessoas de terra firme como para aquelas que vivem junto a um terreno desse tipo.

Uma vez evidenciada a instabilidade da relação e a variabilidade dos dois elementos que a formam, concluiu-se que, para estudar as mutações semânticas e investigar as causas a que obedecem, seria necessário considerar paralelamente à evolução do signo e à evolução do significado, analisando em cada caso a intervenção do indivíduo, consciente ou inconsciente, as relações afetivas ou emocionais que, transcendendo a comu-

5 N.T. As frases originais são as seguintes: “Prefiero el pescado de *maral* de rio”; “hay mucha *mar* para salir de pesca”; “me ha felicitado la *mar* de gente”, e “venia la novia hecha um brazo de *mar*”. Os exemplos “*mar de rosas*”; “*mar de paixões*” e “*de mar a mar*” aparecem no Dicionário Aurélio e foram adaptados pela tradutora.

nidade, dão lugar a valorações antonomásticas, pejorativas, depreciativas etc., e à ação de fatores sociais, culturais e até políticos, cuja concomitância é decisiva em determinados fenômenos. Um caso de influência política nas mudanças de denominação é a que se operou em nossos dias relativamente ao antigo e vernáculo *jornaleiro*, ou seja, a pessoa que faz um trabalho manual mediante um estipêndio diário. Apenas iniciado na Espanha o movimento socialista, o jornaleiro converteu-se em *operário*. A Constituição de 1931 definiu-o como *trabalhador* e o regime triunfante na guerra civil em 1936 proscreeu todos os nomes anteriores para deixar como único vigente o de *produtor*. Compare-se noutra ordem de idéias a série de eufemismos com que sucessivamente se designa uma coisa ignóbil, anulando cada termo o anterior tão logo este comece a solidarizar-se demasiadamente com a coisa representada⁶. Outras vezes é a própria coisa que, sem mudar de função, se transforma materialmente, como ocorreu com *pena*, desde a originária da ave, que era verdadeiramente uma “pena”, até a caneta-tinteiro mais recente, e a chamada “esferográfica”, que nem ao menos conserva a peninha metálica⁷.

Em consequência destas novas orientações, a semântica diversifica-se e vão aparecendo importantes divisões novas, como a que primeiro se chamou “lexicografia comparada” (Tappolet, 1895) e depois tomou o nome de “onomasiologia” (Zauner, 1903). Esta disciplina, partindo de uma coisa determinada, de um objeto ou de uma noção, propõe-se estudar comparativamente os caminhos que essa coisa seguiu até corporizar-se numa palavra, e pretende reconstruir o processo intelectual e imaginativo que determinou tal corporificação. Com propósito semelhante tornou-se comum o uso de grupos homogêneos de noções (partes do corpo, relações de parentesco, hierarquias militares, nomes de números etc.), supondo-se que, dentro de cada grupo, era de esperar que as mudanças de significação se produziriam com frequência e apresentariam certos caracteres comuns (hipótese dos campos semânticos, *Bedeutungsfelder*). Por outro lado, a psicologia moderna, especialmente a psicologia coletiva e social (Wundt), aplicou-se ao estudo da semântica para pôr as conquistas desta ciência a serviço de determinadas teorias, abrindo deste modo aos filólogos múltiplos e fecundos horizontes...

Nada mais fizemos do que olhar de perto alguns ângulos do amplíssimo domínio que a semântica abrange atualmente; porém, bastará esta abordagem rápida para compreendermos a enorme complexidade dos problemas com que se foi enriquecendo esta ciência, que algumas vezes pretende manter-se apoiada na lingüística e, outras, interna-se pelo terreno da filosofia ou da história da cultura, sem que faltem escolas híbridas nem tentativas mais ou menos engenhosas para reduzir a heterogeneidade dos fenômenos semânticos a uma só lei geral. E também bastarão as suscintas indicações, que precedem, para advertir o lexicógrafo do perigo a que se exporia se, em vez de concentrar-se em comprovar e relacionar as mudanças de significado com que se depara, se excedesse em inquirir as causas de tais mudanças partindo desses postulados que pomposamente se chamam “leis”, tais como a de especialização, a de repartição, a de irradiação, a de diferenciação, a de eliminação do supérfluo, a do desgaste e tantas outras que só alcançam efêmera vigência. Tudo isto significa que, quando o lexicógrafo se defronta com os fenômenos semânticos não deve ir além de determinar as diferentes acepções que de fato concorrem no vocábulo e de estabelecer, quando for possível, a relação em que estão umas com as outras. E já é suficiente.

⁶ *Divertimientos filológicos*, pág. 297.

⁷ N.T. Em espanhol denominam-se todos esses objetos com a palavra “pluma” = pena.

4. Reconstituição de significados para relacioná-los com a etimologia.

Já é suficiente, pois esse trabalho deverá começar muitas vezes dentro dos parênteses que, nos verbetes dos dicionários, estão reservados para a etimologia; em outras palavras, o lexicógrafo terá de realizar a tarefa etimológica antes de proceder ao estudo e coordenação das acepções que o étimo tenha desenvolvido numa língua determinada. E aqui tem outra vez aplicação a advertência de Menéndez Pidal, que já recordamos anteriormente: “Quando a relação entre a etimologia e o termo derivado é obscura por se terem perdido acepções intermediárias, é necessário precisar essas acepções”.

Vejamos um exemplo concreto. No linguajar familiar dá-se o nome de *pagão*⁸ ao que paga dinheiro, culpas ou outra coisa, quando era outro quem normalmente deveria pagar. Se consultarmos o Dicionário acadêmico observaremos que atribui ao vocábulo *pagão* três acepções: a primeira aplica-se aos idólatras e politeístas; a segunda, por extensão, a todo infiel não batizado; e a terceira é a familiar que já mencionamos ao princípio. Se quisermos saber a origem desta (acepção), o Dicionário limita-se a dizer-nos que deriva do latim *paganus*, sem acrescentar o que isto significa, com o qual se nos induziu em erro. O leitor leigo sente que esse “pagão”, que paga, guarda relação com o verbo “pagar” como “escrivão” com “escrever”, porém estaria enganado se acreditasse que se trata de uma derivação normal, porque na língua espanhola as formas em *-ano* não derivam diretamente dos verbos, mas dos nomes, adjetivos ou advérbios.

O “*escribano*” do espanhol supõe um *escribanus* do latim. A esta altura o leitor continua perguntando-se o que significa esse *paganus* da etimologia. E se lhe dizem que equivale a “camponês” ou habitante de um “pago” rural, ainda assim não consegue explicar de onde saem os idólatras e politeístas ou os infiéis não-batizados. Convém, pois, que se explique que, quando o Cristianismo se apropriou das cidades, a idolatria refugiou-se durante algum tempo nas aldeias, e que por esta razão o nome de aldeão (*pagão*) resultou equivalente a *não-cristão*. Outro exemplo moderno acabará de ilustrar este ponto. Se num dicionário estrangeiro se define a expressão *quinta-coluna*, que já goza de difusão universal, e se acrescenta como único esclarecimento que procede do espanhol, o leitor não aprenderá nada. Por que “quinta” ou “sexta”? É preciso fazer referência ao momento histórico em que, organizadas militarmente “quatro colunas” para a tomada de Madri no ano de 1936, designou-se com o nome de “quinta” a que, dentro da capital, poderia formar-se com os partidários do exército atacante. Daí que se chame quinta-coluna ao conjunto de cidadãos de qualquer país que, dentro dele, favorecem desígnios estrangeiros. Poderíamos dizer que tudo isto é semântica, e não serei eu quem o contradiga; o que me interessa advertir é que se trata de uma semântica aplicada, ou seja, concebida como meio e não como fim de uma semântica posta egoisticamente a serviço do lexicógrafo. E esta é a que nos vai servir para orientar-nos no dedalo de problemas que nos esperam tão logo transponhmos esse parêntese em que se encerra a etimologia.

5. Conceito de “acepção”.

Dissemos que a palavra isolada, que é como se apresenta ao lexicógrafo, é um nú-

⁸ N.T. Cf. Dicionário Aurélio: *Pagão* (de pagar) = *Pagante*. (Gíria). No *Grande Dicionário da Língua Portuguesa* de Antonio Moraes da Silva aparece: “*Pagão*”, *adj. e s.m. (de pagar)*. *Que ou que paga; a pessoa que corre com as despesas*. Cf. *Antenor Nascentes, A Gíria Brasileira*”. Aqui, encontramos: “*Pagão*, s.m. A pessoa que corre com as despesas: O oferecimento é feito pelo pagão, cristo, sofredor, isto é, a pessoa que faz a despesa. (J. Calasans, *Cachaça, moça-branca*, 17)”.

cleo de possibilidades significantes quase ilimitadas; há algumas, porém, que se realizam com freqüência e se repetem com o mesmo valor em frases ou combinações diversas, até o ponto de se individualizarem. A palavra *ordem*, contém como substrato básico conceptual a idéia de uma “certa disposição de várias coisas no espaço”; v.gr.; “pôr em ordem os livros de uma biblioteca”. Pois bem, essa “certa disposição de várias coisas”, considerada em relação aos elementos arquitetônicos e ornamentais de um edifício, foi a causa de que uma das possibilidades significantes do vocábulo “ordem” se tenha realizado e consolidado com uma equivalência unívoca, como se vê nas combinações “ordem coríntia”, “ordem dórica”, “ordem composta” etc. Já não se trata de uma ordem qualquer, mas de um conceito específico que enriqueceu a palavra *ordem* com um significado novo. Essa mesma palavra, tornamos a encontrá-la em outra série, própria desta vez da terminologia militar: “ordem aberta”, “ordem fechada”, “ordem de parada”, “ordem de batalha”, e já temos outro sentido especial do vocábulo.

Cada um destes sentidos especiais ou gerais é o que em lexicografia constitui uma *acepção*.

Se as quatro ou cinco mil palavras usadas por uma pessoa culta não tivessem que expressar mais que outras tantas idéias, não seria necessário que os vocábulos fossem plurivalentes. A *polissemia*, isto é, a concorrência de significados diversos num mesmo signo verbal, é consequência inevitável da desproporção que existe entre o número destes signos e a enorme quantidade de noções que buscam expressão na linguagem. O lexicógrafo, portanto, deve ter presente sempre a possibilidade de que o vocábulo submetido à sua análise se decomponha em várias acepções, que requerem tratamento especial, e que serão muitas ou poucas conforme o vocábulo de que se trate e conforme também o critério subjetivo do lexicógrafo.

6. Dificuldades que apresenta a separação de acepções.

Se examinarmos em vários dicionários modernos um verbete cujo conteúdo conceptual é substancialmente o mesmo nas principais línguas cultas, observaremos notáveis diferenças no tocante à divisão da matéria. O verbete *centro*, por exemplo, descontados os americanismos, tem no Dicionário acadêmico onze acepções; no Dicionário Histórico abandonado, dezoito; no Dicionário de Oxford⁹ vemos dezoito seções principais que, com as subdivisões internas, dão lugar a vinte e cinco itens, enquanto o *Dictionnaire Général*¹⁰ se contenta com quatro. A palavra lugar (fr. *lieu*) tem vinte e cinco parágrafos no grande *Dicionário de Littré*¹¹, e 172 (!!) na 5.^a edição do *Vocabulário da Crusca*¹² (it. *luogo*). Grande é a discrepância; e embora se tenha que atribuí-la, em boa parte, à diversidade de propósito e à diferente extensão das obras, é ainda mais importante a diferença de critério adotado, porque num caso se analisa até a sutileza os diferentes matizes conceptuais e, no outro, separam-se apenas as concepções que razoavelmente não poderiam fundir-se.

Qual destes critérios extremos será preferível é coisa que não se pode decidir com argumentos e que depende em grande medida do material com que se trabalha e da concepção pessoal dos redatores. Desde logo é preciso convir em que a bifurcação em

9 N.T. refere-se ao *The Oxford English Dictionary* (1.^a ed.: 1884-1928).

10 N.T. Trata-se do *Dictionnaire Général de la Langue Française*. Du commencement du XVII^e e. siècle jusqu'a nos jours. Paris, Librairie Ch. Delagrave, 1905. A impressão foi iniciada em novembro de 1888). 2 vol.

11 N.T. Refere-se ao *Dictionnaire de la Langue Française*, Paris, Hachette, 1863-1872, 5 vols.

12 N.T. A primeira edição do “*Vocabulário della Crusca*” é de 1612 em 6 tomos. Trata-se do mais antigo dicionário do italiano, elaborado pela *Accademia della Crusca*, Florença.

itens, subitens e infra-itens, levada até às últimas conseqüências, prejudica notavelmente a perspectiva de conjunto, embora contribua, por outra parte, para explicar a genealogia de cada uma das acepções. A excessiva condensação, pelo contrário, tem o inconveniente, sobretudo para um dicionário com citações, de obrigar a prescindir de muitas delas, às vezes preciosas, sob pena de juntá-las indistintamente em detrimento de sua eficácia ilustrativa, além de que não permite observar a fase em que se acha o processo de especialização das acepções recentes. Não é factível, pois, estabelecer uma regra aplicável às múltiplas contingências previsíveis razão pela qual devemos dar-nos por satisfeitos se encontrarmos algumas fórmulas empíricas que nos assinalem uma orientação geral. E vamos pedi-las à semântica.

A discussão empreendida por Steinthal e Hermann Paul sobre a existência ou inexistência de palavras com mais de um significado carece praticamente de interesse para o redator de um Dicionário, quando se trata de línguas modernas. Remontando-se mais atrás é possível conjecturar que, num momento determinado, o latim *arcus* só significasse o arco para lançar flechas, e que a curvatura da vara flexível retesada sugerisse a idéia de chamar do mesmo modo a construção arquitetônica de figura curva (arco de ponte, de aqueduto etc.); mas o lexicógrafo que trabalha com uma língua neolatina se encontra com o étimo de duplo significado e não pode prescindir deste fato ao fazer sua composição de lugar. No caso de *musculus* (dim. de *mus*, “rato”) a primeira idéia pode ter sido a do animal, transposta logo às massas carnosas do corpo vivo que se movem sob a pele. Todavia, tudo nos induz a pensar que o latim *lacertus* se aplicou primitivamente a determinados músculos do braço e que a ampliação de significado para designar o “lagarto” operou-se no sentido inverso à proposta para *musculus*, ou seja, do músculo ao réptil. O diminutivo *musculus* não deixou na sua descendência castelhana nenhum indício do “rato”; mas o positivo *mus (-ris)* deu-nos o diminutivo “*murecillo*” em que convivem a acepção de “rato pequeno” e a de “músculo” (esse “*murecillo*” é o atual “*morcillo*” com que as cozinheiras e os açougueiros designam certa parte das reses). Nos casos de *arcus*, *lacertus*, *mus* e tantos outros, a semântica histórica pode colocar-se o problema de indagar qual dos significados que conviveram em cada palavra foi o originário e por que caminho deu este nascimento ao outro. O redator de um dicionário castelhano só tem que se perguntar se passaram a esta língua ambos os significados para relacioná-los com o étimo e proceder em conseqüência.

A solução não é tão expedita quando o desdobramento do significado se produz na língua herdada. As dificuldades crescem à medida que se aproxima de nós a data em que o fenômeno se manifesta, e chegam ao máximo cada vez que nos achamos na presença de um processo inacabado. Quando se construiu o primeiro utensílio para obter nos instrumentos de corda, o som por fricção, esse utensílio tinha exatamente a mesma forma que os arcos de caça, salvo que em lugar da corda ou do nervo destes levava um maço de crinas. Nada mais natural que dar a esse utensílio o nome de “arco”. Tratava-se de uma ampliação ocasional de significado que, dada a propagação do objeto, sua permanência no uso e sua mudança ulterior de forma, adquiriu individualidade suficiente para constituir uma acepção. (Atualmente esse utensílio é uma varinha reforçada com uma ligeira curvatura invertida em relação ao arco de caça, isto é, com a cavidade para fora.) Este emprego “ocasional” de uma palavra, que se converte logo em emprego “usual”, é um dos procedimentos mais freqüentes para a multiplicação de acepções.

Outro processo não menos fecundo consiste na passagem de uma noção abstrata a uma concreta, entendidos estes dois termos de uma maneira diferente daquela que aparece na terminologia gramatical. Os nomes de substâncias, como “água”, são gramati-

calmente concretos. Entretanto do ponto de vista em que agora nos colocamos, “água” aparece como uma abstração de algo que tem em comum coisas tão diversas como a água salgada dos mares, a doce do rio, a destilada, as águas sulfurosas, as termais etc. Se dissermos, contudo, “água de Solares” ficam excluídas todas as demais águas possíveis: individualizamos a coisa, nós a tornamos “concreta”. A palavra *gás* inventou-se expressamente, como se sabe, para denominar certos vapores e designou logo depois toda classe de fluidos aeriformes: porém, com o passar do tempo, ocorreu que um destes fluidos chegou a ter tal importância na vida doméstica e urbana, que se converteu no *gás* por antonomásia. Hoje, quando se fala de “restrições no *gás*”, de “um fogareiro a *gás*” ou da “Companhia de *gás*” todos entendemos que se trata de certo hidrocarbureto que se obtém destilando o carvão de pedra. Esta restrição do significado, esta condensação do conceito geral numa só das infinitas espécies nele contidas, dá origem a acepções independentes que o lexicógrafo deverá registrar separadamente.

E chegamos ao ponto mais difícil. A transição do emprego ocasional ao emprego usual, ou seja, desde que surge a inovação deliberada ou inconsciente que modifica o sentido de uma palavra, até que prospere a compreensão e a aceitação dessa mudança por parte da comunidade, não é sempre um processo tão simples e tão claro como no caso do “arco” do violino. A consciência de uma associação de idéias repentina ou de um estado de ânimo emotivo que pede uma intensidade de expressão desacomumada, ou simplesmente para buscar um efeito cômico ou pelo desejo de distingüir-se, o falante despreza a palavra corrente que lhe vem aos lábios e, lançando mão de outra, substitui aquela por esta. Se a troca dá o resultado desejado, e quem a ouve a repete por sua conta com êxito semelhante, rapidamente a palavra em questão se terá enriquecido com uma nova acepção.

Às vezes, o caminho percorrido desde o significado inicial até o derivado coloca os dois em situação diametralmente oposta. Disto oferecem exemplos abundantes os qualificativos com os quais se pretende expressar excelência ou admiração em grau superlativo. Todos eles, até os mais hiperbólicos desgastam-se em pouco tempo. Assim, quando *soberbo*, *magnífico*, *portentoso*, *prodigioso*, *maravilhoso* etc., começam a parecer formas pálidas e mornas em relação à veemência do que fala, surgem *estupendo*, *bárbaro* ou “bestial”: “Fulano possui um talento “bestial”. A expressão não pode ser mais disparatada, já que a bestialidade é o mais contrário possível a talento e contudo, todos interpretamos essa frase exatamente no sentido em que a propôs o seu autor. Chegará a tomar *bestial* o significado de “excelente”, “insuperável em sua linha”? Estamos diante de um processo inacabado e seria anticientífico predizer o seu fracasso, uma vez que outros processos não menos absurdos se consumaram e consolidaram. *Formidável*, etimologicamente, é o que “causa espanto”, o que “inspira terror”. Os monstros do Apocalipse são propriamente *formidáveis*. Porém, se um espectador, subjugado, atraído ou possuído pela arte ou pela beleza de uma estrela de cinema, ou por ambas as coisas, qualifica-a de “formidável” sabemos bem que estão muito longe de sua imaginação os sentimentos deprimentes que acompanham o conceito de medo. A nova significação não só se emancipou da antiga mas também, até certo ponto, a eclipsou. A passagem tinha desta vez a seu favor uma acepção recolhida no Dicionário, que diz assim: “Excessivamente grande na sua linha,” A noção de grandiosidade, de algo extraordinário, já tinha vencido a noção de temor.

É curioso observar a este respeito que o caso de *formidável* não parece, pelo menos em espanhol, um fenômeno isolado e casual, mas a conseqüência de uma polaridade mental com algumas raízes psicológicas; porque tão logo como *formidável* relegou à penumbra o seu sentido de “terrorífico”, surgiu na linguagem familiar uma locução

que reproduz com todo o seu vigor esse sentido para expressar a admiração ilimitada. — Fulana é tão linda como dizem’?', pergunta um, e o outro responde: — “De medo”¹³. Nos dias de hoje é impossível achar em castelhano uma fórmula expressiva que supere este encômio.

Numa outra ordem de idéias, compare-se a evolução de “miserável” (do latim miserabilis), “o que é digno de compaixão”, até chegar a “perverso”, “abjeto”, “canalha” (Dicionário da Academia).

Se tivéssemos que assinalar todos ou, pelo menos, os principais caminhos por onde as palavras adquirem significados novos, necessitaríamos muito mais tempo que o previsto para o presente trabalho. Remetemos aos tratados de semântica; e só mencionaremos, para terminar esta exposição prévia, o caso em que o determinante absorve o conteúdo do determinado e o elimina como coisa supérflua.

Na expressão “navio encouraçado”, o determinante, que aportava tão-somente a noção de “couraça” absorveu logo a de “navio”, e quando hoje se fala de um *encouraçado* entende-se, sem mais, que se trata de um navio de guerra de armamento especial, uma de cujas características, que comparte com outros navios, é a sua forte blindagem. O mesmo aconteceu com *vapor*, que já não é somente o fluido aeriforme cuja pressão se utiliza como força motriz, mas também o “barco” movido desta maneira.

Estes dois casos são bem conhecidos porque se produziram quase ao mesmo tempo em várias línguas modernas; porém, não consta que fora da Espanha a palavra *vapor* tenha tomado a significação de “ferrovia”. Antes de que se começasse a usar este último vocábulo dizia-se “caminho de ferro (calco do francês *chemin de fer*) de vapor”; mas como tal denominação resultava imprópria e muito longa abreviou-se para *vapor*. Por ocasião do centenário da nossa primeira estrada de ferro¹⁴, que se comemora nestes dias, o musicólogo José Subirá ofereceu-nos numa interessante conferência várias partes de obras musicais, contemporâneas da introdução desse meio de transporte, entre as quais figura uma canção intitulada “Rede de Amores”, música do maestro Espín Guillén e letra de R. Larrañaga, onde figura a seguinte estrofe:

Yo soy mosita muy fiera
y tan neta en el querer,
que tras mi chairo corriera
más que el *vapor* de Aranjuez¹⁵.

E no diário madrilenho *La Nación*, pela mesma data, dava-se a notícia do atropelamento de uma pessoa pelo trem, sob a epígrafe: “*Vítima do Vapor*”.

O interessante deste caso é a rapidez com que se operou a condensação de “ferrovia a vapor” em *vapor*, enquanto que a de “navio a vapor” não se registra senão até quase um século depois de tais navios estarem à vista de todo mundo nos portos.

7. Possibilidade de orientação através de algumas normas empíricas.

Aprofundarmo-nos nestas considerações seria o mesmo que entrar de cheio no

13 N.T. Em português diríamos: — “É linda de morrer”.

14 N.T. A primeira estrada de ferro espanhola, ligando Madrid a Barcelona, foi inaugurada no dia 24 de outubro de 1848.

A presente obra foi publicada em 1950; deve ter sido escrita um ou dois anos antes, ou seja, na época do centenário.

15 N.T. Eu sou “mocinha” muito fogosa/e tão firme no meu amor./que atrás de meu “perfume” correria/más que o trem de Aranjuez. As formas “mocinha” e “perfume” são hipotéticas, já que os originais “mosita” e “chairo” não aparecem nos dicionários.

domínio da semântica especulativa quando o que procuramos são pontos de referência para o mister lexicográfico. Resumiremos, pois, os ensinamentos deste capítulo nas seguintes recomendações:

- 1.º) Quando um vocábulo teve na língua de origem duas ou mais acepções e estas, com maior ou menor vitalidade, passaram para a língua herdeira — no nosso caso a espanhola - deverão anotar-se nesta as várias acepções com a conveniente separação. E isto não só pelas razões antes apontadas mas também pensando na respectiva progênie de cada um dos significados. Em espanhol o verbo *ampollar* (port. *empolar*) (levantar ampolas na pele) procede de *ampolla* (port. *empola*) = bexiga, bolha; *ampolleta* (port. *ampulheta*) (relógio de areia) deriva-se da *ampolla* (port. *ampola*) = vasilha; *ampuloso* (port. *empolado*) (usado por Valdés no *Diálogo de la lengua*) nasce provavelmente através do francês *ampoulé de ampulla* (port. *empola*) = expressão inchada.
- 2.º) Quando um vocábulo de grande extensão lógica dá lugar a um significado de extensão mínima e de grande compreensão, tal significado deve ser definido como acepção independente. Sob a denominação genérica de *hierba* ou *yerba* (*erva*) entram espécies de número incontáveis; porém na América do Sul a “*hierba*” (pronunciado “*yerba*”) (= erva), é uma espécie única: a erva mate ou chá do Paraguai. (A este exemplo pode-se acrescentar o de “*gás*”, utilizado mais acima).
- 3.º) Quando um emprego *ocasional* passou a ser *usual* deverá registrar-se tal emprego como acepção à parte. Uma “*perra*” (cachorra) foi durante longo tempo uma moeda de níquel (de 5 ou 10 centavos), que devia esta denominação familiar à interpretação burlesca do leão rampante que aparecia no anverso. Esta acepção de “*perra*” (cachorra, cadela) está a caminho de extinguir-se por ter desaparecido da circulação a moeda de cobre a que dava nome; porém, viveu longo tempo, teve derivações como “*perrona*” (“*perra* grande”) e figura no Dicionário desde 1914. No ano de 1925 cunharam-se pela primeira vez moedas de uma mistura de cobre e níquel que valiam 25 centavos e ostentavam no anverso uma caravela. Não se sabia como designá-las até que alguém sugeriu — reproduzindo o processo das “*perras*” — que se chamassem *caravelas*¹⁶. Logo deixaram de cunhar-se tais moedas e o lexicógrafo deve perguntar-se hoje se o emprego “*ocasional*” de *caravela* chegou a ter estado literário ou se alcançou difusão suficiente para merecer a consideração de “*usual*”.
- 4.º) Uma questão particularmente delicada é a que nós coloca o uso figurado dos vocábulos. Examinemos o texto seguinte: “A obra de Spengler sugere todo um universo de idéias”. Esse *universo* tem o mesmo sentido metafórico de muitos outros termos com que normalmente expressamos e encarecemos a idéia de “abundância”; por ex: “um *mundo* de coisas”, “um *dilúvio* de petições”, “um *enxame* de pretendentes”, etc. Posto que “*mundo*”, “*dilúvio*” e “*enxame*” têm já no Dicionário outras tantas acepções especiais com a equivalência comum de “*multidão* de pessoas ou coisas”, poderemos aplicar a *universo* a sua? E se nos parece que não tem direito à ela, em que fundamentaremos a negativa? A única acepção de “*universo*” (como substantivo) que figura no Dicionário é uma simples remissão a “*mundo*”, o que estabelece uma perfeita igualdade entre esses

16 Se não me engano foi Wenceslao Fernández Flórez quem propôs, humoristicamente, a denominação de *caravela*, que, em pouco tempo, degenerou em *caravana* entre as classes populares. N.T. No Brasil é comum chamar-se a nota de Cr\$ 1.000,00 de *barão* porque reproduz a imagem do Barão de Rio Branco no seu anverso.

conceitos. Por este lado, portanto, não se vê uma distinção que justifique tratamento diferente; porém, creio que a encontraremos se nos permitirmos utilizar como pedra de toque um postulado concebido assim: *Quando a correta compreensão de um vocábulo empregado com valor metafórico, exige que o ouvinte se refira ao significado correto para voltar deste ao figurado, deve inferir-se que se trata de um emprego ocasional do vocábulo e não de uma nova acepção.* Ou seja, que se para entendermos corretamente a expressão “um enxame de pretendentes” começamos por evocar a imagem de umas abelhas amontoadas para tirarmos daí a idéia de “multidão”, este rodeio constitui a prova de que “enxame” não tem na realidade mais do que uma acepção, e vice-versa. Repetindo a experiência com as palavras “universo” e “mundo” observaremos que esta última se emancipou, faz tempo, do veículo que a ligava ao conceito fundamental de “cosmos”, em conseqüência, entre outras coisas, do seu emprego freqüente em sintagmas tais como “meio mundo sabia isso”, “oferecem-lhe este mundo e o outro”, “todo mundo está contra ele”, etc. *Universo*, pelo contrário, é palavra erudita e, portanto, pouco propícia para formar parte de modismos e frases feitas. Por isso, o seu emprego esporádico em sentido figurado é, hoje em dia, uma simples figura de retórica e não constitui acepção especial.

- 5.º) Nos casos de enriquecimento do determinante a expensas do determinado (*navio encouraçado* = *encouraçado*, subst.) importa comprovar se o processo semântico está efetivamente consumado, já que, a não ser assim, a nova acepção resultaria falsa. No jogo de bisca pratica-se uma variedade caracterizada porque nela é obrigatório seguir a carta jogada. O lance “arrasta” e os outros devem seguir o naipe ou pôr trunfo. Daí que esse tipo de bisca se chame “arrastada”. Outra variedade deste jogo é a bisca “sueca”¹⁷, que se joga entre quatro parceiros distribuindo-se dez cartas a cada um. Ora, o adjetivo “sueca” conseguiu substantivar-se, incorporando o significado de bisca, e atualmente, não só na conversa familiar, mas até nos textos literários, fala-se de “jogar uma partida de *sueca*”, porém não sucede o mesmo com *arrastada*, apesar de que esta classe de bisca tem no uso uma vida muito mais longa do que a “sueca”. Um caso semelhante ocorre com a denominação dos trens. Não estamos menos familiarizados com o “tren de pasajeros” do que com o “tren de mercancías”¹⁸ e, contudo, dizemos “um mercancías”¹⁹ e não dizemos, “un pasajeros”. Por isso, falta no Dicionário espanhol uma acepção do artigo “mercancía”: “m.pl. Tren de mercancías”. A estas contradições aparentes não falta uma explicação razoável, mas renunciemos entrar nela para não nos alongarmos demais.

As precedentes recomendações que não são certamente as únicas que poderiam formular-se, não se devem tomar como normas dogmáticas que pretendem resolver “a priori” problemas tão heterogêneos e complicados, como os que apresenta na prática a divisão de um verbete em mais ou menos acepções. Somente podem servir em alguns casos para indicar um rumo aproximado com o fim de evitar que o coeficiente subjetivo, inevitável nestas operações, arraste o lexicógrafo para o escolho da condensação excessiva ou para o outro, não menos perigoso, do desmembramento exagerado.

17 N.T. Tipo descrito no Dicionário Cladas Aulete e que se adapta perfeitamente ao caso de enriquecimento do determinante a expensas do determinado, citado por J. Casares. No original aparece o “*o tute subastado*” (bisca leiloadada), cuja equivalência em português não foi encontrada.

18 N.T. Em português, “trem de carga”.

19 N.T. “Un mercancías” (do esp.) = um trem de carga (em port.)

8. Ordenação das acepções.

Feita esta ressalva, vamos supor que as diferentes acepções de um verbete já estejam convenientemente delimitadas e providas de sua definição correspondente. Já nos poderemos despedir da semântica? De modo nenhum; temos de recorrer a ela para que nos ajude a ordenar convenientemente essas acepções, que às vezes se aproximam de uma centena. Não vamos perder tempo explicando em que se diferenciam, teoricamente, os principais critérios aplicáveis à ordenação de acepções, critérios que se costumam chamar *empírico, genético, lógico e histórico*. Bastará, para o nosso propósito, comparar as vantagens e os inconvenientes dos dois métodos contrapostos que se vêm praticando sem exceção nos dicionários mais importantes da época moderna²⁰.

A Academia Espanhola ateve-se sempre à regra seguinte: “Em cada verbete vão colocadas por esta ordem as diversas acepções dos vocábulos: primeiro as de uso vulgar e corrente, depois as antiquadas, as familiares, as figuradas, as regionais e hispano-americanas, as de gíria e, por último, as técnicas.”²¹

Este não é o momento apropriado para descermos a pormenores, criticando certos extremos dessa ordenação ou assinalando o índice de frequência com que se infringiu o programa adotado. O que nos interessa é o sistema, e este fica bem caracterizado dizendo que nele se parte do mais conhecido, do mais atual dentro da língua comum, isto é, do “uso vulgar e corrente” para terminar com as falas particulares e com os significados específicos. Todos os Dicionários espanhóis modernos têm sido redigidos com igual critério, o qual também prevalece no Dicionário da Academia Francesa e na maioria dos estrangeiros. Este é o método que chamamos “empírico”.

Não se trata de uma antiquilha, de uma rotina que nos envergonha. Prova disso é que muitos lustros depois de ter-se posto em voga entre os lexicógrafos a idéia de considerar cada verbete do Dicionário como uma biografia do correspondente vocábulo, vemos surgir no monumental *New Standard Dictionary*, de Funk e Wagnalls (1913) esse método empírico valorizado e paladinamente defendido como segue:

“A média dos leitores vai buscar no Dicionário a significação *atual* e mais *corrente* de uma palavra²¹. O nosso propósito consistiu em proporcionar-lhes esta informação de modo fácil e seguro, sem permitir que se interpusesse nada entre a palavra e a sua significação mais conhecida e importante. Isto deixa de lado o método usual ou histórico seguido por outros lexicógrafos, que consiste em dar primeiramente a etimologia do vocábulo, depois a acepção mais próxima da origem, embora às vezes seja inusitada, e depois outras acepções, deixando para o último lugar a atual e mais corrente, *quando precisamente esta última acepção é a que mais geralmente se busca* (sublinhado no texto). O leitor médio fica, desta maneira, embaraçado e confuso. Pareceu-nos melhor não seguir determinado sistema, simplesmente porque seja lógica e filosoficamente correto, se isso contribui mais para dificultar que para facilitar a tarefa do consultor”.

9. Vantagens e inconvenientes do método usual.

Já temos aqui, frente a frente, as duas tendências antagônicas, que buscam a sua respectiva justificação: uma, em razão de utilidade prática; a outra, em consideração de ordem científica. O dicionário compilado pelo método empírico preocupa-se com a

²⁰ Na ordenação *analógica* ou por idéias tratei amplamente no *Nuevo concepto del Dicionario...* e no prólogo do meu *Diccionario Ideológico*, Barcelona, 1942.

²¹ *Introductory*, pág. XIII

imensa maioria dos leitores a quem pouco importa a origem e a evolução semântica do vocábulo cujo significado atual é o único que desejam conhecer. O dicionário redigido com base em princípios históricos preocupa-se, pelo contrário, com o reduzido grupo de pessoas cultas que, não contentes com averiguar o significado da palavra, querem saber também o como, o quando e o porquê. Um, o empírico, limita-se a informar, a responder o que se lhe pergunta: o seu ideal é antecipar-se, adivinhar o que provavelmente buscará o leitor e servi-lo em primeiro lugar sem rodeios, sem fadiga e sem perda de tempo. O outro, o histórico, supõe no leitor uma curiosidade intelectual que, convenientemente satisfeita, ampliará a sua cultura, despertará o seu interesse pelas vicissitudes da língua que fala e proporcionar-lhe-á um conhecimento da mesma mais profundo e satisfatório. Ambos os tipos de léxico têm, pois, dos seus respectivos pontos de vista, boas razões em que se apoiar. Ora, como a nós nos interessa muito mais o rendimento científico do trabalho lexicógrafo do que o cômodo manejo dos Dicionários, não podemos passar por alto os inconvenientes do método empírico.

O primeiro consiste na ausência de uma “estatística do uso” que nos permita determinar, entre vários significados *todos usuais*, qual é verdadeiramente o *mais usual*. Fazemos uma experiência com as quatro acepções de *assunto*, que figuram nos textos seguintes:

“O *assunto* que se discute é a baixa dos alugueis” (Assunto = tema de discussão).

“Tenho nas mãos, um *assunto* que pode dar dinheiro.” (Assunto = negócio lucrativo).

“O *assunto* do filme é o mesmo da novela.” (Assunto = argumento).

“João coleciona quadros de *assunto* religioso.” (Assunto = o que se representa na obra de arte).

Escrito cada texto numa ficha, pede-se a várias pessoas, sucessivamente e em separado, que ordenem as fichas começando por aquela em que *assunto* tem hoje o significado mais usual, para terminar com a que oferece o significado menos corrente. O resultado que eu obtive com oito indivíduos de nível intelectual satisfatório, embora de diversa formação cultural, é o seguinte: o “assunto = negócio” obteve invariavelmente o primeiro posto; o “assunto” de um quadro, o último; os outros dois textos deram lugar a flutuações. Pois bem, comparando este resultado com o Dicionário acadêmico, onde por definição o significado de uso vulgar e corrente deve ser o primeiro, veremos que a acepção favorecida no texto como o número 1 é precisamente a última no Dicionário. Que prova isto? Prova aquilo que já tínhamos dado a entender: que a ordenação de acepções pelo método empírico será arbitrária na maioria dos casos. Além do mais, o que hoje possa resultar acertado, não foi ontem e certamente deixará de sê-lo amanhã, posto que as várias acepções de cada termo mantêm entre si uma luta constante para alcançar a primazia. Esse *assunto* = negócio, elegido atualmente por unanimidade para o primeiro lugar, nem sequer aparece no léxico acadêmico até 1914; enquanto o *assunto* = tema, vem figurando à cabeça do verbete desde há mais de dois séculos. Um exemplo ainda mais peregrino da contradança que o uso impõe às acepções de uma palavra, nós o acharíamos examinando as peripécias de *alferes*, palavra que percorreu toda a escala militar, desde a mais alta hierarquia até a mais baixa (cf. Almirante, *Diccionario Militar*, s.v.). De tudo isso se deduz que se os dicionários empíricos tivessem que cumprir o que prometem, teriam que publicar edições frequentes, compassadas com as mudanças e caprichos do uso, refazendo em cada uma delas os verbetes para ordenar as acepções relativamente à maior ou menor frequência com que são empregadas em cada momento histórico; e porque isto não se faz nunca em realidade, mas

acrescenta-se a acepção moderna à continuação da antiga, é por isso que esse “assunto = negócio” figura em último lugar.

É evidente, portanto, que uma obra lexicográfica, concebida com critério científico e que aspire a ter validade para várias gerações, terá de procurar outra maneira de apresentar os materiais léxicos. E essa maneira, enquanto não se medite noutra melhor, é a que tem por fundamento os princípios históricos.

1. — *Aplicação do método histórico à ordenação de acepções — Dificuldades que surgem na prática.*

Já vimos o método histórico, definido por seus opositores, como o procedimento que “consiste em dar primeiramente a etimologia do vocábulo, depois a acepção mais próxima da origem, embora seja inusitada, e depois outras acepções, deixando para o último lugar a atual e mais corrente” (1.8). Desde a posição oposta, que é a dos partidários desse método, não temos inconvenientes em aceitar essa definição se descontarmos a malícia polêmica com que se pretende insinuar que a acepção atual e corrente é a mais afastada da etimologia, coisa que acontece umas vezes sim e outras não. Porém as características, as vantagens e as dificuldades do método histórico se precisarão claramente, melhor que com discussões teóricas, quando tratemos de aplicar esse método a um caso real, como nos propomos fazer mais adiante.

Começaremos pelas dificuldades. Uma imagem que poderia servir-nos para representar o processo das evoluções semânticas, desde o ponto de vista lexicográfico, é a de uma árvore invertida, isto é, com a raiz dirigida para cima, onde está a pré-história e com o tronco e os ramos para baixo. Uma linha horizontal, símbolo da superfície do solo-linha completamente inadmissível para o lingüista e que só tem valor para o lexicógrafo —, separa a raiz, que se submerge na língua de origem, do tronco e dos ramos, que se estendem na língua herdeira. Neste esquema, arbitrariamente simplista, só se considera uma raiz e um tronco, com mais ou menos ramos. Ora há casos, e são os mais freqüentes, em que a raiz também se bifurca e dá pelos lados brotos tão vigorosos que é difícil decidir muitas vezes, a qual deles conviria melhor a denominação de tronco. É claro que se nos fosse possível contemplar o organismo vegetal na sua integridade viva, o problema de precedência dos brotos não seria insolúvel; porém, quando o que se oferece ao nosso exame são partes incompletas da planta, resulta muito difícil determinar se se trata de ramos ou brotos, em que ordem se produziram e a que distância se encontravam da raiz.

Tal distância ou afastamento pode ser de duas classes: temporal e conceptual. Supondo que tivéssemos ante os olhos a série ininterrupta dos significados que foi assumindo um signo verbal desde que apareceu no espanhol até o momento em que se redige o Dicionário, e supondo também que nessa série de significados uns derivassem dos outros formando uma cadeia semântica, teríamos um critério de ordenação que seria ao mesmo tempo lógico e genético. Depois, colocadas as acepções pela ordem cronológica das datas em que cada uma delas apareceu pela primeira vez na língua, poderia acontecer que esta ordem coincidisse com a da série genética e, ao mesmo tempo, tão objetiva e convincente que colocaria o lexicógrafo acima de qualquer discussão.

Infelizmente, este ideal é poucas vezes executável. Muitos significados, que sem dúvida existiram na fala e sem os quais não é explicável a passagem de uma acepção documentada para a imediatamente seguinte na ordem genética, desapareceram sem deixar vestígio nos textos conhecidos. A dificuldade cresce quando a raiz do nosso símile, isto é, a etimologia, está representada por vocábulos que na língua de origem têm já significados múltiplos em consequência de um desdobramento semântico que pode ser para-

lelo ao que logo observamos no espanhol, divergente ou desenvolvido em sentido inverso. Nestes casos não é possível pensar numa série genética única. Pode haver tantas quantas sejam as acepções do étimo transmitidas à língua herdeira, onde às vezes se cruzam ou se bifurcam, dando origem a novas séries.

Por outro lado, a ordem cronológica em que as acepções fazem a sua aparição no espanhol parece que deveria remontar-nos invariavelmente até enlaçar com a etimologia, mas esta suposição é desmentida com grande frequência. Certo vocábulo grego, que através do latim vulgar entrou com significação desfigurada no espanhol, reaparece muitos séculos depois com toda a sua pureza etimológica, ora como um cultismo ora como um termo técnico de direito, de filosofia ou de medicina. Outras vezes é a forma que se afasta do étimo para retroceder ao cabo dos anos até a origem. A palavra latina *creta* = (argila, barro branco, giz) incorpora-se ao espanhol como *greda* (= greda) e assim a vemos na *Crônica General* e no *Lapidario* de Alfonso el Sabio²², no *Fuero de Sepúlveda*²³, nos *Glossários latino-espanhóis* publicados por Américo Castro, em Nebrija²⁴, etc., até que a grafia *creta* aparece no primeiro terço do século XVII, e precisamente num texto científico: na tradução que fez Gerônimo de Huerta da *História Natural* de Plínio.

Fizemos estas observações para que não se pense que no método histórico se tem encontrado uma panacéia ou um procedimento automático que exclui todo elemento conjectural. Nem o sistema resolve todos os problemas que preocupam o lexicógrafo, nem a sua adaptação à variedade dos casos que se apresentam na prática é coisa que se possa regular de antemão mediante uma casuística geral. Assim o reconhece publicamente, depois de uma experiência de cinquenta anos, o mais rigorosamente histórico de todos os Dicionários históricos modernos: o chamado abreviadamente de *Oxford* (*The Oxford English Dictionary*), iniciado sob a direção de Dr. Murrat. A propósito de certos artigos em que a ordem genética difere abertamente da cronológica, lemos no prólogo da referida obra: “É preciso tratar cada palavra da maneira que pareça mais adequada para apresentar os fatos da sua história e do seu uso.”

2. Ensaio de aplicação deste método a um verbete do Dicionário — Etimologia do lat. *ordo*.

E agora vamos examinar o primeiro vocábulo que se nos depare, como se estivéssemos na mesa de trabalho dispostos a redigir sua biografia. Tornaremos a utilizar a palavra *orden* (= ordem) que já apareceu anteriormente.

O substantivo *ordo* (lat.), étimo indubitável de *orden* (esp.), está intimamente relacionado com o verbo deponente *ordior*, *-iri*, que significa “começar a tramar um tecido” (Walde, *Lat. Etym. Wörterbuch*). Trata-se, pois, de vocábulos procedentes da linguagem têxtil e disso ficam incontestáveis vestígios em *redordior*, “devanar” (= dobar, enovelar) (Plínio); em *exordium*, “urdidumbre y comienzo” (urdidura e começo) (Quintiliano); em *togae exorditae*, “togas empezadas a tejer” (togas começadas a tecer) (Festo), etc. Vemos aqui dois conceitos associados, “começar” e “tecer” que com o passar do tempo, iam seguir rumos divergentes. *Ordior* (ou mais propriamente *ordio*, *-ire* com o significado de “tecer”, está documentado em Sêneca e em Plínio; mais tarde em São Jerônimo e Santo Isidoro, e continua nas línguas românticas (esp. e port.: *urdir*); porém no apogeu da latinidade (Cícero, Cornélio Nepote, Virgílio, etc.),

22 N.T.: Afonso X, o Sábio — (1221 — 1284) foi rei da Espanha e destacou-se pelas suas atividades culturais.

23 N.T.: *Fuero* (compilação de Leis) de Sepúlveda: data de 1076 sob o reinado de Alfonso VII.

24 N.T.: Antonio de Nebrija (1444 — 1522), autor da primeira Gramática Castelhana (1492).

a acepção que predomina quase exclusivamente é a de *começar, empreender* (cfr. *orsa*, “empresa”, *orsus*, “começo”).

Quanto a *ordo*, embora partamos do parentesco etimológico com *ordio*, não basta isto para esclarecer seu significado primário, visto que se torna difícil estabelecer uma ponte semântica desde a idéia de “começo”, até a de “fila de coisas”, que é a primeira acepção documentada. Faz-se necessário voltar novamente ao tear. Porque *ordior* significa conjuntamente “começar” e “tecer”? Porque a textura se compõe de duas fases consecutivas correspondentes à textura e à trama. A textura, embora não o digam os Dicionários, chamou-se em espanhol a “urde”. Assim o atesta o seguinte provérbio: “Urde haya que la trama el diablo la caga” (haja urdidura, que a trama o diabo a defe-ca). Correas explica: “que começando as obras, põe-se diligência e remédio para terminá-las”. Aqui temos um reflexo da consciência lingüística primitiva que associava a idéia de “começo” à de “urdir”. A operação de tecer, de fato, “começa-se colocando no tear uma “série de fios” paralelos, e arremata-se atravessando e entrecruzando esses fios com outros, que é o que constitui a “trama”. Ora, nessa série de fios ainda não tramados é onde há que buscar o nascimento do vocábulo *ordo*. A conjectura não é muito arriscada porque a apoiam outros termos que também significam “série”; como “fila”, “fileira”, “fiada”, “enfiada”, etc., onde está bem patente a idéia do “fio”. O próprio vocábulo latino *series*, apesar do seu atual aspecto abstrato, teve na sua origem um sentido material como derivado de *sero, -ere*, “entrelaçar”, “trançar” (*series capillorum*).

3. Polissêmia deste Vocábulo

Se admitirmos, pois, que *ordo* designou originariamente o “conjunto” de fios paralelos que se põem na urdideira antes de serem atravessados pela trama, já não há dificuldade em passar dessa acepção hipotética para a de “fila de coisas”, abundantemente documentada. Cícero fala da fila de oliveiras que limitam em linha reta uma herdade: *Ejus fundi extremam partem oleae directo ordine definiunt*; Virgílio menciona uma embarcação com os remos dispostos em três filas: *termo consurgunt ordine remi*, etc. (Acepção α , cf. p.37).

Daí, por um processo psicológico habitual, que consiste em trasladar para a coordenada do tempo as noções originariamente espaciais, surge o conceito da seqüência, da sucessão cronológica, da série de fatos que se apresentam uns após os outros. Esta acepção (*dzeta*) consideramo-la secundária; não tem descendência própria e amiúde está implícita em outras acepções. Assim, quando as pessoas em fila (*eta*) vão chegando a certo lugar, fazem-no uma após a outra: *singuli per ordinem cum ferro cubiculum intrarunt* (Quintiliano).

E, por último, como um denominador comum a todos os conjuntos de coisas que se situam ou se sucedem guardando entre si certa relação, aparece o conceito abstrato de “ordem”, que Cícero define como segue: *ordinem sic definiunt, compositionem rerum aptis et accomodatis locis* (acep. *lambda*).

Tentemos agora historiar a descendência das duas acepções fundamentais A e B, começando pela série Aa. Entre as muitas coisas que ordinariamente se encontram formando fila, estão os assentos do circo e do teatro, de onde, pelo mecanismo ordinário da especialização, *ordo* passa a significar “fileira”, “fila”, “degrau” (“grada” no esp.) de um local de espetáculos (acep. *beta*). Ora, nem todas as fileiras (“gradas”) do circo romano estavam indistintamente ao alcance do público. Em virtude de um decreto de Augusto, existia, entre as fileiras, uma certa escala de preferência, estabelecida

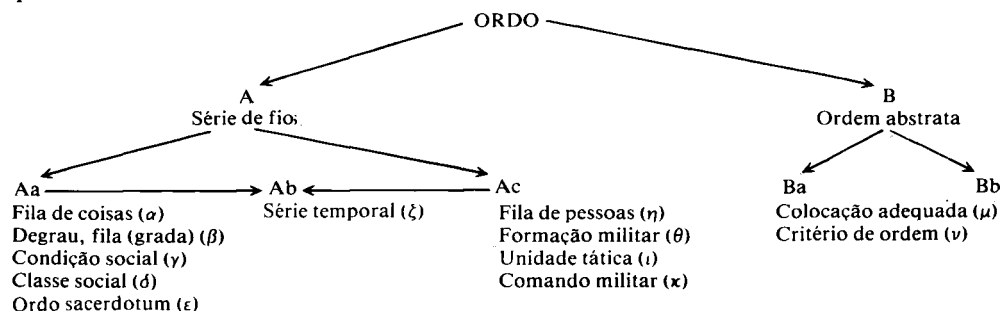
com arranjo à categoria social de quem devia ocupá-las. Havia, por exemplo, catorze fileiras destinadas exclusivamente aos senadores, e por isso, quando se dizia de alguém que se sentava em tais fileiras, *sedisti in quatuordecim ordinibus* (Cícero), era tanto como reconhecer-lhe a sua dignidade senatorial. Outras fileiras estavam reservadas para os soldados, outras para a plebe, etc., e daqui resultou um princípio de classificação hierárquica dos cidadãos e conseqüentemente uma nova acepção de *ordo* no sentido de “condição social” (acep. *gamma*). O processo abstrativo que observamos em “gradus”, desde degrau de escada, *escalarum gradus*, até “categoria hierárquica”: *In altíssimo dignitatis gradu collocati sumus*. (Cícero).

Depois, pelo fato mesmo de estarem assim hierarquizados os componentes da sociedade, as pessoas de cada grupo, consideradas em conjunto, passaram a constituir uma “classe social”, um grêmio ou um colégio: *ordo senatorius*, *equester ordo*, *ordo oratorum*, *mercatorum*, etc. (acep. *delta*). Entre estas ordens ou grêmios nos interessa distinguir, pela importância que mais adiante iria tomar, a classe dos servidores do templo, constituídos em colégios: a “ordem dos sacerdotes” (*summa collegia*); a *ordo haruspicum* e, na Vulgata, a *ordo levitarum* (acep. *epsilon*).

Chegamos até aqui, partindo das “coisas em fila”. Consideremos a fila de pessoas (acep. *eta*). Podemos imaginar as pessoas uma atrás da outra e teremos a noção de “fila” ou de séquito: *Comitum longissimus ordo* (Juvenal); ou ainda alinhadas uma ao lado da outra, como vemos na formação militar: *ordine egredi*, “sair das filas” (Salústio). Esta última maneira de dispor as pessoas em filas (acep. *teta*), que já não é circunstancial, mas que obedece a uma finalidade específica, varia, naturalmente, segundo o número e a colocação dos soldados que se agrupam, e pode ser uma coorte, uma centúria, um manipulo, etc. Tais frações de um exército tomam também o nome de *ordo* (acep. *iota*); e quando se confia a certa pessoa a direção de uma destas unidades táticas, recebe, como é natural, o “comando” dela (acep. *capa*), *ordinem accepit in legione quinta* (Inscrição).

Só nos resta por estudar a série B, “colocação respectiva de várias coisas” (acep. *lambda*). Esta colocação pode ser caprichosa, injustificada, meramente casual, ou pode fazer-se com vista à proporção e à harmonia, isto é, em “boa ordem” (acep. *mi*): *Nihil est pulchrior in omni ratione vitae dispositione atque ordine* (Columela). Esta acepção restringida de ordem pressupõe um critério de ordenação, uma norma, uma predeterminação, nascida em alguns casos, da mesma índole das coisas (leis físicas: ordem da natureza), emanada de uma vontade soberana: *Quid sumus... ordo Quis datus* (Pérsio), ou estabelecida pela razão, pela moral ou pelo costume (acep. *ni*): *extra ordinem crimina probantur* (Digesto).

Apresentadas as deduções anteriores de uma forma gráfica, dariam o seguinte esquema:



Pouco importa para o nosso propósito que o esquema genérico imaginado se ajuste mais ou menos à realidade histórica. Não nos comprometemos a resolver problemas de lexicografia latina nem dispomos para isso dos materiais que oferecerá em ocasião oportuna aos investigadores o *Thesaurus* das Academias germânicas, cujos trabalhos *se renovam felizmente nestes momentos*²⁵. O que nos interessava era apresentar um quadro sinótico das principais acepções que teve *ordo* em latim, antes de que nascessem os dialetos românicos, e identificá-las com algum signo que nos permita seguir-lhes a pista dentro de lexicografia espanhola.

4. Sua descendência em espanhol, francês, italiano e inglês.

Seria certamente muito instrutivo um estudo comparado da descendência semântica de *ordo* nas línguas românicas; porém, embora supondo que existisse — e não existe em muitos casos — toda a informação necessária, esse estudo nos desviaria demais da finalidade concreta que perseguimos. Não obstante isto, cedemos aqui e lá à tentação de expor alguns dados que ilustram a evolução do vocábulo “ordem”, tomando-os do francês, do italiano e do inglês.

Para o francês utilizou-se o Dicionário de Godefroy²⁶, o de Littré²⁷ e o *Dictionnaire Général*²⁸. Os dados referentes ao italiano procedem na sua maior parte do Dicionário de Tomaseo²⁹, que rara vez vai além de Boccaccio ou de Dante, e isto deve-se ter presente, porque o fato de que não contemos com autoridades italianas dos séculos XI ao XIV não significa que não tenham existido³⁰.

As referências ao inglês justificam-se por duas razões: a primeira é que o vocábulo que estudamos ingressou na referida língua, com toda a sua bagagem latina, quase nos mesmos dias em que aparecem os seus rebentos nos mais antigos monumentos das línguas românicas e, convertido em *order* (*ordre*, *wordre*, *ordyr*, etc.), seguiu uma evolução espantosamente paralela àquela que observamos no espanhol. Por outro lado, o Dicionário de Oxford é considerado unanimemente como o protótipo dos Dicionários históricos e contém uma informação só comparável à que oferecerá a seu tempo o *Diccionario Histórico de la Lengua Española*³¹. É esta é a segunda razão que nos moveu a ter o léxico inglês sempre à vista.

Catorze são as acepções de “ordem” que figuram no léxico oficial. Desde já podemos adiantar, sem risco, que quando os redatores do nosso primeiro *Diccionario Histórico* chegarem a desenvolver este verbete terão de triplicar, pelo menos, o número atual de acepções; porém, a fim de não complicar demais a demonstração que ora tentamos, só manejaremos escassamente umas trinta, incluídas naturalmente as do Dicionário Acadêmico, e também as agruparemos em séries semânticas, como fizemos em latim, porém marcadas agora com números romanos.

Série I.

A acepção “fila de coisas” (Aa do latim), aparece documentada em espanhol desde fins do século XIII e continua vivendo até os nossos dias. Em francês produz-se o fenômeno curioso de que *ordo* origina, ao lado da forma semiculta *ordre*, a popular

25 N.T.: Aproximadamente 1949 ou 1950.

26 N.T.

27 N.T.: Cf. (nota 11, p.76)

28 N.T.: Cf. (nota 10, p.76)

29 N.T.: Niccolo Tommaseo — *Dizionario della Lingua Italiana*, Torino, 1917.

30 Parece que a Academia da Crusca se ocupa atualmente de preparar um Dicionário italiano medieval.

31 N.T.: O autor refere-se ao *Diccionario Historico de la Lengua Española*, Madrid, publicado em 1960 pela Real Academia Española.

orne, que convive durante algum tempo com ela e adquire a acepção de que tratamos. *Orne*, como “fila de cepas”, remonta até o século XIV e subsiste ainda em alguns dialetos. O significado de “fila de coisas” registra-se em italiano por volta do último terço do século XVI, quase na mesma data do que em inglês. As citações espanholas mais antigas são as da *General Estoria*³².

Vem a seguir nesta série a sucessão no tempo. Encontramo-la em francês no século XIII (*Roman de la rose*), em espanhol por volta de 1344³³ e em inglês em 1382.

A “grada” (port. degrau, fileira) no sentido material de “fila de assentos”, figura unicamente na *Historia Imperial* de Pedro Mejía (1545)³⁴. Vemo-la também em italiano em data algo posterior.

“Ordem” com o significado translaticio de “grau”, isto é, como escalão de uma hierarquia social, achamo-la pela primeira vez no *Fuero Juzgo*³⁵. Trata-se de tomar declarações a testemunhas ausentes, as quais deverão depor diante do juiz manifestando tudo o que sabem “por ordem”. Este modo adverbial “por ordem” traduz a expressão mais explícita do texto latino *per conditionum seriem*. Cabe, portanto, pensar em um decalque, mas há autoridades dos séculos XV e XVI que não deixam lugar a dúvidas quanto à existência desta acepção. Não a vemos em francês nem em italiano, mas sim em inglês a partir de 1300.

Ao chegar aqui bifurca-se a série I. O que era condição das pessoas traslada-se às coisas, para significar a classe destas, sua qualidade, seu degrau, sua natureza; espanhol, 1520³⁶, francês, 1532; inglês, 1736.

Daqui nasce no século XVIII o uso técnico de “ordem” na língua dos naturalistas, com valor intermediário entre a classe a tribo ou a família³⁷, e depois, na técnica matemática, a “ordem” de uma curva ou de uma equação com a equivalência de “grau”.

Voltando à linha principal da série achamos a acepção de “classe social, grêmio, corporação”, etc. Documenta-se em francês desde o século XII, em inglês desde o XIV, em espanhol³⁸ e italiano desde o XVI.

Já dissemos que entre as várias “ordens” de pessoas que existiram em Roma, senadores, patricios, cavaleiros, plebeus, etc., interessavam-nos particularmente a “ordem” dos sacerdotes. Durante o primeiro período da República, os que haviam de adquirir a dignidade sacerdotal, eram admitidos por cooptação no colégio correspondente. A partir da implantação do Cristianismo, a potestade suprema para a organização da Igreja, criada pelo próprio Jesus Cristo, correspondeu aos Apóstolos. Eles nomearam como seus auxiliares os diáconos, com faculdade para administrar o batismo. Instituíram à continuação o grau dos presbíteros para as funções do culto e da administração de certos sacramentos e, por último, criaram os bispos, a quem transmitiram a plenitude da potestade apostólica. Cada uma destas delegações, que imprimia caráter sagrado a quem as recebia, efetuava-se mediante certos ritos que eram outras tantas, con-

32 “... por medio del campo delas tablas yvan tres *órdenes* de varas e desta guisa fazien-se V *órdenes* dellas”. Ed. Solalinde, pág. 442.

33 “Tal ordem dio el traydor como los descabeçasen uno a uno, en *orden* asi como nasçieron”. *Infantes de Lara*, Ed. M. Pidal, pág. 339.

34 “... los theatros... de tal manera eran obrados y hechos que por sus *órdenes* y assientos eran capacissimos de infinita gente”. Ed. 1547, f.º 62 v.

35 N.T.: “Fuero Juzgo”: compilação legislativa dada por Recesvinto e aprovada em 53/54. Feita unicamente para o serviço dos Tribunais de Justiça.

36 “... según la *orden* de las inclinaciones naturales que los hombres tienen, así es la *orden* de los preceptos de la ley natural”. — Las casas, *Apologética historia*... NB de AE, t. 13, pág. 103.

37 “... hallamos en las métodos clases, *órdenes*, tribus, legiones, cohortes y familias, de los géneros y las especies”. Clavijo. Tard. de la *Hist. Nat.* de Buffon, Ed. 1785, t. 17 pág. 228, nota.

38 “... a la *orden*, equestre, que era media entre al pueblo y los patricios, hizo (Caligula) solemnissimo combite y sala”. Pedro Mejía, loc.cit., f.º 32 r.

sagrações ou “sacramentos”. Chamava-se, pois, “ordem” (esp. “orden”) cada um dos graus da hierarquia e também se chamava “ordem” o conjunto das sucessivas ordenações com que se chegava à plena dignidade sacerdotal. A acepção correspondente aos “graus do sacerdócio” documenta-se como segue: espanhol, 1250³⁹; inglês, 1300. A de “ordem sacerdotal” aparece na França no século XI, na Espanha em 1140, na Inglaterra em 1290 e na Itália em 1556. A autoridade espanhola que alegamos, e que é por certo a mais antiga de todas as acepções de “ordem”, achamo-la no Mío Cid⁴⁰. Para chegarmos a esta conclusão, tivemos que dissentir, com todo respeito, de Menéndez Pidal, que entendeu “regla monástica” (regra monástica) onde nós lemos “ordem sacerdotal”. Trata-se da passagem (verso 2373) em que o bispo Dom Jerome se apresenta ao Cid desejoso de Matar mouros: “mi orden et mis manos querria las onorar” (minha ordem e minhas mãos, queria honrá-las). Segundo a nossa opinião, o que o bispo quer honrar, derramando sangue dos infiéis, é o caráter sagrado de que está investido e não uma “regra monástica”. Esta interpretação é confirmada pela *Crônica de Veinte Reyes*, onde o bispo diz ao Cid: “porque yo oy dezir que siempre guerreavades com mouros, por eso me vine yo de mi tierra para vos, cobdiçando lidiar con ellos; e por ende por honrar mis manos e mis *hordenes*” (porque eu ouvi dizer que vós sempre guerreáveis com os mouros, por isso vim da minha terra para junto de vós, cobiçando lutar com eles, e com o fim de honrar as minhas mãos e as minhas ordens).

A acepção de ordem sacerdotal levava implícitas, como vimos, a noção de sacramento e a de hierarquia. Transposta esta última aos espíritos angélicos, aplicou-se a cada um dos nove coros, enquanto a de sacramento, desvinculada de toda reminiscência hierárquica, pôde designar qualquer dos restantes sacramentos e especialmente o do matrimônio. A acepção de “coro dos anjos” registra-se no inglês em 1225 e no espanhol por volta de 1720⁴¹. A de “sacramento do matrimônio” aparece na França e na Espanha⁴² em meados do século XIII e na Inglaterra no XIV. A extensão a outros sacramentos é peculiar à França, onde a partir do século XIV uma pessoa “ordenada” era a que tinha confessado ou comungado e tinha recebido a extrema-unção.

Pelo mesmo caminho por onde a *ordo sacerdotum* veio a constituir um grupo social dentro do Estado nasceu logo a *ordo monachorum*, isto é, a ordem religiosa como corporação ou comunidade, acepção que aparece no espanhol⁴³, inglês e francês durante o século XIII e na Itália durante o século XIV. Estas instituições religiosas, que existiram já em meados do século VI da nossa Era, tinham de receber a aprovação dos Pontífices para ficarem devidamente constituídas. Posteriormente, outras irmandades, cujos indivíduos se ligavam por votos metade religiosos e metade guerreiros, deram origem às “ordens militares”. Neste caso eram os reis aqueles que confirmavam o correspondente estatuto; e junto a estas ordens funcionavam as de cavalaria, que, embora tivessem finalidade diferente, confundiam-se amiúde com aquelas, pelo menos na denominação. Apesar de que a ordem de Alcântara espanhola foi aprovada em 1156, a ficha mais antiga de nosso arquivo é da primeira metade do século XIII⁴⁴. Na França a

39 “... después que los nuestros cléricos, que son ordenados de *orden* de epistola o dend arriba, se entierran oy cada uno com las vestimentas de su *orden*”. *General Estoria*, Ed. Solalinde, p. 524.

40 N.T.: *Poema de Mio Cid*, a epopéia espanhola mais antiga (1140) conta as façanhas do guerreiro nacional espanhol Don Rodrigo Díaz de Vivar, apelidado Cid pelos árabes e Campeador pelos espanhóis.

41 “Nueve *órdenes* de ángeles ordenó nuestro señor Dios em la elesia celestial”. *Partidas*, 1.ª, tit. 6, pról.

42 “Onde porque esta *orden* del matrimônio establesió Dios mesmo por si...”. *Partidas*, 4.ª, pról.

43 “... e do y mio cuerpo e de mi mugier fuera ent si entráremos en *orden*”. *Doc. Ling.*, Ed. M. Pidal, 77, 49, 10.

44 “Mando que en Çorita non aya sinon un palacio de la *orden* de Calatrava”. *Fuero de Zorita*, Mem. Hist. Esp., t. 44, pág. 50.

acepção de ordem militar já aparece nos textos do século XI, na Inglaterra, no século XIV, e, na Itália, no século XVI, embora conste que a Annunziata data de 1362.

Destas ordens derivam as meramente honoríficas que existem hoje em quase todos os países como distinção outorgada pelo Estado para premiar determinados serviços. E com isto chegamos ao final da série I e das suas ramificações.

Série II

A “fila de pessoas” só tem uma citação antiga no espanhol, correspondente a meados do século XV: “van en pos de la Reyna una luenga *orden* de sus mugeres” (*Mar da Istorias*) (vão atrás da rainha uma longa fila das suas mulheres). Não está documentada nas outras línguas.

A formação militar registra-se a partir do século XVI na Itália, e é duvidoso que antes dessa data tenha existido na Espanha⁴⁵.

A acepção de unidade tática parece a usada no Poema de Afonso XI⁴⁶, porém a autoridade que não oferece dúvida é a de Palência: “assi las alas como los cuneos i quales quier otras disciplinas militares, se llamaron *órdenes*”. Podemos pois, datar esta acepção hipoteticamente meados do século XIV e com segurança no fim do séc. XV. Não ficou vestígio algum em nenhuma das outras línguas que comparamos.

O comando militar só tem a seu favor uma autoridade de Coloma (1625). Existe também no francês e precede um século a nossa citação⁴⁷. E com isto damos por terminada a série II.

Série III

Vemos a acepção de “ordem” como colocação respectiva de coisas ou pessoas, em geral, na Espanha desde meados do século XII⁴⁸; na França usou-a “o recluso de Molliens” em princípio do referido século, e na Inglaterra e na Itália é documentada no século XIV.

Vem a seguir a idéia da boa ordem, da disposição conveniente ou especialmente adequada para certo fim. Espanha, meados do século XIII⁴⁹; Inglaterra, 1382. E como entre as coisas convenientemente dispostas existe certa conexão ou recíproca dependência, origina-se o significado de relação que se consolidou na locução “en orden a” (a fim de, por meio de) muito usado no século XVI na Espanha⁵⁰ e na Inglaterra (*in order to* ou *for*).

Por outro lado, o conceito de colocação adequada e harmoniosa de elementos destinados a um mesmo fim dá origem a uma acepção técnica de arquitetura: *ordem corintia*, *ordem composta* etc., Espanha, 1552⁵¹, Inglaterra, 1563, França, 1694. Esta

45 Inclui-se esta acepção no esquema, mas vistas as variantes que oferece a passagem que servia de autoridade, prescindiu-se de citá-la.

46 Trata-se de uma leitura equivocada da Ed. Riv., cantiga 1.763 — Diz *orden* e deve ser *otero* (outeiro).

47 “... quedó todo el campo á *orden* de monsieur de Rona”. Coloma, *Gerras de los Países Bajos*, ed. Riv. pág. 105.

48 “... cata do cae el cabo del alhizada de los quentos del quarto que son escriptos de uno fata XC et non pares mientes á los quostos si van por *orden* ò non”, Lib. del Sab. de Astron., ed. Rico y S., t.2, pág. 205.

49 Fue ante de medio día el comer aguisado /.../ fue el pueblo venido, por orden asentado, /el enperador sobre todos como omebien enseñado”. *Lib. de Alexandre*, ed. Morel Fatio, copla 2572.

50 “... todos quantos se salvaron desde el principio del mundo se salvaron en *orden* á la Pasión de Cristo”. Fr. J. de los Angeles Ob. núst., ed. NB de AE, t. 20, pág. 508.

51 “A Venus, Flora, Proserpina, yá las Náyades, parece convenir el *Orden* Corintio, porque las fábricas primorosas, y adornadas de flores, hojas y volutas, parecen añadir belleza á la propia de estas deidades”. Ortiz, trad. de Vitrubio, ed., 1787, pág. 12.

acepção parece ter entrado nas línguas modernas com as primeiras traduções de Vitruvio.

O conceito de critério ordenador, que já vimos despontar em latim, verifica-se nas línguas herdeiras dando origem a acepções especiais segundo a natureza desse critério. Assim temos:

1. uma ordem do Universo, que reflete a norma soberana obedecida por tudo o que foi criado, no espanhol e no inglês, meados do século XVI⁵²;

2. uma ordem humana, que se rege pela razão, pela moral ou pelos costumes, significado que se registra no espanhol desde a segunda metade do século XIII⁵³, no inglês, no século XV, e no italiano, no XVI; e

3. outra ordem também humana que se concretiza na regra ou conjunto de regras aplicáveis em certos casos. Espanha, século XV⁵⁴; Inglaterra, século XIV. Como um caso particular desta última acepção podemos considerar a “regra monástica” que é mais ou menos severa e que se observa ou se guarda bem ou mal. Este significado acha-se em Berceo⁵⁵ e não aparece nas outras línguas que estudamos⁵⁶.

Não se esgota aqui esta série semântica, porque a regra coletivamente aceita pode ser substituída pelo desígnio individual na forma de plano, traçado ou projeto: Espanha, século XVI⁵⁷, Inglaterra, século XV. E ainda depois de adotado um plano, fica por determinar o procedimento, o método, a maneira de pô-lo em prática. Esta acepção é recolhida na França pela forma popular *orne* a princípios do século XIV, enquanto o semi-culto *ordre* não é incorporado até muito depois. Na Espanha temos essa acepção desde o século XV⁵⁸, e na Inglaterra e na Itália arrasta-se até o XVIII.

Pode ainda acontecer que a pessoa que imaginou um plano ou um procedimento não seja a mesma que irá executá-lo e neste caso terá que valer-se de outra ou outras pessoas que agirão por delegação, mediante “instruções” verbais ou escritas. Esta acepção de *ordem*, abundantemente documentada no espanhol desde o século XIV⁵⁹, não se vê registrada em outras línguas, embora seja quase certo que terá existido como antecedente do último significado de *ordem* que recolhemos em continuação, a saber: quando aquele que recebeu as instruções está ligado àquele que as deu por um vínculo de obediência e não pode, por conseguinte, deixar de ouvi-las nem alterá-las, o encargo ou a comissão convertem-se “ipso facto” no que o Dicionário acadêmico define como “mandato que se deve obedecer, observar e executar”. Esta acepção, que se delinea na Espanha desde o século XIV⁶⁰, apresenta-se na Itália por volta da mesma época e ainda demora dois séculos a aparecer no inglês. A mais antiga citação francesa (Littré) corresponde à segunda metade do século XVII.

52 “... yo pensaba... que eran tus hechos regidos por alguna orden.” *La Celestina*, act. 21.

53 “Dize Hugo del Pradó que la orden de caridad es: que Dios sea amado sobre todas cosas”. *Evangelios e Epistolas*, ed. collijn y Staaff, pág. 199.

54 “Todo hombre en esta vida debe tener algun orden de vivir”. Pulgar, *Glosa de Mingo Revulgo*, ed. 1787, pág. 8.

55 N.T.: Berceo, Gonzalo de (1195/1264): primeiro poeta espanhol de nome conhecido. Famoso pelas suas obras em louvor a Nossa Senhora.

56 “Grado bueno a Dios, e a Sancta Maria./ Non aviné meior nul monge en la Mongia./ Lo que diçe la regla, facia ét todavía./ Guardaba bien la orden sin ninguna folia”. Berceo, *St. Dom.*, ed. Riv., t. 57, pág. 42.

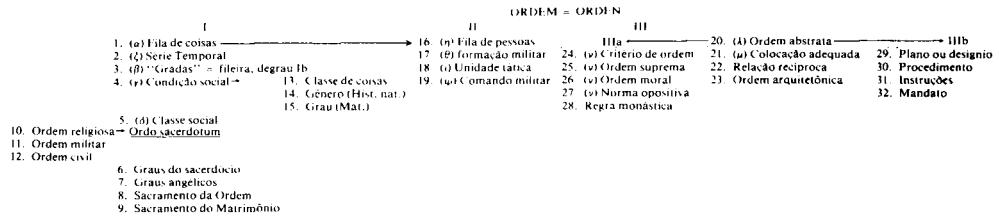
57 “Hecho el recebimiento, no como fué, que sería imposible decirlo, mas como á la orden del libro conviene, movieron todos juntos para la insola”. Amadis, ed. Riv., t. 40, pág. 358.

58 “Mossen Febrer fiço obras notables e algunos afirman aya traydo/ el Dante de lengua florentina en catalán, non menquando punto en la orden del metrificar e consonar”. Santillana. *Ob.*, ed. 1852, pág. 11.

59 “Así que, yo daré orden cómo, compliendo con mi servicio e con vuestra honra, della podais ser excusado”. *Amadis*, ed. Riv., t. 40, pág. 147.

60 “La... puerta de los Deanes, la que se abrió por el alcayde de dicha Santa iglesia en fuerza de la orden que para ello le estaba dada”. *Mem de Fernando IV*, ed. Acad. Hist., t. 2., pág. 869.

O resultado das deduções expostas, apresentado em forma sinótica, pode ver-se no quadro anexo.



5. Ensinamentos que se deduzem da experiência realizada.

Antes de passar adiante, gostaríamos de fazer uma pausa para tirar da cansativa exposição precedente algumas conseqüências de caráter geral. É a primeira sensação confortável de confiança no que se refere à existência e substantividade das acepções enumeradas. O lexicógrafo nunca pode evitar neste ponto a receio de ter pecado por excesso ou por deficiência. Tal acepção, colocada por separado, não seria um matiz insignificante de tal outra? Não existem em certa definição elementos heterogêneos que deveriam receber tratamento individual?... No caso presente, o escrúpulo fica notavelmente atenuado ao comprovar que a nossa divisão em acepções coincide quase exatamente com a que realizaram independentemente uns dos outros e em línguas diferentes os mais acreditados dicionários modernos.

A segunda consideração refere-se às possibilidades do futuro Dicionário histórico espanhol. Já vimos experimentalmente que o nosso material, ainda incompleto¹, supera notavelmente em informação os dicionários franceses e italianos, e até pode medir-se sem deslustre com o formidável arsenal que serviu de base ao monumental Dicionário de Oxford.

Na apresentação das trinta acepções de *orden* identificadas de maneira sucinta, isto é, omitindo em benefício da brevidade definições circunstanciais, procurou-se seguir um encadeamento lógico que pretende ilustrar a forma como cada significado pôde nascer do anterior. Supusemos que a acepção "mandato", última da série IIIb, originou-se da penúltima e esta da anterior, e sucede que a realidade histórica não se opõe a esta conjectura²; porém, que faríamos no caso de um desacordo flagrante entre a ordem genética suposta e a cronológica comprovada? Por mais que nos doa confessá-lo, este conflito, que ocorre com frequência, ainda não encontrou solução satisfatória. Numeradas correlativamente as acepções do Dicionário de Oxford que interessam ao nosso propósito, isto é, as que são iguais ou equivalentes às espanholas, e colocadas depois de acordo com a data mais antiga das autoridades alegadas em cada acepção, a que hoje figura com o número 1 passaria ao 19.º lugar, a 2, ao 7.º, a 3 ao 9.º; a 4, ao 24.º etc. Além disto, há acepções que estão notoriamente fora de lugar, como a de *orden* em sentido arquitetônico, colocada na série dos grupos sociais como continuação das ordens de cavalaria. Desde logo, não é um acerto o verbete *orden* do famoso

¹ As numerosas contribuições de última hora, que correspondem precisamente aos primeiros monumentos da língua, não se puderam utilizar para o trabalho, pelo que se pode prever que, quando se redigir o verbete *orden* do projetado Dicionário Histórico, a disposição das acepções não coincida com a do nosso esquema provisório. Este só tem como finalidade a de apresentar um caso prático como ponto de referência para as observações de índole geral.

² Dizemos que "não se opõe", porque a sutil distinção entre as "instruções" para fazer uma coisa e o "mandato" para executá-la nem sempre resulta evidente nos textos espanhóis (Cf. as abonações das notas, p. 92, notas 59 e 60).

Dicionário inglês, e assim o reconhecem claramente os seus redatores num parêntese preliminar, onde nos advertem que “o método que foi seguido é em muitos pontos puramente provisório”; contudo as nossas observações não visam descobrir defeitos numa obra que é sob muitos aspectos admirável, mas apenas tornar evidente uma considerável margem de incerteza que ainda padecem os mais modernos critérios lexicográficos, o que tira todo aspecto de petulância à tentativa de remediar, no possível, algumas das deficiências observadas.

6. *Enlace das diversas acepções do étimo com as da língua herdeira.*

A primeira inovação introduzida com este fim no nosso plano consiste em ter substituído as quatro grandes séries do Dicionário inglês por três séries fundamentais que, subdivididas com arranjo ao seu conteúdo conceptual, originam um total de sete grupos semânticos suficientemente homogêneos. Deste modo facilita-se em muito a concordância da “ordem genética” com a “ordem histórica” dentro de tais grupos, o que seria impossível conseguir numa longa série heterogênea. Assim, pois, a precedência cronológica não se tem em conta uma só vez e misturando acepções tão diversas como o coro dos anjos, o sacramento do matrimônio, as ordens militares, as ordens arquitetônicas e os graus de uma curva (série II do Dicionário de Oxford), mas só se atende a essa precedência cada vez que se inicia um grupo, e sem preocupar-se com o que sucede nos outros.

Outra inovação, que pode ter conseqüências fecundas, refere-se ao modo de relacionar a etimologia com a progênie comprovada na língua herdeira, sempre que se trata de vocábulos polivalentes na língua de origem, isto é, nos casos de polissemia do étimo. Em vez de nos contentarmos, retomando o nosso exemplo, com escrever na etimologia que “ordem” procede do latim *ordo*, distinguimos e assinalamos as acepções deste vocábulo que passaram para o espanhol: α) “fila de coisas”; β) “sucessão no tempo”; γ) “fileira, degrau do circo” etc. Deste modo, evidenciamos um esquema da evolução semântica do vocábulo latino, evolução essa que pode comparar-se com a observada depois na língua espanhola e dar lugar a interessantes deduções.

Depois, em cada acepção espanhola, fazemos referência à latina, se é que existiu, com o que se estabelece uma rede de conexões que desfaz automaticamente muitos equívocos. Quando encontramos, por exemplo, no século XVI, e uma só vez, “ordem” com o sentido de “fileira do circo”, seja qual for o lugar que tenhamos atribuído a esta acepção dentro de uma série semântica, não temos que preocupar-nos como nem quando surgiu tal significado com relação ao que o precede. Trata-se simplesmente da acepção γ do étimo. Por outro lado, ao verificar apenas que a acepção de “mandato” não se refere diretamente a nenhuma acepção do étimo, somos induzidos a procurar uma explicação para esse fato dentro da evolução neolatina do vocábulo “ordem”. Também se conseguiria diferenciar deste modo aquelas acepções espanholas, que realmente se incorporam à língua, dos latinismos ocasionais sem raízes e sem conseqüências.

Por outro lado, uma vez postas frente a frente a série semântica latina e as correspondências espanholas, é quando nos podemos dar conta da solução de continuidade que se adverte entre umas e outras durante o período que compreende a ínfima latini-dade, e isto permite conjecturar em alguns casos o que sucedeu no referido período. No que se refere a “*ordem*”, vemos que as mais antigas acepções neolatinas, incluídas as que passaram para o inglês, são as de “sacramento”, “coro de anjos”, “ordem monástica” e outras do mesmo tipo, todas elas de caráter religioso, o que demonstra claramente que uma onda de latim eclesiástico se antecipou vários séculos à incorporação dos significados procedentes do latim literário.

7. Dados estatísticos de frequência no uso de cada acepção.

Outra inovação que gostaríamos de ver implantada é de caráter estatístico e estaria representada por um algarismo indicador da frequência com que foi usada cada acepção. Não se trata de uma simples curiosidade, mas de um dado importante, na nossa opinião, para a história das palavras, porque, evidentemente, não pode oferecer o mesmo interesse um desvio de significado na pena de um só autor que uma mutação semântica admitida e patrocinada por um grupo considerável de escritores. No primeiro caso, sempre cabe pensar numa distração ou num equívoco, enquanto no segundo é preciso levar a coisa a sério e buscar-lhe uma explicação. Assim, embora numa primeira impressão nos surpreendamos que “ordem” tenha significado “maneira de fazer as coisas”, temos que render-nos à evidência de umas vinte autoridades contestes pertencentes à mesma época; e posto que a economia do Dicionário não permita relacioná-las todas, poderia pôr-se em continuação às três copiadas (+ 17) para indicar a que número ascendem as fichas que ficaram arquivadas. Inversamente, quando não se pusesse tal parêntese, entender-se-ia que as autoridades apresentadas eram todas as disponíveis.

1. Os compostos intermitentes: Precauções para a sua ordenação

Até agora nós nos preocupamos somente com os vocábulos isolados como material próprio da lexicografia, embora sem esquecer que existem também unidades léxicas constituídas por dois ou mais vocábulos. A elas vamos dedicar agora a nossa atenção; antes porém, para desembaraçar o caminho, temos que falar, embora brevemente, das palavras compostas. Estas abundam mais numas línguas do que em outras. As teutônicas, e especialmente o alemão, como, por outro lado, o grego, têm marcada propensão a formar palavras compostas nas que entram vários núcleos significantes. O espanhol não carece certamente desta possibilidade, mas sempre usou dela com muita moderação e em determinadas condições. Um conglomerado de quatro substantivos como *Eisenbahnfahrpreis* “preço da passagem de (caminho de ferro) trem” é inconcebível no espanhol, e até os “simples” compostos bimembres como *casatienda* (loja-residência), *arquibanco* (banco comprido com gavetas), *maestrescuola* (mestre-escola), *bocacalle* (entrada de rua) etc., são relativamente escassos.

Além destes compostos com dois nomes, temos os de dois adjetivos (*agridulce* = *agridoce*); de adjetivo e nome (*vanagloria* = *vanglória*); de preposição e nome (*trastienda* = loja de fundos; *sobreloja*); de dois verbos (*vaivém* = *vaivém*); de verbo e nome (*ganapán* = *ganha-pão*); de preposição e verbo (*porvenir* = *porvir*); de pronome e verbo (*qualquiera* = *qualquer*); de advérbio e verbo (*malcriar* = educar mal); de conjunção e verbo (*siquiera* = *sequer*); de advérbio e nome (*bienvenida* = boas vindas); de advérbio e adjetivo (*malcontento* = descontente; *maleducado*) etc.

Não nos interessa completar a enumeração, que é tarefa própria da Gramática, nem estudar por agora quais desses compostos são do tipo “coordenado” ou “subordinado”, “progressivo” ou “regressivo”. Do nosso ponto de vista, que é o do lexicógrafo, examinamos uma questão prévia que afeta o grau de consolidação dos vocábulos compostos. Quando a fusão dos componentes pode considerar-se definitivamente consumada, o tratamento do signo verbal resultante e a sua colocação por ordem alfabética, são coisas que não oferecem dificuldade, entretanto, nos outros casos, convém proceder com muita cautela, posto que sempre houve e terá de haver um período flu-

tuante, uma etapa de transição desde que se inicia a aproximação mútua dos dois ou mais vocábulos até que se unam com um vínculo indestrutível, sacrificando um deles, e às vezes ambos, a sua figura individual.

A existência de uma autoridade, como a Academia Espanhola, que decide quando se pode escrever “dieciséis” (dezesseis) numa palavra, em lugar de “diez y seis” (o que não deixa de ter conseqüências para a taxa dos telegramas) evita em muitos casos dores de cabeça, porém se retrocedermos um par de séculos ou se tentarmos prever o que sucederá daqui a algum tempo, já não podemos contar com essa ajuda. *Porvenir* (porvir), segundo a Academia, só é um substantivo masculino. Isto reflete exatamente um uso moderno, mas um uso que provavelmente não vai mais além de fins do século XVIII. Ainda no último terço do século XIX escrevia D. Severo Catalina, “el tiempo por venir” (o tempo por vir), como era normal na época clássica. Contudo, antes que ela terminasse já se começou a escrever “porvenir” (porvir) numa só palavra, preparando o seu emprego como adjetivo, à semelhança de “venidero” (vindouro) ou de “futuro” (futuro). Um pouco mais adiante, assim como se dizia “tiempo pasado” (tempo passado), se disse “tiempo porvenir” (tempo por vir) e, suprimindo logo “tiempo” (tempo), que facilmente se subentendia, obteve-se o jogo completo de adjetivos masculinos substantivados o “pasado” (passado), o “presente” (presente), o “futuro” (futuro) e “el porvenir” (o porvir). A fórmula lógica e correta, “lo por venir” (o por vir), tinha desaparecido. Hoje já ninguém se preocupa com o que nos reserva “lo por venir”⁶³ (o por vir). Todos falamos de “el porvenir” (o porvir), e quando dizemos a um rapaz, com uma acepção ainda não registrada nos léxicos, que não se case sem ter assegurado o seu “porvenir” (porvir), nem sequer pensamos no tempo, mas apenas numa próspera situação econômica. O fato é que em nossos dias o amálgama dos elementos deste composto não pode oferecer dúvida e a sua ordenação alfabética também não. Porém, se estivéssemos no século XVIII, não poderíamos dizer o mesmo. “Porvenir” (por vir) em duas palavras teria que figurar no interior do verbete dedicado ao verbo “venir” (vir) e não em outro lugar. Por isso a Academia, apesar de que a grafia “porvenir” era usada pelos melhores escritores, inclusive Cervantes, não acreditou ter ante si um composto e não o registrou no *Dicionário de Autoridades*⁶⁴. Ela o fez por primeira vez em 1817 com o significado de “sucesso futuro”, sem fazer menção expressa de “tempo” e com a nota de “familiar”. Outros exemplos: “Al rededor” (ao redor), modo adverbial, converteu-se numa palavra a partir da quinta edição e foi parar, por conseguinte, da letra R à letra A. A locução “en seguida” (em seguida, logo depois) teve que esperar até 1947 para poder superar a separação ortográfica, pulando assim da letra S (verbo “seguida”) para a letra E. Faz dois anos, portanto, o candidato que escrevesse “enseguida” numa só palavra teria de lamentar na sua prova um risco vermelho, que poderia ser decisivo para o seu “porvenir” (porvir).

Trata-se em todos estes casos do reconhecimento de um uso geral que existiu durante um certo tempo antes que fosse sancionado. Desse uso tinham-se indícios visíveis quando cada um escrevia ao seu modo, refletindo assim o seu sentimento pessoal em relação à vitalidade ou diminuição do valor conceptual dos termos relacionados e em relação à independência ou unificação dos respectivos acentos. Quem pronunciasse a frase “escuela usted *cual quiera* de estos libros” (escolha o senhor qualquer um (qual quiser) destes livros) poderia pensar esse “quiera” (quiser) como uma forma do verbo

⁶³ N.T.: Em espanhol “por venir” exige o artigo neutro *lo*, enquanto “porvenir” (subst.) possui art. masculino *el*.

⁶⁴ N.T.: O *Dicionário de Autoridades* é o *Dicionário da Academia*, reproduzido periodicamente, sempre aumentado e corrigido, e que recebe esse nome (D. de Autoridades), em virtude de apresentar sempre testemunhas de escritores já consagrados pela fama.

“querer” e neste caso teria tendência a escrever separadamente, mas se dissesse a um amigo “dame un libro *cualquiera* para pasar el rato” (dê-me um livro qualquer para passar o tempo), a falta de concordância desse “quiera” referido à segunda pessoa desfaz toda intenção de que o amigo escolhesse o livro e o que pedia só dava a entender que lhe era indiferente um livro ou outro. “Cual” passava a ser proclítico de “quiera”, com perda do seu acento próprio, e isto se refletia na escrita mediante a união dos dois termos. O mesmo acontece com a expressão “como quiera” (como queira, como quiser), que faz algum tempo aparece nos textos formando uma só palavra, com a noção de “querer” perdida. “Pepito es un alumno aplicado, y no así *comoquiera*, sino el primero de la clase” (Zezinho é um aluno aplicado, e não de qualquer modo, mas o primeiro da classe). O Dicionário não legalizou esta união e a locução usual *como quiera que* aparece registrada no verbo “querer”⁶⁵. “Seguida”, como substantivo que expressa a “ação e efeito de seguir”, deixou de usar-se, segundo minhas informações, antes de terminar o século XVIII. Desde então para cá podemos coligir que todos os que empregaram “en seguida” tiveram consciência de manejar uma fórmula unitária; mas, eles a escreveram numa ou em duas palavras? Não há forma de sabê-lo. Quem tiver mandado papéis à imprensa e se tiver afastado em algum ponto da ortodoxia acadêmica, por descuido ou deliberadamente, ter-se-á visto inexoravelmente conduzido a ela pela mão inquisitorial do tipógrafo.

Tudo isto deve fazer meditar o lexicógrafo sempre que se lhe apresentarem dúvidas desta índole, sobretudo se se propõe realizar uma obra válida para várias gerações. Portanto cada vez que comprove que a seqüência inseparável de dois termos mostra sinais inequívocos de converter-se numa simbiose, e mais ainda quando este resultado possa influir na colocação que terá de dar-se ao composto, deverá anotar esse composto em seu novo lugar alfabético, sem suprimir uma eventual referência no lugar que ficou vago. Sirva de exemplo esse “comoquiera”, que, com razão deveria figurar já como composto na letra C.

2. As Combinações binárias de caráter estável.

Depois dos compostos flutuantes apresentam-se a nossa consideração as combinações binárias, entre as quais convém distinguir aquelas que se formam e desfazem ocasionalmente e aquelas de caráter estável. Estas últimas, são, a nosso entender, as únicas que deve recolher o Dicionário, posto que nelas um ou dois elementos combinados tomam um significado especial pelo fato de se acharem juntos. Na expressão “número primo”, por exemplo, o substantivo “número” poderia desprender-se do adjetivo e continuar sendo “número” como em “número par” ou “número ímpar”; porém “primo” só pode ter a acepção específica que aqui vemos se não se separar de “número”. No caso de “obra prima”⁶⁶ são os dois elementos que se influenciam mutuamente. “Obra” não equivaleria ao trabalho do sapateiro se não levasse o referido adjetivo ao lado, e “prima” não excluiria o trabalho de pôr saltos e meias-solas se não fosse a continuação de “obra”. O sapateiro de “obra prima” é o que faz calçado novo, diferentemente do remendão, que também se chama “zapatero valiente”, embora não o digam os léxicos.

65 N.T.: Realmente o Dicionário acadêmico registra essa expressão no verbete de “querer”, porém o *Dicionário de Uso de la Lengua Española*, de María Moliner, Madrid, Gredos, 1970 registra *comoquiera* num verbete separado.

66 N.T.: “Obra prima” em esp. significa o sapato novo feito pelo sapateiro, em oposição ao trabalho de conserto feito pelo sapateiro remendão. “Obra prima” (port.) diz-se em espanhol “obra maestra”.

Neste terreno como dissemos ao falar dos compostos intermitentes, é necessário andar com precaução. O Dicionário, por exemplo, considera combinação estável a expressão “ordem dórica” com um critério que nos parece discutível. “Ordem” tem já de per si uma acepção específica de arquitetura, e “dórico”, separado de “ordem”, continua significando o mesmo quando qualifica outros substantivos, como “coluna”, “capitel”, “estilo” etc. Um caso especial dentro das combinações binárias está representado pelos pares de substantivos em aposição, como *arco íris*, *pedra imã*, *sal amoniaco*, caso que examinaremos adiante sob outro ponto de vista; mas, em geral, estas combinações, vão parar sem dificuldade nem inconveniente nos verbetes do Dicionário, atendendo aos substantivos que figuram em primeiro lugar, e de lá se faz uma remissão ao substantivo ou adjetivo que ocupa a segunda posição (cf. § 39). Nesse terreno, a regularidade da seqüência sintática espanhola evita os difíceis problemas que traz a colocação das combinações nos dicionários das línguas germânicas. Em inglês, por exemplo, *Sea anemone* tem de alfabetar-se no A, posto que o termo principal é *anemone*, enquanto *sea* tem uma função adjetiva, como na equivalência espanhola “anemone de mar” (anêmona-do-mar). Contudo *Adam's Apple* terá de catalogar-se partindo do primeiro elemento, porque a “maçã-de-Adão” (pomo-de-Adão), não é um tipo de maçã, mas a proeminência que forma a laringe na parte anterior da garganta.

3. Os polinômios verbais. Diferentes critérios para sua ordenação alfabética.

As complicações começam quando passamos dos binômios, considerados até agora, para os polinômios verbais de três ou mais termos. Para estes casos está prevista uma hierarquia semântica que vai por ordem decrescente de importância: nome, verbo, adjetivo, pronome, advérbio etc., e sobre a base deste sistema redigiram-se regras para o uso do Dicionário acadêmico, que se podem ver em suas páginas preliminares. À primeira vista parece que estas regras vão resolver todos os problemas imagináveis, mas na prática não é assim. Quando uma parte da oração se substantiva e se apresenta em competição com um nome, pode arrebatar-lhe a primazia. Na locução “al caer de la hoja” (ao cair da folha), por exemplo, o Dicionário concede mais categoria ao infinitivo “caer” (cair) do que ao substantivo “hoja” (folha). No tocante aos verbos, também cabem vacilações. Os auxiliares não se têm em conta, por princípio; mas ocorre com freqüência que eles assumem um valor transitivo, e assim a frase “no haber más que pedir” (é só pedir) não se cataloga em “pedir”, mas em “haber” (haver). Contudo, outros verbos, que normalmente não são auxiliares, funcionam como se o fossem na formação de tempos compostos: “*tengo pensado escribir un soneto*” (*tenho⁶⁷ pensado em escrever um soneto*); “*ya llevo andado lo peor del camino*” (já tenho andado a pior parte do caminho); “*dejé dicho que no vendría a cenar*” (deixei dito que não viria jantar), etc. Tudo isto dá lugar a hesitações, que seria conveniente evitar.

Ora, se o critério do Dicionário não está a salvo de objeções, como acabamos de ver, não é menos discutível o adotado, em outros trabalhos lexicográficos, depois de madura reflexão. Reunidos recentemente vários filósofos norte-americanos tão capazes como R.S. Boggs, Lloyd Kasten e H.R. Richardson, sob a presidência do ilustre hispanista Hayward Keniston, para redigir um *Ensaio de Dicionário do espanhol medieval*⁶⁸, começaram estudando e fixando minuciosamente as normas a que todos teriam de submeter-se para assegurar a uniformidade do trabalho e, entre essas normas,

67 N.T.: O verbo *ter* em português funciona normalmente como auxiliar, porém, o mesmo não acontece no espanhol, onde o auxiliar dos tempos compostos não-marcado é “haber”. (haver).

68 *Tentative Dictionary of Medieval Spanish*, Chapel Hill, North Carolina, 1946.

está a que diz respeito ao lugar em que devem registrar-se as combinações e as frases. Diz assim: “As locuções não se colocarão atendendo ao verbo se contiverem um substantivo, pronome, adjetivo ou advérbio”. Aqui vemos notavelmente subvertida a hierarquia semântica que fixou a Academia, já que o verbo, que ali vinha em segundo lugar, passa agora a ocupar o último.

Para que se perceba praticamente o resultado diferente que produz a aplicação dos dois critérios que comparamos, bastará dar alguns exemplos. A expressão *tener a menos* (ter em menos), que a Academia registra em “tener” (ter), iria parar, segundo o sistema de Keniston, no advérbio “menos”. A frase *más vale tarde que nunca*, (antes tarde do que nunca), que o Dicionário acadêmico recolhe em “valer”, apareceria no advérbio “màs” (mais), por ser o primeiro, e por certo o menos significativo, dos três contidos na frase espanhola. A verdade é que o procedimento moderno não resulta, em nossa opinião, mais vantajoso do que o tradicional, e a consequência que daí deduzimos é que o problema que coloca a catalogação das expressões pluriverbais não teve até hoje solução satisfatória⁶⁹.

4. Tratamento das combinações (locuções) dentro do verbete correspondente.

Prescindindo no momento das sutis preocupações gramaticais com as quais nos temos deparado, vamos supor que a combinação que nos interessa já tenha encontrado o verbete em que deverá aparecer. Tomemos ao acaso *letra menuda* no sentido figurado de “astúcia e sagacidade”. Vamos ao verbete *Letra* e aí vemos todas as combinações binárias desta palavra ordenadas alfabeticamente, de acordo com a inicial do segundo elemento. Nossa locução precisa acomodar-se entre “letra mensajera” (ant. carta mensageira) e “letra mercantil” (tipo de letra usada antigamente entre comerciantes). O resultado deste sistema de colocação automática dá lugar a algumas séries heterogêneas como: *letra de mano* (a manuscrita), *letra de molde* (a impressa), *letra dental*⁷⁰ (a que se pronuncia aproximando a língua dos dentes), *letra dominical* (a que corresponde aos domingos no cômputo eclesiástico). Outra série: *letra minúscula* (a pequena); *letra muerta* (lei ou convênio que não se cumpre); *letra nasal*⁷⁰ (a que se pronuncia com a intervenção das fossas nasais); *letra negrilla* (a de imprensa mais grossa do que a comum). Vemos aqui misturadas as acepções retas com as transláticas, as familiares com as técnicas, a fonética com a imprensa e o calendário com a caligrafia. É satisfatório este método?

Já sabemos que a ordem alfabética é uma desordem organizada da qual não se pode prescindir por completo; mas se aceitarmos sua tirania para a sucessão dos verbetes, isto não exclui que no interior dos mesmos, tentemos agrupar as locuções de modo um pouco mais lógico do que o atual. Porque, além do mais, acontece que muitas das denominações registradas são sinônimas entre si e, para não repetir a explicação em cada caso, utiliza-se o divertido sistema das remissões que, como dizia Migliorini, é parecido ao jogo-da-glória: A *letra egípcia* está no início de uma coluna e nos remete à parte de baixo *negrilla* (negrito). *Versal* envia-nos a outra coluna, onde está *mayúscula* (maiúscula); *capital* obriga-nos a virar a página e nos remete outra vez a *mayúscula*, e o mesmo sucede com a *letra de caja alta* (letra de caixa alta).

Enquanto se espera uma solução melhor, poderíamos ensaiar a seguinte: já que no verbete mencionado existe uma definição circunstanciada de *Letra* como tipo de im-

⁶⁹ Veja-se o meu *Nuevo concepto del Diccionario...* p.152, nota.

⁷⁰ Hoje se diria mais precisamente: “fonema nasal”, “fonema dental”, etc.

prensa (acepção 4.^a), por que não pôr a continuação todos os nomes desses tipos, evitando as idas e vindas e permitindo ao leitor dar-se conta, numa única olhada, de todas as equivalências existentes? Deste modo teríamos: “*Letra mayúscula, capital, — versal* ou — *de caja alta* (aqui a explicação); — *minúscula* ou — *de caja baja* (idem, id.); — *cursiva, — bastardilla* ou — *itálica* (idem, id.); — *negrita, negrilla* ou — *egipcia* (idem, id.). Não seria isto muito mais razoável e instrutivo?

O mesmo podemos dizer a respeito da acepção fonética de “Letra”, que é a segunda do verbete. Neste caso teríamos uma vantagem adicional não desprezível: a de ver de relance quais são as denominações que faltam e quais são as que sobram..., se é que não sobram todas. Vejamos por quê. Tínhamos concordado em que o Dicionário deve recolher todo tipo de combinações estáveis e, por conseguinte deve excluir as que não o são. Na última edição acadêmica entraram alguns tecnicismos de fonética, como *africada, velar, dorsal, implosiva, apical* etc. As definições correspondentes, devidas a Navarro Tomás⁷¹, seriam cientificamente irrepreensíveis se não resultassem adulterados por um detalhe da técnica lexicográfica. Hoje lemos, por ex., “AFRICADO, DA, adj. Gram. Dícese del sonido o letra cuya articulación consiste...” (AFRICADO, DA. Adj. Gram. Diz-se do som ou letra cuja articulação consiste...) e no fim da definição “Ú. t. c.s.f.” (usa-se também como substantivo feminino). Isto, dito assim, não é verdade. Quando se fala de “uma africada”, não se subentende “som” nem “articulação”, mas “letra”. A redação correta do verbete deveria ter esta forma: “AFRICADO, DA. adj. Fon. Diz-se do som cuja articulação... 2. Diz-se da letra que representa esse som.” E aqui é onde está em seu lugar “Ú.t.e.s.f.”. Mediante este simples retoque e unificadas as definições, que agora não o estão, (numas, se parte do “som”; noutras, da “consoante”; noutras, da “articulação”, e noutras, “do som, da articulação e da letra” ao mesmo tempo), só faltaria levar *apical* para o verbete próprio a *A, dorsal* ao *D*, e assim sucessivamente, com o que se teria clarificado o verbete *letra*, aliviando-o de mais de vinte itens desnecessários... e inexactos.

5. Combinações ocasionais de caráter informativo. Devem figurar no Dicionário?

Dissemos anteriormente (III, 2) que as combinações estáveis são as únicas que devem preocupar o lexicógrafo, posto que, separados os seus elementos, perdem a significação que assumiram no conglomerado; mas vemos que alguns dicionários, entre eles o acadêmico, registram outros tipos de combinações que designaremos com o qualificativo de “ocasionais”. Os seus componentes unem-se e separam-se livremente quando convém e, conservando a sua respectiva significação, podem formar parte de outras combinações. Lembremos o exemplo de “arco dórico” (III, 2).

Se percorrermos o verbete *Buque* (navio) no Dicionário, observaremos que se dá ao trabalho de explicar-nos o que significa “buque de vapor” (navio a vapor), *de vela* (a vela), *de ruedas* (de rodas), *de hélice, de transporte, de cabotaje* (de cabotagem) *mercante, de guerra* etc. Trata-se evidentemente de combinações ocasionais, já que *de vapor* (a vapor) pode unir-se com o mesmo sentido a “buque” (navio), “barco”, “máquina” etc. Porém a combinação “caballo de vapor (cavalo — vapor) entra no grupo das combinações de caráter estável, porque a acepção de unidade de medida de potência equivalente a 75 quilógrâmetros não reside em “caballo” nem em “vapor”, mas na combinação dessas duas palavras.

71 N.T.: Tomás Navarro Tomás, grande foneticista espanhol, é autor entre outras obras de: *Manual de Pronunciación Española* e *Manual de Entonación Española*.

Além disso, o leitor pergunta-se por que toda esta informação está ausente do verbete “barco”, sendo assim que na linguagem comum é tão freqüente aplicar a “barco” como a “buque” (navio) a maioria das especificações citadas; e sua dúvida pode ir além: são necessários esses pormenores informativos numa obra lexicográfica? E nós, referindo-nos mais concretamente a um Dicionário histórico da língua, podemos também perguntar-nos: tem cabimento tantas e tantas explicações que *já não versam sobre as palavras, mas sobre as coisas*? Não se trata mais de um tipo de informação que tem seu lugar adequado nos dicionários de tipo enciclopédico? Não vamos decidir aqui a questão, já que a resposta pertinente às várias perguntas formuladas dependerá do caráter que se queira dar ao Dicionário. A título de antecedentes nos limitaremos a fazer constar que o atual Dicionário da Academia, verbete *Arco*, na acepção de arquitetura, nada menos que vinte e quatro espécies de arcos, desde o *adintelado* (o que acaba por degenerar em linha reta) até o *zarpanel* (arco que consta de várias porções de circunferência tangentes entre si e traçadas desde diferentes centros) enquanto o Dicionário de Oxford, com uma extensão vinte vezes maior que a do Espanhol, só traz a acepção geral correspondente.

ABSTRACT: Semantics and Lexicography interpenetrate mutually as Lexicography is not limited to collecting the words of lexicon, but attempts to describe word meanings and usage. The lexicographer also deals with the evolution of word significations to establish the gamut of meanings (“acceptions”) of a lexical sign. Casares establishes the concept of “acception” and discusses the problem of discrimination of meanings and their structure in polysemic words. Another critical point for the lexicographer is the recognition and correct identification of metaphorical meanings. Casares uses as an illustrative example the word Lat. > Sp. orden (Port. ordem), polysemic sign. He provides a chart of the semantic network concerning the evolution of the meanings of this word from the original Latin etymon to modern Spanish. Casares also treats the problem of lematization, i.e., the technical decision of opposing, as a dictionary entry, one or another lexical form. This choice arouses permanent debate among lexicologists over the manner in which compound words should be entered. How does the lexical categorization of a verbal polynomial take place and when? This problem becomes more complex owing to the chaotic tradition of many scripts, particularly in the case of “lexical syntagmes”. Casares advocates a dictionary which contains a frequency index of the usage of each word and of each meaning of a word.

KEY-WORDS: Semantics; lexicography; meaning; “acception”, onomasiology; semasiology; etymology; ordering of meanings; polysemy; metaphorical meaning; etymon; usage frequency; binary combination; verbal polynomial; treatment of lexical syntagmes.

DEFINIÇÃO LEXICOGRÁFICA EM SEMÂNTICA DESCRITIVA

Por Uriel WEINREICH*
Tradução de Maria Cecília P. Barbosa LIMA**

RESUMO: O autor analisa os vários aspectos do ato de definir uma palavra para um dicionário e os tipos de critério a serem adotados. Uma questão de relevância é a do tipo de metalingua que se deve adotar na redação de um dicionário e a metodologia empregada para definir o "definiendum", podendo-se empregar várias estratégias: o método analítico, o método sintético, o método denotativo, o método ostensivo ou de mostração, o método implicativo, ou contextual. O método escolhido dependerá da natureza do termo a ser definido: um referente concreto, uma noção abstrata, uma ação ou processo verbal, um instrumento gramatical etc. A sinonímia e a antonímia amplamente usadas nas definições têm também grande importância lexicográfica. O lexicógrafo, ou a equipe de dicionaristas que trabalham na confecção de um dicionário, nunca se deve esquecer que as suas definições devem valer para toda a comunidade lingüística a que ele se destina e assim usarem a linguagem comum a todos e não o(s) seu(s) idioleto(s) particular(es).

UNITERMOS: Descrição semântica; lexicografia; dicionário monolingüe; thesaurus; significado; definição; sinônimo; denota; definição ostensiva; definição lexicográfica; metalingua.

1. DESCRIÇÃO SEMÂNTICA E LEXICOGRAFIA

1.1. Os falantes de uma língua sentem intuitivamente uma relação entre certos pares ou conjuntos de palavras, que não se deve a nenhuma similaridade fonológica ou gramatical clara. Como falantes do inglês, podemos afirmar, sem muita hesitação, que em cada um dos seguintes grupos de três palavras, duas se relacionam mais estreitamente que a terceira: *up, high, small; open, eat, close; end, after, grass* [para cima, alto, pequeno; abrir, comer, fechar; fim, depois, grama]. Poderíamos, com toda probabilidade, chegar a um consenso quanto ao modo de completar uma proporção como: *son: daughter:: brother: _____* [filho: filha:: irmão: —]. Presume-se também que concordaríamos a respeito da ambigüidade de expressões como *She couldn't bear children* ou *(Was the weather bad or good?) It was fair.**** Fornecer um relato explícito de tais intuições é um bom modo de começar o estudo da semântica descritiva.

1.2. A descrição semântica de uma língua consiste na formulação, em termos apropriados, dos significados que as formas dessa língua têm, de acordo com o grau de interpessoalidade desses significados (cf. sec. 3). As formas cujos significados devem ser

* Pertenceu ao Departamento de Lingüística da Universidade de Columbia, foi Diretor da Language and Culture Atlas of Ashkenazic Jewry e Editor do The Field of Yiddish.

** Departamento de Letras Modernas — Instituto de Letras, Ciências Sociais e Educação — UNESP — 14800 — Araquara — SP.

*** N. do T. As interpretações possíveis seriam: 1.ª) a) Ela não podia suportar crianças. b) Ela não podia ter filhos. 2.ª) a) O tempo estava ruim ou bom? Estava bom. b) O tempo estava ruim ou bom? Estava mais ou menos.

descritos são de muitas espécies: usualmente são morfemas, freqüentemente, conjuntos de morfemas em construção, e, ocasionalmente, componentes submorfêmicos; mas podem também ser contornos prosódicos, classes de morfemas, classes de construção e processos gramaticais, caso sua seleção não seja automática.

Quando formas de uma língua são gramaticalmente combinadas, seus significados, como bem se sabe, interagem. As formas de tal interação semântica parecem ser, em larga escala, universais; caso, porém, haja uma ou outra específica de uma determinada língua (por ex., aquelas que dependam de formas gramaticalmente diferenciadas de predicação), também estas devem ser incluídas na descrição semântica de tal língua.

1.3. Não há procedimentos de descoberta conhecidos para descrições semânticas corretas. O melhor que se pode esperar é um teste de adequação que nos capacite a escolher entre descrições alternativas. Idealmente, uma descrição é adequada quando nos fornece meios claros para aproximarmos as intuições que os falantes nativos têm sobre as relações semânticas de palavras na sua língua. Vale dizer, exigimos que os significados claramente formulados de termos como *open*, *eat*, *close** — e não nossa intuição como falantes do inglês — nos permitam dizer que *close* é mais semelhante a *open* que a *eat* e assim realizar operações metalingüísticas dessa natureza. Além disso, as descrições semânticas deveriam ser tão completas quanto possível; deveriam ser consistentes; deveriam também ser simples e elegantes, embora os critérios de simplicidade sejam (como em outros campos) obscuros, pelo que se deve certamente cuidar para não sacrificar a consistência à elegância — um sacrifício que é o mal de muitos dicionários existentes.

1.4. A descrição semântica de termos isolados (em contextos apropriados) é o domínio costumeiro da lexicografia. O estudo de padrões recorrentes de relacionamento semântico e de quaisquer expedientes (fonológicos ou gramaticais) pelos quais os padrões possam se representados é, às vezes, chamado de “lexicologia”. Além de problemas descritivos, a lexicologia pode tratar de questões históricas e comparativas.

O produto do trabalho do lexicógrafo aparece de vários modos: dicionários monolíngües e plurilíngües, dicionários de sinônimos, “thesauri”, dicionários enciclopédicos, estudos de campos de palavras e similares. Há também outras abordagens à descrição semântica — testes de associação, testes semântico-diferenciais, contagens de freqüência — em que métodos experimentais e quantitativos têm papel relevante. Um esquema completo para a descrição semântica proporcionaria a integração dessas abordagens variadas. Este artigo, entretanto, se restringe à abordagem lexicográfica, à exceção do cap. 3, que discute um uso mais completo de procedimentos experimentais, até mesmo na lexicografia propriamente dita.

1.5. Falamos de “abordagem” lexicográfica e não de “método”, porque a própria lexicografia usa muitos métodos mas nenhum deles foi completamente explicado. A indiferença que a lexicografia demonstra para com sua própria metodologia é espantosa. Talvez os lexicógrafos sejam complacentes porque seu produto “funciona”. Mas é legítimo perguntar de que modo ele funciona, a não ser pelo fato de que os dicionários vendem bem. O fato de que na nossa cultura particular há grande demanda de dicionários monolíngües de definições, que incluem definições de palavras tão comuns que ninguém possivelmente iria procurá-las, é, por si mesmo, um interessante dado etnográfico. (Em que outro lugar professores secundários da língua materna se esforçam

* N. do T. *abrir, comer, fechar.*

por inculcar em seus alunos o “hábito do dicionário”?) Mas esse fato não é substituto para a investigação metodológica. A literatura existente está longe de ser satisfatória. Este trabalho se preocupa com algumas questões metodológicas em lexicografia sobre as quais os lexicógrafos têm, surpreendentemente, silenciado.

1.6. As reflexões puramente teóricas sobre a descrição semântica deveriam ser suplementadas por uma análise comparativa de dicionários existentes. O campo inglês exerce especial atrativo, uma vez que há tantos dicionários monolíngües concorrentes para se comparar. Deve-se, porém, ter em mente que mesmo os dicionários concorrentes se copiam uns aos outros. No concernente a considerações teóricas, estamos, infelizmente, adstritos às indicações, na maioria das vezes, de uma única tradição lexicográfica, não só nos países de língua inglesa, mas em todo o mundo ocidental. Isso torna difícil distinguir entre similaridades de estrutura semântica em duas línguas e similaridades em convenções lexicográficas. Quão estimulante seria ter como indicações os produtos de uma lexicografia popular em uma sociedade não ocidentalizada!

1.6. Não é só a amostragem de tradições lexicográficas que é pequena, mas as línguas européias mais importantes, assim como seus dicionários, são provavelmente bem atípicos. O crescimento explosivo da quase internacional linguagem especializada da ciência, o prestígio e a difusão do conhecimento científico na sociedade, e a manutenção de acesso a períodos iniciais da língua através do cultivo da literatura, tudo isso fez com que não só os dicionários das línguas ocidentais mais importantes se tornassem maiores, mas também com que essas próprias línguas, como objetos de descrição semântica, se tornassem, também, de certo modo, maiores.

A hipertrofia de uma língua como o inglês pode ser caracterizada de vários modos. Um deles é o simples tamanho de seus dicionários, ou o fato de que seus dicionários estão atravancados de palavras rotuladas de *Arc.*, *Obs.*, *Rare*, *Zool.*, *Astron.**, e assim por diante. Um modo menos óbvio fica ilustrado pela seguinte experiência: apresentou-se a um grupo de estudantes formados este conjunto de oito sinônimos: *crabby*, *gloomy*, *glum*, *morose*, *saturnine*, *sulky*, *sullen*, *surly*** . Todos disseram que “conheciam” as palavras e que não havia nenhum par delas com significado perfeitamente idêntico. Foram então apresentadas a eles as definições correspondentes encontradas no *Merriam Webster New Collegiate Dictionary* com o pedido de que juntassem cada definição ao termo definido. Os resultados foram medíocres.

1.6. Podemos chegar a algumas sugestões recomendáveis para a reforma da lexicografia, mas a tarefa de descrever novamente uma língua como o inglês na proporção da edição não-simplificada do *Webster New International Dictionary* seria tão desorientadora que desencorajaria qualquer empreendimento decisivo. Entretanto, talvez seja justificado pensar-se naquela parte do vocabulário inglês (ou francês, ou russo etc.) que se encontra nos níveis de maior frequência (por ex., as 1.000 ou 10.000 palavras mais frequentes), como sendo análoga ao vocabulário total de uma língua menos hipertrofiada. Com tal redução, a descrição semântica do inglês seria mais semelhante à descrição semântica da maioria das línguas do mundo. (Ver também sec. 2.6.)

1.7. A lexicografia monolíngüe, como qualquer outra abordagem da descrição semântica, pressupõe uma teoria específica do significado. Dentre as diversas variedades

* N. do T. O mesmo pode ser dito dos dicionários de língua portuguesa. Um dicionário de uso contemporâneo da língua como o *Novo Aurélio* está também hipertrofiado com termos já caídos em desuso, ou em vias de desaparecimento, assim como numerosos tecnicismos e regionalismos.

** Todas essas palavras têm um núcleo de significação comum correspondente aproximadamente a sombrio, triste, mal humorado.

de teorias semânticas, os dicionários de definição parecem basear-se num modelo que supõe a distinção entre o significado propriamente dito (“significação”, “compreensão”, “intenção”, em várias terminologias) e as coisas significadas por qualquer signo (“denotação”, “referência”, “extensão” etc.). Essa dicotomia, que é de origem medieval pelo menos, foi discutida na lingüística do século XIX sob o título de “forma interior” e na lingüística (pós) saussuriana, relacionada a substância do conteúdo vs forma de conteúdo (“valeur”,* etc.); aparece em filosofia moderna sob várias roupagens, como, por exemplo, nos trabalhos de J. S. Mill, Frege, Peirce, Morris, como também nos de Carnap e Quine, que estudaram a possibilidade de eliminar tal distinção. Várias teorias semânticas alternativas, por outro lado, parecem ter sido excluídas como bases da lexicografia. Por exemplo:

a. O “significado lingüístico” de um termo é a probabilidade de que ele ocorrerá, calculada a partir do contexto de outras formas do mesmo discurso.

A lexicografia convencional não está interessada em nenhum “significado lingüístico” separado do “significado cultural”. Do ponto de vista de um dicionário, a probabilidade de ocorrência de um termo, se é que se pode calculá-la, mede, quando muito, sua banalidade ou sua falta de sentido. Não seria suficiente analisar o discurso como sinais alternantes de banalidade.

b. Significados como estados psíquicos são inacessíveis à observação, e a semântica descritiva deve esperar até que maiores progressos na neurologia os torne acessíveis.

Esse “reducionismo” messiânico, característico, por exemplo, de Bloomfield, é teoricamente estranho à descrição lexicográfica. Com efeito, mesmo que tivéssemos especificações neurológicas, digamos, das emoções, a descrição semântica de termos relativos à emoção poderia ser continuada independentemente, do mesmo modo que a descrição semântica de termos de cores pode ser altamente independente da já conhecida psicofísica da visão.

c. O significado de um termo é o seu uso na língua.

Esse slogan da filosofia lingüística britânica pode oferecer uma saída para certos impasses lexicográficos em relação à polissemia, mas, como teoria geral, exigiria que renunciássemos aos dicionários e nos satisfizéssemos, no máximo, com concordâncias.

Deve-se mencionar uma quarta teoria do significado que é complementar à da lexicografia convencional (pelo menos quando aparece em suas formas menos agressivas). Essa é a teoria do poder emotivo das palavras, particularmente da capacidade que um termo tem de “evocar” outros termos por meio de um salto através de suas estruturas semânticas interiores. Alguns dicionários reconhecem informalmente essa possibilidade quando rotulam certas palavras de “depreciativo”, “afetivo” etc., mas o uso que se faz disso é casual. Já que os psicólogos demonstraram a existência de impressionantes normas interpessoais de força emotiva de muitas palavras, essas descobertas podem merecer uma inclusão sistemática em descrições amadurecidas de línguas.

1.8 O propósito da discussão que se segue é duplo: explicar e criticar o que os lexicógrafos fazem quando definem palavras.

Passamos agora à análise crítica de definição lexicográfica como um expediente do arsenal da semântica descritiva.

* Em francês no original.

2. DEFINIÇÃO

2.1. *Estrutura de Definições.* Para se fazer uma reconstrução racional daquilo que um dicionário faz, seria útil conceber-se o significado de um termo como o conjunto de condições que devem ser preenchidas para que o termo seja denotativo. Nessa perspectiva, uma formulação do significado requer uma lista dessas condições necessárias à conotação.

Uma objeção óbvia a dicionários é que as condições em tal “conjunto” não são necessariamente discretas, e que qualquer análise é, portanto, muito artificial. É o clássico problema gestaltico. Uma língua natural, por ser articulada, não é uma metalíngua adequada para a análise de suas gestaltes semânticas, e nenhuma metalíngua mais adequada foi estabelecida. (Compare-se nossa eficiência em reconhecer rostos com nossa incapacidade de descrevê-los com palavras). Mas não precisamos restringir a lexicografia, exigindo que uma definição seja a reprodução perfeita de um significado, ou que o “definiendum” seja distinguível, do “definiens” por mera inspeção. Muito menos podemos esperar da lexicografia de uma língua natural que o “definiendum” seja literalmente substituível pelo “definiens” num discurso normal. O que estamos autorizados talvez a exigir de uma lexicografia racionalizada é que a codificação da gestalt do sentido no código descontínuo da metalíngua definidora seja realizada sob certas restrições de forma e que a definição resultante seja aceitável para leigos que sejam representativos da língua e que tenham condições de entender as restrições formais que a governam.

A consistência da lexicografia poderia ser melhorada se os autores de dicionários se ativessem ao pressuposto de que os termos de uma língua são, em geral, complementares. (Ver também sec. 2.5) Essa pressuposição sugere que o caso mais importante para se tratar em descrição semântica é aquele em que a significação de um termo acaba onde a de outro se inicia. Em geral, a descrição semântica deveria ter como objetivo não definições “absolutas”, mas definições que separem o significado de um termo daquele de outros termos de significados semelhantes (sinônimos). A circularidade que daí resulta deveria ser admitida francamente, não como um vício, mas como um princípio diretor da lexicografia.

O agrupamento de sinônimos numa seqüência resulta num “thesaurus” como o de Roget. Sem dúvida, aspectos significantes da estrutura do vocabulário podem ser estudados com base num único “thesaurus”, do mesmo modo que traços de uma área geográfica podem ser estudados mesmo num mapa sem indicações. Por exemplo, a lista de palavras num “thesaurus” pode ser interpretada como uma série de projeções (mapeamentos) unidimensionais da estrutura semântica multidimensional do vocabulário. As dimensões do vocabulário poderiam ser então estudadas em termos de intersecções de linhas de sinônimos (= referências remissivas entre parágrafos do “thesaurus”). Essa lexicografia perfeitamente “formal” seria provavelmente tão insatisfatória quanto uma fonologia completamente formal, que não se refira à substância fônica. Ao contrário de um mapa normal, que é a projeção bidimensional de um espaço tridimensional, o “thesaurus” enquanto descrição semântica mapeia um espaço de muitas dimensões (na realidade, num número desconhecido de dimensões). No caso de mapas comuns, entendemos perfeitamente seu mecanismo semiótico e podemos torná-los tão precisos e unívocos quanto queiramos. No caso de uma língua como um mapa, não temos conhecimento a priori do espaço-objeto, da “escala” elástica do mapa, ou de algumas outras propriedades da representação. É, pois, importante, se quisermos ir longe, “orientar” a descrição para sua substância semântica. Do contrário, podemos

aprender que *verde* fica entre *azul* e *amarelo*, assim como *escarlata* está entre *vermelho* e *azul*, sem saber se *verde* e *escarlata* cobrem segmentos iguais ou desiguais do espectro solar. Para determinar isso, teríamos que orientar a linha lingüística de sinônimos para o aspecto físico.

Indicar as direções de um mapa corresponde, em descrição semântica, à demonstração de alguns “denotata”. Isso também é conhecido como “definição ostensiva”. Para ser eficiente, uma definição ostensiva deve produzir não só exemplos positivos de “denotata” de um termo, mas também contra-exemplos negativos (= exemplos positivos de “denotata” de um termo sinônimo). Assim, uma única amostra de vermelho não dá uma definição ostensiva suficiente de *vermelho*; os limites do significado devem ser estabelecidos mostrando-se várias espécies de vermelho e também amostras do que é semelhante ao vermelho, mas não é vermelho (por ex. laranja, rosa, preto avermelhado). Os “denotata” devem ser apresentados fisicamente ou então por meio de nomes; assim, uma série de amostras coloridas ou os termos *vermelho*, *azul*, *amarelo* etc. constituem duas variedades de definição ostensiva de *cor*.

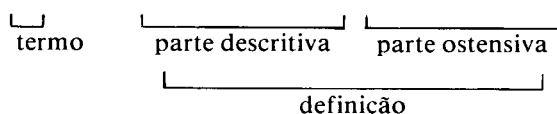
Com base na análise acima, podemos construir uma forma canônica de definição lexicográfica. Seja X o termo de uma língua cujo significado deve ser descrito e sejam X', X'' etc, sinônimos de X (isto é, termos semelhantes, mas não necessariamente idênticos em significado); sejam c₁, c₂ etc., condições que devem ser preenchidas quando se quer que X denote; e sejam d₁, d₂ etc. amostras de “denotata”.

Uma definição lexicográfica tem então a seguinte forma:

X denota se c₁ e c₂ e ...c_n; por exemplo, d₁ ou d₂ ou ...d_n.

X' denota se c'₁ e c'₂ e ...c'_n; por exemplo, d'₁ ou d'₂ ou ...d'_n.

•
•
•



A forma clássica de definição “per genus et differentias” pode ser transformada em forma canônica; ver sec. 2.64.

Podemos agora apreciar algumas propriedades de definições lexicográficas (sec. 2.2.-2.4 e 2.7) e de sistemas de definições (sec 2.5 e 2.6).

2.2. *Especificidade de Definições.* De duas definições, será mais específica a que apresentar um número maior de condições. Consideremos o seguinte par esquemático:

X denota se c₁ e c₂

X denota se c₁ e c₂ e c₃

Podemos dizer que a última definição é a mais específica. Mas só faz sentido comparar a especificidade de definições se elas compartilham as mesmas condições, como o par acima compartilha c₁ e c₂. Esta outra definição, por exemplo:

Y denota se c₄ e c₅ e c₆ e c₇, não é comparável ao par anterior quanto à especificidade.

Os dicionários modernos parecem ser, em geral, cuidadosos em tornar suas definições adequadamente específicas. Definições do tipo exemplificado por:¹

VERST 'a Russian measure of length'
['medida russa de comprimento']

são evitadas por serem insuficientemente específicas, desde que não diferenciam VERST de outras medidas russas de comprimento.

Por outro lado, uma definição como:

TRIANGLE 'a figure that has three sides and three angles, the sum of which is 180°'
[TRIÂNGULO 'figura que tem três lados e três ângulos cuja soma é 180°']

é evitada por ser demasiadamente específica, desde que TRIÂNGULO é suficientemente definido pelo número de lados. Definições excessivamente específicas podem ser chamadas de enciclopédicas.

Um exame dos dicionários Merriam-Webster mostra que houve muito cuidado no controle da especificidade das definições, mas definições científicas especializadas (precedidas de rótulos como Bot., Zool. etc.) são, freqüentemente, enciclopédicas. Com referência a nomes de plantas e animais, definições enciclopédicas ocorrem mesmo sem qualquer rótulo, como por exemplo:

CARROT 'a biennial plant' (*Daucus carota sativa*) with a usually orange-colored, spindle-shaped 'edible root...'

[CENOURA 'planta bienal (*Daucus carota sativa*) com raiz comestível de forma alongada, de cor geralmente alaranjada...']

Aqui, 'biennial' pode bem ser demasiadamente específica para uso leigo (de não-botânicos) do inglês.

Se dois termos têm definições de especificidade adequada mas diferentes, embora comparável, pode-se dizer que os significados dos termos são de especificidade diferente. 2.3 *Crêterios das Condições*. Uma condição para a denotação pode ser considerada como dotada de critério se, uma vez não preenchida, nossos informantes se recusarem a aplicar o termo em questão. Assim, se uma suposta cadeira mostrasse evidência de não ter sido feita para que se possa sentar nela (por ex., só duas pernas, ou pernas de papelão), os informantes poderiam não querer chamá-la de CADEIRA, e concluiríamos que "ter sido feita para que se possa sentar nela" é uma condição de critério para esse termo. Em línguas naturais, os significados freqüentemente incluem condições que não são tão definitivamente baseadas num critério.

Os dicionários indicam informalmente o caráter duvidoso dos critérios por meio de marcadores especiais. Em dicionários da língua inglesa a palavra "specially" [especialmente] precedendo a afirmação de uma condição indica que o que se segue é menos baseado num critério que o resto; "usually" [geralmente] parece indicar um grau ainda mais baixo de critério. Alguns dicionários usam "loosely" [não exatamente] para indicar partes não baseadas em critério de definições ostensivas, como por exemplo:

COLOR'... a particular hue... including loosely black, white...'

[COR'... um tom específico... incluindo indistintamente preto, branco...']

Mais tais marcadores não ocorrem somente no jargão do lexicógrafo. As línguas comuns parecem conter termos para fins relacionados, como por exemplo, o inglês TRUE, como em *a true patriot* (= "tomar as condições de patriotismo criteriosamente").

Na prática, os lexicógrafos determinam o grau dos critérios por meio de introspecção individual, ou por “introspecção coletiva” (o compilador verifica os rascunhos das definições fornecidas pelos elementos da equipe); no caso de termos científicos, eles chegam mesmo a consultar especialistas. Mas se a lexicografia deve ser descritiva, todos os casos duvidosos e interessantes de critérios deveriam ser verificados pelo exame de uma amostra de usuários da língua (cf. sec. 3).

2.4. *Natureza das condições.* Os significados diferem no que diz respeito à natureza das “operações” exigidas para que se certifique de que as condições para a denotação são preenchidas. Infelizmente, a lexicografia parece inclinada a suprir esse fato pela elegância fácil e uniformidade espúria do seu estilo de definição.

As “operações” vão desde a inspeção imediata por meio dos órgãos dos sentidos até procedimentos experimentais ou dedutivos mais complexos. Supondo que CARROT [CENOURA] seja adequadamente definida como ‘edible, tapering, orange-colored root’ [‘raiz comestível alongada, de cor alaranjada’], sua forma alongada e coloração alaranjada são imediatamente aparentes ao sentido da visão; sua edibilidade talvez exigisse operações mais complexas (com certos riscos para os dentes e o estômago das pessoas, o que tornaria preferível dispensar a edibilidade como sendo elemento não baseado em critério, nesse caso). As significações podem ser classificadas de acordo com o órgão do sentido posto em ação em suas várias condições, mas há certamente significados que não recorrem a nenhuma evidência sensorial, ou, se o fazem, será a uma evidência sensorial misturada com outras espécies de evidência, como por ex., EARLY [CEDO] ou WONDER (... to query in the mind) [‘... procurar na mente...’].


Classificar, com base fenomenológica, um vocabulário é uma tarefa absorvente e viável, desde que se tenha um corpus de definições adequadas. O primeiro passo em tal investigação seria escolher um corpus de definições adequadas num dicionário convencional, no qual espaços interessantes e significativos entre porções do vocabulário são preenchidos com palavreado elegante.

Muitos dicionários preferem formular condições sob a forma de experimentos dos físicos, pelo menos como um paralelo às operações de um leigo. Poucos, porém, são consistentes. Desse modo, o *Concise Oxford* define BLUE [AZUL] só ostensivamente, ao passo que COLOR [COR] é definida por procedimentos psicofísicos. Vale a pena ponderar sobre se não se deveria permitir aos falantes de uma língua sugerirem o tipo adequado de condições para os significados dos vários termos de sua língua. Reconhecidamente, na nossa sociedade, as condições sugeridas por leigos se pareceriam, frequentemente, com experimentos físicos, como no caso de EDIBLE [COMESTÍVEL] e OUT-OF-ORDER [ESTRAGADO], termo esse que envolveria como condição ‘não fazer o que deveria fazer quando conectado a um circuito elétrico e depois ligado’. Mas haveria esperança de se evitar a irrelevância óbvia de informação especializada de ótica ao se definir, digamos, AZUL.

2.5. *Continuidade do Sistema de Definição.* A cobertura semântica do mundo por uma língua seria estritamente contínua se, para cada termo X definido como ‘ c_1 e c_2 e ... c_n ’, houvesse nessa língua um sinônimo X’ adequadamente definido como ‘ c'_1 e c'_2 e ... c'_n ’ (isto é, mudando-se uma das condições).² É, no entanto, verdade que uma língua não é semanticamente contínua em todos os níveis de seu vocabulário. Por exemplo, mesmo que a forma alongada de uma cenoura pareça ser uma condição baseada no critério de definição da palavra inglesa CARROT [CENOURA], não parece haver nenhu-

ma palavra inglesa que designe um legume em tudo semelhante a uma cenoura exceto no fato de ser de forma esférica.

Seria difícil encontrar um tópico mais crucial da lexicografia do que a comparação de continuidades e discontinuidades semânticas das línguas. De fato, enquanto algumas discontinuidades, como entre termos para espécies biológicas discretas, parecem ser (muito, mas não totalmente) determinadas pela natureza, muitas outras são específicas de uma língua-e-cultura.

Tem-se descrito uma discontinuidade como um caso em que, se alteramos uma definição mudando alguma c_1 para c'_1 , não há nenhum termo na língua que seja definido pela definição alterada. Isso dá importância ao modo detalhado pelo qual as descrições são formuladas. Considere-se a seguinte definição: X ' c_1 e c_2 e c_3 e c_4 '. Suponhamos que mudando-se c_1 para c'_1 não se produza nenhum termo definível na língua, mas mudando-se c_1 para c'_1 e, ao mesmo tempo, mudando-se c_2 para c'_2 se produza uma definição de um termo existente, X'. Em tal caso, embora por demais esquematizado, considerações intralingüísticas ad hoc sugerem que ' c_1 e c_2 ' deveriam ter sido tomadas como uma única condição. Por exemplo, supondo que BÉET [BETERRABA] seja como CARROT [CENOURA], com exceção do fato de que é vermelha (e não laranja) e tem a forma  (e não alongada), e supondo-se que não haja raízes comestíveis "intermediárias", não deveriam cor-e-forma ser consideradas como uma única condição? Nesse fácil exemplo, a objeção pode ser refutada pelo argumento de que há não só uma razão extralingüística para que se tratem cor e forma separadamente, mas também uma boa razão lingüística, dados os vocabulários distintos em inglês para forma e cor. Em outros casos, porém, especialmente onde o atrativo das condições se exerce não sobre os sentidos, sua diferenciação pode causar problemas e ter que ser resolvida levando-se em conta a congruência do padrão (um problema não estranho aos lingüistas de fonologia e da gramática).

Outro corolário diz respeito à escolha de uma metalingua de definição. Os dicionários de definição mais ambiciosos têm sido provavelmente os monolíngües, i.e., a metalingua na qual as condições para denotação são formuladas (e, se apropriada, na qual são nomeados os exemplos) tem sido a mesma que a língua objeto descrita. A situação não seria diferente num dicionário de definição bilíngüe em que ambas as línguas pertencessem à mesma cultura e estivessem aproximadamente ao mesmo nível de desenvolvimento léxico (cf. sec. 1.62) — e.g., língua objeto-inglês, metalingua-francês. Considerem-se, porém, duas línguas como o alemão e o rético, a segunda das quais é menos desenvolvida (isto é, a grosso modo, tem menos palavras): os problemas de formular condições são muito diferentes para o *Dicziunari Rumansch Grischun* (língua objeto, rético, metalingua, alemão) do que seriam para um dicionário imaginário alemão-rético. Vale a pena indagar em que medida o grau absoluto de desenvolvimento lexical de uma língua, e, em dicionários de definição bilíngües, também o grau de desenvolvimento de uma língua relativa à outra, afetam a eficiência de uma língua como meio de formular definições. Um grau menor de desenvolvimento provavelmente ajuda a evitar a quebra não justificada de certas condições, mas também atrapalha certas quebras necessárias.

Finalmente, o conceito de continuidade e discontinuidade semânticas pode fornecer uma base para o estudo da densidade desigual de vocabulário dentro de uma língua ou entre línguas. Essa noção apareceu relacionada a "espectros" semânticos restritos e amplos, como por exemplo, quando se descobriu que as línguas eram lexicalmente ri-

cas em domínios associados a “temas” culturais (camelos árabes, neve esquimó, cortesia alemã medieval, pobreza iídiche). A riqueza poderia ser definida como um alto grau de continuidade semântica em seqüências de definições de especificidade relativamente alta (cf. sec. 2.2).

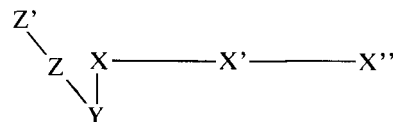
2.6. *Estrutura e Economia Lexicais.*

2.6.1. Na sec. 2.1. apresentamos as noções de sinônimo e condição como sendo interdependentes, uma vez que um sinônimo de X era chamado de termo definível quando uma das condições para X era alterada. Na sec. 2.5., mostramos, ao falar de descontinuidades semânticas, que sinônimo e condição, embora interrelacionados, não são fatores total e necessariamente dependentes um do outro em cada caso individual. Esse aspecto mais sutil pode ser agora levado adiante.

Vamos definir um conjunto sinonímico de primeiro grau como um conjunto de sinônimos que diferem numa única condição. Um par de antônimos seria um conjunto sinonímico de primeiro grau de dois membros. Conjuntos sinonímicos de maiores graus poderiam ser descritos do mesmo modo. Poderíamos, também, esquematizar as diversas e complicadas intersecções entre conjuntos sinonímicos. Por exemplo, dados os seguintes termos:

X 'c₁ e c₂ e c₃'
 X' 'c'₁ e c₂ e c₃'
 X'' 'c''₁ e c₂ e c₃'
 Y 'c₁ e c'₂ e c₃'
 Z 'c₁ e c'₂ e c'₃'
 Z' 'c'₁ e c'₂ e c'₃'

Podemos formular a intersecção dos conjuntos em três dimensões:



Seria, porém, útil ressaltar que a explicação de lexicografia desse modo pode ser proveitosa quando os elaboradores de dicionários aplicarem os mais severos padrões de economia verbal na formulação de definições. A proliferação de vocabulários de definição característica de muitos dicionários torna-os fontes inadequadas de material para a pesquisa lexicográfica crítica. Assim, se examinarmos a definição de END no Webster Collegiate: ‘a limit or boundary; esp., a limiting region or part... The extremity or conclusion of any event or series of events... The extreme or last purpose...’*, há poucas esperanças de se detectar a sinonímia de primeiro grau com BEGINNING ‘the commencement; the start... a point in space or time at which a thing begins...’** ou a sinonímia de primeiro grau, ou de grau um pouco mais alto entre END [FIM] e AFTER [DEPOIS DE], END [FIM] e (NOT) MORE [(NÃO) MAIS] etc.

A lexicografia convencional acredita aparentemente que a metalingua de definição deveria conter, pelo menos, a língua objeto toda. Essa crença pode ser injustificada. Poderíamos, idealmente, almejar uma metalingua “absoluta” que fosse inteiramente

* N. do T. “um limite ou fronteira; esp. uma região ou parte limitadora...”

A extremidade ou conclusão de qualquer escrito ou série de eventos... A finalidade última ou extrema...”

** N. do T. “o começo; o início... um ponto no espaço ou tempo no qual uma coisa começa...”

independente da língua objeto, ou de qualquer língua natural. Mas, uma vez que esse ideal em semântica é ilusório (em contraste com a fonética), deveríamos procurar modos de tornar a metalingua menos rica, e não tão rica, ou mais rica do que a língua objeto.

Um modo de reduzir progressivamente a riqueza da metalingua de definição seria exigir que a definição de um termo X fosse formulada somente com palavras de frequência maior que a de X. Outra limitação, menos rigorosa mas talvez mais prática, seria verificar, por ensaio e erro, em que medida a metalingua pode ser economizada se se concordar em usar o menor número possível de palavras diferentes em definições, mas usar cada uma o mais frequentemente possível. A redução é semelhante à buscada na construção do Inglês Básico, embora essa tarefa puramente pragmática tenha produzido, paralelamente, pouca teoria. Não há motivo para se supor que as aproximadamente 800 palavras do Básico representem a metalingua mínima para definir o inglês; poder-se-ia quem sabe reduzi-lo para alguém dos limites do Básico através de fórmulas como, por exemplo, (y é o) FIM (de x) 'não há x depois de y'; (z) AJUDA (x a fazer y) 'x não pode fazer y sozinho, x pode fazer y com z'; (x está) CHEIO 'alguma coisa está dentro de x, nada mais pode estar em x' etc.

Pode até ser que quanto mais a metalingua for reduzida, mais frequente e mais complexa se torne a necessidade de se recorrer à definição ostensiva. Mas essas relações não foram investigadas, e assim o problema de otimização não pode ser seriamente formulado.

2.6.2. Se uma redução controlada da metalingua fosse realizável pelos caminhos aqui sugeridos, o vocabulário de uma língua poderia considerar-se constituído de estratos, do seguinte modo:

Estrato 0: termos definidos só circularmente e por definição ostensiva.

Estrato 1: termos cujas definições contêm somente termos do estrato 0, mas sem circularidade.

Estrato 2: termos cujas definições contêm termos do estrato 0 e do estrato 1, sem circularidade.

Estrato *n*: termos cujas definições contêm só termos dos estratos 0, 1, 2, ... *n*-1. Poder-se-ia então dizer que a língua de definição, assim estratificada, incluiria a língua objeto toda, exceto o vocabulário do estrato mais externo, *n*.

Esse esquema é provavelmente idealizado, uma vez que há pouca base para se supor que o vocabulário de uma língua natural produzirá estratos discretos — particularmente se a circularidade puder ser excluída de todos os estratos, menos o estrato 0. Mas o esquema tem realmente algo intuitivo e poderia ajudar a reconstruir a organização de um "thesaurus" como o de Roget. Sugere, ainda, um meio de se distinguir entre os elementos mais ou menos "básicos" de um vocabulário.

2.6.3. Vários modos alternativos de se fraccionar um vocabulário já foram sugeridos. Um deles aplicaria à língua natural a distinção lógica entre termos (descritivos) e operadores (lógicos). De fato, um processo de definição austero poderia reduzir algumas das conjunções do inglês a termos mais primitivos (UNLESS [a menos que], IF [se], e talvez mesmo AND [e] a 'not' [não], 'or' [ou]). Mas alguns termos ingleses conteriam, em suas definições, uma mistura de operadores e descritores (e.g. NEGATIVE 'resulting from the *no*-operation' [NEGATIVO 'resultante de uma não-operação], e o estrato 0, como definido acima, conteria, certamente, termos descritivos e também lógicos. Outra dicotomia é a que existe entre símbolos e índices (elemen-

tos dêiticos, “shifters”, “indicações egocêntricas”)*. Esses dois tipos de termos apresentam, de fato, mecanismos semióticos diferentes, mas, por outro lado, a separação entre eles não é sempre nítida (assim, os advérbios HOME [em casa ou para casa] ou ALONG [junto com] contêm um elemento de retro-referência ao sujeito da oração que pode ser “I” [eu]; além disso, seria difícil conceber-se um estrato central que consistisse tão-somente nas palavras índices básicas de uma língua.

2.6.4. Seria conveniente, a esta altura, considerar as definições de gênero-diferença. Um modo de transformar uma definição do tipo clássico numa forma canônica seria considerar o nome do gênero como c_1 , e.g.: BENCH ‘a seat which is long...’ [BANCO ‘um assento que é longo...’], como equivalente a

BANCO denota se (c_1) é um assento, (c_2) é longo, ... Nesse caso, ‘assento’ acrescenta informação de critério e, o que é mais, ASSENTO poderia, por sua vez, ser definido sem ‘banco’ como “denotatum” de amostra. Por outro lado, se definirmos o inglês AZURE como ‘azul celeste’, estamos usando um nome de gênero, ‘azul’, que é, ele próprio, definido ostensivamente em AZUL ‘a cor do céu, do mar profundo’ etc... Em outras palavras, alguns nomes de gênero são muito menos úteis como condições que outros. Expressões como ‘uma coisa que ...’, ‘pessoa que...’ etc. ajudam tão pouco a transformar em forma canônica as definições que as contêm que melhor seria deixá-las de lado.

É provável que línguas diferentes difiram quanto à “profundidade” em que as hierarquias de gêneros possam ser construídas. Em inglês, a economia da definição é diferente da de uma língua que não tenha nenhum termo correspondente à *qualidade*, e na qual COR não possa ser definida como “uma qualidade que...”

2.7. *Sintaxe da Metalingua de Definição*. Além do problema de vocabulário metalingüístico, podemos considerar a sintaxe da metalingua. A lexicografia convencional observa certas regras nesse campo. Por exemplo, os dicionários estão limitados pela restrição de que a definição deve ser um sintagma endocêntrico, sujeito às regras da sintaxe comum da língua-objeto, um sintagma funcionalmente equivalente ao termo definido. Assim, FLOW n. ‘act of flowing’, (to) FLOW v. ‘to move or circulate, as liquid...’, [Flutuação n. “ato de flutuar”, Flutuar v. “mover-se ou circular como um líquido...”] etc. Embora essa convenção possa conduzir à elegância, ela parece dever-se a uma pretensão à permutabilidade entre o termo e a sua definição, o que é absurdo para as línguas naturais.

No entanto, apesar da desnecessária pretensão de interpermutabilidade, os dicionários, inesperadamente, passam para uma metalingua diferente, e.g., NOT ‘an adverbial particle expressing negation’ (New Collegiate Webster) [NÃO ‘partícula adverbial exprimindo negação’], ou GOOD ‘a term of general or indefinite commendation’ (Shorter Oxford) [Bom “termo de apreciação geral ou indefinida”]. Uma abordagem crítica à lexicografia exige que determinemos se e em que ocasiões tais mudanças de metalingua são legítimas.

Uma conseqüência indesejável das limitações que os lexicógrafos impõem à sua metalingua é que termos relacionais são tratados como se fossem absolutos; num fraseado lógico, predicados que requerem mais de uma variável são tratados como predicados de um só lugar. Assim, a definição de BETWEEN [ENTRE] encontrada no *Webster Collegiate* não dá nenhuma indicação de que esse termo exige um objeto da forma x e y , ou, pelo menos, um substantivo plural; tampouco fica isso evidente no

* N. do T. O *Dictionnaire de Linguistique* (Larousse, 1973) traz um artigo sobre *embrayeurs* (shifters), embreadores, onde esse termo é definido como aquelas unidades do código que remetem necessariamente à mensagem e seu contexto para serem plenamente descodificadas tais como: *eu, papai, aqui* os tempos verbais etc.

fraseado da definição: 'in the space or interval which separates' ['no espaço ou intervalo que separa']. ('Separates' pode ter um único objeto). Na definição de END [FIM] a que nos referimos no fim da sec. 2.6.1., não há indicação da natureza relacional do termo (*end of...*) [fim de]. Algumas das economias frisadas neste trabalho poderiam ser alcançadas se se desse aos dicionários maior liberdade para mostrar a natureza relacional de muitos termos de uma língua, e.g., usando variáveis algébricas (x, Y, \dots) no termo e também na definição, como foi ilustrado para END [FIM], HELP [AJUDAR], FULL [CHEIO], na sec. 2.6.1.

2.8. *Uma Classificação de Definições*. Uma das discussões mais completas de definição em forma não técnica está contida numa monografia de Richard Robinson (*Definition*, Oxford, 2.^a impressão, 1954). Poderia ser útil concluir a presente análise por meio de uma comparação de seus resultados com os de Robinson. Consideremos os sete métodos de definição de "palavra-coisa" que Robinson delineou, colocando-os na estrutura do nosso modelo.

a) O Método de Sinônimos. Consiste em dizer "que a palavra definida significa o mesmo que alguma outra palavra..." Sinônimos perfeitos são, porém, raros em línguas naturais, especialmente na maioria não-hipertrofiada, e são um meio de descrição ineficiente e não-fidedigno. Por outro lado, o arrolamento linear de sinônimos não-perfeitos de um termo não limita efetivamente seu significado. Portanto, o método a), embora comum em dicionários elegantes, deveria ser descartado.

b) O Método de Análise e c) O Método de Síntese. Essa distinção não é suficientemente clara para ser aplicável a línguas naturais; mesmo em sua própria apresentação, Robinson sente que é necessário refutar uma visão semelhante, mas distinta (de C.I. Lewis) da mesma dicotomia. De acordo com a teoria de definição lexicográfica que nós esboçamos, a distinção entre o analítico e o sintético é, na melhor das hipóteses, uma questão de grau, visto que depende da espécie ou talvez da dificuldade das operações necessárias para se poder afirmar se as condições para denotação foram preenchidas ou não (ver sec. 2.4.). Talvez somente termos sabidamente não denotativos (e.g. CENTAUR [CENTAURO]) seriam a rigor analíticos, embora mesmo CENTAURO pudesse (verdadeiramente) denotar um animal imaginário da espécie pretendida.

d) O Método denotativo e e) O Método ostensivo. Eles são um só e o mesmo, exceto em que d) apresenta denotata de amostra descrevendo-os em palavras, ao passo que e) apresenta exemplos físicos.

f) O Método Implicativo (ou Contextual). Aqui, o termo e a definição são fundidos; não se pretende nenhuma equivalência ou permutabilidade; o termo é mostrado em uso, num contexto em que só esse termo pode ocorrer (por exemplo, DIAGONAL é "definido" pela oração: 'A square has two diagonals and each of them divides the square into two rightangled isosceles triangles' ["Um quadrado tem duas diagonais, e cada uma delas divide o quadrado em dois triângulos isósceles de ângulos retos"]). Parece, porém, que essa categoria é baseada só numa peculiaridade sintática superficial, desde que as definições implicativas podem ser transformadas em forma canônica: DIAGONAL 'that which a square has two of... etc.' ["aquilo de que um quadrado tem dois... etc."]

g) O Método Produtor de Regras. Exemplo: "a regra para a palavra "EU" é que cada falante deve empregá-la para indicar a si próprio. Como todas as definições são regras semânticas, é só uma questão de estilo de definição. Em nossa discussão, amenizamos o caráter de regra das definições ao omitirmos a expressão "denota se". Ela pode ser reformulada à vontade, em muitas variantes. Assim, a definição de BLUE [AZUL] pode-se dar a forma explícita de uma regra se dissermos: BLUE 'an adjective

applied to color such that.' (ou 'such as...') [AZUL "um adjetivo aplicado à cor, tal que..." ou "como..."]. Temos dúvidas de que g) seja uma subclasse útil de definição.

3. VALIDAÇÃO EMPÍRICA DE DESCRIÇÕES SEMÂNTICAS

Foi dito na sec. 1.3. que não há procedimento de descoberta para descrições semânticas *corretas*. Isso não equivale a dizer, no entanto, que não existam modos de se obter descrições *experimentais* dentre as quais a mais adequada possa ser selecionada ou sintetizada.

Um grupo de processos pode ser caracterizado como "extensional". Alguns escritores acreditam que, num processo "extensional", o descritor fica restrito a um papel de observador: ele observa as "coocorrências" de certos termos com determinados traços do contexto não-lingüístico. Mas, obviamente, isso pode fazer justiça só a um fragmento de uma língua; por exemplo, a situação que ocorre com NAPOLEÃO seria geralmente a de sala de aula, e não a situação particular da França por volta de 1800. Um processo extensional que envolvesse o descritor num papel mais ativo fa-lo-ia dar ao informante "denotata" experimentais de amostra; uma definição ostensiva seria escolhida a partir das aprovações ou desaprovações do informante. Talvez seja esse o meio pelo qual um modelo pseudobehaviorista, envolvendo "disposições para responder" (e.g. Quine, Morris), pudesse ser traduzido para testes behavioristas. Uma outra modificação do processo faria com que o descritor não só apresentasse exemplos físicos de "denotata", mas também descrevesse exemplos imaginários ao informante. Carnap acredita ser essa uma investigação intensional e não extensional, mas, de acordo com a nossa teoria, ela ainda lida com definições ostensivas.

Para aqueles que não estão satisfeitos com definições ostensivas — elas são insuficientes, de acordo com a teoria aqui desenvolvida — fica o problema de fazer derivar definições descritivas ("intensionais") de definições ostensivas. É aparentemente um fato biológico que seres humanos são capazes de fazer derivar definições intensionais a partir de fatos ("percepção dos universais"): não só lexicógrafos, mas todas as crianças o fazem, e o fazem bem. Parece, portanto, desperdício atribuir toda a carga do lexicógrafo, ou a qualquer outro descritivista semântico isolado. Por que não recorrer à ajuda de uma amostra de falantes da língua?

Gostaríamos de afirmar que a operação metalingüística de definir é um universal cultural. Quer dizer, todas as línguas fornecem um modo de perguntar: "O que é um X?" e em todas as culturas, pelo menos as crianças fazem uso desse artifício. Ao usá-lo, elas obtêm respostas significativas, algumas das quais são bem satisfatórias. Nós, portanto, acreditamos que é possível obter tentativas de definição de falantes ingênuos de qualquer língua.

Mas gostaríamos de fazer uma afirmação ainda mais arrojada: as definições que podem ser obtidas de informantes ao perguntar-se lhes na sua língua "O que é um X?" (i.e., obtidas a partir das respostas ao estímulo "O que é um X?") não são completamente fortuitas; pelo contrário, elas mostrarão um certo padrão recorrente para qualquer X, e é esse padrão que constitui a estrutura culturalmente compartilhada do significado de X. A reprodução desse padrão pode ser uma das tarefas do descritor semântico, embora devesse ser determinado de que modo o objetivo de tal reprodução pode ser acomodado a outros critérios para a adequação de descrições semânticas (cf. sec. 1.3.).

A filosofia da linguagem continua esbarrando no significado estruturado como um dado bruto. Frege e Peirce enfrentaram esse fato com relação a problemas lógico-modais e ele continua surgindo na literatura mais recente sob títulos que o citam indire-

tamente: “discurso oblíquo”, “opacidade de descrição” (Quine), “estrutura intensional” e “isomorfismo intensional” (Carnap). Os psicólogos demonstraram a existência de normas por meio de métodos não lexicográficos de descrição semântica (associações, diferencial semântico). Gostaríamos de supor que uma “definição popular” também manifestaria normas culturais, e essas seriam de particular interesse para um lexicógrafo preocupado com as bases empíricas da sua ciência.

Sob esse ponto de vista, a definição do lexicógrafo pode ser considerada como a hipótese de um homem ou de um grupo, que deve ser validada mostrando-se que ela é aceitável para os falantes da língua ou para um número representativo de falantes. A descrição de uma língua às vezes trata de idioletos, mas, em geral, ela tem aspirações mais altas — ela visa às características comuns a muitos idioletos. A lexicografia como disciplina descritiva (e não-normativa) deve também levar a sério o critério de interpessoalidade. Cada aspecto da produção do lexicógrafo pode e deveria ser submetido a testes de interpessoalidade: a escolha de termos a serem definidos (o que excluiria, ou pelo menos, separaria palavras raras e portanto amplamente sem sentido); a especificidade de definições (sec. 2.2.) e a determinação de critérios (sec. 2.3.) e a natureza (2.4.) de condições para denotação. Possivelmente, mesmo a estrutura de *sistemas* de definição (como discutido nas sec. 2.5. — 2.6.) pode, em parte, ser testada quanto à aceitabilidade.

Não seria difícil delinear um programa de pesquisa interdisciplinar prático para testar as sugestões acima apresentadas a fim de tornar a lexicografia mais científica do que ela tem sido até aqui.³

NOTAS

1. Daqui por diante, simbolizaremos o termo por letras maiúsculas a definição por meio de aspas simples e omitiremos as palavras “denota se”.
2. Para X’ devemos excluir uma expressão composta como “X EXCETO QUE NÃO c MAS c’”, que pode sempre ser formulada “ad hoc”. É essa possibilidade de formar tais expressões que faz das línguas os instrumentos semanticamente flexíveis que elas são.
3. Certas idéias contidas neste trabalho são mais desenvolvidas no meu artigo “On the Semantic Structure of Language”, in *Universals of Language* (ed. J.H. Greenberg), 2.^a ed., Cambridge, Mass., 1966, pp. 142-216.

POST-SCRIPTUM

Desde a publicação deste artigo, continuei a elaborar minhas idéias sobre definição lexicográfica. São de particular relevância para este assunto meus artigos: “Webster’s Third: A Critique of Its Semantics”, *IJAL* 30.405-409 (1964), e “Explorations in Semantic Theory”, in *Current Trends in Linguistics*, vol. III (ed. T.A. Sebeck), The Hague, 1966, 395-477. Nesse último artigo, exploro a possibilidade de se ligar a descon-tinuidade entre a língua objeto e a metalingua de definição que lingüistas modernos to-maram por empréstimo, à lógica simbólica, talvez sem justificação suficiente.

Sobre a questão de se validar empiricamente afirmações descritivas em semântica, eu gostaria de citar agora dois estudos que tratam, em parte, desse tópico. São o trabalho de Karl E. Zimmer *Affixal Negation In English and Other Languages: An Investigation of Restricted Productivity* (Suplemento de *Word*, vol. 20, n.º 2), N. York, 1964; e a obra de Edward H. Bendix: *Componential Analysis of General Vocabulary: The*

WEINREICH, U. — Definição lexicográfica em Semântica Descritiva. Trad. de Maria Cecília P. Barbosa Lima. *Alfa*, São Paulo, 28(supl.):103-118, 1984.

Semantic Structure of a Set of Verbs in English, Hindi an Japanese, suplemento de *IJAL*, vol. 32, n.º 2, 1966.
(Janeiro de 1967)

WEINREICH, U. — Lexicographic definition in descriptive semantics. Trad. de Maria Cecília P. Barbosa Lima. *Alfa*, São Paulo, 28(supl.):103-118, 1984.

ABSTRACT: Weinreich analyses several aspects of the definition of a word and the criteria to be adopted. A relevant issue is the metalanguage to be used when writing a dictionary and the correspondent methodology to define the "definiendum". One may utilize several strategies: the analytical, syntetic, denotative, ostensive, implicative or contextual methods. The chosen method will depend on the nature of the term to be defined: concrete referent, abstract notion, verbal action or process, grammatical instrument, etc. Synonyms and antonyms largely used in definitions are very important in Lexicography. The dictionary team must remember that definitions should be valid for the whole community to whom these definitions are addressed. Lexicographers should, therefore, use a language common to everyone and not their individual idiolect.

KEY-WORDS: Semantic description; lexicography; monolingual dictionary; thesaurus; meaning; definition; synonym; denotata; ostensive definition; lexicographic definition; metalanguage.

A INTEGRAÇÃO DOS NEOLOGISMOS POR EMPRÉSTIMO AO LÉXICO PORTUGUÊS

Ieda Maria ALVES*

RESUMO: Os termos estrangeiros empregados em uma língua podem constituir estrangeirismos e empréstimos. Com base em um corpus extraído de revistas e jornais brasileiros contemporâneos, procuramos estudar a fase neológica do empréstimo e a integração desse elemento à língua portuguesa.

UNITERMOS: Neologia; neologismo por empréstimo; estrangeirismos.

1. NEOLOGISMOS. DEFINIÇÃO E TIPOLOGIA

O sistema lexical de uma língua está constantemente se inovando. A este propósito, afirma B. Quemada (19, p. 137-8) que “se tornou evidente, para a maioria dos usuários, que uma língua de cultura moderna, necessariamente científica e técnica, não deve ver na neologia lexical apenas um mal inevitável. É a primeira condição a partir da qual o idioma pode permanecer um instrumento de comunicação nacional, mesmo internacional, e não ser apenas uma língua viva. Deve até considerar a criatividade lexical como parte responsável pela sua riqueza imediata, como o sinal evidente de sua vitalidade. Uma língua que não conhecesse nenhuma forma de neologia seria uma língua morta e, em suma, a história de todas as nossas línguas constitui a de sua neologia”.

Neologismo é definido por Boulanger (4, p.65-6), como uma unidade do léxico, palavra, lexia ou sintagma, cuja forma significante ou a relação significante-significado /.../ não estava realizada no estágio imediatamente anterior de um determinado sistema lingüístico. Assim, neologismo constitui uma unidade lexical de criação recente, uma acepção nova que se atribui a uma palavra já existente ou, então, um termo recentemente emprestado a um outro código lingüístico. Com base nesta definição, o A. estabelece três tipos de neologia: *formal*: neologismos criados por meio de derivação, composição, siglas, redução de palavras ou pela articulação de uma ou diversas sílabas que possuem um valor significante inédito; *semântica*: neologismos criados pela atribuição de um novo significado a um mesmo segmento fonológico; *por empréstimo*: neologismos que resultam da adoção de um lexema estrangeiro**.

Ao ser criado, o neologismo deve ser integrado ao sistema de uma língua. Esta integração deve obedecer aos critérios ortográficos, fonológicos e morfossintáticos desse idioma. Limitar-nos-emos, neste trabalho, ao estudo dos neologismos por empréstimo

* Departamento de Lingüística — Instituto de Letras, História e Psicologia — UNESP — 19.800 — Assis — SP.

** Alguns autores, como M.T. Biderman (1, p.158), consideram o empréstimo um tipo de neologia formal e classificam os neologismos em apenas duas categorias: formais e semânticos. K. Bochmann (3, p. 79) refere-se a certas escolas lingüísticas que, sobretudo na Rússia e na Romênia, tendem a considerar neológica apenas a inovação formal.

empregados na língua portuguesa e de sua integração a esse sistema lingüístico. Para tanto, usaremos exemplos recenseados em revistas e jornais brasileiros contemporâneos.*

2. NEOLOGISMOS POR EMPRÉSTIMO

Ocorre a neologia por empréstimo quando um elemento estrangeiro (expressão, conteúdo ou ambos) é utilizado numa determinada língua e passa a ser codificado por ela (15, p. 52).

Nem sempre, porém, o emprego de um lexema num outro sistema lingüístico supõe sua integração a esse idioma. E. Haugen (14, p. 213-5) propõe uma tipologia formal do neologismo por empréstimo. Denomina *modelo* a unidade lexical da língua A que é empregada pelos usuários da língua B. Assim, pode ocorrer a *importação* ou a *substituição* do elemento modelo. Dá-se a importação quando a unidade lexical recebida mantém-se inalterável formalmente; em caso contrário, há substituição. Haugen divide os neologismos por empréstimo em: *loanswords* — importação morfêmica sem substituição; *loanblends* — substituição morfêmica e importação; *loanshifts* — substituição morfêmica sem importação. A integração do termo estrangeiro ocorre na fase dos *loanblends* — quando o elemento externo se incorpora a uma classe morfológica e também na dos *loanshifts* — que correspondem aos decalques e aos casos de evolução semântica causados pela influência estrangeira.

Baseando-se na capacidade receptiva de um sistema lingüístico aos empréstimos, Vocadlo (22, p. 170) classifica as línguas em: *homogêneas* (teutônicas, celtas) — pouco acolhedoras às expressões estrangeiras; *amalgamadas* (neolatinas, grego moderno) — recebem empréstimos de idiomas aparentados; *heterogêneas* (inglês, romeno, persa) — abertas a elementos estrangeiros.

Preferimos adotar a classificação proposta por Guilbert (11, p. 92-3), segundo a qual o lexema externo à língua constitui um *estrangeirismo* ou um *empréstimo*. No primeiro caso, incluem-se os nomes próprios, patronímicos, termos que exprimem realidades sem correspondência na língua receptora. O empréstimo constitui o elemento já integrado ao sistema lingüístico adotante, como diz Weinreich (23, p. 11): “Quando um falante da língua x emprega uma forma de origem estrangeira não como um empréstimo eventual da língua y, mas porque ele ouviu o termo estrangeiro sendo usado por outros falantes em x — enunciados, este elemento externo pode ser considerado, do ponto de vista descritivo, como parte da língua x.”**

A fase neológica do termo estrangeiro situa-se entre o estrangeirismo e o empréstimo e corresponde à sua instalação no sistema de uma língua. Tal fase, de caráter transitório, é denominada *peregrinismo* (5, p. 224). Nesse período de difusão, o elemento estrangeiro pode adaptar-se ao idioma que o recebe. L. Guilbert (11, p. 96-8) adota três critérios, por meio dos quais essa unidade lexical é considerada em fase de integração a uma língua: morfossintático, semântico e fonológico.

** Os exemplos citados foram extraídos das revistas *Capricho* (Ca), *Cláudia* (Cl), *Desfile* (De), *Isto É* (IE), *Manchete* (Ma), *Veja* (Ve), *Visão* (Vi) e dos jornais *O Diário de São Paulo* (DSP), *o Estado de São Paulo*, (ESP), *Folha de São Paulo* (FSP).

** Uma outra tipologia, proposta por Schank (21, p. 67-88), distingue entre empréstimos não integrados, parcial e totalmente integrados, o que corresponderia a uma divisão entre *palavras de citação*, *estrangeirismos* e *empréstimos*. Segundo esta classificação, o estrangeirismo representa a unidade lexical já em vias de integração a um sistema lingüístico.

2.1. Critério Morfossintático

2.1.1. Composição e Derivação

Quando o lexema estrangeiro constitui a base de uma derivação ou de uma composição de acordo com a morfossintaxe de uma língua, ele está se integrando ao léxico desse sistema. Assim, podemos afirmar que um termo emprestado faz parte de uma comunidade lingüística desde que seja susceptível de derivação e de composição, tal como os elementos autóctones.*

Em português, registramos alguns neologismos** compostos de termos estrangeiros e prefixos (ou elementos de composição) vernáculos: “Durante os últimos dias foram presas 25 pessoas nas manifestações *antiapartheid*” (<prefixo *anti* — + base inglesa *apartheid*) (FSP, 07.12.84, p. 15, c.6); “Porque é o único que se transforma em Toca-Fitas: basta ligá-lo num *Auto-Deck* Motorádio” (< elemento de composição *auto* — + base inglesa *deck*) (Ve, 06.08.81, p.10); “... toca-discos belt drive semi-automático e *auto-repeat*...” (< elemento de composição *auto* — + base inglesa *repeat*) (Ve, 16.12.81, p. 147); “... ela usa *maxipull* e meiões em malha canelada, ...” (< elemento de composição *maxi* — + base inglesa *pull*) (De, 05.81, p. 56, c. 1); “... a calça comprida, também em seda, que tem corte *semibaggie*, da *Decan-Deux*” (< elemento de composição *semi* — + base inglesa *baggie*) (De, 04.81, p. 92); “... o resultado da pesquisa sobre o “Craque do Ano” e um *super-poster* surpresa” (< prefixo *super* — + base inglesa *poster*) (VE, 16.12.81, p. 141).

Alguns neologismos resultam da composição de uma ou duas unidades lexicais vernáculos e de uma estrangeira: “... saia reta em lã cor de tomate e *blazer-jaquetão* com botões dourados, ...” (< base inglesa *blazer* + base portuguesa *jaquetão*) (De, 05.81, p. 88); “Ela usa três peças básicas: *calças* — *jogging* em malha, ...” (< base portuguesa *calças* + base inglesa *jogging*) (De, 05.81, p. 72); “*Calças* — *training* em malha, pull em molleton com grafismo, ...” (< base portuguesa *calças* + base inglesa *training*), (De, 05.81, p. 68, c.2); “... calças levemente afuniladas e spencer com *gola* — *summer*...” (< base portuguesa *gola* + base inglesa *summer*) (De, 04.81, p. 50); “Frente para o mar. *Piscina* — *deck*. Playground” (< base portuguesa *piscina* + base inglesa *deck*) (Ve, 11.02.81, p. 6); “Logo no número de Jorge Amado você ganha um *poster* — *calendário*” (< base inglesa *poster* + base portuguesa *calendário*) (Ve, 12.08.81, p. 119, c.3); “... junto ao *pull-jaquetão* em malha usado com colete acolchoado, ...” (< base inglesa *pull* + base portuguesa *jaquetão*) (De, 05.81, p. 69); “... a *saia* — *calça* — *Kilt* com detalhes em pelica, ...” (bases portuguesas *saia* e *calça* + base inglesa *Kilt*) (De, 05.81, p. 83); “O tecido xadrez pode ser adotado, desde que miúdo e em cor escura /.../, no *vestido* — *housse* em sarja Flezal, ...” (< base portuguesa *vestido* + base francesa *housse*) (De, 06.81, p. 79, c.2).

Observamos também alguns neologismos criados por meio da derivação sufixal nominal, em que a base estrangeira, um nome próprio, contribui para a formação de substantivos comuns: “Contudo, uma curta frase faz vacilar, na constituição *breshneviana*, o edifício dos direitos e garantias” (< bases russas *Breshnev* + sufixo — ano) (ESP, 19.06.77, p. 174, c.7); “O *gaullismo* funciona hoje, na França, como uma doutrina teológica dispersada entre diversas religiões” (< base francesa /De/ *Gaulle* +

* Cf. Franolic (8, p. 247), George (10, p. 64-6), Dubois (6, p. 15).

** Consideramos neológicas as unidades lexicais não dicionarizadas pelo *Novo dicionário da língua portuguesa*, de Aurélio Buarque de H. Ferreira (edição de 1975). (7).

suf. — *ismo*) (FSP, 14.06.77, p. 10, c.4); "...lideranças confessadamente totalitárias, *maotsetungistas* ..." (< base chinesa *Mao Tse Tung* + suf. *-ista*) (ESP, 17.06.77, p. 3 c.4); "... num clima de confusão política e de divisão entre *gaullistas* e *giscardianos*, socialistas e comunistas" (< bases francesas /De/ *Gaulle* e *Giscard* /d'Estaing/ + suf. — *ista*) (ESP, 14.06.77, p. 9, c.1).

A integração de uma unidade lexical recebida por empréstimo pode ser bastante significativa, de modo que o sufixo estrangeiro pode gerar derivados com uma base vernácula. É o que ocorre em francês, em que os sufixos *-ing* e *-man* ocasionam a criação de falsos anglicismos, constituídos por elementos ingleses inexistentes na língua inglesa: *lifting*, *planking*, *printing* (suf. — *ing*), *comingman*, *crossman*, *limitman*, *scratchman*, *tennisman*,... (suf. — *man*) (10, p. 65; 13, p. 109). Fenômeno semelhante é citado por Franolic (9, p. 356) em relação a empréstimos franceses providos do mesmo sufixo que entraram em grande número na língua croata, de tal modo que o sufixo, reconhecido como tal pelos falantes, desligou-se de sua base e tornou-se capaz, por sua vez, de formar palavras novas com radicais autóctones. Nesse caso, o sufixo integrou-se ao sistema sufixal do croata. Em *Language*, Bloomfield (2, p. 429) refere-se ao emprego, em inglês, de termos de origem francesa constituídos com o sufixo *-able*, *-ible*, (*agreeable*, *excusable*, *variable*), o qual se estendeu a outras bases inglesas, como *bearable*, *drinkable*, *eatable*,...*

2.1.2. Classe Gramatical

A maior parte dos neologismos recebidos por empréstimo no português pertence à classe gramatical dos substantivos, mais raramente à dos adjetivos e à dos verbos. Este fato ocorre não somente em nossa língua, mas é comum em todos os sistemas linguísticos**. Bastante raros são os empréstimos de palavras gramaticais.

Em geral, o elemento recebido por empréstimo conserva a classe gramatical que possuía na língua de que procede: "Chega ao Brasil, num salto de liderança, o '*basket*' mais completo e avançado do mundo!" (*basket* — substantivo no inglês e no português) (Ma, 21.08.82, p. 91). Algumas vezes mudam de classe gramatical, como os substantivos ingleses *direct-drive*, *holding* e *quick-vision*, empregados adjetivamente: "... além do primeiro toca-disco *direct-drive* que, virtualmente, não apresenta nenhum wow e flutter" (Ve, 29.09.76, p. 93, c.1-2); "Com a criação da Portobrás, os portos terão uma empresa '*holding*'..." (Ve, 22.12.76, p. 83); "Graças ao sistema '*Quick-Vision*', exclusivo da Philips, ..." (Ma, 18.12.76, p. 46, c.2).

2.1.3. Gênero e Número

Os neologismos por empréstimo introduzidos em uma língua integram-se normalmente a seu sistema de flexão em gênero e número.

Quando emprestado a línguas que se flexionam quanto ao gênero, o termo estrangeiro quase sempre segue a língua original, como os elementos franceses *fourreau* e *tour* e o espanhol *tribunales*: "... aí temos *um fourreau* em veludo lavrado..." (De, 06.81, p. 80); "Comece *seu tour* pela Europa com o esplendor..." (Ma, 27.11.76, p. 8).

* Segundo Mattoso Câmara Jr., (18, p. 258), tais tipos de empréstimos são numerosos e frequentes, mesmo sem a existência de contato espacial de dois sistemas linguísticos. Basta apenas que alguns lexemas, com uma determinada estrutura mórfica, sejam emprestados para que seus sufixos constitutivos se tornem produtivos na língua importadora.

** Cf. Jespersen (16, p.211-2), Machado (17, p.272), Biderman (1, p.165).

54, c.1); “Esta tarde, o episódio deverá repetir-se outra vez nos escuros e estreitos corredores do edifício dos ‘tribunales’ buenaireses” (ESP, 30.06.77, p.9, c.1).

Se a língua doadora não possui flexão em gênero, o elemento emprestado tende a receber um, ao integrar-se a um sistema lingüístico provido dessa flexão (5, p. 257). Em português, é comum à maioria dos nomes receberem o gênero masculino, o não — marcado (1, p. 165)*: “No Brasil é o designer Nilo de Almeida quem cria esta moda especial:...” (CI, 10.82, p. 9); “... que há 40 anos vem determinando a qualidade e o ‘blend’ de cada xerez ...” (Ve, 18.08.82, p.12, c.1); Mas para o pessoal que quer mais opções, a Santista tem os *Colors* e os *Blacks*” (Ve, 18.05.83, p. 23).

Em relação à flexão em número, o termo estrangeiro geralmente conserva em português o número original: “Essas palavras significam que Old Smuggler é o blend de 38 dos mais finos *malt whiskies* produzidos na Escócia” (Ma, 13.11.76, p. 74); “Você vai entender porque o Passat não é apenas um campeão de *rallies*” (Ma, 11.12.76, p. 59); “E descubra tudo aquilo que os outros *scotches* prometeram” (Ma, 21.08.76, p. 135).

Na verdade, o neologismo por empréstimo pode guardar a flexão em número da língua de que provém por um período limitado. Na medida em que se integra à língua importadora, tende a adaptar-se ao seu sistema de formação de plural**. “Especializada em pratos típicos da região sul da Itália, a casa lembra as genuínas *trattorias* de Bari, ...” (Ve, 23.09.81, p. 50, c.2).

2.1.4. Decalque

Um modo mais insidioso de integração da expressão estrangeira processa-se por meio do decalque, que consiste na versão literal desse elemento externo para a língua receptora: “Abbud — *Alta costura* masculina” (< francês *haute couture*) (Ve, 07.01.81, p. 51). Pode ocorrer alteração na ordem dos termos, ao serem transpostos para a língua portuguesa. Assim, *lojas de departamentos*, sintagma calcado no inglês *department stores*, obedece à ordem determinado e determinante, mais usual no português: “Um dos importantes setores da economia /.../, caracterizando-se por possuir hoje uma das mais importantes cadeias de *lojas de departamentos* em nosso país” (IE, 28.01.81, p. 98).

2.2. Critério Semântico

Segundo o critério semântico, a instalação do termo estrangeiro ocorre quando tal elemento introduzido na língua receptora com um único significado, torna-se polissêmico.

Assim, o francês *détente*, introduzido na política internacional para significar a distensão entre os Estados Unidos e a Rússia, é também aplicado à política brasileira: “Se o próprio presidente propôs a *détente*, cabe aos políticos dos dois partidos aceitá-la e dar curso a esta proposição” (ESP, 17.06.77, p. 4, c.1). *Gulag*, substantivo próprio russo que intitula uma obra de A. Solzenitzin (*O Arquipélago Gulag*), é empregado em

* Guilbert. (12, p. 1584) observa que a atribuição do gênero masculino ou feminino a um termo estrangeiro pode ser condicionada por fatores de sentido e forma. Em francês, *pacha* e *odalisque* pertencem ao masculino e ao feminino, respectivamente, embora o turco, de onde se originam tais palavras, ignore a flexão em gênero. Esta constatação também se aplica ao português, em que *odalisca* pertence ao gênero feminino e *paxá* ao masculino. Os elementos ingleses *girl*, *milady* e *miss* são integrados ao gênero feminino do português (cf. 7) por se referirem a pessoas do sexo feminino.

** Dubois (6, p. 14) exemplifica tal afirmação com a palavra italiana *dilettante* que, empregada no francês, guardou durante um certo tempo o plural *dilettanti*, antes de adaptar-se ao plural francês *dilettantes*.

referência às ditaduras militares na América Latina: “Pelos cálculos da IADB, 170 de seus diplomados hoje ocupam cargos de destaques no ‘gulag’ de ditaduras militares na América Latina” (FSP, 19.06.77, p.1, c.2). As unidades lexicais *blue jeans* (ou *jeans*) e *lingerie*, de origem inglesa e francesa, respectivamente, foram introduzidas em português para designar peças do vestuário. Passaram a ter, posteriormente, uma nova acepção: a de nomear os tecidos que servem para confeccionar tais peças: “Você está sempre elegante com esta saia de *jeans* da Queene” (Cl, 06.81, p. 103); “... como este conjunto de calça e blusa de *lingerie*, ...” (Cl. 05.81, p. 106).

2.3. Critério Fonológico

Segundo este critério, um termo estrangeiro começa a fazer parte do léxico de uma língua à medida que se integra fonologicamente a ele. Na verdade, o elemento estrangeiro tende a adaptar-se ao sistema fonemático da língua receptora*.

Algumas vezes, como observa Weinreich (23, p. 28), podemos precisar que a integração fonológica do empréstimo é posterior à introdução desse elemento por via escrita: é o caso, por exemplo, do inglês *gasoline*, que no japonês coloquial do Havaí é pronunciado/g'aswrin/.

Assim, o termo emprestado não é adaptado fonologicamente à língua importadora; ao contrário, recebe uma pronúncia de acordo com o sistema fonológico desse idioma.

A adaptação fonológica do termo estrangeiro revela-se por uma adaptação ortográfica: “O *birô* político o nomeou ministro da Administração Interna” (< francês *bureau*) (FSP, 20.06.77, p.2, c.4); E isto é apenas um simples exemplo das maravilhas de uma cidade com 40 teatros de primeira classe, com *butiques* coloridas...” (< francês *boutique*) (Ve, 17.11.76, p. 116, c.1); “Como este conjunto de calça e blusa de *lingerie*, com jaqueta de tafetá em *degradês*” (< francês *dégradê*) (Cl, 05.81, p. 106); “Não faltaram fé nem confiança ao rei don Juan nem ao *premiê* Adolfo Suárez” (< francês *premier*) (FSP, 19.06.77, p.2, c.1).

Entretanto, mesmo que uma unidade lexical já tenha sido dicionarizada e aportuguesada, podemos encontrá-la grafada de acordo com a língua de que procede. Eis alguns exemplos, tomados ao francês: “... para quando a ocasião exige um vinho de sabor e *bouquet* ligeiramente frutê.” (Vi, 09.05.77, p. 2); “No alto da página, duas versões do *chemisier*, peça indispensável no guarda-roupa da mulher elegante” (De, 04.81, p.84); “Ao lado, vestido Crylor estampado em suaves ondas azul/cinza, cintorolotê e mangas *raglan*, da Deblu” (De, 04.81, p. 60); “*Mousse* de pêssego. Bata no liquidificador uma lata de leite Moça...” (Ma, 25.12.76, p. 65, c.1); e ao inglês: “O Pas-sat disputou quatro campeonatos de *rallye*” (Ma, 11.12.76, p. 59); “100% *nylon*. Exclusividade Tabacow” (Ma, 14.05.77, p. 101). O *Novo dicionário da língua portuguesa* (7) registra as formas mencionadas respectivamente como: *buquê*, *chemisiê*, *raglã*, *musse*, *rali* e *náilon***.

2.4. Critério de Frequência.

Nem sempre a expressão estrangeira empregada na língua portuguesa adapta-se à

* Cf. Mattoso Câmara Jr. (18, p. 263), George (10, p. 63), Weinreich (23, p. 14).

** Este fato significa, para Guiraud (13, p. 103), que tais palavras traduzem denotações semânticas estrangeiras ou são portadoras de conotações estilísticas.

sua fonologia e ortografia. Em alguns casos, a forma original permanece. Parece-nos, entretanto, que a frequência do termo emprestado, ainda que empregado na sua forma nativa, constitui um critério para a sua aceitabilidade na língua portuguesa. Eis alguns exemplos de lexemas ingleses (*design*, *jeans*) e francês (*griffe*) frequentes em português*: “Uma grande coleção de jóias e relógios, em ouro 18 K de *design* exclusivo, espera por você na Natan” (Ma, 07.05.83, p. 2); “Etiqueta francesa com qualidade brasileira no *design* e no conforto” (Cl. 10.82, p. 9); “O lado feminino do *jeans*” (Ca, 12.82, p. 3, c.3); “*jeans* produzidos por confecções Zopa Ltda.” (Ve, 01.12.82, p. 78); “Conheça a nova *griffe* Berta” (Ma, 30.04.83, p. 52); “Na Tilty’s encontro de tudo o que preciso para vender em minha butique. E com tudo o que se possa desejar de uma *griffe*, ...” (De, 10.82, p. 29, c.3).

3. CONSIDERAÇÕES FINAIS

Depois de criado, o elemento neológico faz parte da dinâmica da língua. Como bem assinala Saussure (20, p. 133), os termos de uma língua são solidários e o valor de cada um resulta da presença simultânea dos outros. Assim, toda criação neológica contribui para alterar o funcionamento da língua e é alterada pela criação ou desaparecimento de outros elementos.

Entretanto, sabemos que a existência de um neologismo é ratificada pela aceitação da sociedade em que ele está inserido, pelo seu uso efetivo nessa comunidade. E não podemos, *a priori*, determinar a possibilidade que tem uma unidade lexical neológica de ser aceita e de ser realmente integrada ao código da língua. B. Quemada (19, p. 141) manifesta-se a respeito: “O uso, este mestre inquieto na medida em que ele soube guardar seu mistério, é de fato uma síntese de forças muito diversas, oriundas de tendências psicológicas, econômicas, culturais, de regras de economia funcional da educação, etc., fatores que condicionam o aparecimento, a difusão e as oportunidades de sobrevivência de um neologismo. A multiplicidade destes fatores, tanto quanto seu caráter flutuante, explica a aparente dificuldade de apreensão e, para alguns, a arbitrariedade do uso”.

Uma vez consagrado pelo uso, o elemento neológico é geralmente inserido num dicionário. Constitui o dicionário o critério final, segundo o qual um neologismo é integrado ao léxico da língua. Atribui-se assim, ao lexicógrafo, o poder de decidir sobre a aceitabilidade ou não de um novo termo ou expressão e sobre sua incorporação à língua. Observa L. Guilbert (11, p. 54) que, por causa desse poder que é concedido ao lexicógrafo, ele se cerca de garantias. Trabalha em equipe, a fim de evitar uma responsabilidade pessoal quanto à aceitação ou rejeição de um neologismo. Baseia-se em obras lexicográficas anteriores e daí resulta um certo parentesco freqüentemente observado no gênero. A defasagem observada entre o emprego de um neologismo na imprensa ou em escritores, cuja notoriedade na comunidade constitui uma espécie de julgamento de aceitabilidade, e seu registro num dicionário é conseqüência da prudência do lexicógrafo. Mas essa prudência é necessária e solidifica o registro do neologismo, tornando-o efetivamente aceito pela comunidade.

* Consideramos frequentes os elementos diferentemente repartidos entre os vários textos analisados.

ALVES, I.M. — A integração dos neologismos por empréstimo ao léxico português. *Alfa*, São Paulo, 28(supl.):119-126, 1984.

ALVES, I.M. — L'intégration des néologismes par emprunt au lexique portugais. *Alfa*, São Paulo, 28(supl.):119-126, 1984.

RÉSUMÉ: Les termes étrangers employés dans une langue peuvent être des xénismes et des emprunts. Ayant comme source un corpus constitué d'exemples pris dans des revues et des journaux brésiliens contemporains, on a essayé d'étudier la phase néologique de l'emprunt et son intégration à la langue portugaise.

UNITERMES: Néologie; néologie par emprunt; xénisme.

REFERÊNCIAS BIBLIOGRÁFICAS

1. BIDERMAN, M.T. — *Teoria lingüística (lingüística quantitativa e computacional)*. Rio de Janeiro, LTC, 1978.
2. BLOOMFIELD, L. — *Le Langage*. Trad. do inglês de J. Gazio. Paris, Payot, 1970.
3. BOCHEMANN, K. — Conditions sociales et conditions linguistiques de la néologie lexicale et leur place dans un modèle de la production lexicale. *Philologica Pragensia*, Praga, 20(2): 79-87, 1977.
4. BOULANGER, J.C. — Néologie et terminologie. *Néologie en Marche*, Montréal, 4: 5-128, 1979.
5. DEROY, L. — *L'emprunt linguistique*. Paris, Les Belles Lettres, 1956.
6. DUBOIS, J. — L'emprunt en français. *L'Information Litteraire*, Paris, 10-6, 1963.
7. FERREIRA, A.B.H. — *Novo dicionário da língua portuguesa*. Rio de Janeiro, Nova Fronteira, 1975.
8. FRANOLIC, B. — Adaptation secondaire ou la dérivation des emprunts d'origine française en croate. *Lingua*, Amsterdam, 40: 247-61, 1976.
9. FRANOLIC, B. — L'intégration des suffixes étrangers (français) dans le système suffixal croate. *Lingua*, Amsterdam, 27: 355-66, 1971.
10. GEORGE, K.E.M. — Anglicisms in contemporary French: II — Linguistic aspects. *Modern Languages*, London, 57: 63-8, 1976.
11. GUILBERT, L. — La créativité lexicale. Paris, Larousse, 1975.
12. GUILBERT, L. — L'emprunt. In: *Grand Larousse de la langue française*. Paris, Larousse, 1971 — 8.T.II.
13. GUIRAUD, P. — *Les mots étrangers*. Paris, PUF, 1971.
14. HAUGEN, E. — The analysis of linguistic borrowing. *Language*, Baltimore, 26: 210-31, 1950.
15. HUMBLET, J. — Vers une typologie de l'emprunt linguistique. *Cahiers de Lexicologie*, Besançon, 25(2): 46-70, 1974.
16. JESPERSEN, O. — *Language. Its nature, development and origin*. 10. ed. London, G. Allen & Unwin, 1954.
17. MACHADO, J.P. — Notas soltas sobre a influência arábica na língua portuguesa. *Boletim Mensal da Sociedade de Língua Portuguesa*, 14: 262-72, 1963.
18. MATTOSO CÂMARA JR., J. — *Princípios de lingüística geral*. 4. ed. Rio de Janeiro, Acadêmica, 1970.
19. QUEMADA, B. — A propos de la néologie. *La Banque des Mots*, Paris: 2, 137-50, 1971.
20. SAUSSURE, F. de — *Curso de lingüística geral*. Trad. de A. Chelini, J.P. Pais e I. Blikstein. São Paulo, Cultrix, 1970.
21. SCHANK, G. — Vorschlag zur Erarbeitung einer perationalem Fremdwortdefinition. *Deutsche Sprache*. 2: 67-88, 1974. Apud por KIRKNESS, A. Sobre a lexicologia e lexicografia das palavras estrangeiras. In: *Problemas da lexicologia e lexicografia*. Trad. e intr. de M. Vilela. Porto, Civilização, 1979, p. 236-7.
22. VOCADLO, O. — Some observations on mixed languages. 4. th. *International Congress of Linguistics*: 169-76, 1936. Apud MAURIS, J. — Problématique de l'emprunt lexical. *Travaux de Terminologie et de Linguistique*. Québec, 1:87-103, 1982.
23. WEINREICH, U. *Languages in contact*. 8. ed. The Hague Paris, Mouton, 1974.

A PREFIXAÇÃO NEOLÓGICA NO VOCABULÁRIO DA PROPAGANDA CONTEMPORÂNEA

Rony Farto PEREIRA*

RESUMO: Um exame dos neologismos ocorrentes na publicidade, com um corpus constituído pelas revistas Veja e Isto É, demonstrou que o processo da derivação denominado prefixal é o mecanismo mais utilizado na formação das palavras novas nesse domínio. Após a indicação do quadro geral dos prefixos encontrados, apresentam-se exemplos, em seus respectivos contextos, acompanhados de comentários sobre certas peculiaridades.

UNITERMOS: Neologismos; prefixação; derivação; publicidade; vocabulário.

1. NEOLOGIAS INTERNAS E EXTERNAS

São bastante variadas as tipologias de neologismos surpreendidas nas obras do gênero, que recentemente pudemos consultar. Oferecem, apesar disso, alguns pontos comuns, os quais remetem quase sempre a determinados processos, normalmente mais produtivos na língua.

Registra-se, em alguns casos, apenas uma divergência terminológica, que pode ser considerada irrelevante ou mesmo inócua. Às vezes, as classificações refletem pequenas oscilações nos rótulos que certos estudiosos empregam, eles próprios, em seus diferentes textos, escritos em épocas distintas.

Uma análise do quadro geral de neologismos** oferecido pelas diferentes obras, à luz dos anúncios publicitários da atualidade (destacados das edições semanais das revistas *Veja* e *Isto É*, do ano de 1981), permitiu-nos chegar a um esquema relativamente simples (10, p. 54).

Foram encontrados exemplos de neologismos originários de dois processos distintos: a) as neologias internas, ou seja, criação de elementos novos a partir de componentes do próprio vernáculo e b) as neologias externas, isto é, transferência lexical de elementos já formados em uma língua estrangeira determinada.

Apesar de apresentarem essa natureza diversa, já que os empréstimos estrangeiros não são *criações verdadeiras*, mas apenas *adoções* de termos preexistentes no idioma estrangeiro (2, p. 95), não se poderá verificar nessa separação um caráter nitidamente ortodoxo e, obviamente, muito rígido. As mencionadas adoções podem ter um desenvolvimento posterior na língua vernácula, ocasião em que se utilizam sufixos pertencentes ao idioma de chegada, por exemplo. É o que ocorre na formação de termos que

* Departamento de Linguística — Instituto de Letras, História e Psicologia — UNESP — 19800 — Assis — SP.

** Chamar-se-á *neologismo*, neste trabalho, todo vocabulário que não estiver registrado no *Novo dicionário da língua portuguesa*, de Aurélio Buarque de Holanda Ferreira, conquanto se tenha consciência das limitações desse critério. Prefere-se usar *neologia* quando se tratar do processo de formação dos vocábulos.

derivam de palavras (como *cartoon* > *cartunista*) ou de siglas estrangeiras (como *ufo* > *ufologia*).

De outro lado, criações muitas vezes se realizam com base em formas propriamente estrangeiras ou externas: é o que acontece na derivação ou composição em que entram afixos alógenos ou eruditos. Nesses casos, a incorporação dos componentes ao vernáculo se dá por haver longa tradição de uso — o que não impede tratar-se de constituintes estrangeiros (como o *anti-*, de *antigermes*, ou o *ultra-*, de *ultra-avanzado*).

No interior desses dois blocos, constata-se um entrelaçamento em vários pontos. Nota-se que um mesmo neologismo poderá surgir da ampliação do âmbito da significação (semântico, portanto) e conter em si, simultaneamente, a formalização dos recursos peculiares à estrutura morfo-sintática do vernáculo (mecanismo formal); não é outra a formação de vocábulos como *colunável*, *porta-objetos* ou *segundeiros* (cf. 10, p. 141-2).

Este trabalho objetiva examinar, entre os neologismos constituídos por processos internos à língua portuguesa e utilizando seus próprios elementos, alguns casos de derivação prefixal, tomada em seu sentido corrente (9, p. 136). Trata-se do mecanismo mais produtivo encontrado, dentre os vários processos de constituição de neologismos, exatamente por permitir à publicidade cumprir seu papel fundamental de preencher as necessidades de novidade e de familiaridade (8, p. 83).

2. DERIVAÇÃO PREFIXAL

Afirma Jean Dubois que “o procedimento da prefixação se revela particularmente apto para traduzir as relações dinâmicas que caracterizam as estruturas sociais e políticas” (3, p. 140). As razões para uma preferência que se dá à prefixação, sobre os demais mecanismos derivacionais, estão ligadas ao fato de o elemento prefixal poder “traduzir uma relação fundamental, porque o componente radical, o significante, conserva o aspecto morfológico; o conteúdo semântico é imediatamente surpreendido” (3, p. 140).

Certos prefixos são nitidamente predominantes nos anúncios, uma vez que seu valor “convém particularmente às tendências da língua publicitária”: entre eles, podem contar-se “*auto-* (comodidade de um aparelho que trabalha sozinho); *anti-* (luta contra inimigos de toda sorte); *super-*, *extra-*, *ultra-*, equivalentes a superlativos”. (5, p. 281).

A partir do exemplo de Guilbert (6, p. 219), que acha importante aproximar uma categorização segundo o valor sintático de uma correspondente distribuição semântica, resolvemos efetuar uma classificação dos neologismos por prefixação que considerasse o valor semântico dos prefixos empregados. Aproveitamos, também, o esquema elaborado por Ieda Maria Alves (1, p. 10-13).

Conforme o valor semântico corrente, os prefixos registrados puderam ser reunidos como mostra o Quadro I; trata-se de um esquema útil, especialmente para uma ordenação da exposição das unidades lexicais inventariadas.

Por questão de espaço, não apresentaremos neste trabalho todos os neologismos encontrados no *corpus* analisado. Um exemplo de cada tipo será registrado, em seu respectivo contexto, ao qual se acrescentou algum comentário, quando esse procedimento poderia ilustrar a caracterização do processo em exame.

Quadro I: Prefixos encontrados no *corpus*

<i>prefixos</i>	<i>valor semântico</i>
<i>a(n)- des- in-</i>	negação
<i>anti-</i>	oposição
<i>pro-</i>	favorecimento
<i>extra- max- re- sobre- super-</i>	intensidade
<i>bi- mono- multi- poli- tri-</i>	quantificação
<i>co- inter- pós- pré- retro- trans-</i>	espacialidade-temporalidade
<i>auto-</i>	inerência (ação ou movimento próprio)

2.1. Prefixos de negação

*a(n)-
anecóica:*

“Um bom exemplo é a maior e mais rigorosa câmara *anecóica* da América Latina. Usada para testar caixas acústicas (...)” (VE 689, 18.11.81: 75, 1)*

*des-
desimportar:*

“Mas aquilo que conseguirmos *desimportar*, aquilo que nos aparelharmos para produzir nós mesmos, é nosso (...)” (IE 221, 18.03.81: 9, 1)

*in-
insonorizado:*

“Teto em alumínio *insonorizado*.” (VE 664, 27.05.81: 95, 2)

Merece destaque o vocábulo *anecóica*, pouco corrente. O contexto em que aparece mostra, claramente, uma referência ao seu significado, diretamente relacionado à palavra já dicionarizada *ecóico* (“referente a eco”), de que é antônimo. Isso se percebe — e fica patente o caráter de negação do prefixo — quando se compreende que *anecóica* é determinante de uma câmara destinada a testar caixas acústicas, ou seja, de uma câmara à prova de eco.

Chama a atenção, igualmente, a formação do termo *desimportar*, que substitui a expressão “não importar” (empregada para expressar a negação do processo de importar). Evidentemente, *desimportar* parece estilisticamente mais ativo, do ponto de vista

* Os contextos serão sempre localizados nessa ordem: revista (VE = Veja; IE = Isto É), número da edição, data, página e coluna.

dos ideais e objetivos da propaganda, fundamentados na inovação e na surpresa, de um lado. Por outro, substitui um segmento menos funcional, porque também se caracteriza por parecer familiar à massa consumidora, especialmente por constituir um vocábulo soldado e unificado em um só segmento lexical.

2.2. Prefixo de oposição

anti-

anti-corrosão:

“Seções de aço inox garantem a mais perfeita anticorrosão.” (VE 662, 13.05.81: 72)

O número de neologismos formados a partir do prefixo *anti* — foi muito grande. Não se registrou qualquer outro prefixo, no *corpus* estudado, que tivesse uma natureza opositiva.

Pode-se ressaltar que um aspecto interessante, na série de exemplos com *anti-*, foi o fato de se constatar uma relativa instabilidade no uso (ou omissão) do hífen. Pelo que se verificou, não se segue na propaganda qualquer tipo de regra nesse sentido, tanto que o mesmo segmento aparece duas vezes seguidas, uma com hífen e outra sem esse sinal gráfico.

2.3. Prefixo de favorecimento

pro-

proálcool:

“Países desenvolvidos estão acompanhando com grande interesse o trabalho que a iniciativa privada vem fazendo no *Proálcool*.” (IE 251, 14.10.81: 45, 1-2)

Uma vez que se trata de projetos de apoio a determinados setores da vida nacional, realidades como a do *Proálcool* (assim como a da *Pró-cana*, *Pró-várzeas* ou *Pronazém*) implicam simultaneamente o aproveitamento da significação do prefixo *pro-* (em proveito de) e o truncamento sintomático da palavra “projeto” ou “programa”, geralmente incluída no sintagma de que se originou essa espécie de sigla. A interpretação será menos contestável e mais ratificada quando se observa o mesmo mecanismo de redução na segunda parte de alguns desses vocábulos, como em *Pronazém* = em favor da armazenagem; *Propec* = em favor da pecuária, etc.

2.4. Prefixos intensivos

extra-

extrapesados:

“E só poderia ser do forno da maior usina brasileira de fundidos e forjados *extrapesados* de aço: a Vibasa.” (VE 654, 18.03.81: 85, 1)

maxi-

maxivalorizar:

“O Nacional está lançando um novo produto para *maxivalorizar* o seu cruzeiro (...)” (VE 679, 09.09.81: 115, 1)

re-

rerrefinarmos:

“Mas, essa importação pode ser evitada se *rerrefinarmos* o lubrificante que foi usado (...)” (VE 691, 02.12.81: 123)

sobre-

sobredimensionados:

“Os produtos Falk são *sobredimensionados* com reserva de resistência, o que faz com que eles agüentem o mais duro trabalho (...)” (VE 658, 15.04.81: 83)

super-

superfacilitadas:

“Condições de pagamento *superfacilitadas*: 10% de entrada e o saldo financiado em até 20 anos (...)” (IE214, 28.01.81:97, 1)

ultra-

ultrabaixos:

“Você vai conhecer o primeiro cigarro brasileiro *ultrabaixos* teores.” (IE 222, 25.03.81: 38-a)

Bastante empregados na propaganda, dado o seu grande valor distintivo — que intensifica a qualidade ou as propriedades dos produtos — os prefixos intensivos encontram sua expressão mais apreciada no segmento *super-*. Assim, no *corpus* analisado, para apenas uma dezena ou menos de neologismos com outros prefixos intensivos, o conjunto com *super-* chegou a quase cinquenta unidades.

Isso confirma as palavras de Galliot, para quem o prefixo *super-* se torna “um prefixo exclusivamente publicitário” e mostra, nesse domínio, incrível vitalidade, violentamente maior “que todos os outros sobre nossos muros e em nossos jornais. Desde o século XVIII, serviu para criar adjetivos de valor superlativo (...), antes de criar os substantivos de mesmo valor, como nos modernos *super-chocolates*.” (5, p. 289)

2.5. Prefixos quantificadores

bi-

biapoiada:

“A Treliça M-150 pode ser usada *biapoiada* ou em balanço.” (VE 661, 06.05.81: 125, 1)

mono-

mono-cassete:

“A sofisticada tecnologia National criou o RQ-2211 M, o *mono-cassete* de maior potência em sua classe (...)” (VE 671, 15.07.81: 106,2)

multi-

multimedição:

“Esse novo sistema, conhecido por *multimedição*, funciona das 8 da manhã às 8 da noite (...)” (IE 243, 19.08.81: 37, 1)

poli-

poliuretano:

““(...) graças ao melhor isolamento térmico que se conhece, a espuma rígida de *poliuretano*.” (VE 681, 23.09.81: 104)

tri-

tri-jato:

“E os nossos modelos de longo alcance das Séries 30 e 40 oferecem a economia do *tri-jato* em rotas internacionais (...)” (VE 658, 15.04.81: 36, 2)

2.6. Prefixos de valor espaço-temporal

co-
coaxial:

“Você pode usar o conjunto em sua residência, ou desmembrar o RG 800 estéreo *coaxial* para ter um maravilhoso som estereofônico em viagens ou passeios.” (VE 689, 18.11.81: 64)

inter-
intercambiáveis:

“Condicionar com porta reversível de painéis *intercambiáveis*.” (IE241, 05.8.81:7,3)

pós-
pós-fixada:

“Têm juros e correção monetária, podendo ser pré-fixada ou *pós-fixada*.” (VE 668, 24.06.81: 163, 2)

pré-
pré-combustão:

“Esse desempenho (...) é mais uma vantagem do projeto dos motores Caterpillar de câmara de *pré-combustão*.” (IE 215, 04.02.81: 75, 1)

retro-
retro-escavadeiras:

“Elas equipam as máquinas que geram o progresso, como os guindastes, empilhadeiras, pás-carregadeiras, *retro-escavadeiras*, tratores florestais (...)” (IE 220, 11.03.81: 8)

trans-
transceptor:

“O TT/109 é o *transceptor* mais vendido para entidades governamentais (...)” (VE 649, 11.02.81: 61, 1)

Pelo menos duas observações podem ser feitas a respeito desses neologismos constituídos a partir de prefixos de valor espaço-temporal. A primeira se refere aos elementos com *pré-*, os quais sintetizam até mesmo o tipo de civilização em que vivemos. A pressa, o tempo sempre curto, faz com que as pessoas privilegiem os produtos que possam oferecer vantagens e economia de tempo: uma casa *pré-fabricada*, um alimento *pré-cozido* — tudo isso tem muito valor para o homem contemporâneo. No *corpus* analisado, encontramos um extenso número de palavras formadas por *pré-*, contra apenas uma (*pós-fixada*) com *pós-*.

A segunda diz respeito ao vocábulo *transceptor*. Poder-se-ia limitar a análise a um simples caso de derivação prefixal que une *trans-* à base; parece-nos, entretanto, que existe no termo um pouco mais, aproximando-o quase de uma invenção, de uma espécie de aglutinação. O produto serve para transmitir e receber mensagens: é um rádio que exerce as funções de transmissor e de receptor, portanto um *trans(missor)(re)ceptor*.

2.7. Prefixo indicativo de inerência (ação ou movimento próprio)

auto-

auto-retráteis:

“Os cintos de segurança dianteiros de três pontos são *auto-retráteis* e inerciais.” (IE 236, 01.07.81: 27)

De significado especial, o prefixo *auto-* é muito utilizado no vocabulário da publicidade, por envolver a marca daquilo que faz por si mesmo, característica bastante apreciada em nossos dias.

Marcel Galliot afirma que o prefixo *auto-* é, “primeiro para a língua da tecnologia, depois para a língua do reclame, um elemento expressivo de escolha. Que mérito admirável, com referência ao aparelho que se inventa (ou se vende), que ele faça por si mesmo, tudo sozinho, o trabalho para o qual foi concebido, economizando assim todo o esforço ao comprador eventual!” (5, p. 284).

3. CONCLUSÃO

A comparação da quantidade de neologismos formados pela prefixação com os originados pelo processo de sufixação evidencia, como já se frisou, uma produtividade muito maior dos primeiros.

Algumas conclusões relevantes, que procuram explicar essa preponderância dos elementos prefixados, podem ser resumidas abaixo (cf. 5, p. 296):

a) o prefixo é *menos soldado* à palavra que o sufixo; aquele continua a ser sentido como elemento separado, na maioria dos casos;

b) assim sendo, a significação do prefixo é *mais nítida e estável* que a dos sufixos, apesar de algumas vezes os prefixos também comportarem mais de uma acepção;

c) o número de prefixos é *menos limitado* que o dos sufixos; em teoria, toda palavra pode vir a tornar-se um prefixo;

d) essa liberdade de criação tem como conseqüência a passagem para a língua corrente de certos prefixos particularmente empregados na publicidade, assim como a integração de determinados prefixos ditos “clássicos”, como *neo-* e *anti-*, para o domínio usual dos falantes comuns, por intermédio da alta freqüência na língua do anúncio.

Finalmente, deve-se frisar que a análise dos neologismos ocorrentes nos textos publicitários permitiu constatar que, apesar do enorme prestígio das línguas estrangeiras, é o próprio português que alimenta — para a maioria de suas designações — o anúncio português. Nesse sentido, o número de neologismos chamados internos foi muito superior aos denominados externos, confirmando a observação de Galliot, referente ao francês. (5, p. 224).

REFERÊNCIAS BIBLIOGRÁFICAS:

1. ALVES, I. M. — Observações sobre a prefixação intensiva no vocabulário da publicidade. *Alfa*, São Paulo, 24: 9-14, 1980.
2. BOULANGER, J. C. — Néologie et terminologie. *Néologie en marche*, Montreal, 4,B: 7-126, 1979.
3. DUBOIS, J. — *Le vocabulaire politique et social en France de 1869 à 1872*. Paris, Larousse, 1962.
4. FERREIRA, A. B. de H. — *Novo dicionário da língua portuguesa*. Rio de Janeiro, Nova Fronteira, 1975.
5. GALLIOT, M. — *Essai sur la langue de la réclame contemporaine*. Toulouse, Privat, 1955.
6. GUILBERT, L. — *La créativité lexicale*. Paris, Larousse, 1975.
7. HAMPEJS, Z. — Para o estudo da linguagem da imprensa brasileira contemporânea. *Revista de Filologia Portuguesa*. Rio de Janeiro, 6(1): 51-114, 1961.
8. LAGNEAU, G. — La course au trésor (Prolégomènes à une analyse raisonnée du langage publicitaire). *Communications*, Paris, 17: 82-98, 1971.
9. MARTINET, A. — *Elementos de lingüística geral*. Tradução e adaptação para o português de Jorge Morais Barbosa. 6. ed. Lisboa-Rio de Janeiro, Martins Fontes-Presença, 1975.
10. PEREIRA, R. F. — *Neologismos na mensagem publicitária*. Assis, Instituto de Letras, História e Psicologia, Unesp, 1983. (Dissertação de Mestrado)

GLOSSÁRIO

Maria Tereza Camargo BIDERMAN*

abonação

Frase ou enunciado, extraído de um bom autor da língua, onde ocorre a palavra que está sendo definida e/ou explicada no dicionário. Ilustra o uso da palavra no contexto. Cf. documentação.

acepção

Cada um dos sentidos ou significados de uma palavra polissêmica, isto é, que tem vários valores semânticos. Cf. polissemia.

algoritmo

Conjunto de regras operatórias, ou seqüência de raciocínios e/ou operações que podem fornecer a solução de um problema. Em Linguística Computacional chama-se de *algoritmo* um processo de cálculo que associa de uma maneira unívoca um conjunto de dados lingüísticos a um conjunto de resultados. Esse processamento se faz por um conjunto finito de operações e de regras explícitas e obrigatórias, que devem ser compreendidas e executadas numa determinada ordem por um computador. Um *algoritmo* pode, por exemplo, executar automaticamente a decomposição morfológica de um enunciado. Cf. analisador sintático e semântico.

analisador sintático (parser) e semântico

Um *analisador sintático e semântico* é constituído de um conjunto de algoritmos, elaborados e ordenados de tal forma que um computador possa analisar qualquer frase ou enunciado que se lhe apresente, desde que a máquina siga as instruções do analisador. Em princípio, um *analisador sintático e semântico* foi criado para analisar uma determinada língua. Como as línguas são diferentes estruturalmente, faz-se necessário criar um analisador para cada língua, ou adaptar um modelo que se tenha revelado eficiente em outra língua.

antônimo

Palavra que se opõe a uma outra pelo sentido. Ex.: *frio* é antônimo de *quente*; *bonito*, de *feio*.

arquilexema

Palavra superordenada semanticamente em relação a outras palavras porque o seu significado é mais geral; assim, pode aplicar-se a vocábulos mais específicos, podendo ser usado como um substantivo para esses últimos. Ex.: as palavras *coisa*, *objeto* são arquilexemas; não contêm nenhum sema (unidade de significação) específico. Podem, portanto, substituir palavras como: *cadeira*, *banco*, *automóvel*. Nesse tipo de unidade

* Departamento de Linguística — Instituto de Letras, Ciências Sociais e Educação — UNESP — 14800 — Araraquara — SP.

léxica (arquilexema) estão neutralizados todos os traços específicos dos lexemas individualizantes como *cadeira*, *banco*, *automóvel*.

arquivo

Conjunto de dados informativos armazenados, depois de terem sido classificados e ordenados, a fim de serem utilizados posteriormente. Em Lexicografia é o conjunto dos dados léxicos que servirão para a redação de um dicionário. A documentação sobre esses dados, classificados em fichários, tem, como unidade de informação, a ficha. Em Informática é qualquer conjunto de dados (banco de dados), classificados e ordenados para poderem ser recuperados facilmente quando se fizerem necessários. Em Linguística Computacional tem o mesmo significado; nesse caso, os dados do arquivo ou banco de dados são relativos à língua, ou a várias línguas.

binário

Composto de duas unidades ou dois elementos. Em matemática um sistema binário é aquele que opera apenas com duas unidades nas suas regras operatórias. O computador é uma máquina que opera na base binária.

binarismo

Sistema baseado no princípio da oposição binária. Muitos lingüistas reconhecem nas línguas a presença e atuação deste princípio. Assim, encontramos em fonologia claras evidências de oposição binária entre os fonemas. No português, por exemplo, surdos (/p/, /t/, /k/) versus sonoros (/b/, /d/, /g/). No léxico também esse princípio parece atuar na estruturação dos significados, mas de maneira menos sistemática. Por exemplo: *homem X mulher*; *frio X quente*; *menino X menina etc.*, *campo léxico ou lexical*, *campo semântico*.

Conjunto de unidades léxicas que se referem mutuamente e cujos valores semânticos se delimitam reciprocamente. Pode-se compor tanto de significantes derivados de uma mesma base formal (*estudo*, *estudante*, *estudar*, *estudioso*) como dos significados das palavras de valor semântico próximo (*trabalho*, *labor*, *tarefa*, *carga*, *labuta*). Quando se trata do conjunto dos significados, o *campo léxico* pode ser também denominado *campo semântico*. Cf. rede semântica.

código

Um *código* é um sistema de signos (ou sinais) que serve para transmitir mensagens e/ou informações. Ele é constituído por um conjunto de elementos e um conjunto de regras para combinar esses elementos. Ex.: o código de trânsito, uma língua qualquer.

cognato

Palavra que tem a mesma raiz de outra, ou outras, e com que constitui família de palavras. Ex.: *claro* e *clareza* são cognatos. Outros elementos dessa mesma família de cognatos: *claridade*, *claramente*, *clarear*, *clarificar*, *esclarecer*.

concordância

A *concordância* é um conjunto de citações ordenadas de uma dada palavra. A palavra pode aparecer à direita ou à esquerda da citação, ou contexto, e pode vir na forma invariável (subst. masc. sing.; verbo, infin.) ou nas suas formas flexionadas em ordem alfabética. As palavras que são postas em concordância seguem-se em listagem alfabética. As citações extraídas de obras são acompanhadas de códigos que permitem localizá-las no texto de origem. Existem concordâncias de autores, concordâncias de

obras (as *concordâncias* da bíblia, p.ex.). Há séculos zelosos lexicólogos vêm elaborando *concordâncias*. As bíblicas, por exemplo, são muito antigas. Hoje elas podem ser executadas automaticamente por um computador, a partir de um programa especialmente construído para este fim. Constituem uma fonte riquíssima de informações para a tarefa lexicográfica, poupando aos dicionaristas um gigantesco trabalho braçal na coleta de abonações para as entradas do seu dicionário. Em nossos dias as *concordâncias* textuais substituíram o arquivo do lexicógrafo nos grandes centros de Lexicografia.

conjunto aberto e conjunto fechado

Um *conjunto fechado*, ou *finito*, é um conjunto de elementos que podem ser circunscritos, enumerados. Em Linguística os fonemas de uma língua, os pronomes de um idioma constituem um *conjunto fechado*. Inversamente, um *conjunto aberto*, ou *infinito*, é aquele cujos elementos não podem ser perfeitamente enumerados e circunscritos. O léxico constitui um conjunto aberto. De fato, ele se caracteriza por sua permanente possibilidade de ampliação, à medida que uma cultura se expande e avança o conhecimento. O léxico só deixaria de ser um sistema aberto se a língua correspondente morresse, isto é, deixasse de ser falada por grupos sociais como ocorreu com o latim.

conotação — conotativo

Elemento da significação de uma palavra que não é comum a todos os falantes e pode variar segundo os contextos em que a palavra em questão está inserida. Um vocábulo como *luz* terá um valor semântico básico (denotação) e poderá ter inúmeros valores *conotativos* (abstrato, literário, emotivo, religioso etc.), dependendo do contexto em que ele aparecer e da pessoa que emite ou interpreta a mensagem.

corpus

Conjunto geralmente grande de dados lingüísticos que servem como matéria-prima para a pesquisa e a análise lingüística. Este termo começou a ser utilizado inicialmente na Estatística Léxica, onde os estudiosos levantaram grande volume de palavras em uma língua, para estudar o seu comportamento quantitativo, a sua distribuição na língua e no discurso. Hoje é usado em todos os domínios da Linguística com o significado acima definido.

definição ostensiva ou mostrativa

É um tipo de definição que utiliza a mostra do referente da palavra como recurso, em virtude da dificuldade de a linguagem verbal identificar claramente o *definiendum*. É muito freqüente nos dicionários. Ex.: *azul* = que é da cor do céu, do mar profundo, da água das piscinas. Esse tipo de definição torna-se muito apropriado quando o referente é um objeto, um ente, ou uma qualidade perceptível no mundo extralingüístico. O caso típico é o dos animais e plantas, em que os dicionários maiores e melhores geralmente mostram o referente da palavra através da sua imagem: um desenho, uma fotografia etc.

Cf. *definiendum*.

definiendum

A palavra-entrada do dicionário que deve ser definida pelo dicionarista.

definiens

A palavra (ou palavras) usada(s) para definir o *definiendum*.

denotatum — pl. denotata

Referente do signo lingüístico. No caso do signo *árvore*, a planta que se vê na natureza. Os *denotata* também podem ser abstratos sem correspondente no mundo extralingüístico: *constituição, sabedoria*. L.Zgusta (V. O dicionário unilingüe) usa muito estes termos (*denotatum, denotata*): a maioria dos lingüistas, porém, prefere usar referente.

designata

A coisa designada, a palavra designada, o que se refere, isto é, o referente.

dicionário bilíngüe

Dicionário que trata da correspondência das palavras entre duas línguas, procurando indicar a tradução de um termo de uma língua na outra língua considerada. Ex.: dicionário português-inglês, dicionário inglês-português.

dicionário unilingüe

Dicionário que contém as palavras de uma língua e as explica e/ou define através de paráfrases nessa mesma língua. Também chamado de *dicionário monolíngüe*. Contudo, a formação etimológica *unilingüe* é preferível por não ser heterogênea.

documentação

Enunciados, frases, parágrafos, onde está documentado o uso de uma palavra determinada numa obra literária ou científica, ou em qualquer outro tipo de escrito (jornalismo, p.ex.).

empréstimo

Palavra que foi tomada emprestada de uma outra língua por não existir termo correspondente no idioma que acolheu o empréstimo. Esse *empréstimo*, com freqüência, incorpora-se ao vocabulário vernáculo.

entrada

Cada uma das palavras explicadas por um dicionário. Em geral as entradas são alistadas em ordem alfabética e na forma canônica ou lema (Cf. essas palavras).

estatística léxica

Ciência que estuda o vocabulário do ponto de vista quantitativo. Os numerosos estudos estatísticos feitos sobre o léxico das mais variadas línguas evidenciaram que existem certas características estatísticas que aparecem em todas as línguas. A freqüência do uso de determinadas palavras (os instrumentos gramaticais, p.ex.) tem uma tal estabilidade que poderia ser considerada como um traço distintivo dessas palavras ao lado de outras características morfossintáticas. Outra característica que se repete em todas as línguas sobre as quais se fizeram estatísticas léxicas rigorosas é que as 1.000 palavras mais freqüentes de uma língua compõem mais de 80% de qualquer texto nessa língua. Esse índice pode chegar a quase 89%.

etiqueta

Rótulo qualificativo e explicativo que se atribui a cada entrada, ou a cada sentido, ou acepção de uma palavra. Essas *etiquetas* são geralmente abreviadas no contexto para não sobrecarregar o texto do verbete. Por exemplo: *abst.* (= significação abstrata);

met. (= significado metafórico); *col.* (= usado na linguagem coloquial); *reg.* (= regionalismo, só usado em determinada região); *arc.* (= arcaico, desusado, antigo) etc.

exemplário

Conjunto de exemplos (= frases, enunciados, contextos) utilizados para exemplificar o(s) significado(s) e o(s) uso(s) de uma palavra-entrada. Cf. abonação, documentação.

fonte

Autor ou obra que forneceu ao dicionarista um exemplo, ou uma citação com que ele abonou o significado e/ou o uso de uma palavra.

forma canônica

Forma de uma palavra considerada como forma-base para dar a ela entrada em um dicionário. As palavras que sofrem flexões morfossintáticas variadas e variáveis no discurso têm sido tradicionalmente representadas por uma das variantes do seu paradigma. Por exemplo: no caso do adjetivo, o masculino singular (*belo, bonito, amável, feliz*); no caso do verbo, a forma infinitiva (*amar, vender, partir*). Cf. lema.

glossário

Pequeno vocabulário, ou relação de palavras, em que se explica o significado das mesmas, para ajudar o leitor na compreensão do texto que lê. Modernamente são comuns os glossários de linguagem técnica como este.

índice de palavras

Também chamado *index-verborum*.

É uma lista das palavras que ocorreram em um determinado *corpus* (cf. essa palavra), ordenadas segundo dois critérios: 1) em ordem decrescente (ou crescente) da frequência com que ocorreram no *corpus*; 2) em ordem alfabética sendo indicada a frequência com que cada palavra ocorreu no *corpus*, à direita do vocábulo.

key-word-in-context

O mesmo que concordância (cf. essa palavra).

lema

Unidade léxica ideal que representa um paradigma de formas flexionadas. Essa unidade constitui a típica entrada de dicionário e representa todas as demais formas do paradigma. Por exemplo: a forma infinitiva *cantar* que aparece nos dicionários de português representa todas as outras variantes deste paradigma verbal: *canto, cantara, cantasse, cantando* etc. É, portanto, um *lema*. Cf. forma canônica.

lematização

O ato de reduzir uma forma que ocorreu num texto ao lema que passa a representá-la numa lista de palavras, ou num dicionário. Consideremos a oração: "Os dias correm céleres." A *lematização* dessa seqüência produziria os lemas: *o, dia, correr, célere*. Cf. forma canônica.

lexema

Unidade léxica abstrata que faz parte do léxico de uma língua. Ele se atualiza no discurso na forma de uma palavra flexionada com todas as marcas gramaticais exigidas pelo contexto. Cf. lexia, palavra.

lexia

Forma que um lexema assume no discurso. Ex.: “O dia está claro.” Temos aí quatro lexias. Em Lexicologia esse termo técnico se opõe a *lexema*; foi cunhado para evitar a imprecisão e ambigüidade de termos da língua comum, tais como: *palavra*, *vocábulo*.

lexicalização

O processo formal de categorizar, ao nível do léxico, uma forma, ou uma combinatoria de formas usadas no discurso. Ocorre lexicalização na evolução histórica de uma língua: 1) quando um *empréstimo* estrangeiro passa a fazer parte do léxico da língua (ex.: *coquetel*, *futebol*, *butique*); 2) quando uma criação vernácula nova é definitivamente aceita pela comunidade de falantes do idioma; 3) quando uma seqüência de formas passa a ser compreendida e usada como se fosse uma forma única, uma unidade léxica (ex.: *mercado negro*, *capa-de-chuva*).

lexicologia

Ciência que se ocupa do estudo do vocabulário de uma língua. Ela procura estudar o léxico enquanto sistema, e os seus elementos constitutivos nas suas peculiaridades. Opõe-se à Lexicografia que procura descrever o léxico, elaborando dicionários para consulta dos usuários da língua. A *Lexicologia* tem objetivos teóricos e especulativos. Tenta responder a questões como: o que vem a ser o léxico de uma língua quando confrontado com a gramática dessa mesma língua? como ele se estrutura? como ele se expande? como se dá o processo de categorização lexical? como se processa a aquisição e o enriquecimento do léxico de um indivíduo? como se faz o armazenamento do léxico nas memórias individuais? etc.

língua culta padrão

Modelo de linguagem considerado como padrão ideal numa comunidade lingüística, ou numa nação, se a língua for considerada relativamente às suas funções socioculturais e políticas. Coincide com a linguagem do grupo que possui maior competência lingüística e, por essa razão, goza de prestígio na sociedade: os intelectuais e os escritores.

memória do computador

Numa unidade central de processamento computacional existem vários equipamentos com finalidades distintas e com capacidade de operação diversa. Entre elas a *memória* principal que pode se subdividir em duas: 1) a *memória* que armazena os dados que serão analisados e processados; 2) a *memória* que armazena o programa (ou programas), isto é, a lista ordenada de instruções (algoritmos) que deverão ser seguidas na análise dos dados. Essas *memórias* podem ser ampliadas. Nos grandes centros de computação como, por exemplo, num sofisticado centro de pesquisa, ou num centro que opera milhões de dados, outras unidades de memória podem ser acopladas à memória principal a fim de ampliar a capacidade de armazenamento da máquina, tanto com relação a dados como com relação às instruções operatórias. Cf. algoritmo, analisador sintático e semântico.

metalingua

Um *metalingua* ou *metalinguagem* é um sistema de significação (uma linguagem), cujo objeto de descrição é um código e os seus signos. Um exemplo típico de *metalinguagem* é a definição e/ou descrição de uma palavra fornecida pelo dicionário.

Ex.: *crise* = / período difícil onde há dúvidas, incertezas e dificuldades antes de se poder chegar a uma solução que resolva a crise./ A *metalinguagem* dos dicionários usa os próprios signos da língua natural para explicar um signo do código lingüístico. Um outro tipo de *metalingua* (metalinguagem) se situa num nível mais abstrato. Opera com conceitos e signos lógico-matemáticos e se serve de uma teoria para analisar um objeto (uma língua natural, p.ex.), sem se reportar necessariamente a manifestações observáveis.

neologia

Segundo Louis Guilbert, grande especialista em Lexicologia, “a *neologia* lexical se define pela possibilidade de criação de novas unidades léxicas, em virtude das regras de produção existentes dentro do sistema léxico”.

neologismo

Unidade léxica (palavra) que é sentida como recente na língua pelos falantes do idioma. Existem *neologismos* em que o signo lingüístico todo é novo (forma e significado) — *jeans, fríser* — e outros em que só o significado é novo — *computador, átomo*. O estudo dos *neologismos* se relaciona diretamente com a história da língua; podemos estudar os *neologismos* do século XVI, os *neologismos* do século XIX, bem como os *neologismos* do século XX, ou das décadas 1950-80.

nomenclatura

A seqüência dos itens lexicais (as entradas do dicionário), ordenados geralmente em ordem alfabética, é o que os lexicógrafos chamam de nomenclatura.

obsoleto

Qualidade da palavra que não faz mais parte do vocabulário corrente, usual. Portanto, desusado, arcaico.

ocorrência

Unidade léxica que se atualiza no discurso produzido (oral ou escrito). Ex.: No verso de Fernando Pessoa “Ah, canta, canta sem razão” — temos duas ocorrências do lexema <cantar>

palavra

Unidade psico-sociológica fundamental da língua, essencial tanto no processo de comunicação, como no processo simbólico de apreensão do universo pelos sujeitos. É termo da língua comum, sendo pouco rigoroso para o uso técnico da Lexicologia e da Lexicografia. Charles Muller, em várias de suas obras, tentou atribuir aos termos *mot* (= palavra) e *vocable* (= vocábulo) valores técnicos na língua francesa para os seus trabalhos de Estatística Léxica (cf. essa palavra). A meu ver, essa tentativa se frustrou pelo mero fato de que essas palavras estão desgastadas pelo uso comum.

palavra instrumental

Palavra de valor e uso gramatical como preposições, conjunções, artigos etc. É um instrumento de articulação do discurso.

palavra lexical

Palavra de significação plena, não gramatical. Dentro do universo vocabular é

aquela que faz parte do conjunto de elementos exclusivamente lexicais, já que o outro conjunto é híbrido, pois contém elementos de natureza gramatical. Cf. palavra instrumental.

palavra plena

O mesmo que *palavra lexical*.

polissemia

Fenômeno lingüístico da acumulação de vários significados em um único lexema (palavra). Assim, um significante como *cavalo* pode significar: 1) animal quadrúpede e mamífero; 2) peça do jogo de xadrez; 3) tronco de uma planta onde se faz um enxerto. Toda língua possui um número enorme de palavras polissêmicas, pois esse fenômeno resulta do princípio da economia operante na língua. Em geral, quanto mais usada e mais antiga for uma palavra na língua, tanto maior costuma ser a pluralidade dos seus significados. É o caso de *ato*, *fazer* e *fundo*, p.ex.

recuperação de dados

Em Lingüística Computacional, Informática e Lexicografia, chama-se *recuperação de dados* o ato de retirar de um arquivo, ou de um banco de dados, determinados dados que se deseja num determinado momento.

Quando os dados estão estocados num disco que pode ser posto na memória do computador (cf.essa palavra), podemos recuperar informações que tinham sido anteriormente armazenadas no disco, através de um programa que dá as instruções adequadas ao computador.

recuperação de informação

O mesmo que recuperação de dados.

rede semântica

Uma rede semântica é um conjunto articulado e estruturado de campos *léxicos* (cf. essa palavra) que se comunicam entre si. Os campos léxicos vão-se integrando uns nos outros, através de nós sêmicos comuns (unidades de significação), a fim de compor a grande cadeia de uma rede semântica.

referência cruzada

No texto de um verbete (entrada de dicionário) é freqüente o dicionarista remeter a outra palavra. Essa prática se explica da seguinte forma: além de economizar espaço no dicionário, evita-se repetir informações que já foram dadas em outro verbete e por isso se remete a ele. Por outro lado, para que o consulente compreenda bem o significado e o uso de uma palavra, é preciso contrapô-la a outras palavras de significação próxima, ou oposta. Ninguém aprende ou apreende uma unidade de léxico isoladamente, uma vez que o léxico é formado de campos léxicos e grandes redes semânticas integradas. Muitas vezes, é só através da compreensão de um campo léxico, ou de uma rede, que o usuário poderá tomar consciência exata do valor de uma palavra.

referente

Em Lexicologia e Semântica chama-se de *referente* o objeto ou conceito que é expresso através de um signo lingüístico (uma palavra ou uma locução); a coisa, a noção à qual é feita referência. Ex.: o *referente* de *cavalo* (1.º significado, cf. polissemia) é o

animal quadrúpede; o *referente* de *saudade* é o sentimento que se experimenta e que essa palavra refere.

registro

Um registro lingüístico é uma variante da língua (com gramática e vocabulário próprios), utilizado em função de determinados fatores humanos e sociais. São fatores determinantes para a seleção do registro a ser usado pelo falante: 1) o contexto e a situação; 2) os interlocutores em presença; 3) o tema do discurso. Os registros derivam da natureza social da linguagem e das atitudes sociais dos indivíduos em relação à língua.

repertório léxico (vocabular)

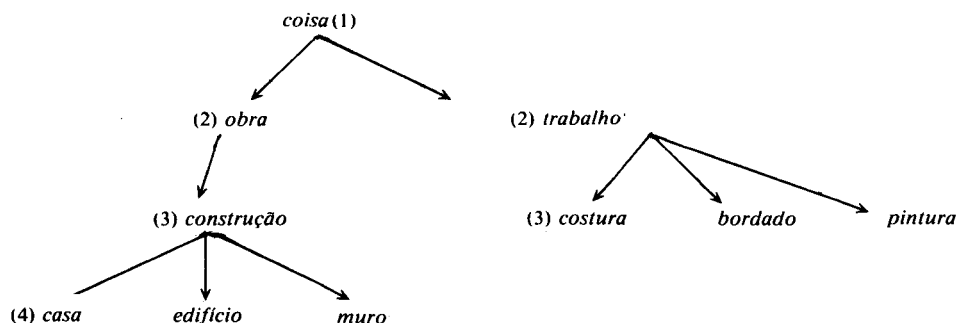
Conjunto das palavras de uma língua, ou universo léxico. Pode-se falar também no *repertório léxico* (vocabular) de um indivíduo, de uma obra, de um grupo social, dando a esse termo um significado mais restrito.

semasiologia

A *semasiologia* é uma área da Semântica que estuda os significados e a sua estruturação interna, a partir dos signos lingüísticos (das palavras). Divide-se assim o signo lingüístico nas suas duas faces — significante e significado. O método semasiológico considera os significantes para indagar sobre os significados, ou investigar o fenômeno da significação. O contrário da *semasiologia* é a *onomasiologia*, que parte da significação em busca da designação lingüística dos conceitos ou objetos considerados.

superordenado

Os lexicólogos admitem que o léxico de uma língua está estruturado não só em campos léxicos (cf. essa palavra) e redes semânticas (cf. essa palavra), mas também de maneira hierárquica. De acordo com essa última característica, o vocabulário se ordenaria em cadeias hierarquizadas, onde alguns termos estão situados teoricamente acima de outros, isto é, são superordenados em relação aos que estão mais abaixo na cadeia em árvore. Ex.:



(1) é superordenado em relação a (2), (3) e (4)

(2) é superordenado em relação a (3) e (4)

(3) é superordenado em relação a (4)

tesouro ou thesaurus

Dicionário completo que deveria conter todas as palavras de uma língua, descrevendo o universo vocabular desse idioma na sua totalidade.

Não só a sua nomenclatura deverá ser exaustiva, como a redação dos verbetes, incluindo-se neles todas as denotações e conotações possíveis da palavra-entrada. Ainda mais: deverá conter uma exemplificação e documentação completa dos valores e usos de cada palavra; para tanto, a equipe responsável pela elaboração de um *thesaurus* deverá compilar todos, ou quase todos, os textos escritos da língua para evitar omissões. Uma tarefa dessa é evidentemente gigantesca. Só é imaginável nos tempos contemporâneos quando grandes equipes de especialistas se reuniram em centros de pesquisa especialmente criados para este fim, como é o caso do *Trésor de la langue française*, que se elabora há mais de vinte anos em Nancy (França) com o concurso de computadores e de grandes verbas do governo francês. No caso de uma língua morta como o latim, todo dicionário deveria ser do tipo *thesaurus*, já que o universo léxico dessa língua se fechou, pois não existem mais falantes desse idioma e, portanto, não se criam mais neologismos (cf. essa palavra).

verbeta

O texto de uma palavra-entrada de um dicionário, inclusive ela própria. Os dicionários são formados de seqüências de verbetes.

vocábulo

Sinônimo de *palavra* na língua comum; portanto, é ambíguo e não tem os requisitos de rigor técnico necessários à linguagem científica, a saber: um significado unívoco. Para completar as informações sobre *vocábulo* consulte o verbeta *palavra*.

unidade léxica

Unidade vocabular tanto com respeito à significação como com respeito à forma gramatical, que tem um uso característico no discurso. Sinônimo de *lexema* (cf. essa palavra), ou de *lexia* (cf. essa palavra) conforme o contexto.

usuário

1) O consulente de um dicionário; o indivíduo que usa o dicionário para informar-se sobre o significado e o uso de uma palavra. 2) O falante da língua, aquele que usa a língua para comunicar-se.

SUBJECT INDEX V. 28(supl.)

- Computer
machine dictionary
organization, p.1
- Contemporary advertising
new words formation
neological prefixation, p.127
- Designation
words and things relations, p.45
- Dictionaries — Typology
English Language
Latin languages, p.1
- Dictionaries writing
“Definiendum” definition
metalinguage
words definition, p. 103
- Foreignism
loan translation neologism
Portuguese language, p.119
- Grammar
basic concepts
lexicology, p.45
- Lematization
lexie
verbal polynomial
lexical categorization, p.71
- Lexicography
“acception” concept
meaning discrimination
polysemic words structure
metaphorical meanings, p.71
- Lexicography — History
French dictionary
Portuguese dictionary
Spanish dictionary
XVIth to XXth century, p.1
- Lexicology
lexicon and grammar
basic concepts, p.45
- Loan translation neologism
foreignism
Portuguese language, p.119
- Machine dictionaries
organization
computers, p.1
- Neological prefixation
new words formation
contemporary advertising, p.127
- Pattern dictionaries
organization
compilation of a corpus
homonymy
lexical entries
lexicographical definition
polysemy
entries writing, p.27
- Portuguese language
loan translation neologism
foreignism, p.119
- Unilingual dictionaries
organization
compilation of a corpus
homonymy
lexical entries
lexicographical definition
polysemy
entries writing, p.27

ÍNDICE DE ASSUNTOS V. 28 (supl.)

- Computação**
dicionários de máquina
elaboração, p.1
- Designação**
relação palavras e coisas, p.45
- Dicionários**
de máquina
elaboração
 computação, p.1
padrão
elaboração
 compilação de corpos
 definição lexicográfica
 entradas lexicais
 homonímia
 polissemia
 redação de verbetes, p.27
redação
 definição das palavras
 definição do "Definiendum"
 tipo de metalingua, p.103
tipologia
 língua inglesa
 línguas latinas, p.1
unilingüe
 elaboração, p.27
- Estrangeirismo**
neologismo por empréstimo
 língua portuguesa, p.119
- Gramática**
conceitos básicos
 lexicologia, p.45
- Lematização**
lexias
 polinômio vocabular
 categorização lexical, p.71
- Léxico**
conceitos básicos
 lexicologia, p.45
- Lexicografia**
conceituação da acepção
 discriminação das acepções
 ordenação de palavras polissêmicas
 valores metafóricos, p.71
- Lexicografia — História**
dicionários espanhóis
dicionários franceses
dicionários portugueses
 séculos XVI ao XX, p.1
- Lexicologia**
léxico e gramática
 conceitos básicos, p.45
- Língua portuguesa**
neologismo por empréstimo
 estrangeirismo, p.119
- Neologismo por empréstimo**
estrangeirismo
 língua portuguesa, p.119
- Prefixação neológica**
formação das palavras novas
 propaganda contemporânea, p.127
- Propaganda contemporânea**
formação das palavras novas
 prefixação neológica, p.127

ÍNDICE DE AUTORES
SUBJECT INDEX
V. 28 (supl.)

ALVES, I.M.	p. 119
BIDERMAN, M.T.C.	p. 1,27,135
CASARES, J.	p. 71
FEIJÓO-HOYOS, B.L., trad.	p. 71
LIMA, M.C.P.B., trad.	p. 103
MORAIS, C.B. de, trad.	p. 45
PEREIRA, R.F.	p. 127
REY-DEBOVE, J.	p. 45
WEINREICH, U.	p. 103

NORMAS PARA APRESENTAÇÃO DE ORIGINAIS

ALFA publicação da UNESP tem por finalidade divulgar trabalhos inéditos, comunicações e resenhas sobre lingüística redigidos em português ou em língua estrangeira (espanhol, francês, italiano, inglês ou alemão) elaborados por docentes da Universidade ou por outros especialistas. Só serão resenhados livros que tenham sido publicados nos dois últimos anos, em se tratando de obras nacionais e quatro anos para as publicações estrangeiras, cabendo à Comissão de Redação a seleção dos trabalhos para publicação.

Os trabalhos deverão ser encaminhados diretamente ao Diretor da publicação em três vias, em uma só face, espaço duplo e em folha de papel tamanho ofício, evitando cortar as palavras no final da linha, mesmo que a margem fique irregular. Deverão obedecer a seguinte seqüência: Título, subtítulo (quando necessário); Autor(es) por extenso e apenas o sobrenome em maiúsculo; Filiação científica em nota de rodapé, indicada por asterisco; Texto.

Citação no texto — Os autores referenciados serão indicados pelo número da referência. Acrescenta-se o número da página, em caso de citação textual ou quando o autor julgar necessário. Caso a clareza do texto o exigir, o articulista poderá mencionar, entre parêntesis, também o sobrenome do autor. Ex.: (ANDRADE, 5, p. 8).

Tabelas — Serão numeradas consecutivamente com algarismos arábicos e encabeçadas pelo seu título.

Ilustrações — (fotografias, gráficos, desenhos, mapas etc.) serão designados no texto como "figuras" (fig.) e numerados seqüencialmente com algarismos arábicos.

Desenhos e gráficos — Deverão permitir perfeita reprodução em clichês com redução de 6,5cm. Os desenhos serão feitos em papel vegetal e tinta nanquim. Deverá ser indicada no texto a localização das ilustrações. Quando as ilustrações excederem a quatro a Comissão de Redação reserva-se o direito de solicitar a redução de seu número.

Observações, aditamentos e pormenores do texto poderão aparecer em notas de rodapé, indicados por asterisco.

MATERIAL DE REFERÊNCIA

Resumos: Constará do resumo: citação bibliográfica, resumos em português antecedendo o texto e outro em inglês no final do trabalho, antes da referência bibliográfica. Não deverão exceder a 100 palavras. Para o resumo em inglês deverá ser adotada a palavra ABSTRACT.

Unitermos — Palavras ou expressões que identifiquem o conteúdo do artigo. Os três principais unitermos serão escritos em primeiro lugar. Para o termo em inglês deverá ser adotada a palavra KEY-WORDS.

Quando o artigo for em língua estrangeira o resumo final do trabalho deverá ser em português.

REFERÊNCIAS BIBLIOGRÁFICAS — Devem ser pela ordem alfabética do sobrenome do autor, numeradas consecutivamente e apresentadas de acordo com os exemplos abaixo:

Artigos: MÉNARD, N. — Richesse lexicale et mots rares. Le Français moderne, Paris 46(1):33-43, 1978.

Livros: CARDOSO, E. — Guia de conversação português-inglês. 2. ed. Lisboa, Betrand, 1971.

FRANÇOIS, D. — A noção de norma em lingüística. In: MARTINET, J. et alii — *Da teoria lingüística ao ensino da língua*. Rio de Janeiro, Ao Livro Técnico, 1979. p. 87-97.

Os dados e conceitos emitidos nos trabalhos bem como a exatidão das referências bibliográficas são de responsabilidade exclusiva dos autores.

SEPARATAS — Serão distribuídas 25 (vinte e cinco) separatas ao primeiro autor do trabalho.

Os trabalhos que não se enquadrarem nessas normas serão devolvidos aos autores para as necessárias adaptações que serão indicadas em carta pessoal.