

UM PONTO DE INTERSEÇÃO PARA A DIALECTOLOGIA E A LEXICOGRAFIA: A PROPOSIÇÃO DE ELABORAÇÃO DE UM DICIONÁRIO DIALETAL BRASILEIRO COM BASE NOS DADOS DO ALiB¹

A POINT OF INTERSECTION FOR DIALECTOLOGY AND LEXICOGRAPHY: THE PROPOSAL OF DEVELOPMENT OF A BRAZILIAN DIALECTAL DICTIONARY

Américo Venâncio Lopes MACHADO FILHO

Universidade Federal da Bahia-CAPES/Université Paris 13-COFECUB

RESUMO

O presente artigo visa discurrir sobre as bases metodológicas gerais de elaboração do Projeto **Dicionário Dialectal Brasileiro (DDB)**, uma obra de verve coletiva e interinstitucional, que pressupõe a congregação de uma equipe de especialistas, tanto no domínio da dialectologia, quanto no domínio da lexicografia e das ciências da informação.

Palavras-chave: Lexicografia dialetal, Atlas Linguístico do Brasil; Dicionário Dialectal Brasileiro

ABSTRACT

The present article aims at tracing the general methodologic basis to the Project **Dicionário Dialectal Brasileiro – DDB** (Brazilian Dialectal Dictionary), a collective and interinstitutional research work that will congregate a team of specialists in the domains of dialectology, lexicography and computer science.

Keywords: Dialectal lexicography; Linguistic Atlas of Brazil; Brazilian Dialectal Dictionary.

PREÂMBULO

(...) *un dictionnaire n'est et ne peut être a notre époque une oeuvre individuelle* (CATACH et al., 1971, p. 20).²

Embora tenha razão Piel (1991 [1976], p. 235) ao afirmar que nunca, em uma língua de cultura,

será possível reconstituir todas as fases (...) percorridas [pelo léxico] e destrinçar a contribuição das muitas gerações que nele colaboraram até se constituir o magno edifício que hoje se nos depara nos grandes dicionários modernos,

é certo que a Lexicografia contemporânea tem caminhado no sentido de perscrutar, a passos largos e de forma incontestavelmente sólida, esse movediço alicerce que serve de base às variegadas gramáticas das línguas naturais: o léxico.

Depois de longo tempo do que se poderia chamar de obscuridade científica, passou, nos últimos anos, essa entidade teórica ou, como preferem alguns chamar, nível de análise, a ser considerada pela linguística contemporânea – ao menos para alguns pesquisadores –, como "elemento central da língua" (VILELA, 1979, p. 17), tirando da sintaxe a hegemonia dos estudos linguísticos.

¹ Trabalho elaborado durante realização de Estágio Pós-Doutoral na *Université Paris XIII*, sob o financiamento do Governo Brasileiro, através de bolsa de estudos de quatro meses concedida pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES, com base no acordo de cooperação firmado entre esse organismo e o *Comité Français d'Evaluation de la Coopération Universitaire avec le Brésil* – COFECUB, da França.

² Tradução livre: "Um dicionário não é, nem pode ser, na nossa época, uma obra individual".

Não obstante esse arrojado ponto de vista em defesa do léxico (sobre o qual não se pretende aqui dar margens a uma discussão que se sabe longa), com o implemento dos estudos linguísticos, que passaram a se concentrar sobretudo em direção aos trabalhos de natureza variacionista – que têm em Weinreich, Labov e Herzog (1968) as bases teórico-filosóficas inaugurais –, os trabalhos de pesquisa de viés lexicográfico no Brasil pouco se têm desenvolvido no sentido de privilegiar a variação, mesmo porque, como bem registra Quemada (apud CATACH et al., 1971, p. 22): "*La présentation alphabétique des adresses suppose l'existence préalable d'une orthographe fixée ou d'un système de conventions défini*".³ O que se pode hoje observar em relação ao registro da variação no dicionários contemporâneos, publicados no Brasil, se refere meramente a marcas de uso, que normalmente refletem uma certa carga de preconceito em face do padrão ortográfico que neutraliza quaisquer outras atualizações linguísticas que se possam insinuar na nomenclatura.

Aliás, têm os dicionários de língua, preferencialmente – porém não exclusivamente –, utilizado textos escritos, na composição dos *corpora*, e como condicionador maior, obviamente, a norma-padrão, excluindo do processo de lematização os elementos que consideram os lexicógrafos como erros ortográficos, posição que não se pode considerar inapropriada na estrita perspectiva metodológica, embora seja em algum grau politicamente incorreta, que se diga, já que veladamente despreza outros usos linguísticos.

Entretantes, a questão do tratamento da variação lexical, no modelo variacional antes definido, tem, porém, cabido mormente à dialectologia no plano horizontal e, nesse âmbito, ao registro cartográfico das realizações "desviantes" ao presumido *standard* linguístico nacional. Mas como assegurar a difusão mais ampla do conhecimento da realidade linguística de uma nação, no que tange aos usos lexicais, de forma a atender, de maneira econômica e com maior amplitude, o acesso aos dados dialetais, por parte de seus supostos principais grupos de interesse ou público-alvo, isto é, os estudiosos da língua e da cultura e a população em geral?

Segundo Picoche (1973, p. 10-11):

*L'intérêt actif des spécialistes de la philologie et de la linguistique pour la dialectologie remonte aux alentours de l'année 1880, en France comme en Allemagne. (...) A partir de ce moment, les grands courants de l'évolution de la science linguistique influencèrent la recherche dialectologique, et, réciproquement, la dialectologie contribua, dans une certaine mesure, à cette évolution.*⁴

Esse interesse renovado pela dialectologia tem há muito se espreado para o espectro dos estudos lexicográficos, a ponto de, para Ronco (2004, p. 441-442),

les grands atlas linguistique nationaux de la première génération ont désormais perdu terrain par rapport à cette heuristique synthétique méthodologique qu'est le 'dictionnaire-atlas',

³ Tradução livre: "A apresentação alfabética das entradas supõe a existência prévia de uma ortografia estável ou de um sistema de convenções definido".

⁴ Tradução livre: "O interesse ativo de especialistas da filologia e da linguística pela dialectologia remonta aos anos de 1880, na França e na Alemanha. (...) A partir desse momento, as grandes correntes de evolução da ciência linguística influenciaram a pesquisa dialectológica e, reciprocamente, a dialectologia contribuirá, em certa medida, a essa evolução".

*c'est-à-dire un dictionnaire possédant certaines caractéristiques propres aux atlas ou, plus exactement, un atlas qui conserve certaines caractéristiques formelles d'un dictionnaire.*⁵

Sem concordar com a alegada "perda do terreno" do trabalho dialectológico, pelas óbvias razões de que não se pode imaginar a análise linguística sem o concurso da dialectologia, nem sem o devido registro de suas pesquisas de campo no formato próprio a essa atividade científica, de cariz cartográfico por excelência, vê-se, porém, na atitude de maior aproveitamento da base de dados dos projetos de elaboração de atlas linguísticos, uma oportunidade transdisciplinar de bom termo, nomeadamente para o avanço da lexicografia moderna.

Embora não seja a lexicografia dialetal uma atividade cuja especialidade já se registre em dicionários de linguística, sequer mesmo nos gerais de língua, já conta com uma larga produção, bastante considerável em alguns países, sobretudo na Europa, em que muito dialetos correm o risco de desaparecer com as pressões econômico-sociais do mundo moderno. Sobre isso e em defesa da construção de dicionários dialetais, dizem Barbato e Varvaro (2004, p. 431), em relação ao cenário na Itália:

*The more a dialect loses ground to Italian, until it is reduced to a second language, the more the dictionary becomes a necessary instrument, to learn it or at last to keep it alive.*⁶

Bastante clara, para as populações europeias enquanto elemento de identificação social, a noção de dialeto não parece ser, entretanto, muito transparente no Brasil para as suas populações, podendo confundir-se às vezes com a noção de desvio e erro ou com o conceito intuitivo que tem o falante sobre norma linguística. Isso talvez se deva ao fato de algumas características normais terem se difundido verticalmente nos diversos dialetos nacionais de forma quase uniforme, notadamente algumas marcas mais estigmatizadas socialmente, sobretudo nas populações de baixo letramento formal, fazendo com que o sentido de unidade linguística se tenha em algum grau se obliterado. Não obstante, para Cardoso (2008, p. 13), o Brasil,

na sua ampla extensão territorial — país-continente —, apresenta-se como uma terra de grandes contrastes, marcada pela heterogeneidade cultural, social e econômica que se vai refletir, também, na língua portuguesa, hoje majoritariamente falada. A diversidade da língua está, pois, vinculada à diversidade cultural tomada nos seus mais diferenciados aspectos.

E é à busca dessa realidade que se desenvolve hoje no Brasil, coordenado por um comitê nacional que integra diversas instituições de ensino superior, o Projeto *Atlas Linguístico do Brasil* (ALiB) que tem a sua sede na Universidade Federal da Bahia, sob a presidência e direção executiva das professoras/pesquisadoras Suzana Alice Marcelino da Silva Cardoso e Jacyra Andrade Mota, respectivamente, e com a participação das universidades federais do Ceará, da Paraíba, do Pará, de Juiz de Fora, de Ouro Preto, de

⁵ Tradução livre: "os grandes atlas linguísticas nacionais da primeira geração têm doravante perdido terreno em relação a essa feliz síntese metodológica que é o 'dicionário-atlas', isto é, um dicionário que possui certas características próprias aos atlas, ou mais exatamente, um atlas que conserva características formais de um dicionário".

⁶ Tradução livre: Quanto mais um dialeto perca terreno para o italiano, até que seja reduzido a uma L2, quanto mais o dicionário se torna em um instrumento necessário, para se aprender sobre sua realidade, ou pelo menos para o conservar vivo.

Mato Grosso do Sul, de Santa Catarina, do Rio Grande do Sul e da Universidade Estadual de Londrina. Colaboram, ainda, com o Projeto diversas outras instituições de ensino, públicas e privadas.

O projeto prevê a aplicação dos inquéritos em duzentos e cinquenta pontos do Brasil, estando já concluídos 77,6% das localidades previstas⁷. Os questionários utilizados são: a) Questionário Fonético-Fonológico (QFF), com cento e cinquenta e nove questões, em que se incluem questões para apuração de diferenças prosódicas; (b) Questionário Semântico-Lexical (QSL), com duzentas e duas questões, engloba catorze áreas temáticas ou conceituais; (c) Questionário Morfossintático (QMS), com quarenta e nove questões, busca desvendar fenômenos de variação diatópica nos usos morfossintáticos, como a própria denominação denuncia. O cronograma de execução encontra-se em fase de alimentação do banco de dados informatizado com as informações coletadas dos inquéritos, o que deve possibilitar um vasto registro da variação lexical existente no Brasil.



Figura 1: Capa da edição dos questionários utilizados na coleta dos dados (COMITÊ NACIONAL DO PROJETO ALiB, 2001)

⁷ Os dados atualizados podem ser encontrados em <http://www.alib.ufba.br/index.asp>.

Note-se, entretanto, que o primeiro atlas linguístico a ser elaborado na história do Brasil, o *Atlas Prévio dos Falares Baianos*, de autoria de Nelson Rossi *et al.*, imprimiu-se em 1963, estreitando uma tendência de investigação sobre a língua que iria, em poucos anos, contagiar diversos pesquisadores e grupos do País. Em pouco mais de quatro décadas depois da publicação desse atlas precursor, diversos mestres e doutores foram formados nessa área de investigação em universidades brasileiras, fazendo com que o sonhado *Atlas Linguístico do Brasil* passasse a ser, hoje, uma realidade em curso.

A ideia do aproveitamento dos dados do Projeto ALiB na perspectiva dos estudos lexicográficos começou sua trajetória com a primeira visita, à Universidade Federal da Bahia, do Diretor do LDI (*Lexique, Dictionnaire, Informatique* – CNRS UMR nº 7187), da *Université Paris 13*, Salah Mejri, que imediatamente propôs um acordo de cooperação entre as duas universidades, assim como a participação das equipes do LDI e do ALiB no Programa CAPES/COFECUB.

Com a aprovação pelas instâncias competentes dos dois países de ambas as empresas e dos respectivos projetos que foram submetidos a análise, deu-se continuidade à discussão da melhor utilização desse material, a partir da metodologia adotada no LDI, no que concerne ao tratamento informatizado do léxico. Inicialmente, se cogitou a elaboração de um dicionário especializado, tendo o léxico do futebol como tema. Essa ideia foi, posteriormente, substituída pela da construção de um dicionário dialetal brasileiro, cujas bases de concepção se discutem a seguir.

BASES PRÉVIAS PARA A ELABORAÇÃO DO DICIONÁRIO DIALETAL BRASILEIRO (DDB)

Como se sabe,

les dictionnaires ont dû bénéficier, au début du XVI^e siècle, des avantages de l'imprimerie et de ses conséquences économiques, sociales et culturelles. L'avènement de l'informatique et de ses circonstances socio-économiques et culturels exerce aujourd'hui une action de la même sorte sur le devenir des travaux et des produits lexicographiques (QUEMADA, 1983, p. 102).⁸

Essa nova realidade tem indissociado a construção de dicionários impressos da de dicionários eletrônicos, assim como tem sido implausível se conceber, na atualidade, o tratamento de qualquer base de dados linguísticos de forma meramente mecânica ou manual. Essa posição estabelece logo a princípio a necessidade de projeção de um dicionário que, para além de pressupor um produto a ser impresso, permita a confecção de um instrumento eletrônico que ofereça, aos possíveis interessados por uma obra dessa natureza, outras formas de consulta, a exemplo de pesquisas reversas, que possam, ademais, ser acessadas remotamente, por intermédio da Internet, uma das vias de socialização da informação mais importantes do mundo moderno.

⁸ Tradução livre: "os dicionários se beneficiaram, no início do século XVI, do surgimento da imprensa e de suas consequências econômicas, sociais e culturais. O advento da informática e de suas circunstâncias socioeconômicas e culturais exerce hoje uma ação idêntica sobre os trabalhos e produtos lexicográficos".

Esse condicionamento requer, inicialmente, que os dados que venham a alimentar o banco informatizado do ALiB sejam devidamente sistematizados na perspectiva do aproveitamento das respostas aos inquiridos, para posterior extração automatizada, sem a contaminação com outras informações linguísticas que não aquelas próprias a cada um dos informantes, em relação a cada um dos questionários antes referidos.

Um diagnóstico inicial dos formulários ora utilizados pelo Projeto ALiB para digitação dos resultados dos inquiridos sinalizou para uma alteração na estratégia de registro atual. Contrariamente à digitação sequencial em planilha de texto adotada, em que as intervenções dos inquiridores, assim como as identificações protocolares pertinentes, antecedem cada resposta do informante, poder-se-ia pensar em uma nova estratégia de digitação, em que se utilizasse o padrão XML⁹ como fundamento de todo o processo de registro e codificação. Isso traria imensa vantagem no sentido de aproveitamento da base dicionarística já existente em programas abertos, voltados para o tratamento informatizado de línguas naturais, a exemplo do UNITEX, ou mesmo do CORPINDEX,¹⁰ que objetiva a partir de modificação programática do UNITEX, dotá-lo de ferramentas mais sofisticadas de utilização, possibilitando, inclusive, a geração de relatórios de correspondência lexical com outras línguas que já disponham de trabalho de pesquisa semelhante ao que se propõe aqui.

Ademais, deve-se ter em vista que, os programas informáticos antes sugeridos para serem utilizados na manutenção da base de dados, em sua extração e em seu posterior tratamento, atentam para a não-desformatação das notações fonéticas, sendo capazes de migrar as fontes IPA, diacríticos e símbolos, para qualquer uma das plataformas finais, seja para editoração do dicionário em papel, seja para o desenvolvimento da plataforma de consulta "online", sem qualquer prejuízo formal.

A concepção macroestrutural do **Dicionário Dialectal Brasileiro (DDB)**, por seu turno, deve se alinhar à de trabalhos precedentes que viram na abordagem onomasiológica uma mais-valia em sua construção. Foi assim que Weijnen (cf. RYS; KEYMEULEN, 2009) procedeu na coordenação da confecção do *Dicionário dos Dialectos Brabanteses (Woordenboek van de Brabantse Dialecten – WBD)*¹¹, 1967–2005, do *Dicionário dos Dialectos Luxemburgueses (Woordenboek van de Limburgse Dialecten – WLD)*, 1983–2008¹² e do, ainda em curso, mesmo após seu falecimento, o *Dicionário dos Dialectos Flamengos (Woordenboek van de Vlaamse Dialecten – WVD)*, 1979–¹³.

Para Rys e Keymeulen (2009, p. 130-131):

The arrangement of the three dictionaries shows the interest that Weijnen took in dialect geography and etymology. The effect of a thematical arrangement is that geographically differentiated lexemes denoting one concept are pulled together, thus allowing for a word map and for etymological insights since (chronologically) related words appear side by

⁹ O padrão XML, sigla utilizada para *Extensible Markup Language*, é uma metalinguagem de largo uso hoje no mundo informático moderno e objetiva descrever com precisão e economia qualquer tipo de dado, com base em um parâmetro sintático que pode ser ampliado, consoante às necessidades do projeto a que procure servir.

¹⁰ O CORPINDEX se encontra em fase de desenvolvimento no LDI (*Lexique, Dictionnaire, Informatique* – CNRS UMR n° 7187), da *Université Paris XIII*, sendo uma criação de um de seus pesquisadores, o Senhor Fabrice Issac.

¹¹ <https://fuzzy.arts.kuleuven.be/rewo/wbd.htm>

¹² <https://fuzzy.arts.kuleuven.be/rewo/wld.htm>

¹³ <http://www.wvd.ugent.be/>

side. A simple glance at a list of Flemish dialect words like butoor, putoor, putoor, puitaard, puitreiger for the bittern (Dutch 'roerdomp', Botaurus stellaris) for instance, immediately brings to light the intermediate phases between the French loan butoor (5Fr. butor 'bittern') and the folk etymology puitreiger, lit. 'frog heron' (WVD III,1:219). In an alphabetically arranged dictionary, etymological relations are more difficult to uncover since the different words appear at their alphabetical positions.¹⁴

Considerando a própria organização temática a que se submete o Questionário Semântico-Lexical (QSL) do ALiB, parece que o DDB deva atender às mesmas áreas conceituais pensadas para esse questionário, abaixo reproduzidas, no caso de se detectar um volume de variação que o justifique, ou, já de posse dos resultados dos levantamentos, reduzi-lo a áreas conceituais mais amplas, em que alguns dos temas possam ser conjugados, sobretudo na elaboração dos fascículos impressos, haja vista que para a plataforma informatizada não pareça se justificar tal economia:

1. Acidentes geográficos
2. Fenômenos atmosféricos
3. Astros e tempo
4. Atividades agro-pastoris
5. Fauna
6. Corpo humano
7. Ciclos da vida
8. Convívio e comportamento social
9. Religião e crenças
10. Jogos e diversões infantis
11. Habitação
12. Alimentação e cozinha
13. Vestuário e acessórios
14. Vida urbana

Como, na construção de dicionários, "*aux problèmes de l'ordre alphabétique sont nécessairement liés ceux des variantes graphiques et des renvois*" (CATACH et al., 1971, p. 23)¹⁵, o **DDB** deve privilegiar um sistema de remissão bastante eficiente e complexo, capaz de dar conta de todas as co-referências possíveis, incluindo-se aí o sistema de identificação isoglósica da unidade lexical sobre o que adiante se referirá na composição microestrutural a ser sugerida. No caso do **DDB** impresso (já que a programação de um dicionário eletrônico já prevê o que se segue), antecedendo cada um dos fascículos temáticos deve ser gerada uma lista de palavras, em ordem alfabética,

¹⁴ Tradução livre: "A organização dos três dicionários mostra o interesse que Weijnen teve pela geografia dialetal e etimologia. A opção por uma estruturação temática faz com que lexias que denotem o mesmo conceito, mesmo que geograficamente distantes, possam ser arroladas em conjunto, permitindo inferências, quer na perspectiva etimológica (cronológica), quer sua identificação cartográfica. Um breve olhar na lista de palavras do dialeto flamengo como *butoor, putoor, putoor, puitaard, puitreiger*, para *bittern* (um tipo de pássaro) (holandês 'roerdomp', *Botaurus stellaris*), imediatamente esclarece as fases intermediárias entre o empréstimo francês *butoor* (5Fr. *butor* 'bittern') e a forma etimológica popular *puitreiger*, lit. 'frog heron' (WVD III,1:219). Em dicionários de estruturação alfabética, as relações etimológicas são menos evidentes, já que as diferentes palavras aparecem em sua estrita posição alfabética.

¹⁵ Tradução livre: "aos problemas de ordem alfabética são necessariamente relacionados os das variantes gráficas e das remissões".

grafematicamente apresentadas, com a devida remissão à entrada correspondente, para servir de guia aos consulentes. Note-se que possíveis realizações variacionais do tipo <abroba> ou <abobra>, relacionadas ao padrão *abóbora*, independentemente de não serem ortograficamente reconhecidas, devem ocupar sua posição alfabética nessa lista de palavras.

Convém aqui se fazer um alerta de que, enquanto em um dicionário geral da língua – "*le serviteur de la tradition*"¹⁶ (CATACH *et al.*, 1971, p. 167) por excelência – não pareça haver espaço para a inclusão de variantes que fujam ao *status quo* linguístico, pelas também óbvias razões de economia e insuficiência de dados, um dicionário dialetal, ao contrário, deve abarcar toda a instabilidade gráfica que os usos reais da fala possam em si fomentar, mesmo que esses itens não venham a constituir um cabeça de verbete na nomenclatura principal, senão lemas secundários na sua microestrutura, além de elementos integrantes do índice de palavras, como antes referido. Naturalmente, as unidades lexicais que não correspondam ao padrão ortográfico podem ser apresentadas em itálico ou por qualquer outro indicador estrutural que melhor sinalize essa condição, mesmo porque na perspectiva do programador sempre talvez haja uma melhor opção. O sistema de transcrição grafemática deve, todavia, ser o mesmo que tem sido adotado pela equipe do ALiB até hoje.

Conquanto a dimensão estrutural do **DDB** só possa ser considerada a partir da avaliação efetiva da base de dados, todas as respostas dos informantes aos três questionários, isto é, ao QFF, ao QSL e ao QMS, em todos os 250 pontos que são abrangidos pelo ALiB, devem ser consideradas em sua composição, ganhando-se com isso volume de informação, o que se refletirá naturalmente na riqueza vocabular do dicionário que se objetiva construir. Caberá ao programa informático a ser utilizado realizar as devidas correspondências temáticas e variacionais, a partir da etiquetagem de todo o material, inclusive de possíveis lexias e/ou colocações que possam ser identificadas, assim como a sua localização geográfica e realização fonética de maneira dinâmica. Outra questão macroestrutural a definir será a da adoção do sistema de abreviaturas, que deverá ser unívoco e invariante, em prol do perfeito processamento dos dados.

Quando se sugere o aproveitamento dos dados de todos os questionários, pressupõe-se que, como o foco é o léxico", muitos dados que não estejam presentes no QSL podem ser identificados nas respostas aos outros inquéritos dos outros questionários, nomeadamente os signos lexicais candidatos a lema secundário. Como a frequência não pode ser considerada como critério de lematização no caso do **DDB**, senão para dicionários gerais de língua, a quantificação não conta para a composição dos verbetes, isto é, não importa se um item apareça em mais de um questionário, já que a codificação XML dará conta de identificar qualquer redundância de um dado tem por ponto ou por informante. Ademais, quanto ao QFF (Questionário Fonético Fonológico) em especial, este servirá para assegurar o caráter variacional pretendido, garantindo possíveis variantes fonéticas de um item que, porventura, não exiba variação lexical no território nacional.

Enquanto proposta inicial, poder-se-ia pensar em um microestrutura, isto é,

¹⁶ Tradução livre: "o servidor da tradição".

l'ensemble des informations ordonnées qui suivent l'entrée; cet ensemble a une structure constante qui répond à un programme et à un code d'information applicable à n'importe quelle entrée"¹⁷ (REY-DEBOVE, 1971, p. 151),

que obedecesse à seguinte composição, em que os itens e indicadores estruturais fossem contemplados na ordem em que se apresentam. Estes últimos, ou seja, os indicadores, são sugeridos mormente para a versão impressa, ficando a critério do programador a alteração sistemática desses elementos na versão eletrônica do dicionário:

- **Lema principal.** Forma ortográfica canônica do item lexical ou da lexia composta ou complexa, em minúscula e negrito, com recuo à esquerda do restante do corpo do verbete, na edição impressa, e em qualquer outro destaque na versão eletrônica.
- **Lemas secundários.** Toda variação fônica identificada em relação ao lema principal, em transcrição fonética, sem negrito, antecedida por vírgula, com o número respectivo do ponto em que foi identificada a variante em questão, entre parênteses. A lista de pontos, assim como o mapa geral, integrará, obviamente, o *front matter* do dicionário impresso. No dicionário eletrônico, estabelecimento de *hiperlink* de cada realização fônica, para audição e para a identificação cartográfica da isoglossa, por botões de acesso.
- **Lemas múltiplos.** Em caso de particularidades de flexão ou derivação que porventura provoquem alomorfia lexemática, devem ser registrados na cabeça do verbete. A forma de apresentação no dicionário eletrônico pode ser indicada por botões de acesso, sinalizando femininos, plurais ou superlativos, por exemplo.
- **Classificação gramatical.** No sentido de permitir o "diálogo" futuro do DDB eletrônico com dicionários dialetais de outras línguas naturais, prevê-se a mesma classificação adotada pelos dicionários-de-base dos programas que serão utilizados na sua concepção.
- **Etimologia.** Considerando que *"la dialectologie a contribué à faire avancer très efficacement une science essentiellement historique, celle de l'étymologie"*,¹⁸ como afirma Picoche (1973, p. 12), e que a informação etimológica se constitui para a lexicografia um dado de extremo interesse, sobretudo em casos de homonímia, deve-se prever seu levantamento em todas as unidades lexicais passíveis de apuramento, apondo-se, inclusive, as discrepâncias detectadas entre os diversos etimólogos consultados. Em caso de impossibilidade de precisão, fazer constar o termo "ignorada", ou similar.

¹⁷ Tradução livre: "o conjunto de informações ordenadas, seguintes à entrada; correlacionada a uma estrutura constante que se refira a um programa e a um código de informação a serem aplicados a todas as entradas, indistintamente".

¹⁸ Tradução livre: "a dialectologia contribuiu bastante para o desenvolvimento eficaz de uma ciência essencialmente histórica, a etimologia".

- **Definição.** Como se sabe, "*one of the main reason why people consult dictionaries is that they want to get information about meaning*"¹⁹ (Moerdijk, 2003, p. 273). Não obstante ser essa máxima muito mais relacionada a dicionários gerais do que propriamente a dicionários dialetais, já que para estes o centro do interesse é a própria identificação da variação espacial, prevê-se a elaboração de definições lexicográficas para todos os itens selecionados, isto é, definições que comportem uma paráfrase sêmica, por assim dizer, de compromisso, em que se explicitem o *genus proximum* e as diferenças específicas da lexia em foco, evitando-se com isso a perpetuação de uma prática condenável, por vezes muito utilizada pela lexicografia contemporânea, que é a da definição sinonímica.
- **Abonação.** Embora nem sempre seja possível apresentar um contexto frasal real que possa abonar cada um dos lemas selecionados, em razão da própria característica dos questionários (que por vezes prevêem como resposta apenas uma lexia que represente um dado conceito que se quer perscrutar), prevê-se no DDB, quando possível, o registro de todas as concordâncias relacionadas ao item, existentes no *corpus*. Em caso de sua ausência, por completa indisponibilidade de registro na resposta ao questionário, repetir o signo lematizado como aparece na base de dados.
- **Remissão.** O sistema de remissão deve dar conta quer dos casos de variação fonética, quer dos casos de variação lexical, assim como de possíveis relações de antonímia e homonímia, aí incluídos os homógrafos e homófonos, obviamente. Estas últimas relações podem ser consultadas no dicionário eletrônico através de botões de acesso. No caso do dicionário impresso, a indicação pode fazer parte da microestrutura final do verbete. Deve-se atentar ainda para o estabelecimento de uma rede remissiva que permita a visualização cartográfica integrada de uma determinada ocorrência lexical, em todos os pontos geográficos em que esta ocorra, gerando-se, aí, as isoglossas correspondentes. No dicionário impresso, essa relação deverá ser feita pela indicação do ponto que é apresentado logo após a variante, entre parênteses. O consulente que deseje conhecer maiores detalhes deverá consultar, naturalmente, o *Atlas Linguístico do Brasil*, em seu formato cartográfico, que, provavelmente, deverá estar disponível por acesso remoto.
- **Marcas de uso.** Após a análise dos dados gerais do ALiB, proceder ao inventário das marcas identificadas, com vistas a seu registro, quer no dicionário eletrônico, quer no impresso.

CONCLUSÃO

Pretendeu-se, de forma geral, apresentar as bases preliminares de elaboração do **Dicionário Dialectal Brasileiro (DDB)**, obra de verve coletiva e interinstitucional, que

¹⁹ Tradução livre: "uma das principais razões por que as pessoas consultam dicionários é em função de desejarem informação sobre o significado".

envolverá diversos especialistas, quer na área da dialectologia, quer nas áreas da lexicografia e das ciências da informação, do Brasil e da França.

Sua concepção não está voltada ao tratamento isolado de dialetos brasileiros, mas visa permitir uma visão pandialetal da realidade variacional do léxico no Brasil, com base no dados do Projeto *Atlas Linguístico do Brasil* (Projeto ALiB).

Considerando que, "sem o conhecimento da história, da cultura e do momento presente, é impossível se pensar na elaboração de um dicionário", como bem assinalou, oralmente, Jean Pruvost em uma recente Jornada de Dicionários, na Universidade de Cergy-Pontoise, na França, e que

ce n'est pas à travers un terme seul que l'on pourra donner toute la dimension culturelle dont il est question, mais en utilisant les différents reseaux de nomenclature traversant l'ensemble de l'ouvrage²⁰ (FERRARA, 2009, p.191-192),

o **DDB** objetiva conjugar de forma abrangente língua e cultura, no sentido de permitir o conhecimento mais abrangente possível da relação dialetal que se estabelece, através do léxico, nas comunidades que o utilizam.

REFERÊNCIAS BIBLIOGRÁFICAS

- BARBATO, Marcello; VARVARO, Alberto. Dialect dictionaries, **International Journal of Lexicography**, v. 17, nº 4, p. 429-439, dec 2004
- CARDOSO, Suzana Alice. **O papel das pesquisas sociolinguísticas e dialectológicas para os estudos do português brasileiro**. Salvador, 23 p. Trabalho não publicado.
- CATACH, Nina; GOLFAND, Jeanne; DEVUX, Roger. **Orthographe et lexicographie**. Tome I. Paris: Didier, 1971. 333 p.
- COMITÊ NACIONAL DO PROJETO ALiB. **Atlas Linguístico do Brasil**: questionários. Londrina, Editora UEL, 2001.
- FERRARA, Alice. Le 'dictionnaire d'un pays, le pays d'un dictionnaire' un type particulier de dictionnaire lexicoculturel, **ÉLA**, nº 154, p. 191-203, avril-juin 2009.
- MOERDIJK, Fons The codification of semantic information. In: STERKENBURG, Piet van (ed.). **A practical guide to lexicography**. Amsterdam: John Benjamins, 2003. p. 273-293
- PICOCHÉ, Jacqueline. Les monographies dialectales (domaine gallo-roman), **Langue Française**, v. 18, nº 1, p. 8-41, 1973.
- PIEL, Joseph Maria. Origens e estruturação histórica do léxico português. In: CASTRO, Ivo et al. **Curso de história da língua portuguesa**: leituras complementares. Lisboa: Universidade Aberta, 1991[1976]. p. 233-242.
- PRUVOST, Jean. Quelques perspective lexicographiques à mesurer à l'aune lexicoculturelle, **ÉLA**, nº 154, p. 137-153, avril-juin 2009.
- QUEMADA, Bernard. Bases de données informatisées et dictionnaires, **Lexique**, nº. 2 p. 101-120, 1983.
- VILELA, Mário. **Estruturas léxicas do português**. Coimbra: Almedina, 1979
- REY-DEBOVE, Jacqueline. **Etude linguistique et sémiotique des dictionnaires français contemporains**. La Haye/Paris: Mouton, 1971.

²⁰ Tradução livre: "não é através de um termo isolado que se poderá dar toda a dimensão cultural da qual faz parte, mas utilizando as diferentes redes de nomenclatura que perpassa o conjunto da obra.

- RONCO, Giovanni. Au delà des dictionnaires: les atlas linguistiques. **International Journal of Lexicography**, v. 17, n. 4, p. 441-455, dec. 2004.
- ROSSI, Nelson et al. **Atlas prévio dos falares baianos**. Rio de Janeiro: INL, 1963.
- RYS, Kathy; KEYMEULEN, Jacques van. Intersystemic correspondence rules and headwords in Dutch Dialect Lexicography. **International Journal of Lexicography**, vol. 22, n° 2, p. 129-150, april 2009.
- WEINREICH, Uriel, LABOV, William & HERZOG, Marvin. Empirical foundations for a theory of language change. In: LEHMANN, W. & MALKIEL, Yakov (Eds.) **Directions for Historical Linguistics: A Symposium**. Austin: University of Texas Press, 1968. p. 95-195.