

AVANÇOS DA GEOLINGÜÍSTICA NO BRASIL

Hilda Gomes Vieira
Universidade Federal de Santa Catarina

1. Introdução

Minha proposta, nesta oportunidade, é apresentar os resultados da pesquisa na área de Dialetoologia, voltada para o desenvolvimento da Geolingüística no Brasil nas últimas décadas. Inicialmente, faço uma retrospectiva sobre a evolução dos recursos disponíveis aos pesquisadores desde o primeiro trabalho de campo de Gilliéron (1896), tão bem delineado na obra de Chambers & Trudgil (1998). A seguir, enfatizo os recursos disponíveis na elaboração dos primeiros atlas lingüísticos publicados no Brasil (cf. Vieira, 1998) e demonstro como os avanços metodológicos da Geolingüística têm registrado um período de crescimento na última década, em função do advento da Informática.

Objetivo proceder a uma análise da situação da Geolingüística no Brasil, desde o primeiro atlas lingüístico publicado na Bahia, intitulado *Atlas Prévio dos Falares Baianos* de autoria de Nelson Rossi (1963). A partir de um balanço dos cinco primeiros atlas publicados no Brasil, e de alguns outros em andamento, examino as dificuldades que costumam suscitar atrasos nas publicações, em razão dos recursos tecnológicos disponíveis em cada época, e ofereço algumas sugestões para a constituição de um atlas lingüístico em muito menos tempo, incentivando a elaboração de Bancos de Dados Geolingüísticos em todas as universidades do Brasil, a partir de modernos recursos da Informática como pré-requisito para o desenvolvimento de um bom atlas lingüístico.

2. Retrospectiva

A história da Geografia Lingüística no mundo, conhecida como Geografia Dialetoal, ou Geolingüística, apresenta um período de crescimento e expansão na primeira metade do século 20, seguido de um declínio nas décadas seguintes, e de uma recente revitalização nos últimos anos. Um histórico, apresentado por Chambers & Trudgil (op.cit), mostra que o primeiro trabalho de campo começou na França em 1896, sob a direção de Jules Gilliéron. O questionário revisado com 1500 perguntas foi aplicado pelo não lingüista Edmond Edmont, que havia sido treinado para trabalhos de transcrição fonética. A publicação final do último volume se deu em 1910. Outros grandes Atlas foram publicados e encontram-se como jóias raras nas bibliotecas dos países respectivos. De um modo geral, levam décadas para serem publicados, dadas as características próprias de trabalho de tamanha envergadura.

Os recursos disponíveis em cada época foram, aos poucos, evoluindo e o método da Geografia Dialetoal, a que passo a denominar de Geolingüística, foi-se consolidando paulatinamente, incluindo os seguintes procedimentos:

- 1) questionário formal, depois informal, questões de completar;
- 2) aplicação de respostas indiretas (escritas), perguntas diretas (respostas orais), transcritas *in loco*, depois gravadas, e questões de completar;
- 3) cartas lingüísticas descritivas e analíticas — interpretativas — com símbolos fonéticos manuscritos e personalizados, hoje, com fontes em programas de computador para uso do Alfabeto Fonético Internacional;

- 4) seleção de informantes com critérios específicos, considerando-se sexo, idade, escolaridade, nível social, etc;
- 5) conclusão, com publicação, dos sempre demorados atlas lingüísticos, em que grande parte das cartas é elaborada manualmente, transformando-se, muitos deles, em verdadeiras obras de arte, pouco consultadas nas bibliotecas.

3. Os atlas lingüísticos no Brasil

Em Vieira (1998), apresentei um balanço crítico dos Atlas Lingüísticos publicados no Brasil, visando examinar os problemas que costumam atrasar as publicações em face dos recursos tecnológicos disponíveis em cada época e, assim também, apresentar sugestões para elaboração e publicação dos volumes com as Cartas Lingüísticas em tempo mais curto, através do emprego de recentes recursos da Informática. Foram abordados, entre outros, os seguintes tópicos: a) análise dos atlas lingüísticos em referência ao número de perguntas, cartas publicadas, número de pontos, número de informantes por ponto, tipo de alfabeto fonético usado, datas do projeto inicial e da publicação dos atlas; b) sugestões para informatização dos dados através de um banco de dados geolingüísticos; no elenco das sugestões, procurei demonstrar que, apesar das diferenças entre sistemas para informatizar os de dados de um projeto destinado à elaboração de um atlas lingüístico, seja ele nacional, regional, estadual, ou local, muita coisa há de comum; logo, torna-se possível aglutinar os resultados de todos os projetos estaduais em uma única carta nacional, para, assim, chegar-se à tão sonhada definição de áreas dialetais no Brasil.

Passo, então, a apresentar alguns tópicos tratados naquela oportunidade que vem ao encontro da nossa proposta de trabalho atual:

Os estudos de Geolingüística no Brasil ganham terreno ao final do presente século. Já se pode contar com cinco atlas lingüísticos publicados, em todo, ou em parte. Esses Atlas, relacionados aqui em ordem crescente pelo ano de publicação, são os seguintes: 1) *Atlas Prévio dos Falares Baianos* (Rossi, 1963); 2) *Esboço de um Atlas Lingüístico de Minas Gerais* (Ribeiro e outros, 1977); 3) *Atlas Lingüístico da Paraíba* (Aragão & Menezes, 1984); 4) *Atlas Lingüístico de Sergipe* (Ferreira e outros, 1987); 5) *Atlas Lingüístico do Paraná* (Aguilera, 1994).

O Atlas Lingüístico do Brasil, ALIB, que tem sido preocupação de pesquisadores desde 1950, está em vias de se concretizar por uma equipe de pesquisadores coordenada por Suzana Alice Marcelino Cardoso, da Universidade Federal da Bahia, sendo que a elaboração de seu projeto está em fase bem adiantada (Cardoso, 1996). Tem-se notícias de que mais cinco atlas estão em andamento há alguns anos, lutando contra as dificuldades características de trabalho de tal envergadura.

3.1 Análise do volume de cartas publicadas

Um exame cuidadoso dos atlas lingüísticos já editados revela que, na maioria deles, o número de cartas é bem pequeno em comparação com o imenso *corpus*, resultante de longos inquéritos aplicados aos informantes em todas as localidades.

Uma análise na Tabela I, organizada a partir da leitura dos Atlas, evidencia a ocorrência e percentagem da freqüência das questões aplicadas aos informantes cujas respostas não foram aproveitadas para as Cartas nos Volumes dos respectivos Atlas:

Tabela 1 - Ocorrência e percentagem de freqüência de questões sem carta dos cinco Atlas Lingüísticos, publicados no Brasil até 1998

Atlas	Questões	Cartas	Quest. s/Carta	% s/carta
Bahia	164	154	10	6,10
MinasGerais	415	73	342	82,41
Paraíba	877	149	728	83,01
Sergipe	700	182	518	74,00
Paraná	325	191	134	41,23
Total	2.481	749	1.732	69,81

Esses dados nos mostram que um total de 69,81% das questões não foram aproveitadas para as cartas, enquanto o percentual de respostas sem aproveitamento em cartas chega a alcançar 83,01% em um só Estado. Esse material, se não estiver sendo aproveitado para outros volumes, poderá estar se perdendo, ou talvez, estar sendo considerado sem planejamento adequado para ser incluído (e/ou disponibilizado) em outras pesquisas.

3.2 Análise do volume de respostas gravadas

Considerando-se que cada questão é respondida por muitos informantes, o volume de respostas no *corpus* gravado que fica sem aproveitamento atinge proporções acentuadas, como atesta a Tabela 2:

Tabela 2 - Freqüência de respostas sem carta dos cinco Atlas Lingüísticos, publicados no Brasil até 1998

	Quest. s/ Carta	Informant.	Resp. s/ Carta
Bahia	10	50	500
Minas Gerais	342	50	17.100
Paraíba	728	107	77.896
Sergipe	518	30	15.540
Paraná	134	130	17.420
Total			128.456

Tendo em vista que os informantes são muitos em cada Estado, as respostas sem carta chegam a atingir um total de 128.456 nos cinco Atlas. A tabela 2 permite que o leitor observe o detalhamento dessa perda por Estado.

Pode-se imaginar, então, quanto material lingüístico existe armazenado nas universidades, sem tratamento adequado para aproveitamento em outras cartas. Isto porque as fitas, em que são tradicionalmente gravadas as entrevistas, têm curta duração, sendo que, nem sempre, os dados foram transcritos. E mesmo, se fosse esse o caso, na maioria das vezes, as transcrições manuais, geralmente feitas a lápis, não constituem fonte

adequada para consultas freqüentes, dada a pouca durabilidade do material, em razão de seu manuseio, e a dificuldade de controle das consultas.

Após a leitura dos Volumes de Cartas e à procura de informações que me fornecessem notícias sobre o destino desse material, encontrei algumas respostas em publicações posteriores.

O Atlas da Bahia publicou, em 1965, um Volume Introdutório, que apresenta o elenco das respostas transcritas impressas, mas apenas das que foram motivo de cartas, fato que considero de grande valor para proteção do *corpus*.

Como na Bahia o número de perguntas sem carta é pequeno, estaria sem tratamento pouco material. Porém, a leitura do Volume Introdutório, editado 3 anos depois da publicação do Volume de Cartas, evidenciou que todo o material não aproveitado para cartas fora registrado no Questionário Comentado, ou está “devidamente arquivado e catalogado” no Laboratório de Fonética da Universidade Federal da Bahia “com livre acesso à documentação” (Rossi, 1965: 49). Como não tive o prazer de conhecer tal arquivo, não posso opinar sobre o tipo de tratamento dado a essa documentação.

No próprio volume de cartas de Minas Gerais, é apresentado um plano da obra em 4 volumes a serem publicados (Ribeiro e outros, 1977:17). Porém, até o presente, não o segundo volume não saiu. Em trabalho discutido em Seminário, Zagari (1996:16) declara que, em dez anos, o resultado das entrevistas totaliza cinco mil horas de gravação; contudo, ele não se refere ao tipo de tratamento dedicado às respectivas gravações.

O Atlas da Paraíba publicou o Volume 2 no mesmo ano do Volume 1, contendo uma análise das formas e estruturas lingüísticas observadas nas entrevistas. Em 1996, doze anos após a publicação dos volumes, uma das autoras (Aragão, 1996:29) afirma que a riqueza do material coletado poderá alimentar trabalhos de pesquisadores e professores por muito tempo. Entretanto, não encontrei registros do tipo de tratamento conferido ao material lingüístico não utilizado para cartas.

Por outro lado, o Atlas do Paraná, segundo Aguilera (1996:45), em publicação dois anos após o volume das cartas, informa que a transcrição completa de cada entrevista se encontra datilografada, compondo um acervo de 4.000 laudas em 13 volumes. A autora manifesta intenção de preservar o acervo para consulta de outros pesquisadores, transferindo para CD as 300 horas de gravação dos dados.

O que se pode observar de concreto do exame de todo o material, mesmo dos trabalhos publicados após o volume das cartas, é que os projetos para a elaboração de um atlas lingüístico se arrastam por décadas, com perspectivas de retorno somente a longo prazo, até que seja publicado o primeiro volume de cartas.

3.3 Análise do tempo de publicação dos primeiros volumes

Não surpreende a ninguém verificar que um volume de cartas lingüísticas leva de cinco a vinte e quatro anos para ser publicado (uma média de 11,4 anos), como demonstra a Tabela 3, uma vez que o trabalho enfrenta incontáveis dificuldades que não cabe aqui enumerá-las.

Tabela 3 - Tempo médio em anos, despendido para publicação do primeiro volume de cartas dos cinco Atlas Lingüísticos do Brasil até 1998

Atlas	Início	Publicação	Tempo em anos
Bahia	1958	1963	5
Minas Gerais	1969	1977	8
Paraíba	1974	1984	10
Sergipe	1963	1987	24
Paraná	1984	1994	10
Média			11,4

Ao leitor cabe analisar os demais dados dessa tabela que apresenta, por Estado e por total, o tempo despendido, desde a elaboração do Projeto, ou dos trabalhos iniciais, até a publicação do Primeiro Volume de Cartas.

Entretanto, conhecendo-se, atualmente, recursos capazes de diminuir esses prazos, é indispensável que comecemos a nos preocupar com eles.

A preocupação aumenta quando se sabe que os atlas lingüísticos em andamento contam com alguns anos em atividade, desde a concepção dos seus respectivos projetos. São eles: Atlas Lingüístico de São Paulo, Atlas dos Pescadores da cidade do Rio de Janeiro, Atlas Lingüístico de Mato Grosso, Atlas Lingüístico do Ceará e Atlas Lingüístico da Região Sul do Brasil (ALERS) envolvendo os estados de Paraná, Santa Catarina e Rio Grande do Sul.

Dos três primeiros, não tenho conhecimento de seu estágio atual. O ALERS, iniciado em 1989, está com seus três primeiros volumes praticamente prontos, mas, ainda não se publicou nem um volume de cartas, passados quase dez anos de trabalho árduo de uma equipe de pesquisadores das três universidades Federais do Sul do País.

O Atlas do Ceará, já publicou o volume do questionário no ano de 1982. Dezesesseis anos se passaram sem que se publicasse seu primeiro volume de cartas.

Tem-se notícias de que o questionário do Atlas Lingüístico de São Paulo está pronto desde 1961 e serviu de base para o questionário do Atlas do Paraná, mas ainda não se editou seu primeiro volume de cartas.

Essa situação tem sido compreensível, uma vez que há bem pouco tempo não se dispunham dos recursos modernos da informática, que vêm oferecendo tecnologia cada vez mais capaz de modificar toda uma situação que aí está.

Além disso, os trabalhos de Geolingüística lidam com um enorme volume de dados que precisam ser transcritos em alfabeto fonético por pessoas altamente especializadas, e são passíveis de intermináveis revisões por parte do pesquisador.

Sem material informatizado, não sobra tempo para publicações paralelas ao trabalho de elaboração do Atlas, essenciais à vida do pesquisador e do próprio atlas. Dessa forma, a publicação do atlas lingüístico, na maior parte das vezes, ou não se realiza, ou “só acontece”, dentro de uma ou duas décadas.

3.4 Os tipos de cartas lingüísticas

As cartas lexicais ocupam a grande maioria dos atlas, apresentando-se sob várias formas, ora descritivas com legendas e símbolos e/ou transcrição fonética na própria carta, ora em forma de isoléxicas.

As cartas lexicais, tipo descritivas, contêm símbolos para representar as variantes lexicais nos diversos pontos de inquérito. A legenda de cada carta registra cada variante em sua forma ortográfica. Esse tipo de carta, apresentada com uma **freqüência** total de 321 pontos nos cinco atlas lingüísticos, tem sido a mais corrente, talvez por não precisar de símbolos fonéticos, que têm dificultado, não só a impressão, mas também a sistematização dos dados lingüísticos, mesmo com o advento de modernos recursos da informática. Os dados fonéticos precisam de tratamento específico, que exige preocupação do pesquisador desde a elaboração do projeto inicial. Tais requisitos aplicam-se a uma série de procedimentos como: tipo de questionário específico com orientações aos inquiridores; sistemas especiais de computador para emissão de relatórios especiais; seleção das fontes de símbolos fonéticos e, ainda, seleção de símbolos para codificação automática dos dados de forma inteligente e econômica, visando à emissão de todas as respostas das cartas descritivas básicas. Desse modo, torna-se possível atender ao objetivo básico de um bom atlas lingüístico.

Continuando a análise dos atlas, outro tipo de carta, ora denominada de carta léxica, ora fonética, apresenta a transcrição fonética de cada variante dentro da carta em cada ponto do inquérito respectivo. Trabalho de artista com desenhistas especializados, como é o caso da Bahia que faz uso de normógrafos e de delicadas artimanhas para combinar pedaços de símbolos com partes dos desenhos dos normógrafos, recortes, colagens, etc.

Cada carta desse tipo, refere-se, muitas vezes, a uma das variantes lexicais de cartas anteriores, e não descreve um único fato fonético, pois o tipo de questionário lexical dá margem a uma variedade muito grande de itens lexicais diferentes na sua forma fonética, impedindo que se trate, numa mesma carta, de um só fato.

Uma versão um pouco diferente da anterior, encontra-se nos atlas em que a transcrição fonética de cada variante é codificada com símbolos, mas, do mesmo modo, cada carta não trata de apenas um só fato fonético, pelo mesmo motivo supracitado, dificultando a leitura das mesmas.

O último tipo de cartas (cf. Tabela 4), o da fonética propriamente dita, que descreve apenas um fato em toda a carta, aparece com uma freqüência 17, diante de um volume de 441 cartas que usam dados fonéticos e de 749 cartas publicadas nos cinco Atlas. Dessas 17 cartas, apenas 4 são registradas com símbolos e transcrição fonética na legenda, enquanto as demais, em número de 13, apresentam-se em forma de cartas de isófonas.

Tabela 4 - Ocorrência e percentagem da freqüência de cartas fonéticas em cinco Atlas Lingüísticos, publicados no Brasil até 1998

Atlas	Cartas	c/ fonética	transcrição %	c/1 só fonético	fato
Bahia	154	135	87,66	0	
Minas Gerais	73	28	38,36	3	
Paraíba	149	71	47,65	4	
Sergipe	182	134	73,63	0	
Paraná	191	43	22,51	10	
Total	749	411	54,87	17	

Considerando os resultados dessa análise nos atlas do Brasil — aliados à minha experiência desde 1989, em Santa Catarina, em desenvolver um Projeto Integrado objetivando a publicação do Atlas Lingüístico-etnográfico da Região Sul do Brasil, (ALERS/RS) — e levando em conta as inevitáveis necessidades que estarão sendo geradas pelos trabalhos recentes dedicados à implementação do Projeto Atlas Lingüístico do Brasil, ALIB, não poderia me furtar a oferecer alguma contribuição com base em nossas descobertas em Santa Catarina.

Defendo a tese de que, em qualquer projeto de elaboração de um atlas lingüístico, seja ele estadual, regional, ou nacional, é necessário incluir, como primeira etapa, uma série de providências para se construir um Banco de Dados Informatizados.

Modernamente já se vem observando iniciativas para o desenvolvimento de sistemas informatizados, como o que está sendo implementado para o ALERS Regional do Sul do Brasil. Esse programa, elaborado em 1995 aproximadamente, permite reunir os dados lingüísticos resultantes das entrevistas realizadas nos três estados, Paraná, Santa Catarina e Rio Grande do Sul.

Mas é necessário levar em conta as diferentes necessidades em cada tipo de atlas lingüístico, seja estadual, regional, ou nacional.

Os atlas estaduais que permitem maior densidade em seus pontos de inquérito têm como finalidade principal apresentar todas as cartas descritivas, sem exceção, isto é, sem abandonar os itens que não dão boas “cartas”, pelo simples fato de eles não apresentarem variantes, como é costume se pensar em trabalhos de Geolingüística.

Os atlas regionais, ou o nacional, poderão se beneficiar de toda essa informação, não precisando incluir as cartas descritivas, pelo menos com todos os pontos do inquérito. Um atlas regional ou nacional pode conferir maior ênfase a cartas bastante sofisticadas para elaboração de áreas dialetais mais significativas, com definição de isófonas, isoléxicas e isomórficas, sempre a partir das cartas descritivas estaduais.

Um fato fonético em um Estado, que não dá “boa carta” na concepção do lingüista, por não apresentar variantes, poderá contrastar com outra variante do mesmo fato em outro Estado, que poderia também ter sido abandonado por não dar “boa carta”.

O registro desses dados em uma carta descritiva será de grande importância para a confecção dos atlas regionais ou nacionais. Deve sobrar tempo para os atlas nacionais se dedicarem a análises ainda mais sofisticadas - como a aplicação do sofisticado método da Dialetrometria- visando obter uma análise acurada para a determinação de áreas dialetais, com localização de semelhanças e diferenças mais ou menos acentuadas.

4. Sugestão para informatização de Atlas Lingüístico

Como disse anteriormente, a experiência como pesquisadora do CNPq na Universidade Federal de Santa Catarina, integrando a equipe de pesquisadores, coordenada por Walter Koch, da Universidade Federal do Estado do Rio Grande do Sul, para elaborar o Atlas Lingüístico-Etnográfico da Região Sul do Brasil, e a disseminação, nos centros de pesquisas, dos recursos da informática, no decorrer desses quase dez anos, permitiram-nos que fôssemos, paulatinamente, descobrindo novos caminhos capazes de agilizar a publicação dos Atlas Lingüísticos estaduais e, conseqüentemente, de criar a oportunidade de implementar atlas regionais com mais recursos, contribuindo, assim, para o tão sonhado Atlas Nacional.

Nesta secção trato aqui das cartas fonéticas, por dois motivos principais, intimamente relacionados. Primeiro, este é o tipo de carta que não está muito presente nos 5 Atlas Lingüísticos publicados no Brasil, por razões anteriormente citadas. Segundo, os

recursos modernos estão paulatinamente oferecendo tecnologia efetivamente válidas para elaboração de cartas descritivas fonéticas, com muita perfeição e em tempo ideal, para a consecução do objetivo principal de um atlas estadual, que é o de oferecer, em tempo hábil, cartas descritivas para os atlas regionais e/ou o nacional, selecionando fatos relevantes a esse nível, nacional ou regional.

As preocupações que devem cercar a confecção de boas cartas fonéticas começam a partir da proposição de um bom projeto, em que se definam aspectos intimamente relacionados, tais como: a) a organização e a manutenção de um Banco de Dados Geolingüísticos em cada Estado, em que fiquem registrados não só o material para aproveitamento das cartas, mas toda e qualquer resposta do informante que permita identificar o contexto fonético de cada elemento fônico estudado; b) o estabelecimento de uma rede de pontos, o mais densa possível para os atlas estaduais; deve-se levar em conta o fato de que há estados pequenos, mas com variedade de línguas, trazidas por imigrantes nas várias épocas de colonização, como é o caso de Santa Catarina, com nada menos de 6 idiomas – o italiano, o alemão, o eslavo, o japonês, o polonês, o árabe, somados ao português luso e o açoriano; c) a confecção de questionário fonético específico (QFF), cujas respostas dêem margem à emissão de um mesmo vocábulo por todos os informantes, permitindo analisar cada fato num mesmo contexto; d) a seleção de fatos fonéticos variáveis, não só da língua da comunidade a ser estudada, mas também daqueles relevantes aos demais estados da nação; e) a forma de publicação das cartas em que os dados sejam apresentados os dados o mais objetivamente possível, um só fato fonético por carta; f) a publicação dos quadros de variantes, de modo a fornecer informações adicionais sem prejudicar a objetividade das cartas.

Pensando nesses pontos básicos, está sendo testado em Santa Catarina, um *software* para informatizar os dados em um Banco de dados Geolingüísticos, visando atender às necessidades acima destacadas.

5. Conclusão parcial

Acredito não ter ainda esgotado a aprendizagem dos ensinamentos contidos nos atlas lingüísticos a que me propus examinar. Entretanto, posso afirmar, com convicção, que os pesquisadores que nos antecederam fizeram verdadeiros milagres diante das dificuldades de cada época, enfrentadas e superadas, para nos legar verdadeiras obras de arte no registro da Língua Portuguesa falada no Brasil. Acredito que nosso maior compromisso em contribuir com avanços mais acelerados se deve ao fato de que dispomos atualmente de maiores fontes de recursos tecnológicos e, conseqüentemente, encontramos menos dificuldades, além de contarmos com esse verdadeiro legado dos pesquisadores que nos antecederam e em quem nos apoiamos para dar continuidade à nobre tarefa que temos pela frente.

As cartas dos atlas lingüísticos estaduais devem primar por fazer, em primeiro lugar, a documentação da língua, através de cartas descritivas, deixando para um segundo momento as cartas de análise mais elaboradas. Os atlas fornecerão, dessa forma, em primeiro lugar, material descritivo adequado para o atlas, nacional, ou regional, poder selecionar os fatos mais relevantes e suficientes para montar, com maior segurança e em prazo curto, o quebra-cabeças das áreas fonéticas, cuja delimitação é tão almejada a nível nacional. Todas essas exigências demonstram o quanto precisamos contar com os modernos recursos da informática, desde os momentos iniciais da elaboração do projeto de constituição de um atlas.

Como conclusão, apresento, a seguir, uma proposta de trabalho em cursos de Letras, a ser aperfeiçoada pelo grupos de pesquisa interessados.

5.1 Diversidade lingüística e ensino da língua materna: uma proposta para os cursos de Letras

Considerando que as línguas do mundo são essencialmente dinâmicas, submetidas que são ao longo da história a processos de variação e mudança, justifica-se uma preocupação para com o estudo das variantes lingüísticas como medida de se superar as dificuldades de ensino advindas dessa diversidade.

Nas reflexões sobre a necessidade imperativa de se conhecer melhor as variantes da língua que se deseja ensinar, somadas à minha experiência como pesquisadora no Projeto ALERS, busquei subsídios para uma proposta de ensino nos cursos de Letras.

O objetivo central da proposta em questão é destacar a necessidade de se criar um banco de dados informatizados com as variantes da língua falada em cada localidade, incorporando a participação dos Cursos de Letras, incentivando alunos e professores a aceitarem esse desafio.

Pode-se iniciar o trabalho a partir de pequenos projetos de descrição da língua na disciplina de Lingüística (na graduação e/ou pós-graduação). Os alunos serão orientados para, paralelamente ao ensino teórico, se debruçarem sobre a língua de sua comunidade, procedendo à descrição e à análise das variantes mais representativas.

A preocupação primeira será com um Banco de Dados da Língua Falada. Pequenos projetos poderão ser destinados à constituição de um banco que irá fazendo parte do acervo da biblioteca da escola.

O trabalho pode ser organizado de forma que, cada turma de alunos, sob a tutela do professor da disciplina de Lingüística, se ocupe com informantes diversos, a fim de cobrir diferenciada gama de fatores, como idade, escolaridade, sexo, registros, etc, em todos os bairros de cada cidade mais próxima onde se situam os cursos de Letras.

Uma tal disciplina de Lingüística, na graduação, terá como objetivos principais: a) iniciar estudos para projetar um Banco de Dados informatizados, com a descrição das variantes encontradas nos municípios, priorizando, primeiramente, o aspecto fonético-fonológico, com vistas a oferecer subsídios mais consistentes e, em tempo mais curto, para aprimorar o ensino/aprendizagem da língua; b) incentivar a realização de trabalhos desse tipo nos cursos de Graduação e Pós-Graduação em Letras; c) efetuar análises do *corpus* gravado nas entrevistas; d) treinar os alunos para a elaboração de material para publicação e de material didático.

Posteriormente, poder-se-á descrever outros níveis da língua, como o morfo-sintático, o léxico-semântico e o textual, de forma que, a cada 10 ou 20 anos se possa recomençar pelo nível fonético-fonológico novamente. Aos poucos, iremos escrevendo a história da língua para os que nos seguirem.

Para a recolha dos dados e a transcrição do *corpus*, recomenda-se a metodologia da Geolingüística tradicional, mas, na apresentação dos resultados e na descrição e análise das variantes para fins de produção de material didático, já se pode contar com resultados mais rápidos e de divulgação mais acessível e mais barata, com os recursos da Informática. (Na secção, a seguir, serão apresentados, como resultado de 10 anos de pesquisa, informações sobre um *software* criado para agilizar a descrição das variantes de língua em qualquer localidade, estado ou país).

O questionário fonético-fonológico poderá ser organizado, juntando-se fatos lingüísticos de outros Atlas de uma determinada língua já produzidos. No Brasil, por

exemplo, um projeto que deseja descrever a língua portuguesa falada, de qualquer município, poderá reunir, aproximadamente, 60 fatos de um questionário prévio para o Projeto ALERS (cf. Mercer, 1989). A estes, poderão ser acrescentadas mais 180 perguntas a partir de vocábulos estudados no Atlas Lingüístico da Paraíba (cf. Aragão, 1998: 55-77), totalizando 210 perguntas. Outros fatos fonéticos poderão ainda ser acrescentados a partir da observação do pesquisador na área a ser pesquisada.

Elaborado o questionário dessa forma é possível se confrontar grande número de variantes de pontos estratégicos do Brasil: o Sul com o Nordeste, e posteriormente, o Norte com o Sul, etc. Eliminaremos, assim, os motivos da crítica feita aos primeiros cinco atlas lingüísticos do Brasil, em que apenas quatro das perguntas dos respectivos questionários são coincidentes, impossibilitando que sejam feitas comparações para determinação de áreas dialetais no Brasil.

As perguntas devem ser apresentadas com o número do item, seguido do vocábulo que é objeto de estudo, acompanhado do objetivo (variante alvo dentro do vocábulo) e da pergunta. Cada uma delas deve ser formulada de modo a ensejar que todos os informantes respondam com o mesmo vocábulo escolhido para objeto de estudo em uma determinada questão.

Trabalhos posteriores podem incluir outros tipos de questionários com características próprias, como questionário morfossintático, semântico-lexical e de textos, pois o aspecto fonético-fonológico apenas inicia o trabalho que está por se fazer em cada município.

Quanto aos informantes, a experiência recomenda inicialmente adultos, de nível de escolaridade entre 4ª a 8ª série, sendo um de cada sexo, com duas faixas de idade cada: a primeira entre 20 a 35 anos e a segunda, entre 46 a 60, deixando-se uma faixa intermediária de transição dos 36 aos 45 anos, ou incluindo-a para pesquisa mais detalhada.

O mesmo procedimento pode ser seguido para os outros dois níveis de escolaridade, a depender do bairro e do interesse da comunidade. O próximo passo pode ser trabalhar com informantes de cada sexo com idade entre 4 a 8 anos, seguindo-se os de 12 a 16 anos.

Quanto à transcrição fonética, recomenda-se que sejam feitas a lápis para resistir as rigorosas revisões necessárias nesta fase da pesquisa.

Para a definição da rede de pontos, recomenda-se atingir todos os bairros do município, em que se detectem influências de informantes descendentes de etnias diversificadas.

Quanto ao perfil dos inquiridores, estes devem ser alunos de Letras, cursando, de preferência, o segundo ano da disciplina de Lingüística, devidamente treinados pelo professor da disciplina para:

- a) discutir o Projeto e as perguntas formalizadas;
- b) fazer transcrição fonética;
- b) aplicar os questionários, preencher as fichas dos informantes e das localidades e gravar as respostas dos informantes;
- c) distribuir as localidades aos inquiridores (dois alunos por localidade para cada dois informantes), dando-se preferência, sempre que possível, ao bairro onde reside o inquiridor, ou ao bairro próximo de sua residência, caso necessário;
- d) selecionar informantes e aplicar questionário com gravação das respostas em fita cassete em um bom gravador;
- e) transcrever os dados em fichas, utilizando os símbolos do Alfabeto Fonético Internacional;
- f) proceder à informatização dos dados em um computador, no Sistema SPDGL, ou em outro disponível;
- g) participar de seminários de avaliação e planejamento, visando à publicação dos resultados da descrição das variantes da língua falada nos bairros.

Trabalho idêntico ao aqui sugerido foi aplicado na UNIVILLI (Universidade da Região de Joinville), nos primeiros e segundo anos do Curso de Letras. O material gravado pelos alunos está sendo informatizado no SPDGL para publicação.

5.2. O Sistema de Processamento de Dados Geolingüísticos (SPDGL)

Nesta última sessão, detalho o sistema de informatização do Banco do SPDGL, que vem sendo utilizando por estudantes e pesquisadores da área de Geolingüística.

Uma vez revisada a transcrição fonética dos dados gravados nas entrevistas, estes são digitados no SPDGL para obtenção do banco. O sistema permite que se obtenham vários tipos de relatórios: a) relatório de cada item-resposta em todos os pontos; b) relatório de cada ponto com todos os itens-resposta; c) relatório de cada variante em todos os pontos; d) relatório de cada ponto com todas as variantes; e) relatório do quadro de variantes fonéticas, objeto de estudo e, finalmente, das cartas lingüísticas descritivas.

Cada relatório, portanto, cumpre a função de subsidiar a implementação do projeto, seja para facilitar a revisão do dados com maior propriedade, seja para publicar os resultados em banco de dados ou em forma de atlas lingüístico.

A título de ilustração, incluo, neste artigo, alguns tipos de relatórios de dados lexicais, que foram alimentados com dados de Santa Catarina no Projeto ALERS (Atlas Lingüístico da Região Sul do Brasil), a saber: a carta lexical de “mutirão” em Santa Catarina, a carta lexical de “capinar” e a carta origem do povoamento de Santa Catarina.

BIBLIOGRAFIA

- AGUILERA, Vanderci Andrade. *Atlas Lingüístico do Paraná*, Curitiba: Editora da Universidade Estadual de Londrina, 1995.
- AGUIRELA, Vanderci de Andrade. *Geolingüística no Brasil: caminhos e perspectivas*. Editora da UEL, Londrina, 1998. p. 55-78.
- ARAGÃO, Maria do Socorro Silva de & MENESES, Cleusa Palmeira Bezerra de. *Atlas Lingüístico da Paraíba*, Brasília: Universidade Federal da Paraíba/CNPQ, 1984.
- CARDOSO, Suzana Alice Marcelino. O atlas lingüístico no Brasil: uma questão de política. In: *Caminhos e perspectivas para a Geolingüística no Brasil*, Salvador: Universidade Federal da Bahia - Instituto de Letras (1996: 87-100), 1996.
- CHAMBERS, Jak and TRUDGILL, Peter. *Dialectology*. New York/USA: Cambridge University Press, 1998.
- COMITÉ NACIONAL DO PROJETO ALiB. *Atlas Lingüístico do Brasil: Questionários*. Londrina: Editora da UEL, 1998.
- FERREIRA, Carlota da Silveira ; MOTA, Jacyra de Andrade ; FREITAS, Judith Mendes de Aguiar; ANDRADE, Nadja Maria Cruz de ; CARDOSO, Suzana Alice Marcelino ; ROLEMBERG, Vera Lúcia Sampaio & ROSSI, Nelson. *Atlas Lingüístico de Sergipe*, Salvador: Universidade Federal da Bahia / Fundação Estadual de Cultura de Sergipe, 1987.
- GOEBEL, Hans. Dialectometrie interponctuelle. In: *Revue Linguistique Romane*. Tome 47, Strasbourg, 1983. p. 353-412.
- KOCH, Walter. Atlas Lingüístico-Etnográfico da Região Sul do Brasil (ALERS). *Seminário Nacional Caminhos e perspectivas para a Geolingüística no Brasil*, Salvador: Universidade Federal da Bahia, 1996. p.55-77.

- KOCK, W.; KLASSMANN, M.; MERCER, J. L. V.; AGOSTINI, B.; FURLAN, O. A.; VIEIRA, H.G. & ALTENHOFEN, C.V. *Projeto Elaboração de um Atlas Lingüístico-Etnográfico da Região Sul (ALERS)*, 1989. (3 volumes no prelo)
- MOLLICA, Maria Cecília. *Influência da fala na alfabetização*. Rio de Janeiro: Tempo Brasileiro, Berlim, 1998.
- RADTKE, EDGA e THUN, Harald. Nuevos caminos de la geolingüística románica. In: Dialectologia pluridimensionalis românica 1. RADTKE/THUN (org.), *Neue Wege der Romanischen Geolinguistik : Akten des Symposiums zur Empirischen Dialectologie Heidelberg/Mains*, 21 - 24. 10, 1996: 1-49.
- RIBEIRO, José ; ZAGARI, Mário Roberto Lobuglio; PASSINI, José; GAIO, Antônio Pereira. *Esboço de um Atlas Lingüístico de Minas Gerais*, Rio de Janeiro: Fundação Casa de Rui Barbosa, 1977.
- ROSSI, Nelson. *Atlas Prévio dos Falares Baianos*, Rio de Janeiro: Instituto Nacional do Livro/MEC, 1963.
- VIEIRA, Hilda Gomes. Sistemas de transcrição e elaboração de cartas lingüísticas. In: *Caminhos e perspectivas para a geolingüística no Brasil*, Salvador: Universidade Federal da Bahia, 1996. p. 97-107.
- _____ Resultados do ALERS em Santa Catarina: africção palatal da oclusiva /t/ nas zonas rural e urbana. *Anais do I Encontro de Variação Lingüística do CONESUL*. Porto Alegre: UFRGS, set. 1996c. 10 p.
- _____ Fundamentos para organizar, implementar e manter um banco de dados geolingüísticos. In: AGUILERA, Vanderci de Andrade(org.) , *A geolingüística no Brasil*, Londrina: Editora UEL, 1998. p. 207-224.
- _____ Um balanço crítico da geografia lingüística no Brasil. In: *Estudos de geolingüística do Português Americano*, Ed. Frankfurt am Main: TMF-Frankfurt, 2000. p. 15-25.
- _____ *Sistema de processamento de dados geolingüísticos (PDGL 1.6)*. Software elaborado no Departamernto de Língua e Literatura Vernáculas da Universidade Federal de Santa Catarina, alimentado com dados do Projeto ALERS, 1999.
- _____ Importância do atlas lingüístico no ensino/aprendizagem da escrita. In: *Estudos Lingüísticos* , v. 28. Jaurú/SP : Editora da Universidade do Sagrado Coração de Jesus, 1999. p.263-268.

ANEXO 1

Carta n° 1: Carta lexical das variantes *mutirão/puxirão*, na fala gravada em 1990, de 80 informantes da zona rural do Estado de Santa Catarina. (*Banco de Dados Geolinguísticos - ALERS/SC-UFSC*). Aqui, estão registradas apenas as duas variantes mais frequentes:



ANEXO 2

Carta n° 2: Carta lexical das variantes de *capinar/carpir*, fala gravada em 1990, de 80 informantes da zona rural do Estado de Santa Catarina (*Banco de Dados Geolingüísticos – ALERS/SC-UFSC*):



ANEXO 3

Carta nº 3 : Origem do povoamento do Estado de Santa Catarina.

