

PESQUISAS LINGÜÍSTICAS: NOVOS RECURSOS

Alfredo Maceira Rodríguez (UCB)

1. INTRODUÇÃO

Embora haja notícias de preocupações com estudos lingüísticos muito anteriores, os estudos que permitiram que as ciências lingüísticas pudessem vir a ter *status* de verdadeira ciência, com objeto de estudo definido e metodologia própria, ocorreram no século XIX, no auge do Evolucionismo, estudos que nas ciências da linguagem propiciaram pesquisas em línguas tão afastadas no tempo e no espaço como o sânscrito e várias protolínguas que remontam ao proto-indoeuropeu. O método que dominou quase todo o século XIX e parte do XX foi o denominado Histórico-comparativo. Nele, os elementos lingüísticos das mais diversas línguas em estudo são recolhidos, analisados, classificados e comparados. Trata-se de demonstrar a existência de protolínguas, numa época em que o Romantismo era o ideal em que se buscava a valorização das nacionalidades, com base numa mitificada tradição medieval. Por outro lado, procurava-se provar que as línguas eram organismos vivos, que seguiam as leis dos demais organismos biológicos: nascimento, crescimento, reprodução e morte. Daí a terminologia ainda hoje em uso: famílias lingüísticas, língua-mãe, língua materna, etc.

2. RECURSOS DE PESQUISA

2.1. Período Historicista

Ainda no século XVIII, William Jones e outros estudiosos europeus visitavam países distantes como a Índia e faziam observações e anotações sobre as línguas ali encontradas, mantendo, muitas vezes, correspondência entre si.

Os recursos então disponíveis eram principalmente as fontes escritas, tendo que se conformar quase sempre com fragmentos de textos em qualquer tipo de suporte que se pudessem obter graças ao esforço de preservação de gerações sucessivas para finalidades religiosas, como os referentes aos Vedas¹, ou ainda possíveis inscrições, geralmente em

¹ Os textos védicos datam de 1500 a 500 a C. A literatura era preservada pela tradição oral. Atribuía-se um texto a cada família, que se encarregava de memorizá-lo e transmiti-lo aos

templos, túmulos ou outros monumentos do passado, conservados desde a Antigüidade pela tradição. O máximo que esses devotados estudiosos conseguiam obter era algum documento original ou, mais comumente, sua cópia, penosamente transcrita à mão. No tocante à língua oral, pode-se imaginar a dificuldade para tentar transcrever fonemas de línguas desconhecidas ou não cultivadas literariamente. Mas os únicos recursos eram esses, já que não havia qualquer possibilidade de conservar o som, recurso inventado e popularizado muito tempo depois.

2.2. A Geografia Lingüística

No final do século XIX e inícios do XX ganharam interesse os estudos dialetológicos. Aqui também era necessária a transcrição fonética, com a diferença de que as variantes deviam ser conhecidas dos pesquisadores. O principal problema ainda se situava na transcrição. Agora não só interessavam as variantes fonológicas, mas também as fonéticas. A Geografia lingüística é uma consequência do interesse pelos estudos dialetais levados a cabo de início por vários estudiosos europeus, entre os que se destaca o italiano G. I. Ascoli, mas a Geografia lingüística popularizou-se realmente na França, onde Gilliéron publicou o *Atlas Lingüístico da França*, em Paris, (1902-1910). Este método consiste na representação cartográfica das variedades dialetais de uma determinada região. A dialetologia estuda as variantes dialetais de determinada área, mas à Geografia lingüística cabia, além da coleta, a análise e interpretação dos dados. Os recursos usados eram em geral a interrogação direta dos falantes e a transcrição do que o entrevistador ouvia do informante. Em geral não se estudava a língua das grandes cidades porque era onde as inovações ocorriam em primeiro lugar, devido, em grande parte, ao cruzamento de culturas e à concentração de falantes de diversas áreas geográficas. Assim os pesquisadores se expandiam por todo o território determinado para a pesquisa, procurando que a malha da área estudada fosse o mais densa possível para cobrir todo o território objeto da pesquisa, sem esquecer as pequenas localidades mais afastadas, que costumam ser as mais conservadoras.

Depois do mapa lingüístico da França, foram elaborados outros na Europa, em áreas menores como Córsega, Andorra e outros territórios, procurando aperfeiçoar os critérios de elaboração, dando-se quase sempre

descendentes. Estes textos acabavam tornando-se conhecidos pelo nome da família encarregada de conservá-los. Conservaram-se 5 livros de hinos, também com partes em prosa.

preferência ao material vocabular. O procedimento podia variar do interrogatório direto, feito pelo inquiridor ao informante, à remessa de questionários via postal.

Outras metodologias de pesquisa lingüística de alguma forma relacionadas com a Geografia lingüística, como a Onomasiologia, empregavam iguais ou parecidos processos de pesquisa e tinham as mesmas dificuldades.

3. A GRAVAÇÃO

A Geografia lingüística e estudos similares, atingiram seu apogeu nas décadas de 1910 a 1950, embora ainda se continuem fazendo alguns estudos nessa direção, ocupando-se primordialmente com o léxico. Posteriormente vem-se dando ênfase à Fonética e à Fonologia. O estudo destas disciplinas foi muito facilitado com a gravação sonora. Na década de 1960 foi inventada a fita cassete em consequência do que logo surgiram gravadores portáteis cada vez mais aperfeiçoados que permitiram copiar, reproduzir e conservar a voz com absoluta fidelidade para posteriores estudos. Não só permitiram maior precisão nos estudos fonéticos como nos lexicais e sintáticos, ou seja, em tudo o relacionado com a comunicação oral. A partir de então, os pesquisadores lingüísticos ficaram de posse de um excelente recurso que lhes abriu um amplo campo no estudo da fala.

4. A LÍNGUA ESCRITA

Os modernos recursos de gravação da fala não ocasionaram o desinteresse pelo estudo da língua em sua modalidade escrita. Ao contrário, desenvolveram-se recentemente várias linhas de estudo sobre a organização do discurso como a coesão e a coerência textuais. Por outro lado, o estudo de textos antigos permanece em evidência. Como vimos, a língua do passado conservou-se na forma gráfica. Para deduzir do texto sua possível pronúncia recorre-se aos estudos comparativos, como ocorre principalmente com o latim para quem o estudo da fonética das línguas neolatinas permitiu que se reproduzisse sua pronúncia hipotética. Esses textos são minuciosamente estudados pelos filólogos, constituindo uma disciplina conhecida como Ecdótica ou Crítica textual, que longe de estar esgotada, ainda oferece um vasto campo de estudos dentro da complexidade dos estudos lingüísticos. Embora sem possibilidade de qualquer registro sonoro, a Ecdótica conta atualmente com modernos processos de fotocópia e

microfilmagem, que facilitam a leitura e elucidação de textos antigos, geralmente manuscritos, assim como com os recursos da Informática para organização, tabulação e confronto dos elementos lingüísticos dos textos em estudo.

5. O COMPUTADOR

Se o gravador permitiu ampla coleta de dados relativos à fala, o computador, inicialmente (década de 1970) já podia ser utilizado pelos lingüistas para o processamento dos dados recolhidos nas pesquisas.

Com o surgimento dos computadores pessoais (microcomputadores), os pesquisadores das ciências lingüísticas tiveram seu trabalho muito facilitado. Na Lexicografia tornou-se desnecessário o manuseio de fichas, que era prática tradicional indispensável. A dificuldade ainda permanece muitas vezes no pesquisador, principalmente quando já está habituado a exercer sua atividade de acordo com os meios tradicionais e com os quais vem obtendo bons resultados. Este pesquisador pode relutar ou ter dificuldade a adaptar-se em um meio tão novo e que evolui com tanta rapidez e do qual ainda não sente necessidade. A verdade é que a Informática não é um modismo. Veio para ficar e está alterando muitas características tradicionais do cotidiano, até mesmo do homem comum, em todo o mundo, para melhor ou para pior. É verdade que nem tudo pode ser resolvido com o computador, pelo menos até os dias de hoje. No que diz respeito aos estudos lingüísticos, um interessante trabalho de Gary F. Simons², na Internet, analisa a situação dos pesquisadores lingüísticos em relação a seu uso da Informática e destaca a falta de programas específicos para esse fim. Assinala que os editores de texto existentes são deficientes (em 1993) devido a seu restrito número de caracteres, (256) no código ASCII. Com esta limitação tornava-se difícil trabalhar com textos multilíngües, geralmente os mais empregados pelos pesquisadores, devido, entre outras peculiaridades, à existência de diacríticos. Por exemplo, apenas para codificar o alfa minúsculo do grego são necessários 12 caracteres compostos diferentes, levando-se em conta marcações de espírito e de acentos.

Estas restrições já estavam em fase de superação com o desenvolvimento de novos programas que estavam em estudo e que

² SIMONS, Gary F. *The Nature of Linguistic Data and the Requirements of a Computing Environment for Linguistic Research*. Summer Institute of Linguistics. Dalas. TX., 1993.
GarySimons@sil.org

prometiam satisfazer as expectativas dos pesquisadores de poder dispor para seu uso de um verdadeiro ambiente multilingüístico.

Outro problema que oferecia dificuldades era a seqüência que define a ordem alfabética para listagem em cada língua. Por exemplo, a seqüência de caracteres *ll* em inglês situa-se entre *li* e *lo*, mas em espanhol encontra-se entre *lu* e *ma*.³ No espanhol o dígrafo *ch* tem registro independente no dicionário, no que difere do português e de outras línguas. Recentemente houve polêmica sobre a permanência ou não do *ñ* na linguagem da Internet, o que acabou sendo aceito.

Outro problema apresentado no referido estudo era a dificuldade do estabelecimento da hierarquia de dados lingüísticos, porque os editores de texto até então disponíveis tratam os dados necessários para uma entrada lexical como uma seqüência de parágrafos. Assim não haveria como organizar a estrutura hierárquica de seus elementos: palavra-base, etimologia, subentrada e exemplos dentro da entrada.

Estes problemas estão hoje superados por novos programas que estão continuamente sendo elaborados e aperfeiçoados. Problemas que ainda possam existir podem ser solucionados. Compete aos usuários lingüistas apresentar sugestões aos especialistas de Informática para que possam criar e aperfeiçoar programas úteis e práticos para facilitar o trabalho que se quer realizar. Os estudos lingüísticos tornam-se cada vez mais interdisciplinares. No futuro será quase impossível realizar pesquisas sem a concorrência de outras ciências e disciplinas, afins ou não.

6. AS REDES MUNDIAIS

O computador pessoal desde o início trouxe grandes vantagens para o pesquisador lingüístico, mas ele ainda era uma ferramenta isolada: permitia a transcrição de dados gravados, facilitava seu processamento (cálculo, colocação em ordem alfabética, possibilidade de sua localização, substituição, etc.). Tudo isto pode ser feito com um processador de texto, ferramenta indispensável para quem trabalha com línguas. Isto já é muito, porém a evolução não pára. Faz bem poucos anos, o computador individual passou a ligar-se e comunicar-se com outros, próximos ou distantes. Logo foram surgindo redes de computadores, até ficar disponível a mais internacional e popular de todas: a Internet. O número de computadores ligados a essa rede em todo o mundo conta-se por dezenas (ou já centenas) de milhões em todo o mundo e este número aumenta constantemente.

³ Exemplo: Ing. *lizard*, *llama*, *load*. Esp. *luz*, *llama*, *maca*.

Para o pesquisador lingüístico, sua importância parece que ainda não foi totalmente avaliada. É fácil imaginar o que se pode fazer: ter acesso a bibliotecas, universidades e especialistas em qualquer parte do mundo, enviar e receber mensagens, informações e documentos com extraordinária rapidez, trabalhar em equipe com pesquisadores, sem preocupação com a distância, etc. Algo disto já se fazia ou ainda se faz por meio do correio, mas pode-se calcular a dificuldade que isto supunha e, principalmente, a demora. Por meio da Internet é possível enviar textos de qualquer tamanho, até mesmo anexados a uma simples mensagem. Já existe possibilidade de transmitir, copiar e reproduzir arquivos sonoros. Podemos imaginar o que isso representa para a pesquisa da fala. Suponhamos um estudo que requeira informações sobre os dialetos de uma pequena ilha distante e dos quais existam poucos especialistas e ainda que estes poucos estejam dispersos em diversas partes do mundo, além do que suas ocupações e outras dificuldades não lhes permitam contatos com os colegas, a não ser por telefone ou correio. Por meio da Internet podem receber o material gráfico e/ou sonoro requerido para estudo. Os programas de computador hoje disponíveis já lhes permitem processar esses documentos com bastante rapidez e precisão. O que antes era um processo demorado e paciente, agora pode ser realizado num período de tempo muito menor e sem tantos problemas inerentes ao processo tradicional.

7. O FUTURO DA PESQUISA LINGÜÍSTICA

Agora se impõe uma pergunta: O pesquisador lingüístico será dispensável? Tudo leva a crer que não. Mudam os processos, mas os problemas persistem. Os temas lingüísticos a demandarem estudo aprofundado são múltiplos. Grande parte deles tinha sido deixada de lado porque não havia condições de executá-los e agora os novos recursos de pesquisa e análise os tornam realizáveis. Além disso, as línguas evoluem sempre e essa evolução agora pode ser acompanhada de perto. Quando se concluía um estudo de um determinado traço dialetal e se publicavam seus resultados, estes já podiam diferir bastante de quando foram realizados. Hoje é possível concluir estudos e publicar seus resultados numa fração do tempo anteriormente requerido.

O computador, inicialmente concebido para o cálculo, tornou-se uma máquina indispensável na vida moderna. No decorrer de algumas décadas (a partir dos anos 40) desenvolveu-se tanto que originou uma nova ciência: a Informática. Hoje seria difícil imaginar a vida, principalmente nas cidades,

sem sua participação. Como funcionaria o sistema bancário e até o comércio em geral?

A rápida evolução e o amplo emprego da Informática de certa forma chocou as pessoas metódicas, com formação tradicional. Instintivamente o homem resiste às inovações e de início duvida de sua eficácia. Alguns até se recusam a tomar conhecimento de qualquer inovação, seguros que estão do que bem aprenderam e solidificaram ao longo da vida. Os jovens são mais abertos a inovações, por isso se diz que aprendem melhor na área de Informática do que os mais velhos. O que ocorre é que os jovens já se criaram expostos às inovações tecnológicas e, por outro lado, estão desprovidos de resistência ao novo, ao contrário, para eles é um desafio próprio da idade. Mas está se provando que o aprendizado da Informática, como qualquer outro, não depende de idade e sim de interesse e dedicação. No caso dos que se dedicam a estudos lingüísticos, a Informática veio a facilitar-lhes muito o trabalho. Ninguém pensará em organizar um dicionário, por exemplo, usando as fichas tradicionais, se pode fazê-lo muito melhor criando um arquivo no computador. Além da facilidade de processamento, tem a de cópia, revisão, etc. O mesmo se pode dizer de qualquer trabalho com textos. Embora possam existir casos em que a Informática seja de pouca ajuda, as tarefas mais maçantes e repetitivas podem ser facilmente executadas.

Uma área em que a Informática ainda não se tem tornado eficaz é na tradução, sobretudo na literária. As sutilezas do sentido em textos em que predomina a conotação fazem com que a tradução sem tradutor humano (com programa de tradução) ainda não seja satisfatória. É um dos grandes desafios dos criadores de programas destinados a traduzir de uma língua para outra. Eles já existem no mercado, mas seu resultado ainda parece não ser confiável, apesar de ter melhorado muito na área técnico-científica, onde predomina a denotação e um certo internacionalismo no campo lexical.

Se tentarmos traduzir qualquer texto um pouco longo entre duas línguas historicamente próximas e bem conhecidas, deparamo-nos com casos em que não nos satisfaz qualquer solução que se nos ofereça. Há casos em que temos que valer-nos de circunlóquios ou de notas explicativas para transmitir a idéia da língua-fonte. É fácil supor o que fará o tradutor automático no estágio em que hoje o conhecemos. Então parece que esse tipo de tradução nunca será possível e não falta quem afirme isso, mas aqui cabe a pergunta. Apenas trinta ou quarenta anos atrás seria fácil conceber o que hoje conseguimos realizar com a Informática? Supomos que não. Parece que nessa área todos não haverá solução impossível em um tempo relativamente curto, levando em conta os avanços obtidos em umas poucas décadas.

Aos pesquisadores lingüísticos e a todos os que trabalham com textos não lhes devem faltar atividades e desafios, pois as línguas cada vez se tornam mais complexas e o campo de estudos continua ampliando-se. As técnicas e os métodos de pesquisa e trabalho mudam com a mesma velocidade da Informática. Num futuro próximo, a julgar pelos novos recursos que diariamente estão sendo postos à disposição dos pesquisadores lingüísticos, teremos muito maior conhecimento da linguagem humana, de seu comportamento e mutabilidade assim como de todas as línguas e dialetos de que se tem notícia. Vale a pena engajar-se e persistir neste campo de estudos.

8. BIBLIOGRAFIA

- IORDAN, Iorgu. *Introdução à lingüística românica*. 2. ed. Trad. Júlia Dias Ferreira. Lisboa : Calouste Gulbenkian, [1982].
- LYONS, John. *Linguagem e lingüística: Uma introdução*. Trad. Marilda Winkler Averbug & Clarisse Sieckenius de Souza. Rio de Janeiro : Zahar, 1982.
- MIAZZI, Maria Luísa Fernandez. *Introdução à lingüística românica: Histórico e métodos*. São Paulo : Cultrix, 1976.
- ROBINS, R. H. *Lingüística geral*. 2. ed. Porto Alegre–Rio de Janeiro : Globo, 1982.
- . *Pequena história da lingüística*. Rio de Janeiro : Ao Livro Técnico, 1979.