

RETROSPECTIVA

ESTUDOS SOCIOLINGÜÍSTICOS NO RIO DE JANEIRO

Giselle Machline de Oliveira e SILVA e Sebastião Josué VOTRE (Universidade Federal do Rio de Janeiro).

*ABSTRACT. The present paper offers a general view of the origins and development of the research group Census of Linguistics Variation in Rio de Janeiro. We present the projects that have been implemented as well as the papers, thesis and dissertations accomplished on the basis of the Census data.*

1. Introdução

Na década de 70 consolidaram-se dois grupos de pesquisa sobre o uso da língua na Universidade Federal do Rio de Janeiro: o grupo do projeto NURC, sob a coordenação do saudoso Professor Celso Cunha, e o grupo do Projeto CENSO, tendo como coordenador o professor Anthony Julius Naro. Na mesma época outros grupos de estudos sociolingüísticos estavam em atuação, dentre os quais citamos os coordenados por J. Heye, H. Gryner e D. Callou.

Desde então, tem sido múltipla e variada a produção de pesquisas sobre a fala do Rio e no Rio. Neste trabalho, todavia, vamos nos concentrar apenas na produção do grupo, aqui identificado como grupo CENSO.

2. O grupo CENSO

Em 1979, sob a coordenação do Professor Anthony Julius Naro, organizamo-nos sob a forma de um grupo de pesquisa lingüística de caráter interuniversitário, com base física no Departamento de Lingüística e Filologia

da Faculdade de Letras da UFRJ. Nosso interesse principal era estudar o uso da língua falada no Brasil, com ênfase para o município do Rio de Janeiro.

Nosso primeiro projeto financiado pela FINEP foi o Censo da Variação Lingüística no Rio de Janeiro, do qual derivou o rótulo do grupo. Desse projeto, resultou a gravação e transcrição de 48 horas de entrevista oral, com informantes adultos cariocas, procedentes de todas as áreas geográficas do município do Rio de Janeiro, estratificados por sexo, idade e escolarização. Armazenamos os dados em meios eletrônicos e procedemos à análise preliminar de alguns fenômenos variáveis da fala carioca, nos níveis da fonologia, da morfossintaxe e do discurso.

O projeto CENSO foi amplamente discutido, antes e durante sua execução, tanto dentro como fora do país.

Na linha de buscar subsídios financeiros práticos e voltados para a educação, oferecemos ao INEP o projeto Estruturas da Fala e Aquisição da Língua Padrão. Nesse projeto, concentramo-nos sobretudo na tarefa de dilatar a amostra CENSO com o registro da fala de 16 crianças, e analisar fenômenos variáveis que interessassem ao ensino da língua padrão culta, no 1º e 2º graus, com atenção para estratégias de ensino e de aprendizagem dessas mesmas variantes.

Concluído o projeto CENSO, encaminhamos à FINEP dois anos depois, o projeto intitulado Subsídios Socio lingüísticos do Projeto CENSO à Educação. Nesse projeto, o grupo trabalhou sobretudo no sentido de identificar, em cada fenômeno específico que estava sendo investigado, as implicações pedagógicas das descobertas sobre o funcionamento da língua falada.

Em 1987, submetemos à FINEP o projeto intitulado Mecanismos Funcionais do Uso Lingüístico, concluído em 1989. Apresentamos agora o nosso quarto projeto, já aprovado, mas com verba ainda não liberada, Uso da Língua na Comunicação (ULC). Seguem-se algumas características do projeto ULC.

### 3. Projeto atual

#### Problemas e necessidades

Nossas necessidades mais imediatas e palpáveis estão associadas aos projetos e estudos em desenvolvimento. Elas se relacionam com: a) a confirmação de hipóteses que levantamos em projetos anteriores, que testamos preliminarmente e que precisamos validar com novos dados e novos testes; b) a avaliação e teste de novas hipóteses, seja com os dados já disponíveis em nosso programa de pesquisas, seja em outros.

Em primeiro lugar, sentimos necessidade de continuar as análises que estamos desenvolvendo: alguns dos temas que ora investigamos merecem aprofundamento teórico e empírico, quer por sua importância intrínseca, quer por suas aplicações. Assim, o projeto ULC se caracteriza por ser, essencialmente, uma continuação das pesquisas já em andamento.

Outro aspecto que nos interessa na definição desse objetivo, ao empreender estudos sobre o uso da língua, é o de comparar os resultados por nós obtidos com os de outros pesquisadores de outras partes do país. Entendemos como problema grave a falta de compatibilidade entre os diversos corpora existentes no Brasil, sobretudo por falta de informação sobre os critérios adotados na coleta dos dados e na metodologia de sua análise e interpretação. Faltam condições ideais para a análise contrastiva das variações dialetais em diferentes regiões do país, o que impede a tarefa de avançar no traçado de um perfil sociolinguístico da comunidade de fala coloquial não-culta do Brasil. Assim, o Grupo de Trabalho de Sociolinguística da ANPOLL tem como um de seus objetivos fazer um censo dos corpora no Brasil.

De fato, um problema comum às pesquisas na área da linguagem é o da precariedade dos dados. O trabalho com dados linguísticos é muitas vezes dificultado pela natureza mesma do objeto de estudo: a produção de fala com

que tomamos contacto é um texto, mais ou menos longo, mais ou menos audível, quase transcrito, armazenado em meios precários, perecíveis e pouco confiáveis. Para sanar essa deficiência, e contribuir decisivamente na organização de uma base de dados eletrônicos, propomo-nos a armazenar dados de diferentes fontes faladas e escritas, de outras regiões do país, e otimizar os processos de acesso e pesquisa, para facilitar a consulta e a portabilidade desses dados. Num primeiro momento, prevemos que o acesso será privativo dos membros do grupo, mas assim que os arquivos estejam organizados, pretendemos abri-los a todos os pesquisadores, nacionais e estrangeiros, como fazemos atualmente com os dados do arquivo CENSO.

Sentimos especialmente necessidade de intercâmbio de dados, para ratificar nossos achados com dados de corpora similares de outras regiões do país. Mais especificamente, precisamos testar, em outras amostras, os resultados e achados com dados do Rio de Janeiro, em fenômenos do tipo: concordância verbal e nominal, nós VS a gente, regência de verbos específicos (como ir), uso de seu VS dele, presença de artigos frente a possessivo e patronímios, ordenação dos constituintes na oração, processos de redução das sonorantes, distribuição dos marcadores conversacionais, indeterminação e definitude.

Vemos também como problema empírico complexo a integração da análise multivariacional com os modelos de análise da conversação, devido às diferenças nos critérios para coleta e transcrição e, especialmente, aos métodos de tratamento dos dados. Pretendemos, nesta questão, contribuir positivamente, através da coleta de dados, de acordo com a abordagem interacional, para comparar os resultados com os da metodologia multivariacional. Interessa-nos descrever a correlação entre as variedades lingüísticas e múltiplas funções interacionais e discursivas que os elementos lingüísticos desempenham como parte de um sistema sócio-comunicativo.

Nosso objetivo nesse ponto é, pois, aumentar o acervo, de modo a permitir as análises contrastiva e não-contrastiva. A fim de atender a esta carência, propomo-nos a criar um banco interacional de dados, de dimensões modestas neste primeiro momento, com 20 informantes, gravados em distintos contextos comunicativos.

Também constatamos a necessidade de oferecer à comunidade científica e pedagógica uma concordância, e um índice completos do nosso corpus de fala. Estes dois documentos são o subproduto natural de corpora de textos armazenados eletronicamente. Concordância é um documento eletrônico que lista em ordem alfabética e freqüencial todas as ocorrências dos itens do corpus, e dá a freqüência absoluta e relativa de cada ocorrência. Seu mérito maior consiste em oferecer, para cada ocorrência, o contexto lingüístico imediato à esquerda e à direita, o que permite distinguir os homônimos, estudar as variantes semânticas, sintáticas e pragmáticas. O índice é um dicionário de freqüência que distribui os itens léxicos, com a freqüência de ocorrência de cada um, em três ordenações: ordem decrescente de freqüência, ordem alfabética convencional, ordem alfabética inversa (tomando como referência a(s) última(s) letra(s) de cada palavra).

Por fim, sentimos necessidade de testar o efeito da modalidade escrita nos fenômenos que estamos estudando, pois é comum que em nossas análises sobre dados da fala surjam perguntas do tipo: - Como se comporta a escrita formal e a informal? Como se passava o fenômeno na escrita do século passado? Como se manifestam as estratégias de embalagem da informação na fala e na escrita? Como se dá o desempenho de crianças e adolescentes, e de pessoas de outras idades, que estão em processo de aquisição da escrita?

Tendo então constatado que é problema grave o das diferenças estruturais entre fala e escrita, propomo-nos a criar um banco específico de dados da escrita informal, a começar por cartas pessoais de escritores consagrados.

#### 4. Organização Estrutural do Grupo

O grupo de estudos sobre o uso da língua constitui um organismo de pesquisa com a finalidade de concentrar esforços em torno de um interesse geral, o de desenvolver estudos sobre a língua a partir de dados reais, com vistas ao desenvolvimento e aprimoramento da pesquisa no campo da ciência lingüística e, em especial, em cumprimento das etapas previstas em ULC.

O grupo se organiza numa estrutura modular dinâmica, em setores que podem ser ativados simultaneamente ou em momentos diferentes, de acordo com a necessidade e as prioridades no cumprimento de objetivos previamente estabelecidos (a cada etapa do ULC como um todo e às etapas nele contidas). A coordenação se incumbem de congregar os diversos módulos e oferecer condições para o bom andamento das atividades de pesquisa. O módulo acadêmico concentra as atividades de pesquisa lingüística propriamente dita. O módulo de acervo de dados se ocupa da matéria-prima a ser analisada. Nele acha-se a amostra CENSO, constituída com o auxílio de diferentes acordos firmados entre FINEP/FUJB em 01 de outubro de 1981 e entre FINEP/FUJB em 15 de abril de 1983 e 10 de dezembro de 1984. Aqui estão previstas também a ampliação dos dados já existentes e a anexação de outras amostras como o banco interacional, o banco da escrita e os bancos de intercâmbios. O módulo computacional compreende atividades relacionadas à computação, contribuindo para o desenvolvimento da área interdisciplinar da lingüística computacional. Aqui se incluem tarefas como armazenamento e recuperação automática, bem como digitação e pesquisa em bancos de dados.

O módulo de formação e aperfeiçoamento visa a atender às necessidades dos pesquisadores em geral. Assim, promove o estímulo à formação de novos pesquisadores, treinando alunos de graduação e pós-graduação para cumprirem tarefas de auxiliar técnico de pesquisa, com

vistas à sua eventual atuação como pesquisadores futuros. Esse módulo ocupa-se também com a atualização e o aperfeiçoamento da equipe científica, de modo a assegurar o bom nível de qualidade dos trabalhos. Estimula nossos pesquisadores a passarem temporada no exterior e a partilharem, na volta, suas experiências com o grupo.

O projeto ULC é original em sua forma de conduzir os trabalhos de pesquisa: todos os pesquisadores participam ativamente das decisões administrativas e acadêmicas, mas cada pesquisador é responsável autônomo por um ou mais projetos individuais, que leva adiante com a ajuda dos auxiliares de pesquisa e com a colaboração regular e sistemática dos colegas, uma vez que o grupo se reúne quinzenalmente com a função de avaliar e discutir o andamento dos projetos individuais. Cada projeto individual apresenta um ponto na grande tapeçaria da realidade sociolinguística da fala do Rio de Janeiro.

## 5. Orientação Teórica e Metodológica

No primeiro momento da pesquisa, os integrantes do grupo CENSO partiram de uma orientação teórica e metodológica relativamente homogênea, inspirada nos trabalhos de William Labov e Gillian Sankoff, e nesse quadro desenvolveram seus primeiros projetos.

Motivados por difentes desafios teóricos e metodológicos, vêm procurando novas alternativas, inspiradas, sobretudo, nos trabalhos de funcionalistas como W.Chafe, Paul Hopper, S.Thompson, T.Givón e R. Tomlin; e nos analistas da conversação, como Sacks & Schegeloff, e nos adeptos da difusão lexical, como Wang & Scheng.

O grupo todo tem a tendência de buscar fora da língua propriamente dita e no contexto de seu uso os fatores que motivam e condicionam a sua estrutura.

Consolidou-se no grupo um método de pesquisa de

tipo retrodutivo (que engloba traços indutivos e dedutivos), e que representamos por uma espiral contínua. Com efeito, movimentamo-nos num vaivém constante entre o que nos oferecem os modelos teóricos e o que encontramos nos dados. Somos, então, a todo momento, solicitados - quer pelos dados, quer pela teoria - a postular e testar novas hipóteses, com novas variáveis. Uma vez confirmadas, as hipóteses passam a figurar no modelo teórico, que é sempre provisório, mas cumulativo. Verificamos, por exemplo, que em vários fenômenos em estudo, onde imaginávamos que havia variação, o que houve foi uma distribuição: diferentes formas cumprem diferentes funções, em diferentes contextos.

Dada a natureza do grupo e a diversidade de interesses por correntes distintas de análise, não faz sentido falar em uma metodologia única nos seus projetos. Antes, devemos falar em metodologias convergentes, que convivem com as tarefas de propor e testar fatores (estruturais, sociais, pragmático-discursivos e outros) que possam vir a ser relevantes para a interpretação da variação e mudança, em diferentes níveis de manifestação, no português brasileiro. Em seu aparente sincretismo, as tendências que utilizamos revelam influência marcante da Teoria da Variação, da Linguística Funcional e da Análise Conversacional.

A metodologia da teoria da variação é o fio condutor principal da maioria dos projetos individuais. Diferentes aspectos dessa metodologia vêm sendo testados e consolidados, desde 1975, por integrantes do grupo, em estudos individuais, dissertações de mestrado e teses de doutorado, bem como nos três projetos anteriores.

As linhas gerais da Metodologia da Variação por nós utilizadas supõem a operacionalização e o desdobramento de conceitos e procedimentos de trabalho relacionados à sistematicidade da variação linguística nas comunidades de fala. Essa sistematicidade pode ou não estar associada a mudanças rápidas ou lentas. A teoria postula que uma regra é variável quando representa o efeito, não



categórico, da coatuação simultânea de fatores estruturais (fonético-fonológicos, morfossintáticos, semânticos), pragmáticos e sociais. As hipóteses supõem tanto a autonomia da atuação de cada fator, como a possibilidade real de interação entre fatores de efeitos contrários ou divergentes. Na prática, codificamos os dados como se cada fator tivesse efeito autônomo, e procedemos à análise atentos a possíveis contextos de interação.

A metodologia atribui pesos idênticos a fatores estruturais e sociais, mas os programas analíticos ponderam cada grupo de fatores em termos de seu efeito relativo na aplicação de cada regra.

Entre os fatores sociais, nosso grupo está dando atenção especial para os parâmetros controlados na coleta de dados da amostra CENSO (idade, sexo e escolarização), pretendemos continuar testando o efeito de variáveis novas na área da sociolinguística (mercado ocupacional, exposição aos meios eletrônicos e sensibilidade linguística), e tentaremos encontrar meios discretos de codificar a variável etnia.

Entre os fatores estruturais, optamos conscientemente pela procura de condicionamentos em todos os níveis linguísticos: na fonética e na fonologia, na morfossintaxe, no léxico (em classes lexicais ou em itens léxicos isolados, nos casos de difusão lexical) e no texto.

Utilizamos um aparato eletrônico suficientemente testado e validado em outras áreas de investigação sobre a coatuação de fatores de distinta natureza. O instrumento principal é o programa VARBRUL, desenvolvido por David Sankoff na Universidade de Montreal. Nossa versão do programa está acoplada ao programa SWAMINC, desenvolvido por Anthony Naro. Projetado para realizar as tarefas preliminares, ele analisa, retira dados categóricos, faz amalgamações ou separações. Desenvolvemos um sistema específico de manipulação dos dados, combinando os dois programas, a que denominamos SWAVA (Votre e Naro, 1980). Além disso, contamos com novos programas computacionais da série VARBRUL, em versões para micro tipo

PC, McINTOSCH e VAX. Mais sofisticados, eles permitem operar com variáveis enéarias (fenômeno que não tem duas variantes apenas (como singular e plural), mas várias, como por exemplo: pronomes de tratamento (você, tu, senhor).

Mais poderosos, eles permitem: 1) operar com variáveis binárias e enéarias; 2) redefinir variáveis dependentes no decorrer do trabalho, sem novo formato de codificação e 3) calcular pesos relativos com número diferente de dados pelas diversas variáveis interdependentes.

A proposta de uma metodologia de orientação discursiva, ou funcional, em desenvolvimento por membros desse grupo de pesquisa, é um desdobramento da nossa preocupação em postular, testar e comprovar o efeito de fatores pragmático-discursivos, ao lado dos fatores estruturais e sociais propriamente ditos. Entre esses fatores, prestamos atenção especial para o "status" do referente, para o grau de transitividade da oração, sem grau de sequencialidade e para o nível de gramaticalização do fenômeno.

Estamos reelaborando de forma sistemática as categorias referentes ao "status" de informação do referente. A oposição velho-novo foi acrescida das categorias inferível, disponível, novo parcial, único, irrelevante, e há determinados fenômenos que parecem requerer novos refinamentos.

A variável grau de transitividade da oração, proposta por Hopper & Thompson (1980), envolvia inicialmente os traços seguintes: número de participantes, cínese, aspecto do verbo, punctualidade do verbo, volicionalidade, afirmação do predicado, modo, agentividade do sujeito, e sua individuação. Estamos reavaliando os traços em termos ponderados, e acrescentando novos traços ao complexo; é o caso da oposição tético VS categórico, de Sasse (1985), que distingue entre orações inanalísáveis, ou téticas, que enunciam um evento holístico, e orações categóricas, binárias, em que o predicado dá informações a respeito de uma base de predicação ou sujeito/tópico.

A categoria seqüencialidade está estreitamente vinculada à oposição figura VS fundo, da psicologia, e presente nos estudos sobre teoria literária. Figura cor responde aos eventos seqüenciais, enquanto o fundo dá conta de comentários, ilações, detalhamentos simultâneos aos eventos-figura. Nossa contribuição, além de testar a variável nos dados do português, está consistindo em testar novas propostas, não mais binárias, e sim eneárias, em que o continuum figura-fundo se possa manifestar; outra contribuição consiste em ver o efeito da oposição em outros gêneros discursivos, e não apenas na narrativa, em que os estudos estão mais avançados.

Nível de gramaticalização de um fenômeno é um item difícil de avaliar, porque muitas vezes não sabemos se o fenômeno está em processo de mudança, nem qual é a direção da mudança. Mesmo com essa limitação, temos alguns indícios de que é possível estabelecer estágios de gramaticalização para alguns fenômenos de natureza sintática, como ordenação vocabular, e distribuição das relações de causa e consequência, em que coatuam de forma mais ou menos transparentes as dimensões iconicidade e economia da construção.

Em relação à análise conversacional, iniciamos a coleta de uma amostra de fala que incorpora diversas situações de encontros assimétricos (discursos narrativos, descritivos, conversas formais e informais, relatos de experiências, casos e discursos institucionalizados). Prestamos atenção especial à transcrição dos dados, tentando desenvolver um sistema notacional apropriado para o registro de padrões entonacionais.

Sintetizando nossa atividade de pesquisa, comportamo-nos como segue: utilizamos o discurso como fator na análise; procuramos fazer uma análise no discurso (e não do discurso), partindo sempre do uso. Por fim, admitimos que somente fora da estrutura linguística podemos buscar explicação para o modo de ser de cada fenômeno.

Postulamos uma estratégia de investigação de indicadores de mudança lingüística, que leve em conta a estrutura da língua, em covariação com a faixa etária dos membros da comunidade. Antes de afirmar que determinado fenômeno está variando e mudando, temos que investigar cuidadosamente a distribuição deste fenômeno em grupos de distintas faixas etárias, que convivem simultaneamente na mesma comunidade lingüística, e acompanhar o desenvolvimento "diacrônico" desse fenômeno pelo recontacto com cada grupo testado, ou pelo contacto com novos grupos da mesma idade. Além disso, seria conveniente encontrar indicadores externos aos grupos analisados, como testemunhos de gramáticos, da literatura, e de outros documentos escritos e gravados.

## 6. Alguns Resultados

Até agora os resultados coletivos dos estudos do grupo dizem mais respeito às variáveis sociais, não por privilegiarmos essas variáveis, mas porque se prestam mais a comparações e generalizações.

Verificamos que a Escolaridade é a mais atuante entre as variáveis sociais, ela é a que influencia o maior número de fenômenos e sempre no mesmo sentido. Poderíamos afirmar que se um fenômeno é socialmente condicionado, atua a escolaridade, até mesmo quando a forma não parece ser alvo de ensino ou correção. Mais importante, entretanto, é termos verificado que os fenômenos se correlacionam com a escola de modo diverso. Assim, por exemplo, podem ocorrer casos em que há falantes que entram na escola usando com grande frequência a forma padrão, enquanto outros não a usam, mas a escola põe o uso não-padrão (ou põe talvez a própria criança). Nesse esquema se enquadram fenômenos como a Regra de Concordância Nominal (Scherre), a de regência do verbo ir, de movimento (Mollica), ambos com formas estigmatizadas pela escola. Podem ocorrer casos em que a maioria dos falantes entra na escola sem usar a variante padrão que

será adquirida durante a escolaridade sem que desapareça todavia a variante não-padrão. Contam-se nesse caso os fenômenos de aquisição das variantes nós e seu, contrastando respectivamente com a gente e dele. Esses últimos são praticamente as únicas formas a existirem no "primário", mas não são erradicadas pela escolaridade. De fato, as variantes a gente e dele não chegam a ser estigmatizadas.

Embora o sexo tenha se mostrado pouco atuante de per si, há uma tendência em nossos resultados a que mulheres mais velhas favoreçam mais a forma padrão do que homens da mesma faixa, neutralizando-se essa diferença entre os mais jovens. Também notou-se uma influência maior da escola sobre mulheres do que sobre homens.

A idade atuou em três dos sete fenômenos estudados e, nos três, no sentido dos mais velhos prestigiarem a forma culta.

Quanto às variáveis sociais experimentais, o mercado ocupacional mostrou um bom desempenho, embora menos espetacular do que em Montreal. Evidenciou-se que os homens tendem a ser mais sensíveis a essa pressão do que as mulheres, ela atua, principalmente, na faixa etária de 15 a 50 anos, época em que os falantes são mais influenciados pelas experiências do seu tipo de trabalho. Naqueles fenômenos cuja variante não-padrão não é estigmatizada pela escola, a ação dessa variável mostrou-se fraca ou até nula. O contato com a mídia (e particularmente a televisão; único meio de comunicação a chegar de fato a nossos falantes não-universitários) mostrou-se levemente atuante. Por exemplo, os falantes mais expostos à televisão mostram tendência a usarem a variante padrão, naqueles fenômenos em que os atores de novela a utilizam.

## 7. Outras atividades

Temos oferecido cursos relacionados às linhas teóricas e metodológicas do Projeto nas Universidades Federais da Paraíba, do Piauí, do Ceará, de Sergipe e nas Universidades Inglesas de New Castle, upon Tyne e Essex.

Conjuntamente com o grupo de trabalho de sociolinguística da ANPOLL, promovemos o primeiro Seminário Regional de Estudos sobre o Uso da Língua, na Faculdade de Letras da UFRJ, nos dias 10 e 11 de novembro de 1988. O Seminário permitiu intenso intercâmbio de experiências em pesquisa empírica e de propostas teóricas e inovações metodológicas em Sociolinguística. Dele participaram pesquisadores de orientações diferentes de outras regiões do país, tais como: Pernambuco, Brasília, São Paulo, Minas Gerais, Santa Catarina e Rio Grande do Sul. Houve também representação do Uruguai.

Em todos os projetos temos como meta prioritária atuar decisivamente na formação de novos recursos humanos para a pesquisa. Desde o início, tentamos atribuir aos estagiários funções de pesquisa, a par de funções meramente rotineiras. Envolvermos graduados e alunos de graduação e pós-graduação em atividades de coleta, transcrição, codificação e análise dos dados. Temos atraído novos investigadores com auxílios da FINEP, do CNPq, da FAPERJ e da própria Universidade.

Dentre os pesquisadores do Projeto, as Professoras Maria Marta Pereira Scherre e Maria Cecília de Magalhães Mollica defenderam teses de doutorado: sobre Reanálise da Concordância Nominal, a primeira; a segunda, utilizando dados por nós coletados. Maria Conceição Paiva e Nelize Pires de Omena estão desenvolvendo suas teses de doutorado, com os mesmos dados. Eles também serviram de fonte para a mestranda Mônica Rio Nobre elaborar sua dissertação sobre a ordem adjetivo-nome. Nos últimos anos, temos verificado um interesse consistente, e crescente, da parte de mestrandos e doutorandos, em desenvolver suas pesquisas a partir do nosso banco de dados da fala.

Os membros do grupo vêm prestando assessoria, na qualidade de consultores, a vários grupos de pesquisa de diferentes regiões do país: UFC, UFSE, UFSC, PUCRS, UFPB/Campina Grande, UFPR e UFRS.

## 8. Trabalhos Pessoais

Como todo conjunto, somos compostos de elementos que não abdicam de sua individualidade e escolhem, com total liberdade, suas linhas de pesquisa. A seguir, resumimos os projetos individuais para os próximos dois anos.

Anthony Naro e Sebastião Votre estão trabalhando na análise da ordenação vocabular do sujeito e verbo em narrativas do corpus Censo, do corpus Dialetos Sociais Cearenses, do Português Fundamental de Portugal e da Literatura Brasileira. Vão estudar também a ordenação vocabular de constituintes adverbiais, quer na forma nominal, quer na forma de orações, com atenção especial tanto para a ordenação quanto para as combinações que se verificam entre esses constituintes e as orações a que se agregam.

Nelize Pires de Omena vai estudar a indeterminação lingüística na variação sintática. O aspecto do uso da língua a que se propõe estudar - o da indeterminação - constitui um projeto de pesquisa a longo prazo. Presentemente, o que está realizando é o estudo das estratégias sintáticas de indeterminação do referente, que é agente da ação verbal, sob uma perspectiva diacrônica. Serão as seguintes as formas observadas: a) verbos na terceira pessoa do plural, sem sujeito expresso; b) verbos na terceira pessoa do singular, sem sujeito expresso; c) verbos na terceira pessoa do singular, acompanhados do pronome se, sem sujeito expresso; d) a passiva sintética; e) a passiva analítica, sem agente expresso; f) as formas nominais do verbo sem sujeito ou agente expresso.

Utilizando a metodologia e técnicas da teoria da variação, Omena objetiva: a) analisar a gênese das formas b) e c), citadas no parágrafo acima; b) observar a distribuição das formas concorrentes, através dos séculos; c) explicitar as condições linguísticas favoráveis a cada uma das formas.

Vera Lúcia Paredes Pereira da Silva estuda dos aspectos sintáticos e discursivos da correspondência de escritores brasileiros contemporâneos, compreendendo: a variação do sujeito (para confronto com os resultados de Paredes Silva, 1988) e a variação no objeto indireto de 2ª pessoa; a estruturação interna de cartas pessoais, em termos de uso de parágrafos e delimitação de tópicos discursivos; e a classificação de tipos de cartas.

Maria Marta Pereira Scherre e Anthony Julius Naro estudam o paralelismo formal na concordância de número no português popular do Brasil. No presente momento, estão desenvolvendo uma pesquisa que tem como objetivo principal discutir a influência do paralelismo formal na concordância de número no português popular do Brasil. Pretendem: a) definir com mais precisão a caracterização da variável paralelismo formal; b) discutir implicações teóricas advindas da análise da variável paralelismo formal, na sua ligação estreita com as "ocorrências sucessivas de uma mesma variável dependente"; c) discutir a direcionalidade da influência da variável paralelismo formal; d) discutir a relação da variável paralelismo formal com outras variáveis, principalmente com relação a variáveis que possam estar em competição com ela; e) e ver o comportamento de variáveis linguísticas e sociolinguísticas com dados seriados, ou seja, com dados que envolvem diversas ocorrências sucessivas de uma mesma variável dependente e com dados isolados.

Estão desenvolvendo o trabalho com base nos dados do projeto Censo. Pretendem também trabalhar com dados do Ceará. Suas reflexões vão envolver a concordância



de número no sintagma verbal, no sintagma nominal, nos predicativos e nos participios passivos.

Giselle Machline de Oliveira e Silva está aperfeiçoando seus trabalhos sobre definitude, seja através do estudo de artigos frente a possessivos e patronímicos, seja da variação entre seu e dele, seja ainda através da quantidade de artigos em receitas culinárias, além de continuar investindo na análise discursiva através dos marcadores conversacionais.

Alzira Tavares de Macedo, atualmente no Canadá, onde está estagiando com David Sankoff, está trabalhando em marcadores conversacionais, além de se dedicar a outros fenômenos discursivos, tais como as diferentes estratégias apresentadas pelos discursos direto e indireto.

Cecília Mollica vem atuando em três flancos que podemos assim classificar: a) rastreamento de dados tendo como foco a questão da mudança lingüística; b) Sociolingüística descritiva, salientando a variação e seus aspectos envolvidos; c) Sociolingüística aplicada, revelando as possíveis contribuições da área para o ensino da língua materna.

Continuando sua pesquisa sobre dequeísmo, um trabalho extensivo de busca de dados está sendo feito em diferentes fontes sonoras, para cobrir as primeiras décadas do nosso século, assim como em inúmeros textos de língua escrita, para abranger outros séculos. Tem como objetivo identificar a entrada de estruturas dequeístas no português, a exemplo do que foi realizado pelo espanhol.

Paralelamente, vem desenvolvendo pesquisas nas linhas variacionista, difusionista e de atitude a respeito de alguns fenômenos fonéticos e sintáticos variáveis do português moderno, tendo como base a amostra Censo. Os processos de prótese e aférese, ou de assimilação, encontrados no português moderno falado, são foco de análise, da mesma forma variantes alternativas em estrutura comparada.

A terceira frente de investigação volta a atenção para o diagnóstico de fatores responsáveis pelo controle da aquisição da língua padrão no processo escolar. Formas mais estigmatizadas na língua falada e erros mais salientes na língua escrita estão sendo escolhidos para a análise, de modo a medir-se o peso de variáveis do ponto de vista do aluno, do professor e do exemplo de aula no ajuste à norma.

Sebastião Votre está investigando os processos de regularização morfo-fonológica na fala, com atenção especial para as duplas nominais menos-menas, todo-toda, meio-meia, pouco-pouca, e para as duplas verbais seja-seje, esteja-esteje. Está sendo desenvolvido um teste de atitude para verificar, com informantes de diferentes níveis de escolaridade, o grau de consciência da variabilidade das duplas, e a opção de cada falante por uma das formas, bem como a justificativa que o informante oferece para a escolha efetuada.

Maria da Conceição Paiva examina dois aspectos relativos à estruturação da relação causa-consequência: a ordenação das causas e a presença ou ausência de conectores interligando-as. Procura mostrar que os dois processos são relacionados, e analisa os condicionantes semânticos e discursivos que atuam sobre ambos os processos.

Maria Luíza Braga investiga sentenças clivadas no português coloquial oral do Rio de Janeiro. São identificados 6 possíveis tipos de sentenças clivadas. Estas construções são analisadas sob uma perspectiva discursiva e interacional. Mostra a contribuição das mesmas na estruturação do tópico discursivo, bem como o papel que desempenham na interação que se desenvolve entre os interlocutores. Evidencia também o papel decisivo dessas sentenças na transmissão de contraste.

Cláudia Nívia Roncarati de Souza com Maria Luíza Braga vem desenvolvendo um estudo sobre as estruturas paralelas de lista e de repetição no discurso semi-coloquial oral do Rio de Janeiro, objetivando determinar o papel que estas estruturas-descritas quanto ao padrão

entonacional- desempenham quer na estrutura de tópicos discursivos, quer na interação que se constitui entre os interlocutores.

Dentro da filosofia do Projeto ULC de integrar a análise multivariacional com modelos de orientação discursiva, Roncarati vem coordenando a implementação de um Banco de Dados Conversacionais, que tem por objetivo favorecer: (1) a explanação de fenômenos morfo-sintáticos e fonológicos coletados em situações propiciadoras da fala natural, em que os usuários da língua estejam envolvidos, em diferentes encontros assimétricos e simétricos; (2) a descrição mais acurada das variedades lingüísticas quanto às funções interacionais e discursivas que essas variedades podem desempenhar como parte de um sistema sócio-comunicativo e (3) a comparação das análises feitas em dados da Amostra Censo.

## 9. Conclusão

Na forma como se concebe, o Projeto ULC é relevante no contexto da sociolingüística brasileira, porque o grupo de pesquisa que o apresenta assumiu uma posição de liderança na proposição de novas metodologias de análise do uso da língua, disseminou essas propostas por outros centros de pesquisa, e agora precisa comparar, com outros corpora e outras abordagens, que os modelos são adequados à análise diastrática e diatópica do português do Brasil. O Projeto se justifica pelo trabalho que presta na análise da variabilidade que se manifesta na fala do Brasil.

O Projeto ULC é relevante também no sentido de permitir um conhecimento mais realista e objetivo da variação geográfica, por comparar dados de corpora de distintas regiões do país, coletados e armazenados segundo a metodologia que desenvolvemos nos projetos anteriores, sobretudo no Projeto Censo: é o caso, inicialmente, dos projetos Dialetos Sociais Cearenses, e Fronteira Brasil-Argentina. Com efeito, o grupo do Rio implementou nova

política de intercâmbio de dados, que pretende consolidar no correr de ULC, com vistas a ter mostras de todos os bancos que foram organizados com orientação sociolinguística.

O Projeto é relevante, por fim, pelo seu papel como centro de treinamento e aperfeiçoamento de pesquisadores, porque continuará envolvendo pessoas de diferentes níveis de formação nas tarefas de coleta, desenvolvimento de sistemas de acesso e análise automática, transcrição e correção de dados, bem como na análise propriamente dita.

Finalmente, notamos que este texto é coletivo, no sentido seguinte: cada parágrafo aqui é de alguma forma intersubjetivo, de tanto que compartilhamos, com os demais membros do grupo, as idéias e as palavras que selecionamos, frutos de numerosos relatórios e artigos longamente discutidos entre nós. Essa coesão no grupo não foi planejada e nem é planejável, nasceu e foi paulatinamente alimentada principalmente pelo respeito à enorme diversidade de seus elementos, diversidade essa que talvez seja a maior característica do grupo.

Embora tendo certeza de que a força de um projeto não provém das verbas recebidas, mas sim do grupo humano que o constitui, não podemos dar a receita de como compor um bom grupo. Seria o mesmo que querer explicar um bom casamento. Só sabemos que tentativas de elaborar grupos de cima para baixo, através da chefia do departamento, por exemplo, não são bem sucedidas. Há, porém, um anverso da medalha que é preciso mostrar e que lutamos para remediar: nosso principal defeito provém também de nossa coesão. Queremos nos referir à falta de publicação dos resultados do grupo. Por um lado a gana de publicar esfria após tanta discussão interna. Por outro lado, a necessidade de discutir, que às vezes se torna um vício, não é apropriada para elaborar um livro, quando todos querem dar palpites que embora válidos são muitas vezes incompatíveis com estilos e modos de pensar divergentes.

Assim mesmo, pensando os numerosos prós e contras, ainda vale a pena.