

## UM ESTUDO SISTEMÁTICO SOBRE AS VARIEDADES LINGUÍSTICAS DO NORTE FLUMINENSE

LUQUETTI. Eliana Crispim França  
FERREIRA. Dhienes Charla  
PAES JÚNIOR. Gelson Caetano

Universidade Estadual do Norte Fluminense Darcy Ribeiro- UENF

[elinaff@gmail.com](mailto:elinaff@gmail.com)  
[dhienesch@hotmail.com](mailto:dhienesch@hotmail.com)  
[gelcapaio@yahoo.com.br](mailto:gelcapaio@yahoo.com.br)

**Resumo:** O objetivo do presente projeto foi apresentar a elaboração de um *corpus* representativo da região Norte Fluminense. A elaboração de um *corpus* dessa natureza é de suma importância para futuras pesquisas linguísticas na região, não somente para os pesquisadores de diversas áreas. Além disso, este banco de fala permitirá registrar e resgatar a memória, as características de fala de seus habitantes em oposição aos traços linguísticos típicos dos habitantes de outras regiões. Pretende-se que esta proposta seja ampliada para toda a região Norte Fluminense, a fim de atender as demandas e particularidades das localidades. Em primeiro lugar, realizou-se um estudo mais aprofundado da estrutura social das cidades que compõem a região supracitada. Posteriormente, foi feita uma coleta de amostras da cidade de Campos para a elaboração de uma análise da fala da região. Sua exatidão, característica necessária ao meio acadêmico, favorece ao estudo da variedade linguística e da oralidade, pois apresenta situações reais de uso da fala de forma dinâmica. Buscou-se adaptar as estratégias de elaboração do *corpus* D & G Gramática já existente, constituído de professores da UFRJ, da UFF, da UERJ e da UFRN, para os objetivos que no momento nortearam a presente proposta.

**Palavras-chaves:** variedades linguísticas; dialetologia; memória; fala.

### Introdução

Este trabalho foi realizado pelo *Grupo de Estudos Linguagem e Educação* sob a coordenação da professora Eliana Crispim França Luquetti, da Universidade Estadual do Norte Fluminense Darcy Ribeiro (UENF), que atuou na supervisão da coleta e na organização do *corpus* “A língua falada e escrita na região Norte Noroeste Fluminense”.

A questão da constituição de um banco de dados de língua falada pressupõe uma vinculação teórica que, de um lado, estabelece o uso linguístico numa comunidade de fala como objeto de estudo e, de outro, apresenta um conjunto de procedimentos metodológicos para que este objeto possa ser estudado.

Nosso objetivo, aqui, é oferecer materiais linguísticos, que caracteriza a concepção de um banco de dados de língua oral em situação real de uso e língua escrita, com marcas da oralidade. Além disso, esse banco de dados representa um recorte de uma comunidade de fala, que pode ser aprendida de forma mais abrangente, constituindo uma amostra estratificada de acordo com determinados parâmetros sociais, ou representar apenas um determinado

segmento. Como, os parâmetros sociais estabelecidos por William Labov, que iniciou os estudos referentes à Sociolinguística, foram idade, sexo, classe social, estilo de fala e etnia. O tipo de distribuição de dados variáveis em função de faixa etária e da classe social tem servido nos estudos clássicos como indicadores da direção do processo: se de mudança ou de variação estável. Optamos por utilizar o parâmetro escolaridade ao invés de classe social devido a diversos *corpora* constituídos no Brasil (Amostra Censo, por exemplo) utilizarem a mesma variável.

Dessa maneira, um banco de dados é capaz de prover um conjunto considerável de dados sobre a variação estruturada de uma determinada comunidade de fala. Para tanto, precisa ser constituído de acordo com determinados procedimentos metodológicos, que procuram assegurar que o comportamento que ele teria se não estivesse sendo observado nas situações reais de uso. A entrevista deve ser conduzida de maneira que, por meio de técnicas específicas, seja possível apreender o comportamento em todos os sentidos do falante. Um dos objetivos da entrevista é atingir o vernáculo, ou seja, aquela situação comunicativa em que nenhuma atenção é prestada à fala.

A constituição de um banco de dados de língua falada, fundamentados nos procedimentos da sociolinguística variacionista, constitui valiosa fonte de dados para o pesquisador, que pode aliar a explicação de estudos de caso de fenômenos específicos tanto sintáticos quanto fonológicos à discussão de questões teóricas mais amplas, assim como, outras áreas de conhecimentos podem se apropriar desses dados a fim de compreender aspectos diversos da linguagem, da comunidade, da cultura local e outros.

As entrevistas foram realizadas em Campos dos Goytacazes no período de junho a dezembro de 2011 e quatro foram os objetivos principais que nortearam o levantamento deste *corpus*, a saber:

- a) Analisar o comportamento da iconicidade (semelhança existente entre a forma e a coisa representada), por meio de diferentes fenômenos linguísticos, em situações reais de uso da língua;
- b) Criar um banco de dados com correspondência de conteúdo entre fala e escrita, de modo a viabilizar a comparação mais rigorosa entre essas duas modalidades da língua;
- c) Testar em diferentes subgêneros textuais o modo de codificação da informação;
- d) Comparar o comportamento dos canais da fala e da escrita em relação a esses subgêneros e outros que possivelmente emergiram de análise futuras.

## **1. A constituição do *corpus***

A relevância deste trabalho é legitimada com base em pesquisas anteriores realizada pelo grupo Discurso & Gramática que trabalha com estudos na área de linguística funcional, voltado especialmente para os processos de mudança linguística e gramaticalização. O Grupo de Estudos Discurso & Gramática é constituído de professores da UFRJ, da UFF, da UERJ e da UFRN.

O Grupo foi fundado no Departamento de Linguística e Filologia da Faculdade de Letras em 1991, pelo professor Sebastião Josué Votre. Seu primeiro projeto integrado de pesquisa, apoiado pelo CNPq, foi Iconicidade na fala e na escrita, com duração de dois anos. À base desse projeto integrado, assim como das demais pesquisas desenvolvidas pelo grupo, está um *corpus* em que se organizaram amostras de língua falada e escrita com informantes em quatro cidades brasileiras: Rio de Janeiro, Natal, Rio Grande e Juiz de Fora. Quatro foram os objetivos principais que nortearam o levantamento do *corpus* Discurso & Gramática:

- a) Analisar o comportamento da iconicidade, através de diferentes fenômenos lingüísticos, em situações reais de uso da língua;
- b) Criar um banco de dados com correspondência de conteúdo entre fala e escrita, de modo a viabilizar a comparação mais rigorosa entre essas duas modalidades da língua;
- c) Testar em diferentes subgêneros textuais (narrativa de experiência pessoal, narrativa recontada, descrição de local, relato de procedimento e relato de opinião) o modo de codificação da informação;
- d) Comparar o comportamento dos canais da fala e da escrita em relação a esses subgêneros.
- e)

Nesse sentido, o que se pretendeu foi adaptar as estratégias de elaboração do corpus D&G da UFRJ aos objetivos pertinentes deste trabalho. Assim, O *corpus* “A língua falada e escrita na região norte-noroeste fluminense” é composto por depoimentos de 106 informantes. Cada um destes produziu cinco tipos distintos de textos oral e escrito, que são os seguintes tipos de textos: 1) Narrativa de experiência pessoal; 2) Narrativa de experiência recontada; 3) Descrição de local; 4) Relato de procedimento; e 5) Relato de opinião.

Visando verificar se os fenômenos investigados poderiam sofrer influência do grau de escolarização do falante, selecionamos falantes que estivessem cursando diferentes níveis de escolaridade: como: Educação de Jovens e Adultos (EJA), Ensino Fundamental I e II, Ensino Médio, Ensino Superior completo e incompleto. Além disso, entrevistamos informantes de diferentes grupos sociais.

Controlamos a variável sexo, distribuindo informantes femininos e masculinos em cada um dos grupos pesquisados. Há correlação estreita entre idade e escolaridade. Assim, definem-se, em linhas gerais, as seguintes faixas etárias: de 9 a 20 anos e acima de 21 anos.

## 2. Procedimentos da coleta do material

A entrevista para a coleta dos dados é altamente estruturada, e o informante já sabe, de antemão, quais são os cinco itens que irá abordar. Tanto na fala como na escrita, ele sabe também a finalidade da coleta e qual a destinação acadêmica e social da mesma. Além disso, tem garantido o seu anonimato e negocia livremente com o pesquisador os melhores horários e locais em que vai conceder novos dados.

A coleta foi feita, essencialmente, pelos bolsistas. Nesta tarefa, cada um deles ficou encarregado de um determinado tipo de informante.

Iniciamos nossas atividades com o treinamento dos bolsistas. Nesse sentido, organizamos sessões de trabalho em que apresentamos e discutimos os procedimentos de coleta, assim como as dificuldades desse tipo de tarefa.

Na coleta de dados, foram seguidas algumas instruções necessárias para garantir uma condição de comunicação que se aproximasse o máximo possível de uma situação real e espontânea de interação. Para tanto, cada entrevistador deveria adotar os seguintes processos:

- Esclarecer o informante sobre a finalidade da coleta: o projeto de pesquisa e sua destinação acadêmica;
- Obter do informante a concordância em participar da entrevista, resguardando-se o anonimato deste por meio do uso de somente seu primeiro nome;
- Orientar o informante a respeito dos temas e questões que seriam propostos no depoimento (oral) de sua expressão;
- Ser preciso e explícito quanto ao tipo de texto a ser coletado: 1) Narrativa de experiência pessoal: conte uma história que tenha ocorrido com você que tenha sido

interessante, triste ou alegre; 2) Narrativa de experiência recontada: conte uma história, que tenha ocorrido com alguém que você conheça, que tenha sido interessante, triste ou alegre. Solicitamos, para este último tipo de texto, que fossem evitados comandos do tipo conte alguma coisa, assim como narrados filmes ou novelas que o informante tivesse visto; 3) Descrição de local: descreva, diga como é o lugar onde você mais gosta de ficar, passear ou brincar; 4) Relato de procedimento: você sabe fazer alguma coisa? O quê? Conte como se faz isto; e 5) Relato de opinião: a) Para os professores do 1º ciclo do Ensino Fundamental de Campos dos Goytacazes: o que você acha sobre o ensino de Língua Materna?; e b) Para os demais informantes residentes na cidade de Campos: o que você acha da cidade de Campos dos Goytacazes?;

- Escolher um lugar bem silencioso;
- Gravar uma fala sua para verificar a qualidade da gravação;
- Deixar o gravador ligado desde o início e durante toda a entrevista;
- Gravar no início da entrevista os dados pessoais do informante;
- Ser cordial, procurar vez por outra, dirigir-se ao informante pelo nome;
- Planejar com o informante o encontro para a coleta de dados;
- Coletar os textos orais em uma única sessão;
- Coletar os textos orais em sessões individuais;
- Recomeçar a coleta oral, se o texto produzido não estivesse de acordo com o tipo de texto definido no projeto;
- Não tornar a entrevista um interrogatório;
- Evitar as interrupções; no entanto, em qualquer momento do registro oral, todo informante pôde solicitar interrupção de gravação, tanto para se esclarecer, com relação ao tipo de texto que estava sendo coletado, quanto para se reorientar ou por motivo de natureza outra;
- Preencher a ficha de identificação de cada informante (Apêndice A); e
- Preencher as fichas de coleta de dados escritos (Apêndice B).

Todo material coletado passou por três fases de validação. A primeira validação ocorreu após a coleta do material oral, foram ouvidos todos os áudios para que se pudesse proceder à transcrição. Alguns informantes, depois de feita a entrevista, optaram por não fazer o texto escrito. Por esse motivo, os textos orais destes encontram-se disponíveis em arquivo.

Depois da primeira validação os textos orais foram transcritos, utilizando-se algumas das normas utilizadas pelo Projeto NURC/SP e outras estabelecidas pelo Grupo Discurso & Gramática.

A segunda validação ocorreu após a transcrição dos textos foi, então, procedida sua validação da transcrição feita. Para realizar esta etapa, o material foi trocado entre os bolsistas, devendo cada um deles ouvir os áudios, comparando os textos orais com as transcrições feitas pelos colegas. Essa fase provocou muita alteração.

Na validação final, os textos digitados passaram por uma nova revisão, com consulta à versão original das redações, e com nova escuta dos áudios. Em vista desse novo controle, a parte escrita foi representada na digitação exatamente como estava na ficha de coleta de dados escritos. Quanto à parte oral, procurou-se representar a fala do modo mais fiel possível.

Este projeto será fundamental nas pesquisas do *Grupo de Estudos Linguagem e Educação/UENF* e para outros estudiosos da linguagem das mais diferentes áreas do conhecimento.

### 3- Pesquisas anteriores que serviram de base a este trabalho

#### 3.1- O Corpus Discurso & Gramática

O Grupo de Estudos Discurso & Gramática é constituído de professores da UFRJ, da UFF, da UERJ e da UFRN.

O Grupo foi fundado no Departamento de Linguística e Filologia da Faculdade de Letras em 1991, pelo professor Sebastião Josué Votre. Seu primeiro projeto integrado de pesquisa, apoiado pelo CNPq, foi *Iconicidade na fala e na escrita*, com duração de dois anos.

À base desse projeto integrado, assim como das demais pesquisas desenvolvidas pelo grupo, está um *corpus* em que se organizaram amostras de língua falada e escrita com informantes em quatro cidades brasileiras: Rio de Janeiro, Natal, Rio Grande e Juiz de Fora. Quatro foram os objetivos principais que nortearam o levantamento do *corpus* Discurso & Gramática:

- Analisar o comportamento da iconicidade, através de diferentes fenômenos lingüísticos, em situações reais de uso da língua;
- Criar um banco de dados com correspondência de conteúdo entre fala e escrita, de modo a viabilizar a comparação mais rigorosa entre essas duas modalidades da língua;
- Testar em diferentes subgêneros textuais (narrativa de experiência pessoal, narrativa recontada, descrição de local, relato de procedimento e relato de opinião) o modo de codificação da informação;
- Comparar o comportamento dos canais da fala e da escrita em relação a esses subgêneros.

Ao todo, o *corpus* do grupo D & G é composto por depoimentos de 93 informantes. Cada um destes produziu cinco tipos distintos de textos orais (narrativa de experiência pessoal, narrativa recontada, descrição de local, relato de procedimento e relato de opinião) e, a partir destes, cinco textos escritos, para assim garantir a comparabilidade entre os canais: falado e escrito, o que totaliza 928 registros.

Visando verificar se os fenômenos investigados poderiam sofrer influência do grau de escolarização do falante, foram selecionados falantes que estivessem cursando diferentes séries da escola regular. Os informantes apresentam então os seguintes níveis de escolaridade: Classe de alfabetização infantil – Classe de alfabetização adulta, 4<sup>a</sup> série do primeiro grau, 8<sup>a</sup> série do primeiro grau, 3<sup>a</sup> série do segundo grau e último ano do 3<sup>o</sup> grau. Foi também controlada a variável sexo, com a distribuição equilibrada dos informantes femininos e masculinos em cada um dos graus de escolarização pesquisados.

A entrevista para a coleta dos dados é altamente estruturada, e o informante já sabe, de antemão, quais são os cinco itens que irá abordar. Tanto na fala como na escrita, ele sabe também a finalidade da coleta, e qual a destinação acadêmica e social da mesma. Tem garantido o seu anonimato, e negocia livremente com o pesquisador os melhores horários e locais em que vai conceder novos dados.

A coleta foi feita, essencialmente, pelos bolsistas de Iniciação Científica. Nesta tarefa, cada um deles ficou encarregado de um grau de escolarização. As atividades foram iniciadas com o treinamento dos bolsistas. Nesse sentido, foram organizadas sessões de trabalho em que os procedimentos de coleta, assim como as dificuldades desse tipo de tarefa foram apresentados e discutidos.

Então, na coleta de dados, foram seguidas algumas instruções necessárias para garantir uma condição de comunicação que se aproximasse o máximo possível de uma situação real e espontânea de interação.

### 3.2- Outros estudos

William Labov, em 1963, elaborou uma interessante pesquisa linguística na ilha de Martha's Vineyard, um lugar relativamente isolado, situado perto da costa do estado de Massachusetts, nos Estados Unidos. A pesquisa observou a influência do contato dialetal entre moradores da ilha e veranistas. Pelo resultado das análises, pôde-se perceber que os fatores *distribuição geográfica* e *faixa etária* foram importantes no sentido de que a zona rural e os falantes entre 31 e 45 anos, exatamente a faixa populacional que, apesar do processo de recessão econômica na ilha, decidiu por lá permanecer, mantiveram traços linguísticos típicos da ilha. Esses informantes constituíam os que mais reagem à invasão dos veranistas. Essa reação revelava, através da demarcação linguística, um sentimento positivo em relação à ilha. A partir daí, Labov propôs a noção de *orientação para a identidade*. Assim, diz LABOV (1995):

Nosso desejo é entender a estrutura interna do inglês vineyardense, incluindo as diferenças sistemáticas que já existem e as mudanças que estão ocorrendo agora na ilha. Para tanto, selecionaremos para estudo um aspecto linguístico característico de Martha's Vineyard com mais amplo espectro possível de variação e o mais complexo padrão de distribuição.  
(LABOV, 1995, p. 25).

Essa noção indica uma atitude positiva em relação ao local, contrariamente àqueles que, possuindo um sentimento negativo em relação à ilha, rejeitavam a norma local, adotando uma pronúncia mais próxima ao inglês padrão.

Estudos semelhantes já foram elaborados no Brasil. Um deles foi elaborado por Bortoni (1985), que na década de 1980, analisou a fala de imigrantes originários da zona rural do Alto Paraíba em Minas Gerais, radicados em Brasilândia, cidade satélite de Brasília. Esse trabalho demonstrou que há alteração em certas características do dialeto regional dos migrantes que chegam à cidade já adultos. Esse processo é influenciado pelos padrões de socialização no novo ambiente.

Outro estudo de natureza semelhante (Ferrari: 1994) analisou a fala da comunidade, relativamente isolada, do Morro dos Caboclos, na cidade do Rio de Janeiro. O objetivo era verificar os reflexos que o isolamento e, por outro lado, o contato com a cidade que poderiam ter nos traços linguísticos dos informantes. Foi constatado que os falantes que trabalhavam na cidade apresentavam traços linguísticos diferentes daqueles que nunca desciam o morro.

### 4 – Algumas considerações

Esta análise da comunicação nos permitiu depreender o quão profunda é a diversidade linguística presente na amostra selecionada, que se reflete em suas diversas construções discursivas. Construções que caracterizam um grupo/comunidade de fala, por meio das divergências na expressão comunicativa estabelecida entre falantes de dialetos distintos no processo de interação.

Portanto, um projeto de elaboração de um banco de fala é de fundamental importância para estudos posteriores e possibilitando a reflexão e análise de traços da fala de habitantes de diferentes cidades da região Norte Fluminense. Sua exatidão, característica necessária ao meio acadêmico, favorece ao estudo da variedade linguística e da oralidade, pois apresenta situações reais de uso da fala de forma dinâmica.

## 5. Anexos

## ANEXO A – Quadro com as normas para transcrição das falas

OCORRÊNCIAS	SINAIS	EXEMPLIFICAÇÃO
Incompreensão de palavras ou segmentos	( )	eu me amarro ficar ( ) olhando no espelho
Hipótese do que ouviu	(hipótese)	paramos (num) posto
Truncamento de sílaba e/ou quebra de sequência	/	não/ sabe que eu não tenho... é::... não foi nem muito intere/ não foi nem muito triste...
Qualquer tipo de pausa, substituindo todos os sinais específicos da língua escrita que desempenham tal função: ponto-e-vírgula, ponto final, dois pontos e vírgula.	...	não é o que era antigamente... onde a gente não... sabia de nada...
Interrogação	?	sabe o que que é?
Qualquer alongamento	::	ou então no:: congelador
Comentário do transcritor	((minúsculas))	((risos)) ((pigarro))
Discurso direto	“ ”	ela “vamos? eu tenho que ir a Petrópolis... você vai comigo? eu “tá bom... vamos...”
Superposição, simultaneidade de vozes	[texto]	I: meu tio também... ele faz quadros... [e isso...] E: [e como é que é?] que... você faz?
Números	por extenso	meia quatro dois... décimo quarto andar...
Nomes comuns estrangeiros	itálico	ah::... o <i>mousse</i> é super fácil...
Onomatopeia e siglas	caixa alta	no que eu me joguei pro lado... ela foi pro outro... eu PUFF... bati na árvore... se uma universidade do porte da PUC...
Nomes próprios	iniciais maiúsculas	mas... eu fui a Petrópolis com uma amiga...
Nomes de profissão, cursos em geral	minúsculas	desenho industrial, agronomia, engenharia etc.

Fonte: <<http://www.discursoeagramatica.lettras.ufrr.br/>>. Acesso em: 06 maio 2012.

Informações adicionais:

- Os exemplos usados são das entrevistas do *corpus* Discurso & Gramática.
- A fala do entrevistador é marcada com os mesmos critérios do informante:  
fala do entrevistador - E: e... você estuda:: que curso?  
fala do informante - I: direito... na Cândido Mendes...
- Não se indica o ponto de exclamação (frase exclamativa).
- Fáticos: ah, ahn, uhn, tá.
- Podem-se combinar sinais. Por exemplo: assim::... (alongamento e pausa), não porque::/ (alongamento e truncamento de sílaba e/ou quebra de sequência).

## 6 – Referências bibliográficas

- BORTONI, Stella Maris. 1985. A migração rural- urbana no Brasil: uma análisesociolingüística. In: TARALLO, Fernando (org.). *Fotografias sociolingüísticas*. São Paulo: Brasiliense.
- BRANDÃO, Silva Figueiredo. 1991. *A geografia lingüística no Brasil*. São Paulo: Ática.
- ELIA, Sílvio. 1979. *A unidade lingüística do Brasil*. Rio de Janeiro: Padrão.
- FERRARI, Lílian Vieira. 1994. *Variação lingüística e redes sociais no Morro dos Caboclos*. Rio de Janeiro: UFRJ. Tese de Doutorado.
- LABOV, W. 1995. *Principles of linguistic change. Vol. I: Internal factors*. Cambridge: Blackwell.
- MARTINET, André. 1993. *Funcion y dinámica de las lenguas*. Madrid: Gredos.
- NASCENTES, Antenor. 1953. *O linguajar carioca*. Rio de Janeiro: Organização Simões.
- NASCENTES, Antenor. 1958. *Bases para a elaboração do Atlas lingüístico do Brasil*. Vol I. Rio de Janeiro: Fundação Casa de Rui Barbosa.
- ORLANDI, E.P., GUIMARÃES, E. e TARALLO, F. 1989. *Vozes e contrastes: discurso na cidade e no campo*. São Paulo: Cortez.
- ROSSI, Nelson. 1963. *Atlas prévio do falares baianos*. Rio de Janeiro: INL.
- SILVA NETO, Serafim da. 1970. *História da língua portuguesa*. Rio de Janeiro: Livros de Portugal.
- ZÁGARI, Mário R. L. 1988. *Esboço de um Atlaslingüístico de Minas Gerais*. Rio de Janeiro: Fundação Casa de Rui Barbosa.