

Edenize Ponzo Peres

**O Uso de *Você, Ocê e Cê* em Belo Horizonte:
Um Estudo em Tempo Aparente e em Tempo Real**

Tese apresentada ao Programa de Pós-Graduação em Letras: Estudos Lingüísticos da Faculdade de Letras da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do Título de Doutor em Letras: Lingüística.

Área de Concentração: Lingüística

Linha de Pesquisa: Estudo da Variação e Mudança
Lingüística

Orientador: Prof^a Dr^a Jânia Martins Ramos

Co-orientador: Prof. Dr. José Olímpio de Magalhães

Belo Horizonte
Faculdade de Letras

2006

**Tese aprovada em 27/03/2006 pela Banca Examinadora
constituída pelos Professores Doutores**

**Jânia Martins Ramos – UFMG
Orientadora**

**José Olímpio de Magalhães – UFMG
Co-orientador**

Lilian Coutinho Iacovenco – UFES

Maria Eugênia Lamoglia Duarte – UFRJ

Eunice Maria das Dores Nicolau – UFMG

Lorenzo Vitral – UFMG

A meus pais, que, onde quer que se encontrem,
ficarão felizes com a conclusão deste trabalho.

Agradeço a todos aqueles que, direta ou indiretamente, contribuíram para a realização e conclusão desta Tese, especialmente,

- a **Alberto**, pela inestimável colaboração nas diversas fases do Doutorado;
- a **Ana Paula Huback**, amiga muito querida, pelo auxílio na obtenção de bibliografia, pela concessão de parte do *corpus* de 2002 e, sobretudo, pelo incentivo constante;
- a **Camila Tavares Leite**, pela conversão dos dados em CD e pela colaboração nos testes espectrográficos;
- à Prof^a Dr^a **Célia Regina dos Santos Lopes**, da UFRJ, pela gentileza do envio de seus textos para a complementação desta pesquisa;
- à Prof^a Dr^a **Evelyne Dogliani Madureira**, pelo constante interesse, carinho e disposição em ajudar;
- a **Gianni e Paula**, pelo grande auxílio com respeito ao Programa Varbrul;
- aos **informantes** dos dois *corpora*, que, mesmo sem saber, tornaram possível esta pesquisa;
- à Prof^a Dr^a **Jânia Martins Ramos**, minha orientadora, pela leitura atenta das versões deste trabalho, seus comentários e sugestões;
- ao Prof. Dr. **José Olímpio de Magalhães**, meu co-orientador, pela orientação amigável e, principalmente, pelo apoio, incentivo e encorajamento nas horas difíceis;
- às Prof^{as} Dr^{as} **Lilian Coutinho Yacovenco** e **Maria da Penha Pereira Lins**, e ao Prof. **Marcelo Lopes**, da UFES, pelo inestimável auxílio – direto e indireto – a este estudo;
- à Prof^a Dr^a **Maria do Carmo Viegas**, pela concessão do *corpus* de 1982 e pelas sugestões, quando do Exame de Qualificação;
- aos **professores** da Faculdade de Letras/UFMG, pelas lições teóricas e também de vida, e aos **ex-professores** dos antigos Departamentos de Lingüística e de Letras Vernáculas/UFMG, que, sem imaginar, me ensinaram tanto;

e, finalmente, agradeço à **Silvinha** a compreensão por todas as vezes em que, por causa da tese da mãe, era obrigada a ficar em casa.

RESUMO

O uso de *você*, *ocê* e *cê* em Belo Horizonte: um estudo em tempo aparente e em tempo real

Esta pesquisa procura investigar o uso das formas *você*, *ocê* e *cê* em Belo Horizonte, em duas épocas distintas: 1982 e 2002. Para tanto, foram analisados dados de fala, obtidos através de entrevista sociolingüística, coletados nas duas épocas, levando-se em conta seis fatores lingüísticos – função sintática das formas nas frases; contigüidade em relação ao verbo; comportamento nas construções de tópico; expressão da referência; tipo de frase em que aparecem e foco – e três fatores extralingüísticos – idade (três e cinco faixas etárias, respectivamente, nos *corpora* de 1982 e de 2002), gênero e classe social (média e baixa) dos informantes. Assim, o objetivo central desta tese é fazer-se um estudo sincrônico e diacrônico das três formas, descrevendo seu comportamento tanto no nível gramatical quanto no nível social, de acordo com a Teoria da Variação.

Dos nove fatores internos e externos analisados, nem todos se mostraram significativos nas duas amostras, como a contigüidade, o tipo de frase e o tópico, enquanto que os fatores função sintática e idade mostraram-se significativos em ambos os *corpora*. Quanto aos demais fatores, mostraram-se significativos em apenas um *corpus*. Os principais resultados encontrados indicam que está havendo mudança em progresso com relação à forma *cê* e também uma especialização das formas com respeito à expressão da referência, sendo *você* preferencialmente usado nas referências indefinidas e *cê*, nas referências definidas. Por outro lado, os contextos marcados, como as funções de objeto de verbo e de preposição e o foco, apresentam restrições ao uso de *cê*, embora não impeçam sua ocorrência, o que pode representar um lento início de mudança em curso.

ABSTRACT

The use of *você*, *ocê* and *cê* in Belo Horizonte: A study in apparent and real time

This study seeks to investigate the use of the forms *você*, *ocê* and *cê* in Belo Horizonte during two distinct time periods: 1982 and 2002. Speech data obtained from sociolinguistic interviews conducted in the two different time periods were analyzed, taking six linguistic factors into consideration – the syntactic function of each form within phrases, contiguity in relation to the verb, behavior in the construction of topic, the expression of reference, the type of phrase in which the item appears, and focus – as well as three extralinguistic factors – age (three and five age groups in the 1982 and 2002 corpora respectively), gender and social class (middle and low) of the informants. The central objective of this thesis is to conduct a synchronous and diachronic study of the three forms, describing their behavior at the grammatical as well as social level in accordance with Variation Theory.

Of the nine internal and external factors analyzed, some were significant in neither of the two samples, such as contiguity, the type of phrase and the topic, whereas the factors of syntactic function and age were shown to be significant in both of the corpora. As for the other factors, they were shown to be significant in only one corpus. The principal results obtained indicate that a change is occurring in relation to the form *cê* and there is also a specialization of the forms with respect to the expression of reference, with *você* being preferable for indefinite references and *cê* preferable for definite references. On the other hand, marked contexts, such as the object of the verb, the object of the preposition, and focus, presented restrictions on the use of *cê*, while not impeding its use, which could represent a slow beginning of a change.

LISTA DE TABELAS

TABELA 1	As formas <i>você</i> , <i>ocê</i> e <i>cê</i> no processo de indeterminação	11
TABELA 2	Uso das formas de tratamento ao rei (em %)	87
TABELA 3	Uso das formas pronominais e de tratamento nas relações hierárquicas entre os personagens das peças teatrais no PB e no PE	92
TABELA 4	Distribuição das formas de tratamento no PB – 1833 – 2 ^a metade do século XX (em %)	94
TABELA 5	Distribuição das formas de tratamento mais freqüentes em português (Séc. XVIII – XIX)	95
TABELA 6	Distribuição das formas pronominais e de tratamento em função do destinatário da carta	95
TABELA 7	Total de ocorrências de <i>você</i> , <i>ocê</i> e <i>cê</i> – <i>corpus</i> de 2002	117
TABELA 8	Distribuição de <i>você</i> , <i>ocê</i> e <i>cê</i> – 2002 – Fatores não-significativos	118
TABELA 9	Distribuição de <i>você</i> , <i>ocê</i> e <i>cê</i> , conforme a Função Sintática – 2002	119
TABELA 10	Distribuição de <i>você</i> , <i>ocê</i> e <i>cê</i> , conforme a Contigüidade – 2002	125
TABELA 11	Distribuição de <i>você</i> , <i>ocê</i> e <i>cê</i> , conforme a Referência – Todas as funções – 2002	131
TABELA 12	Distribuição de <i>você</i> , <i>ocê</i> e <i>cê</i> , conforme a Referência – Função de Sujeito – 2002	132
TABELA 13	Distribuição de <i>você</i> , <i>ocê</i> e <i>cê</i> , conforme a Referência e a Idade – Função de Sujeito – 2002	134
TABELA 14	Distribuição de <i>você</i> , <i>ocê</i> e <i>cê</i> , conforme o Foco – 2002	136
TABELA 15	Distribuição de <i>você</i> , <i>ocê</i> e <i>cê</i> , conforme a Idade – 2002	140
TABELA 16	Distribuição de <i>você</i> , <i>ocê</i> e <i>cê</i> , conforme o Gênero – 2002	143
TABELA 17	Distribuição de <i>você</i> , <i>ocê</i> e <i>cê</i> , conforme o Gênero e a Idade	145

TABELA 18	Distribuição de <i>você</i> , <i>ocê</i> e <i>cê</i> , para crianças e adolescentes – Testes	149
TABELA 19	Distribuição de <i>você</i> , <i>ocê</i> e <i>cê</i> – Entrevistas	149
TABELA 20	Distribuição de <i>você</i> , <i>ocê</i> e <i>cê</i> – <i>Corpus</i> de 2002 – Faixas etárias III, IV e V	152
TABELA 21	Total de ocorrências de <i>você</i> , <i>ocê</i> e <i>cê</i> – <i>Corpus</i> de 1982	156
TABELA 22	Distribuição de <i>você</i> , <i>ocê</i> e <i>cê</i> – Fatores não-significativos – 1982	157
TABELA 23	Distribuição de <i>você</i> , <i>ocê</i> e <i>cê</i> , conforme a Função Sintática – 1982	158
TABELA 24	Distribuição de <i>você</i> , <i>ocê</i> e <i>cê</i> , conforme a Idade – 1982	162
TABELA 25	Distribuição de <i>você</i> , <i>ocê</i> e <i>cê</i> , conforme a Classe Social – 1982	164
TABELA 26	Total de ocorrências de <i>você</i> , <i>ocê</i> e <i>cê</i> – <i>Corpus</i> de 2002 – Faixas III, IV e V	185
TABELA 27	Distribuição de <i>você</i> , <i>ocê</i> e <i>cê</i> , conforme a Função Sintática – 2002 – Faixas III, IV e V	187
TABELA 28	Distribuição de <i>você</i> , <i>ocê</i> e <i>cê</i> , conforme a Idade – 2002 – Faixas III, IV e V	190
TABELA 29	Distribuição de <i>você</i> , <i>ocê</i> e <i>cê</i> , conforme a Classe social – 2002 – Faixas III, IV e V	193
TABELA 30	Distribuição de <i>você</i> , <i>ocê</i> e <i>cê</i> , conforme a Referência – 1982	196
TABELA 31	Distribuição de <i>você</i> , <i>ocê</i> e <i>cê</i> , conforme a Referência – 2002 – Faixas III, IV e V	196
TABELA 32	Distribuição de <i>você</i> , <i>ocê</i> e <i>cê</i> , conforme a Focalização – 1982	198
TABELA 33	Distribuição de <i>você</i> , <i>ocê</i> e <i>cê</i> , conforme a Focalização – 2002 – Faixas III, IV e V	198

TABELA 34	Distribuição de <i>você</i> , <i>ocê</i> e <i>cê</i> , conforme o Gênero – 1982	200
TABELA 35	Distribuição de <i>você</i> , <i>ocê</i> e <i>cê</i> , conforme o Gênero – 2002 – Faixas III, IV e V	200

LISTA DE GRÁFICOS

GRÁFICO 1	Distribuição de <i>ocê</i> , <i>ocê</i> e <i>cê</i> , conforme a Função Sintática – 2002	120
GRÁFICO 2	Distribuição de <i>ocê</i> , <i>ocê</i> e <i>cê</i> , conforme a Contigüidade – 2002	126
GRÁFICO 3	Distribuição de <i>ocê</i> , <i>ocê</i> e <i>cê</i> , conforme a Referência – Todas as funções – 2002	131
GRÁFICO 4	Distribuição de <i>ocê</i> , <i>ocê</i> e <i>cê</i> , conforme Referência – Função de sujeito – 2002	133
GRÁFICO 5	Distribuição de <i>ocê</i> , <i>ocê</i> e <i>cê</i> , conforme o Foco – 2002	137
GRÁFICO 6	Distribuição de <i>ocê</i> , <i>ocê</i> e <i>cê</i> , conforme a Idade – 2002	141
GRÁFICO 7	Distribuição de <i>ocê</i> , <i>ocê</i> e <i>cê</i> , conforme o Gênero – 2002	144
GRÁFICO 8	Distribuição de <i>ocê</i> , <i>ocê</i> e <i>cê</i> , conforme a Função Sintática – 1982	159
GRÁFICO 9	Distribuição de <i>ocê</i> , <i>ocê</i> e <i>cê</i> , conforme a Idade - 1982	162
GRÁFICO 10	Distribuição de <i>ocê</i> , <i>ocê</i> e <i>cê</i> , conforme a Classe Social – 1982	165

LISTA DE QUADROS

QUADRO 1	Padrões de mudança lingüística nos indivíduos e na comunidade	36
QUADRO 2	Correlação Classe Social X Ocupação	52
QUADRO 3	Perfil dos informantes – 1982	108
QUADRO 4	Perfil dos informantes – 2002 – Crianças e adolescentes	113
QUADRO 5	Perfil dos informantes – 2002 – Faixas etárias III, IV e V	115

SUMÁRIO

INTRODUÇÃO	01
CAPÍTULO 1 - O QUADRO TEÓRICO	07
1.1. Estudos anteriores sobre as formas <i>você, ocê e cê</i>	07
1.2. O estudo da variação e da mudança lingüística	23
1.2.1. A Teoria da Variação	23
1.2.2. Variação e mudança lingüística	24
1.2.3. O estudo da mudança em tempo aparente	31
1.2.4. O estudo da mudança em tempo real	34
1.2.5. Os fatores externos ou extralingüísticos	37
1.2.5.1. Idade	38
1.2.5.1.1. A infância	39
1.2.5.1.2. A adolescência	39
1.2.5.1.3. A fase adulta	43
1.2.5.2. Gênero	45
1.2.5.3. Classe Social	51
1.3. O fator Frequência dos itens lexicais e a mudança lingüística	58
1.3.1. Fidelholtz	59
1.3.2. Phillips	61
1.3.3. Bybee	65
1.4. Os aspectos fonético-fonológicos do português e as formas <i>você, ocê e cê</i>	71
1.4.1 A estrutura da sílaba	72
1.4.2 O acento	74
1.4.3 O ritmo e a velocidade de fala	76
1.4.4 Processos fonológicos do português	77

1.4.4.1 A elisão	78
1.4.4.2 A redução de encontros consonantais	80
1.4.4.3 O apagamento de sílabas	81
1.5 Conclusão	83
CAPÍTULO 2 – AS FORMAS <i>VOCÊ, OCÊ E CÊ</i> – PERCURSO HISTÓRICO	84
2.1 As formas de tratamento do latim ao português antigo	84
2.2 A evolução de <i>Vossa Mercê</i>	88
2.3 A forma <i>você</i> na língua portuguesa	92
2.4 As formas <i>ocê</i> e <i>cê</i> na língua portuguesa	97
2.5 Conclusão	99
CAPÍTULO 3 – PROCEDIMENTOS METODOLÓGICOS	100
3.1 Os fatores internos e externos analisados	101
3.2 Os dados	103
3.2.1 Os dados do <i>corpus</i> de 1982	105
3.2.1.1 A coleta de dados	105
3.2.1.2 O perfil dos informantes	105
3.2.2 Os dados do <i>corpus</i> de 2002	108
3.2.2.1 A coleta de dados de crianças e adolescentes	109
3.2.2.2 O perfil dos informantes crianças e adolescentes	111
3.2.2.3 A coleta de dados dos demais informantes	114
3.2.2.4 O perfil dos informantes das demais faixas etárias	114
3.3 A análise dos dados	116
3.4 Conclusão	116
CAPÍTULO 4 – A ANÁLISE SINCRÔNICA DOS DADOS – O <i>CORPUS</i> DE 2002	117
4.1 As ocorrências de <i>você, ocê e cê</i> , conforme a Função Sintática	119

4.2	As ocorrências de <i>você, ocê</i> e <i>cê</i> , conforme a Contigüidade	125
4.3	As ocorrências de <i>você, ocê</i> e <i>cê</i> , conforme a Referência	130
4.4	As ocorrências de <i>você, ocê</i> e <i>cê</i> , conforme o Foco	135
4.5	As ocorrências de <i>você, ocê</i> e <i>cê</i> , conforme a Idade	140
4.6	As ocorrências de <i>você, ocê</i> e <i>cê</i> , conforme o Gênero	143
4.7	Os testes dos informantes crianças e adolescentes	147
4.8	Conclusão	154
CAPÍTULO 5 – A ANÁLISE DOS DADOS – CORPUS DE 1982		156
5.1	As ocorrências de <i>você, ocê</i> e <i>cê</i> , conforme a Função Sintática	158
5.2	As ocorrências de <i>você, ocê</i> e <i>cê</i> , conforme a Idade	161
5.3	As ocorrências de <i>você, ocê</i> e <i>cê</i> , conforme a Classe Social	164
5.4	A análise das formas e os processos fonético-fonológicos do português	167
5.4.1	A função de Sujeito	169
5.4.2	A função de Objeto de Preposição	171
5.4.3	A função de Objeto de Verbo	177
5.5	Conclusão	182
CAPÍTULO 6 – A ANÁLISE DIACRÔNICA DOS DADOS: O ESTUDO DAS FORMAS <i>VOCEÊ, OCÊ</i> E <i>CÊ</i> EM TEMPO REAL ...		184
6.1	O fator Função Sintática	186
6.2	O fator Idade	189
6.3	O fator Classe Social	193
6.4	O fator Referência	196
6.5	O fator Focalização	197
6.6	O fator Gênero	200
6.7	Os fatores não-significativos	203
6.8	Conclusão	206

CAPÍTULO 7 - DISCUSSÃO FINAL DOS RESULTADOS E CONCLUSÕES	207
REFERÊNCIAS BIBLIOGRÁFICAS	219

INTRODUÇÃO

Este trabalho trata do uso das formas *você*, *ocê* e *cê* em Belo Horizonte, em dois períodos de tempo distintos - 1982 e 2002 -, tendo por objetivo analisar, com base na Teoria da Variação, os fatores lingüísticos e extralingüísticos que condicionam esse uso e verificar, no tempo real, se está havendo mudança em progresso em relação às três formas.

A forma *você* originou-se da forma de tratamento *Vossa Mercê*. Essa expressão surgiu no século XIV, firmando-se no século XV, sendo usada inicialmente para o tratamento ao rei. As formas “Vossa + Nome”, como *Mercê*, *Senhoria*, *Alteza* etc., ao exaltar as qualidades morais e a superioridade do monarca, acabaram por substituir o pronome *vós*, utilizado indistintamente tanto para o rei quanto para outros nobres.

Entretanto, já na segunda metade do século XV, a expressão *Vossa Mercê* para tratamento ao rei foi substituída por outras formas nominais, ao mesmo tempo em que aquela se ia vulgarizando, chegando, por fim, a ser usada para qualquer pessoa não íntima, a quem se deveria respeito.

Em conseqüência da vulgarização de seu uso, essa expressão, um tanto longa, sofreu consecutivas reduções fonéticas, dando origem a diversas variantes, e, dentre elas, a forma

você. No início do século XVI, era essa a situação da língua portuguesa, a qual foi trazida para o Brasil.

Durante o processo de gramaticalização pelo qual *Vossa Mercê* passou, originando o pronome *você*, as duas formas percorreram caminhos distintos, na língua: enquanto *Vossa Mercê* tinha mobilidade na frase, *você* era usado apenas na função de sujeito pré-verbal, expandindo-se lentamente para outros contextos.

Outra consequência do mesmo processo é a reorganização do quadro pronominal da língua, com a entrada de uma forma que fazia referência à segunda pessoa, mas que fazia a concordância verbal com a terceira. Essa situação, por sua vez, originou a necessidade de preenchimento do sujeito, o que fez aumentar a frequência de uso desse pronome, gerando novas reduções e dando origem a outras formas, usadas atualmente no português brasileiro (doravante, PB), como *ocê* e *cê*.

As formas *você*, *ocê* e *cê* já foram alvo de estudo de alguns pesquisadores, com base em dados de fala coletados em algumas localidades do Brasil e sob diversas perspectivas teóricas. Um dos aspectos mais interessantes do uso dessas três formas é sua distribuição irregular: *cê* normalmente não ocorre exercendo as funções de objeto de verbo e de preposição, e tem ocorrência nula – ou quase – nos casos de foco enfático.

Um outro fato interessante é o comportamento da forma *ocê*: embora, dentre as três formas, seja aquela que apresente a menor porcentagem de ocorrências no total, é bastante favorecida nas funções de objeto de verbo e de preposição, o que indica que outro(s) fator(es) está(ão) envolvido(s) nesse uso. E, realmente, quando se observam esses contextos, vê-se que diante de *ocê* é muito comum apresentar-se uma palavra terminada em vogal átona, favorecendo, assim, os processos fonológicos da língua, que transformam essas expressões em construções não-marcadas do português.

Do que foi exposto até aqui, observa-se que uma mudança lingüística não é um fato isolado na língua, mas atua em conjunto com outros processos lingüísticos e com fatores de ordem

social. Ademais, a mudança não ocorre de maneira rápida nem uniforme, pelo contrário: no seu início é um processo bastante lento, em que a forma inovadora aos poucos vai expandindo-se para novos contextos, ocupando os espaços anteriormente exclusivos da forma padrão. Assim sendo, determinar esse caminho, identificando os contextos que mais favorecem e os que mais restringem essa expansão deve estar entre os objetivos dos trabalhos que busquem investigar os processos de mudança lingüística.

Outra característica inerente a esses processos é sua heterogeneidade, que decorre da própria heterogeneidade social. Assim, caracterizar o grupo social que favorece o uso da forma inovadora é necessário para se descrever o próprio processo de mudança. Portanto, é preciso identificar as pessoas que formam esse grupo: sua idade, sua classe social, seu gênero, seu nível de escolaridade etc.

Portanto, a escolha de uma perspectiva teórica que caracterize e explique os processos de mudança envolve a concepção que se tenha do que seja uma língua e do seu funcionamento. É uma corrente teórica que concebe a língua como um sistema intimamente ligado à sociedade que a fala e que sistematiza sua inerente heterogeneidade é a Sociolingüística ou a Teoria da Variação. Daí ela ser tomada como marco teórico desta pesquisa.

Portanto, neste trabalho serão analisados os fatores lingüísticos e extralingüísticos que podem auxiliar na compreensão dos processos de variação de uso e de mudança lingüística que envolve essas três formas: a **função sintática** exercida pelas formas; o **tipo de frase** em que aparecem; sua **contigüidade** em relação ao verbo; e seu comportamento nas construções com **tópico** e com **foco**, assim como na expressão da **referência**. Com relação aos fatores extralingüísticos, foram tomados a **idade**, o **gênero** e a **classe social** dos informantes.

Nesta investigação, foram analisados dois *corpora*: um, de 2002, composto por dados de fala de informantes de cinco faixas etárias: faixa I (08 a 11 anos), II (12 a 15 anos), III (de 16 a 30 anos), IV (de 31 a 47 anos) e V (acima de 47 anos). O outro *corpus* contém dados

de fala de informantes de três faixas etárias (de 16 a 30, de 31 a 47 e acima de 47 anos), coletados em 1982. Apesar de serem destinados originalmente a propósitos bastante diferentes dos objetivos deste estudo, como se verá ao longo desta tese, os dois *corpora* puderam ser aproveitados para a presente pesquisa.

Em resumo, este trabalho tem por objetivo geral avançar com os estudos já realizados sobre o uso das formas *você*, *ocê* e *cê*, investigando os fatores que condicionam a sua variação e a indicação de mudança, e por objetivos específicos os listados abaixo.

OBJETIVOS ESPECÍFICOS

1. Determinar os fatores lingüísticos e extralingüísticos que condicionam o uso das três formas, nos dois períodos de tempo pesquisados.
2. Verificar se está havendo mudança em progresso das formas com respeito à expressão da referência.
3. Verificar, no tempo real, se está havendo mudança em progresso com respeito às formas *você*, *ocê* e *cê*, em Belo Horizonte.

Com base no exposto, cinco hipóteses nortearam este estudo, as quais aparecem a seguir.

HIPÓTESES

- 1) A forma *cê* apresenta um comportamento sintático peculiar, não ocorrendo nos mesmos contextos que *você* e *ocê*.
- 2) O uso das formas *você*, *ocê* ou *cê* é condicionado por fatores de natureza fonética e fonológica da língua portuguesa.

- 3) A forma *ocê* deve ter relativamente pouca frequência nos dados.
- 4) Está havendo, no PB, um processo de especialização em relação ao par *você* e *cê*, com a preferência de *cê* com referência indefinida.
- 5) A comunidade de Belo Horizonte, com relação às três formas, apresenta o padrão de mudança geracional (cf. Labov, 1994, p. 83).

Esta tese está dividida em sete partes, ou capítulos:

No primeiro capítulo, serão apresentados os pressupostos teóricos que norteiam esta pesquisa. Em primeiro lugar, serão resumidos alguns dos mais importantes trabalhos já realizados acerca do tema em questão. Em segundo, serão apresentados os tópicos da Teoria da Variação mais relevantes para este estudo. Em seguida, serão tratados dois assuntos relacionados às questões discutidas nesta pesquisa, que são a importância da frequência dos itens lexicais para os processos de mudança e os aspectos fonéticos e fonológicos da língua portuguesa que estão relacionados à variação de uso das três formas.

No segundo capítulo, será feita uma abordagem histórica das três formas em estudo, destacando-se a ascensão e a decadência de *Vossa Mercê*, o processo de gramaticalização pelo qual a expressão passou e o uso das formas *você*, *ocê* e *cê* na língua portuguesa.

No terceiro, serão apresentados os procedimentos metodológicos utilizados na pesquisa: a seleção e o perfil dos informantes, bem como a coleta de dados das duas amostras.

No quarto, os dados coletados em 2002 serão apresentados e analisados quantitativa e qualitativamente, tendo em vista os fatores internos e externos considerados significativos para o processo em estudo. Além desses aspectos, serão analisados os resultados obtidos por meio dos dados dos informantes crianças e adolescentes deste *corpus*.

No quinto, será feito o mesmo com respeito aos dados de 1982, bem como serão apresentados e discutidos os resultados referentes à análise fonético-fonológica envolvendo o uso das formas, nos dois *corpora*.

No sexto, será apresentado o estudo em tempo real, contrapondo-se os dados de 2002 aos dados de 1982, a fim de que sejam respondidas as perguntas que originaram este trabalho, caracterizando-se melhor, desse modo, o uso das três formas na comunidade de Belo Horizonte e a mudança por que passa o par *você* e *cê*.

No sétimo e último, será feita uma discussão geral dos resultados encontrados para as três formas, a fim de que se possa descrever sua trajetória no português brasileiro.

CAPÍTULO 1 – O QUADRO TEÓRICO

O objetivo deste estudo é analisar o comportamento lingüístico das formas *você*, *ocê* e *cê* sincrônica e diacronicamente. Para tanto, neste capítulo, serão apresentados os aportes teóricos que embasaram essas análises e que propiciaram a compreensão deste fenômeno.

Primeiramente, serão apresentados os resultados de trabalhos anteriores que trataram deste mesmo assunto. Em seguida, serão abordados alguns tópicos importantes acerca da variação e da mudança lingüística. Posteriormente, serão resumidos os trabalhos de Fidelholtz, Phillips e Bybee, que tratam da frequência de uso dos itens lexicais e suas conseqüências para o processo de mudança, um assunto de fundamental importância para a análise das formas *você*, *ocê* e *cê*. Por último, serão elencados alguns processos fonéticos e fonológicos que, ao que parece, também exercem influência sobre o comportamento das três formas e, por isso, são importantes para este estudo.

1.1 Estudos anteriores sobre as formas *você*, *ocê* e *cê*

As formas *você*, *ocê* e *cê* despertaram e continuam despertando o interesse de alguns estudiosos do PB, como será mostrado agora.

Vitral (1996), estudando o comportamento sintático de *você*, *ocê* e *cê*, verifica, no *corpus* analisado, que a forma *cê* não é usada em determinados ambientes, como sujeito em posição pós-verbal, objeto de verbo e de preposição, em posição de tópico e recebendo

foco, ou seja, *cê* somente ocorre como sujeito pré-verbal. A partir desses resultados, o autor conclui que a forma *cê* está dando continuidade a um processo de gramaticalização¹ iniciado com o pronome de tratamento *Vossa Mercê*: *cê* corresponde à etapa da cliticização, conforme aponta Vitral (1996, p. 119)²:

Item com significado lexical: Vossa Mercê > item gramatical: você > clítico: ê > afixo flexional

Ramos (1997), por sua vez, com o objetivo de testar a hipótese de cliticização de *cê* e também de descrever o uso de *você*, *ocê* e *cê*, explicitando os fatores lingüísticos e extralingüísticos que o condicionam, analisa 342 dados obtidos através de entrevistas sociolingüísticas com informantes belo-horizontinos. Os fatores levados em conta na análise são:

1) Internos

- (i) função sintática da forma (sujeito, objeto de verbo ou objeto de preposição);
- (ii) focalização da forma (se assume ou não a posição de foco);
- (iii) tipo de referência da forma (específica ou não-específica);
- (iv) tipo de oração em que a forma aparece (interrogativa simples, interrogativa ‘que que’ ou declarativa);
- (v) posição da forma em relação ao verbo (contígua ou não-contígua);
- (vi) topicalização (se a forma pode ou não ser tópico).

¹ Gramaticalização, segundo Castilho (1997, p. 31), é “o trajeto empreendido por um item lexical, ao longo do qual ele muda de categoria sintática (=recategorização), recebe propriedades funcionais na sentença, sofre alterações morfológicas, fonológicas e semânticas, deixa de ser uma forma livre, estágio em que pode até mesmo desaparecer, como consequência de uma cristalização extrema”.

² Em trabalho posterior (CIRÍACO, VITRAL e REIS, 2004, apresentado adiante), o autor reformula sua hipótese inicial e propõe que *cê* ainda não é um clítico pleno, encontrando-se ainda em processo de cliticização.

2) Externos

- (i) sexo;
- (ii) idade (jovens, medianos e velhos).

Os principais resultados obtidos pela autora são estes:

- a) as três formas ocorrem na posição de sujeito em 94,4% das frases;
- b) o uso de *você* está diminuindo, de acordo com a faixa etária; a forma *cê* é a preferida pelas três faixas etárias, mas não apresenta aumento de ocorrências; já o uso de *ocê* teve um pequeno aumento;
- c) a forma *cê* não aparece como complemento de verbo nem como complemento de preposição;
- d) nas construções topicalizadas, a forma *cê* aparece em 88% dos casos;
- e) a forma *cê* aparece não-contígua ao verbo, mas a probabilidade de que isso ocorra é pequena (PR=.32). A não-contigüidade ocorre com advérbios curtos, como *não*, *já*, *só* etc.
- f) nas orações interrogativas simples, há a preferência pela forma *você*; nas orações interrogativas em que aparece a estrutura ‘que que’ e nas orações declarativas, há a preferência pela forma *cê*;
- g) os três casos de focalização acontecem somente com a forma *você*, ou seja, *cê* não aparece recebendo foco;
- h) com relação ao tipo de referência, os dados mostram que a referência indefinida ocorre mais na fala dos jovens (55%) que na fala dos medianos (36%) e velhos (10%). Esses resultados indicam uma mudança em progresso, em relação ao tipo de referência: está havendo uma preferência pelo uso de referência indefinida;
- i) ainda com respeito à referência, quando se trata de referência definida, “os jovens e medianos preferem usar o pronome *você* (67%); quando é indefinida, preferem *cê* (56%). O inverso se observa no comportamento dos velhos: quando a referência do item é indefinida, a preferência é por *você* (53%); quando definida, a preferência é por *cê*

(65%)” (RAMOS, 1997, p. 51). Esses resultados levaram a autora à hipótese de que está havendo uma especialização quanto ao uso dessas formas, ou seja, uma outra mudança em progresso: o *você* está sendo preferido para expressar referência definida, e o *cê*, para referência indefinida;

- j) os resultados encontrados confirmam a hipótese de cliticização da forma *cê* (RAMOS, 1997, p. 57)

Em outro trabalho sobre *você*, *ocê* e *cê* (RAMOS, 2000), a autora descreve, com base no tempo aparente, o uso das três formas em Belo Horizonte e Ouro Preto. Para isso, analisa dados extraídos de entrevistas sociolinguísticas, entrevistas de TV e transcrição de uma aula de nível universitário. Os informantes estão divididos em três faixas etárias – jovens (de 15 a 25 anos), medianos (de 52 a 63 anos) e velhos (de 77 a 91 anos). São considerados, na análise, fatores internos (posição das formas na oração e tipo de referência) e fatores externos (escolaridade, sexo, idade e região geográfica). Os resultados a que a autora chegou são:

- a) a forma *cê* aparece em 30,8% dos casos nas entrevistas de televisão, um índice alto, tendo em vista o estilo cuidado de fala ali empregado, o qual privilegia a forma *você*;
- b) a distribuição das três formas com relação à função sintática que exercem nas frases confirma a hipótese de cliticização;
- c) com respeito à referência do item, os dados de Ouro Preto reforçam a hipótese de especialização semântica das formas. Nessa cidade, a probabilidade de uso de *você* com referência [-determinada] e de *cê* com referência [+determinada] vai aumentando conforme a idade dos informantes vai diminuindo, ou seja, enquanto a probabilidade de uso de *você* [-det] e *cê* [+det] pelos velhos é de, respectivamente, 0.42 e 0.44, a probabilidade com relação aos medianos é de 0.56 e 0.63, respectivamente, e, com relação aos jovens, 0.59 e 0.76, respectivamente. Globalmente, há uma leve preferência pelo uso de *você*, quando o referente é [-determinado] (RAMOS, 2000, p. 188). Assim, a autora conclui, com base no tempo aparente, que está havendo mudança em progresso.

Alves (1998) é outro autor que estuda as formas *você*, *ocê* e *cê*. Em seu trabalho, procura descrever os recursos preferidos de indeterminação do sujeito pelos belo-horizontinos.

Da análise de 287 dados, constituídos de gravações de entrevistas e de programas de TV e rádio, além de entrevistas sociolinguísticas, o autor constata que as formas tradicionais de indeterminação do sujeito, reconhecidas pela Gramática Tradicional - o verbo na terceira pessoa do plural e verbo intransitivo ou transitivo indireto + SE -, estão sendo substituídas principalmente pelo uso de *você*, *ocê* e *cê* (45,6% dos casos), *a gente* (25,4%) e *as pessoas* (6,6%). Esses resultados parecem indicar “um processo de mudança no português brasileiro, relativo à construção de estruturas que expressam indeterminação do sujeito” (ALVES, 1998, p. 82).³

Com relação às formas *você*, *ocê* e *cê*, particularmente, o autor encontra os resultados expostos na tabela abaixo, transcrita de Alves (1988, p. 57, Quadro III):

TABELA 1
As formas *você*, *ocê* e *cê* no processo de indeterminação

Formas Pronominais	Frequência	%
Você	78	59
Ocê	09	07
Cê	44	34
Totais	131	100

Os dados revelam que a forma *você* é a preferida para expressar indeterminação, vindo em segundo lugar a forma *cê* e, por último, a forma *ocê*. Esses dados vão de encontro aos obtidos por Ramos (1997, p. 50, tabela 7), que atestam a preferência dos informantes jovens

³ Duarte (2003), analisando dados de fala do Rio de Janeiro no tempo aparente e no tempo real, encontra resultados semelhantes para a indeterminação do sujeito: a forma *você* aparece em primeiro lugar, seguida por *a gente* e as formas com categoria vazia, e a 3ª pessoa do plural.

e medianos pelo uso de *cê* para expressarem referência indefinida, ou seja, para expressarem indeterminação. Alves justifica essa diferença pelo fato de a maior parte de seu *corpus* ser constituída de entrevistas gravadas de rádio e TV, não sendo possível saber a idade dos informantes. Assim, declara o autor, é provável que a maioria de seus sujeitos seja constituída por pessoas classificadas como velhas. (ALVES, 1998, p. 67).

Um outro trabalho sobre as formas *você*, *ocê* e *cê* é o de Coelho (1999), que procura identificar os fatores lingüísticos e extralingüísticos que condicionam o uso dessas formas no município de São Francisco/MG. Os fatores considerados pela autora são:

1. Lingüísticos

- (i) tipo de contexto em que as três formas pronominais aparecem (interpretação definida ou indefinida);
- (ii) ambiente fonológico que precede as formas (vogal, consoante ou pausa);
- (iii) função sintática das formas pronominais (sujeito, objeto ou complemento de preposição);
- (iv) tipo de frase em que as formas ocorrem (afirmativa, negativa ou interrogativa).

2. Extralingüísticos

- (i) procedência geográfica dos informantes (urbana ou rural);
- (ii) classe social (socialmente privilegiada ou não-privilegiada);
- (iii) idade (de 15 a 25 anos; de 26 a 49 anos; e de 50 anos ou mais);
- (iv) sexo;
- (v) escolaridade – baixa (de 0 a 04 anos), média (de 05 a 11 anos), ou alta (superior a 11 anos).

A autora, nesse estudo, encontra os seguintes resultados, dentre outros:

- a) a variante *cé* lidera sobre as demais, seguida de *ocê* e *você*, o que significa que *cé* já está implementada naquela comunidade;
- b) a forma *cé* é favorecida pela função de sujeito da frase (61% das ocorrências);
- c) a variante *você* é mais favorecida por contextos indefinidos (22% das ocorrências) que por contextos definidos (19% das ocorrências);
- d) na função de objeto de preposição, a forma preferida pelos informantes foi *ocê* (45% das ocorrências); já na função de objeto do verbo, a forma mais utilizada foi *você* (54% das ocorrências);
- e) a forma *você* ocorre mais na zona urbana (28% das ocorrências) que na rural (8%).

Abaixo seguem transcritas as principais conclusões da autora:

- a) “É notável a preferência pelas formas com interpretação definida, usadas em 73% dos 622 casos. (...) Das três formas, a variante *cé* é a mais usada, tanto com interpretação definida (55%) quanto com interpretação indefinida (63%)” (p. 62);
- b) “Os resultados (...) mostram que: os jovens usam mais as variantes não-padrão do que a padrão, mas os velhos também usam mais as variantes não-padrão; apenas os medianos é que usam a padrão mais do que as não-padrão. Portanto, não se pode falar em mudança em progresso mas, pode-se dizer que a variante *cé* já está implementada na língua dessa comunidade” (p. 72).
- c) “A distribuição das três variantes *você*, *ocê* e *cé* segundo a Procedência Geográfica mostra a forma *você* como sendo tipicamente urbana, onde a forma *ocê* é pouco freqüente e a forma *cé* é a preferencialmente usada. De acordo com essa distribuição, a forma *ocê* ocorre menos na área urbana do que na área rural, onde a forma *cé* é, também, a preferencialmente usada.” (p. 81).
- d) “No que diz respeito ao uso das variantes com interpretação definida e com interpretação indefinida, os resultados contrariam a hipótese de “especialização” das formas” (p. 81).

Andrade (2004) também pesquisa as formas *você*, *ocê* e *cê*, analisando 1258 dados de fala de quinze crianças de 10 a 14 anos do Distrito Federal e de uma informante de 25 anos que já residia ali há 19 anos. O trabalho tem o objetivo de:

- (i) determinar os condicionamentos sintáticos, semânticos, discursivos e sociais que atuam na ocorrência das formas;
- (ii) testar a hipótese de cliticização de *cê* (cf. VITRAL, 1996; e RAMOS, 1997) e também a hipótese de que *cê* é um pronome fraco (cf. KATO, 1999);
- (iii) comparar os dados do Distrito Federal com os de Minas Gerais, através dos resultados de Ramos (1997) e de Coelho (1999). Para atingir esses objetivos, a pesquisadora analisa os seguintes fatores:

1. Internos

- (i) função sintática da forma;
- (ii) tipo de frase em que as formas aparecem;
- (iii) contigüidade das formas em relação ao verbo e o tipo de elemento que aparece entre eles;
- (iv) paralelismo discursivo;
- (v) referência das formas.

2. Externos

- (i) procedência geográfica dos informantes;
- (ii) gênero;
- (iii) variação estilística.

Os principais resultados deste estudo são:

- a) a forma *você* foi a preferida tanto pelas crianças quanto pela informante jovem, vindo em segundo lugar a forma *cê*, em terceiro lugar *zero*⁴ e, em quarto, *ocê*: 62,1%; 21,8%; 14,1% e 2%, respectivamente;
- b) as formas *você*, *ocê* e *cê* apresentam um comportamento sintático diferenciado: a forma *cê* ocorre apenas como sujeito das orações, é favorecida pelas frases do tipo “que que” e não aparece em sentenças marcadas, como nas orações clivadas, nas topicalizadas e nas orações com sujeitos coordenados;
- c) com respeito à referência, os dados mostram que, para a referência definida, *você* é a forma preferida, com 63% dos casos, vindo depois *cê*, com 21%. Para a referência indefinida, *você* também é a forma preferida, com 71% dos casos, vindo em seguida *cê*, com 24%;
- d) com relação ao tipo de elemento interveniente entre a forma *cê* e o verbo, os dados mostram que a contigüidade e a interposição de clítico (*me*, *te*, *se*, *num*) favorecem levemente a forma *cê* (PR = .53 e .54, respectivamente); a interposição de monossílabo tônico desfavorece ligeiramente *cê* (PR = .45) e a interposição de elemento não-clítico desfavorece fortemente essa forma (PR = .22);
- e) levando-se em consideração os fatores externos, a forma *cê* é mais usada na comunidade rural, é preferida pelos homens e é usada em situações informais de fala. Em situações formais, *você* é a forma preferida;

O fato de a forma *cê* admitir a interposição de elementos entre ela e o verbo – o que não é característico dos clíticos – faz a autora concluir que *cê* “apresenta tanto características de pronome fraco como de clítico” (ANDRADE, 2004, p. 130) e ela justifica sua posição: “Partindo do pressuposto de que a gramaticalização é um processo lento e gradual, pode-se admitir que *cê* pode estar adquirindo comportamento clítico, sem, no entanto, apresentar todas as características próprias de um clítico verdadeiro” (ANDRADE, 2004, p. 130).

O estudo de Andrade é valioso por permitir ver com mais clareza os fatores lingüísticos e extralingüísticos que estão relacionados ao uso das três formas, especialmente *cê*, no Distrito Federal. Entretanto, há que se ressaltar o fato de que, nessa cidade, o uso de *cê* é

⁴ Andrade (2004) chamou de “zero” os casos de omissão do sujeito ou do objeto.

bastante reduzido, caso se comparem os dados de Brasília com os de Belo Horizonte, o que implica que não se podem generalizar as conclusões de Andrade. E, de fato, os dados desta tese confirmam essa suposição, como será visto em capítulo adiante.

Um outro estudo sobre formas reduzidas, incluindo-se a partícula *cê*, é realizado por Ciríaco, Vitral e Reis (2004), cujo objetivo é atestar foneticamente as reduções de *cê* (pronome *você*), *ez* (pronome *eles*) e *num* (*não*), além de comprovar-lhes, por meio de análise acústica, o processo de cliticização. Para isso, foi desenvolvido “um experimento composto de treze sentenças cuidadosamente escolhidas, observando-se os contextos e propriedades fonéticas de duração total, ritmo e acento das mesmas.” (CIRÍACO et al., 2004, p. 145)

O experimento consistiu no seguinte: as frases foram lidas por alunos do primeiro período do curso de Letras da UFMG, nascidos e residentes em Belo Horizonte, sendo dois de cada sexo. Cada informante fez três leituras⁵ de cada uma das treze sentenças do experimento, sendo que seis foram lidas três vezes pelos informantes, cada uma pronunciada com uma velocidade diferente: normal, rápida e rapidíssima. As outras sete frases foram pronunciadas apenas uma vez a cada leitura. No total, o *corpus* consta de 75 realizações das sentenças por cada informante, num total de 300 realizações.

As frases do experimento são as seguintes (CIRÍACO et al., 2004, p. 147):

1. Diga *Cê* disse que *Zé* vai viajar. (*Cê* = Cecília)
2. Diga *eles* falou que *Zé* vai viajar.
3. Diga *Jô não* disse nada não.
4. Diga *cê* disse que *Zé* vai viajar. (*cê* = pronome)
5. Diga *cê* falou que *Zé* vai viajar. (*cê* = pronome)
6. Diga o estilista vestiu *Cê* bem. (*Cê* = Cecília)

⁵ Antes de pronunciar as frases, os informantes foram instruídos a guardá-las rapidamente na memória.

7. Diga *eles* disse que Zé vai viajar.
8. Diga o estilista vestiu-*se* bem.
9. Diga *Cê* falou que Zé vai viajar. (Cê = Cecília)
10. Diga *Jô não* falou nada não.
11. Diga o estilista vestiu *cê* bem. (cê = pronome)
12. Diga *ex* diz muita coisa.
13. Diga *eles* diz muita coisa.

Após a gravação das leituras, procedeu-se à edição e à análise acústica das reduções, atestando sua duração e intensidade. Assim, embora os autores reconheçam que o estudo esteja em fase preliminar, contando com poucos dados, estes confirmam que a forma *cê* apresenta características fonéticas intermediárias entre o apelido CÊ e o clítico “*se*”, o que vem a comprovar o processo de cliticização por que passa a forma *cê*.⁶ Desse modo, os resultados obtidos levaram os autores às seguintes conclusões:

- a) Pode-se atestar foneticamente a redução de *você* (= *cê*);
- b) *cê* apresenta menor duração que a forma plena correspondente, CÊ (= Cecília), na posição de sujeito. Na posição de objeto de verbo, *cê* apresenta duração menor que “CÊ” e maior que o clítico “*se*”;
- c) *cê* (pronome) apresenta intensidade normalmente menor que CÊ (nome) e igual à de “*se*”. Portanto,
- d) o processo de cliticização de *cê*⁷ pode ser confirmado, tanto pelo parâmetro da duração quanto da intensidade.

⁶ Contudo, de acordo com as análises feitas, a intensidade das formas reduzidas não é muito diferente das formas plenas, o que não era esperado. O Prof. Dr. José Olímpio de Magalhães, fazendo uma observação a esse fato, indica que, ao que parece, a intensidade está compensando a duração das formas reduzidas (2004, p. 154).

⁷ As conclusões valem também para as outras reduções estudadas, “*num*” e “*ez*”; entretanto, somente se fará referência à forma *cê*, dados os objetivos deste trabalho.

O último trabalho a ser mencionado neste item é o de Barbosa (2005), que pretende determinar o *status* da forma *cê*: clítico ou palavra plena. Para tanto, a autora, lançando mão de dados introspectivos e reais⁸, realiza testes sob as perspectivas da Cliticização (testes fonológicos, acentuais, morfológicos e sintáticos)⁹ e sob a perspectiva da Fonologia Prosódica (testes prosódicos). Abaixo segue um resumo de cada um deles e seus respectivos resultados.

1) Testes fonológicos

Segundo a autora, “Os testes fonológicos de Zwicky (1985) indicam o comportamento do elemento investigado atentando-se à sua formação no constituinte prosódico palavra fonológica.” (BARBOSA, 2005, p. 97). O exemplo analisado é, segundo notação da autora, o de número 96:

(96) [Cê,]ω]C]Φ]I [que está de blusa amarela,]I [[[aproxime-se] ω]C]Φ]I]U¹⁰

A partir desse exemplo, Barbosa analisa (2005, p. 97):

*Em (96) temos três frases entoacionais cuja organização prosódica é justificada pelo fato de orações relativas explicativas formarem seu próprio domínio entoacional. Nesse exemplo, o pronome *cê* é uma única palavra plena (possui um acento primário) que forma uma palavra fonológica; esta forma uma unidade prosódica superior: grupo clítico. Este grupo clítico forma uma frase fonológica não-ramificada (formada apenas por uma única palavra fonológica), que por sua vez constitui uma frase entoacional. Como o clítico pronominal se ancora num elemento hospedeiro verbal, é impossível analisar *cê* como clítico, pois não há este elemento hospedeiro necessário no grupo clítico e nem no constituinte prosódico imediatamente superior: frase fonológica.*

⁸ Os dados analisados por Barbosa (2005) são os mesmos de Coelho (1999), coletados em São Francisco/MG.

⁹ Os testes fonológicos, acentuais, morfológicos e sintáticos são os de Zwicky (1985). Também são aplicados os testes sintáticos de Kayne (1975).

¹⁰ Os símbolos usados no exemplo significam: ω = palavra fonológica; C = grupo clítico; Φ = frase fonológica; I = frase entoacional; U = enunciado fonológico.

2) Testes acentuais

A análise deste aspecto é feita com base na distinção acentual presente nas duas frases abaixo:

“Foi *cê* o culpado de tudo” e

“Foi-se o culpado de tudo” (2005, p. 98)

Pelo fato de um clítico depender do acento primário de outra palavra, seu “hospedeiro”, - o que acontece com o pronome *se*, mas não ocorre com *cê* -, a autora conclui que não há semelhança acentual entre *se* e *cê*, ou seja, *cê* não é clítico.

3) Testes morfológicos¹¹

Os testes realizados são os de Zwicky (1985), que determinam se um elemento é clítico. São eles: de ligação, de fechamento, de ordem e de distribuição, sendo que, em todos eles, Barbosa conclui que *cê* comporta-se como palavra plena (2005, p. 102-110).¹²

4) Testes sintáticos

Nesses testes, são analisados os seguintes fatores: a) apagamento sob identidade (se um dos dois elementos $x + y$ puder ser apagado, os dois são palavras plenas); b) substituição (se um elemento puder ser substituído por outro, os dois são formas plenas); e c) de movimento (se um de dois elementos combinados puder ser movimentado, os dois serão formas plenas).

¹¹ A autora, nesse subitem, usa a seguinte denominação (BARBOSA, 2005, p. 99): “Testes usando similaridades entre afixos flexionais e clíticos”, já que estes se comportam como afixos flexionais de uma palavra plena.

¹² Pelo fato de que a descrição desses testes ficaria muito extensa, somente foi apresentado seu resultado.

Todos esses testes, de acordo com Barbosa (2005), provam que *cê* não é um clítico. Além deles, a autora apresenta exemplos de que: i) entre *cê* e o verbo pode haver intercalação de vários elementos; ii) *cê* pode ser modificado; iii) *cê* pode aparecer coordenado a um NP lexical e a um pronome tônico; e iv) *cê* pode ser topicalizado, o que também comprova que *cê* não pode ser um clítico. Os exemplos dados pela autora para confirmar os testes acima são, respectivamente:

- a) “Eu vou à festa e **cê vai** também.” e
“Eu vou à festa e **cê**, também.”
- b) “Duas horas e **cê** só falou isso agora!” e
“Duas horas e **tu** só falou isso agora!”
- c) “Eu acho que **cês** não passaram de ano, não.” e
“**Cês**, eu acho que não passaram de ano, não.”
- d) “Cê **nunca mais** me verá.”
- e) “Cê é o mais sem-vergonha.” e
“**De todos**, cê **só** é o mais sem-vergonha.”
- f) “**Cê e Joaquina** vão viajar juntas.” e
“**Cê e ele** serão felizes.”¹³
- g) “**Cês**, eu acho que não passaram de ano, não.”

¹³ Barbosa (2005), entretanto, afirma que os sintagmas “Joaquina e cê” e “Ele e cê” são agramaticais. Para a autora, isso se explica pela “preferência de *cê* pela posição precedida de pausa, conforme sua distribuição fonológica” (op. cit., p. 110)

5) Testes prosódicos

Os testes prosódicos foram realizados para confirmar a hipótese da autora de que *cê* é uma palavra fonológica, tal como outros monossílabos tônicos do português, como **dê**, **nu** e **dá**. Comparando-se essas palavras com suas respectivas formas átonas, têm-se:

dê/de

nu/no

dá/da

Tendo em vista que esses pares de monossílabos se distinguem apenas pela tonicidade, sendo os primeiros elementos tônicos, a autora afirma que, com a forma *cê*, acontece o mesmo, como se vê abaixo:

cê/se

Assim, Barbosa conclui que *cê* é um monossílabo tônico, e não um clítico.

De acordo com a autora, isso é comprovado também pelo fato de que *cê*, ao contrário de *se*, não sofre redução vocálica nem semivocalização ou apagamento, como em

“**Cê** ia” [e] / *[i] / *0 (zero)

Um exemplo com *se* seria:

“*Se* una” [e] / [i]

Para explicar a fraca intensidade de *cê* em certos contextos, Barbosa cita Major (1985). Esse autor observa que “um acento no nível da palavra pode ser alterado no nível da sentença, isto é, padrão rítmico do nível da sentença pode alterar padrão rítmico do nível da palavra para acomodar uma tendência acentual.” (BARBOSA, 2005, p. 115-116). No caso da forma *cê*, a autora acredita que, dependendo do contexto em que esteja – uma unidade prosódica acima da palavra fonológica -, ela poderá sofrer uma perda de seu acento. Em resumo, a autora afirma:

*O que ocorre é que a palavra plena *cê* sofre perda acentual no nível da sentença (no constituinte prosódico frase entoacional) em determinados ambientes, recebendo acento fraco. Mas isso não quer dizer que perdeu seu acento de palavra, apenas que não recebeu acento frasal. Outros ambientes são indicadores de sua presença com acento forte, também no nível da sentença (no constituinte frase entoacional). (BARBOSA, 2005, p. 119).*

Concluindo-se a apresentação do trabalho de Barbosa, todos os testes aplicados por essa autora com relação à forma *cê* comprovam que esta é uma palavra plena, e não um clítico.

Apresentar esses estudos, neste subitem, teve por objetivo demonstrar a complexidade da análise das formas *você*, *ocê* e *cê* no PB atualmente. E, de fato, os trabalhos listados até aqui sobre essas formas apresentam claras divergências, especialmente no tocante a três aspectos:

- a) o fato de *cê* somente aparecer na função de sujeito pré-verbal (VITRAL, 1996; RAMOS, 1997 e 2000; ANDRADE, 2004) ou em outras posições (COELHO, 1999; BARBOSA, 2005); conseqüentemente,
- b) o *status* de *cê* como clítico (VITRAL, 1996; RAMOS, 1997 e 2000), como forma fraca ainda ou em processo de cliticização (ANDRADE, 2004; CIRÍACO et al., 2004), ou como palavra plena (BARBOSA, 2005);
- c) a expressão da referência, havendo preferência pela forma *cê* para expressar referência indefinida (RAMOS, 1997) ou a preferência de *você* para expressar essa referência (ALVES, 1998; COELHO, 1999).

As divergências acima podem originar-se dos diferentes *corpora* utilizados, ou mesmo devido às diferentes regiões geográficas onde os dados foram coletados. Mas, como os resultados são bastante distintos, faz-se necessário proceder a um número maior de investigações, com *corpora* mais abrangentes, para se tentar dirimir as dúvidas que restaram dos estudos anteriores. Portanto, esta tese pretende avançar nas discussões acima, apresentando dados não contemplados pelos trabalhos citados, ou seja, a inclusão de informantes crianças e adolescentes de Belo Horizonte e de um *corpus* de 1982, a fim de verificar o comportamento das três formas no tempo real. Nos capítulos IV e V, esses dados serão apresentados e, posteriormente, confrontados com os já expostos neste item.

A seguir, será apresentado outro tópico pertinente a este trabalho.

1.2 O estudo da variação e da mudança lingüística

Neste item, serão discutidos temas referentes à heterogeneidade inerente às línguas e suas conseqüências no campo lingüístico e social: os pressupostos da Teoria da Variação, a variação e a mudança lingüística, a mudança em progresso em tempo aparente e em tempo real, e a importância dos fatores extralingüísticos para a variação e a mudança.

1.2.1 A Teoria da Variação

É fato conhecido, há muito tempo, que as línguas são heterogêneas; entretanto, com poucas exceções, as correntes teóricas que estudaram tanto a língua quanto a mudança lingüística não levaram em consideração essa heterogeneidade, concebendo as línguas como um sistema homogêneo, cujo foco de estudo deveria ser sua estrutura interna.

A concepção de língua como um sistema homogêneo, centrado no indivíduo, ou seja, desvinculado do grupo social que usa essa língua em suas interações diárias, foi adotada pela lingüística histórica, pelo estruturalismo e pelo gerativismo.

Por outro lado, houve aqueles que consideraram a heterogeneidade da língua, não a vendo dissociada de sua comunidade de falantes – por exemplo, Meillet, Schuchardt, Sapir e os lingüistas do Círculo Lingüístico de Praga (cf. WEINREICH, LABOV e HERZOG, 1968). Entretanto, esses estudiosos não elaboraram nenhum método para se estudar sistematicamente a complexidade dos dados de fala e para se pesquisar a mudança lingüística.

De acordo com Chambers (1995, p. 15-6), foi William Labov¹⁴, em seu trabalho sobre Martha's Vineyard, de 1963, e, principalmente, em seus trabalhos sobre a língua falada na cidade de Nova York, o primeiro lingüista que reuniu evidências da variação lingüística e que demonstrou que ela é ordenada, padronizada e sistemática. A partir daí, foi possível estudar mais profundamente – e com maior propriedade – os fatores envolvidos na mudança lingüística. Essa nova corrente de estudos da língua ficou conhecida como Sociolingüística¹⁵ e, mais tarde, como Teoria da Variação.

1.2.2 Variação e mudança lingüística

A língua está intimamente relacionada à cultura e ao modo de ser e de viver de uma sociedade; assim, é evidente que a língua deverá ser heterogênea, pois ela irá refletir os contrastes, os confrontos, os desejos de afirmação e de identidade de cada indivíduo e de cada grupo social.

¹⁴ Chambers faz referência a um estudo anterior, de Fischer (1958), que trabalhou com a maioria dos conceitos da Sociolingüística.

¹⁵ O próprio Labov argumenta que é um equívoco falar em Sociolingüística, já que a língua não pode ser concebida fora de seu contexto social. Entretanto, foi este o nome como ficou conhecida essa corrente teórica.

Se as relações sociais dentro de uma determinada comunidade não são simples, a língua refletirá essa complexidade e será, obviamente, heterogênea. Entretanto, essa heterogeneidade pode ser sistematizada, e um dos grandes méritos da Teoria da Variação foi demonstrar que a variação lingüística, existente em todas as épocas, em todas as comunidades de fala e em todos os níveis do sistema lingüístico – fonético, lexical, sintático e semântico –, ocorre de forma organizada e regular. Outra grande contribuição dessa corrente foi o desenvolvimento de técnicas que permitiram sistematizar essa variação. (cf. HINSKENS, HOUT e WETZELS, 2000, p. 20).

As variações lingüísticas podem ser divididas em dois grandes grupos: a variação regional e a variação social, mas é esta que será tomada como objeto deste estudo. Ela está relacionada ao uso de duas ou mais variantes de uma mesma variável por integrantes de uma comunidade. Esse uso, porém, não se dá ao acaso, fortuitamente, mas sim de acordo com fatores lingüísticos (ambiente fonológico, função sintática das variantes, sua posição no sintagma etc.) e extralingüísticos (estilo de fala, contexto, gênero, idade e classe social dos falantes, entre outros), que atuam de forma interdependente em qualquer situação de fala.

Com relação às conseqüências da variação lingüística, como foi dito acima, a língua está direta e intimamente ligada aos grupos sociais que a falam e a sua cultura. Pelo fato de a estrutura social não ser nem uniforme, nem simples, essa heterogeneidade irá se refletir na linguagem, tornando-a igualmente complexa e fazendo surgir variados dialetos¹⁶, usados diferentemente por cada grupo social e em situações de uso distintas. Entretanto, é a língua falada, usada nos contextos informais de interação social, que interessa mais de perto aos variacionistas, tendo em vista que é aí que o vernáculo aparece.

A variação é, segundo os sociolingüistas, a causa primária das mudanças lingüísticas, e ela decorre de fatores de ordem sociocomunicativa ou mesmo lingüística. Para outros pesquisadores, no entanto, a causa das mudanças advém da aquisição imperfeita da língua

¹⁶ Neste trabalho, está-se usando o termo dialeto como sinônimo de língua, uma vez que ele também se constitui num sistema lingüístico que serve à comunicação de seus falantes.

pelas crianças. Marchello-Nízia (2003, p. 372-3) resume as causas mais apontadas pelos pesquisadores:

- *The child's acquisition of its mother tongue: the "desire for clarification" leads to reorganization of the grammatical system;*
- *A first socio-pragmatic factor: pressure of the "high norm";*
- *A second socio-pragmatic factor: the subjective effort of distinction or emphaticisation sought by the speaker. It forces him to create turns of phrase or new forms which surprise or seduce the addressee;*
- *A socio-political factor might have repercussions on the language: orthographic or terminological modifications coming from a political decision might bring a change in the language;*
- *A fifth factor, coming from social pressure, is that of technical and scientific necessity: a new object or concept created in society will be a necessary source of lexical neologism.*

Além desses,

- *The origin of a change can reside in another change occurring elsewhere in the system: it concerns chain shifts;*
- *A change can occur among the grammatical distinctions of the language;*
- *Self-regulation of the grammatical system can occur, considering grammar as a dynamic system. The extension of certain distinctions or features can lead to the creation of new forms;*
- *Finally, at the intersection of internal and external factors, the fact that the speaker sets himself as the referential pivot in the expression of spatial, temporal, and relational relationships can modify the stock of grammatical forms. This appears particularly in the development of deictics, prepositions and adverbs.*

Com relação à primeira causa citada da mudança, a aquisição da língua pela criança, a autora afirma: "This factor is crucial, but it has been proven that it is not the only one, and it must be defined according to other factors." (2003, p. 372). O mesmo já diziam Weireich et al. (1968, p. 188), para quem "linguistic change is transmitted within the community as a whole; it is not confined to discrete steps within the family." Opinião semelhante tem Kroch (2003)¹⁷, para quem "...a estabilidade de algumas línguas ao longo de períodos de

¹⁷ O texto de Kroch lido é uma tradução de seu artigo "Syntactic Change", em Baltin, Mark and Collins, Chris (eds.) (2001). A tradução foi feita por Silvia Regina Cavalcante (UNICAMP), em 2003.

tempo¹⁸, mesmo considerando pequenos detalhes, sugere que a corriqueira aquisição da linguagem não pode em geral ser imprecisa.” (2003, p. 5-6). Para esse autor,

Os mais estudados casos de desvio de mudança sintática de longo tempo são casos mais plausíveis de competição de gramáticas (ou seja, diglossia sintática), na qual as formas em competição podem se diferenciar no registro social, com uma variante vernacular lentamente guiando uma variante escrita fora de uso. (op. cit., p. 6)

Concordando-se com Kroch (2003), Weinreich et al. (1968) e Chambers (1995), entre outros, nesta pesquisa será adotada a perspectiva sociolingüística das causas da mudança, que são a variação e todos os seus fatores correlacionados. Assim sendo, para que a mudança lingüística ocorra, é preciso que duas (ou mais) variantes estejam em competição¹⁹, ou seja, que elas possam substituir-se uma(s) à(s) outra(s) num mesmo contexto lingüístico²⁰. A maioria dessas variações ocorre uma única vez e não volta a acontecer, mas umas poucas, sim, e se espriam ao ponto de serem imitadas por outros integrantes do mesmo grupo social ou de outros grupos. Portanto, a nova forma entrará em competição com a antiga. Numa etapa posterior, uma das variantes vencerá a “luta” e, então, a regularidade será alcançada de novo.

No início, a mudança é lenta²¹, pois a nova forma não está exposta a todos os falantes da comunidade, mas somente a um grupo dela. Nos estágios intermediários da mudança, quando o contato dos falantes de outros grupos sociais com a variante inovadora torna-se mais intenso e esta passa a ser adotada por mais indivíduos, a mudança ocorre

¹⁸ O autor dá o exemplo do japonês, cuja estrutura sintática manteve-se estável do período medieval até o atual, ao contrário da sintaxe inglesa, em que se observam grandes mudanças durante esse mesmo intervalo de tempo.

¹⁹ É ponto pacífico entre os sociolingüistas que toda mudança pressupõe a variação, numa comunidade; entretanto, o contrário não se verifica, ou seja, nem toda variação se transforma em mudança, podendo ocorrer uma variação estável, como será visto adiante.

²⁰ O que foi dito acima aplica-se para variantes fonológicas. Paredes Silva (2003) faz um breve mas interessante resumo sobre uma polêmica travada entre Labov e Henrieta Cedergren, sobre a possibilidade de aplicação desse postulado quando se trata de variantes não-fonológicas.

²¹ Bailey (1973, p. 77) afirma: “A given change begins quite gradually; after reaching a certain point (say, twenty per cent), it picks up momentum and proceeds at a much faster rate; and finally tails off slowly before reaching completion.”

rapidamente, para então, em seu estágio final, ela voltar a acontecer lentamente²². Colocada graficamente, a mudança apresenta a forma de um S (cf. BAILEY, 1973, p. 77, LABOV, 1994, p. 65; CHAMBERS, 1995, p. 203).

Percebe-se, com isso, que a mudança lingüística está intimamente associada às inter-relações que ocorrem com os grupos de indivíduos dentro de uma comunidade. Milroy (1992, p. 33), por exemplo, estabelece:

I shall attempt to show that is the multiple speaker-function of language in use that makes linguistic change possible and suggest that we must look at these speaker-functions if we are to make progress in understanding the nature of language change.

Mais adiante, à página 44, ele acrescenta:

... it is possible that in the course of a conversation innovations may be introduced and, as they may be unfamiliar to some participants, they may be miscomprehend; however, as we have seen, the principles of conversation allow for clarification of these miscomprehensions. For this reason, and for other contextual reasons, the principles of conversation permit linguistic changes to be negotiated between speakers and thus admitted into the language system: it is conversational context that provides the conditions for change to be accomodated.

Assim, nota-se que, para se compreender a mudança lingüística, é preciso empreender uma análise não só do comportamento da(s) variante(s) dentro do sistema lingüístico, mas também proceder a uma análise do comportamento dos membros das comunidades de fala. Para tanto, é preciso descrever esses membros: sua faixa etária, sua escolaridade, sua

²² De acordo com Labov (1994, p. 66), "Two competing sounds are associated with the social values characteristic of the speakers who use them, and the progress of the change is associated with the adoption of the values of one group by members of the other. Thus, at the beginning of the change, speakers of the older form are rarely exposed to the newer form, so that little change or transfer can take place. The rate of change will be greatest when contact between speakers is greatest, that is, at midpoint."

posição social, seu gênero, sua etnia, sua procedência geográfica etc., a fim de determinar a influência desses fatores externos no processo de mudança. Para a explicação da mudança, os fatores lingüísticos e extralingüísticos não podem ser vistos isoladamente. Conforme afirma Labov (1972, p. 02),

The contribution of internal, structural forces to the effective spread of linguistic changes (...) must naturally be of primary concern to any linguist who is investigating these processes of propagation and regularization. However, an account of structural pressures can hardly tell the whole history. Not all changes are highly structured, and no change takes place in a social vacuum. Even the most systematic change chain shift occurs with a specificity of time and place that demands an explanation.

Para se proceder à análise da mudança, levando-se em conta os fatores lingüísticos e extralingüísticos envolvidos, Weinreich et al. (1968, p. 183-7) apresentam cinco princípios, ou cinco problemas, a serem resolvidos. Tarallo (1994, p. 73-74), discutindo esses mesmos princípios, transforma-os em perguntas, as quais seguem abaixo:

Fatores condicionadores. *Quais são os fatores gerais efetivos para a mudança – se é que existem – que determinam e distinguem possíveis mudanças de mudanças impossíveis do sistema e que, ao mesmo tempo, apontam direções de mudança?*

Encaixamento. *Como uma determinada mudança lingüística se encaixa no sistema circundante de relações sociais e lingüísticas?*

Avaliação. *Como os membros de uma determinada comunidade avaliam a mudança lingüística e, em especial, quais são os efeitos dessa avaliação sobre o processo de mudança em si?*

Transição. *Como e por quais caminhos a língua muda?*

Implementação. Por que, quando e onde determinada mudança ocorreu?

O método básico para o estudo da mudança lingüística é a combinação da observação do tempo aparente com o tempo real²³. Entretanto, Labov (1994, p. 63) afirma que

In many speech communities this method [a análise em tempo real] can't be realized (...) In such cases, we are strongly motivated to study the present in greater depth, and see what can be deduced about change in progress from the actual distributions in apparent time.

Portanto, na falta de estudos de tempo real, as questões a respeito das mudanças lingüísticas podem ser respondidas observando-se a mudança em tempo aparente, ou seja, analisando-se a distribuição das formas lingüísticas em variação pelos falantes de uma comunidade divididos por faixas etárias. Uma mudança acontece lentamente e, assim, esses falantes apresentarão diferentes padrões de fala com relação à mudança em questão. Portanto, se a frequência de uso de uma forma lingüística estiver caindo com relação à faixa etária mediana e, principalmente, jovem, poder-se-ia ter aí uma mudança em progresso.

Para um melhor entendimento acerca desse processo, a seguir serão expostas algumas características do estudo da mudança em tempo aparente e em tempo real, assim como as dificuldades que este último apresenta ao pesquisador.

²³ Mais adiante serão tratados estes dois temas: o tempo aparente e o tempo real.

1.2.3 O estudo da mudança em tempo aparente

Foi Labov, em seu estudo de 1963, sobre Martha's Vineyard, quem primeiro utilizou e difundiu o construto do tempo aparente (cf. BAILEY et al., 1991, p. 241), atestando sua validade. Para Labov, “the first and most straightforward approach to studying linguistic change in progress is to trace change in apparent time: that is, the distribution of linguistic variables across age levels.” (1994, p. 45-6). Assim, conforme dito acima, se a variante inovadora for mais comum entre os jovens, pode estar havendo mudança em progresso. Entretanto, a estratificação por idade pode estar indicando também um outro processo, diferente da mudança em progresso: a *gradação por idade (age-grading)*. Segundo Boberg (2004, p. 251),

While this hypothesis [do tempo aparente] is intuitively appealing, it can only be accepted in cases where the opposite possibility, called AGE-GRADING, can be discounted. In age-grading, people do change the way they speak over their adult lifetimes, so that generational differences represent the effect of aging rather than change in the language.

Vê-se, com isso, que a gradação por idade é uma mudança no comportamento lingüístico dos indivíduos que se repete a cada geração. Portanto, numa certa idade – provavelmente as mais jovens – os falantes de uma comunidade usam determinada forma lingüística, normalmente estigmatizada socialmente. Entretanto, quando crescem, a abandonam e passam a usar a mais valorizada. Um exemplo do fenômeno de gradação por idade é a pronúncia da letra *z* pelos habitantes de Ontário (Canadá), citado por Chambers (1995, p. 189). Quando crianças, os falantes pronunciam-na como *zee*, seguindo o padrão dos Estados Unidos. Conforme crescem, abandonam essa pronúncia e passam a adotar a socialmente valorizada, *zed*.²⁴

²⁴ Entretanto, o estudo de Boberg (2004) sobre a mesma pronúncia em Montreal, também no Canadá, evidencia que esse fenômeno é um caso de mudança em progresso, nessa localidade.

Subjacentes às hipóteses do tempo aparente e da gradação por idade estão as idéias de que, respectivamente: i) a gramática do indivíduo não muda ao longo de sua vida; ii) o comportamento lingüístico dos membros de uma sociedade pode mudar no decorrer de sua vida. Nas palavras de Boberg (2004, p. 256-7):

This model [do tempo aparente] assumes that grammars stabilize after acquisition, so that people's linguistic behavior at any point in their adult lives reflects the form of the language that they learned as children. From this model we can infer that generational differences in language reflect ongoing change, whereby each new generation of speakers pushes the change lightly further ahead.

Quanto à gradação por idade, afirma:

The second model is that of the age-grading hypothesis (...). This model assumes that grammars continue to evolve after acquisition, so that people's linguistic behavior in later life may not be an accurate reflection of how they spoke as children or of the state of the language when they acquired it. Postacquisition change has generally been assumed to take the form of increasing conservatism with age, meaning that older people discard some of the innovations they adopted in their youth. These changes are presumably conditioned by age-related shifts in social values and orientations, from innovation and counterculture in youth, to social ambition and increasing conformity in middle age, to conservatism in old age²⁵.

O mesmo autor observa que os níveis mais abstratos da gramática, como a fonologia e a sintaxe, são menos suscetíveis de sofrerem mudanças²⁶ ao longo da vida do falante que os níveis menos abstratos, como o léxico e, talvez, certos aspectos da fonética. Para ele, o nível lexical é muito menos estável e, por isso, muito mais sujeito às mudanças após a

²⁵ Nesse estudo (2004), Boberg propõe uma outra hipótese de variação lingüística baseada na idade, a da Adoção Tardia da Língua (Late Adoption Hypothesis), a qual envolve a adoção, pelos falantes mais velhos, das formas inovadoras que surgem na comunidade. Esse comportamento dos membros adultos, segundo o autor, acelera o processo de mudança.

²⁶ O estudo de Naro e Scherre (1991), sobre a concordância verbal entre informantes do MOBREAL, no Rio de Janeiro, confirma que a gradação por idade pode ocorrer no nível sintático.

aquisição da língua que os outros níveis, já que o léxico está menos propenso a sofrer preconceito. A observação de Boberg confirma o que Chambers (1995) já havia dito sobre a gradação por idade: ela está relacionada basicamente ao vocabulário dos indivíduos.

As evidências de que os falantes podem mudar seu comportamento lingüístico durante sua vida colocam em xeque a validade do construto do tempo aparente, gerando polêmica entre os lingüistas que estudam a mudança. Entretanto, embora haja consenso quanto à existência da gradação por idade e que ela não pode ser descartada dos estudos sincrônicos de mudança lingüística, a hipótese do tempo aparente é confirmada por vários lingüistas, tais como Labov (1972, 1994), Trudgil (1988), Bailey et al. (1991), Chambers (1995), Boberg (2004) etc., sendo a adotada nesta pesquisa.

Apesar dessa confirmação, alguns dos lingüistas mencionados fazem objeções ao estudo do tempo aparente. Bailey et al., por exemplo, observam: “We cannot be certain (...) that apparent-time differences universally reflect language change in progress for all kinds of linguistic features” (1991, p. 263), reconhecendo, com isso, que as mudanças são possíveis nos subsistemas acima dos níveis fonológico e morfossintático. Para Boberg, a gradação por idade “has generally been held to be the main obstacle to apparent-time analyses of age variation in language: we cannot conclude that a change is underway until we have ruled it out.” (2004, p. 257). Até mesmo Trudgil, que afirma que “the apparent-time methodology is an excellent sociolinguistic tool for investigating linguistic changes in progress” (1988, p. 48), observa que “there are (...) cases where differences between older and younger speakers are repeated in every generation, and we need to guard against this eventuality wherever possible when using the apparent-time methodology.” (1988, p. 33)

Portanto, embora a gradação por idade seja relativamente rara, o pesquisador não pode descartá-la imediatamente, tendo, para testar a hipótese do tempo aparente, que recorrer ao estudo da mudança em tempo real, isto é, ao estudo da comunidade em dois períodos de tempo distintos. Qualquer diferença nos resultados pode servir de resposta à pergunta se há ou não mudança em progresso. Labov (1994, p. 46) conclui: “Indeed, such real-time

differences are what we mean by linguistic change, in the simplest and most straightforward definition of the term.”

1.2.4 O estudo da mudança em tempo real

A discussão apresentada até aqui demonstrou que a forma mais eficaz de se testar a validade dos estudos de tempo aparente é o estudo de tempo real. Bailey et al. (1991, p. 243), por exemplo, afirmam que “without real time data, no test of the construct [o tempo aparente] is possible.” Entretanto, apesar da importância do estudo da comunidade em tempo real, sabe-se que nem sempre é fácil ou mesmo possível realizá-lo. As dificuldades são várias: falta de pesquisas anteriores da comunidade em estudo, falta de tempo e/ou de interesse do pesquisador em voltar a ela 10 ou 20 anos depois para coletar novos dados, falta de textos que atestem o vernáculo do lugar etc.

Labov (1994, p. 73-76) afirma que o modo mais simples e eficiente de se fazer uma observação em tempo real é procurar estudos já realizados anteriormente na comunidade que se quer pesquisar e comparar os dados antigos com os mais novos. Porém, esta não é uma tarefa simples, haja vista que a sociolinguística quantitativa desenvolveu-se a partir da década de 60, com Labov, e as diferenças metodológicas entre as pesquisas atuais e os estudos anteriores podem dificultar ou mesmo impossibilitar quaisquer comparações. Para que estas sejam possíveis, são necessários pelo menos dois estudos que tenham as seguintes características (cf. BAILEY et al., 1991, p. 244):

- 1) eles precisam ter sido feitos com o espaço de, no mínimo, uma década ou mais entre eles;
- 2) eles precisam analisar algumas das mesmas variáveis;
- 3) eles precisam estudar a mesma comunidade de fala; e

4) eles precisam incluir dados que tenham sido coletados da mesma maneira.

Um outro método – mais difícil e elaborado que o primeiro – é retornar à mesma comunidade depois de um certo período de tempo e repetir o estudo, seguindo-se a mesma metodologia adotada no primeiro. Entretanto, Labov (1994, p. 73-6) observa que, para que os resultados sejam válidos, a comunidade não pode ter sofrido alterações bruscas, como ter recebido um grande número de imigrantes ou ter-se tornado um pólo industrial. Este segundo tipo de estudo recebe o nome de *Estudo de Tendência (Trend Study)*.

Um terceiro tipo de método é denominado *Estudo de Painel (Panel Study)*. Neste caso, o pesquisador voltará à comunidade e entrevistará os mesmos informantes da primeira pesquisa. Labov (1994, p. 73-6) afirma que, neste tipo de estudo, se corre o risco de haver perda de informantes da primeira pesquisa e de o número de dados coletados ser pequeno; entretanto, as respostas obtidas serão muito valiosas.

Abaixo, o autor expõe as limitações de cada tipo de pesquisa para que sejam caracterizadas as mudanças lingüísticas (LABOV, 1994, p. 84-5):

If we confine our observations to distributions in apparent time, we will detect those conditions that lead to a differentiation of generations: that is, age-grading and generational change. We will not be able to distinguish between these two. Nor will we be able to detect the existence of communal change, since it will be indistinguishable from stability.

A panel study, on the other hand, will detect conditions where the individual either changes or is stable: age-grading and communal change. But a panel study by itself will not differentiate between these two, or between stability and generational change, since it provides no view of the community except through the behavior of the same individuals.

(...)

Since a trend study includes two studies of apparent time, it will both detect unstable behavior of individuals and distinguish stable from unstable communities, differentiating all four of these patterns. In that sense, the trend study would be the best possible approach to gathering data on linguistic change. It has only one limitation: it produces no

information on the behavior of individuals over time. That information, provided by panel studies, is essential for interpreting the many studies where little or no real-time data are available.”

Por fim, Labov sugere que os dados do tempo aparente sejam relacionados com os do tempo real, para se reconstruir a cronologia das várias etapas da mudança e para se correlacionar essa cronologia com as características sociolinguísticas de cada estágio (1994, p. 78).

No Quadro I, abaixo, Labov (1994, p. 83, Tabela 4.1.) faz uma relação entre tempo aparente e tempo real, e mostra os quatro padrões básicos de mudança linguística, levando em consideração a comunidade ou seus indivíduos isoladamente.

QUADRO 1

Padrões de mudança linguística nos indivíduos e na comunidade

	Indivíduo	Comunidade
1. Estabilidade (<i>Stability</i>)	Estável	Estável
2. Gradação por idade (<i>Age-grading</i>)	Instável	Estável
3. Mudança geracional (<i>Generational change</i>)	Estável	Instável
4. Mudança da comunidade (<i>Communal change</i>)	Instável	Instável

Segundo o autor, se não houver mudança no comportamento linguístico do indivíduo ao longo de sua vida, nem no comportamento da comunidade, então haverá estabilidade.

No caso do segundo padrão, ou seja, de os falantes mudarem seu comportamento linguístico ao longo de suas vidas, mas a comunidade como um todo não mudar, o padrão será de *gradação por idade*.

O terceiro padrão – *mudança geracional* – é o tipo de mudança lingüística mais comum, que os autores tratam por mudança em progresso. Aqui, o falante mantém a freqüência de uma determinada variável ao longo de sua vida; entretanto, há um aumento nos valores adotados por outros falantes, incrementados pelas gerações subseqüentes, até haver a mudança.

O quarto padrão – *mudança da comunidade* – é o padrão em que todos os membros da comunidade variam a freqüência de uso de uma variável conjuntamente, ou adquirem novas formas simultaneamente.

Viu-se, neste item, que, embora o construto do tempo aparente seja validado por vários estudiosos da mudança lingüística, faz-se importante proceder ao estudo de tempo real para que a mudança seja mais detalhada e caracterizada. No caso desta tese, foi possível realizá-lo, já que foram utilizados *corpora* que satisfazem às condições estipuladas nas páginas anteriores. No capítulo referente aos procedimentos metodológicos empregados neste estudo, este assunto retornará.

O próximo subitem tratará de outro tema relevante para a compreensão do fenômeno da variação e da mudança lingüística: os fatores extralingüísticos.

1.2.5 Os fatores externos ou extralingüísticos

No item 1.2.2 deste capítulo, viu-se que, para se entender a mudança lingüística, é necessário levar em conta, além dos fatores lingüísticos, alguns aspectos sociais da comunidade lingüística: são os fatores extralingüísticos ou externos.

Os fatores extralingüísticos que podem desempenhar papel preponderante tanto na variação quanto na mudança lingüística são: idade, gênero, classe social, nível de escolaridade, ocupação profissional, etnia, (LABOV, 1972), região geográfica (urbano/rural) e posição

nas redes sociais²⁷ (LABOV, 2002). Apenas os três primeiros serão analisados, nesta pesquisa, tendo em vista que esta se limitou à cidade de Belo Horizonte e que nela há informantes crianças. Quanto ao fator etnia, analisá-lo num país como o Brasil é uma tarefa bastante complexa, razão pela qual não foi levado em conta. Quanto às redes sociais, elas não foram analisadas neste trabalho, já que, na coleta de dados dos dois *corpora* que o compuseram, não se considerou esse fator.

A seguir, serão discutidas algumas questões importantes sobre os fatores extralingüísticos analisados.

1.2.5.1 Idade

Os estudos que tratam da mudança lingüística atestam que as línguas variam de acordo com a faixa etária dos falantes, o que implica que os indivíduos que constituem uma comunidade apresentam, conforme sua idade, diferentes características de fala, dependendo do(s) grupo(s) a que eles pertencem e a que se associam, da habilidade que possuem para pôr em prática as normas lingüísticas da sociedade etc. (Labov, 2001, p. 101).

Segundo Labov (2001), para se proceder à divisão do *continuum* de idade de um indivíduo em fases, deve-se levar em conta os diferentes estágios da vida em uma sociedade. Na americana atual, o autor cita seis (2001, p. 101):

the alignment to the pre-adolescent peer group (8-9) [anos], membership in the pre-adolescent peer group (10-12), involvement in heterosexual relations and the adolescent group (13-16), completion of secondary schooling and orientation to the wider world of work and/or college (17-19), the beginning of regular employment and family life (20-29), full engagement in the work force and family responsibilities (30-59), retirement (60s).

²⁷ Mais adiante esse tema será abordado.

Entretanto, Labov não dá maiores detalhes a respeito de cada uma dessas fases. Por sua vez, Chambers (1995) estabelece três períodos lingüisticamente significantes na vida dos indivíduos: a infância, a adolescência e a vida adulta:

1.2.5.1.1 A infância

A criança, nos primeiros anos de vida, tem seu círculo de relações restrito aos familiares e aos amigos próximos. Assim, a aquisição de sua língua²⁸ e o uso das normas lingüísticas se dão sob a influência dos adultos mais próximos a ela, e esta passa a falar como eles. Entretanto, à medida que a criança cresce e seus contatos aumentam e se diversificam, o que geralmente acontece, nos países industrializados, quando ela entra para a escola – em torno de quatro ou cinco anos – (CHAMBERS, 1995, p. 168), a criança passa a sofrer a influência de outros adultos - principalmente professores - e dos amigos. A partir daí, iniciam-se para ela os grupos e as redes sociais a que estará permanentemente associada e, assim, sua linguagem passará a ser influenciada pelos amigos²⁹. Chambers (1995, p. 159) resume esse fato: “Although the family circle normally provides the first speech models for infants, within a few years it is replaced by a more significant one, the circle of friends.”

1.2.5.1.2 A adolescência

No mundo ocidental, a adolescência, que vai dos treze aos vinte e um anos³⁰ (CHAMBERS, 1995, p. 172), é um período da vida marcado pela ousadia, rebeldia e contestação dos valores vigentes na sociedade. É uma fase também de insegurança pessoal

²⁸ Chambers (1995, p. 163) chama a atenção para uma capacidade peculiar dos indivíduos dessa faixa etária: a de aprender e falar uma outra língua ou um outro sotaque como se fosse um falante nativo: “Children seven or under will almost certainly acquire a new dialect perfectly.(...) People over the age of 14 almost certainly will not. (...) In between the ages of seven and 14, people differ in their ability to acquire accents, apparently unpredictably.”

²⁹ Conforme foi visto anteriormente, existe uma polêmica acerca dessa posição: enquanto alguns autores postulam que a mudança é originada pela aquisição imperfeita da linguagem pela criança, Weinreich et al. (1968), Chambers (1995), Marchello-Nizia (2003) e Kroch (2003), por exemplo, defendem que a heterogeneidade e a variação são causas principais das mudanças.

³⁰ Não existe um consenso a respeito da idade em que termina a adolescência. Por exemplo, para Labov (2001, p. 101), na sociedade americana, a adolescência vai até os 19 anos.

e de incertezas, em que os indivíduos buscam a independência dos pais e a auto-afirmação como pessoas e como cidadãos.

Para compensar a insegurança decorrente dessa independência, o adolescente é marcado por um forte sentimento de união e de identificação com seus pares. Assim, para ele, o grupo geralmente se torna mais forte, mais influente, enfim, mais importante que os pais.

A rebeldia e a irreverência do comportamento, por sua vez, espelham-se na linguagem. A adoção de gírias serve não só como símbolo da adolescência, mas também como marca registrada dos membros de determinados grupos; elas distinguem quem pertence a eles e quem está de fora, e, a fim de preservar essa distinção, elas mudam com rapidez. Sociolinguisticamente, a combinação entre esse desejo de separação dos mais velhos e a união com os pares faz com que os adolescentes sejam o principal agente de variação e mudança lingüística (CHAMBERS, 1995; LABOV, 1994).

Antes de passar ao próximo item, referente à fase adulta, é importante aprofundar um pouco mais as observações postas por Chambers (1995) quanto à infância e à adolescência. Segundo Labov (1994, p. 47), “adolescents and preadolescents, are the leading edge in the progress of a sound change, and (...) any study must be sure to include extend recordings of their speech.” Assim, nesta pesquisa foram incluídos dados de informantes dessas duas faixas etárias.

Entretanto, para a análise dos resultados obtidos através deles, é necessário conhecer mais a respeito dessas duas faixas etárias, tais como seu raciocínio, afetividade e socialização, os quais também embasarão as análises posteriores. Esse detalhamento não é dado nem por Chambers (1995), nem por Labov (2001), mas existem claras e importantes diferenças que marcam os indivíduos dessas faixas de idade, sendo que é preciso tê-las em mente, para se proceder às análises que se seguirão. Para tanto, foi necessário buscar em outro autor essas informações: Piaget (2003 [1976]). Abaixo se encontra um resumo da proposta do autor sobre este assunto.

Piaget (2003), ao estudar o desenvolvimento da inteligência, da vida afetiva e das relações sociais das crianças desde seu nascimento até a adolescência, estipula-lhes seis estágios: o primeiro, dos reflexos e das primeiras tendências instintivas e das primeiras emoções; o segundo, dos primeiros hábitos motores, percepções organizadas e sentimentos diferenciados; o terceiro, da inteligência senso-motora ou prática e das primeiras fixações exteriores da afetividade. Esses três primeiros estágios acontecem até um ano e meio ou dois anos de vida.

O quarto estágio é o da inteligência intuitiva, dos sentimentos inter-individuais espontâneos e das relações sociais de submissão ao adulto (acontece dos dois aos sete anos); o quinto, das operações intelectuais concretas – o começo da lógica – e dos sentimentos morais e sociais de cooperação (dos sete aos onze-doze anos); o sexto, das operações intelectuais abstratas, da formação da personalidade e da inserção afetiva e intelectual na sociedade (dos onze-doze aos quinze anos) (Piaget, 2003, p. 15).

Explicitando um pouco mais as faixas etárias pertinentes a este estudo, ou seja, dos sete anos à adolescência, tem-se:

1) O quinto estágio: a infância (de sete a onze-doze anos)

- a) *O pensamento e o raciocínio.* Nesta fase, as crianças adquirem as noções de conservação e de permanência de substâncias, peso, volume, comprimento, tempo, espaço e velocidade, entre outros. De acordo com o autor (2003, p. 48), é depois dos sete anos que um espaço racional começa a se construir, ou seja, elas passam da intuição para as operações racionais. É a fase do pensamento concreto.
- b) *A afetividade.* Aparecem nas crianças desta fase novos sentimentos morais, como o respeito mútuo, o acatamento de regras estabelecidas em comum acordo, a honestidade, o companheirismo e o sentimento de justiça. Esse sentimento de justiça em relação aos companheiros, aponta o autor, é tão grande que influencia as

relações entre elas e os adultos, podendo modificar suas atitudes em relação aos pais.

- c) *A socialização.* As crianças, nesta fase, tornam-se capazes de cooperar com os amigos, distinguem seu ponto de vista do ponto de vista dos outros, já argumentam e pensam antes de agir. Socialmente, já sabem estipular e obedecer a regras comuns ao grupo no qual estão inseridas.

2) O sexto estágio: a adolescência (dos onze-doze anos até os quinze anos)

a) *O pensamento e o raciocínio.* Por volta dos onze ou doze anos, há uma mudança consideravelmente importante no pensamento dos adolescentes: a passagem do pensamento concreto para o formal, ou seja, do pensamento centrado na própria realidade concreta das crianças para o pensamento hipotético-dedutivo, em que os adolescentes abstraem do real hipóteses e idéias. A partir do momento em que o pensamento formal começa, os adolescentes são capazes de libertar-se do real e, conseqüentemente, de fazer reflexões e construir teorias. De acordo com o autor, a adolescência é o último período do desenvolvimento mental (2003, p.63).

b) *A afetividade.* A personalidade começa a ser construída no fim da infância (dos oito aos doze anos) e, na adolescência, ainda está em formação. Devido a isso, sua personalidade ainda apresenta desequilíbrios, que farão com que os adolescentes sejam, ao mesmo tempo, egocêntricos e também extremamente altruístas, e tenham sentimentos generosos, querendo transformar a sociedade e o mundo dos adultos, melhorando-os.

c) *A socialização.* A conturbação dos adolescentes tem implicações na sua vida social. Ao mesmo tempo em que eles parecem ser anti-sociais, porque se encerram em seu mundo, o convívio com outros adolescentes acaba por fortalecer os laços de amizade e de cooperação mútua, mesmo que haja alguma

divergência de opinião entre eles. Assim sendo, nesta fase, os seu grupo de colegas e amigos tem imensa importância para eles.

Como será visto na apresentação e na análise dos dados de crianças e adolescentes desta pesquisa, as observações acima irão ajudar na compreensão dos resultados obtidos para essas duas faixas etárias.

Com base no resumo acima, nesta tese adotou-se a divisão dada por Piaget. Assim, os informantes crianças e adolescentes ficaram divididos em dois grupos:

- (i) faixa etária I, de oito³¹ a onze anos;
- (ii) faixa etária II, de doze a quinze.

A partir deste ponto, pode-se, então, retornar à descrição das faixas etárias estipuladas por Chambers (1995), com a fase adulta.

1.2.5.1.3 A fase adulta

De acordo com Chambers (1995, p. 177), depois dos anos turbulentos da adolescência,³² os indivíduos tendem a buscar a realização profissional e afetiva. Seus interesses, então, passam a ser a carreira, o casamento e a família. Os jovens adultos também estabelecem suas preferências em termos de roupas, recreação, condição de vida, política e ética profissional. Essas preferências podem mudar um pouco, mas quase nunca mudam radicalmente.

Assim como acontece com os adolescentes, essas características dos jovens adultos marcam sua linguagem. Portanto, vê-se que alguns deles parece moldarem sua forma de falar à sua ambição profissional: quanto mais desejam ascender na escala social, quanto

³¹ Labov (1994, p. 49), baseando-se em seus próprios estudos e de outros pesquisadores, estipula que a idade mínima para a inclusão de crianças nas pesquisas sociolinguísticas é de oito anos. Assim sendo, nesta pesquisa, será acatada também a proposta de Labov.

³² Conforme visto, para Chambers (1995), a adolescência vai dos treze aos vinte e um anos.

mais aspiram a um alto cargo ou a uma profissão valorizada socialmente, mais tentam adequar sua fala à linguagem padrão. É o que Chambers (1995, p. 178) chama de “pressão do mercado” (*market pressure*)³³: a pressão que certas profissões exercem sobre os indivíduos para que padronizem sua fala. O autor lembra que essa pressão não está diretamente relacionada à classe social. Uma pessoa – de qualquer camada social – que deseje ascender socialmente tentará ajustar sua forma de falar à linguagem mais valorizada socialmente.

Passada essa fase, os adultos mais velhos “cristalizam” sua fala, e não se verificam mudanças significativas em sua linguagem. Nas palavras de Chambers (1995, p. 184):

For the stages of life beyond young adulthood, our best evidence indicates that once the features of the sociolect are established in the speech of young adults, under normal circumstances those features remain relatively stable for the rest of their lives. Even when linguistic changes take root in the speech of younger people in the same community, the older people usually remain impervious to it, or nearly so.

O que foi dito acima força a retomada de um ponto discutido anteriormente, sobre o estudo da mudança lingüística. O exemplo mencionado, de mudanças ocorridas devido a pressões sociais e/ou profissionais, apresenta-se como uma contra-prova à hipótese de que os indivíduos cristalizam sua forma de falar aos 15 anos, mais ou menos (NARO, 2003, p. 44).

A idéia de cristalização da língua nos indivíduos é assumida pela maioria dos lingüistas e, por isso, é conhecida como “hipótese clássica”. De acordo com essa hipótese, baseada na psicologia desenvolvimentista, “a gramática do indivíduo não pode sofrer mudanças significativas porque o acesso aos dispositivos cognitivos que possibilitam a sua manipulação (a chamada faculdade da linguagem) fica bloqueado” (NARO, 2003, p. 44).

³³ O autor baseia-se na idéia de “mercado lingüístico”, “marché linguistique”, de Pierre Bourdieu, presente em seu trabalho “L’économie des échanges linguistiques”, de 1977.

Entretanto, há dados empíricos que mostram que a hipótese clássica não se sustenta: além dos casos de pressão social ou profissional, de que fala Chambers (1995), estudos de painel levados a cabo por pesquisadores brasileiros têm mostrado que alguns indivíduos realmente apresentam mudança de seu comportamento lingüístico durante sua vida, enquanto que outros o mantêm (BRAGA, 2003; RONCARATI, 2003; NARO e SCHERRE, 2003; OMENA, 2003; etc.)

Apresentar aqui essa discussão se faz importante, na medida em que cada uma dessas posições valida um pressuposto teórico essencial para o estudo da mudança lingüística. A hipótese da cristalização da língua aos quinze anos de idade, aproximadamente, permite estudar a mudança basicamente através do tempo aparente; a hipótese da variação ao longo da vida do falante aponta para a dúvida com relação ao construto do tempo aparente e, por conseguinte, para a necessidade de se fazer um estudo de tempo real, principalmente um estudo de painel, para se estudar a mudança.

Neste trabalho, apesar de se acreditar na “hipótese clássica”, foi realizado, como já se mencionou, o estudo em tempo real, para se complementar o estudo em tempo aparente, embora não tenha sido possível proceder ao estudo de painel. Com isso, pretende-se descrever com maior clareza o caminho percorrido pelas formas *você*, *ocê* e *cê* na comunidade belo-horizontina de 1982 até os dias atuais.

1.2.5.2 - Gênero

Antes de se iniciar este tópico, é necessário esclarecer a adoção, nesta pesquisa, do termo “gênero”. Dos autores consultados para a elaboração deste item, apenas James Milroy (1992) e Guy (1988) usam o termo “sexo”, quando diferenciam homens e mulheres. Os demais autores preferem usar o termo “gênero”. Essa discussão não é meramente terminológica, mas envolve a questão dos papéis que homens e mulheres exercem na sociedade. Sundgren (2001, p. 117) afirma:

It is essential to distinguish between sex as a biological category and gender as the social construction of sex. If you correlate variables with sex there is an obvious risk of looking at gender simply as a binary opposition and of letting sex represent gender without due attention to how the gender roles are construct.

Esse questionamento indica que, ao se adotar o termo *gênero*, dever-se-á observar atentamente esses papéis, tanto os masculinos quanto os femininos. Concordando com o fato de que, em sociedades desiguais, homens e mulheres desempenham papéis diferentes, o que incidirá sobre seu comportamento lingüístico, neste trabalho será usado o termo “gênero”.

Voltando agora ao tema deste subitem, é fato bastante conhecido que o gênero desempenha um importante papel entre os fatores externos envolvidos na variação e na mudança lingüística. Isso porque há um extenso número de trabalhos, na literatura sociolingüística, que comprova um comportamento lingüístico diferenciado entre homens e mulheres.

Em se tratando de variáveis estáveis³⁴, por exemplo, segundo Paiva (2003), Gordon (1997), Milroy, L. (1987), Milroy, J. (1992), Chambers (1995), Sundgren (2001) e Labov (2001; 2002), as mulheres têm a tendência de usar mais as formas padrão da língua – evitando as estigmatizadas e não-padrão – que os homens de mesma classe social e nas mesmas circunstâncias. Labov (2001, p. 266) estabeleceu esse padrão em forma de um princípio, assim transcrito: “For stable sociolinguistic variables, women show a lower rate of stigmatized variants and a higher rate of prestige variants than men”³⁵.

O autor, entretanto, faz a ressalva de que, para que esse princípio seja válido, é preciso que as mulheres tenham acesso às normas-padrão da língua. Por isso, esse princípio geralmente não se aplica às camadas sociais mais baixas.

³⁴ De acordo com Chambers (1995, p. 107), “stable variables are those that are well established as indicators in a community and are not undergoing change”.

³⁵ Entretanto, também há casos de homens privilegiando as formas de prestígio mais que as mulheres, principalmente nos países muçulmanos. (LABOV, 2001, p. 274)

A restrição do autor remete a um fato que já se tornou consenso entre os autores citados acima: o fator gênero – assim como qualquer outro fator extralingüístico - não pode ser analisado isoladamente³⁶. De acordo com a variável estudada, o gênero irá atuar de forma mais ou menos significativa, dependendo de sua relação com outros fatores, como o estilo usado, a classe social e a idade dos falantes. Assim, os estudos sociolingüísticos mostram que as diferenças entre homens e mulheres tendem a ser maiores na classe média baixa³⁷, no estilo de fala mais cuidado e nas faixas etárias mais avançadas. Entre os jovens, as diferenças são reduzidas ou até inexistentes.

As diferenças entre os gêneros acontecem não só na variação estável, mas também nas mudanças em progresso. Labov (2001) divide-as em dois tipos: “change from above” e “change from below”. Segundo ele, nas mudanças “from above”,

They [as variáveis lingüísticas] may take the form of importation of a new prestige feature from outside the speech community, or the redistribution of forms with known prestige values within the community. Changes from above take place at a relatively high level of social consciousness, show a higher rate of occurrence in formal styles, are often subject to hypercorrection, and sometimes form overt stereotypes similar to stable sociolinguistic variables. (2001, p. 274)

Já as mudanças “from below”, também de acordo com o autor, são

... the primary form of linguistic change that operates within the system, below the level of social awareness. These include the systematic sound changes that made up the major mechanism of linguistic change. (LABOV, 2001, p. 279)

³⁶ A interdependência entre os fatores externos não acontece com os fatores internos: “While internal constraints on variation are typically independent of each other, it is normal to find strong interaction among the external factors.” (LABOV, 2002, p. 11)

³⁷ Labov (2001, p. 272) afirma: “The tendency to avoid stigmatized forms and prefer prestige forms is greatest for the women of the lower middle class, and is often minimal for the lower class and upper middle class. (...) In general, the second highest status group shows the greatest gender differential, along with the highest degree of linguistic insecurity and the sharpest slope of style shifting.”

Em ambos os casos de mudanças – “from above” e “from below” –, as mulheres estão na liderança, desde que se trate de formas não-estigmatizadas na sociedade. Elas dão início ao processo e, posteriormente, os homens adotam a forma inovadora. Esse “atraso” dos homens na adoção da forma inovadora é de uma geração. Exemplificando este fato com o estudo do desvozeamento do [ʒ] em Buenos Aires, feito por Wolf e Jiménez (1979), Labov (2002, p. 13) estabelece³⁸:

Males 24 to 15 years old approximate the values of women 38 to 55 years old, and 12 year old males have values in the range of women 24 to 35 years old. It is equally important to note that as the change nears completion the differences between males and females seem to disappear.

Em resumo, as mulheres, ao mesmo tempo em que privilegiam as formas de prestígio, padrão, também lideram as mudanças – se não se tratar de formas estigmatizadas socialmente –, ou seja, ao mesmo tempo em que elas assumem uma posição conservadora, também assumem uma posição inovadora na língua. Essa atitude contraditória feminina Labov denominou “Paradoxo do Gênero”, assim descrito: “Women conform more closely than men to sociolinguistic norms that are overtly prescribed, but conform less than men when they are not.” (2001, p. 293)

O que foi dito até agora a respeito das diferenças do comportamento de falantes homens e mulheres é ponto pacífico na literatura sociolingüística: comprovam-no vários estudos levados a cabo em diversas partes do mundo, citados principalmente por Chambers (1995) e Labov (2001). Mas, por outro lado, há muita divergência entre os autores quando se tenta explicar o porquê dessas diferenças.

³⁸ Labov (2002, p. 15) explica esse fato com base no sentimento feminino de identidade e também na assimetria da aquisição da linguagem com respeito ao gênero.

Tradicionalmente, as tentativas de explicação para o comportamento lingüístico de homens e mulheres se baseiam em aspectos sociais: as mulheres, mesmo conseguindo sua emancipação social e econômica em algumas sociedades,³⁹ ainda pertencem ao segundo grupo de poder e *status*, cabendo-lhes a tarefa de educar os filhos (LABOV, 2001, p. 262). Assim, as mulheres, conscientes de sua posição de grupo de menor poder socioeconômico, tentam, através da adoção das normas lingüísticas de prestígio, superar ou mesmo diminuir sua desvantagem em relação aos homens em outras áreas.

Posicionamento semelhante tem Sundgren (2001), em cujo estudo sobre Eskilstuna, Suécia, observou que a construção do gênero, ou seja, a distribuição dos papéis masculinos e femininos é diferente nas sociedades modernas, mesmo naquelas que parecem mais igualitárias. Acerca dessa questão, diz ela (SUNDGREN, 2001, p. 121-2):

Women's linguistic competence becomes a means of power in an unequal society where men have more social and political power. In the social norm for a woman, from whom it is more expected than of a man to make an attractive appearance, language is also include as a means of gaining respect and status. A woman who uses many non-standard variants is probably judged more negatively than a man who does the same.

Gordon (1997) é outra autora para quem as causas da diferença de gênero são sociais. Em sua pesquisa, a autora constata que a sociedade neozelandesa ainda associa, como há mais de um século, a linguagem e a vestimenta da classe social mais baixa à promiscuidade e, conseqüentemente, a um julgamento moral negativo. Assim, ao preferir as formas-padrão da língua, as mulheres não estão pretendendo ser mais do que elas são, mas evitar que sejam associadas a um estereótipo negativo que sofrem as mulheres da classe baixa.

³⁹ Labov (2001, p. 279) ressalta que as mudanças que se verificaram nos últimos anos com as mulheres nas sociedades modernas não foram suficientes para que seu comportamento lingüístico mudasse rapidamente, haja vista que a posição de debilidade da mulher se insere num padrão cultural bastante antigo.

Para Chambers (1995), por outro lado, nas sociedades industrializadas, os papéis masculinos e femininos se sobrepõem, se equivalem. Com isso, as diferenças lingüísticas entre homens e mulheres se devem mais a causas biológicas que sociais. Reportando-se a vários estudos psicológicos e neurológicos, Chambers lista uma série de vantagens e habilidades lingüísticas das mulheres em relação aos homens, como a fluência verbal, a compreensão auditiva de material escrito e falado, o vocabulário etc. (1995, p. 132)⁴⁰. Essas vantagens seriam uma decorrência da compleição do cérebro feminino: este parece ser mais globalmente organizado para funções específicas, enquanto o cérebro masculino é altamente lateralizado, com as funções verbais no hemisfério esquerdo e as funções espaciais no hemisfério direito.

O autor conclui explicando o fato de as habilidades verbais das mulheres serem maiores que as dos homens, em qualquer comunidade: “(...) if it is true that women generally have an innate verbal facility different from men, that difference should override cultural, political or social bound.” (1995, p. 135). É isso o que os dados expostos pelo autor mostram.

Resumindo o que foi dito até agora a respeito do fator “gênero”, há um consenso entre os autores de que homens e mulheres têm comportamento lingüístico diferente: estas privilegiam as formas-padrão da língua e, quando se trata de mudança lingüística envolvendo variantes não-estigmatizadas, estão na liderança. Por outro lado, as causas dessas diferenças dividem os estudiosos em dois grupos: os que acreditam que elas tenham origem social e Chambers, que aponta primordialmente uma causa biológica para elas.

Apesar de este assunto específico – as causas das diferenças – não fazer parte dos objetivos centrais deste estudo, pensar no problema auxilia a interpretação dos resultados e a análise de como os fenômenos lingüísticos estão inseridos nos complexos grupos sociais.

⁴⁰ Chambers (1995, p. 132) afirma, entretanto, que as vantagens femininas nesses testes são reduzidas e não devem ser exageradas.

1.2.5.3 - Classe Social

Dentre os três fatores extralingüísticos analisados neste trabalho, a classe social é o mais complexo e, portanto, o mais difícil de ser definido. Entende-se por classe social a estratificação da sociedade tomando-se por base semelhanças quanto a alguns aspectos. Entretanto, não há um consenso entre os lingüistas sobre os fatores que devam ser levados em conta na determinação das diferentes classes sociais. Os mais comuns, citados por Macaulay (1976, p. 184) e Labov (2001, p. 60-7), por exemplo, são ocupação, educação, residência, tipo e área de moradia, renda e mobilidade social.

Todavia, segundo Macaulay (1976), um grande número de indicadores, em vez de facilitar, pode dificultar a análise das classes sociais. Assim, em seu estudo sobre Glasgow, a ocupação foi o critério adotado para a distribuição dos indivíduos em classes, pelo fato de o autor acreditar que ela seja o principal indicador das classes sociais (1976, p. 174)⁴¹. Além desses fatores, nas distinções de classe, outros aspectos também são levados em conta, como a vestimenta, maneiras, recreação e gostos de um modo geral.

Nas nações industrializadas, são basicamente três as classes sociais: alta, média e baixa. Na primeira, encontram-se as pessoas que herdaram fortuna e privilégios; na segunda, os trabalhadores não-manuais, também chamados *white collars*, que geralmente têm um maior nível de escolaridade, exercem funções de supervisão ou mando e têm salários mais altos; e, na terceira, os trabalhadores manuais, conhecidos por *blue collars*, que normalmente têm menor escolaridade, exercem funções subordinadas e têm salários mais baixos.

Entretanto, nessas sociedades altamente complexas, foi preciso fazer uma distinção mais pormenorizada das classes média e baixa, subdividindo-as também em alta, média e baixa. O quadro abaixo, de Chambers (1995, p. 37, Table 1), evidencia a correlação entre as subdivisões das classes sociais com as respectivas profissões.⁴²

⁴¹ Chambers (1995, p. 47), citando esse mesmo artigo, afirma que a hipótese de Macaulay precisaria ser confirmada em localidades fora dos países industrializados.

⁴² Algumas denominações das ocupações do quadro, se traduzidas para o português, apresentariam discordância quanto ao seu sentido original, como “*Professional*”, que engloba profissionais qualificados e

QUADRO 2

Correlação classe social x ocupação

MIDDLE CLASS (MC)	Upper (UMC) Middle (MMC) Lower (LMC)	Owners, directors, people with inherited wealth Professionals, executive managers Semi-professionals, lower managers
WORKING CLASS (WC)	Upper (UWC) Middle (MWC) Lower (LWC)	Clerks, skilled manual workers Semi-skilled manual workers Unskilled labourers, seasonal workers

As divisões e subdivisões acima se aplicam aos países industrializados. Em se tratando de países em desenvolvimento, como o Brasil, as classes não são tão estratificadas, o que implica que não se podem adotar esses mesmos critérios sem que lhes sejam feitas algumas adaptações.

Para se definirem essas classes, em países como o nosso, devem ser tomados não só indicadores como ocupação, mas também origem social, renda, acesso a bens materiais e culturais, grau de inserção em redes sociais etc. (cf. MOLLICA, 2003, p. 29), sendo eles bastante fluidos e complexos. Nesta pesquisa, foram adotadas as seguintes divisões: classe média e classe baixa. Os critérios adotados para se estabelecerem essas divisões nos *corpora* seguem o que diz Mollica (op. cit.) e serão apresentados no Capítulo III, sobre os procedimentos metodológicos empregados neste estudo.

Porém, independentemente do país e de seu grau de desenvolvimento, é fato que as diferenças sociais trazem conseqüências importantes para a variação lingüística, como apontam diversos autores. Sabe-se que a linguagem chamada culta, a variante padrão da língua, está estreitamente relacionada às pessoas mais escolarizadas e/ou pertencentes às classes mais altas da sociedade, enquanto a linguagem não-padrão se apresenta como a

também os chamados profissionais liberais. Assim, optou-se por manter o quadro em inglês, para que ele fosse fiel à idéia do autor.

marca das classes mais desfavorecidas socialmente. Aquela é associada às normas de prestígio e ao poder social de seus falantes, sendo requerida pelos postos de trabalho mais altos⁴³ e pelas situações formais de fala. É a divulgada pela imprensa e pelos órgãos de públicos oficiais. A linguagem não-padrão, por sua vez, sofre a mesma discriminação que sofrem as pessoas que a falam, sendo alvo de correção pela escola.

Dessa estreita relação entre língua e classe social resultam conseqüências lingüísticas e sociais importantes, já observadas por vários estudiosos da língua.

O primeiro fato é o *code-switching*⁴⁴, ou seja, a mudança das variantes não-padrão para as padrão, quando os falantes se encontram em uma situação formal de fala ou quando escrevem, o que implica o julgamento – e a aceitação – das variantes-padrão como corretas, de prestígio, ou, pelo menos, mais adequadas que as variantes não-padrão, o que demonstra uma avaliação muito comum – no sentido de ser compartilhada por muitos – no que diz respeito a prestígio e a normas de comportamento (cf. GUY, 1988, p. 50).

Mas esse fato também traz uma conseqüência particularmente importante para os estudos sociolingüísticos: uma tendência dos informantes de abandonar, quando são entrevistados, o vernáculo, preferindo a linguagem padrão, tida como a mais adequada em situações de entrevista, ou seja, em situações mais formais. Tem-se aí, então, o Paradoxo do Observador, de que trata Labov (1972, p. 209), com o qual o pesquisador tem de lidar para conseguir bons dados. No capítulo sobre os procedimentos metodológicos desta Tese, este assunto será retomado.

Se a primeira conseqüência da relação direta entre língua e classe social é a adoção temporária da linguagem de prestígio, a segunda conseqüência é a adoção mais constante da linguagem padrão, devido à ascensão social. Os indivíduos que ascendem socialmente

⁴³ O que está sendo dito está associado ao que se chamou anteriormente de “Mercado Lingüístico”.

⁴⁴ O termo *code-switching* é muito empregado em estudos sobre bilingüismo, referindo-se à mudança de línguas (códigos) que certos indivíduos fazem, em determinadas situações. Essas situações estão estreitamente relacionadas aos costumes adotados em comunidades ou famílias bilingües, sendo orientadas social e discursivamente. Entretanto, aqui o termo *code-switching* está sendo empregado em sentido um pouco diverso, de mudança de variantes.

tendem a estabelecer relações com os integrantes de sua nova classe, afastando-se de seus antigos companheiros. Com isso, esses indivíduos também abandonam as marcas de seu dialeto de origem, passando a adotar as de seu novo grupo social, inclusive usando menos variantes não-padrão que os integrantes da nova classe. Chambers (1995, p. 57) atribui a esse fato uma maior atenção à linguagem por parte daqueles que ascendem, ou seja, trata-se de um caso de hipercorreção.

Por outro lado, esse “abandono das raízes” traz várias conseqüências para quem ascende. Nas palavras de Chambers (1995, p.48):

Various stereotypes are associated with upward mobility: people who rise might be classed, depending upon who is judging them, as people who put on airs or who forget their roots (...) These stereotypes suggest the discomfort that can be involved in breaching social barriers, and to some extent they may function as mild deterrents or vague social pressures to keep people “in their place”.

Essa propensão social explica por que a variante padrão da língua, sendo considerada o veículo legítimo do poder social, político e econômico de uma nação, não é adotada por todos os seus membros, principalmente pelas classes destituídas desses poderes. A resposta a essa questão é, na verdade, a terceira conseqüência da relação entre língua e classe social: a solidariedade entre seus membros. Devido ao fato de que a linguagem é uma marca bastante forte de uma classe, o uso da variante padrão por um integrante da classe trabalhadora pode ser visto por seus pares como esnobismo, afetação e mesmo um ato de hostilidade com sua família, amigos e vizinhos (cf. GUY, 1988, p. 51). Este autor conclui dizendo que:

The fact that nonstandard speakers have not historically rushed to adopt the dominant linguistic norms shows that these do not have the same force for all classes, and that different classes may have different social and linguistic goals.(1988, p. 51)

Em vista do que foi exposto, percebe-se que as barreiras sociais e as conseqüentes barreiras lingüísticas não só dificultam a mobilidade social, mas também explicam a manutenção da variação dialetal numa mesma comunidade de fala, pois a adoção de certas normas lingüísticas envolve a solidariedade ou não de um indivíduo ao grupo a que pertence.

Por outro lado, a variação lingüística leva a mudanças, mas, para isso, é preciso haver contato entre os diferentes grupos sociais, e, como afirma Chambers (1995, p. 49), “*partly by choice and partly by chance, then, the social classes are not in constant or close contact.*”

Entretanto, conforme afirma Labov em vários de seus estudos, há duas subdivisões de classes que, em função de seu trabalho, estão em permanente contato com as classes imediatamente acima e imediatamente abaixo da sua: a trabalhadora alta e a classe média baixa. O Quadro 02, acima, revela que, nessas classes, estão pessoas cujas ocupações propiciam esse contato. Devido a isso, são essas classes, segundo o autor (por exemplo, Labov, 1972, p. 172; 2001, p. 31-2; 2002, p. 06), que lideram as mudanças lingüísticas consideradas típicas, ou seja, *from below*, desenvolvidas internamente, sem envolver empréstimos.

Por outro lado, os agentes das mudanças *from above* são as classes mais altas. São estas também que oferecem resistência às mudanças *from below*, com o objetivo último de assegurar seu *status quo*. Nas palavras de Guy (1988, p. 59-60):

(...) conflicts over the sociosymbolic significance of linguistic innovations are seen as a consequence of the conflicting interests of different social classes. The working class (broadly defined to include the lower-paid and lower-status level of the middle class such as secretaries, clerical workers, book-keepers, etc.) are the basic source of untargeted innovations. (...) Higher-status groups, however, not belonging to the working class and wishing to defend their social position, will naturally resist and denigrate such innovations.

Portanto, o conservadorismo lingüístico das classes altas nada mais é que o reflexo de seu conservadorismo social e político.

Resumindo-se o que foi exposto até aqui, neste item, as diferenças de classes – que englobam diferenças de renda, de escolarização, de trabalho etc. – sua estruturação e seus conflitos favorecem a ocorrência de um complexo jogo de forças que leva tanto à homogeneização quanto à diversidade das línguas, o que ocasiona, por sua vez, a variação e, por fim, a mudança lingüística.

Por último, é necessário fazer-se uma observação a respeito da relação entre classe social e mudança lingüística. Apesar do que foi mencionado anteriormente, algumas mudanças não são explicáveis em termos de a classe X ou Y favorecê-las (LABOV, 2002, p. 09). Nas palavras de Chambers (1995, p. 66-7):

Although social class has been the primary social variable in sociolinguistics, linguists are well aware that some social groups are not class differentiated and nevertheless show linguistic differentiation. Clearly, within tightly structured, relatively homogeneous social clusters – neighborhoods, parishes, institutions – individual further demarcate themselves by patterns of linguistic variation. These micro-level social clusters are called networks.

Os estudos de rede surgiram na Antropologia, com Bott (1958, apud MILROY e MILROY, 1992, p. 02) e depois com vários outros estudiosos que não estavam satisfeitos com as análises baseadas em estruturas sociais, políticas e econômicas altamente abstratas para explicar o comportamento dos indivíduos. Daí surgiu o interesse dos estudiosos das línguas em pesquisar as redes sociais.

Se o estabelecimento de critérios para a determinação das classes sociais é uma tarefa bastante complexa, fazer o mesmo para as redes sociais também não é fácil. Alguns estudiosos estabeleceram critérios para medir a estrutura das redes sociais (cf. MILROY, L., 1987, p. 141-2; e SUNDGREN, 2001, p. 118), o que permite identificar se os elos entre

os indivíduos e sua comunidade são fracos (*weak-ties*) ou fortes (*strong-ties*)⁴⁵. Estabelecer a força dos vínculos entre os membros de uma comunidade é importante para os estudos sociolinguísticos, na medida em que, segundo Nevalainen (2000, p. 264), há evidências empíricas de que as mudanças linguísticas ocorrem mais rapidamente quando os vínculos numa comunidade são fracos⁴⁶.

Os trabalhos feitos com redes mostram que elas, embora independentes, são, de certa forma, paralelas às classes sociais. Nos dois casos, há concordância de regras e normas de diversos tipos por parte de seus membros, diferindo apenas quanto ao seu alcance e extensão. Por isso Milroy e Milroy (1992, p. 02) afirmam:

Social class is fundamentally a concept designed to elucidate large-scale social, political, and economic structures and processes, whereas social network relates to the community and interpersonal level of organization.

Não é fácil determinar em quais circunstâncias o estudo de rede é mais apropriado que o estudo de classe social. Milroy e Milroy (1992, p. 05), após citarem alguns problemas de estudos sociolinguísticos realizados com base em classes sociais, apresentam uma tentativa de integrá-las com as redes – apontando suas inter-relações –, por meio do conceito de “modos de vida” (*life-mode*) de Højrup (1983, apud MILROY e MILROY, 1992, p. 18-23). Para os Milroy, ambas as idéias – classes sociais e redes – são complementares, em vez de conflituosas. Opinião semelhante tem Labov (2001, p. 59), para quem os dados dos estudos de rede são um adendo importante para a análise socioeconômica, mas não substitutos dela.

⁴⁵ Milroy e Milroy (1992) esclarecem que os antropólogos geralmente dividem as redes sociais em dois tipos: a) *estruturais*, que pertencem à forma e ao padrão das redes, e que se dividem em *close-knits* e *loose-knits*; e b) *interacionais*, que pertencem ao conteúdo dos vínculos, dividindo-se em *strong-ties* e *weak-ties*.

⁴⁶ A “*Weak-tie Hypothesis*” não é pacificamente aceita por todos os autores lidos, havendo mesmo uma polêmica entre eles quanto a este assunto. Para um aprofundamento desta questão, consulte-se, além dos já mencionados neste item, o texto de Bot e Stoessel (2002), que trata da relevância dos estudos de rede para os processos de variação e mudança linguística.

No caso desta pesquisa, não será feito o estudo de redes, mas de classes sociais. Duas explicações para isso são:

a) nenhum dos *corpora* deste estudo foi composto com base em redes, e sim em classes sociais;

b) para se realizarem estudos de rede, é necessário entrar em determinada(s) comunidade(s)⁴⁷, estudá-la(s) e entrevistar certas pessoas-chave – os chamados *core members*, segundo Chambers (1995, p. 88) –, além dos outros integrantes centrais do grupo, o que dificulta a coleta de dados, tornando-a mais demorada e, por conseguinte, mais difícil em caso de trabalhos com prazo determinado para a sua conclusão, como é o caso de dissertações e teses.

Como já foi dito, no Capítulo III serão expostos os critérios adotados nesta pesquisa para se fazer a divisão dos informantes em classes sociais.

O próximo item deste trabalho tratará de outro tema importante para o embasamento das análises dos dados.

1.3 O fator Frequência dos itens lexicais e a mudança lingüística

Neste item, serão apresentados estudos que, embora tratem a mudança lingüística sob outras perspectivas teóricas, diferentes da Teoria da Variação, ajudam a entender o processo de transformação pelo qual passaram as formas *Vossa Mercê* e *ocê*. São eles os de Fidelholtz (1975), Philips (1984, 1998, 1999, 2000 e 2001) e Bybee (2001, 2002a,

⁴⁷ Um famoso estudo de rede foi o de Labov (1972a), com os adolescentes do Harlem, Nova York. É interessante notar que, nesse estudo, dois dos entrevistadores dos jovens eram negros como os informantes, e conheciam a fundo a comunidade em questão e seus costumes. Esse fato foi fundamental para que as diferenças raciais não prejudicassem os resultados. Daí percebe-se a atenção que deve ter o pesquisador ao fazer um estudo de rede, para que seus dados sejam válidos. Em Chambers (1995, p. 69-70), encontra-se uma lista com alguns importantes estudos sociolingüísticos de redes realizados até então.

2002b e 2003). Esses autores abordam a importância da frequência de uso de itens lexicais ou de formas gramaticais nas mudanças linguísticas⁴⁸.

1.3.1 Fidelholtz

Em seu trabalho de 1975, Fidelholtz objetiva demonstrar que a familiaridade de determinadas palavras, em certos contextos fonéticos, pode exercer uma grande influência em sua pronúncia – mais especificamente, a redução de vogais em sílabas pretônicas.

A familiaridade tem estreita relação com a frequência de uso dessas palavras: se são familiares, a tendência é que elas sejam pronunciadas mais frequentemente.⁴⁹ A essa influência da frequência de uma palavra na sua produção oral Fidelholtz chamou “Regra de Frequência” (“*Frequency Rule*”, 1975, p. 200). Para comprovar sua hipótese, o autor fornece diversos exemplos de palavras com um mesmo ambiente fonético, mas com variação de redução. Alguns exemplos citados por ele são: San Francisco e °Francisco⁵⁰; *m*istake e °mistook; *a*stronomy e °gastronomy; *p*urdue e °purgation; etc. Nesses pares, sempre a palavra que sofre a redução vocálica é mais frequente que a outra.

Outros exemplos da validade da regra são as palavras *trombone* e *Austrália*, que somente têm a vogal pretônica reduzida quando são pronunciadas, respectivamente, por

⁴⁸ Bybee (2002b, p. 01) afirma que foi Schuchardt (1972 [1885]) quem primeiramente observou que as palavras mais frequentes eram as primeiras afetadas pela mudança sonora.

⁴⁹ Entretanto, à p. 207, Fidelholtz observa: “Familiarity is the intuitive concept associated with frequency of occurrence, but not all instances of using a word are equally effective in producing ‘familiarity’.” Mais à frente, ele complementa seu pensamento: “Some rare but ‘noticeable’ words can seem familiar – the word ‘bersek’ eg is moderately familiar to most people, yet does not occur in even the most extensive word counts.” Assim, a familiaridade e a frequência de uso estão diretamente relacionadas, mas não podem ser tomadas uma pela outra.

⁵⁰ Segundo sua notação, “◌” diante de uma palavra significa que sua vogal pretônica não se reduz; a ausência desse sinal significa que a vogal se reduz; e “(◌)” significa que algumas vezes as vogais sofrem redução e outras vezes não sofrem. Por outro lado, as vogais pretônicas aparecem no texto em itálico. Como neste trabalho as citações do autor já aparecem em itálico, as vogais pretônicas foram sublinhadas, para não haver dúvidas quanto a elas.

trombonistas e por australianos. Assim também sucede com vários nomes de localidades, como Manhattan e San Antonio, cujas vogais são reduzidas pelos habitantes desses lugares.

As causas da diferença de frequência estão, segundo o autor, nas diferenças entre a fala e a escrita:

The most probable reason for the rather sharp division in frequency between words which show reduction and ones which do not is the following. While words have roughly the same relative frequency in spoken and written language, the more frequent words are used even more frequently in speech than in writing (and perhaps more frequently in less formal language use in general); and conversely, less frequent words are used even less frequently orally. Differences between spoken and written style are one reason for this tendency. When speaking, one has less time to search for the “best” word (ie, one relatively rare or inaccessible), but typically one states the premise, restates it, and sums up the point. In (English) writing, however, a predominant stylistic device is to avoid redundancy – don’t repeat ideas, structures, or words. (p. 201)

Por outro lado, Fidelholtz afirma que a Regra de Frequência não atua sozinha, mas é condicionada pelo ambiente fonético em que a vogal se encontra. Quando este não é favorável à redução, existe uma tendência de aquela não atuar. Na realidade, os dois fatores – frequência e contexto fonético – estão bastante relacionados, um atuando diretamente sobre o outro. Nas palavras do autor,

Phonological environment sometimes cancels out high familiarity (eg. V before V⁵¹ never reduces: °duet) – and sometimes cancels out low familiarity (V before C; V always reduces: pharyngeal [unless this environment is in turn cancelled out by high ‘foreignness’: °Camay]); and high familiarity sometimes cancels out phonological context (astronomy, Manhattan, etc.)⁵² (p. 207-8)

⁵¹ “V antes de `V” caracteriza uma vogal tensa.

⁵² Por outro lado, o autor afirma que, conforme a frequência de uma palavra vai aumentando, o peso do ambiente fonético vai diminuindo, ocasionando, assim, a redução.

Abaixo seguem alguns contextos que desfavorecem – ou mesmo bloqueiam – a redução da vogal pretônica, de acordo com o autor:

- a) as vogais pretônicas iniciais imediatamente anteriores a outra vogal normalmente não se reduzem. Ex.: °coyote, °duality, °aorta;
- b) as vogais imediatamente seguidas de encontro de duas ou mais consoantes geralmente não se reduzem. Ex.: °October, °magnetic, °gastronomy, etc.⁵³;
- c) ainda com relação ao caso (b), acima, a redução é bloqueada quando a primeira consoante do encontro consonantal é /l/ ou nasal. Ex.: °salvation, °Milwaukee, °ambition, °campaing, °umbrella.

A explicação de Fidelholtz para isso é de ordem cognitiva, de processamento de informações e armazenamento de palavras no léxico pelos indivíduos. Entretanto, o autor atenta para o fato de que, qualquer que seja a explicação para esse fato, está claro que existe uma estreita inter-relação entre a frequência das palavras e a redução vocálica em certos ambientes fonéticos: vogais em palavras usadas com frequência se reduzem mais que vogais em palavras pouco usadas, embora essa relação admita exceções. (p. 208)

1.3.2 Phillips

Phillips é outra autora que relaciona as mudanças sonoras à frequência de palavras. Em seu trabalho de 1984, buscando caracterizar essa relação, ela estabeleceu o que chamou de Hipótese da Frequência-Atuação (Frequency-Actuation Hypothesis) (1984, p. 336): “Physiologically motived sound changes affect the most frequent words first; other sound changes affect the least frequent words first.”

⁵³ Há casos em que ocorre redução mesmo nesses contextos, como em *astronomy*. Segundo o autor, aí atua a Regra de Frequência.

Em trabalhos posteriores, a autora apresenta um refinamento da Hipótese da Frequência-Atuação, de 1984. Essa reformulação, por unir frequência de palavras e implementação da mudança, foi intitulada por ela de Hipótese da Frequência-Implementação (Frequency-Implementation Hypothesis):

Changes which require analysis – whether syntactic, morphological, or phonological – during their implementation affect the least frequent words first; others affect the most frequent words first. (2001, p. 134)

As mudanças que atingem as palavras mais frequentes primeiro estão baseadas na fisiologia da fala,⁵⁴ que atuam nas formas fonéticas superficiais (1984, p. 320). Envolvem a redução de vogais e o apagamento e a assimilação tanto de vogais quanto de consoantes. Segundo Phillips (2000, p. 04),

... changing affect the most frequent words first ignore information in the lexical entry, allowing features of segments to blur, producing assimilations and reductions, or allowing stress rules to reset to default position.

Como exemplos dessas mudanças, a autora cita:

- a) a mudança de *-an* [an] para *-on* [ón], no inglês antigo (OE) (1999, p. 107);
- b) a ditongação do inglês médio (ME) do tipo [ey] para [ei] (2000, p. 04); e
- c) a mudança de acento dos verbos dissílabos ingleses terminados por *-ate*, como “*frústrate*” para “*frustráte*” e “*díctate*” para “*dictáte*” (2001, p. 125-6).

⁵⁴ Phillips (1984, p. 336) observa que não se pode confundir “mudanças fisiologicamente motivadas”, que têm a ver com alterações sonoras ocorridas no trato vocal, com “mudanças foneticamente condicionadas”. Estas serão mencionadas adiante.

Estes, por serem usuais, mudaram muito antes de outros pouco comuns, como “*láctate*” ou “*lústrate*”. A explicação para isso, de acordo com a autora, é que, no caso dos verbos muito usuais, os falantes ignoraram a natureza não-sufixal de *-ate*, ou seja, não analisaram a palavra com um sufixo verbal, e sim como uma palavra de um único constituinte. De acordo com Phillips (2001, p. 125),

With frequent -ate verbs, (...) the suffixal nature of the -ate was being ignored, not analyzed as a verbal suffix. That is, the more frequent verbs in -ate was being ignored, not analyzed as a verbal suffix. That is, the more frequent verbs in -ate had lost the analysis of stem + suffix and were being treated like monomorphemic verbs, allowing the stress rules of English to apply automatically, without hindrance.

Por outro lado, as mudanças que atingem as palavras menos frequentes primeiro são motivadas por fatores diretamente relacionadas com o léxico, atuando nas suas formas subjacentes (1984, p. 337). São as mudanças que envolvem a análise dos constituintes das palavras pelos falantes, levando-os a fazer uma analogia entre elas e o sistema da língua. Por isso essas mudanças se dão no sentido da padronização, da regularização da língua.

A autora dá como exemplos dessas mudanças:

- a) a regularização do passado de certos verbos pouco usuais, como “*creep*”, “*leap*” e “*weep*”, que, seguindo a analogia da língua, formam o passado através de *-ed*: “*creeped*”, “*leaped*” e “*weaped*” (1984, p. 323);
- b) a fusão do /*ö(:)*/ com /*e(:)*/, no ME (1984, p. 326);
- c) a formação das palavras diatônicas no inglês moderno⁵⁵; e
- d) o apagamento dos *glides* no inglês americano do Sul, como em “*duke*”, “*nude*”, “*tune*” etc. (1984, p. 323; 1999, p. 107; 2000, p.04).

Nessas mudanças, os falantes fazem a análise dos constituintes das palavras, quer no nível fonológico, quer no morfológico, quer no sintático. Por exemplo, a fusão de /*ö(:)*/ com

⁵⁵ Nas palavras de Phillips (1984, p. 332), diatônicos são “homographic word pairs which receive final stress if they are verbs (*permit*) but initial stress if they are nouns (*pérmít*).”

/e(:)/ ocorreu por uma restrição fonológica: línguas sem vogais arredondadas frontais altas geralmente não têm vogais arredondadas frontais médias; assim, a fusão aconteceu para que o sistema fonológico da língua não fosse violado. No caso da formação diatônica das palavras, as pouco freqüentes levam o falante a fazer uma análise morfológica antes de sua produção (2001, p. 132)⁵⁶: se forem nomes, receberão o acento na primeira sílaba; se forem verbos, na segunda.

Um outro aspecto importante abordado por Phillips é a influência que a classe gramatical exerce sobre a mudança sonora. Ela afirma que, ao que parece, a importância da classe gramatical precede a importância da freqüência (2001, p. 134): “And it does seem that in determining which words are affected first in a sound change, word class takes precedence over word frequency.”

Levando-se em conta a relação entre mudança e classes de palavras, no caso das mudanças que envolvem análise segmental e que atingem os itens menos freqüentes primeiro, são os nomes os primeiramente atingidos, por serem estruturalmente mais complexos. Por outro lado, as palavras funcionais são as primeiramente atingidas no caso das mudanças fisiologicamente motivadas, tanto por serem estruturalmente mais simples quanto por serem muito freqüentes. Segundo Phillips (1983, apud 2001, p. 128-9),

Function words have been used here to identify that wide range of words which normally receive low sentence stress. Bollinger's (1975, p. 121-2) list provided the basis for deciding whether to include a word as a function word or not: the linking verb 'to be'; prepositions; determiners; quantifiers; coordinating conjunctions; relative pronouns; adverbial conjunctions; intensifiers; auxiliary verbs; pronouns [grifo meu]; proadverbs, and other proverbs.

Por último, é importante notar que a autora, embora dê explicações para a mudança centradas no indivíduo e sua relação com a estrutura da língua, reconhece que a mudança

⁵⁶ Phillips afirma: “Speakers access word class before they access phonological structure.” (2001, p. 132). Assim, essa análise é comumente feita para todos os casos, e não só para as formações diatônicas.

não tem a ver apenas com fatores estruturais, mas que a comunidade de fala exerce uma grande influência nesses processos. Se assim não fosse, afirma (1984, p. 322-3), não haveria variação nas comunidades. Segundo ela (1998, p. 370), depois que há a inovação – o que acontece numa subcomunidade –, toda mudança se implementa por meio do contato entre falantes, que se encarregam de difundi-las.

1.3.3 Bybee

Bybee (2001, 2002a, 2002b e 2003) também demonstra a importância da frequência de uso para os processos de mudança. Há dois tipos de frequência relevantes para o estudo da mudança: a de palavras e a de tipos. Segundo ela (2003, p. 03):

Token or text frequency is the frequency of occurrence of a unit usually a word or morpheme, in running text.

(...)

Type frequency refers to the dictionary frequency of a particular pattern, e.g. a stress pattern, an affix etc.

Os exemplos que a autora apresenta para a frequência de palavras isoladas são os verbos *broke*, passado de *break*, que ocorreu muito mais vezes que *damaged*, passado de *damage*, no *corpus* de Francis e Kucera (1982, apud BYBEE, 2003, p. 03). Já os exemplos de frequência de tipo é o passado dos verbos ingleses, que, em sua maioria, é formado por meio do sufixo –ed, mas também pode ocorrer através de outra palavra, no caso dos irregulares, como *have: had*.

Tanto a frequência de palavras como a frequência de tipo têm um papel preponderante para a mudança lingüística, mas há uma diferença entre elas. De acordo com Bybee, quando se trata de mudanças sonoras condicionadas foneticamente, os itens mais frequentes mudarão primeiro e com maior rapidez. (BYBEE, 2002b, p. 23; 2003, p. 15). Já nas mudanças analógicas, os itens menos usuais e que não estejam de acordo com os padrões gerais da

língua tenderão a se regularizar primeiro⁵⁷. É o caso da regularização do passado dos verbos ingleses. Nas palavras da autora (BYBEE, 2002a, p. 04),

In the case of irregular verbs regularizing, there are two effects of token frequency: (a) The form of the stem that is less frequently used will be lost, and the more frequent form will remain (Bybee, 1985), and (b) the less frequent verbs will regularize, whereas the more frequent ones (e.g., keep, kept) will retain their irregularities. (...) the stronger forms will remain, whereas the weaker ones have a tendency to be lost over time.

Entre os dois tipos de frequência acima citados, o mais importante para esta pesquisa – a mudança de *Vossa Mercê* para *você* e desta para *cê* – é o que diz respeito à frequência de palavras. Assim, ela será aqui mais detalhada.

Segundo Bybee (2003, p. 04; 19), a grande frequência de uso dos itens lexicais traz conseqüências importantes para a mudança lingüística. Uma delas é a manutenção de estruturas conservadoras, mesmo estando a língua sofrendo os processos de regularização, como é o caso, já mencionado aqui, de certos verbos que mantêm sua irregularidade na forma do pretérito, mesmo com a enorme produtividade do passado formado pelo sufixo – *ed*. Verbos muito comuns, como *eat/ ate, brake/ broke, write/ wrote*, são exemplos disso.

Outra conseqüência é a sintaxe particular dos verbos modais do inglês – *can/could, may/might, will/would, shall/should* e *must*. De acordo com Bybee (2003, p. 19), as propriedades sintáticas desses verbos eram comuns a todos os verbos do inglês. Entretanto, em razão da grande frequência dos verbos listados acima, essas propriedades ficaram

⁵⁷ Bybee (2003, p. 22) observa: “It might seem contradictory that repetition could both encourage innovation in one domain and enhance conservatism in another. This paradox is also found in the lexical diffusion of phonetics vs. morphophonemic change. In Hooper (1976), I pointed out that sound change affects high frequency items first, while analogical leveling affects low frequency items first. The substantive properties of words or phrases, their meaning and phonetic shape, are modified, usually reduced, with use. The ritualization or automatization process has an on-line effect of compressing and reducing; this is a processing effect. In contrast, the structural properties of words and phrases, i.e, the morphological structure of words and the syntactic properties of constructions are preserved by repetition; this is a storage effect.”

estritamente relacionados a eles, enquanto todos os outros verbos se adequaram à mudança sintática da língua inglesa.

Uma terceira consequência é a mudança sonora, e a autora comprova essa afirmação citando diversos estudos feitos sobre variação em diversas línguas. Ela aborda alguns processos fonológicos que atuam em palavras muito frequentes, tais como a fusão de segmentos (2002a; 2003) e a mudança vocálica (2002b)⁵⁸, mas aqui será tratado o que mais interessa a este estudo: a redução (incluindo-se aí o apagamento).

De acordo com Bybee (2002a, p. 01; 2003, p. 16), um grande número de mudanças fonológicas é de natureza redutiva, e elas acontecem mais e em primeiro lugar com as palavras mais frequentes. Como exemplo, a autora cita o apagamento das oclusivas /t/ e /d/ finais do inglês, do /d/ intervocálico e do /s/ final do espanhol, algumas mudanças vocálicas do inglês, a redução de vogais em *schwa* no holandês e no inglês, e o apagamento do *schwa* no inglês (2002b, p. 08). Bybee (2001, p. 07-10) cita o caso do subjuntivo dos verbos franceses, que está caindo em desuso, com exceção dos verbos mais frequentes e que ocorrem em contextos também mais frequentes.

Nesses casos, documentados por vários autores, observou-se que os itens mais frequentes eram atingidos primeiro pela mudança. Assim, Bybee conclui: “One might therefore predict that in general reductive changes tend to occur earlier and to a greater extent in words or phrases of high frequency.” (2002b, p. 08).

Neste ponto, é necessário fazer-se uma observação. Bybee (2002b, p. 13; 23), baseando-se em estudo próprio e de outros pesquisadores, confirma o que diz Phillips: a frequência não é o único fator que propicia as mudanças, embora estas sejam bastante favorecidas por

⁵⁸ Em Bybee (2002b, p. 6-7), a autora faz uma objeção à explicação neogramática de Labov (1994) para a mudança vocálica e argumenta a favor da tese de que a frequência das palavras é relevante nesses casos. Ela afirma: “It appears (...) that some evidence can be found that high frequency words undergo vowel shift before low frequency words.” (2002b, p. 7). Em seguida ela aponta: “... it may be more difficult to discern frequency effects in vowel shifts because of the effects of the preceding and following environments, which narrow each phonetic class to a small number of words.” (2002b, p. 7).

aquela. O contexto em que se encontra uma forma poderá acelerar, retardar ou mesmo bloquear uma mudança.

Com a mudança morfossintática ocorre o mesmo. Segundo Bybee (2002b, p. 27):

The morphological structure of words also plays a role from the initial stages of a change, but less because morphemes have some special status with respect to change and more because of the contexts in which they appear. Alternating contexts retard change while uniform ones allow change to hurry ahead.

Bybee dá explicações de ordem cognitiva para os fatos mencionados acima, mas elas não serão abordadas aqui⁵⁹. Entretanto, é importante expor algumas observações da autora sobre a relação entre a frequência dos itens e as mudanças lingüísticas, especialmente as reduções fonéticas.

A autora afirma que as palavras e expressões muito freqüentes têm maior propensão à redução devido à automatização e a sua predizibilidade (2002a, p. 03). Segundo ela, “Changes that affect high frequency words first are a result of the automation of production, the normal overlap and reduction of articulatory gestures that comes with fluency.” (2002b, p. 26). Assim, se as reduções são o resultado da automatização da fala, esse processo será mais desenvolvido quanto mais for praticado, ou seja, nas palavras e expressões mais freqüentemente usadas.

Além disso, as reduções também têm causas de ordem discursiva e pragmática:

If the speaker knows that the word will be easily accessed in the context, because it or related words have already been activated, the reductive automating processes will be allowed to advance. If the word is less

⁵⁹ Uma breve caracterização do que ela chama *Associative Network Framework* pode ser encontrada em Bybee (2003, p. 09-10).

predictable in discourse, the speaker is likely to suppress the reductive process and give the word a more explicit articulation. (2002b, p. 09).

Se as mudanças que atingem os itens mais freqüentes primeiro têm sua origem na automatização da produção, as que atingem os itens menos freqüentes se devem, de acordo com Bybee (2002b, p. 09), a uma menor experiência dos falantes com esses itens. Como foi dito anteriormente, estas últimas mudanças conformam os itens pouco freqüentes ao padrão normal da língua. A razão disso é que o único modo de se preservarem as propriedades idiossincráticas dos itens de uma língua é mantê-las na memória (2001, p. 04). Assim sendo, se a freqüência desses itens é baixa, eles podem não estar suficientemente disponíveis na experiência diária para serem aprendidos e fortalecidos, sofrendo, então, mudanças no sentido de se adaptarem às normas gerais da língua.

Por fim, cabe mencionar algumas das conseqüências da alta freqüência de palavras e expressões para os processos de mudança. De acordo com Bybee (2003, p. 01), a freqüência não é somente um resultado, mas também o fator fundamental que contribui para o processo de mudança lingüística ⁶⁰.

Para Bybee (2003, p. 02), o processo de gramaticalização implica a transformação de uma seqüência de palavras ou morfemas muito freqüentes na língua em uma unidade só, o que ocorre exatamente por meio da repetição. Segundo a autora (2003, p. 02),

...repetition (...) leads to the reduction of form through the weakening of the individual gestures comprising the act, and through the reorganization of a series or formerly separate gestures into one automated unit; and emancipation, which occurs as the original symbolic function inferred from the context in which it occurs.

⁶⁰ A partir desse ponto, os apontamentos de Bybee são sempre relativos à gramaticalização, mas, sendo esta um tipo específico de mudança e levando-se em conta o que foi visto até aqui sobre o papel da freqüência, pode-se aplicar essa idéia para quaisquer outros tipos de mudança.

A repetição gera uma outra característica importante para o processo de gramaticalização: a generalização, ou o enfraquecimento do significado original de uma palavra ou expressão. A idéia é que essa palavra ou expressão vai se tornando cada vez mais generalizada e se expandindo para novos contextos, com novas associações pragmáticas (2003, p. 03-04). Por outro lado, nem sempre há total perda semântica de um morfema, palavra ou expressão. Às vezes, seu sentido original pode coexistir com o sentido generalizado, persistindo em certos contextos.

Um exemplo dessa generalização e conseqüente mudança, analisado em Bybee (2003), é o processo de gramaticalização pelo qual passou o antigo verbo *cunnan*, atualmente o modal *can*: ele gradualmente expandiu seu uso para outros tipos de verbos e de sujeitos, levando a uma generalização de seu significado e de seu uso. Atualmente, não há nenhum sinal de seu significado original, que era “saber”.

Pelo que foi exposto dos trabalhos de Fidelholtz, Philips e Bybee, nota-se que há um consenso entre eles em relação à importância e à atuação da frequência nos processos de mudança lingüística. Os itens mais frequentes serão atingidos primeiramente se se tratar de mudanças sonoras foneticamente condicionadas; os itens menos frequentes mudarão primeiro se se tratar de mudanças que envolvam a análise dos seus constituintes. Outro ponto de concordância entre os autores é quanto ao papel do contexto onde se encontram os itens lexicais: um ambiente desfavorecedor poderá retardar ou mesmo impedir a mudança.

O que foi apresentado neste item é válido para se pensar no processo pelo qual passou a forma *Vossa Mercê*, dando origem ao pronome *você*, chegando a *cê*. Viu-se que as reduções (incluindo-se o apagamento) são favorecidas pela repetição de palavras ou expressões, ou seja, são favorecidas pela frequência de uso dos termos. E parece ser esse o caso dessas formas. Entretanto, essa discussão ficará adiada para a parte final deste estudo.

A seguir se passará ao último item deste capítulo, um apontamento sobre alguns aspectos da fonética e da fonologia do português. Como se observou neste item 1.3, o ambiente

fonético poderá levar adiante ou bloquear uma mudança. Assim, analisar os contextos onde se encontram as formas *você*, *ocê* e *cê* ajuda a explicar os resultados obtidos nesta pesquisa.

1.4 Os aspectos fonético-fonológicos do português e as formas

você, ocê e cê

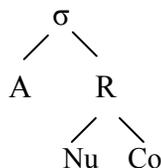
Um dos questionamentos feitos na Introdução desta tese se referia aos dados discrepantes em relação às formas *você*, *ocê* e *cê*: por que elas não têm a mesma distribuição sintática, ou seja, por que a forma *cê* apresenta tão pouca ocorrência como objeto, se é a favorita dos falantes na função de sujeito, com mais de 74% das ocorrências nos dois *corpora*? Uma hipótese levantada, a de número 02, foi que essa escolha está relacionada a aspectos fonéticos e fonológicos da língua portuguesa.⁶¹

Assim sendo, aqui serão apresentados, resumidamente, alguns aspectos das teorias fonética e fonológica do português – a estrutura da sílaba, a acentuação, o ritmo e a velocidade de fala – bem como processos fonológicos – a elisão e a redução de segmentos e sílabas – que serão utilizados para embasar a análise do ambiente em que se encontram as formas em estudo. Ao final deste item, a hipótese acima será especificada.

⁶¹ Para responder a essa pergunta, inicialmente, meu co-orientador, Prof. Dr. José Olímpio de Magalhães, e eu pensamos que o processamento da produção determinasse a ocorrência das três formas: se a velocidade de fala fosse bastante rápida, a redução da(s) palavra(s) anterior(es) – especialmente os conectivos – levaria automaticamente à redução das formas, no caso, o uso de *cê*; se a palavra anterior não fosse reduzida, a forma utilizada seria *você*. Esse pensamento, baseado na Teoria de *Garden Path* (cf. MAIA & FINGER, 2005; MAGALHÃES, 2005), se deveu aos dados obtidos dos *corpora* de 1982 e 2002, que demonstravam que as formas *ocê* e *cê* seguiam palavras – especialmente conectores – reduzidos. Entretanto, dois fatos posteriores nos fizeram mudar de idéia: i) a descoberta de uns pouquíssimos casos de *ocê* e *cê* sem a redução do conectivo; ii) o fato de que, em algumas comunidades, o uso de *cê* como objeto de verbo e de preposição não ocorre. Assim, independentemente da taxa de elocução, formas como “p’cê” – atestadas nos dados deste estudo e em São Francisco/MG (Coelho, 1999) – não aconteceriam, o que implica que o processamento da produção não é o fator mais importante nessa escolha. Adiante este assunto será retomado.

1.4.1 A estrutura da sílaba

A estrutura de uma sílaba consiste em um ataque (A) e uma rima (R), a qual se divide no núcleo (Nu) e na coda (Co), conforme o esquema abaixo, retirado de Collischonn (2001, p. 92)



Qualquer uma das categorias acima pode ser vazia, com exceção do núcleo, o qual, segundo Bisol (2001, p. 231), é sempre ocupado por uma vogal.

Collischonn (2001, p. 107) lista as possibilidades da sílaba em português:

V	<u>é</u>
VC	<u>ar</u>
VCC	<u>instante</u>
CV	<u>cá</u>
CVC	<u>lar</u>
CVCC	<u>monstro</u>
CCV	<u>tri</u>
CCVC	<u>três</u>
CCVCC	<u>transporte</u>
VV	<u>aula</u>
CVV	<u>lei</u>
CCVV	<u>grau</u>
CCVVC	<u>claustro</u>

O padrão CV é universal, aparece em todas as línguas do mundo, sendo, portanto, não-marcado ou canônico (SCHANE, 1973, p. 52-3; WEISS, GORDON e LILLYWHITE, 1987, p. 67; BARLOW e GUIERUT, 1999, p. 1455; e CRYSTAL, 2000, p. 238). Como será visto adiante, esse padrão silábico também exerce influência na escolha dos falantes por uma das formas *você*, *ocê* ou *cê*.

A partir do exposto, percebe-se que a estrutura silábica que será gerada com as formas átonas dependerá da escolha feita pelos falantes por uma das formas *você*, *ocê* ou *cê*, havendo preferência pela forma canônica, menos marcada. Entretanto, outros padrões silábicos também ocorreram, como pode ser visto pelo enunciado (186/02), abaixo. A seguir, aparecem algumas frases retiradas dos *corpora* de 1982 e 2002, com seu respectivo padrão silábico, em que a preposição (átona) forma um grupo clítico com o pronome (tônico):

(E 186/02) “Ela nunca vai pisar na bola c o m v o c ê.”⁶²

CVC CVCV

(E 691/02) “É, eu tamém concor<do>⁶³ c’ o c ê.”

CVCV

(E 460/82) “... acabou o filme p r a v o c ê.”

CCV CVCV

(E 529/82) “Só pra ficar olhan<do> p r ‘o c ê.”

CC VCV

⁶² Seguindo Câmara Jr. (1984, p. 52) e Collischon (2001, p. 107), considerar-se-á aqui a nasalização como uma consoante.

⁶³ Os segmentos entre colchetes angulares <> ou não foram pronunciados pelos informantes, ou foram muito reduzidos.

Pelos exemplos acima, vê-se que tanto a escolha de *com* e *pra* + a forma plena *você* como a de *c'* e *pr'* + a forma reduzida satisfazem ao padrão silábico do português, tendendo para a sílaba canônica universal CV. A partir desses exemplos – e de outros que serão analisados posteriormente – percebe-se também que o uso da forma reduzida *ocê* sempre vem acompanhado de uma preposição igualmente reduzida, o que indica que a escolha de uma das três formas levará o falante a adequar o conectivo, utilizando-se, para isso, de processos fonológicos como a elisão e a redução. Assim, quer se opte pela forma reduzida ou não, ambas as escolhas satisfazem a forma CV ou CCV.

Por outro lado, como será visto ao longo desta tese, houve pouquíssimas ocorrências de *cê* como objeto de preposição, mas esta discussão ficará suspensa por enquanto. Mais adiante ela será retomada.

A seguir, a abordagem de outro tema deste item.

1.4.2 O acento

“Acento” é o “termo usado em fonética para se referir ao grau de força ou intensidade ao se produzir uma sílaba” (CRYSTAL, 2000, p. 15). Como se sabe, no português, o acento recai sempre em uma das três últimas sílabas das palavras: se recai na antepenúltima sílaba, a palavra será proparoxítona; se recai na penúltima, será paroxítona; se recai na última, teremos uma oxítona. A grande maioria das palavras em português é paroxítona, vindo em segundo lugar as palavras oxítonas e, em terceiro, as proparoxítonas, com um número bastante reduzido de casos.

Em português, cada palavra fonológica tem somente um acento primário – a vogal tônica –, independentemente de seu tamanho; já as demais vogais recebem o acento secundário⁶⁴. Quando a palavra recebe um sufixo, na maioria das vezes o acento primário muda para o sufixo e ocorre um deslocamento dos acentos secundários, para se evitar um choque, como pode ser visto no exemplo abaixo, de Lee (2002, p. 150):

*cá*sa > *casé*iro, mas *café* > *câ*fez*in*ho

Deste modo, o acento secundário, no interior das palavras do português, ocorre em intervalos regulares, a cada duas sílabas a partir do acento primário, como o demonstram os exemplos abaixo, também de Lee (2002, p. 150):

- a. pindamonhangába
- b. fortalecimento
- c. irresponsabilidade
- d. oportunidade

Por outro lado, quando se trata de uma seqüência de palavras, Lee (2002, p. 151-2) afirma que no português ocorre o choque entre acentos primários, conforme demonstram os exemplos seguintes:

- e. jogador líndo
- f. computador lénto
- g. mulhér álta

⁶⁴ Enquanto Cristóvão-Silva (1999, p. 183) e Lee (2002, p. 156) afirmam que as vogais postônicas não recebem qualquer acento, Hernandorena (2001, p. 75) estipula o coeficiente 2 para elas, quando se trata de palavras isoladas, e o coeficiente 3, quando se trata de uma expressão. O coeficiente 1 é dado ao acento primário de uma palavra.

Entretanto, de acordo com Collischonn (2001, p. 153) e Hernandorena (2001, p. 78), um dos aspectos da estrutura métrica das línguas é a rejeição à seqüência de sílabas com acento primário⁶⁵. Como será visto mais adiante, esse aspecto é importante para a explicação de alguns resultados encontrados neste estudo. Por exemplo, quando a forma *cê* está depois de uma sílaba tônica, parece haver uma tendência a evitá-la: “ *Vi *cê*”. No entanto, se o falante retirar a tonicidade do *cê*, deslocando-a para outra palavra, evitará o choque de acentos: “Vi *cê* ontem”, o que pode ser confirmado pelos exemplos abaixo, retirados do *corpus* de 2002⁶⁶.

(E 614/02) “... ne<go> te arras<ta> *cê* lá dentro.”

(E 142/02) “... dinheiro suj<o> te lev<a> *cê* à morte...”

Nesses exemplos, vê-se que a forma *cê* como objeto de verbo não aparece sozinha seguindo a sílaba tônica; sempre há material fonético depois de *cê*, o que desloca seu acento.

1.4.3 O ritmo e a velocidade de fala

Ritmo e velocidade de fala são freqüentemente confundidos, mas, de acordo com Moraes e Leite (1993, p. 67), essas noções devem ser cuidadosamente distinguidas. Segundo os autores, “a rigor, a velocidade de fala independe em grande parte do padrão rítmico da língua: qualquer enunciado, de padrão rítmico acentual ou silábico⁶⁷, emitido numa velocidade “neutra”, pode ser produzido com maior ou menor velocidade” (1993, p. 68).

⁶⁵ Isso é conhecido em Fonologia como OCP (Obligatory Contour Principle), “Princípio do Contorno Obrigatório”, que diz que as línguas tendem a evitar elementos adjacentes idênticos, em diversos níveis de análise. Por exemplo: *Universidade de Minas Gerais* é falado *Universida’de Minas Gerais*.

⁶⁶ No *corpus* de 1982 não houve nenhum caso de *cê* como objeto acusativo. No de 2002 houve mais um caso, “*Eu vou ma:tá: cê, fdp!*”, que será discutido no Capítulo V.

⁶⁷ Moraes e Leite (1993, p. 68) afirmam que “línguas como português, o inglês ou o árabe têm como unidade rítmica o pé acentual, que se caracteriza por um acento forte na sua primeira sílaba, acento esse que retorna a intervalos de tempo constantes”. Já o ritmo silábico é caracterizado por sílabas com duração estável, independentemente de elas serem tônicas ou átonas. São exemplos de línguas com esse tipo de ritmo o francês, o espanhol, o japonês e o iorubá.

O ritmo é “uma regularidade percebida nas unidades proeminentes na fala.” (Crystal, 2000, p. 230). Essa regularidade pode dar-se em termos de sílabas acentuadas ou não-acentuadas; sílabas longas ou breves; e *pitch* alto ou baixo, sendo que pode ocorrer também uma combinação dessas variáveis.

Já a velocidade de fala, que é uma tradução da expressão inglesa *speech rate*, poderia ser traduzida também como “taxa de articulação”, expressando melhor seus efeitos. Quando se fala rapidamente, não se completa a articulação do segmento, isto é, as taxas caem, e isso leva o falante a omitir segmentos ou mesmo sílabas inteiras, quando estas são pré ou postônicas⁶⁸. Assim, a conseqüente redução de partes do enunciado devido à redução dos gestos articulatórios (pela velocidade da fala) influenciará diretamente as diversas produções de formas das sentenças, dentre elas, o pronome *você*, que se reduz para *ocê* ou *cê*. Os exemplos abaixo, dos *corpora* deste estudo, demonstram esse fato.

(E 703/02) “Ó, p’*cê* ver!”

(E 708/02) “Xô perguntá: p’*cê*.”

(E 496/82) “O que qu’*cês* tão levantan<d>o?”

(E 476/82) “S<e> *cê* ver a Bíblia, s<e> *cê* conhecer...”

1.4.4 Processos fonológicos do português

A seguir, serão apresentados alguns processos fonético-fonológicos do português que têm relação direta com a realização das três formas pronominais.

Com o objetivo de inter-relacionar o desenvolvimento e as desordens da fala à teoria fonológica, Weiss, Gordon e Lillywhite (1987) listam e descrevem os vários processos

⁶⁸ Em Moraes e Leite (1993, p. 74-76), discutem-se as possíveis motivações para a supressão de sílabas e segmentos nas línguas de ritmo acentual, como o português.

fonológicos das línguas naturais. Dentre eles, um interessa de perto a este estudo, o qual é, de acordo com Bisol (1993, p. 26), o desencadeador mais freqüente de ressilabação: a elisão⁶⁹. Além dele, veremos o processo de redução de encontros consonantais e, por último, a redução de segmentos e sílabas.

1.4.4.1 A elisão

A elisão é a omissão de fonemas - consoantes, vogais ou mesmo sílabas inteiras – no discurso corrido. Enquanto Collischonn (2001, p. 117) e Bisol (2001, p. 235) afirmam que a elisão somente ocorre em fronteiras de palavras, em Crystal (2000, p. 92) há a menção de que a elisão pode ocorrer no interior da palavra⁷⁰. Sem entrar no mérito da questão, para este estudo interessa a elisão de fonemas em fronteiras de palavras, e é sobre isso que esta parte da abordagem irá tratar.

Segundo Collischonn (2001, p. 117), a elisão afeta principalmente a vogal baixa /a/ e ocorre geralmente quando a vogal inicial da palavra seguinte é posterior - /o/ ou /u/, portanto -, e, opcionalmente, quando se trata de vogal frontal. Uma outra condição para que haja a elisão é a obrigatoriedade de que as vogais envolvidas sejam átonas. Os exemplos abaixo foram transcritos de Bisol (1993, p. 23-25; 2001, p. 235):

a – o > o:	“menina orgulhosa”	[meninorgulhosa]
a – u > u:	“menina humilde”	[meninumilde]
a – e > e:	“menina elegante”	[meninelegante]

⁶⁹ Bisol (1993) ainda cita a degeminação e a ditongação. A degeminação é o processo que transforma duas vogais idênticas numa só, desde que a segunda vogal seja átona. Esse processo pode ocorrer tanto no interior da palavra – “veemente” passa a “v[e]mente” e “coordenador”, a “c[o]rdenador”⁶⁹ – quanto em fronteiras de palavras – “vovô horroroso” passa a [vovorrórózu]. Já a ditongação “é o processo de formação de ditongos com a vogal final de um vocábulo e a inicial de outro, desde que uma das vogais seja alta (restrição segmental) e átona (restrição rítmica)” (COLLISCHONN, 2001, p. 107). Assim, pelo fato de o /o/ de *ocê* ser uma vogal média, a ditongação não se aplica a este estudo e, por isso, não será tratada aqui.

⁷⁰ Segundo Crystal (2000, p. 92), “Nas palavras polissílabas, as vogais e as consoantes de sílabas átonas geralmente desaparecem na fala em velocidade normal. No inglês, *camera* e *probably* são pronunciadas /kamrə/ e /prəbl/, respectivamente. Em português, as sílabas átonas costumam ser omitidas na fala, e uma expressão como “você precisa” pode acabar sendo pronunciada “cê picisa” (/sepsizΛ/).”

a – i > i: “pela idade eu era pequena” [pelidadi]

As frases abaixo, retiradas do *corpus* de 2002 deste trabalho, são exemplos de elisão:

(E 172/02) “(...) ela acaba jogand<u> ocê pro buraco.”

(E 188/02) “... te jog<a> ocê pra trás.”

(E 158/02) “Ah, deix<a> ocê.”

Neste ponto, é necessário fazer-se uma observação. Embora a elisão seja considerada processo fonético/fonológico, há que se notar que ela é o fruto da velocidade da fala, ou seja, da taxa de articulação do processo de produção. Veja-se que os exemplos acima não ocorreriam na fala em ritmo pausado, em que houvesse a necessidade ou a intenção de se enfatizar uma ou as duas palavras envolvidas. Assim, pode-se afirmar que a elisão ocorre como consequência do processamento da fala, como afirmam Callou & Leite e Schane, respectivamente:

Tratamos aqui apenas de alguns processos [fonológicos], mas queremos deixar assinalado que a maioria deles pode ser explicada por fenômenos articulatorios e perceptuais. (CALLOU e LEITE, 2001, p. 45)

Most phonological processes can be explained as articulatory or perceptual phenomena. Assimilation has a natural explanation in coarticulation. (...) Other phonological processes can be explained through perception. (SCHANE, 1973, p. 61)

No caso deste estudo, esta discussão não é muito importante. É suficiente saber que, nos dados, quando a palavra anterior às três formas terminava em vogal átona, a forma *ocê* foi a preferencialmente usada, tendo em vista que, desse modo, os falantes obteriam duas vantagens: a redução do tempo na produção do enunciado e a estruturação silábica canônica CV. No capítulo V este assunto voltará a ser discutido.

1.4.4.2 A redução de encontros consonantais

A redução dos encontros consonantais é o processo em que ocorre a omissão de uma das consoantes desse encontro, obtendo-se, normalmente, o padrão silábico CV. A consoante omitida pode ser uma soante ou uma obstruinte (cf. WEISS et al., 1987, p. 68). Os exemplos dados pelos autores são /mok/, para “smoke”, e /bek/, para “break”.

Cristófar-Silva (2002), fazendo o mesmo estudo com dados do português, observa que a omissão das líquidas é freqüente, como atestam os seguintes exemplos, que se encontram nas páginas 94 e 95:

ou[tr]o	~	ou[to]
sem[pr]e	~	sem[pe]
[fla]mengo	~	[fa]mengo
exem[pl]o	~	exem[po]

Dentro desse quadro, podem-se juntar dois exemplos mais, que são comumente ouvidos na fala espontânea:

(1) [ps]isa, de “precisa”

(2) [ps]ê, de “pra (vo)cê”

A forma [ps]isa pode ser ouvida até mesmo em programas de televisão; já a forma [ps]ê foi atestada nos informantes deste estudo – conforme os dados que serão apresentados e outros, detectados na fala espontânea de moradores de Belo Horizonte –, além de São Francisco/MG, segundo Coelho (1999).

Observe-se que, nestes dois últimos casos, a redução do encontro consonantal [pr], em nível de processamento, gerou [p], que tanto pode aceitar a forma *ocê*, originando “p’ocê”, quanto a forma *cê*, originando “p’cê”. Este último é um ataque silábico complexo possível, já que obedece a uma escala ascendente de sonoridade [p] + [s]. Em capítulo adiante este ponto será retomado.

1.4.4.3 O apagamento de sílabas

O apagamento de sílabas átonas ou fracas, citado por Weiss et al. (1987, p. 67-8), refere-se à omissão da(s) sílaba(s) átona(s) de uma palavra multissilábica. Esse apagamento se dá no nível de processamento da produção, e o resultado será a redução de palavras.⁷¹ Os exemplos de redução em inglês, dados pelos autores, são:

/teto/, para “potato”, e

/næna/, para “banana”

Os dados deste estudo mostraram uma imensa quantidade de sílabas átonas muito reduzidas ou mesmo apagadas. Abaixo seguem alguns exemplos:

(E 142/02) “(...) dinheiro su<jo> te lev<a> *cê* à morte.”

(E 167/02) “(...) mas às vê<ze>s te jog<a> *ocê* pro buraco.”

(E 454/02) “Tá mui<to> difi<cil> pr<a> *ocê*, né?”

(E 614/02) “*Cê* vai assistir o futebol *cê* ... ne<go> te arras<ta> *cê* lá den<tro>”

(E 717/02) “É, eu tamém concor<do> c<om> *ocê*.”

⁷¹ Relacionado a esse assunto, Magalhães (2005), em comunicação proferida no GT de Sociolinguística da ANPOLL, sobre POSSIBILIDADES FONÉTICAS/FONOLÓGICAS PARA AS VARIAÇÕES DE *ELE*, *ELA*, *ELES*, *ELAS*, diz o seguinte: “Na perda de massa fônica, a vogal tônica tende a permanecer, assim como as marcas morfológicas de gênero e número. Pela hipótese, é de se esperar que todas as formas que mantiverem a sílaba tônica possam ocorrer como variantes da forma plena”.

(E 382/82) “Se quiser ficar com ess<e> pr<a> ocê po<de> levar.”

(E 529/82) “Só pra ficar olhan<do> pr<a> ocê.”

Mas este processo também é importante para este estudo por uma outra razão. Ele explica a redução da forma “Vossa Mercê”. Essa expressão sofreu, ao longo do tempo, inúmeras reduções de segmentos e sílabas átonas até chegar a sua sílaba tônica: a forma *cê*. No próximo capítulo será feita essa discussão.

Em resumo, neste item foram expostos alguns tópicos acerca da fonética e da fonologia do português que, ao que tudo indica, influenciam a ocorrência das formas *você*, *ocê* e *cê*. Viu-se que a escolha dessas formas parece ser determinada pela busca dos falantes por padrões característicos, ou seja, não-marcados da língua em termos de estrutura silábica, acentuação e ritmo. Para tanto, eles se utilizam de alguns processos fonológicos, como a elisão e a redução de segmentos e sílabas, incluindo-se aí o apagamento.

Tendo sido expostos os apontamentos acima, pode-se, então, retomar a hipótese de número 02, a fim de que ela possa ser detalhada.

Hipótese 2: O uso das formas *você*, *ocê* ou *cê* é condicionado por fatores de natureza fonética e fonológica da língua portuguesa.

Assim, essa hipótese pode ser mais bem especificada, conforme consta abaixo:

Hipótese 2.1: O uso das formas *você*, *ocê* e *cê* busca adaptá-las aos padrões fonéticos e fonológicos da língua portuguesa, no tocante à estrutura silábica, acentuação e ritmo.

No Capítulo V, essa hipótese será testada.

1.5 Conclusão

Neste capítulo, foram vistos os pressupostos teóricos que embasaram as análises desta pesquisa. Primeiramente, foram resumidos alguns dos estudos já realizados sobre as formas *você*, *ocê* e *cê*. As diversas análises realizadas, com diferentes enfoques e seguindo variadas concepções teóricas, levantaram polêmicas e dúvidas que este estudo, por meio de seus resultados, irá discutir.

Em seguida, foram apresentados alguns dos principais aspectos da corrente teórica adotada neste trabalho, a Teoria da Variação, como a mudança lingüística e os construtos do tempo aparente e do tempo real, bem como os fatores externos relacionados à variação e à mudança lingüística.

Posteriormente, apresentaram-se alguns estudos que comprovam a relação entre a frequência de uso dos itens lexicais e a mudança, auxiliando na compreensão dos caminhos percorridos pelo pronome de tratamento *Vossa Mercê* até nossos dias. Ainda com a intenção de explicar os resultados obtidos nesta pesquisa, na última parte deste capítulo, foram vistos alguns tópicos referentes à fonética e à fonologia do português, que trarão maior clareza principalmente aos dados discrepantes encontrados na distribuição sintática da forma *ocê*.

No próximo Capítulo, serão abordados temas relacionados à história e aos usos desses itens lexicais: do pronome de tratamento *Vossa Mercê* ao *cê* de nossos dias.

CAPÍTULO 2 – AS FORMAS *VOCÊ*, *OCÊ* E *CÊ* – PERCURSO HISTÓRICO

Neste capítulo, será apresentada, resumidamente, uma descrição das formas em estudo: *você*, *ocê* e *cê*, desde sua origem até seu uso na língua portuguesa atualmente.

2.1 As formas de tratamento do latim ao português antigo

No latim, para as formas de tratamento, havia os pronomes *tu* – para um tratamento informal a um único interlocutor – e o *uos* (*vós*), usado em dois casos: (i) para a referência direta a mais de um interlocutor e (ii) para o tratamento respeitoso a um único interlocutor. (CÂMARA Jr., 1976; FARACO, 1996)

Ao lado dessas formas, no latim também se usava a forma indireta de referência, pela qual se expressavam as qualidades morais e o *status* social do ouvinte. Assim, os imperadores romanos não eram tratados simplesmente por *uos*, mas sim por *Uestra Maiestas*⁷², com o verbo na terceira pessoa do singular, destacando-se, dessa forma, sua importância e sua superioridade naquela sociedade.

Por outro lado, em Portugal, no início, a situação era diferente. Nos primeiros tempos da monarquia, o rei mal se distinguia dos outros nobres, já que seu poder não era

⁷² Faraco (1996, p. 58) cita ainda *Uestra Serenitas*, *Claritudo*, *Excellentia* e *Alternitas*, como formas de tratamento ao imperador.

suficientemente forte para destacá-lo, e os riscos das guerras contra um inimigo comum e a familiaridade imposta pela vida militar aproximavam-no de seus vassallos. É a partir do século XIII que o rei começa a distinguir-se das outras classes e somente no século XV ele consegue eliminar qualquer autoridade contrária à sua (LUZ, 1956, p. 274-5).

Nesse início – século XIII –, o soberano era tratado por *vós*, havendo esse tratamento perdurado até o século XVI (LUZ, 1956, p.282). Entretanto, aos poucos, outras formas de tratamento foram surgindo, e o pronome *vós* foi lentamente sendo substituído pelas formas “Vossa + Nome”. Primeiramente, as formas indiretas de tratamento começaram a substituir os pronomes oblíquos *vos*, *a vós*; em seguida, o pronome sujeito *vós*; e, por último, o possessivo *seu* (e suas flexões) substituiu o possessivo *vosso*. Segundo Luz (1956, p. 303-307), data de 1455 os primeiros exemplos de uso de *lhe*, em vez de *vos*, dirigidos ao rei. Já na função de sujeito, nas cortes de 1490, a terceira pessoa é mais freqüente que a segunda pessoa do plural. Quanto ao emprego de *seu*, nas cortes de 1481-82, ainda não aparece, mas, na primeira metade do século XVI, substitui em grande parte o uso do pronome *vosso*.

As formas de tratamento com a estrutura Vossa + Nome, como *Vossa Mercê*, por exemplo, foram introduzidas na língua portuguesa no século XIV e, especialmente, no XV. Cintra (1972, p. 18) afirma que a expressão *Vossa Mercê* aparece pela primeira vez nas *Actas das Cortes*, em 1331, pronunciada principalmente por castelhanos para dirigir-se ao seu rei ou ao rei de Portugal.

Não há um consenso, porém, entre alguns dos autores lidos, sobre esse dado. De acordo com Luz (1956, p. 300; 359), a expressão aparece duas vezes nas cortes de 1331, mas é provável que já existisse antes dessa época. Segundo Cintra (1972, p. 17), até o século XIV, “notamos antes de mais nada a total (grifo meu) ausência de tratamento de tipo nominal”. No entanto, Faraco (1996, p. 60) afirma que, em textos de Fernão Lopes, sobre o período de 1357 a 1433, os aristocratas já se tratavam por *Vossa Mercê*. Por outro lado, Ali (1976, p. 93) e Nascentes (1956, p. 115) afirmam que, no século XIV, *Vossa Mercê* ainda não se havia cristalizado como expressão pronominal.

Se não há um total consenso com relação à data do surgimento de *Vossa Mercê*, alguns autores, como Faraco (1996, p. 58) e Cintra (1972, p. 19) concordam com o fato de que a forma foi importada do castelhano *Vuestra Merced* e era usada para o tratamento ao rei. De acordo com Lopes (2002a, p. 1-2), o tratamento com *merced* já podia ser encontrado em textos espanhóis do século XIII, tendo surgido, assim como em Portugal, em virtude da decadência do uso de *vós* como tratamento de cortesia.

Desse modo, como no Império Romano o imperador era nomeado indiretamente, tomando-se por base sua importância dentro da sociedade, em Portugal, igualmente, o tratamento ao monarca era indireto, fazendo-se referência à sua generosidade e a seu poder de conceder favores, graças, proteção, justiça etc.

Esse tratamento diferenciado para o rei se justificava socialmente. A partir do século XII, Portugal e mesmo a Europa começaram a sofrer profundas transformações econômicas, políticas e sociais. O crescimento das atividades artesanais e comerciais nas cidades originou o fortalecimento da burguesia, que começou a ter, a partir de 1254, representantes nas Cortes, assim como a nobreza e o clero (FARACO, 1996, p. 55-6).

A ascensão da burguesia e o conseqüente enfraquecimento da nobreza feudal – juntamente com o fato de Portugal transformar-se numa grande potência mundial, graças às grandes descobertas e ao comércio marítimo –, originaram fortes mudanças na vida social e cultural da Corte e, principalmente, transformaram o rei numa figura única. Houve, assim, a necessidade de se destacar essa importância, o que fez com que o tratamento dispensado ao rei fosse também único. Portanto, *vós* já não era suficiente para nomeá-lo, surgindo então as formas nominais de tratamento.

No início, era *Vossa Mercê* a forma mais usual para alguém se dirigir a ele, mas logo outras concorrentes a suplantaram: primeiramente *Vossa Senhora*, vindo depois *Vossa Alteza*⁷³ e,

⁷³ De acordo com Luz (1956), o primeiro exemplo conseguido da forma *Vossa Senhora* é de 1434 (p. 324); *Vossa Alteza* aparece nas obras de Fernão Lopes (escrivão de D. João I e do infante D. Duarte, teria nascido

finalmente, *Vossa Majestade* (ALI, 1965, p. 93; COOK, 1997, p. 452). Conforme Luz (1956, p. 320), estas três últimas formas evidenciam outras qualidades do rei, “mais conformes com a nova concepção da dignidade real”; por isso substituíram *Vossa Mercê*, por exprimirem melhor sua magnificência.

Luz (1956, p. 362) nos fornece um panorama da frequência de uso das três primeiras formas acima para o tratamento ao rei.

TABELA 2

Uso das formas de tratamento ao rei (em %)

	1455	1459	1468	1472-3	1475	1477	1481-2	1490
V.Mercê	19	49	32	37	37	18	7	0
V.Alteza	44	31	63	50	48	54	69	99
V.Senhoria	37	20	5	13	15	28	24	1

O declínio de *Vossa Mercê* para o tratamento real deveu-se à expansão de seu uso para outras figuras da aristocracia portuguesa: os filhos do rei, o condestável, duques de alto estado e condes. “Fidalgos e fidalgotes começaram a aceitar e exigir igual tratamento dos seus criados e subalternos” (ALI, 1976, p. 93). Posteriormente, passou a ser usado como tratamento para a burguesia, chegando, por fim, a significar um tratamento respeitoso para qualquer português a quem não se poderia tratar por *Vossa Senhoria*, por *tu* (considerado bastante íntimo) ou por *vós*.⁷⁴

Cintra (1972, p. 25-36) assim descreve a decadência de *Vossa Mercê*:

em 1380 e morrido em 1460), (p. 326); e o exemplo mais antigo conseguido de *Vossa Majestade* é de 1442 (p. 334)

⁷⁴ Nascentes (1956, p. 117) afirma que, no estilo oficial, o tratamento de *Vossa Mercê* perdurou até 1889.

A fórmula Vossa Mercê, inicialmente aplicada ao rei e à rainha, foi posteriormente destronada desse cargo – no qual foi substituída por, sucessivamente, Vossa Alteza e Vossa Majestade. O honorífico decadente passou a ser aplicado à nobreza, depois à burguesia, e continuou a descer na escala social, circulando em variantes morfofonológicas tais como vossancê e você, a última emergindo no século XVII, e chegou, principalmente em Portugal, ao extremo de ser percebido como ofensivo em certas camadas da população.

Ao mesmo tempo em que *Vossa Mercê* baixava na escala social, essa forma, por ser repetida a toda hora e por todo mundo, sofreu transformações fonéticas, como será visto a seguir⁷⁵.

2.2 A evolução de *Vossa Mercê*

A forma de tratamento *Vossa Mercê* substituiu o pronome de tratamento *vós*, empregado em sinal de distinção à figura real ou a algum membro da nobreza (OLIVEIRA e RAMOS, 2002)⁷⁶. Já no século XVIII, o uso de *vós* para a referência a um único interlocutor cai em desuso, atingindo 25% das ocorrências de tratamento (LOPES e DUARTE, 2003, p. 67) e chega perto de zero no fim do século XIX (LOPES, 2003a, p. 12). Com relação à forma *Vossa Mercê*, esta aparece na primeira fase do século XVIII com 33% de ocorrências, sofrendo um declínio acentuado a partir do fim desse século.

Mas *Vossa Mercê*, devido à frequência de uso e por ser um pouco longa, sofreu uma simplificação fonética resultante da redução de segmentos e sílabas átonas, originando diversas variantes⁷⁷.

⁷⁵ Luz (1956, p. 271) afirma: “... as fórmulas de tratamento cortesias são expressivas, por vezes exageradas, e o valor expressivo das palavras atenua-se rapidamente, pelo uso frequente que dela se faz.”

⁷⁶ Nesta parte do trabalho, faz-se menção a um texto de Oliveira e Ramos (2002) e a outro de Ramos e Oliveira (2002). Não obstante as semelhanças, são dois textos diferentes.

⁷⁷ No último capítulo desta tese, este assunto retornará, relacionando-se o que está sendo dito aqui ao que se mencionou no item 1.3, do capítulo anterior.

Nascentes (1956, p. 119-121) aponta dezoito registros de formas simplificadas de *Vossa Mercê*, além de *você*, podendo haver mais:

Cê, mecê, mincê, ocê, oncê, sucê, suncê, vacê, vainicê, vancê, vansmincê, vassuncê, voncê, vosmecê, vossemecê, vosmincê, vossuncê, ucê.

Apesar de algumas dessas formas ainda existirem, especialmente nos dialetos rurais do Brasil (AMARAL, 1976; ALI, 1996; COELHO, 1999) e de Portugal (OLIVEIRA S., 1996; BASTO, 1931; COOK, 1997), foi a forma *você* que se fixou na língua portuguesa, transformando-se em pronome de tratamento. O primeiro registro da forma *você* aparece em texto do Padre Francisco Manuel de Melo, publicado em 1644 (RAMOS e OLIVEIRA, 2002)⁷⁸, e vai aos poucos ganhando espaço.

A mudança de *Vossa Mercê* para *você*, como acontece normalmente com os processos de mudança lingüística, não foi um fato isolado. Como aponta Lopes (2003a, p. 20):

... a gramaticalização de Vossa Mercê>você não foi um processo isolado, mas uma conseqüência de uma mudança encaixada lingüística e socialmente. Há uma emergência gradativa de formas nominais de tratamento que passam a substituir o tratamento cortês vós, a partir do século XV em português, num primeiro momento pela ascensão da nobreza e mais tarde da burguesia, que exigia tratamento diferenciado. Essa propagação, que começa de cima para baixo, se dissemina pela comunidade como um todo e as formas perdem sua concepção semântica inicial, gramaticalizando-se – algumas de forma mais acelerada que outras, como é o caso de Vossa Mercê>vosmecê>você.

Uma conseqüência extremamente importante do processo de gramaticalização de *Vossa Mercê>você* é o rearranjo da estrutura da língua, especialmente de seu sistema pronominal, como aponta Lopes (2003). Em primeiro lugar, *você* entrou na língua portuguesa na sua forma plural, no lugar de *vós*, que caiu em desuso. Porém, por se haver originado de uma

⁷⁸ Faraco (1996, p. 63) cita uma outra data para o surgimento de *você*: 1666.

forma nominal – *Vossa Mercê* –, que fazia a concordância com a 3ª pessoa, *você* também passa a ter esse comportamento, mesclando a 2ª pessoa com a 3ª.

Em segundo lugar, essa combinação da 2ª pessoa com a 3ª alterou o uso dos possessivos e dos pronomes-complemento. Assim, ao lado de *teu(s)/tua(s)*, encontramos *seu(s)/sua(s)* e *de você(s)*, e ao lado de *te* encontramos *lhe*, muitas vezes na mesma frase, como atestam cartas escritas no Brasil no século XIX (cf. LOPES e DUARTE, 2002, p. 10-11). Uma consequência direta deste último fato é que a forma *dele* passou a ser cada vez mais usada, para se evitar a ambigüidade gerada pelo possessivo *seu*, agora correspondente tanto à 2ª. pessoa quanto à 3ª.

Em terceiro lugar, essa “mistura” entre a 2ª. e 3ª. pessoas atingiu o imperativo, tornando possíveis frases como esta: “**Vem** pra Caixa **você** também”, veiculada na imprensa (cf. FARACO, 1996, p. 78)⁷⁹.

Em quarto lugar, a entrada de *você(s)* no quadro dos pronomes da língua, especialmente no PB, trouxe consequências importantes também com relação ao seu paradigma verbal, em que temos, atualmente, três formas básicas, conforme aponta Lopes (2003a, p. 1): eu **falo**, tu/você/ele/ela/a gente **fala**, vocês/eles **falam**. Com a perda da distinção de pessoa através das desinências verbais, o uso do sujeito passa a ser cada vez mais necessário. Assim, o PB estaria deixando de ser uma língua de sujeito nulo para ser uma língua de sujeito pleno.

A esse respeito, exatamente, Duarte (1996 [1993]), ao analisar peças teatrais populares de sete períodos (1845, 1882, 1918, 1937/38, 1955, 1975, 1922), observa que, nos três primeiros, há uma nítida preferência pelo sujeito nulo. Entretanto, a partir da peça de 1918, principalmente na de 1938, de Humberto Couto, há uma reversão desse quadro, ou seja, tem-se o início de uma tendência crescente ao preenchimento do sujeito, que coincide com a diminuição do uso da 2ª. pessoa direta – *tu* (op. cit., p. 111-113).⁸⁰

⁷⁹ Em Faraco (1996), há uma discussão sobre as consequências da entrada da forma *você* na língua portuguesa, especialmente no tocante ao imperativo. Consulte-se principalmente a 4ª seção, p. 65.

⁸⁰ Duarte, entretanto, cita os estudos de Oliveira (1990) e Paredes da Silva (1988), cujos dados mostram, respectivamente, que havia uma grande quantidade de sujeitos nulos de 1ª e 3ª pessoas mesmo na segunda

Lopes (2003, p. 02) resume o que foi dito aqui a respeito da inserção de *você(s)* no PB:

Constatam-se, dessa forma, várias alterações morfossintáticas: introdução de novas formas pronominais, simplificação do paradigma verbal, preenchimento obrigatório do sujeito e ordem mais rígida na sentença.

Esta última característica citada pela autora representa um aspecto interessante do processo de gramaticalização de *Vossa Mercê* para *vocês*: esta última forma apresenta um comportamento sintático diferente de sua forma desenvolvida: enquanto *Vossa Mercê* apresenta maior mobilidade na frase, a forma *você(s)* começa a ocupar posições mais fixas na sentença, especificamente exercendo a função de sujeito pré-verbal (LOPES, 2003a, p.19; 2002c, p. 10).⁸¹ Uma outra característica desse processo de mudança da forma *você* é que, de expressar unicamente referência definida, ela passa a expressar também a referência indefinida.

Além disso, a mudança teve relação com os aspectos sociais da comunidade. A forma *você* tornou-se mais produtiva nas camadas populares e nas relações assimétricas de superior para inferior, ao passo que o contrário se deu com *Vossa Mercê*, como atestam peças teatrais brasileiras e portuguesas (LOPES e DUARTE, 2003) escritas nos séc. XVII e XVIII. A tabela abaixo, extraída de Lopes e Duarte (2003, p. 68), explicita o que foi dito até aqui com relação a essa diferença de uso das formas nominais e pronominais.

metade do século XX; e que o sujeito nulo ocorre preferencialmente na 1ª pessoa, e não na 2ª. Como ambas as amostras constituem-se de cartas pessoais, Duarte observa que é importante manter-se a uniformidade dos *corpora* sob análise, já que diferentes gêneros textuais podem levar a resultados igualmente diferentes.

⁸¹ Conforme aponta Lopes (2003a, p. 12), baseando-se em Hopper (1991), nos processos de gramaticalização, uma forma não substitui a outra imediatamente; durante um período de tempo, as duas convivem juntas na língua.

TABELA 3

Uso das formas pronominais e de tratamento nas relações hierárquicas entre os personagens das peças teatrais no PB e no PE

Tipo de relação entre informantes	Você	V.M.	Vós	Tu	Outros
De superior para inferior	24/376 6%	16/376 4%	42/376 11%	276/376 73%	18/376 5%
De inferior para superior	09/235 4%	74/235 31%	20/235 9%	58/235 25%	74/235 31%
Membro de um mesmo grupo social (popular)	57/318 18%	17/318 5%	15/318 5%	218/318 69%	11/318 3%
Membros de um mesmo grupo social (não popular)	21/423 5%	31/423 7%	76/423 73%	276/423 73%	64/423 15%

Em resumo, a forma *você* inicia seu processo de gramaticalização no século XVII, sendo usada pelas classes populares. No século XIX, entretanto, já era de uso corrente, especialmente no Brasil. Este é o tema do próximo subitem.

2.3 A forma *você* na língua portuguesa

Em Portugal, à época da colonização do Brasil, a partir do século XVI, o uso tanto de *Vossa Mercê* quanto de *vós* pela população das classes mais baixas – de onde saiu a maior parte de nossos colonizadores – estava já em declínio, ao passo que a simplificação fonética de *Vossa Mercê* estava adiantada, ocorrendo as variantes dessa forma ao lado do pronome *tu*.⁸²

De acordo com Menón (2000), embora não se saiba ao certo que tipo de português era falado aqui no início do período da descoberta, pode-se pensar que essa situação lingüística foi trazida para o Brasil pelos colonizadores. Estes vieram de todas as partes de Portugal, o que deu origem a uma grande diversidade de dialetos convivendo juntos. Entretanto,

⁸² Domingos (2001), entretanto, estudando as formas de tratamento no português do século XVI, através de peças de Gil Vicente, observa o aparecimento apenas de TU (para o tratamento íntimo) e VÓS (para tratamento cerimonioso e de polidez).

rapidamente essa diversidade foi neutralizada, tendo o nosso falar perdido alguns dos traços mais marcantes do português lusitano, diferenciando-se dele.

Além disso, essas diferenças lingüísticas entre a metrópole e a colônia tenderiam a aumentar rapidamente, visto que aqui não havia escolas nem imprensa, ficando a educação a cargo apenas dos jesuítas. Assim, como afirma Menón (2000, p. 131),

Sem escolas para impingir normas e corrigir erros, sem imprensa para fixar visualmente padrões empregados na escrita, a língua poderia perfeitamente ter se modificado mais rapidamente que em Portugal no tocante ao uso de vosmecê, sobretudo na grande massa da população.

Com o objetivo de investigar a implementação do pronome *você(s)*⁸³ no PB, Menón (2000) analisa cartas pessoais de autores de diversas localidades do Brasil. O uso mais longínquo de *você* encontrado pela autora é da década de 1880, em cartas oriundas de Minas Gerais, donde se conclui que, na língua oral, essa forma já existia há muito mais tempo.

Os pronomes de tratamento no PB, em textos escritos do século XIX, também foram alvo de outras pesquisas. Salles (2001), por exemplo, estudou as formas de tratamento no PB do século XIX, sendo que seu *corpus* era constituído de cartas pessoais e de uma peça de teatro – *Caetaninho* ou *O tempo colonial*, de Paulo Antônio do Valle. De um total de 595 ocorrências de tratamentos de terceira pessoa, apenas 37 foram de *você* e três de *vosmicê* – a variante culta de *você* -, tendo essas duas formas sido usadas nas relações entre iguais, ou seja, hierarquicamente simétricas.

Cyrino e Brito [199_] analisaram, através de peças de teatro escritas no Brasil, as formas de tratamento de quatro períodos distintos: 1º) 1833 - 1842 (comédias de Martins Pena); 2º) 1911 (peça de Graça Aranha); 3º) 1933-1937 (peças de Oswald de Andrade); e 4º) segunda

⁸³ Menón afirma que a implementação de *vocês* é anterior à de *você*: “...podemos crer que a hipótese da penetração da forma plural do pronome ter se efetivado antes da forma singular pode ser demonstrada pelas ocorrências de *vocês* mesmo quando no singular aparece o *tu*. Pelo menos no material examinado até agora, *vós* não apareceu nesse contexto.” (2000, p. 145)

metade do século XX (teatro de Gianfrancesco Guarnieri). Nesse estudo, as autoras apresentam o seguinte resultado⁸⁴:

TABELA 4
Distribuição das formas de tratamento no PB - 1833 - 2ª metade séc. XX
(em %)

	VOCÊ	TU	SR./SR^a	VOSSA + N
1º Período (1833-1842)	11	43	40	6
2º Período (1911)	-	97	3	-
3º Período (1933-1937)	21	52	27	-
4º Período (2ª metade séc. XX)	83	10	6	-

Outros trabalhos sobre o uso de formas de tratamento em peças teatrais portuguesas (LOPES, 2003a), cartas escritas no Brasil⁸⁵ (LOPES e DUARTE, 2004) e peças brasileiras e portuguesas (LOPES e DUARTE, 2003) atestam que o uso das formas de tratamento e dos pronomes obedecia à estrutura social vigente aqui e lá, e ao grau de intimidade existente entre os interlocutores. A tabela 5, extraída de Lopes (2003a), e 6, extraída de Lopes e Duarte (2004), abaixo, além da tabela 4, acima, apresentam essa distribuição.

⁸⁴ O quadro acima foi reproduzido a partir do Gráfico 1 – Ocorrências das formas de tratamento empregadas nos quatro períodos, de Cyrino e Brito (199_, p. 3).

⁸⁵ A amostra é constituída de cartas de Minas Gerais, Paraná, Rio de Janeiro e Bahia, escritas nos séculos XVIII e XIX.

TABELA 5
Distribuição das formas de tratamento mais frequentes em português
(Séc. XVIII-XIX)

Tipo de relação entre informantes	Século XVIII	Século XIX
De superior para inferior	Você – 41%	Tu – 89%
De inferior para superior	Vossa Mercê – 59%	Sua Mercê – 60%
Membros de um mesmo grupo social (popular)	Tu – 42%	Você – 59%
Membros de um mesmo grupo social (não popular)	Vossa Mercê – 54%	Vossa Mercê – 50%

TABELA 6
Distribuição das formas pronominais e de tratamento em função
do destinatário da carta

Destinatário da carta	Você	V.M.	Tu	Outras formas	Total
Para o Rei	-	-	-	7 – 100%	7
Para amigos/primos	09 – 14%	01 – 2%	46 – 72%	08 – 13%	64
Para tios	-	19 – 100%	-	-	19
Para filhos	17 – 94%	-	-	01 – 6%	18
Para netos	12 – 30%	-	28 – 70%	-	40
Para pais	-	10 – 28%	-	26 – 72%	36
Sem intimidade	-	80 – 61%	-	52 – 39%	132

Entretanto, nas décadas de 20 e 30 do século XX, a coexistência das duas formas, como mostram as tabelas acima, desaparece, predominando o uso quase exclusivo de *você*. Vê-se, assim, o surpreendente crescimento dessa forma em apenas poucas décadas. Atualmente, a predominância de *você* com relação a *tu*, no Brasil, é incontestável⁸⁶. Este

⁸⁶ Entretanto, Oliveira G. (2000, p. 371) afirma que o *tu* tem-se expandido para algumas áreas do Centro Oeste, onde se falava o *você*. Também Paredes Silva (2003) atesta o aumento do uso do pronome *tu* (com o verbo na terceira pessoa do singular) no Rio de Janeiro.

último pronome é usado (às vezes com o verbo concordando com a terceira pessoa do singular, como em “Tu fez”, “Tu quis” etc.) principalmente na Região Sul (em Santa Catarina, Rio Grande do Sul⁸⁷ e numa minúscula área do Paraná) e em algumas áreas do Nordeste e do Norte⁸⁸ do país. Em todo o restante do país, acontece o uso de *você*.

Situação distinta acontece com o *você* no português europeu, atualmente. Lá, ainda hoje, *tu* é o pronome normalmente usado para o tratamento íntimo (CINTRA, 1972, p. 14-5; FARACO, 1996, p. 63-4; MENÓN 1995, p. 95), destinando-se o pronome *você* ao tratamento de superior para inferior, inclusive de idade, e ao tratamento entre iguais, mas de quem se quer manter uma certa distância (SANTOS, 1985, p. 2; CINTRA, 1972, p. 14-15).

Uma das diferenças mais marcantes quanto ao uso de *você* no Brasil e em Portugal é sua conotação negativa lá, ao contrário daqui. Luft (1957, p. 202-3) chega mesmo a afirmar que ser tratado por *você* é sentido por alguns como um insulto. Basto (1931, p. 187) afirma que *vossemecê* (e, em certas localidades, o *vossecê*), para os aldeãos, é mais respeitoso que *você*. Oliveira, S. (1996, p. 132), em seu estudo sobre a avaliação das formas de tratamento por estudantes da Universidade de Évora, verificou que os informantes valorizam mais o tratamento “*o vizinho/a vizinha*” que *você*. Porém, ao que parece, a conotação de *você* está mudando, em Portugal: o seu sentido depreciativo está desaparecendo, e “você aparece adotado como tratamento afectuoso, mais íntimo do que tu” (CINTRA, 1972, p. 40-1).

⁸⁷ Em Santa Catarina e no Rio Grande do Sul também há regiões onde se fala *você* (OLIVEIRA G., 2000, p. 370).

⁸⁸ O Nordeste foi citado por Oliveira G. (2000); já o Norte, por Cunha e Cintra (2001).

2.4 As formas *ocê* e *cê* na língua portuguesa

As formas *ocê* e *cê* chamam a atenção pela pouca presença nos textos lidos – relacionados neste capítulo - e obras consultadas.⁸⁹ Dos poucos autores que mencionam pelo menos uma das duas formas, Basto (1931, p. 192) afirma que, no indo-português do Norte, usam-se as formas *oscê*, *ocê* ou *ucê* e *cê*, em vez de *você*, e Soares (1955, p. 59) diz que *ocê* é usado na ilha de Santo Antão, pertencente ao arquipélago de Cabo Verde.

No português do Brasil, Nascentes (1956) atesta a existência de *ocê* em Minas Gerais e em Goiás, fazendo uma rápida menção à existência da forma *cê*: “Aparece numa frase típica do linguajar da malandragem. Quando um malandro se lembra de ameaçar outro, este responde *Cê é besta*. Existe em Goiás” (1956, p. 119).

Se a forma *cê* é registrada como sendo de uso da “malandragem”, pode-se deduzir que, em 1956, o uso de *cê* – na zona urbana, pelo menos, – não era usual como é atualmente. Por outro lado, Amaral (1976) cita *ocê* e *cê* como variantes de *Vossa Mercê* já em 1920. E, como afirma Salles (2001), se em 1920 a forma *cê* foi atestada, pressupõe-se que ela já existisse no século XIX.

Paredes Silva (1998) atesta que, em 1990, na cidade do Rio de Janeiro, a forma *cê* já suplantava *você* e a forma zero, ou seja, o sujeito nulo: enquanto *cê* representava 38% do total de ocorrências de seu *corpus*, na função sujeito, *você* representava 32% e zero, 30% (p.128).

Outros estudos também atuais atestam um uso acentuado de *ocê* na região de São Francisco, Minas Gerais, (COELHO, 1999) e que a forma *cê* suplanta *você* em Belo Horizonte e Ouro Preto (RAMOS, 1997 e 2000, respectivamente). Os dados de 1982 e de

⁸⁹ As obras consultadas foram as seguintes gramáticas: Almeida (1999), André (1999), Bechara (2003), Cegalla (2000), Cunha (1985), Cunha E Cintra (2001), Faraco E Moura (1997), Mesquita (1999), Pasquale & Ulisses (2003), Paschoalin & Spadoto (1989), Rocha Lima (2000), Sacconi (1999) e Terra (2000). Além dessas gramáticas, Elia (1992) e Teyssier (2001). Dentre esses autores, apenas Sacconi (op.cit., p. 196) menciona algumas variantes de *Vossa Mercê*, mas não cita *ocê* nem *cê*.

2002 deste estudo corroboram os encontrados por Ramos: a ocorrência da forma *cê*, em Belo Horizonte, na língua falada, é bastante superior a *você*, em todas as faixas etárias analisadas, sendo que o número de ocorrências de *ocê* é muito pequeno, nas duas amostras.

Há, por outro lado, fatos curiosos a respeito da forma *cê*: um deles é a sua ausência – ou quase – em contextos em que *você* e *ocê* ocorrem normalmente. Um outro fato interessante é a presença quase nula da forma *cê* na língua portuguesa falada em outros países que não seja o Brasil, como se pode perceber pelo “descaso” com que ela é tratada por grande parte dos autores citados neste capítulo que estudaram as formas de tratamento na língua portuguesa. Uma razão para a pouca ocorrência de *cê* fora do Brasil pode ser o fato de o pronome *você* ser menos usado que *tu*. De acordo com os autores listados no item 1.3, a frequência de uso favorece a redução dos itens lexicais.

Pode-se perceber um outro dado curioso, em relação à forma *cê*, quando se compara a evolução de *Vossa Mercê* e *Vuestra Merced*. *Vuestra Merced* também sofreu um processo de simplificação fonética, dando origem a diversas formas variantes, citadas por Nascentes (1956, p. 122) e Lopes (2003a, p. 10). Dessas variantes⁹⁰, duas chamam a atenção pela semelhança com formas portuguesas: *Vuced* e *Ucé*, mas não há um *ce* ou um *ted* no espanhol.⁹¹

Resumindo o que foi exposto neste Capítulo, viu-se que a forma *Vossa Mercê*, antigo tratamento dado ao rei, foi-se popularizando e, portanto, começou a deixar de ser dirigida aos monarcas a partir de 1468. O aumento da frequência de uso de *Vossa Mercê* deu origem às suas diversas variantes, resultantes da redução fonética que essa forma de tratamento sofreu.

Essa situação lingüística foi trazida para o Brasil, no início da colonização – século XVI. Em nosso país, a forma *você* foi atestada em textos escritos – cartas pessoais – pela

⁹⁰ De acordo com Nascentes (1956), o último estágio desse processo de simplificação – *Usted* - aconteceu depois do século XVII e é usado regularmente na língua espanhola como forma de tratamento respeitoso.

⁹¹ Consulte-se Lopes (2003), para uma comparação entre a evolução de *Vossa Mercê*, no português, e *Vuestra Merced*, no espanhol.

primeira vez na década de 1880 e, em peças teatrais, no período de 1833-1842. A partir daí, *você* expandiu-se, tendo suplantado a concorrente *tu* em boa parte do país, o que, segundo estudo de Cyrino e Brito [199_], teria ocorrido em meados do século XX. Quanto à forma *cê*, não apareceu nenhuma vez nos dados das pesquisas citadas neste capítulo, referentes aos séculos XVII a XIX, sendo atestada pela primeira vez em 1920.

Diante disso, pode-se perguntar: a partir de quando, realmente, a forma *cê* surgiu? Se ela se originou de *você*, por que é desfavorecida em certos ambientes em que *você* ocorre (cf. VITRAL, 1996; RAMOS, 1997)?

Apesar de haver trabalhos mais recentes a respeito de *cê*, como os já mencionados aqui, nota-se que há ainda pontos obscuros e perguntas a serem respondidas sobre essa forma, principalmente no tocante a suas origens. Alguns desses aspectos serão discutidos ao longo desta tese, mas não todos, por questões de tempo e espaço, mas também por não se relacionarem propriamente aos objetivos desta pesquisa.

2.5 Conclusão

Neste capítulo, foi apresentado um resumo de trabalhos que tiveram por objetivo traçar o percurso histórico das formas sob estudo, desde sua origem, como tratamento ao rei, até seu uso nos nossos dias, tanto no Brasil quanto em Portugal.

Outro importante aspecto abordado diz respeito às implicações lingüísticas e sociais da inserção do pronome de tratamento *Vossa Mercê* e, principalmente, do pronome *você* para a língua portuguesa.

Tendo sido feito um resumo da história e das características de *Vossa Mercê* e suas formas resultantes, serão apresentados, no próximo capítulo, os procedimentos metodológicos seguidos neste estudo.

CAPÍTULO 3 - PROCEDIMENTOS METODOLÓGICOS

A mudança lingüística é um tema que há muitíssimo tempo vem despertando o interesse de pesquisadores de diversas correntes teóricas, desde a Grécia antiga, passando por Bacon, na Idade Média, pelos neogramáticos, estruturalistas, gerativistas, variacionistas e difusionistas (DA HORA, 1997). Em cada uma dessas correntes, a língua e a mudança lingüística eram concebidas de uma maneira, algumas mais semelhantes entre si, mas todas se diferenciavam.

Assim, ao se adotar um marco teórico para este trabalho, foi levado em conta que:

- (i) a concepção de língua aqui adotada é a de que esta não pode entendida fora de seu contexto social, ou seja, sem estar vinculada às interações dos falantes de uma comunidade;
- (ii) o objeto do presente estudo é o uso das três formas *você*, *ocê* e *cê*, observando-se os fatores lingüísticos e extralingüísticos que estão condicionando esse uso.

Portanto, a presente investigação foi realizada tomando-se como base a Teoria da Variação, pelo fato de ser ela a que fornece os subsídios teóricos necessários para se responder às questões pertinentes a esta tese, conforme exposto anteriormente.

A seguir, encontram-se listados os fatores analisados neste trabalho.

3.1 Os fatores internos e externos analisados

Conforme mencionado anteriormente, um dos objetivos deste trabalho é investigar o comportamento das formas *você*, *ocê* e *cê* em tempo aparente e em tempo real. Além desse objetivo, um outro visa a aprofundar os estudos sincrônicos a respeito das três formas, haja vista os resultados divergentes encontrados por outros estudiosos deste tema, os quais foram apresentados no primeiro capítulo. Portanto, nesta pesquisa, a seleção dos fatores internos teve por meta atingir os seguintes objetivos:⁹²

- a) analisar os fatores que favorecem o uso de *você* e *ocê* e desfavorecem o uso de *cê*, como a função sintática e a focalização;
- b) analisar o fator *referência*, cujos resultados se mostraram conflitantes em estudos já realizados;
- c) comparar os dados deste estudo a outros feitos anteriormente, com relação à posição de tópico e tipo de oração em que as formas aparecem, fatores estes não selecionados pelo Programa Goldvarb em qualquer dos *corpora*.

A análise desses fatores permitirá uma melhor descrição do comportamento dessas três formas na cidade de Belo Horizonte.

Com relação aos fatores externos, os mais citados e utilizados em estudos sociolinguísticos são: idade, gênero, classe e/ou redes sociais, escolaridade, ocupação, etnia e procedência geográfica. Nesta tese, como já mencionado anteriormente, somente a idade, o gênero e a classe social serão considerados.

⁹² Um outro objetivo desta tese – determinar a influência dos aspectos fonéticos e fonológicos do português na ocorrência das formas *você*, *ocê* e *cê* – somente foi estipulado no último ano de trabalho, a partir da entrada do Prof. Dr. José Olímpio de Magalhães na co-orientação do mesmo, à época do Exame de Qualificação. Assim, devido à falta de tempo, os aspectos fonéticos e fonológicos não foram tomados como fatores internos pesquisados.

Em resumo, nesta tese, foram analisados os seguintes fatores internos e externos:

1. Fatores internos

a) Função sintática da forma:

- Sujeito
- Objeto de verbo
- Objeto de preposição

b) Contigüidade da forma em relação ao verbo:

- Contígua
- Não-contígua

c) Referência da forma:

- Definida
- Indefinida

d) Focalização da forma⁹³:

- Foco
- Não-foco

e) Posição da forma em construções de tópico⁹⁴:

- Tópico
- Não-tópico

f) Tipo de oração em que a forma aparece:

- Declarativa

⁹³ Com respeito à focalização, foram considerados apenas os casos de acento contrastivo. Um exemplo de frase com foco é: (E 589/02) “*Você é diferente da sua colega*”, em que a forma *você* é pronunciada com mais intensidade que o restante da frase.

⁹⁴ Foram considerados como construções de tópico os deslocamentos à esquerda (cf. Pontes, 1976, Liberato, 1980; e Callou et al., 1996).

- Interrogativa simples
- Interrogativa ‘que que’

2 Fatores externos

a) Idade:

- Faixa etária I (de 08 a 11 anos)
- Faixa etária II (de 12 a 15 anos)⁹⁵
- Faixa etária III (de 16 a 30 anos)
- Faixa etária IV (de 31 a 47 anos)
- Faixa etária V (acima de 47 anos)

b) Gênero:

- Feminino
- Masculino

c) Classe Social:

- Baixa
- Média

3.2 Os dados

No item 1.2 deste trabalho, observou-se que, apesar de muitos autores concordarem com a validade dos estudos de tempo aparente, eles sozinhos não podem dirimir as dúvidas quanto à existência da gradação por idade, sendo necessário, portanto, que se proceda ao estudo de tempo real.

⁹⁵ Apenas no *corpus* de 2002 há dados de fala de crianças e adolescentes.

No caso desta pesquisa, foi possível realizar a análise em tempo real, tendo em vista que os dois *corpora* utilizados – de 1982 e de 2002 – satisfazem às exigências mencionadas por Bailey et al. (1991) e Labov (1994), listadas no subitem 1.2.4. Para maior clareza, elas aparecem repetidas abaixo:

- 1) Os estudos precisam ser realizados com o espaço de, no mínimo, uma década ou mais: os *corpora* deste estudo estão distantes vinte anos no tempo.
- 2) Eles precisam analisar algumas das mesmas variáveis: os dois estudos enfocam fenômenos diferentes – o alçamento das vogais pretônicas, o de 1982; o cancelamento de /r/ final de nomes, o de 2002 –, mas nos dois *corpora* é possível encontrar, com abundância, as formas em estudo nesta tese.
- 3) Eles precisam estudar a mesma comunidade de fala: os dois *corpora* referem-se à cidade de Belo Horizonte. Embora haja diferenças entre as localidades de residência dos informantes, está-se tomando aqui, por hipótese, que, em uma metrópole, com uma interação constante entre seus habitantes, não há, necessariamente, diferenças lingüísticas tão importantes entre informantes oriundos de bairros socioeconomicamente coincidentes a ponto de distorcerem ou mesmo invalidarem os resultados encontrados. Em outras palavras, está-se assumindo que, em Belo Horizonte, as diferenças dialetais significativas encontradas se dão principalmente no nível social, e não no geográfico⁹⁶.
- 4) Eles precisam incluir dados que tenham sido coletados da mesma maneira: os dois trabalhos utilizaram-se da pesquisa sociolingüística para a coleta de dados.
- 5) A comunidade não pode ter sofrido alterações bruscas: apesar de, obviamente, Belo Horizonte ser menor em 1982 que em 2002, na década de 80 já era uma metrópole e um pólo industrial.

Portanto, os *corpora* de 1982 e de 2002 podem ser comparados, e o estudo de tempo real pode ser feito. Abaixo, são apresentadas as duas amostras.

⁹⁶ Apesar do reconhecimento da existência e da importância das redes sociais, não se está trabalhando com elas, nesta tese, e sim com classes sociais. Por isso tomar-se a hipótese acima, de uma certa semelhança entre bairros distintos de uma mesma cidade.

3.2.1 Os dados do *corpus* de 1982

3.2.1.1 A coleta de dados

O *corpus* de 1982 constitui-se de dezesseis entrevistas sociolingüísticas, realizadas pela Prof^a Dr^a Maria do Carmo Viegas (VIEGAS, 1987), que gentilmente cedeu os dados para esta pesquisa.

Depois de selecionados, os informantes foram contatados e convidados a participar da pesquisa, sabendo que teriam sua fala gravada. As entrevistas foram realizadas na residência dos informantes e versaram sobre assuntos do cotidiano. Com isso, a pesquisadora procurou obter conversas descontraídas, para que o Paradoxo do Observador (cf. Labov, 1972) fosse minimizado e que o vernáculo dos entrevistados fluísse. Em seguida, as fitas foram transcritas pela pesquisadora, sendo que, para este trabalho, foi necessário ouvi-las repetidas vezes, para que não houvesse dúvidas quanto à ocorrência das formas *você*, *ocê* e *cê*.

3.2.1.2 O perfil dos informantes

Os informantes eram moradores de dois bairros socioeconomicamente distintos de Belo Horizonte: Colégio Batista – de classe média média – e Barreiro – de classe média baixa –, selecionados pela pesquisadora.⁹⁷

Segundo Viegas (1987, p. 20),

ambas [as áreas] são de formação recente (anos 40/50). Uma delas, o Colégio Batista, é uma extensão de um bairro antigo, a Floresta, habitado inicialmente por chefes de seção e funcionários graduados (cargos relacionados à função de Capital do Estado). A outra, o Barreiro, é um bairro formado quando do incentivo à criação dos pólos industriais na região, habitado principalmente por operários.

⁹⁷ Ver, no Anexo 3 desta tese, mapas de Belo Horizonte, nos quais aparecem os bairros citados.

Abaixo segue uma descrição mais detalhada dos dois bairros (VIEGAS, 1987, p. 21-22):

Hoje, o Barreiro é um bairro populoso e sua população é composta, na maioria, por operários que trabalham na MANNESMANN ou na Cidade Industrial. (...) Os moradores contam com serviços legalizados de água e luz e alguns desfrutam dos serviços de esgoto e telefone. As residências são, na maioria, pequenas casas que possuem, por vezes, um pequeno quintal. Muitas delas têm barracões alugados, abrigando, assim, várias famílias, outras são moradias “parede-meia”, divididas por duas ou mais famílias.

(...)

Os moradores do bairro geralmente possuem aparelhos de rádio e televisão (um ou outro, a cores); alguns possuem automóveis. Uns são donos de suas próprias casas, outros “moram de aluguel”. A educação fica, normalmente, ao encargo de estabelecimentos gratuitos e os jovens, com frequência, estudam no curso noturno, pois trabalham durante o dia, já as mulheres adultas são donas-de-casa ou fazem serviços relacionados às tarefas domésticas (“salgadeiras”, “passadeiras”, “faxineiras” etc.)

Com respeito ao bairro Colégio Batista, segundo Viegas (1987, p. 23-4):

... teve grande parte do seu território pertencente à Igreja Batista, que, nos últimos tempos, loteou, aumentando assim a sua população.

(...)

As residências são, na maioria, prédios, construídos recentemente com muitos apartamentos. Por vezes, encontramos conjuntos de prédios residenciais. Existem algumas casas de construção mais antigas (reformadas ou não) que aos poucos estão cedendo lugar aos prédios de apartamentos. No Colégio Batista todos os moradores desfrutam dos serviços de água, luz, esgoto e coleta de lixo; a maioria possui telefone, assim como carro próprio. As residências são, de modo geral, alugadas ou financiadas pelo BNH. Quanto à educação, é bastante significativa a participação de estabelecimentos particulares, como o próprio nome do bairro denuncia. Alguns jovens trabalham e estudam simultaneamente, cursando, muitas vezes, a universidade. Os moradores desta região são, geralmente, funcionários públicos, vendedores, bancários, pequenos

comerciantes, professores, engenheiros assalariados etc. As mulheres adultas exercem, também, tais profissões.

Nesta área existem mercearias, padarias, bazares, lanchonetes, farmácias, consultórios médicos e odontológicos, escritórios, salões de beleza, igrejas e suas associações. Mas, não existe, por ser um bairro muito próximo do centro da cidade e do centro comercial da Floresta, um centro comercial com agências bancárias ou um comércio mais intenso.

Assim, o nível social dos informantes selecionados é:

- a) Barreiro: os informantes são “operários, filhos ou esposas de operários, com uma renda familiar baixa” (VIEGAS, 1987, p. 24).
- b) Colégio Batista: os informantes têm uma renda familiar mensal maior do que a do outro grupo, mas não exercem uma atividade de prestígio na sociedade.

Aqui cabe uma observação. O *corpus* de 1982 foi originalmente coletado tomando-se duas faixas etárias: jovens e adultos, com uma divisão uniforme dos informantes por classe social e gênero, além da idade. Entretanto, neste estudo, para que fosse possível analisar o processo de mudança em progresso das três formas, foi necessário proceder à divisão dos informantes em três faixas etárias – Faixa I (de 16 a 30 anos), Faixa II (de 31 a 47 anos) e Faixa III (acima de 47 anos). Devido a esse fato, não há equivalência entre o número de informantes mais jovens, mas isso teria como consequência a diminuição de um *corpus* que já é pequeno, se comparado ao de 2002. Portanto, optou-se por manter todos os informantes, sabendo-se que as análises foram feitas com base em porcentagens e probabilidades de ocorrência de cada forma.

Tendo sido feita essa observação, pode-se voltar ao perfil desses informantes. No quadro a seguir, encontram-se seus dados.

QUADRO 3
Perfil dos informantes - 1982

Número	Faixa Etária	Gên.	Origem	Profissão
1	I (16-30 anos)	Fem.	Barreiro	Estudante
2	I	Fem.	Barreiro	Estudante
3	I	Masc.	Barreiro	Estudante
4	I	Masc.	Barreiro	Operário
5	I	Fem.	C.Batista	Estudante
6	I	Fem.	C.Batista	Estudante
7	I	Masc.	C.Batista	Estudante
8	I	Masc.	C.Batista	Comerciário
9	II (31-47 anos)	Fem.	Barreiro	Dona de casa
10	II	Fem.	Barreiro	Dona de casa
11	II	Masc.	Barreiro	Policial
12	II	Fem.	C.Batista	Professora
13	III (+ 47 anos)	Masc.	Barreiro	Operário
14	III	Masc.	C.Batista	Bancário
15	III	Masc.	C.Batista	Comerciário
16	III	Fem.	C.Batista	Professora

3.2.2 Os dados do *corpus* de 2002

O *corpus* de 2002 compõe-se de duas amostras: uma coletada pela Prof^a Ana Paula Huback (HUBACK, 2003), que também gentilmente a cedeu para este trabalho, com dados de informantes três faixas etárias: de 16 a 30 anos; de 31 a 47 anos e acima de 47 anos. A outra amostra, que complementa a primeira, é constituída dos dados de crianças (de 08 a 11 anos) e adolescentes (12 a 15 anos), foi coletada por esta pesquisadora.

A seguir, a coleta de dados e o perfil dos informantes serão descritos, primeiramente das crianças e adolescentes; em seguida, dos demais informantes.

3.2.2.1 A coleta de dados de crianças e adolescentes

Com o intuito de facilitar a coleta de dados de crianças e adolescentes, foi escolhida, como ponto de partida, a Escola Municipal Aurélio Pires⁹⁸, localizada no bairro Liberdade, em Belo Horizonte, que atende a alunos de diversas classes sociais, especialmente as mais baixas – moradores da favela situada ao redor dos bairros Santa Rosa, Jaraguá, Liberdade e São Francisco, chamada de Vila Santa Rosa.

Em março de 2002, a Diretora e as Coordenadoras da Escola foram contactadas, indicando os alunos-informantes dos quais se necessitaria. Entretanto, não foi possível conseguir, nessa Escola, todos os informantes, principalmente alguns de classe média. Assim, duas crianças e alguns alunos e ex-alunos da Escola Fundamental do Centro Pedagógico/UFMG, participaram da pesquisa. Além deles, uma outra informante adolescente de classe média, que foi entrevistada pela Prof^a Ana Paula Huback, compôs o *corpus*.

Após ter em mãos uma lista fornecida pelas coordenadoras dos turnos matutino e vespertino com os possíveis informantes, os alunos foram convidados a fazer as entrevistas, sabendo que elas seriam gravadas, e aceitaram o convite. As entrevistas foram realizadas na sala da Coordenação, um lugar mais reservado, mais silencioso, sendo, portanto, mais propício para esse fim. Os alunos eram chamados a essa sala, quando então eram informados sobre a pesquisa⁹⁹ e diziam se queriam ou não participar dela.

⁹⁸ Foi escolhida uma escola por ser um local onde se encontram vários informantes em potencial, pelo tipo de entrevista que se iria realizar e também por ser um lugar conhecido deles, onde se sentiriam um pouco mais à vontade para participar das entrevistas.

⁹⁹ Como sugere Labov (1972), o motivo real da entrevista foi omitido. Assim, os alunos pensaram que o objetivo da entrevista era uma análise de sua capacidade de fazer e responder a perguntas de forma natural e sincera.

Tendo em vista a faixa etária dos informantes e também para que se obtivessem as formas em estudo, foi pensada uma forma de entrevista¹⁰⁰ (Anexo 1) que parecesse mais uma brincadeira, para que os efeitos dessa situação formal fossem minimizados. Assim, os encontros eram feitos a três: um menino e uma menina¹⁰¹ da mesma idade, de mesmo nível social e da mesma turma na escola faziam perguntas um ao outro, orientados pela entrevistadora.

Ao final das perguntas diretas, quando os informantes já estavam mais acostumados ao gravador, à entrevistadora e a essa situação, estes respondiam a perguntas ou contavam casos vivenciados por eles ou por parentes e amigos (cf. Labov, 1972), ou simplesmente conversam sobre assuntos que iam surgindo ao longo da entrevista. Assim, cada entrevista teve duas partes: a primeira constou de perguntas diretas, conforme explicitado acima; a segunda, de conversas informais entre os informantes e a entrevistadora, e de relatos narrados por eles.

Devido ao tipo de pergunta dirigida aos alunos e à situação mais formal de entrevista, foram analisados por separado os dados obtidos nas duas partes: considerou-se a primeira parte como “teste”, e a segunda, como entrevista, propriamente dita.

As entrevistas com os alunos da Escola Fundamental do Centro Pedagógico/UFMG seguiram a linha das realizadas na E.M. Aurélio Pires, com apenas uma diferença: como os informantes eram ou já haviam sido alunos da entrevistadora, as conversas foram todas mais espontâneas. Assim, a segunda parte da entrevista girava em torno de assuntos da escola e geralmente era muito mais longa que a primeira, ao contrário do que ocorreu com os informantes da Escola Aurélio Pires. Já a coleta de dados da informante pertencente ao *corpus* da Prof^a Ana Paula Huback foi feita somente através de entrevista sociolinguística. Assim, a informante não participou do “teste” e, na análise deste, seus dados não foram computados.

¹⁰⁰ O roteiro da entrevista é de autoria dos professores do II Ciclo da E. F. do Centro Pedagógico/UFMG. Nesta pesquisa, nem todos os itens desse roteiro foram utilizados, mas a maioria sim.

¹⁰¹ A escolha de informantes de sexos distintos deveu-se à facilidade de identificação dos sujeitos, quando fosse feita a transcrição das gravações.

No início, foram contactados dezesseis informantes. Porém, o número de dados extraídos dessas entrevistas foi muito reduzido, o que forçou a inclusão de mais oito, totalizando, então, 24 informantes.

3.2.2.2 O perfil dos informantes crianças e adolescentes

Os informantes foram selecionados de acordo com os seguintes critérios:

- a) faixa etária: de 08 a 11 e de 12 a 15 anos;
- b) gênero: masculino e feminino;
- c) classe social¹⁰²: *média* - moram em casa própria, em bairros de classe média; a família possui pelo menos um automóvel de passeio; em casa, têm acesso a computador, Internet, revistas e jornais; o lazer também compreende cinema e/ou teatro; os pais têm nível universitário, mesmo incompleto; e classe social *baixa* - moram em casa própria, alugada ou cedida por familiares, em bairros pobres; a família não possui automóvel de passeio; não têm acesso, em casa, a computador, Internet, revistas, jornais, teatro ou cinema; os pais, especialmente a mãe, já que muitos informantes não convivem com o pai, estudaram até a 4^a. série do Ensino Fundamental e, ou não trabalham fora, ou trabalham em serviços mal remunerados, como servente de pedreiro, faxineiro(a), empregada doméstica etc.

Todas as crianças e adolescentes da classe baixa viviam na Vila Santa Rosa. Conforme dados de 2000, fornecidos pela Prefeitura de Belo Horizonte – Regional Pampulha, a Vila Santa Rosa teve suas primeiras ocupações no início da década de 60, sendo que, até o início dos anos 80, não existia qualquer processo de urbanização na Vila.

¹⁰² Os critérios referentes à classe social dos informantes seguiram parcialmente Paiva e Scherre (1999), *apud* Mollica (2003, p. 30).

A partir de 1980, quando foi criada a primeira associação de moradores, vários benefícios, além de água e luz, foram obtidos, como esgoto, posto policial, telefones públicos, escolas e creches. Atualmente¹⁰³, 521 famílias – 2059 habitantes - vivem na Vila Santa Rosa. Dados fornecidos pelo Programa BH Vida, da Prefeitura Municipal de Belo Horizonte, mostram que, dessas famílias, 4,2% não têm renda; 19,4% ganham até um salário mínimo (SM); 28,4% ganham de 1,01 a 2 SM; 25,3%, de 2,01 a 3 SM; 17,8%, de 3,01 a 5 SM; 4,4%, de 5,01 a 10 SM; e 0,4%, de 10,01 a 20 SM. Assim, pode-se perceber que a grande maioria dos moradores é de classe baixa.

Esses informantes convivem com a miséria e suas conseqüências: com o vício do álcool e das drogas, e com a violência dentro e fora de casa. Isso se reflete em suas entrevistas: falavam das cenas de violência as quais presenciavam com uma desconcertante naturalidade. Alguns eram muito tímidos, enquanto outros eram agressivos e se orgulhavam disso. Muitos se vestiam com uniformes já velhos e surrados, e, para vários, a merenda escolar era uma das principais refeições do dia.

Com relação aos informantes de classe média, estes moravam nos bairros Planalto, Liberdade, Jaraguá, Santa Rosa, Ouro Preto ou Centro, todos eles de classe média. Seus pais possuíam nível superior de escolaridade - mesmo incompleto - e tinham acesso aos bens citados anteriormente, característicos de classe média.

O quadro abaixo mostra com maior clareza a distribuição desses informantes, conforme o gênero, a classe social e a escola a que pertenciam.

¹⁰³ Os dados mencionados são do ano de 2000. Atualmente, os números citados deverão ser pelo menos um pouco diferentes.

QUADRO 4

Perfil dos informantes - 2002 - crianças e adolescentes

Número	Idade	Gênero	Classe Social	Escola
1	08/11 anos	Fem.	Baixa	A.Pires
2	08/11 anos	Fem.	Baixa	A.Pires
3	08/11 anos	Fem.	Baixa	A.Pires
4	08/11 anos	Masc.	Baixa	A.Pires
5	08/11 anos	Masc.	Baixa	A.Pires
6	08/11 anos	Masc.	Baixa	A.Pires
7	08/11 anos	Fem.	Média	- ¹⁰⁴
8	08/11 anos	Fem.	Média	A.Pires
9	08/11 anos	Fem.	Média	A. Pires
10	08/11 anos	Masc.	Média	-
11	08/11 anos	Masc.	Média	A. Pires
12	08/11 anos	Masc.	Média	A. Pires
13	12/15 anos	Fem.	Baixa	A. Pires
14	12/15 anos	Fem.	Baixa	A. Pires
15	12/15 anos	Fem.	Baixa	A. Pires
16	12/15 anos	Masc.	Baixa	A. Pires
17	12/15 anos	Masc.	Baixa	A. Pires
18	12/15 anos	Masc.	Baixa	A. Pires
19	12/15 anos	Fem.	Média	CP/UFMG
20	12/15 anos	Fem.	Média	CP/UFMG
21	12/15 anos	Fem.	Média	- ¹⁰⁵
22	12/15 anos	Masc.	Média	CP/UFMG
23	12/15 anos	Masc.	Média	CP/UFMG
24	12/15 anos	Masc.	Média	CP/UFMG

¹⁰⁴ Os informantes F e V, filhos de colegas de trabalho desta pesquisadora, eram vizinhos e tiveram sua entrevista gravada na residência de V, no bairro Planalto, em Belo Horizonte.

¹⁰⁵ A estudante I, de 15 anos, pertencia originalmente ao *corpus* da Prof^a Ana Paula Huback, sendo entrevistada por ela.

3.2.2.3 A coleta de dados dos demais informantes

Os dados de fala dos informantes das outras faixas etárias foram coletados pela Prof^a Ana Paula Huback e, posteriormente, cedidos para este estudo. Este *corpus* é originalmente constituído de trinta entrevistas sociolinguísticas, sendo que, para esta tese, foram aproveitadas 23.

Os informantes deste subgrupo eram amigos da entrevistadora, ou seus colegas de Mestrado, ou pessoas conhecidas ou indicadas a ela por amigos em comum. Isso fez com que a seleção e a coleta dos dados da maioria dos informantes fosse fácil e rápida. Entretanto, como afirma Huback (2003), foi-lhe bastante difícil encontrar e entrevistar pessoas acima de 47 anos nascidas e moradoras em Belo Horizonte – seu critério de seleção dos informantes. Daí seu *corpus* não ser uniformemente constituído nessa faixa etária, com apenas um informante masculino na classe baixa.

As entrevistas foram agendadas por telefone com antecedência e realizadas em maio de 2002, na residência dos informantes, em vários bairros de Belo Horizonte, ou em seu local de trabalho, para que os informantes se sentissem mais à vontade e seu vernáculo fluísse.

Para esta pesquisa, as fitas foram novamente ouvidas e transcritas. Em seguida, os dados foram analisados com base nos fatores internos e externos já citados.

3.2.2.4 O perfil dos informantes das demais faixas etárias

Conforme dito acima, os informantes deste subgrupo eram pessoas conhecidas ou indicadas à pesquisadora por amigos em comum. De modo semelhante ao que aconteceu nos outros *corpora*, os informantes deste subgrupo foram convidados e aceitaram participar da entrevista, sabendo que teriam sua fala gravada. No quadro a seguir, encontra-se uma descrição desses informantes.

QUADRO 5**Perfil dos informantes – 2002 – Faixas etárias III, IV e V**

Número	Faixa Etária	Gênero	Classe Social
1	III (16-30 anos)	F	Média
2	III	F	Média
3	III	M	Média
4	III	M	Média
5	III	F	Baixa
6	III	F	Baixa
7	III	M	Baixa
8	III	M	Baixa
9	IV (31-47 anos)	F	Média
10	IV	F	Média
11	IV	M	Média
12	IV	M	Média
13	IV	F	Baixa.
14	IV	F	Baixa
15	IV	M	Baixa
16	IV	M	Baixa
17	V (acima de 47)	F	Média
18	V	F	Média
19	V	M	Média
20	V	M	Média
21	V	F	Baixa
22	V	F	Baixa
23	V	M	Baixa

3.3 A análise dos dados

Após a transcrição das entrevistas, os dados foram preliminarmente analisados e passados para o Programa Goldvarb, que foi utilizado para se obterem os cálculos necessários às análises posteriores: quantidade, porcentagem e probabilidade de ocorrência de cada uma das três formas.

Esse programa, entretanto, apresenta uma limitação quanto ao cálculo de probabilidade para três variantes, como é o caso desta pesquisa. Assim sendo, esse dado foi obtido por meio de três análises binárias: primeiramente, o Programa foi rodado computando-se os resultados da forma *você* juntamente com *cê*, obtendo-se daí seus pesos relativos, os quais constarão das tabelas deste estudo. Em seguida, o Programa foi rodado computando-se os resultados de *ocê* e *cê*, cujos resultados para a forma *ocê* também constarão das tabelas. Por último, rodou-se novamente o Programa, agora com os resultados de *ocê* e *você*, a fim de que sejam comparados com o peso relativo da forma *ocê* obtido anteriormente. Esses resultados serão apresentados em notas de rodapé, em seguida à apresentação de cada tabela.

De posse desses resultados, procedeu-se às análises quantitativa e qualitativa dos dados, tendo em vista os fatores internos e externos descritos anteriormente.

3.4 Conclusão

Neste Capítulo, foram descritos os procedimentos metodológicos adotados nesta pesquisa: um resumo acerca do referencial teórico que a norteou, a descrição da seleção e do perfil dos informantes, bem como a coleta e o tratamento dos dados dos dois *corpora* – 1982 e 2002.

Tendo em vista o exposto, no próximo Capítulo, será feita a apresentação e a análise dos dados obtidos através do *corpus* de 2002.

CAPÍTULO 4 - A ANÁLISE SINCRÔNICA DOS DADOS - O *CORPUS* DE 2002

Neste capítulo, será feita a análise sincrônica dos dados. Os resultados totais encontrados para as formas *você*, *ocê* e *cê* constam da tabela abaixo:

TABELA 7

Total de ocorrências de *você*, *ocê* e *cê* – *corpus* de 2002

Formas	Ocorrências	%
Você	342	23,5
Ocê	56	3,9
Cê	1055	72,6
Total	1453	100

Tendo sido utilizado o programa Goldvarbrul 2001 para a tabulação dos dados, procedeu-se ao teste de significância, o qual indicou que os seguintes fatores não são significativos para a análise: Classe Social, Tipo de Frase em que aparecem as variantes e posição das variantes nas construções de Tópico.

Quanto ao fator Classe Social, este será retomado e analisado no Capítulo VI, quando do estudo de tempo real, já que ele se mostrou significativo para o *corpus* de 1982. Com

respeito aos fatores Tópico e Tipo de Frase, estes não se mostraram significativos para nenhum dos dois *corpora*, mas, na discussão final dos dados, eles serão comparados aos resultados encontrados por outras pesquisas, com o objetivo de se descrever o comportamento das três formas com respeito a esses fatores, não só na comunidade de Belo Horizonte. No quadro abaixo, apresentam-se os resultados encontrados para cada um deles.

TABELA 8

Distribuição de *você*, *ocê* e *cê* – 2002 – Fatores não-significativos

FATORES NÃO-SIGNIFICATIVOS		VOCÊ		OCÊ		CÊ		TOTAL	
		N	%	N	%	N	%	N	%
Classe Social	Média	130	22,6	20	3,5	424	73,9	574	39,5
	Baixa	212	24,1	36	4,1	631	71,8	879	60,5
Tipo de frase	Declarativa	324	25,1	52	4,0	917	70,9	1292	88,9
	Interrogativa	16	11,0	03	2,0	126	87,0	145	10,0
	Interr. 'Que-Que'	04	25	0	0	12	75	16	1,1
Tópico	Tópico	01	33,3	01	33,3	01	33,3	03	0,2
	Não-tópico	341	23,5	55	3,8	1054	72,7	1450	99,8

A seguir, a apresentação dos fatores considerados significativos no *corpus* de 2002.

4.1 As ocorrências de *você*, *ocê* e *cê*, conforme a Função Sintática

No item 1.1, foram apresentados alguns dos mais importantes trabalhos, adotando-se enfoques teóricos variados, sobre os usos das formas *você*, *ocê* e *cê* em Belo Horizonte, Ouro Preto, São Francisco e Brasília. Todos os trabalhos que analisaram o fator Função Sintática foram unânimes em apontar uma distribuição diferenciada para as três formas. Assim, objetiva-se, neste estudo, aprofundar a análise desta questão, observando-se o comportamento das formas em tempo aparente e em tempo real. Primeiramente, neste Capítulo, serão apresentados os resultados do *corpus* de 2002. No próximo, os resultados do *corpus* de 1982.

Para se avaliar este item, vejam-se os resultados abaixo:

TABELA 9
Distribuição de *você*, *ocê* e *cê*, conforme a Função Sintática - 2002

FUNÇÃO SINTÁTICA	VOCÊ			OCÊ			CÊ			TOTAL	
	N	%	PR	N	%	PR ¹⁰⁶	N	%	PR	N	%
Sujeito	312	22,4	.48	35	2,5	.47	1049	75,1	.52	1396	96,1
Objeto de verbo	04	33,3	.91	05	41,7	.98	03	25,0	.17	12	0,8
Objeto de preposição	26	57,8	.97	16	35,5	.97	03	6,7	.03	45	3,1
TOTAL	342	23,5		56	3,9		1055	72,6		1453	100

Visualizando-se os mesmos resultados através de gráfico¹⁰⁷, tem-se:

¹⁰⁶ A análise binomial de *você* e *ocê* apresenta o seguinte peso relativo para *ocê*: sujeito: .47; objeto de verbo: .89; objeto de preposição: .65. Observe-se que, mesmo se contrapondo à forma padrão da língua, a forma *ocê* é favorecida nas funções de objeto de preposição e, principalmente, de objeto de verbo. Mais adiante, quando da análise da importância dos fatores fonéticos e fonológicos para o uso das três formas, este resultado será explicado.

¹⁰⁷ Os valores constantes para a forma *ocê*, em todos os gráficos desta tese, referem-se à análise binomial feita com relação a *cê*.

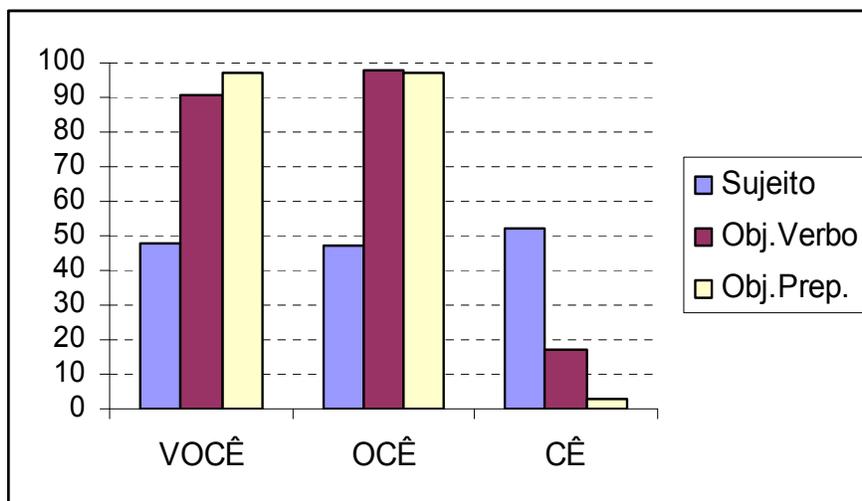


GRÁFICO 1 – Distribuição de *você*, *ocê* e *cê*, conforme a Função Sintática - 2002

O primeiro fato que chama a atenção na tabela e no gráfico acima é realmente a distribuição irregular das formas *você*, *ocê* e *cê* entre as três funções sintáticas. Se 96,1% das ocorrências das três formas se dão na função de sujeito e apenas 3,9% nas outras duas, isso significa que a sua função originária, “natural”, é a de sujeito, ou seja, essa função é não-marcada. Adiante essa questão será discutida.

Esse mesmo resultado revela, por sua vez, que, nas outras duas funções, ou seja, funções de objeto de verbo e de preposição, as três formas sofrem concorrência de outro(s) item(ns) – pronomes ou a forma zero –, como pode ser visto nos exemplos abaixo, retirados do *corpus* de 2002, e em suas possíveis substituições:

- (E 1288) Eu passo e nem soudo *você*.
 - Eu passo e nem *te/o/a* soudo¹⁰⁸.
 - Eu passo e nem soudo \emptyset .¹⁰⁹

¹⁰⁸ É possível ainda a ocorrência de *lhe*, não aceita pela Gramática Normativa.

¹⁰⁹ Nesse caso, falante e ouvinte já sabem quem será saudado.

- (E 967) Não falei *c'ocê* que nós ia ser interrompida?
 - Não *te/lhe* falei que nós ia ser interrompida?
 - Não falei \emptyset que nós ia ser interrompida?

- (E 200) Vou mandar pr'*ocê*.
 - Vou *te/lhe* mandar.
 - Vou mandar \emptyset .

Há que se notar, porém, que nem sempre a substituição é possível, como se vê nos exemplos abaixo:

- (E 266) *Cê* nunca achou... que ela gostasse de *você* não?
 - ... que ela gostasse \emptyset não?
 - * ... que ela *te/lhe* gostasse não?

- (E 233) Depois eu converso *c'ocê*, tá?
 - Depois eu converso \emptyset , tá?
 - * Depois eu *te/lhe* converso.

Por meio desses exemplos, vê-se que, embora alguns contextos não permitam o uso das estratégias acima, elas podem ser concorrentes das formas *você*, *ocê* e *cê* nas duas funções, o que provavelmente justifica a baixa ocorrência das três formas nessas funções sintáticas. Devido a este fato, essas funções são marcadas para as três formas.

Retomando-se agora os resultados da tabela acima, pode-se observar que, na função de sujeito, embora a forma *cê* seja a grande preferida dos informantes, com 75,1% do total das

ocorrências, seu peso relativo é 0.52, ao passo que o de *você* é 0.48 e o de *ocê* é 0.47. Assim, não se pode afirmar que a função de sujeito de fato favoreça a forma *cê*.

Com respeito à função de objeto de verbo, o fato de *cê* ser a forma com o menor número de ocorrências, sendo suplantada até mesmo por *ocê*, contradiz bastante os resultados gerais encontrados para as três formas (cf. Tabela 7). Como se pode perceber pela Tabela 9, acima, realmente essa função desfavorece muito o aparecimento de *cê* (PR = 0.17), favorecendo amplamente *você* e *ocê*, o que comprova os estudos anteriormente citados e a hipótese de que *cê*, neste *corpus*, também apresentaria um comportamento sintático peculiar.

Com respeito à função de objeto de preposição, os resultados indicam que também essa função desfavorece altamente a forma *cê* (PR = 0.03), revelando-lhe assim, igualmente, um comportamento sintático diferenciado. Note-se, porém, que a ocorrência de *cê* como objeto de preposição não era esperada, como atestam alguns dos estudos apresentados no capítulo I, mas aconteceu. Essa função realmente desfavorece seu aparecimento, mas não impossibilitou sua ocorrência, nestes dados.

As frases em que *cê* ocorre como objeto de verbo e como objeto de preposição aparecem a seguir.¹¹⁰

1. Objeto de verbo¹¹¹

(E117) “Eu vou ma:tá:¹¹² *cê*, fdp!” (hwIII)

(E142) “E hoje dinheiro suj<o> te lev<a> *cê* à morte” (hwIII)

(E614) “(...) *cê* vai assistir o futebol *cê* ... nego te arras<ta>¹¹³ *cê* lá dentro” (hwIV)

¹¹⁰ No Anexo 2 desta tese, encontra-se uma lista com todas as ocorrências de *você*, *ocê* e *cê* nessas duas funções sintáticas, nos dois *corpora*.

¹¹¹ As frases contêm os seguintes dados dos informantes: sexo (b = feminino e h = masculino), classe social (r = média e w = baixa) e faixa etária (I = de 08 a 11 anos; II = de 12 a 15; III = de 16 a 30; IV = de 30 a 47; V = acima de 47 anos).

¹¹² O informante pronunciou pausada e enfaticamente a palavra “matar”, alongando suas duas vogais.

2. Objeto de preposição

(E182) “(...) guardar aquilo ali p’*cê*, dentro de si (...)” (hrIII)

E duas frases pronunciadas por uma mesma informante, uma mulher de classe média e da faixa etária IV:

(E683) “Eu vou com [kũ] *cê*”.

(E734) “Xô (Deixe eu) perguntar p’*cê*”.

Neste ponto, cabe uma explicação: todas as seis frases acima passaram por análise acústica, realizadas por meu co-orientador, Prof. Dr. José Olímpio de Magalhães, pela estagiária Camila Tavares Leite e por mim, no Laboratório de Fonética da Faculdade de Letras da UFMG, para se comprovar se realmente os informantes usaram a forma *cê*. Em alguns casos, antes percebidos como *cê*, a análise revelou tratar-se de *ocê*, tendo sido o fonema /o/ pronunciado muito rápida e debilmente. Já com relação ao enunciado (E683), a forma anteriormente entendida e registrada foi *ocê*. Entretanto, o teste acústico não deixou dúvidas: a informante falou realmente “[kũ] *cê*”.¹¹⁴

Com relação à ocorrência de “p’*cê*” nos dois outros enunciados, as análises espectrográficas identificaram um /u/ curtíssimo entre [p] e [se], de 08 ms, no enunciado (734). Tendo em vista que o fonema [o], de *ocê*, medido da mesma informante teve a duração de 60 ms¹¹⁵, não se pôde classificar essa forma como *ocê*, e sim como *cê*, em que houve a assimilação de [u] na labialização de [p].

¹¹³ O informante, ao produzir o enunciado, suprimiu a sílaba “ta”.

¹¹⁴ As análises foram feitas tomando-se como base o estudo de Russo e Behlau (1993).

¹¹⁵ O fonema medido aparece na frase “É, eu também concordo c’*ocê*.” Outros [o] de *ocê* medidos tiveram duração mais ou menos semelhante a esse: 57,1 ms em “(...) mas às vezes te jog’*ocê* pro buraco” e 44,7 ms em “(...) te jog’*ocê* pra trás.” (homem, faixa III, classe baixa); e 44,5 ms em “Vou pegar um pouquinho p’*ocê*” (homem, faixa III, classe baixa).

Já com relação ao “p’cê” do enunciado (182), acima, a labial [p] assimilou totalmente o [u], sendo que as análises espectrográficas não identificaram qualquer dos formantes dessa vogal. Assim, a expressão pronunciada pela informante é, realmente, [pse].

Ainda sobre a forma *cê* como objeto de preposição, merece ser destacado o fato de que isso também ocorre no Espírito Santo, onde esta pesquisadora reside atualmente, ao que parece independentemente de gênero, idade e profissão dos falantes. Algumas das várias frases ouvidas são estas:¹¹⁶

- 1) “Eu olho ele p’cê.” (homem, jovem, contador, residente em Vila Velha)
- 2) “Deix’eu embrulhar p’cê.” (mulher, jovem, farmacêutica, Vila Velha)
- 3) “Eu levo ele lá p’cê” (homem, adulto, lavador de carros, Vila Velha)
- 4) “Eu vou trocar p’cê” (homem, adulto, professor universitário, Vitória)
- 5) (A retribuição a um agradecimento) “Obrigad’cê” (mulher, adulta, vendedora de água de coco, Vila Velha)
- 6) “Eu vou botar a Internet p’cês” (homem, jovem, técnico em informática, Cachoeiro de Itapemirim)
- 7) “Peraí que eu vou olhar p’cê” (mulher, adulta, assistente administrativo, Vila Velha)
- 8) “Eu olho p’cê” (homem, adulto, gerente de banco, Vila Velha)
- 9) “Eu vou pegar um copinho p’cê” (mulher, adulta, professora universitária, Cachoeiro de Itapemirim)
- 10) “Deix’eu me despedir d’cê” (mulher, jovem, professora universitária, Cachoeiro de Itapemirim)¹¹⁷

¹¹⁶ As frases foram ouvidas e registradas por mim. Como foram ouvidas em interações cotidianas, não foi possível gravá-las, apenas transcrever sua pronúncia o mais rápido possível.

¹¹⁷ Após essa jovem professora ter dito tal frase, perguntei-lhe se ela havia realmente dito “d’cê”, ao que confirmou. É importante mencionar que essa moça é natural de Santo André/SP, residindo em Cachoeiro de

Essas frases demonstram que o uso de *cê* como objeto de preposição não ocorre somente em Belo Horizonte e São Francisco (cf. Coelho, 1999), embora não ocorra em outros lugares, como Brasília (cf. Andrade, 2004). Mais adiante, na análise qualitativa dos dados, este ponto será retomado.

4.2 As ocorrências de *você*, *ocê* e *cê*, conforme a Contigüidade

O fator Contigüidade foi analisado para que os resultados desta pesquisa sejam comparados aos estudos anteriores, apresentados no item 1.1. Primeiramente serão apresentados os resultados do *corpus* deste estudo, os quais se encontram na tabela abaixo; em seguida, os resultados de outras pesquisas anteriores. Antes, porém, cumpre observar que a análise deste fator levou em consideração apenas as ocorrências das formas na função de sujeito, tendo em vista que interessa avaliar a contigüidade das formas em relação ao verbo.

TABELA 10
Distribuição de *você*, *ocê* e *cê*, conforme a Contigüidade - 2002

CONTIGÜIDADE	VOCÊ			OCÊ			CÊ			TOTAL	
	N	%	PR	N	%	PR ¹¹⁸	N	%	PR	N	%
Contíguo	245	21,2	.48	30	2,6	-	881	76,2	.52	1156	82,8
Não-contíguo	68	28,3	.58	05	2,1	-	167	69,6	.42	240	17,2
TOTAL	313	22,4		35	2,5		1048	75,1		1396	100

Abaixo tem-se um gráfico com os mesmos resultados:

Itapemirim há apenas três anos. Esse fato indica que é possível que em Santo André também seja encontrada a expressão “d’cê”.

¹¹⁸ Em nenhuma das duas análises binomiais com a forma *ocê* o fator Contigüidade foi selecionado como significativo, pelo Programa Goldvarb.

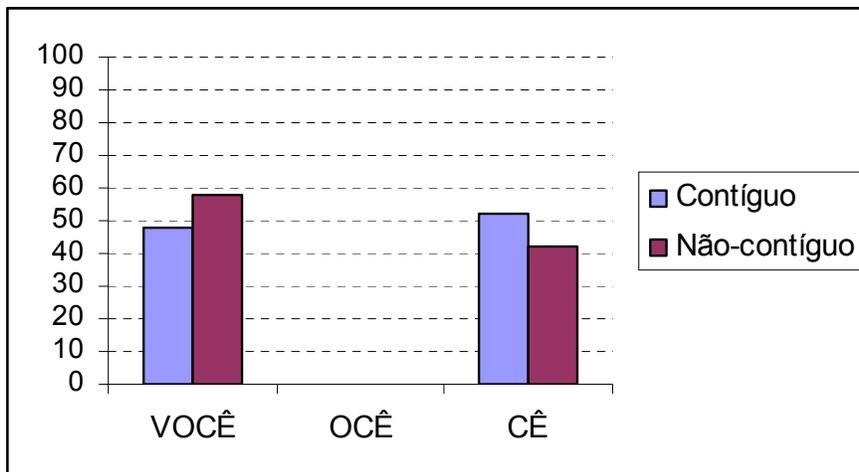


GRÁFICO 2 – Distribuição de *você*, *ocê* e *cê*, conforme a Contigüidade – 2002

Para se fazer a comparação entre todos os dados, é interessante que se recordem os resultados dos trabalhos anteriores que se ocuparam deste fator.

Os dados de Ramos (1997) e Andrade (2004) evidenciam que *cê* é favorecida quando a forma se encontra contígua ao verbo e é desfavorecida quando não está contígua a ele. No *corpus* de Ramos, os elementos interpolados são advérbios curtos, como *não*, *já*, *só* e *sempre*. Já Andrade encontrou clíticos (*me*, *te*, *se*, *num*), monossílabos tônicos, com e sem ditongação (*já*, *só*, *lá*; *nem*, *mais*) e elementos de natureza não-clítica¹¹⁹ (*não*, *nunca*, *menos*, *sempre*, *mesmo*, *também*, *aqui*, *então*, locução adverbial). Por sua vez, Barbosa (2005), analisando o *corpus* de Coelho (1999), também encontra vários elementos interpolados entre *cê* e o verbo, como os exemplos abaixo (2005, p. 55):

“*Cê* nunca mais me chame assim, hein?”

“*Cê*, que tá pra ganhá bebê e tem pressão alta, deve repousar bastante”

¹¹⁹ Andrade (2004, p. 51-2) dá os seguintes exemplos de frases com a forma *cê* apresentando elementos intervinientes de natureza não-clítica: “*Cê* já foi também?”; “*Cê* nem protege sua irmã?”; e “*Cê*, lá na sua escola, custuma contá istória?”

“*Cê*, que é filho de Deus, saberá o que fazer quando chegar a hora.”

Pode-se então, agora, proceder à análise comparativa. A tabela 10, acima, permite observar que, no caso de as formas estarem contíguas ao verbo, não há propriamente o favorecimento de nenhuma delas, ou seja, apesar de *cê* ser a forma preferida nesses casos (76,2%), o seu PR = .52 não indica, de fato, que essa forma seja favorecida. Quanto a isso, é interessante notar que os pesos relativos de *você* e *cê* são absolutamente os mesmos, tanto para os casos de contigüidade quanto de função sintática de sujeito, e muito semelhantes aos casos de não-focalização, isto é, em casos de contextos não-marcados. Mas esta discussão será deixada para a análise final dos resultados deste *corpus*.

Voltando-se aos dados exibidos na tabela 10, o que chama a atenção é o desfavorecimento, embora sutil, da forma *cê* e o leve favorecimento de *você*, confirmando assim os estudos de Ramos (1997) e Andrade (2004). Para a forma *ocê*, o Programa Goldvarb não selecionou a Contigüidade como fator significativo.

Para a forma *cê*, no *corpus* de 2002 deste trabalho, foram encontrados, como elementos intercalados:

(i) advérbios, alguns curtos e outros mais longos:

JÁ:	(E 208)	“E multiprocessador, <i>cê já</i> viu?” (brIII)
NUM:	(E 590)	“Claro que <i>cê não</i> mistura homem com mulher.” (hwIV)
NEM:	(E 121)	“... <i>cê nem</i> sabe se vai voltar ali.” (hwIII)
SÓ:	(E 1118)	“A gente ia bem vestido, <i>cê só</i> tomava Cuba Libre.” (hrV)
AINDA:	(E 732)	“ <i>Cê ainda</i> vai falar assim.” (brIV)
TAMBÉM:	(E 223)	“ <i>Cê também</i> acha?” (brIII)
SEMPRE:	(E 854)	“(...) as palavras <i>cê sempre</i> assimila.” (hrIV)

NUNCA: (E 665) “*Cê nunca acampou?*” (brIV)

(ii) operadores

(E 408) “*Cê, né?, tem que ter uma coisa assim...*” (hrIII)

(iii) clíticos

(E 561) “*Se cê me perguntar...*” (hwIV)

(E 1233) “*O pior é que agora cê me pegou.*” (hrI)

(iv) dois elementos

(E 256) “*Cê não me conta nada!*” (brIII)

(E 1223) “*Cê já me perguntou essa.*” (brIII)

(E 320) “*Cê não é mais um... cê já não é mais um leigo.*” (hrIII)

E, além desses,

(v) pronomes relativos

(E 641) “*Ainda mais cê que é branquinha, cheia de pano assim...*” (brIV)

(E 672) “*Cê que mora na região do Ouro Preto, ela faz vela de citronela.*” (brIV)

(E 787) “*Não, cê que vai me falar isso.*” (brIV)

Já com relação à forma *você*, encontram-se os advérbios listados acima e também elementos intercalados de extensão um pouco maior que os encontrados para a forma *cê*, como se pode observar na frase abaixo:

(E 769) “Se *você simplesmente* decreta: o aluno não pode repetir o ano...” (brIV)

e nas seguintes, pronunciadas por (hrV):

(E 1027) “*Você, dentro de um ônibus*, não conversava com quem ia ao lado.”

(E 1033) “*Você, às duas da manhã*, tá com um movimento de trânsito tão grande...”

(E 1094) “Se *você, em um ano*, deixar de tratar... *você* deixou de atender uma cidade de 60.000 habitantes.”

Com a forma *ocê*, aparecem apenas os advérbios curtos *não* e *já*, como os exemplificados abaixo:

(E 87) “Novela, se *ocê não* acompanhar...” (bwIII)

(E 664) “*Cê não* pode não, q’*ocê já* tá novo demais!” (brIV)

O desfavorecimento de *cê* em casos de não-contigüidade confirma a hipótese de cliticização dessa forma, pois entre um clítico nominativo e o verbo não deve ocorrer qualquer elemento (RAMOS, 1997, p.55). Por outro lado, nos dados dessa autora ocorreram advérbios curtos como elementos intercalados entre a forma *cê* e o verbo, mas isso não desconfirma a hipótese de cliticização de *cê* (RAMOS, 1997, p.56). Por sua vez, essa não-contigüidade levou Andrade (2004) a afirmar que *cê* ainda não seria propriamente um clítico, mas sim um pronome fraco (2004, p. 129)¹²⁰. Já Barbosa (2005) afirma que a não-contigüidade comprova que *cê* não é clítico.

¹²⁰ Vitral, em Ciríaco, Vitral e Reis (2004), como já foi dito, reformulou sua hipótese inicial, estipulando que a forma *cê* ainda não é um clítico pleno.

Não é objetivo deste estudo testar a hipótese de cliticização da forma *cê*, mas os dados deste *corpus* confirmam que, embora as intercalações mais longas se dêem preferentemente com a forma *você*, não é impossível que *cê* também possa ser usado nesses casos, como também mostraram os dados de Andrade (2004) e Barbosa (2005). Assim, os dados deste estudo com respeito à Contigüidade das formas em relação ao verbo também apresentam um problema para essa hipótese.

4.3 As ocorrências de *você*, *ocê* e *cê*, conforme a Referência

Conforme foi visto no item 1.1, alguns dos trabalhos já realizados sobre as formas *você*, *ocê* e *cê* encontraram resultados conflitantes em relação à expressão da referência. Por isso, um dos propósitos estipulados para esta pesquisa é ampliar essa discussão quanto à cidade de Belo Horizonte, apresentando um número maior de dados que os anteriores e apresentando também dados do tempo real.

Tendo em vista esses resultados, na Introdução desta tese estipulou-se a hipótese de que as formas *você* e *cê* estariam passando por um processo de especialização, em que a forma *cê* estaria sendo preferencialmente usada para expressar referência indefinida e *você*, a referência definida.

Com esses objetivos em mente, apresentam-se os resultados encontrados neste *corpus*, constantes da tabela abaixo:

TABELA 11

Distribuição de *você*, *ocê* e *cê*, conforme a Referência – Todas as funções - 2002

REFERÊNCIA	VOCÊ			OCÊ			CÊ			TOTAL	
	N	%	PR	N	%	PR ¹²¹	N	%	PR	N	%
Definida	108	15,1	.36	36	5,1	-	569	79,8	.64	713	49,1
Indefinida	234	31,6	.63	20	2,7	-	486	65,7	.37	740	50,9
TOTAL	342	23,5		56	3,9		1055	72,6		1453	100

Observando-se os resultados em forma de gráfico, tem-se:

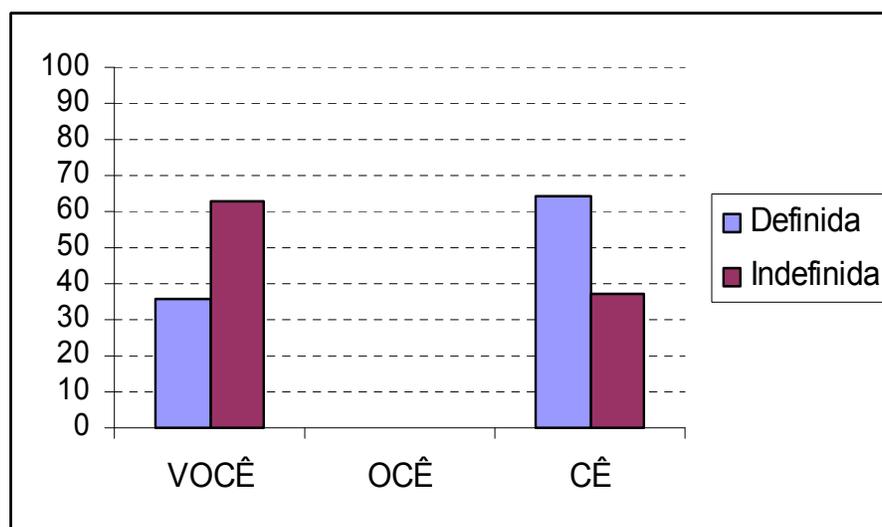


GRÁFICO 3 – Distribuição de *você*, *ocê* e *cê*, conforme a Referência – Todas as funções - 2002

¹²¹ Na análise binomial de *cê* e *ocê*, o Programa não selecionou o fator Referência. Entretanto, na análise de *você* com *ocê*, o fator foi selecionado e apresentou os seguintes PR para *ocê*: referência indefinida: .43; referência definida: .62. Percebe-se, assim, o leve desfavorecimento de *ocê* para expressar referência indefinida e o seu favorecimento para a definida.

Pela tabela acima, vê-se que a forma *cê* é a preferida tanto para indicar referência definida quando indefinida (79,8% e 65,7%, respectivamente). Entretanto, quando se observa o peso relativo das formas, vê-se que a referência definida favorece a forma *cê* (PR = 0.64) e desfavorece *você* (PR = .36).

O contrário se dá quando a referência é indefinida, a qual favorece a forma *você* (PR = 0.63) e desfavorece razoavelmente a forma *cê* (PR = 0.37). Esses resultados gerais vão, assim, de encontro à hipótese assumida nesta tese.

Entretanto, as diferenças com respeito à referência encontradas nos trabalhos citados no item 1.1 deveram-se ao fato de que, enquanto Ramos (1997) analisou a referência das formas apenas na função de sujeito, Alves (1998) e Coelho (1999) procederam à análise levando em conta todas as funções sintáticas. Portanto, neste estudo, também é preciso analisar as três formas somente na função de sujeito. Os resultados encontrados estão na Tabela 12, a seguir:

TABELA 12
Distribuição de *você*, *ocê* e *cê*, conforme a Referência – Função de sujeito - 2002

	VOCÊ			OCÊ			CÊ			TOTAL	
	N	%	PR	N	%	PR ¹²²	N	%	PR	N	%
Definida	88	13,1	.36	17	2,5	-	565	84,3	.64	670	48,0
Indefinida	225	31,0	.63	18	2,5	-	483	66,5	.37	726	52,0
TOTAL	313	22,4		35	2,5		1048	75,1		1396	100

Os resultados também podem ser vistos no gráfico a seguir.

¹²² Na análise binomial de *cê* e *ocê*, este fator não foi selecionado pelo Programa. Entretanto, o fator foi selecionado na análise de *você* e *ocê*, apresentando os seguintes pesos relativos para *ocê*: Referência indefinida: .44; definida: .64. Observa-se assim, que os valores para *ocê* praticamente não se alteram, observando-se esse fator com todas as funções e somente na função de sujeito.

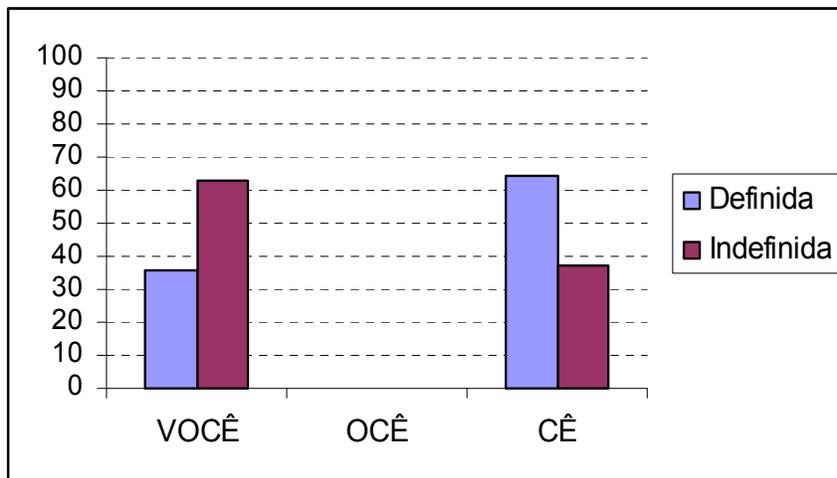


GRÁFICO 4 – Distribuição de *você*, *ocê* e *cê*, conforme a Referência – Função de Sujeito - 2002

Pela tabela 12, vê-se que as posições praticamente não se alteram quando se analisa a referência das formas apenas na função de sujeito. Os pesos relativos são absolutamente iguais para *você* e *cê*, modificando-se minimamente apenas com respeito à forma *ocê*, na análise binomial de *ocê* e *você*. Assim, por esses resultados, apenas, pode-se concluir que os dados deste *corpus* realmente não confirmam a hipótese de que *você* está sendo preferencialmente usado para expressar referência definida e *cê*, referência indefinida, que foi assumida nesta pesquisa.

Entretanto, é importante verificar como se comportam as formas *você* e *cê* – as que estão sendo particularmente discutidas aqui – de acordo com cada faixa etária. Visualizando-se os mesmos resultados e observando-se apenas a função de sujeito, tem-se:

TABELA 13
Distribuição de *você*, *ocê* e *cê*, conforme a Referência e a Idade
Função de Sujeito – 2002

Faixa Etária	VOCÊ						CÊ						TOTAL	
	Ref. Definida			Ref.Indefinida			Ref. Definida			Ref.Indefinida			D	I
	N	%	PR	N	%	PR	N	%	PR	N	%	PR		
V (+ 47)	21	18,4	.56	96	50,8	.68	93	81,6	.44	93	49,2	.32	114	189
IV (31-47)	24	12,6	.49	61	29,7	.47	167	87,4	.51	144	70,3	.53	191	205
III (16-30)	08	5,7	.34	53	19,8	.37	133	94,3	.66	215	80,2	.63	141	268
II (12-15)	10	22,2	.63	08	20,5	.37	35	77,8	.37	31	79,5	.63	45	39
I (08-11)	25	15,4	.52	07	100	-	137	84,6	.48	0	0	-	162	07
Total	88	13,4		225	31,8		565	86,6		483	68,2		653	708

A tabela 13 traz resultados interessantes. Com respeito à referência definida, observa-se um crescente aumento da probabilidade de uso de *cê* até a faixa III, o que indica mudança em progresso. A faixa II, por sua vez, apresenta um aumento significativo tanto do percentual quanto da probabilidade de ocorrência da forma *você*, provavelmente devido à situação de entrevista, como será discutido à frente. Com respeito à faixa I, as probabilidades de *você* e *cê* voltam a ficar equilibradas, não havendo nenhuma forma realmente favorecida aí. Diante disso, pode-se pensar que a forma *cê* está sendo preferencialmente usada para expressar referência definida, já que o favorecimento de *você* pela faixa II diminui bastante na faixa I. Como já se disse, os resultados das crianças e adolescentes serão discutidos mais adiante.

Com relação à referência indefinida, os dados não deixam dúvidas de que a forma *cê* está ocupando gradativamente o espaço de *você*, inclusive com o seu favorecimento pela faixa

II, o que não havia acontecido anteriormente – compare-se a probabilidade de uso de *você* (PR = .36) e de *cê* (PR = .64). Outro dado curioso quanto a esta referência é seu reduzidíssimo número de ocorrências na faixa I. Das doze crianças entrevistadas, apenas uma menina de oito anos, de classe média, produziu enunciados com referência indefinida, o que indica que esse uso não é comum nas crianças. Se pensarmos que esta referência é uma abstração da realidade, uma análise sobre uma situação hipotética, e, segundo Piaget (2003, p. 63), o pensamento hipotético-dedutivo se inicia a partir dos doze anos, justifica-se, então, o número praticamente inexistente deste tipo de referência entre as crianças.

Concluindo, os resultados gerais para o fator Referência indicam que a forma *cê* é usada preferencialmente para indicar referência definida, sendo a forma *você* deixada para expressar referência indefinida. Entretanto, analisando-se o comportamento de cada faixa etária, vê-se um crescimento da forma *cê* também para expressar referência indefinida. Se atentarmos para os resultados da faixa V, notaremos uma probabilidade maior de uso de *cê* com referência definida, o que indica que a forma era preferencialmente usada com essa referência. Assim, sua trajetória parece assemelhar-se à da forma *você*, a qual, inicialmente, era usada para expressar referência definida, passando posteriormente a expressar também a indefinida. Em suma, os resultados apontam para a expansão do uso da forma *cê* e sua conseqüente vitória sobre a forma *você*.

4.4 As ocorrências de *você*, *ocê* e *cê*, conforme o Foco

Conforme foi visto no capítulo I, os dados de alguns dos trabalhos já realizados sobre *você*, *ocê* e *cê* demonstram que *cê*, que era a favorita das formas em contextos de não-focalização, não aparecia recebendo foco. Já Barbosa (2005) apresenta o seguinte exemplo¹²³ de focalização da forma *cê* (2005, p. 82):

¹²³ Os exemplos de Barbosa (2005) baseiam-se em dados reais, mas também intuitivos, como foi citado anteriormente, e a autora não especifica quando ocorre cada um desses dois casos. Assim, não há como saber se o exemplo acima é de um informante ou seu mesmo, e, pela formalidade do clítico em “o conheço”, parece tratar-se de um dado intuitivo.

“ – Maria, nós vamos comprar esse livro.

- *Cê* vai, eu não. Eu já o conheço e não gosto nadinha dele.”

Tendo em vista mais um fator em que *cê* apresenta comportamento peculiar, optou-se por analisá-lo neste trabalho. Os resultados encontrados neste *corpus* são os seguintes:

TABELA 14

Distribuição de *ocê*, *ocê* e *cê*, conforme o Foco - 2002

FOCALIZAÇÃO	VOCÊ			OCÊ			CÊ			TOTAL	
	N	%	PR	N	%	PR ¹²⁴	N	%	PR	N	%
Foco	08	72,7	.93	02	18,2	-	01	9,1	.07	11	0,8
Não-foco	334	23,2	.50	54	3,7	-	1054	73,1	.50	1442	99,2
TOTAL	342	23,5		56	3,9		1055	72,6		1453	100

Visualizando-se a tabela acima em forma de gráfico, tem-se:

¹²⁴ Em nenhuma análise binomial com a forma *ocê* o Programa selecionou o fator Foco.

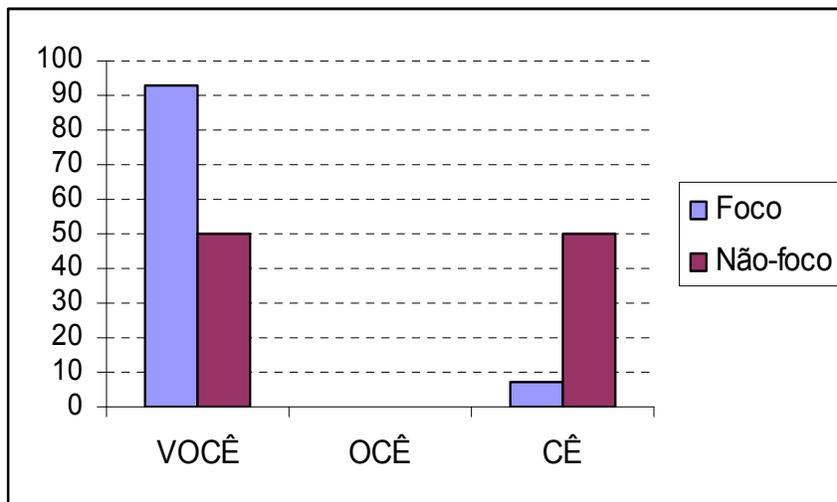


GRÁFICO 5 - Distribuição de *você*, *ocê* e *cê*, conforme o Foco – 2002

A tabela 13 mostra que, em situações normais, de não-focalização, que representam 99,2% dos dados, tanto *você* quanto *cê* têm ocorrências muito próximas do total geral encontrado para ambas as formas (23,2% e 73,1%, respectivamente). Assim, no caso de não haver focalização, seus pesos relativos indicam que nenhuma delas é favorecida. Já a forma *ocê* não foi considerada significativa pelo Programa, nem com a forma *você*, nem com a forma *cê*.

Com relação a este fator, novamente o que chama a atenção é a ocorrência das formas nos pouquíssimos casos de focalização (apenas 0,8% do total). Nota-se aí que a forma *cê* é altamente desfavorecida: seu peso relativo é 0,07, o que confirma, mais uma vez, a hipótese de que *cê* apresenta um comportamento diferenciado das demais, como alguns dos estudos anteriores já mostraram.

Os casos de focalização, neste *corpus*, foram estes:

Com você:

- (E 36) “Esse negócio de **VOCÊ** conscientizar (...)”
- (E 553) “Pra **VOCÊ** ter uma família (...)”
- (E 589) “**VOCÊ** é diferente da sua colega, (...)”
- (E 860) “(...) pra ter conhecimento a **VOCÊ**, eu teria (...)”
- (E 910) “Mas, veja **VOCÊ**, o que mais mudou (...)”
- (E 1038) “(...) de um lugar pra outro, **VOCÊ**, tem um lugar de passagem.”
- (E 1199) “Eu não vou deixar nada pra **VOCÊ**.”

Com ocê:

- (E 63) “Se **OCÊ** faltou uma hora pr'ocê sair (...)”
- (E 724) “Cê tá falando que eu menti pr'**OCÊ**?”

Com cê:

- (E 117) “Eu vou matar **CÊ**, fdp!”

Antes de comentar esses exemplos, é preciso mencionar uma observação feita por Labov (1972). Nesse trabalho, o autor relata sua pesquisa sobre a realização do (r) em final de sílaba e de palavra (*fourth floor*, respectivamente) em três lojas de departamentos de Nova York – Saks Fifth Avenue, Macy’s e S. Klein -, com dois tipos de pronúncia: casual e enfática.

Os resultados desse estudo mostram que os falantes, quando enfatizam a pronúncia, tendem a evitar a forma não-padrão – o apagamento do (r). Assim, o autor afirma: “It will be

important to bear in mind that this attitude – that (r-1) is the most appropriate pronunciation for emphatic speech – is shared by at least some speakers in all three stores” (LABOV, 1972, p. 52). Sendo o (r-1) a realização do (r), ou seja, a pronúncia padrão no inglês americano, percebe-se que os falantes preferem, nos casos de ênfase, o uso de variantes-padrão.

Neste trabalho, pode-se fazer uma análise semelhante. Sendo *você* a forma padrão, ela seria a preferencialmente usada nos casos de foco enfático. Além disso, dentre as três formas, *você* é a prototípica da 2ª pessoa; por conseguinte, ela irá ocorrer tanto no contexto não-marcado – de não-focalização – quanto no marcado – de focalização. E, realmente, a forma *você* foi maioria e é fortemente favorecida aí (PR = .93).

A forma *ocê* não foi considerada significativa pelo Programa Goldvarb, nem na análise binária com *você*, nem com *cê*. É importante mencionar que, de suas duas únicas ocorrências (18,2% dos casos), uma encontra-se na posição de objeto de preposição, que a favorece muito. Por outro lado, nesse mesmo contexto – diante de uma preposição –, a forma *você* foi usada num maior número de casos que a forma *ocê*. Observem-se os enunciados (36), (553) e (1199). Esse fato pode indicar que a forma *ocê* é evitada pelos informantes nos casos de foco ou ênfase.

Em relação ao enunciado (117), a ocorrência de *cê* é surpreendente, tendo em vista que: i) a focalização implica a pronúncia de um item com maior intensidade e, por conseguinte, com um pouco mais de lentidão, e que a forma *cê* é uma redução do pronome *você*, sendo o resultado do processamento da fala, normalmente rápido; ii) a forma *cê*, no enunciado (117), exerce a função de objeto de verbo, que a desfavorece bastante.

Isso posto, o enunciado (117) seria uma situação esdrúxula, um “equivoco”, uma expressão inusitada do informante? Ou seria o início de um processo de mudança, com a forma *cê* espraiando-se para outros contextos? Por meio deste único dado não é possível responder a essas perguntas, já que a mudança lingüística envolve a adoção da forma inovadora por

pelo menos um grupo da comunidade. Portanto, esta discussão ficará suspensa, por enquanto. No capítulo destinado à análise da mudança em tempo real, ela retornará.

Tendo sido expostos os resultados deste estudo quanto aos fatores internos, serão apresentados, a seguir, os fatores externos analisados.

4.5 As ocorrências de *você*, *ocê* e *cê*, conforme a Idade

Ao se analisar o fator Idade, busca-se verificar se os dados deste *corpus* revelam mudança em progresso com relação às formas *você*, *ocê* e *cê*. Para isso, observe-se a tabela abaixo:

TABELA 15
Distribuição de *você*, *ocê* e *cê*, conforme a Idade - 2002

FAIXA ETÁRIA	VOCÊ			OCÊ			CÊ			TOTAL	
	N	%	PR	N	%	PR ¹²⁵	N	%	PR	N	%
V (mais de 47)	119	37,2	.64	13	4,0	-	188	58,8	.33	320	22,0
IV (de 31 a 47)	95	22,7	.51	11	2,6	-	313	74,7	.49	419	28,8
III (de 16 a 30)	72	16,1	.34	23	5,2	-	351	78,7	.66	446	30,7
II (de 12 a 15)	22	25	.51	0	0	-	66	75,0	.49	88	6,1
I (de 08 a 11)	35	19,4	.58	08	4,5	-	137	76,1	.42	180	12,4
TOTAL	342	23,5		56	3,9		1055	72,6		1453	100

Visualizando-se a tabela em forma de gráfico, tem-se:

¹²⁵ Em nenhuma das análises binomiais com *ocê*, o fator Idade foi selecionado como significativo pelo Programa.

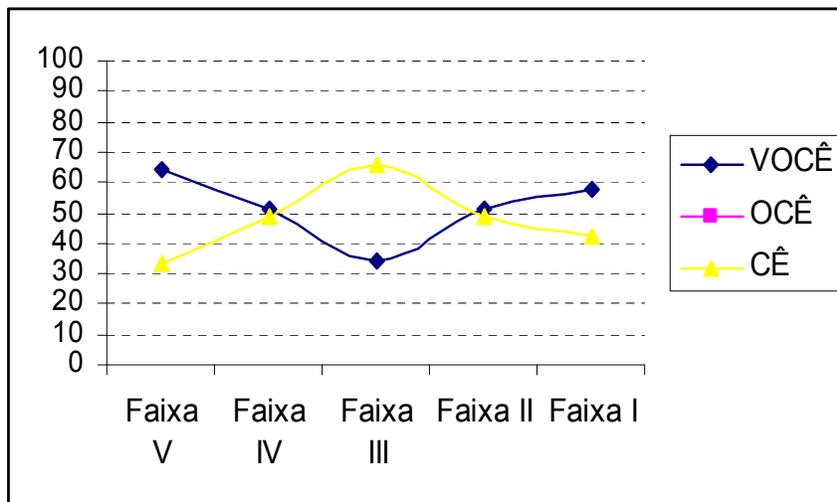


GRÁFICO 6 - Distribuição de *você*, *ocê* e *cê*, conforme a Idade - 2002

Os dados da tabela 15 mostram que a forma *cê* ocorre em maior número que *você* em todas as faixas etárias. Entretanto, seu uso é bastante desfavorecido na faixa V (PR = .33) e levemente desfavorecido na faixa I (PR = .42). Já nas faixas II e IV, não há necessariamente um favorecimento (PR = .49), ao contrário dos informantes de 16 a 30 anos, que a favorecem (PR = .66). Observando-se a probabilidade de ocorrência de *cê* nas faixas V a III, nota-se que ela aumenta progressivamente, indicando uma mudança em curso.

Por sua vez, os resultados das faixas I e II parecem provar o contrário, já que há uma diminuição gradativa da probabilidade de uso de *cê*. Entretanto, uma observação atenta dos percentuais das faixas I a IV permite ver que são bastante semelhantes, variando de 74,7% para a faixa IV – a menor das quatro –, a 78,7% para a faixa III – a maior delas –, ou seja, uma diferença de apenas 4%. Assim, não se pode dizer que as crianças e os adolescentes estejam efetivamente retardando a mudança ou mesmo fazendo-a retroceder.

Com respeito à forma *você*, nota-se exatamente o oposto: a diminuição progressiva tanto dos percentuais quanto da probabilidade de ocorrência nas faixas V a III. Nas faixas II e I,

vê-se o aumento da porcentagem e da probabilidade, cujos índices aproximam-se ou até suplantam os da faixa III. Como explicação para o fato, pode-se pensar que esse comportamento seja decorrente da formalidade que a situação de entrevista sugere e, muito provavelmente, da insegurança que isso lhes gerou. Mais adiante este assunto voltará a ser discutido.

Já com respeito à forma *ocê*, esta, de novo, demonstrou um comportamento bastante irregular: seu percentual de ocorrência passa de 4% (faixa V) para 2,6% (faixa IV), volta a subir para 5,2% (faixa III), cai para 0 (faixa II) e sobe novamente para 4,5% (faixa I). Assim, não há o menor indício de que esteja sofrendo qualquer mudança em curso, nem de aumento, nem de diminuição de seu uso.

Antes de encerrar este subitem, seria importante analisar os resultados obtidos para cada faixa etária separadamente, como já se fez brevemente com as faixas I e II.

Como disse Chambers (1995), o desejo de ascensão e o exercício de certas profissões mais valorizadas socialmente levam o jovem adulto a mudar sua linguagem, preferindo as formas-padrão da língua. Depois de passada essa fase, a linguagem do indivíduo se cristaliza, não se percebendo alterações mais profundas nela. Se isso ocorrer, será preferencialmente no nível lexical (cf. capítulo 1 desta tese). Assim, pode-se perguntar se os resultados desta pesquisa podem comprovar pelo menos uma dessas informações, já que não foi possível fazer-se um estudo de painel.

É fato que, dentre as três formas, *você* é a padrão, como atestam as Gramáticas Normativas da língua. Assim sendo, os informantes das faixas III e IV, seguindo os princípios mencionados no parágrafo anterior, deveriam usar mais a forma *você* que os outros informantes. Entretanto, não é isso o que os dados demonstram: a forma *você* está progressivamente caindo entre as faixas V e III, e o inverso se dá com a forma *cê*¹²⁶.

¹²⁶ No caso das faixas I e II, a hipótese estipulada é de que os informantes podem ter-se ressentido da situação de entrevista e formalizado mais sua fala, como já foi mencionado.

O que esses resultados indicam com relação ao comportamento desses informantes? Ao que parece, eles não percebem *cê* como uma forma que deva ser evitada na língua oral, ou seja, não a percebem como uma forma estigmatizada. E, realmente, *cê* é o resultado de um processo natural na língua, que é a redução fonética de itens lexicais constantemente usados, como já foi abordado anteriormente e como será discutido no final deste estudo. Assim, se os jovens adultos deste *corpus* não usam mais *você* que *cê*, não é por falta de ambição profissional, ou de pressão do mercado pelo uso de formas-padrão, mas sim porque a forma *cê* já está implementada em Belo Horizonte, sendo usada indistintamente pela comunidade.

4.6 As ocorrências de *você*, *ocê* e *cê*, conforme o Gênero

Uma análise do fator Gênero busca verificar se o que a literatura sociolinguística afirma a respeito do comportamento de homens e mulheres em relação à língua, conforme mostrado no subitem 1.2.5.2, se aplica à variação de uso de *você*, *ocê* e *cê*. Assim, espera-se conhecer um pouco mais a respeito desse uso por parte dos informantes femininos e masculinos desta pesquisa. Para tanto, vejam-se os resultados obtidos, constantes da tabela abaixo:

TABELA 16
Distribuição de *você*, *ocê* e *cê*, conforme o Gênero - 2002

GÊNERO	VOCÊ			OCÊ			CÊ			TOTAL	
	N	%	PR	N	%	PR ¹²⁷	N	%	PR	N	%
Feminino	123	17,2	.43	33	4,6	-	559	78,2	.57	715	49,2
Masculino	219	29,7	.57	23	3,1	-	496	67,2	.43	738	50,8
TOTAL	342	23,5		56	3,9		1055	72,6		1453	100

¹²⁷ O fator Gênero não foi selecionado como significativo pelo Programa Goldvarb, em nenhuma das análises binomiais envolvendo a forma *ocê*.

Observando-se essa distribuição em forma de gráfico, tem-se:

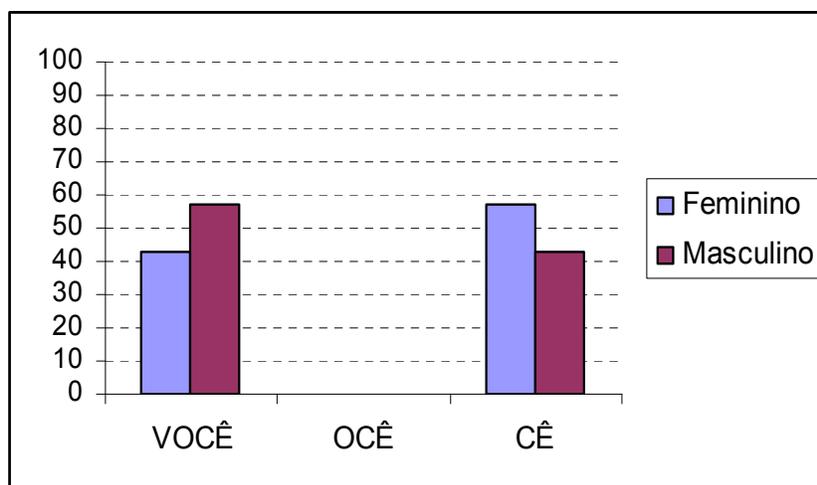


GRÁFICO 7 – Distribuição de *você*, *ocê* e *cê*, conforme o Gênero – 2002

Pelos resultados acima, vê-se que o uso da forma *cê* é ligeiramente favorecido pelas mulheres (PR = .57), da mesma forma que os homens favorecem ligeiramente a forma *você*, com PR igualmente .57. Sabendo-se que as mulheres tendem a rejeitar as formas estigmatizadas, devido a sua posição socialmente desfavorecida na sociedade, como se mencionou no capítulo I, conclui-se que a forma *cê* realmente não sofre rejeição pelos falantes, sendo normalmente aceita e usada em Belo Horizonte.

Com relação à forma *ocê*, embora ela não tenha sido selecionada pelo Programa Goldvarb como significativa, em nenhuma das análises binomiais, a porcentagem de uso de *ocê* pelas mulheres é maior que seu uso pelos homens: 4,6% contra 3,1%. Já a forma *você* é relativamente desfavorecida pelas mulheres (PR = .43) e relativamente favorecida pelos homens (PR = .57). É claro que não se pode pensar que a forma *você* sofra qualquer tipo de rejeição ou estigma na comunidade. Esses resultados apenas indicam que a forma *cê* está realmente vencendo a competição com *você*, e as mulheres estão liderando esse processo.

Outro aspecto mencionado na literatura sociolinguística diz respeito ao fato de a diferença entre homens e mulheres ser mais acentuada na faixa etária mais alta, sendo que, quando se trata de mudança em progresso, na geração seguinte, os homens tendem a seguir o mesmo padrão das mulheres. Para verificarmos se essa afirmação é válida para a mudança de *você* para *cê*, veja-se a tabela a seguir, em que aparecem os resultados de acordo com o gênero e a idade dos informantes.

TABELA 17
Distribuição de *você* e *cê*, conforme o Gênero e a Idade - 2002

Gênero e Faixa Etária	VOCÊ		OCÊ		CÊ		TOTAL	
	N	%	N	%	N	%	N	%
Fem. V (+ 47)	17	12,6	09	6,6	109	80,8	135	9,3
Masc. V (+ 47)	102	55,4	05	2,7	77	41,9	184	12,6
Fem. IV (31-47)	63	24,1	09	3,5	189	72,4	261	18,0
Masc. IV (31-47)	32	20,1	02	1,3	125	78,6	159	11,0
Fem. III (16-30)	14	7,6	13	7,1	157	85,3	184	12,6
Masc. III (16-30)	58	22,1	10	3,8	194	74,1	262	18,0
Fem. II (12-15)	10	17,2	0	-	48	82,8	58	4,0
Masc. II (12-15)	11	36,7	0	-	19	63,3	30	2,1
Fem. I (08-11)	19	24,7	02	2,6	56	72,7	77	5,3
Masc. I (08-11)	16	15,5	06	5,8	81	78,7	103	7,1
Total	342	23,5	56	3,9	1055	72,6	1453	100

Pelos resultados da tabela acima, percebe-se que há uma tendência de os homens da geração seguinte acompanharem o padrão feminino da geração anterior. Com exceção dos homens da faixa etária II em relação às mulheres da faixa etária III, observa-se a tendência de os informantes masculinos seguirem os informantes femininos. Com relação ao

comportamento destoante dos adolescentes homens, ele já foi discutido no subitem anterior e voltará a sê-lo adiante.

Também pela tabela 17 vê-se que a maior diferença entre homens e mulheres encontra-se na faixa etária mais avançada: 80,8% das mulheres acima de 47 anos preferem *cê*, contra 41,9% dos homens dessa faixa etária, ou seja, uma diferença de 38,9 %, a maior entre todas. Por outro lado, nota-se que, nas faixas etárias I e IV, não foram as mulheres que privilegiaram mais a forma *cê* que os homens, havendo uma semelhança bastante grande entre os percentuais: 72,4% para mulheres da faixa etária IV e 72,7% para as da faixa I; e 78,6% para os homens da faixa IV e 78,7% para os da faixa I. Entretanto, essas diferenças são bastante reduzidas, em torno de 6%, o que indica uma variação normal entre esses falantes, não justificando uma idéia de rejeição à forma *cê* por parte desses informantes femininos.

Com respeito à observação de que, perto do final da mudança, as diferenças entre homens e mulheres diminuem ou desaparecem, os resultados não a confirmam. Assim, pode-se pensar que a mudança de *você* para *cê*, embora esteja em processo adiantado, ainda não está no seu final.

Por último, os estudos variacionistas afirmam que as diferenças entre homens e mulheres são mais salientes na classe média baixa e no estilo cuidado. Porém, o fator Classe Social não foi considerado significativo pelo Programa Goldvarb, o que indica que essa observação não se aplica ao caso da forma *cê*. Quanto ao estilo de fala, somente foi possível avaliá-lo para os informantes das faixas I e II, conforme foi visto. Para as demais faixas, não se poderá analisar este fato, tendo em vista que seus dados foram coletados para uma outra pesquisa, bastante diferente desta.

Fazendo-se um resumo dos resultados encontrados para este fator, pôde-se notar que a forma *cê* não sofre preconceito quanto ao seu uso na linguagem informal em Belo Horizonte, haja vista que as mulheres a favorecem. Entretanto, não se pode dizer que as mulheres de todas as faixas etárias estejam liderando o processo de mudança de *você* para

cê, pois as pertencentes às faixas I e IV estão um pouco atrás dos homens. Na busca para uma explicação para esse fato, é proveitoso lembrar o que disse Trudgil (1988, p. 45) a respeito dos processos de mudança lingüística: as mudanças não ocorrem de forma simples, nem direta, nem uniforme, não sendo diferente com as formas *você* e *cê*.

4.7 Os testes dos informantes crianças e adolescentes

Neste subitem, apresentam-se e discutem-se os resultados da primeira parte das entrevistas realizadas com as crianças e adolescentes, chamadas de “teste”.

Sabe-se que as crianças e os adolescentes, quando em situação natural de interação, são geralmente expansivos e falantes, principalmente se estão em companhia de seus colegas e amigos, os quais exercem uma forte influência em seu comportamento e mesmo em sua linguagem (cf. capítulo 1 deste trabalho). Entretanto, devido à sensibilidade e à insegurança próprias de sua idade, uma situação de entrevista com um adulto poderia deixá-los tensos e tímidos, alterando sua linguagem e, conseqüentemente, prejudicando a coleta dos dados.

Diante disso, conforme já se disse no capítulo 3, foi pensada uma estratégia para que os informantes das faixas I e II se sentissem mais à vontade diante de um gravador e de uma pessoa estranha fazendo-lhes perguntas sobre sua vida. Assim, foi criada uma brincadeira em que dois colegas de sexos diferentes, da mesma sala ou pelo menos da mesma idade entrevistassem um ao outro, orientados pela entrevistadora. Depois disso, seria feita a pesquisa sociolingüística propriamente dita.

A hipótese inicial era de que, durante a primeira parte das entrevistas – os testes –, com a companhia de uma outra criança/adolescente conhecida(o) – alguns até amigos e parentes –, esses informantes se acostuariam mais a essa estranha situação e que fosse possível conseguir, no final, algo próximo ao seu vernáculo.

Pelo fato de que esses testes são constituídos de perguntas feitas diretamente um ao outro, somente ocorreu aí a referência definida. Uma outra característica sua é que, devido ao fato de terem sido aplicados logo no início da entrevista, quando os informantes estavam mais tensos, seu estilo é mais formal que a segunda parte. Assim, neste subitem, busca-se responder à seguinte pergunta: O que muda no comportamento lingüístico de crianças e adolescentes com respeito às formas *você*, *ocê* e *cê* em situações mais formais de fala?

Para se proceder à análise do uso dessas três formas, vejam-se os resultados de crianças e adolescentes na primeira parte das entrevistas (os testes) e na segunda. Tendo em vista que uma informante adolescente foi entrevistada pela Prof^a Ana Paula Huback, ela não participou do teste e, por isso, seus dados foram retirados da segunda parte das entrevistas.¹²⁸

Dos testes, o Programa Goldvarb selecionou como significativos apenas os seguintes fatores: gênero, classe social e função sintática. Das entrevistas, apenas classe social e função sintática. Os resultados constam das tabelas abaixo.

¹²⁸ Na entrevista dessa informante adolescente ocorreram nove dados, todos da forma *cê*.

TABELA 18

Distribuição de *você*, *ocê* e *cê*, para crianças e adolescentes – Testes¹²⁹

		VOCÊ			OCÊ			CÊ			TOTAL	
		N	%	PR	N	%	PR ¹³⁰	N	%	PR	N	%
Função Sintática	Sujeito	320	43,8	.47	13	1,8	.49	398	54,4	.53	731	96,6
	Obj.	22	84,6	.96	03	11,5	.98	01	3,8	.04	26	3,4
	preposição											
Gênero	Feminino	146	39,8	.43	02	0,5	-	219	59,7	.57	367	48,5
	Masculino	196	50,3	.56	14	3,6	-	180	46,1	.44	390	51,5
Classe Social	Média	205	59,9	.66	09	2,6	-	128	37,4	.34	342	45,2
	Baixa	137	33,0	.37	07	1,7	-	271	65,3	.63	415	54,8
	Total	342	45,2		16	2,1		399	52,7		757	

TABELA 19

Distribuição de *você*, *ocê* e *cê* para crianças e adolescentes – Entrevistas

		VOCÊ			OCÊ			CÊ			TOTAL	
		N	%	PR	N	%	PR ¹³¹	N	%	PR	N	%
Função Sintática	Sujeito	50	20,0	.48	06	2,4	-	194	77,6	.52	250	96,5
	Obj.verbo	02	66,7	.97	01	33,3	-	0	0	.09	03	1,2
	Obj.	05	83,3	.97	01	16,7	-	0	0	.09	06	2,6
Classe Social	preposição											
	Média	42	33,3	.66	03	2,4	-	81	64,3	.34	126	48,6
	Baixa	15	11,3	.34	05	3,7	-	113	85,0	.66	133	51,4
	Total	57	22,0		08	3,1		194	74,9		249	

¹²⁹ Nos testes, não houve nenhuma ocorrência de *você*, *ocê* e *cê* como objeto de verbo.

¹³⁰ Na análise binomial de *você* e *ocê*, foram encontrados os seguintes PR para a forma *ocê* quanto à Classe Social: classe média: .24; classe baixa: .85. Vê-se, assim, que a classe média, em contextos mais formais, prefere usar a forma padrão, como é de se esperar. Ainda nesta análise, além do fator Classe Social, o Programa selecionou os seguintes: Idade (crianças: PR de *ocê* = .74; adolescentes: PR = .20) e Foco (foco: PR = .98; não-foco: PR = .50)

¹³¹ Na análise binomial de *cê* e *ocê*, nas entrevistas, apenas o fator Tópico foi selecionado pelo Programa como significativo (PR de *ocê* como tópico = .98; não-tópico = .49). Já na análise de *você* e *ocê*, o fator Função Sintática foi selecionado, apresentando os seguintes pesos relativos para *ocê*: função de sujeito: .28; objeto de verbo: .98; objeto de preposição: .97. Comprova-se, assim, também nas entrevistas, que *ocê* é favorecida nas funções de objeto de verbo e de preposição.

Pelas tabelas apresentadas, um primeiro fato que chama a atenção é a uniformidade dos dados: foram poucos os fatores selecionados como significativos, sendo que dois foram escolhidos tanto para a primeira parte da entrevista quanto para a segunda: Classe Social e Função Sintática.

Quanto ao fator Função Sintática, nota-se que a formalidade da situação praticamente não altera os resultados: na função de sujeito não há favorecimento de nenhuma das formas; na função de objeto de preposição, *você* e *ocê* são as mais favorecidas (PR = .97 para as duas), sendo que *cê* não foi usada; quanto à função de objeto de verbo, ela não teve nenhuma ocorrência nos testes e apenas três nas entrevistas. Esses dados confirmam os resultados dos informantes mais velhos de que a função sintática não-marcada das três formas é a de sujeito. Nessa função, de novo *você* e *ocê* são altamente favorecidas, exatamente ao contrário de *cê*, que teve apenas uma ocorrência nos testes e nenhuma nas entrevistas. De novo, o desfavorecimento de *cê* nestas duas funções confirma os resultados gerais já vistos.

Com relação ao fator Classe Social, vê-se que, para a classe média, as posições não se alteram muito entre as entrevistas e os testes. Em ambas as situações, a classe média favorece a forma *cê*, exatamente com os mesmos pesos relativos. Já a classe baixa favorece a forma *cê* nas duas situações, mas um pouco mais nas entrevistas, ou seja, esses informantes se ressentiram um pouco da situação de teste e, durante sua aplicação, usaram mais a forma padrão, *você*.

Por outro lado, esses mesmos resultados mostram que as crianças e os adolescentes da classe média estão mais atentos a uma situação “anormal”, ou seja, diante de uma pessoa gravando sua fala, mudam seu comportamento lingüístico, optando pela forma padrão *você*. Em outras palavras, esse resultado mostra que esses informantes têm mais consciência do que seja a adequação da fala ao contexto: diante de situações formais, utilizam mais a forma padrão – *você*.

Com respeito ao fator Gênero, há diferenças significativas entre a linguagem feminina e masculina somente na situação de teste. Nessa situação mais formal, os informantes do sexo masculino alteram um pouco seu comportamento lingüístico, favorecendo levemente a forma padrão *você*, ao contrário do sexo feminino.

Esses resultados confirmam que *cê* não é uma forma estigmatizada na sociedade, haja vista que as informantes têm preferência por essa forma, conforme atestam inúmeros estudos sociolingüísticos. Com relação às crianças e adolescentes do sexo masculino, ou estão mais atentos a um comportamento lingüístico supostamente adequado a uma situação formal de fala, ou alteraram sua fala por insegurança, talvez, diante de uma situação “atípica”.

Por fim, pode-se fazer uma pequena comparação entre o uso das três formas por crianças e adolescentes, nas entrevistas, e pelas demais faixas etárias¹³², para se verificar o comportamento das três formas quanto aos fatores analisados, pois, como já foi mencionado antes, Labov (1994, p. 47) afirma que são os adolescentes e pré-adolescentes que lideram os processos de mudança lingüística. Para isso, observe-se a tabela 20, em que constam os fatores significativos somente para as faixas etárias III, IV e V.

¹³² Primeiramente, no *corpus* de 2002, os dados das cinco faixas etárias foram tomados em conjunto. O que se pretende, ao separar os dados de crianças e adolescentes dos dados das demais faixas etárias, é verificar se existe alteração quanto ao comportamento das três formas, de acordo com os fatores selecionados como significativos pelo Programa Goldvarb. Para tanto, este foi rodado mais uma vez, agora somente com as três maiores faixas etárias. Os resultados encontram-se na Tabela 20.

TABELA 20

Distribuição de *você*, *ocê* e *cê* – Corpus de 2002 – Faixas etárias III, IV e V

	Fatores	VOCÊ			OCÊ			CÊ			TOTAL	
		N	%	PR	N	%	PR	N	%	PR	N	%
Função Sintática	Sujeito	263	23,1	.48	29	2,6	.47	845	74,3	.52	1137	95,9
	Obj.verbo	02	22,2	.96	04	44,5	.98	03	33,3	.04	09	0,8
	Obj.prep.	21	53,8	.98	15	38,5	.98	03	7,7	.02	39	3,3
Referência	Definida	70	14,2	.35	28	5,6	-	396	80,2	.65	494	41,7
	Indefinida	216	31,3	.61	20	2,9	-	455	65,8	.39	691	58,7
Focalização	Foco	06	66,7	.91	02	22,2	-	01	11,1	.09	09	0,8
	Não-foco	280	23,8	.50	46	3,9	-	850	72,3	.50	1176	99,2
Faixa Etária	+ 47	119	37,3	.67	14	4,4	.66	186	58,3	.33	319	26,9
	31 – 47	95	22,6	.53	11	2,6	.33	314	74,8	.47	420	35,4
	16 - 30	72	16,1	.35	23	5,2	.56	351	78,7	.65	446	37,6
Gênero	Feminino	94	16,2	.41	31	5,3	.72	455	78,4	.59	580	48,9
	Masculino	192	31,7	.58	17	2,8	.37	396	65,5	.42	605	51,1
Total		286	24,1		48	4,1		851	71,8		1185	

Confrontando-se os resultados da tabela 19 com os da tabela 20, vê-se que os dados das faixas etárias III, IV e V apresentam mais fatores significativos que os das faixas I e II, ou seja, os dados das maiores faixas etárias são bem menos uniformes.

O único fator significativo para os dois grupos etários foi a função sintática, e vê-se que, na função de sujeito, a probabilidade de ocorrência de *você*, *ocê* e *cê*, nos dois grupos etários é exatamente a mesma: .48, e .52, respectivamente. Quanto às funções de objeto de verbo e de preposição, vê-se um pequeníssimo aumento do peso relativo da forma *cê*. Na função de objeto de verbo, a probabilidade de ocorrência passou de .04, nas faixas III, IV e V, para .09, nas faixas I e II, o que pode indicar que as crianças e adolescentes estão levando à frente a expansão de *cê* para outros contextos.

Na função de objeto de preposição também houve um aumento muito reduzido dessa probabilidade: passou de .02 para .09, sendo que o peso relativo de *cê* nessas duas funções, nas faixas etárias menores, ainda é extremamente baixo. Porém, como o início e o fim das mudanças lingüísticas são normalmente lentos, como foi visto anteriormente, é possível que esse pequeno aumento dos pesos relativos de *cê* nas funções de objeto de verbo e de preposição indiquem um início de mudança, em que *cê* lentamente está ocupando os espaços de *você*.

Com relação aos fatores Gênero, Focalização e Referência, o fato de serem significativos para as faixas III, IV e V e não serem para as faixas I e II pode indicar uma mudança, com uma uniformização maior no uso das três formas quanto a esses aspectos.

Por outro lado, conforme se viu ao longo da apresentação e da discussão dos dados, parece não haver dúvidas de que a situação em que se realizaram as entrevistas e os testes influenciou os resultados das crianças e adolescentes. Além da tensão que a presença de um gravador durante a conversa pode provocar, um possível agravante dessa situação seria o fato de que as entrevistas foram conduzidas por esta pesquisadora, sendo que alguns desses informantes estudavam no Centro Pedagógico/UFMG e sabiam que ela era professora de português. Entretanto, essa hipótese pode ser descartada, por quatro motivos:

- (i) todos os informantes que a conheciam usaram, uns mais, outros um pouco menos, a forma *cê*;
- (ii) ela própria, durante as entrevistas e testes, usou muito mais a forma *cê* que *você*;
- (iii) os outros informantes demonstraram comportamento similar ao dos alunos do Centro Pedagógico, com relação ao uso das três formas;
- (iv) os alunos do Centro Pedagógico correspondem a apenas 20,8% do total de informantes, número este insuficiente para exercer uma influência significativa nos resultados.

Assim, pode-se pensar numa outra hipótese. Como já se disse, segundo Labov (1972, p.117; 126-7), a hipercorreção está diretamente ligada a uma insegurança lingüística por parte do falante. Tendo em vista que a infância e, principalmente, a adolescência são fases de turbulências, de inseguranças, e, conseqüentemente, de busca de auto-afirmação por excelência (Chambers, 1995, p. 169-70), pois sua personalidade ainda está em formação (Piaget, 2003, p. 62), não é de se estranhar que a insegurança pessoal passe para o campo da linguagem. Portanto, tendo em vista que os informantes têm consciência da existência da variação lingüística conforme o contexto, numa situação inusitada, artificial – uma pessoa gravando sua fala –, optaram por usar, mais do que normalmente o fazem, a forma padrão, como uma forma de valorizar sua fala e, por conseguinte, a eles próprios.

Uma forma de contornar essa situação seria a de gravar a fala dos informantes sem que eles se dessem conta disso. Entretanto, haveria um problema de ordem prática a ser resolvido: onde esconder um gravador num pátio de escola, ou sala de aula (e obter “bons” dados), ou – o que seria pior - na casa deles? Assim sendo, esses mesmos dados foram analisados, contando-se com a ajuda dos dados dos outros informantes, para se fazer uma análise mais confiável dos resultados.

4.8 – Conclusão

Neste capítulo, foram apresentados e discutidos os resultados do *corpus* de 2002, composto por 1453 dados obtidos de 47 informantes dos dois gêneros, divididos em cinco faixas etárias e duas classes sociais. Quatro fatores lingüísticos – função sintática, contigüidade em relação ao verbo, referência e focalização contrastiva – e dois fatores extralingüísticos – faixa etária e gênero –, considerados significativos pelo Programa Goldvarb 2001, foram analisados, com vistas a ampliar a discussão já iniciada em trabalhos anteriores sobre as três formas, comprovando alguns desses resultados e desconfirmando outros.

Por último, foram discutidos os resultados dos testes das crianças e adolescentes, comparando-os com os obtidos nas entrevistas. Além disso, analisou-se seu comportamento lingüístico com base nas suas características emocionais e afetivas, observando-se que o mesmo é afetado pela formalidade do contexto lingüístico. Em seguida, esses resultados foram contrapostos aos das demais faixas etárias, para se poder observar com maior clareza o comportamento lingüístico dos informantes mais jovens quanto às três formas.

Tendo sido feito o estudo sincrônico dos dados, no próximo capítulo serão apresentados os resultados do *corpus* de 1982, para, no seguinte, se fazer a análise diacrônica do comportamento de *você*, *ocê* e *cê*.

CAPÍTULO 5 - A ANÁLISE DOS DADOS – *CORPUS* DE 1982

Neste capítulo, será feito o estudo das formas *você*, *ocê* e *cê* na comunidade de Belo Horizonte, através do *corpus* de 1982.

O resultado geral desse *corpus* se encontra na tabela abaixo.

TABELA 21

Total de ocorrências de *você*, *ocê* e *cê* - *corpus* de 1982

Formas	Ocorrencias	%
Você	93	15,9
ocê	39	6,6
Cê	455	77,5
Total	587	

Do resultado geral de 1982, o Programa Goldvarb 2001 selecionou apenas três fatores como sendo significativos, um lingüístico – Função Sintática – e dois extralingüísticos – Idade e Classe Social. Esse resultado indica a importância dos fatores sociais para o fenômeno em estudo.

Já os fatores que não foram selecionados como significativos para o *corpus* de 1982 – Tipo de frase em que aparecem as formas, Contigüidade em relação ao sujeito da frase, Focalização, Referência, Tópico e Gênero – têm seus resultados expressos na tabela 21, abaixo.

TABELA 22
Distribuição de *você*, *ocê* e *cê* - Fatores não-significativos – 1982

FATORES		VOCÊ		OCÊ		CÊ		TOTAL
		Nº	%	Nº	%	Nº	%	
Tipo de frase	Afirmativa	83	16,1	38	7,4	395	76,5	516
	Interrogativa	48	81,4	1	1,7	10	16,9	59
	“Que que”	0	0	0	0	12	100	12
Contigüidade	Contíguo	60	13,0	24	5,2	378	81,8	462
	Não-contíguo	15	15,5	08	8,2	74	76,3	97
Tópico	Tópico	01	33,3	01	33,3	01	33,3	03
	Não-tópico	92	15,8	38	6,5	454	77,7	584
Foco	Foco	03	60	02	40	0	0	05
	Não-foco	90	15,5	37	6,3	455	78,2	582
Refer.	Definida	32	16,0	11	5,5	157	78,5	200
	Indefinida	61	15,8	28	7,2	298	77,0	387
Gênero	Feminino	37	17,2	15	6,9	164	75,9	216
	Masculino	56	15,1	24	6,5	291	78,4	371

A análise dos resultados considerados significativos será iniciada pelo fator Função Sintática.

5.1 As ocorrências de *você*, *ocê* e *cê*, conforme o fator Função Sintática

Neste subitem, procura-se descrever o comportamento sintático das formas *você*, *ocê* e *cê* em 1982. No próximo capítulo, estes resultados são confrontados com os do *corpus* de 2002. Os dados de 1982 podem ser vistos na tabela abaixo:

TABELA 23

Distribuição de *você*, *ocê* e *cê*, conforme a Função Sintática – 1982

FUNÇÃO SINTÁTICA	VOCÊ			OCÊ			CÊ			TOTAL	
	N	%	PR	N	%	PR ¹³³	N	%	PR	N	%
Sujeito	75	13,4	.46	30	5,4	.48	454	81,2	.54	559	95,2
Obj. Verbo	02	66,7	.97	01	33,3	.90	0	0	.04	03	0,5
Obj. Prep.	16	64	.99	08	32	.99	01	04	.01	25	4,3
Total	93	15,9		39	6,6		455	77,5		587	

Os mesmos resultados encontram-se dispostos em forma de gráfico, a seguir:

¹³³ Na análise binomial de *você* e *ocê*, nenhum fator foi selecionado pelo Programa Goldvarb.

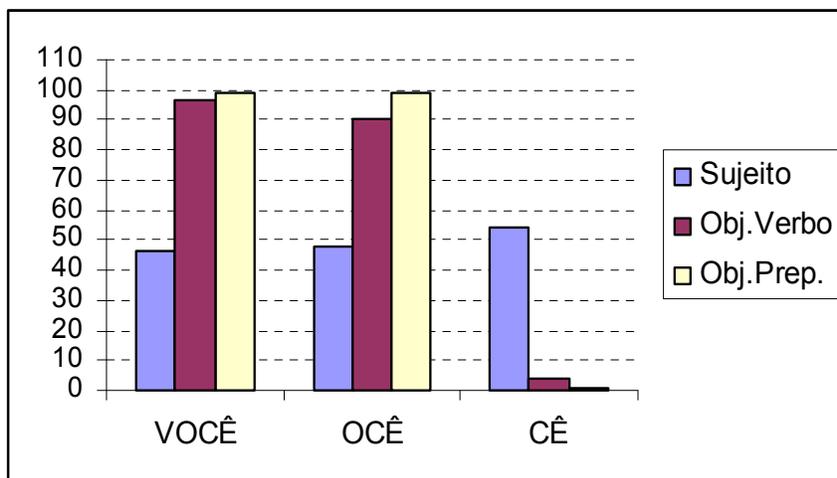


GRÁFICO 8 - Distribuição de *você*, *ocê* e *cê*, conforme a Função Sintática - 1982

A tabela 23 de novo comprova que sujeito é a função não-marcada para as três formas, com 95,2% do total de ocorrências. Em segundo e terceiro lugares, mas muitíssimo atrás, aparecem, respectivamente, objeto de preposição, com 4,3%, e objeto de verbo, com apenas 0,5%. No Anexo 2 desta tese, encontram-se as frases do *corpus* de 1982 em que *você*, *ocê* e *cê* ocorreram como objeto de verbo e de preposição.

Pelos resultados acima, vê-se que, na função de sujeito, *cê* é a franca favorita, com 81,2%, seguida muito atrás pelas outras duas: *você*, com 13,4%, e *ocê*, com apenas 5,4%. Entretanto, a probabilidade apresenta índices interessantes: embora não haja exatamente um favorecimento de nenhuma das três formas, na função de sujeito, a forma *cê* é a de ocorrência um pouco mais provável (PR = .54), vindo *ocê* em segundo lugar, na frente de *você* (PR = .48 e .46, respectivamente). Isso não deixa de significar uma curiosa inversão, na função de sujeito.

Na função de objeto de verbo, *ocê* é amplamente favorecida, mas com probabilidade de ocorrência um pouco menor que *você* (PR = .90, para *ocê*, e .97, para *você*). Mas o que mais chama a atenção aí é, além do reduzidíssimo número de ocorrências – apenas três –, a ausência de *cê* nessa função. Ao que parece, o espriamento de *cê* para esse contexto, a

essa época, ainda não havia acontecido ou estava muito no início, a ponto de não aparecer nos dados.

Os casos de objeto de verbo neste *corpus* são:

- (E 18) “Tia, gostei muito de ter *você* de novo.”
 (E 117) ... *cê cê* confunde muit’*ocê* com a outra pessoa e...
 (E 548) Eu tô querendo *você* mesmo.

Com respeito à função de objeto de preposição, observa-se agora a forma *você* na dianteira, com 64% das ocorrências, seguida da forma *ocê*, com 32%, e *cê*, com apenas uma ocorrência, representando 4% dos dados. Quanto à probabilidade, as formas *você* e *ocê* são imensamente favorecidas, com idênticos pesos relativos: .99. Já a probabilidade de ocorrência de *cê* é ínfima: .01. Estes resultados indicam que, em 1982, se *cê* começava a espalhar-se para esse contexto, ainda estava num processo muito inicial, de tal forma que não os dados não registraram esse uso.

Alguns exemplos das três formas exercendo a função de objeto de preposição são:

- (E 66) Lógico que se ele gosta de *você* ele nunca vai abusar.
 (E 456) Se eu fiz eu chego e falo com *você*.
 (E 269) Aí eu posso tentar arrumar isso p’*ocê*.
 (E 285) É mais suficiente pr’*ocê*.
 (E 161) Eu vou lá olhar p’*cê*.

Em resumo, os dados referentes à função sintática confirmam que as formas *você*, *ocê* e *cê* não desempenham normalmente as funções de objeto de verbo e de preposição,¹³⁴ haja vista sua pequeníssima ocorrência aí: 0,5% e 4,3%, respectivamente. Com relação ao comportamento individual das formas, vê-se que *você* e *ocê* podem exercer as três funções, mas são especialmente usadas como objeto de verbo e de preposição, que são as marcadas. Já a forma *cê* está praticamente limitada à função de sujeito, mas começava ocorrer na função de objeto de preposição.

Esses dados confirmam a hipótese de que *cê* apresenta um comportamento sintático peculiar, diferenciado das outras formas. Por outro lado, é interessante pensar no porquê disso, o que será feito no próximo capítulo.

5.2 As ocorrências de *você*, *ocê* e *cê*, conforme o fator Idade

Assim como aconteceu com os dados de 2002, procura-se saber se os dados de 1982 indicavam mudança em progresso com respeito às formas *você*, *ocê* e *cê*. Os resultados são os constantes da tabela abaixo.

¹³⁴ A esse respeito, veja-se o que foi mencionado quanto à concorrência dos pronomes e da forma zero às três formas, quando da análise desse mesmo fator, em 2002.

TABELA 24

Distribuição de *você*, *ocê* e *cê*, conforme a Idade – 1982

IDADE	VOCÊ			OCÊ			CÊ			TOTAL	
	N	%	PR	N	%	PR ¹³⁵	N	%	PR	N	%
Faixa Etária III (+ 47 anos)	32	21,6	.62	06	4,1	-	110	74,3	.38	148	25,2
Faixa Etária II (31 – 47 anos)	16	21,6	.56	05	6,8	-	53	71,6	.44	74	12,6
Faixa Etária I (16 – 30 anos)	45	12,3	.43	28	7,7	-	292	80,0	.57	365	62,2
Total	93	15,9		39	6,6		455	77,5		587	100

Observando-se os resultados em forma de gráfico, tem-se:

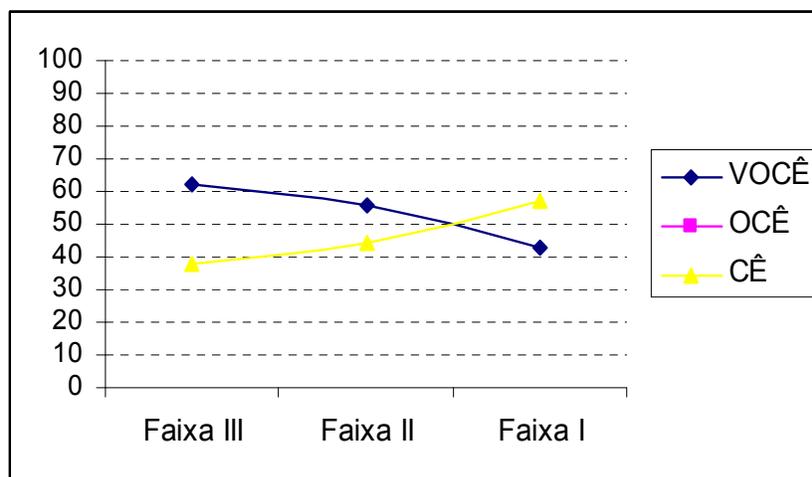


GRÁFICO 9 – Distribuição de *você*, *ocê* e *cê*, conforme a Idade – 1982

¹³⁵ Em nenhuma análise binomial com *ocê* o fator Idade foi selecionado pelo Programa.

Observando-se a tabela 24, acima, nota-se que os percentuais de ocorrências da forma *cê* não indicam mudança em curso, embora exista uma diferença muito pequena entre as faixas II (71,6%) e III (74,3%), ou seja, uma diferença de apenas 2,7%. Entretanto, o aumento gradativo da probabilidade de uso dessa forma não deixa dúvidas de que existe mudança em progresso. Mas veja-se que o maior peso relativo – exatamente da faixa mais jovem – é de apenas .57, ou seja, *cê* não é fortemente favorecida pela faixa etária I. A faixa II a desfavorece levemente, e a III a desfavorece mais acentuadamente.

Com relação a *você*, de novo os percentuais não indicam exatamente mudança – neste caso, diminuição de seu uso, já que, nas faixas II e III, há um curioso empate de 21,6%. Mas os pesos relativos encontrados mostram inequivocamente que *você*, a essa época, perdia espaço para a forma *cê* a cada geração: a faixa etária III a favorecia, ao passo que a faixa II a favorecia levemente e a faixa I a desfavorecia um pouco.

Vê-se, com esses resultados, que, em 1982, embora a forma *cê* já estivesse implementada em Belo Horizonte, a disputa entre as formas *você* e *cê* estava mais ou menos equilibrada, sem a vantagem que vemos hoje para *cê*. Provavelmente esta é a causa de tão poucos fatores serem considerados significativos pelo Programa Goldvarb, neste *corpus*.

Com respeito à forma *ocê*, apesar de não ter sido selecionada como significativa pelo Programa em nenhuma das análises binárias, seus percentuais mostram um aumento de uso com relação a ela, com seu crescimento gradativo de 4,1% (faixa III) para 6,8% (faixa II), até chegar a 7,7% (faixa I).

Pode-se investigar agora o comportamento dos informantes de cada faixa etária, a fim de se verificar a validade das afirmações a respeito das pressões sociais sobre o uso das variantes-padrão da língua sobre os jovens adultos.

Pelos resultados acima, vê-se que a situação apresentada no *corpus* de 2002 já se verificava em 1982. As faixas etárias I e II não se mostraram propensas ao uso de *você*, com percentuais baixos e probabilidades relativamente neutras quanto a esse uso. Assim,

confirma-se o que foi dito para o *corpus* de 2002: os informantes jovens e adultos não sentem propensos a rejeitar a forma *cê*, em situação de fala, mesmo sendo esta “atípica”, como é o caso de uma entrevista.

5.3 As ocorrências de *você*, *ocê* e *cê*, conforme o fator Classe Social

O terceiro e último fator selecionado pelo Programa Goldvarb, para o *corpus* de 1982, foi Classe Social. A tabela abaixo traz os resultados obtidos neste *corpus*.

TABELA 25
Distribuição de *você*, *ocê* e *cê*, conforme a Classe Social – 1982

CLASSE SOCIAL	VOCÊ			OCÊ			CÊ			TOTAL	
	N	%	PR	N	%	PR ¹³⁶	N	%	PR	N	%
Classe Média	45	12,7	.42	26	7,3	-	282	80	.58	353	60,1
Classe Baixa	48	20,5	.62	13	5,6	-	173	73,9	.38	234	39,9
Total	93	15,9		39	6,6		455	77,5		587	100

Esses mesmos resultados podem ser visualizados no gráfico abaixo:

¹³⁶ O fator Classe Social não foi selecionado em nenhuma das análises binárias com a forma *ocê*.

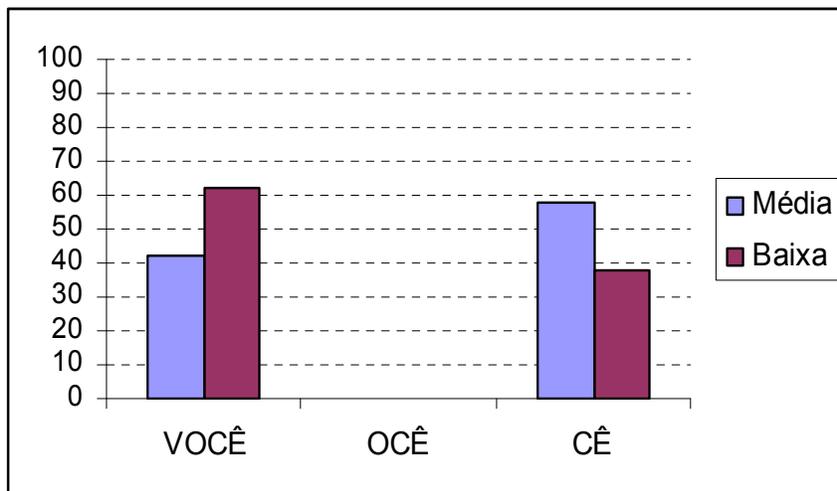


GRÁFICO 10 – Distribuição de *você*, *ocê* e *cê*, conforme a Classe Social – 1982

Os resultados acima mostram que, embora a forma *cê* seja a mais utilizada pelas duas classes sociais, a probabilidade de ocorrência dessa forma não é a mesma. *Cê* é levemente favorecida pela classe média (PR = .58) e é desfavorecida pela classe baixa (PR = .38).

O contrário se dá com relação à forma *você*. Vê-se que ela teve pouca ocorrência nas duas classes sociais, mas, principalmente, na média (12,7%). Assim, os dados mostram que a forma *você* é levemente desfavorecida pela classe média (PR = .42) e é favorecida pela classe baixa (PR = .62).

A forma *ocê* não foi considerada significativa pelo Programa, em nenhuma das análises binárias, mas vê-se que sua ocorrência foi maior entre a classe média do que entre a baixa (7,3% e 5,6%, respectivamente), o que não deixa de ser uma surpresa, tendo em vista que a forma *ocê* é tida como uma marca de pessoas do interior (cf. Ramos, 2000; Coelho, 1999).

Os resultados acima são curiosos. Como se disse anteriormente, segundo Labov (2001), as mudanças encabeçadas pelas classes mais altas são aquelas que envolvem empréstimo de formas prestigiadas e nas quais os falantes têm consciência desse processo. São chamadas *from above*. (LABOV, 2001, p. 272-5).

Já as classes mais baixas lideram as mudanças *from below*, em que os indivíduos não têm consciência desses processos. Elas englobam as mudanças sonoras sistemáticas ou as mudanças lingüísticas propriamente ditas. E parece claro que é este o caso da mudança de *você* para *cê*. Como já foi dito, essa mudança é um processo natural nas línguas, de redução de itens lexicais. Além disso, como também já foi apontado, a forma *cê*, embora ocorra normalmente na língua falada, não é considerada padrão pelas Gramáticas Normativas da Língua Portuguesa – que, aliás, nem a mencionam. Aliado a isso está o fato de que essa forma foi atestada primeiramente nas classes baixas (cf. capítulo II deste trabalho). Assim, esperar-se-ia que a classe baixa favorecesse a forma *cê*, e não *você*.

Então, a que se devem os resultados encontrados no *corpus* de 1982? Se não existem causas lingüísticas nem extralingüísticas que justifiquem esse uso, resta postular a hipótese de que os informantes da classe baixa, sentindo-se numa situação mais formal, não-natural, como é a situação de entrevista, tentaram adequar sua fala ao contexto em questão. É importante lembrar que, ao que parece, esse fato influenciou os resultados de crianças e adolescentes. Com isso, vê-se que os informantes, mesmo os pequenos, sabem distinguir qual a forma “mais adequada” aos diversos contextos de uso da língua.

Assim sendo, o que foi dito indica que o favorecimento da forma *você* pela classe baixa decorre da sensibilidade que esta apresentou quando se viu diante de um gravador e de uma entrevistadora desconhecida, mesmo sendo muito simpática com os informantes. Portanto, atuou aí o Paradoxo do Observador (LABOV, 1972), assim como atuou com as crianças e os adolescentes do *corpus* de 2002.¹³⁷

Em resumo, pode-se concluir que, assim como a situação de entrevista provavelmente acentuou a natural insegurança da adolescência, a mesma situação levou insegurança à classe baixa do *corpus* de 1982, fazendo-a formalizar sua linguagem.

¹³⁷ Entretanto, observe-se a discussão deste mesmo fator com relação aos *corpora* de 1982 e 2002, feita no próximo capítulo.

Tendo sido apresentados e discutidos os resultados dos dois *corpora* desta Tese, pode-se então proceder à análise da importância dos aspectos fonéticos e fonológicos do português para a ocorrência das formas *você*, *ocê* e *cê* nessas duas amostras. Como foi visto no capítulo I, uma explicação para a distribuição irregular das três formas quanto à função sintática que assumem na frase poderia ser encontrada no ambiente fonético em que elas se encontram. Assim, passa-se, a seguir, à discussão deste assunto, testando a hipótese 2.1, estipulada no item 1.4 deste trabalho.

5.4 A análise das formas e os processos fonético-fonológicos do português

Vários estudos já apontaram a irregularidade da distribuição das formas *você*, *ocê* e *cê* quanto à função sintática que exercem nas frases. Por sua vez, os dados desta pesquisa confirmaram o resultado acima, mas indicaram um possível início de espraiamento da forma *cê* para outros contextos, já que seu uso, embora muito timidamente, é verdade, está aumentando nas funções de objeto de verbo e de preposição, de 1982 até hoje.

Já com relação à forma *você*, apesar de sua ocorrência não ser propriamente favorecida na função de sujeito, como se dá nas outras duas funções, essa forma aparece em todos os contextos sintáticos, e é natural que isso aconteça, tendo em vista que se configura como a forma padrão da língua, a tradicional. Mais interessante, entretanto, é o comportamento sintático de *ocê*. Como se observou neste e no capítulo anterior, essa forma é muito levemente desfavorecida quando na função de sujeito, mas sua probabilidade de ocorrência é altíssima nas funções de objeto de verbo e de preposição.

Esse comportamento sintático geral das três formas, como se disse, já foi atestado em vários trabalhos que se ocuparam do estudo deste tema, mas as conclusões a que chegaram não são sempre convergentes. Diante disso, na Introdução desta tese, estipulou-se a hipótese de que o contexto fonético onde ocorrem as formas também estaria influenciando esses resultados. Assim, no capítulo 1, foram resumidos alguns aspectos importantes acerca da fonética e da fonologia do português, os quais embasariam as análises posteriores, realizadas neste subitem. Também no capítulo 1 detalhou-se a hipótese acima, levando-se em conta que os falantes, ao escolherem uma das três formas, buscam padrões característicos – não-marcados – da língua. Assim, a hipótese acima foi especificada e é rerepresentada a seguir.

Hipótese 2.1: O uso das formas *você*, *ocê* e *cê* busca adaptá-las aos padrões fonéticos e fonológicos da língua portuguesa, no tocante à estrutura silábica, acentuação e ritmo.

Portanto, o objetivo deste subitem é testar essa hipótese, analisando-se o comportamento de *você*, *ocê* e *cê* nas funções de sujeito, objeto de verbo e de preposição através dos *corpora* de 1982 e 2002. Para maior facilidade do leitor, abaixo reaparecem as tabelas 23 e 9, que trazem esses mesmos resultados.

TABELA 23

Distribuição de *você*, *ocê* e *cê*, conforme a Função Sintática – 1982

FUNÇÃO SINTÁTICA	VOCÊ			OCÊ			CÊ			TOTAL	
	N	%	PR	N	%	PR	N	%	PR	N	%
Sujeito	75	13,4	.46	30	5,4	.48	454	81,2	.54	559	95,2
Obj. Verbo	02	66,7	.97	01	33,3	.90	0	0	.04	03	0,5
Obj. Preposição	16	64	.99	08	32	.99	01	04	.01	25	4,3
Total	93	15,9		39	6,6		455	77,5		587	100

TABELA 9
Distribuição de *você*, *ocê* e *cê*, conforme a Função Sintática - 2002

FUNÇÃO SINTÁTICA	VOCÊ			OCÊ			CÊ			TOTAL	
	N	%	PR	N	%	PR	N	%	PR	N	%
Sujeito	312	22,4	.48	35	2,5	.47	1049	75,1	.52	1396	96,1
Obj. de verbo	04	33,3	.91	05	41,7	.98	03	25,0	.17	12	0,8
Obj.Preposição	26	57,8	.97	16	35,5	.97	03	6,7	.03	45	3,1
Total	342	23,5		56	3,9		1055	72,6		1453	100

A análise será iniciada pela função sintática que apresenta o maior número de ocorrências, a de sujeito. Em seguida, a função de objeto de preposição e, por último, a função de objeto de verbo. Primeiramente serão apresentados os resultados, para, em seguida, proceder-se à análise.

5.4.1 A função de Sujeito

Repetindo-se o que se disse anteriormente, a função de sujeito, nos dois *corpora*, contém quase a totalidade dos dados. É também a função que apresenta índices mais uniformes para as três formas: não há necessariamente uma forma favorecida nesta função, já que o maior peso relativo – da forma *cê* – é de apenas .54, em 1982, e .52, em 2002.

Assim, é possível encontrar muitíssimos exemplos das três formas como sujeito, tanto de oração principal como de oração subordinada ou coordenada, como os que seguem:

(E 09/82) “Mas de repente aparece uns trem desse *você* quer resolver e não pode.”

(E 39/82) “Medicina não tem jeito d’*cê* trabalhar.”

(E 145/82) “S’*cê* errou, qualquer errinho q’*cê* errar pode ser fatal.”

(E 198/82) “Ele dá matéria pra *você* ler.”

- (E 361/82) “*Ocê* parece com a Andréia, um pouquinho.”
- (E 41/82) “... compra um lote, mas s’*ocê* comprar um lote *cê* vai construir uma casa...”
- (E 120/82) “*Cê* trabalha seis dias e folga dois.”
- (E 253/82) “*Cê* tem que ter paciência p’*cê* mexer com venda é o tipo de coisa, né?”
- (E 49/02) “*Cê* sabe que eu tenho alguns erros em português?”
- (E 750/02) “Ele tem, mas eu acho q’ *cê* não ia gostar não.”
- (E 947/02) “*Você* começa a ter um aluno-problema, *você* começa a estudar o caso, quand’*ocê* esbarra, lá na família.”
- (E 1105/02) “Se *você* passasse ali depois do Carlos Prates, *cê* tava perdido.”
- (E 53/02) “O prédio tem horário de entrar e horário de sair. *Ocê* não pode nem passar nem sair antes.”
- (E 65/02) “S’*ocê* não tiver nada pra fazer, aí *cê* vai e faz o serviço.”

Observe-se, em alguns dos exemplos acima, a confirmação da hipótese estipulada. Nos enunciados (41/82) e (65/02), a forma *ocê* foi usada depois do conectivo *se*, para se conseguir uma “economia” no processamento da fala e também o padrão silábico não-marcado CV. Para tanto, o conectivo foi reduzido, com a elisão entre o /e/ e o /o/: S<e> *ocê*, gerando “s’*ocê*”. Veja-se que, na continuação dessas frases, sem o uso do conectivo, a forma usada foi *cê*, e não *ocê*, como antes. Na próxima função, este assunto será mais detalhado.

5.4.2 A Função de Objeto de Preposição

A segunda função sintática com o maior número de ocorrências foi a de objeto de preposição, com 70 casos de uso das formas *você*, *ocê* e *cê*, nos dois *corpora*. Assim, não

serão apresentadas todas as ocorrências, mas o suficiente para que se possa analisar a hipótese acima.

As ocorrências das três formas se deram com as seguintes preposições:

1 - Com a preposição PARA

A preposição “para” foi a que apresentou o maior número de ocorrências, somando um total de trinta, nos dois *corpora*. Nos dados, ela sofreu as seguintes variações: “pra” e “p”.

“Pra” teve 24 ocorrências, sendo a metade diante da forma *você* e a outra metade diante forma *ocê*. Neste caso, houve a elisão da vogal átona da preposição com o [o] de *ocê*. As frases abaixo são exemplos dessas ocorrências:

(E 354/82) “*Cê* faz a pergunta pra *você* mesmo.”¹³⁸

(E 17/82) “Eles escrevem assim cartinha pra *você* cada coisa mais bonitinha.”

(E 529/82) “Só pra ficar olhando pr’*ocê*.”

(E 382/82) “Se quiser ficar com esse pr’*ocê* pode ficar.”

(E 35/02) “(...) vai faltar isso pra *vocês* daqui dez anos”.

(E 200/02) “Vou mandar pr’*ocê*.”

(E 270/02) “Eu vou abrir o portão pr’*ocê*.”

(E 1137/02) “Ele chega pra *você* com um bilhete (...)”

Já a forma “p” teve seis ocorrências, metade com a forma *ocê* e metade com *cê*. Os casos são os seguintes:

¹³⁸ Nessa frase, aparecem as formas *cê* (como sujeito) e *você* (como objeto de preposição). A forma na função de objeto de preposição foi a sublinhada.

- (E161/82) “Eu vou lá olhar p’*cê*.”
- (E269/82) “Aí eu posso tentar arrumar isso p’*ocê*.”
- (E 271/82) “Eu vou olhar isso p’*ocê*.”
- (E 310/02) “Vou pegar mais um pouquĩ<nho> p’*ocê*.”¹³⁹
- (E 182/02) “(...) guardar aquilo ali p’*cê* den<tro> de si.”
- (E 734/02) “Xo [Deixe eu] perguntar p’*cê*.”

2 – Com a preposição COM

Outra preposição que se destaca por ter apresentado um grande número de ocorrências foi “com”, totalizando 26 casos. Ela sofreu as seguintes variações: “com” [kũ] e “c”.

A forma “com” [kũ] teve dezessete ocorrências (65,4%). Deste total, dezesseis (94,1%) estavam acompanhadas de *você* e uma (5,9%) da forma *cê*. Abaixo seguem alguns exemplos:

- (E 410/82) “Na casa dessa moça que estava com *você*, já roubaram na casa lá.”
- (E 456/82) “Se eu fiz eu chego e falo com *você*.”
- (E 532/82) “... tá doida pra encontrar com *você*.”
- (E186/02) “Ela nunca vai pisar na bola com *você*.”
- (E 743/02) “Não sei se assim aconteceu com *você*.”
- (E 831/02) “Eu vou com *cê*.”

¹³⁹ Os segmentos e sílabas reduzidos ou omitidos na fala dos informantes aparecerão entre < >.

A forma “c” apresentou nove ocorrências (34,6%), todas acompanhadas de *ocê*. Vejam-se alguns exemplos:

- (E 549/82) “Eu falei c’*ocê* que não dava.”
 (E 231/02) “Não vou falar nada c’*ocê*, não.”
 (E 959/02) “Sair c’*ocês* dá peso.”

3 – Com a preposição DE

A terceira preposição com o maior número de ocorrências foi “de”, com nove casos, todos acompanhados da forma *você*. Observem-se alguns exemplos abaixo:

- (E 66/82) “Lógico que se ele gosta de *você*, ele nunca vai abusar.”
 (E 262/02) “Ele é o mais bonito depois de *você*, é claro.”
 (E 1285/02) “Sabe essas pessoas que ficam com raiva de *você*?”
 (E 1291/02) “Aquela pessoa que chega perto de *você* e...”
 (E 1410/02) “Eu vou me vingar de *você*, s’*cê* ligar.”

A preposição “de”, nestes *corpora*, não sofreu variação, mas se sabe que isso é possível, principalmente em determinadas regiões, como mostram os exemplos hipotéticos abaixo¹⁴⁰:

“Eu gosto tanto de *você*!” [dʒivo'se]

“Eu gosto tanto d’*ocê*!” [do'se]

¹⁴⁰ Foi assumido aqui que “de” [dʒi] e “d” [d] são formas diferentes de /de/, devido a que, nem *você*, nem *ocê* desencadeiam processos fonológicos que levam a essa variação, como seria um /i/ iniciando a palavra seguinte a essa preposição.

Assim, depara-se com um caso de opção, por parte do falante, pelo uso de uma das formas. Nos dois casos, conseguir-se-ia a forma silábica não-marcada, mas os falantes optaram pela expressão padrão “de *você*”.

4 – Com a preposição EM

A quarta preposição em número de ocorrências foi “em”, apresentando quatro casos. Ela sofreu as seguintes variações: “em”, com três casos (75%), todos acompanhados da forma *você*; e “n”, com um caso (25%), acompanhado de *ocê*. Os exemplos seguem abaixo:

a) Com a forma “em”:

(E 346/82) “*Você* pensa em você.”

(E 860/02) “Se eu pego uma arma e atiro em você, ou se eu atiro em você de costa, o que é pior, (...)”

b) Com a forma “n”:

(E 704/02) “Agora, se baixar um trem n’ocê aí, minha filha ...”

5 – Com as preposições A e POR

As preposições “a” e “por” tiveram apenas uma ocorrência cada, ambas acompanhadas pela forma *você*. Os exemplos foram:

(E 759/02) “Minha carteira não foi assinada nem por *você* nem pelo seu pai.”

(E 910/02) “Pra dar conhecimento a *você*, eu teria que (...)”

Tendo sido apresentados alguns exemplos de ocorrências das três formas na função de objeto de preposição, essas ocorrências serão, agora, analisadas.

Pelos dados acima, nota-se que, nos casos de objeto de preposição, boa parte dos informantes usou a forma *você*. Nesses casos, não foi preciso fazer qualquer adequação, qualquer redução na preposição, já que esta e a forma *você* formam uma expressão adequada aos padrões silábicos do português. Além disso, as frases nas quais essa expressão está inserida seguem igualmente o padrão rítmico-acental da língua.

No caso da forma *ocê*, observa-se sempre o mesmo fato: a preposição diante dela sempre se reduz, através do processo de elisão. Relembrando o que foi dito no Capítulo I, as características desse processo são:

- a) geralmente ocorrem em fronteira de palavras;
- b) envolvem vogais átonas;
- c) a última vogal da primeira palavra geralmente – mas nem sempre – é /a/, e a primeira vogal da palavra seguinte é posterior.

Diante disso, percebe-se que, no encontro das preposições com a forma *ocê*, ocorre naturalmente a elisão das vogais, além da redução dos encontros consonantais, como se observa a seguir.

$$\text{Pr}<\text{a}> + \text{ocê} = \text{Pr}'\text{ocê}$$

$$\text{P}<\text{r}><\text{a}> + \text{ocê} = \text{P}'\text{ocê}$$

$$\text{C}<\text{õ}> + \text{ocê} = \text{C}'\text{ocê}$$

$$\text{N}<\text{e}>^{141} + \text{ocê} = \text{N}'\text{ocê}$$

¹⁴¹ Está-se afirmando que a preposição “em” transforma-se em “ne” [ni] pelo fato de, na linguagem coloquial, ouvirem-se expressões como “[ni] mim” ou mesmo “[ni] você”.

Essas expressões, além de apresentarem estruturas silábicas do português, ou seja, CVCV – p’ocê, c’ocê, d’ocê, n’ocê – ou CCVCV – pr’ocê –, oferecem a “vantagem” de serem mais curtas e, portanto, preferíveis no processamento de fala, normalmente rápido. E, como ocorre com a forma *você*, satisfazem ao padrão rítmico-acental do português, quando inseridas nas frases.

Por outro lado, com a forma *cê*, ocorreram apenas quatro casos, sendo três com a preposição “para”, na sua forma variante “p”, e uma com a preposição “com” [kũ].

Com relação à expressão “p’cê”, como já foi dito no capítulo 1, houve a redução do encontro consonantal [pr], no nível do processamento de fala, originando “p’cê”, [pse]. Essa expressão, nos dados, foi pronunciada de três formas distintas: a) [pise], com um [i] epentético, bastante curto; b) [pose], em que, conforme já se disse, o [u] – igualmente muito curto – é resultante da assimilação da labialização de [p]; c) [pse], sem qualquer formante entre [p] e [s], conforme testes realizados no Laboratório de Fonética da Faculdade de Letras/UFMG.

Esta última forma, [pse], apresenta uma estrutura silábica bastante comum no português, que é o ataque composto CCV, e, embora o encontro consonantal [ps] não seja muito freqüente na língua, é um ataque silábico complexo possível, pois apresenta uma escala crescente de sonoridade, conforme também já se disse no capítulo 1. Além disso, o fonema [s] serve bastante bem como segundo elemento de um ataque composto¹⁴², o que pode ser comprovado pelos dados dos dois *corpora*, em que, nos casos de sujeito de orações subordinadas, diante da forma *cê*, aparecem conectivos reduzidos - d’cê, p’cê, q’cê -, e também pelas expressões “Psiu!” e “Ops!” e pelas palavras “pcisa”, “psicologia”, “psiquiatria” etc.

Com respeito à expressão [kũ’sê], vê-se que se apresenta de acordo com o padrão silábico português CVCV, sem que nenhum dos processos fonológicos vistos tenha entrado na sua

¹⁴² Em comunicação verbal, o Prof. Dr. José Olímpio de Magalhães observou que o fonema [s] apresenta várias características em comum com a vogal [i], o que favorece sua ocorrência nesses contextos.

produção. Com relação ao ritmo, o enunciado (831/02) também apresenta o padrão normal do português:

(E 831/02) “Eu you cum cê.”

Assim, pode-se perguntar por que a frase acima soa estranha, se está dentro dos padrões fonéticos e fonológicos da língua. Conforme será discutido adiante, é a comunidade que determina se uma expressão é adequada ou não, estranha ou não. E, pelo que os dados com relação a *cê* indicam, esta forma parece estar começando a expandir-se para outros contextos sintáticos, além de sujeito. Como, no início, toda mudança é estranha aos ouvidos dos falantes de uma comunidade (HOPPER e TRAUGOTT, 1993), o pouco contato com essa expressão provoca-lhes a sensação de estranheza da frase.

Passar-se-á agora à discussão da última função sintática: objeto de verbo.

5.4.3 Na função de Objeto de Verbo

Na função de objeto de verbo, houve poucas ocorrências das formas *você*, *ocê* e *cê*, nos dois *corpora*: quinze, sendo seis (40%) com a forma *você*, seis (40%) com a forma *ocê* e três (20%) com a forma *cê*. As frases são as seguintes:

a) Com a forma *você*:

(E 18/82) “Tia, gostei de ter *você* de novo.”

(E 548/82) “Eu tô querendo *você* mesmo.”

(E 97/02) “(...) e nunca larguei *vocês* por isso.”

(E 437/02) “Eles podem te... levar *vocês* lá na zona rural.”

(E 1288/02) “Eu passo e nem sou<do> ([^lsowdu] = saúdo) *você*.”

(E 1418/02) “(...) matar sua mãe e *você*”.

b) Com a forma *ocê*:

(E 117/82) “*Cê cê* confunde muit’ocê com a outra pessoa.”

(E 98/02) “Nunca larguei *ocês*.”

(E 167/02) “(...) mas às ve<ze>s te jog’*ocê* pro buraco.”¹⁴³

(E 172/02) “(...) ela acab<a> te jogand’*ocê* pro buraco.”

(E 188/02) “(...) te jog’*ocê* pra trás”.

(E 1396/02) “(...) Ah, deix’*ocê*.”

3 – Com a forma *cê*:

(E 117/02) “Eu vou ma:ta: *cê*, fdp!”¹⁴⁴

(E 142/02) “(...) dinheiro suj<o> te lev’*cê* à morte (...)”

(E 614/02) “*Cê* vai assistir o futebol *cê* ... ne<go> te arras<ta> *cê* lá dentro.

De acordo com os dados acima, vê-se que as frases em que ocorre o uso de *você* não apresenta quaisquer desvios quanto à estrutura silábica do português, como se pode ver abaixo.

(E 97/02) “(...) e nunca [lah'gey vo'ses] por isso.”

(E 437/02) “(...) eles podem te... [le'va: vo'ses] lá na zona rural.”

(E 1288/02) “Eu passo e nem [sowd vo'se].”

(E 1418/02) “(...) matar [swa'mãyivo'se]”.

¹⁴³ Observe-se que os segmentos e sílabas entre < > foram reduzidos ou omitidos, na produção dos informantes.

¹⁴⁴ Conforme dito anteriormente, o informante alongou as duas vogais do verbo “matar”.

Quanto ao ritmo, para garantir o padrão da língua, o informante reduziu a sílaba <do> no enunciado (1288/02). Por outro lado, como já se mencionou anteriormente, *você* é a forma padrão, tradicional entre as outras duas, o que garante normalmente sua ocorrência nos três contextos sintáticos. Para isso, os falantes moldam sua fala, a fim de obter as formas não-marcadas da língua.

Com relação à forma *ocê*, percebe-se o mesmo que ocorreu com a função de objeto de preposição. Vê-se que as palavras anteriores a essa forma normalmente terminam em vogal átona /o/ ou /a/, o que propiciou sua elisão com o [o] inicial de *ocê*. Com isso, conseguiu-se a estrutura silábica e também o ritmo não-marcado da língua, como se nota nos exemplos abaixo:

(E 117/82) “*Cê cê* confunde [‘m u y t’ o ‘s e]...”

C V V C V C V

(E 188/02) “(...) te [‘ʒ ɔ g’ o ‘s e] pra trás”.

C V C V C V

Por último, tem-se a forma *cê*. Pelos enunciados acima, vê-se que um deles apresenta um encontro consonantal não característico da língua, como é o encontro [vs]. Embora o fonema [v] tenha sido pronunciado pelo informante de uma forma muito breve, muito reduzida, ainda assim ele ocorreu.¹⁴⁵

(E 142/02) “(...) dinheiro suj<o> te [‘l ε v s e] à morte (...)”

C V C C V

¹⁴⁵ Quanto à ocorrência do fonema [s] depois de [v], veja-se a observação feita acima, sobre a possibilidade de [s] ser o segundo elemento de um ataque silábico composto.

Já nos outros dois enunciados, tem-se:

(E 117/02) “Eu vou [m a: 't a: s e] ...”

C V CVCV

(E 614/02) “Cê vai assistir o futebol *cê* ... ne<go> te [a' h a s e] lá dentro.”

V C V C V

Nos dois casos, tem-se a estrutura silábica canônica do português. Quanto ao ritmo, no enunciado (614/02), houve a redução da sílaba <go>, de modo a adequar a frase ao ritmo padrão, e o apagamento da sílaba <ta>, de “arrasta”, o que propiciou o uso da forma *cê*. Já o enunciado (117/02) é uma reprodução, por parte do informante, da fala de uma outra pessoa. Assim, ele alongou as duas sílabas do verbo “matar” e enfatizou também a forma *cê*. Daí o ritmo não estar de acordo com o padrão.

Concluindo este item, os dados aqui apresentados indicam que a forma *você*, sendo prototípica, aparece em todas as funções, mesmo naqueles contextos que, teoricamente, favoreceriam a forma *ocê*.

Com relação à forma *cê*, as suas poucas ocorrências normalmente obedecem às estruturas canônicas do português, com exceção do enunciado (142/02): “(...) dinheiro suj<o> te lev' *cê* à morte (...)”, em que se atesta um encontro consonantal atípico, embora ele seja possível.

Realmente, parece não haver restrições de ordem fonética ou fonológica à ocorrência de *cê* como objeto de verbo ou de preposição, haja vista que essa forma apareceu exercendo essas funções sintáticas nos dados deste estudo. Assim, seu reduzido número de ocorrências se

deve a outras razões. No próximo capítulo, quando da discussão dos resultados gerais desta pesquisa, esta questão será discutida novamente.

Com respeito à forma *ocê*, em todas as vezes em que apareceu, o contexto fonético era propício a sua ocorrência: a palavra anterior terminava por vogal átona ou semivogal. Prova disso é que, dos 95 casos de *ocê* nos dois *corpora*, apenas seis (6,3%) eram de sujeito de oração principal, em que normalmente há pausa, ou seja, não há palavra imediatamente anterior à forma. Todos os outros casos foram de objeto de verbo, objeto de preposição e sujeito de oração subordinada/coordenada.

Como a grande maioria das preposições e muitos verbos terminam em vogal átona, o falante escolhe a forma *ocê* para conseguir formas ótimas na língua, aplicando, para isso, os processos fonológicos já vistos, especialmente a elisão. Eis a explicação para a alta probabilidade de ocorrência dessa forma nas funções de objeto de verbo e de preposição.

Assim, pode-se confirmar a hipótese 2.1, acima: as expressões “preposição/verbo + *você*, *ocê* ou *cê*” terão maior propensão a ocorrer se os processos fonológicos vistos puderem atuar de forma a torná-las mais fáceis, mais “econômicas” e também não-marcadas.

É importante notar, entretanto, que o contexto fonético pode facilitar ou favorecer o aparecimento qualquer uma das formas, mas não determina sua ocorrência, haja vista os exemplos abaixo, em que o ambiente fonético favorece *ocê*, mas os falantes usaram a forma *você*.¹⁴⁶

(E 354/82) “*Cê* faz a pergunta pra *você* mesmo.”

“*Cê* faz a pergunta pr’*ocê* mesmo.”

(E 548/82) “Eu tô querendo *você* mesmo.”

“Eu tô querend’*ocê* mesmo.”

¹⁴⁶ A esse respeito, atente-se para o que foi visto no capítulo 1, item 1.3, a respeito da relação entre o contexto fonético, a frequência de uso e a redução de itens lexicais: o contexto fonético pode impedir a redução de palavras frequentes, mas se a frequência for muito alta, esse impedimento pode deixar de existir.

(E 35/02) “(...) vai faltar isso pra *vocês* daqui dez anos”.
 “(...) vai faltar isso pr’*ocês* daqui dez anos”.

(E 1288/02) “Eu passo e nem soudo *você*.”
 “Eu passo e nem soud’*ocê*.”

Assim, vê-se que a escolha de uma das três formas está condicionada ao uso da língua pela comunidade. É esta quem determina, por exemplo, se a forma *cê* pode aparecer como objeto de verbo ou de preposição, haja vista isso ocorrer em algumas áreas geográficas, como Belo Horizonte, São Francisco, Grande Vitória e Cachoeiro de Itapemirim (ES), mas não em outras, como Brasília (cf. Andrade, 2004). Assim, o que é uma frase aceitável para alguns falantes, não será para outros.

5.5 Conclusão

Neste Capítulo, foram apresentados e analisados os dados referentes ao *corpus* de 1982. Viu-se que foram poucos os fatores considerados significativos nesta amostra, bastante menos que o *corpus* de 2002. Esse fato indica que está havendo uma maior complexidade com respeito ao uso das três formas na comunidade de Belo Horizonte, principalmente com relação aos fatores lingüísticos.

Outro aspecto discutido neste capítulo foi a importância do contexto fonético para a ocorrência das formas. Observou-se que um ambiente fonético que contenha uma palavra terminada em vogal átona favorece a ocorrência principalmente de *ocê*, sendo aplicados os processos de elisão e de redução de encontros consonantais. Entretanto, como se disse, o ambiente fonético favorece, mas não determina o uso dessas formas.

Tendo sido apresentados os resultados do *corpus* de 1982, no próximo capítulo eles serão confrontados com os dados de 2002, para se investigarem os fatores mais significativos na implementação da forma *cê* e verificar, igualmente, se há indícios de mudança em progresso no tempo real.

CAPÍTULO 6 – A ANÁLISE DIACRÔNICA DOS DADOS: O ESTUDO DAS FORMAS *VOCÊ*, *OCÊ* E *CÊ* EM TEMPO REAL

Neste capítulo será desenvolvido o estudo diacrônico do comportamento das formas *você*, *ocê* e *cê* na comunidade de Belo Horizonte. Para tanto, serão contrapostos os dados de 1982 aos de 2002, com o objetivo de verificar:

- 1) Se há evidências de mudança em curso, confirmando-se os resultados do tempo aparente, ou se o uso de *você*, *ocê* e *cê* apresenta-se em variação estável nessa comunidade.
- 2) Se a forma inovadora – *cê* – está se expandindo para outros contextos, antes de domínio da forma padrão – *você*.

Antes, porém, de passar à apresentação e à análise dos resultados, é preciso esclarecer que, enquanto o *corpus* de 1982 é composto por informantes de três faixas etárias, o de 2002 é composto por informantes de cinco faixas. Tendo em vista que é necessário que haja equivalência das faixas etárias nas duas amostras, os dados das crianças e adolescentes foram suprimidos desta parte da análise.

Os resultados gerais dos *corpora* de 1982 e de 2002 encontram-se nas tabelas abaixo.

TABELA 21

Total de ocorrências de *você*, *ocê* e *cê* –
corpus de 1982

Formas	Número	%
Você	93	15,9
ocê	39	6,6
cê	455	77,5
Total	587	100

TABELA 26

Total de ocorrências de *você*, *ocê* e *cê* –
corpus de 2002 – Faixas III, IV e V

Formas	Número	%
Você	286	24,1
ocê	48	4,1
cê	851	71,8
Total	1185	100

Do resultado geral de 1982, o Programa Goldvarb 2001 selecionou apenas três fatores como significativos, um lingüístico – Função Sintática – e dois extralingüísticos – Faixa Etária e Classe Social. Desses três fatores, os dois primeiros foram selecionados também no *corpus* de 2002, o que implica que eles influenciam fortemente o uso das três formas na comunidade estudada (cf. Paredes Silva, 2003, p. 99). Portanto, a análise será iniciada por eles. Em seguida, será realizada a análise dos resultados dos fatores selecionados em apenas um dos *corpora*.

Por último, segue a apresentação dos fatores que não exercem qualquer influência no uso das formas. Apesar de estes não serem discutidos, seus resultados serão confrontados com os de alguns estudos realizados anteriormente, com vistas a confirmá-los ou não.

Antes de se passar às análises, é preciso fazer uma observação a respeito do fator Contigüidade. Esse fator não foi considerado significativo para os dados de 1982, mas o foi para os dados de 2002, computando-se as cinco faixas etárias. Entretanto, computando-se os dados referentes somente às faixas III, IV e V, ele não foi selecionado pelo Programa Goldvarb 2001. Assim sendo, a Contigüidade ao verbo não será analisada juntamente com o *corpus* de 1982.

6.1 O fator Função Sintática

O primeiro fator considerado significativo nas duas amostras foi a Função Sintática. Para se proceder a essa análise, vejam-se os resultados abaixo:

TABELA 23

Distribuição de *você*, *ocê* e *cê*, conforme a Função Sintática – 1982

FUNÇÃO SINTÁTICA	VOCÊ			OCÊ			CÊ			TOTAL	
	N	%	PR	N	%	PR	N	%	PR	N	%
Sujeito	75	13,4	.46	30	5,4	.48	454	81,2	.54	559	95,2
Obj. Verbo	02	66,7	.97	01	33,3	.90	0	0	.03	03	0,5
Obj. Preposição	16	64	.99	08	32	.99	01	04	.01	25	4,3
Total	93	15,9		39	6,6		455	77,5		587	

TABELA 27

Distribuição de *você*, *ocê* e *cê*, conforme a Função Sintática – 2002 – Faixas III, IV e V

FUNÇÃO SINTÁTICA	VOCÊ			OCÊ			CÊ			TOTAL	
	N	%	PR	N	%	PR ¹⁴⁷	N	%	PR	N	%
Sujeito	263	23,1	.48	29	2,6	.47	845	74,3	.52	1137	95,9
Objeto de Verbo	02	22,2	.65	04	44,5	.98	03	33,3	.35	09	0,8
Obj. Preposição	21	53,8	.98	15	38,5	.98	03	7,7	.02	39	3,3
Total	286	24,1		48	4,1		851	71,8		1185	100

Os dados das duas Tabelas mostram que houve um aumento muito pequeno com relação ao uso de *você*, *ocê* e *cê* nas funções de sujeito e objeto de verbo. Já a função de objeto de preposição teve um decréscimo mais acentuado, mas de apenas 0,9%, o que não é significativo. Assim, os resultados indicam que as formas continuam sendo usadas preferencialmente na função de sujeito. Nas outras duas, os falantes de Belo Horizonte preferiam e continuam preferindo os pronomes ou outra das opções já mencionadas anteriormente.

Pelos resultados acima, percebe-se que, na função de sujeito, a situação de 1982 para 2002 não mudou muito: a forma *cê* continua sendo a preferida da comunidade, vindo em segundo lugar *você* e, em terceiro, muito atrás, *ocê*. Houve modificações quanto aos percentuais de ocorrências dessas formas, com o aumento do uso, em 2002, da forma *você* e uma diminuição leve de *cê*, mas forte de *ocê*. Por sua vez, uma análise dos pesos relativos revela que muito pouca coisa mudou: houve um pequeno aumento da probabilidade de *você*, assim como uma pequena diminuição da probabilidade de ocorrência das outras duas formas.

¹⁴⁷ Na análise binomial de *você* e *ocê*, os pesos relativos para *ocê* foram: sujeito: .46; objeto de verbo: .93; objeto de preposição: .68. Confirma-se, então, mais uma vez, que a forma *ocê* é favorecida pelas funções de objeto de preposição e, principalmente, pela de objeto de verbo.

Esses resultados poderiam indicar uma tendência de uniformização, de equilíbrio no uso de *você*, *ocê* e *cê* na função de sujeito, mas, com uma diferença tão pequena, assegurar isso se torna arriscado, o mesmo valendo para a forma *ocê*, que sofreu uma levíssima diminuição de seu peso relativo. Assim, é mais prudente afirmar que não houve mudanças de 1982 para cá, não havendo, também, nenhuma forma favorecida nessa função.

Com relação à função de objeto de preposição, percebe-se que também não há uma grande diferença entre os dados de 1982 e os de 2002, nem em termos de percentual, nem em termos de peso relativo: a forma preferida pela comunidade de Belo Horizonte é *você*, seguida de *ocê* e, por último, bem distante das duas, *cê*. Com relação a *você* e *ocê*, embora haja um decréscimo da porcentagem da primeira e um ligeiro aumento da porcentagem da segunda, em 2002, suas probabilidades de ocorrência praticamente mantêm-se inalteradas. Assim, não se pode falar em mudança quanto a esta função para elas.

Por outro lado, o peso relativo de *cê* também se manteve, mas seu percentual, embora ainda muito pequeno, quase dobrou, de 1982 para 2002. O que se pode pensar desses fatos é que o aumento das ocorrências da forma *cê* como objeto de preposição nestes vinte anos ainda não é o suficiente para se manifestar em termos de peso relativo, o que provavelmente se deve ao fato de essa expansão estar em seu início. Como todo princípio de mudança é muito lento, talvez vinte anos não seja um período de tempo suficiente para que ela seja detectada.

Com respeito à função de objeto de verbo, vê-se que as posições se alteram bastante. O percentual de *você*, que era de 75,0% em 1982, caiu drasticamente em 2002, havendo subido o das formas *ocê* e *cê*. Com relação aos pesos relativos, também houve mudanças importantes: a forma *você*, embora ainda favorecida em 2002, com um PR = .66, teve uma diminuição bastante acentuada, assim como acentuado foi o aumento do peso relativo da forma *cê*: de .03, em 1982, passou .34, em 2002. Ainda que essa função desfavoreça bastante o uso de *cê*, não se pode deixar de observar que esse aumento foi significativo. Por

outro lado, com relação à forma *ocê*, vê-se que ela se mantém praticamente estável nesse período de tempo.

Tendo em vista o exposto, pode-se tentar responder às perguntas iniciais deste capítulo. Em primeiro lugar, está claro que a forma *cê* apresentava e continua apresentando um comportamento sintático peculiar. Por outro lado, percebe-se um avanço de *cê* exercendo a função sintática de objeto de verbo, ocupando um espaço que, em 1982, absolutamente não ocupava. Na função de objeto de preposição, percebe-se apenas um pequeníssimo avanço da forma *cê*. Já na função de sujeito, a comunidade de Belo Horizonte não demonstra uma forte tendência de mudança, haja vista os pesos relativos de 2002 serem muito semelhantes aos de 1982.

Passemos agora ao estudo de outro fator selecionado nos dois *corpora*.

6.2 O fator Idade

Neste item, busca-se verificar se está ou não havendo mudança em progresso com relação ao uso das formas *você*, *ocê* e *cê*, além de identificar o padrão em que se encontra a comunidade de Belo Horizonte, conforme o quadro 1, deste trabalho: se estável ou instável com relação a esse processo.

Outra questão que pode ser discutida é se os dados deste estudo validam a hipótese de que o jovem adulto é mais susceptível às pressões sociais e profissionais no sentido de padronizar sua fala. Diante disso, vejam-se as tabelas abaixo, em que se encontram os resultados quanto à faixa etária dos informantes dos *corpora* de 1982 e 2002, respectivamente.

TABELA 24
Distribuição de *você*, *ocê* e *cê*, conforme a Idade – 1982

FAIXA ETÁRIA	VOCÊ			OCÊ			CÊ			TOTAL	
	N	%	PR	N	%	PR	N	%	PR	N	%
V (+ 47 anos)	32	21,6	.62	06	4,1	-	110	74,3	.38	148	25,2
IV (31-47 anos)	16	21,6	.56	05	6,8	-	53	71,6	.44	74	12,6
III (16-30 anos)	45	12,3	.43	28	7,7	-	292	80	.57	365	62,2
Total	93	15,8		39	6,7		455	77,5		587	100

TABELA 28
Distribuição de *você*, *ocê* e *cê* conforme a Idade – 2002 – Faixas III, IV e V

FAIXA ETÁRIA	VOCÊ			OCÊ			CÊ			TOTAL	
	N	%	PR	N	%	PR ¹⁴⁸	N	%	PR	N	%
V (+ 47 anos)	119	37,3	.67	14	4,4	.66	186	58,3	.33	319	26,9
IV (31-47)	95	22,6	.53	11	2,6	.33	314	74,8	.47	420	35,5
III (16-30)	72	16,1	.35	23	5,2	.56	351	78,7	.65	446	37,6
Total	286	24,1		48	4,1		851	71,8		1185	100

Um primeiro fato depreensível de uma comparação entre as duas tabelas acima é que, tanto em 1982 quanto em 2002, o fenômeno em estudo apresenta-se como mudança em progresso: os pesos relativos de *cê* vão aumentando à medida que a idade dos informantes diminui, ao contrário do que ocorre com a forma *você*, o que indica que a forma *cê*, além

¹⁴⁸ A análise binomial de *você* e *ocê* apresentou os seguintes resultados para a forma *ocê*: faixa III: PR = .68; faixa IV: PR = .31; faixa V: PR = .53. Esses resultados, bastante semelhantes aos da análise de *ocê* e *cê* confirmam que não há mudança em progresso com relação à forma *ocê*.

de já estar implementada em Belo Horizonte, continua avançando com respeito à forma *você*.

Um segundo fato depreensível dos resultados acima é que a probabilidade de uso da forma *cê* aumentou entre as faixas etárias III: em 1982, havia um ligeiro favorecimento dessa forma (PR = .57), sendo que, em 2002, passou para .65, o que não deixa de ser um aumento significativo. A probabilidade de ocorrência de *cê* nas faixas etárias IV também aumenta, mas menos, de 1982 para 2002. O dado mais curioso, no entanto, é a diminuição da probabilidade de ocorrência de *cê* na faixa V/2002. Vê-se, assim, a tendência à padronização da fala por esses informantes atualmente.

Já a forma *ocê* não apresenta qualquer indício de mudança em progresso, nem de aumento, nem de diminuição. Assim, os dados demonstram que ela deve continuar tendo uma ocorrência marginal, restrita a alguns casos de sujeito de oração subordinada/coordenada e, principalmente, de objeto de verbo e de preposição, conforme foi visto no capítulo anterior.

Com respeito à forma *você*, observa-se o contrário do que ocorre com a forma *cê*: a faixa etária III/1982 demonstra um leve desfavorecimento à sua ocorrência, fato que se acentua entre os jovens de 2002. Vê-se, assim, que, realmente, a forma *você* está gradativamente perdendo espaço para a forma *cê*.

Com isso, pode-se responder às duas primeiras perguntas acima: os resultados dos *corpora* evidenciam uma mudança em progresso com relação ao uso da forma *cê*. Tem-se, portanto, que Belo Horizonte apresenta o padrão de mudança geracional (*generational change*), ou seja, de instabilidade, conforme o quadro 1 deste estudo.

Com respeito à tendência de padronização da fala dos jovens adultos, mencionada no capítulo 1, viu-se que ela decorre de um desejo de ascensão social e/ou profissional do indivíduo, além de possíveis pressões do ambiente profissional sobre ele, o que se chamou de pressões do mercado lingüístico. Além disso, como afirma Omena (2003, p. 66-7),

“... com a passagem do tempo, os falantes vão adquirindo a forma mais antiga e mais prestigiada na escrita padrão ou usando-a mais frequentemente. (...) Ao mudar de idade, eles [os falantes] adaptam o seu desempenho ao grupo etário a que passam a pertencer.” Conforme foi mencionado no capítulo 1, as mudanças ocorreriam no nível fonético mais superficial e no nível do léxico, e aqui se enquadram as formas *você*, *ocê* e *cê*.

Para verificar a validade da hipótese acima, observem-se de novo os dados constantes das duas tabelas, em especial os dados da faixa III/1982 e IV/2002, já que esta é representativa daquela. Aí temos a chamada pressão do mercado lingüístico.

Esses dados mostram que a faixa III/1982 favorecia levemente *cê* (PR = .57) e desfavorecia também levemente *você* (PR = .43). Entretanto, na faixa IV/2002, essas diferenças foram neutralizadas e mesmo invertidas, estando agora a forma *cê* levemente desfavorecida (PR = .47) e *você* levemente favorecida (PR = .53). O mesmo aconteceu com a faixa IV/1982: a essa época, ela desfavorecia levemente a forma *cê* (PR = .44), mas a mesma faixa etária, hoje, desfavorece muito mais (PR = .33).

Seguindo-se este raciocínio, poder-se-ia perguntar se não haveria aqui um caso de gradação por idade, e não de mudança geracional, com respeito às formas *você* e *cê*. Entretanto, o peso relativo de .65 para *cê* na faixa III/2002 indica uma tendência de aumento de uso dessa forma. Mesmo que esse grupo formalize sua fala daqui a alguns anos, aos poucos a diferença entre as gerações irá diminuir.

Em resumo, estes dados confirmam a hipótese de que os jovens adultos tendem a formalizar sua fala. Tal fato pode ocorrer devido à imagem que fazem do que seja a linguagem adequada ao seu grupo etário ou a uma situação formal de fala, como é o caso de uma entrevista. Isso implicaria que a linguagem coloquial, a fala real do dia-a-dia desses informantes adultos, poderia ser diferente do que mostraram os dados deste estudo, mas não há como se verificar tal fato aqui.

Conforme foi dito anteriormente, apenas dois fatores foram considerados significativos nos dois *corpora*. Os outros, quando são significativos num *corpus*, não o são no outro. Isso indica que o uso de *você*, *ocê* e *cê* tem sofrido modificações importantes tanto no nível lingüístico quanto no extralingüístico e se faz necessária a sua análise. Assim, no próximo item, teremos a análise do primeiro fator, Classe Social, significativo para o *corpus* de 1982.

6.3 O fator Classe Social

Neste item, serão comparados os resultados do *corpus* de 1982 aos do *corpus* de 2002, para se saber quais modificações ocorreram com relação ao fator Classe Social. Os resultados dos dois *corpora* encontram-se nas tabelas abaixo.

TABELA 25

Distribuição de *você*, *ocê* e *cê*, conforme a Classe Social – 1982

CLASSE SOCIAL	VOCÊ			OCÊ			CÊ			TOTAL	
	N	%	PR	N	%	PR	N	%	PR	N	%
Média	45	12,7	.42	26	7,3	-	282	80	.58	353	60,1
Baixa	48	20,5	.62	13	5,6	-	173	73,9	.38	234	39,9
Total	93	15,8		39	6,7		455	77,5		587	100

TABELA 29

Distribuição de *você*, *ocê* e *cê*, conforme a Classe Social – 2002 – Faixas III, IV e V

CLASSE SOCIAL	VOCÊ		OCÊ		CÊ		TOTAL	
	N	%	N	%	N	%	N	%
Média	171	23,1	32	4,3	538	72,6	741	62,5
Baixa	115	25,9	16	3,6	313	70,5	444	37,5
Total	286	24,1	48	4,1	851	71,8	1185	100

Como foi dito anteriormente, o fator Classe Social mostrou-se significativo apenas para o *corpus* de 1982. Com efeito, as diferenças no uso das três formas em 2002 são mínimas, confrontando-se as duas classes: 2,8% (25,9% - 23,1%) para *você*; 0,7% (4,3% - 3,6%) para *ocê*; e 2,1% (72,6% - 70,5%) para *cê*. Essa uniformização com respeito à classe social indica que a variação do uso das três formas, se tinha uma conotação social relevante, deixou de ter.

Tal resultado é esperado, tendo-se em vista o que foi apresentado nos capítulos anteriores acerca do uso de *ocê* e de *cê*: em ambos os casos, atuam processos fonéticos e fonológicos naturais da língua, o que eliminaria mais facilmente as diferenças sociais. Já a forma *você*, como também já se observou, é a forma prototípica, podendo ocorrer em qualquer contexto lingüístico.

Tendo em vista o exposto, é preciso, então, retomar algo que foi dito anteriormente. Quando da discussão deste fator para o *corpus* de 1982, a maior ocorrência de *você* na classe baixa foi analisada como uma formalização da linguagem por parte desses informantes, devido à situação inusitada de ter um gravador à frente, assim como ocorreu com os informantes crianças e adolescentes de 2002.

Entretanto, observando-se as freqüências de uso de *você* e principalmente de *cê* na classe baixa/1982, média/2002 e baixa/2002, nota-se que estão muito próximas. Assim, o comportamento diferenciado é, na realidade, dos informantes de classe média/1982. Portanto, em vez de se pensar numa provável atuação do Paradoxo do Observador com relação à classe baixa/1982, pode-se supor que a classe média/1982 não se ressentiu da maior formalidade da situação de entrevista, provavelmente pelo fato de que vários dos informantes eram amigos de longa data da entrevistadora, o que os desinibiu, conseguindo-se, então, algo mais próximo a seu vernáculo. Assim, por meio destes dados, se pode ter uma idéia mais real do uso tanto de *você* quanto de *cê* na comunidade de Belo Horizonte, a essa época.

Neste ponto, é importante frisar que a situação de entrevista não pode ser um curinga na análise: ela não pode ser utilizada indiscriminadamente nas ocasiões em que não se encontram respostas para o fenômeno em estudo. Assim sendo, deve-se procurar, primeiramente, explicar os resultados de todo estudo sociolinguístico a partir dos pressupostos variacionistas, atestados em diversos trabalhos da área, e justificar determinados resultados incongruentes pelo Paradoxo do Observador somente quando não existem causas lingüísticas e extralingüísticas atestadas.

Por outro lado, com relação a este objeto de estudo, especificamente, trata-se da variação de uso de um item lexical mais formal – *você* – e de dois mais informais – *ocê* e *cê* –, o que implica que a formalidade da situação pode, realmente, enviesar os resultados, favorecendo o uso de termos considerados “mais adequados” a uma ocasião pouco comum, como é a de se ter a fala gravada – às vezes por um estranho. Desse modo, a justificativa para o fato de a situação de entrevista favorecer determinados resultados do uso de *você*, *ocê* e *cê* somente pode ser pensada para alguns casos, como a Classe Social e outros, em que aspectos lingüísticos e extralingüísticos não conseguem explicá-los.

Tendo sido apresentados e discutidos os fatores que se mostraram significativos apenas para o *corpus* de 1982, será feito o mesmo para os fatores significativos apenas para o *corpus* de 2002.

6.4 O fator Referência

Com relação à referência dos itens, os resultados são os constantes das tabelas abaixo.

TABELA 30

Distribuição de *você*, *ocê* e *cê*, conforme a Referência – 1982

CLASSE SOCIAL	VOCÊ		OCÊ		CÊ		TOTAL	
	N	%	N	%	N	%	N	%
Definida	32	16	11	5,5	157	78,5	200	34,1
Indefinida	61	15,8	28	7,2	298	77	387	65,9
Total	93	15,8	39	6,7	455	77,5	587	100

TABELA 31

Distribuição de *você*, *ocê* e *cê* conforme a Referência – 2002 – Faixas III, IV e V

REFERÊNCIA	VOCÊ			OCÊ			CÊ			TOTAL	
	N	%	PR	N	%	PR	N	%	PR	N	%
Definida	70	14,2	.35	28	5,6	.32	396	80,2	.65	494	41,7
Indefinida	216	31,3	.61	20	2,9	.59	455	65,8	.39	691	58,3
Total	286	24,1		48	4,1		851	71,8		1185	100

Os resultados acima mostram que, em 1982, não havia diferenças significativas quanto ao uso das três formas para indicar referência definida ou indefinida. A vantagem da referência definida sobre a indefinida quanto às três formas era tão pequena que não poderia ser considerada e, portanto, esse fator não foi selecionado como significativo para o *corpus* de 1982.

Entretanto, essas diferenças se acentuam muito, em se tratando do *corpus* de 2002: a porcentagem de uso da forma *você* para expressar referência indefinida praticamente dobrou, de 1982 para 2002; daí seu favorecimento (PR = .61). É interessante notar que, assim como o uso de *você* cresceu, o uso tanto de *ocê* quanto de *cê* diminuiu; daí também o desfavorecimento da forma *cê* (PR = .39). Observe-se que as porcentagens da referência definida para as três formas se alteram muito pouco, de 1982 para 2002. O grande impacto se dá, realmente, com a referência indefinida.

A partir desses resultados, pode-se comprovar o que se disse anteriormente, quanto à referência: a forma *cê* está sendo mais usada para indicar referência definida, assim como a forma *você* está sendo preferencialmente usada para expressar referência indefinida, embora, como se viu no capítulo 4, os dados indiquem uma mudança em curso com respeito à forma *cê* e a referência indefinida. Em resumo, os dados não confirmam a hipótese de número 4, estipulada neste estudo. Já a forma *ocê*, mesmo tendo pouquíssimas ocorrências, quando usada, é também para expressar referência definida.

Com isso, através do tempo real, confirma-se uma mudança em curso na comunidade de Belo Horizonte, quanto ao emprego das formas para expressar referência.

6.5 O fator Focalização

O fator Focalização também foi considerado significativo somente para o *corpus* de 2002. Com relação a este fator, observem-se os resultados constantes das tabelas abaixo.

TABELA 32

Distribuição de *você*, *ocê* e *cê* conforme a Focalização – 1982

CLASSE SOCIAL	VOCÊ		OCÊ		CÊ		TOTAL	
	N	%	N	%	N	%	N	%
Foco	03	60	02	40	0	0	05	0,9
Não-foco	90	15,5	37	6,3	455	78,2	582	99,1
Total	93	15,8	39	6,7	455	77,5	587	100

TABELA 33

Distribuição de *você*, *ocê* e *cê*, conforme a Focalização – 2002 – Faixas III, IV e V

FOCALIZAÇÃO	VOCÊ			OCÊ			CÊ			TOTAL	
	N	%	PR	N	%	PR ¹⁴⁹	N	%	PR	N	%
Foco	06	66,7	.91	02	22,2	.99	01	11,1	.09	09	0,8
Não-foco	280	23,8	.50	46	3,9	.49	850	72,3	.50	1176	99,2
Total	286	24,1		48	4,1		851	71,8		1185	100

Observa-se, pelas tabelas, que, quando não se trata de casos de focalização, obviamente os resultados estão de acordo com os totais gerais dos dois *corpora*. Assim, os dados interessantes dizem respeito à focalização. Neste caso, percebe-se que a forma *você* não sofreu alterações significativas no seu uso em 1982 e em 2002, mas sim as formas *ocê* e *cê*. Enquanto aquela teve uma redução acentuada de seu uso – quase 50% -, esta, que não teve nenhuma ocorrência em 1982, teve uma em 2002.

¹⁴⁹ Na análise binomial de *você* e *ocê*, o fator Foco não foi selecionado pelo Programa como significativo.

Pode-se justificar a diminuição do uso da forma *ocê* em virtude de a comunidade, com seu crescimento, passar a evitá-la, por vê-la como uma marca da fala interiorana (cf. Ramos 1997). Assim, uma possível avaliação negativa dos falantes faria com que eles tendessem a não reforçá-la, como acontece nos casos de focalização contrastiva.

Os casos de foco contrastivo com a forma *ocê*, em 2002, já foram mostrados no Capítulo IV. Os casos de 1982 foram:

(E 117) “(...) *cê cê* confunde muit’***ocê*** com a outra pessoa (...)”

(E 500) “***Ocê*** precisa de ver que menina inteligente!”

Já nos casos de focalização contrastiva com a forma *cê*, percebe-se que o aumento foi irrisório, em vinte anos: apenas uma ocorrência em 2002. Assim, poder-se-ia falar em mudança? Para se responder a essa pergunta, a análise não se deve restringir ao número de casos em 1982 e em 2002.

Como se apontou anteriormente, a focalização contrastiva – ou enfática – constitui um reforço na entonação de um termo, o que implica fazê-lo com um pouco mais de lentidão. No caso da forma *cê*, ela é uma redução de *você* e, como tal, está diretamente ligada à velocidade normal de fala, que é rápida. Nesse sentido, a ocorrência de *cê* recebendo foco não é esperada, parecendo mesmo estranha. Talvez por isso tenha havido um só caso nos dois *corpora*.

Por outro lado, esse caso único se mostra muito interessante, pois demonstra que uma mudança pode estar tendo início, ou seja, os falantes podem estar começando a ver essa forma não mais como uma simples redução fonológica de *você* e, assim, podem estar começando a usá-la em outras situações. Se se analisa este dado juntamente com outros, especificamente a função sintática de objeto de verbo e de objeto de preposição, percebe-se, realmente, que a forma *cê* está sendo usada em outros contextos, não somente em Belo Horizonte, como também em São Francisco (cf. Coelho, 1999), o que pode ser um indício

de um lento espraiamento dessa forma para outros ambientes sintáticos, antes exclusivos de *você* e *ocê*.

6.6 O fator Gênero

O último fator analisado é Gênero, que também foi considerado significativo apenas para o *corpus* de 2002. Os resultados deste fator encontram-se nas tabelas abaixo.

TABELA 34
Distribuição de *você*, *ocê* e *cê*, conforme o Gênero – 1982

GÊNERO	VOCÊ		OCÊ		CÊ		TOTAL	
	N	%	N	%	N	%	N	%
Feminino	37	17,2	15	6,9	164	75,9	216	36,8
Masculino	56	15,1	24	6,5	291	78,4	371	63,2
Total	93	15,8	39	6,7	455	77,5	587	100

TABELA 35
Distribuição de *você*, *ocê* e *cê*, conforme o Gênero – 2002 – Faixas III, IV e V

GÊNERO	VOCÊ			OCÊ			CÊ			TOTAL	
	N	%	PR	N	%	PR ¹⁵⁰	N	%	PR	N	%
Feminino	94	16,2	.41	31	5,3	-	455	78,4	.59	580	48,9
Masculino	192	31,7	.58	17	2,8	-	396	65,5	.42	605	51,1
Total	286	24,1		48	4,1		851	71,8		1185	100

¹⁵⁰ O fator Gênero não foi selecionado em nenhuma das análises binárias com a forma *ocê*.

Comparando-se os resultados de 1982 e os de 2002, observa-se que o comportamento das mulheres não sofreu alterações significativas com relação a nenhuma das três formas: *cê* continua a ser a preferida, sendo ligeiramente favorecida por elas em 2002, vindo em seguida *você* e depois *ocê*.

Com relação ao sexo masculino, por outro lado, no *corpus* de 2002, nota-se um também leve favorecimento da forma *você* por eles e um desfavorecimento das outras duas formas. Pode-se pensar, neste caso, que está havendo alguma mudança com respeito ao gênero?

Conforme foi discutido no capítulo 1, os estudos sociolingüísticos têm frisado que as mulheres usam preferencialmente as formas valorizadas socialmente. Além disso, são elas também que lideram as mudanças lingüísticas, desde que estas não sejam estigmatizadas na comunidade, e é este o caso da forma *cê*.

A partir do exposto, pode-se supor que os homens usem mais a forma *você* por estarem normalmente “atrás” das mulheres nos processos de mudança lingüística. Entretanto, não é isso o que se observa pelos dados das duas tabelas. O percentual de uso da forma *cê* por homens e mulheres, em 1982, e pelas mulheres, em 2002, são muito parecidos. O resultado que se destaca é realmente dos homens, em 2002. Assim, qual a razão dessa mudança de preferência, nos homens? Por que eles estariam preferindo, atualmente, a forma *você*, em vez de *cê*?

A princípio, pode-se pensar em duas hipóteses para essa preferência:

- (i) os homens atualmente estão utilizando mais as formas-padrão da língua, sendo mais formais ao falar;
- (ii) aconteceu uma hipercorreção por parte dos informantes do sexo masculino.

Para analisar as hipóteses acima, é preciso chamar a atenção para três fatos:

1) Dos onze entrevistados do sexo masculino, em 2002, 36,4% faziam Mestrado em Linguística; 18,2% tinham Nível Superior, mesmo que incompleto, e 45,5% tinham até o Ensino Médio. Por sua vez, das doze entrevistadas, 33,3% faziam Mestrado; 16,6% tinham Nível Superior e 50% tinham até o Ensino Médio. Ou seja, as diferenças entre o grau de instrução dos homens e das mulheres são mínimos, e isso implica que a maior frequência de *você*, a forma padrão, ou seja, a maior formalidade dos homens não se deve à escolaridade.

2) Os resultados de crianças e adolescentes, vistos no Capítulo IV, demonstraram que, na primeira parte da entrevista, que foi considerada como uma situação de fala mais formal que a segunda, a forma *você* foi levemente favorecida pelos informantes masculinos. Assim, ao que parece, estes informantes relacionam o uso de *você* a uma maior formalidade da fala: numa situação “anormal”, sentem que as formas-padrão são as mais adequadas.

3) Embora a entrevistadora de 1982 fosse amiga de boa parte de seus informantes de classe média, este fato acontecia de uma forma mais ou menos equilibrada para homens e mulheres. Com relação à entrevistadora de 2002, esta era conhecida de um número também mais ou menos equilibrado de informantes do sexo masculino e feminino. Portanto, os resultados não se explicam por uma diferença de envolvimento entre as entrevistadoras e seus informantes homens e mulheres.

Tendo em vista essas três observações, creio que se pode explicar o comportamento masculino a uma maior atenção, a um maior cuidado com a fala em situações formais de uso da língua pelos informantes de 2002 do que os de 1982.

Poder-se-ia especular aqui sobre as causas que levaram somente os informantes masculinos de 2002 a preferirem a forma *você*: seriam elas um efeito das pressões do mercado lingüístico mais fortes agora? Está mais acentuada, atualmente, a correlação entre o uso de uma linguagem mais formal e a imagem de poder, *status* e valorização do homem na

sociedade belo-horizontina? Ou realmente houve diferenças quanto ao nível de interação entre as duas entrevistadoras e seus respectivos entrevistados masculinos, e isso influenciou os dados?

É difícil responder a essas perguntas; para isso seria necessário refinar a análise dos dados, o que demanda tempo extra, além do fato de que essas respostas fogem aos objetivos deste estudo. Assim, pode-se afirmar apenas o que os dados indicam: os homens atualmente importam-se mais em fazer a adequação de sua linguagem de acordo com o contexto, tornando-a mais formal, quando sentem que é adequado.

6.7 Fatores não-significativos

Como foi dito no capítulo 3 desta tese, a escolha dos fatores analisados deveu-se, entre outras razões, ao objetivo de comparar os resultados obtidos neste trabalho aos de outros estudos já realizados sobre as três formas. Dos trabalhos apresentados, aqueles que analisam pelo menos um desses fatores são Ramos (1997), Coelho (1999) e Andrade (2004). Assim, a comparação será feita com eles.

Os fatores “Tipo de frase em que aparecem as formas” e “Posição de tópico” não foram considerados significativos para nenhum dos dois *corpora* deste estudo, mas foram significativos para os *corpora* trabalhados por essas autoras. Deste modo, serão contrapostos esses dados aos dados de 2002 desta pesquisa - por estarem mais próximos no tempo -, com o objetivo de se investigar o comportamento das formas *você*, *ocê* e *cê* nessas amostras. Os resultados de 2002 são os constantes da mesma Tabela 8, a seguir.¹⁵¹

¹⁵¹ Com o objetivo de melhor avaliar o comportamento das três formas, serão analisados os dados de 2002 de todos os informantes.

TABELA 8
Distribuição de *você*, *ocê* e *cê* – 2002 – Fatores não-significativos¹⁵²

FATORES NÃO-SIGNIFICATIVOS		VOCÊ		OCÊ		CÊ		TOTAL	
		N	%	N	%	N	%	N	%
Tipo de frase	Declarativa	324	25,1	52	4,0	917	70,9	1292	88,9
	Interrogativa	16	11,0	03	2,0	126	87,0	145	10,0
	Interr. ‘Que-Que’	04	25	0	0	12	75	16	1,1
Tópico	Tópico	01	33,3	01	33,3	01	33,3	03	0,2
	Não-tópico	341	23,5	55	3,8	1054	72,7	1450	99,8

As conclusões das autoras em relação a esses fatores vêm a seguir.

Ramos (1997) afirma que:

- a) a forma *cê* é a preferida, nas construções topicalizadas;
- b) a forma *cê* é preferencialmente usada nas orações declarativas e nas interrogativas com estrutura “que que”. Nas interrogativas simples, por sua vez, os informantes preferem a forma *você*.

Os resultados de Coelho (1999) que podem ser contrapostos a estes se referem ao tipo de oração em que se encontram as formas. A autora não discriminou as interrogativas simples das interrogativas do tipo “que que”. Assim, não há como fazer comparações a este respeito. Por sua vez, nas orações declarativas (afirmativas + negativas, na análise de Coelho), *cê* é a forma preferida pela comunidade de São Francisco/MG, com 54,9%, vindo, em segundo lugar, *ocê* (23,6%) e, em seguida, muito próxima a esta, *você*, com 21,4%.

¹⁵² Nesta parte da análise, estão sendo confrontados os fatores considerados não-significativos nos dois *corpora*.

Os resultados de Andrade (2004) indicam que a forma *você* é a preferida nos três tipos de orações.

Agora, então, esses dados podem ser confrontados. E é interessante perceber como *corpora* diferentes podem levar a resultados distintos.

Com relação ao Tópico, nos dados de 2002 desta tese, há empate entre as três formas com respeito às construções com tópico, ao passo que os dados de Ramos indicam predomínio da forma *cê*. No caso de não-topicalização – um contexto não-marcado –, os resultados dos dois trabalhos são idênticos, e isso não é surpresa, tendo em vista que a forma *cê* está implementada em Belo Horizonte, tendo 72,6% de uso entre os informantes desta pesquisa.

Já com relação ao Tipo de Frase, também não há consenso entre os trabalhos. Com relação às orações declarativas, os dados deste estudo e de Coelho confirmam a preferência dos informantes pela forma *cê*. Por sua vez, nos dados de Andrade, vence a forma *você*. Com respeito às interrogativas simples, os resultados de Andrade confirmam os de Ramos, com a preferência daqueles informantes pela forma *você*, ao contrário dos informantes deste estudo, que preferiram a forma *cê*. Por fim, nas interrogativas com a estrutura “que que”, os dados deste estudo confirmam os de Ramos, mas o contrário ocorre com os dados de Andrade.

Em suma, os dados deste estudo, de Coelho e de Andrade apontam para uma preferência da forma mais freqüente nas respectivas comunidades: *cê* em Belo Horizonte e em São Francisco, e *você* em Brasília. Os dados de Ramos seguem o mesmo padrão acima, com exceção das orações interrogativas simples.

Os resultados discordantes dos quatro trabalhos apontam para o fato – comumente citado em estudos sociolinguísticos – de que os usos linguísticos são determinados pela comunidade de fala. Desse modo, o que é válido - ou até gramatical - numa localidade, pode não ser para outra. Por conseguinte, esses resultados indicam que é preciso cautela na generalização de resultados encontrados para uma comunidade específica.

6.8 Conclusão

Neste capítulo, foram contrapostos os resultados dos *corpora* de 1982 e de 2002, com o objetivo de analisar o comportamento das três formas nestes vinte anos e também de verificar se está havendo mudança de comportamento das três formas em relação aos fatores analisados.

Primeiramente, foram discutidos os fatores considerados significativos nas duas amostras e, em seguida, os considerados significativos apenas em uma. Os resultados indicam mudança com relação à função sintática de objeto de verbo, idade e referência indefinida, com o aumento da probabilidade de ocorrência de *cê* de 1982 para 2002. Com respeito aos outros fatores, não se observaram mudanças significativas.

Tendo sido expostos e discutidos os resultados das duas amostras nos três últimos capítulos, agora, então, pode-se passar à discussão geral dos resultados e às conclusões deste estudo. É o que vem a seguir.

CAPÍTULO 7 - DISCUSSÃO DOS RESULTADOS E CONCLUSÕES

Este trabalho teve por objetivo investigar o comportamento lingüístico e extralingüístico das formas *você*, *ocê* e *cê*. Tendo em vista os resultados discordantes obtidos por outros estudos, que analisavam *corpora* diversos, pretendeu-se aqui ampliar e aprofundar a discussão a respeito do uso dessas formas na comunidade de Belo Horizonte. Para tanto, foram analisados dados de fala de duas épocas distintas – 1982 e 2002 –, a fim de se saber se, nesse espaço de tempo, estaria havendo mudanças com relação a alguma dessas três formas e também a fim de se caracterizar o comportamento da comunidade belo-horizontina, conforme quadro apresentado em Labov (1994, p. 83, Table 4.1).

Foram analisados seis fatores lingüísticos: a Função Sintática que as formas exercem nas frases, seu uso quanto à expressão da Referência, o Tipo de Frase em que aparecem, seu comportamento nas construções com Tópico e quanto à Contigüidade ao verbo, e se podem ou não receber Foco contrastivo ou enfático. Além desses, três fatores extralingüísticos: Idade, Gênero e Classe Social dos informantes. Através da análise desses fatores, pretendia-se caracterizar a variação/mudança dentro da estrutura lingüística do PB, assim como dentro da estrutura social de Belo Horizonte.

Como foi visto ao longo desta tese, alguns desses fatores se mostraram relevantes para os dois *corpora*. Foi o caso da Função Sintática das formas e da Idade dos informantes. Outros fatores foram relevantes para apenas um dos *corpora*, como Classe Social, Referência, Focalização, Contigüidade e Gênero. E outros fatores não foram significativos

para nenhuma das amostras, como o Tipo de frase e o Tópico. Assim, nos Capítulos IV, V e VI, procedeu-se à análise desses fatores, para que se confirmassem ou não as hipóteses constantes da Introdução deste trabalho, levantadas a partir dos estudos realizados por outros autores sobre as três formas. Seria interessante iniciar a discussão dos resultados por esta parte, para, em seguida, se analisar o comportamento – nos níveis lingüístico e extralingüístico – das formas. A título de comodidade, as hipóteses vêm repetidas abaixo:

HIPÓTESES

- 1) A forma *cê* apresenta um comportamento sintático peculiar, não ocorrendo nos mesmos contextos que *você* e *ocê*.
- 2) O uso das formas *você*, *ocê* ou *cê* é condicionado por fatores de natureza fonética e fonológica da língua portuguesa.
- 3) A forma *ocê* deve ter relativamente pouca freqüência nos dados.
- 4) Está havendo, no PB, um processo de especialização em relação ao par *você* e *cê*, com a preferência de *cê* com referência indefinida.
- 5) A comunidade de Belo Horizonte, com relação às três formas, apresenta o padrão de mudança geracional (cf. Labov, 1994, p. 83).

Entretanto, no Capítulo I, a segunda hipótese foi mais detalhada, resultando em 2.1, abaixo:

Hipótese 2.1: O uso das formas *você*, *ocê* e *cê* busca adaptá-las aos padrões fonéticos e fonológicos da língua portuguesa, no tocante à estrutura silábica, acentuação e ritmo.

Pode-se continuar, então, a discussão dos resultados.

A hipótese mais rápida e facilmente comprovada foi a de número 3, já que as primeiras tabelas dos dois *corpora* demonstram que *ocê* não conseguiu nem 7% do total de ocorrências. O melhor desempenho de *ocê* foi no *corpus* de 1982, com apenas 6,6% dos dados.

Entretanto, não resta dúvida de que esse baixo percentual se deve ao fato de que, nas posições de objeto de verbo e de objeto de preposição, que favorecem *ocê* enormemente, as três formas têm pouco uso. Como se disse anteriormente, principalmente na função de objeto de verbo, essas formas sofrem a concorrência de outros pronomes e da forma zero. Assim, elas têm ocorrência reduzida nessas funções.

Por outro lado, foram exatamente esses resultados discrepantes de *ocê* que comprovaram a hipótese 2.1 acima. Viu-se que, nas funções de objeto de verbo e de preposição, o ambiente fonético anterior – normalmente vogal átona – leva o falante a escolher *ocê*, utilizando-se dos processos fonológicos vistos no item 1.4. Já as formas *você* e *cê* não são favorecidas por esses processos, mas *você* tem *ocorrência* garantida nessas funções, ao passo que a probabilidade de *cê* exercê-las é muito reduzida. Entretanto, isto não será discutido agora. Mais adiante este ponto será retomado.

Uma outra hipótese facilmente comprovada foi a número de 01. Realmente, *cê* não ocupa os mesmos ambientes sintáticos que *você* e *ocê*, sendo ela bastante desfavorecida como objeto de verbo e de preposição. Como já foi dito, esse fato não se deve especificamente a um ambiente fonético desfavorecedor, por uma razão: mesmo em contextos fonéticos desfavorecedores, a forma *cê* ocorreu como objeto de verbo e de preposição. Por sua vez, a forma *você* também apareceu exercendo essas funções, em ambientes que favoreceriam a forma *ocê*.

Daí poder-se ver que o contexto fonético pode favorecer a ocorrência de uma forma, mas não impede ou, por outro lado, não obriga o seu uso. Como será discutido à frente, a pouca ocorrência de *cê* exercendo essas funções, bem como recebendo foco, está relacionada ao próprio processo de mudança e aos usos lingüísticos da comunidade. Prova disso é que as

ocorrências de *cê* são escassas, mas não nulas, e aumentaram um pouco de 1982 para cá. Mas, como foi dito antes, esta discussão será deixada para mais tarde.

Quanto à hipótese de número 04, os dados desta tese não a confirmaram, nem se computando todos os dados, nem os computando apenas na função de sujeito. Assim, computando-se os dados globais, *você* é preferencialmente usado para expressar referência indefinida, ao passo que *cê* é mais favorecido nas referências definidas. Entretanto, a análise da probabilidade de uso para cada faixa etária não deixa dúvidas de que está aumentando o uso de *cê* para expressar também a referência indefinida. Pensando-se que esses resultados se referem ao *corpus* de 2002 e que, em 1982, não havia diferenças significativas entre a expressão da referência pelas formas *você* e *cê*, pode-se concluir que esses resultados indicam uma tendência de mudança, mas não no sentido do que se havia pensado inicialmente, neste estudo.

Por último, tem-se a hipótese de número 05. A comprovação dessa hipótese é a mais complexa, já que os resultados desta pesquisa foram obtidos através de muitos grupos etários, sendo que alguns apresentaram comportamento peculiar. Assim, os dados de 1982 indicaram mudança em curso: a forma *cê* estava sendo gradativamente preferida pelas faixas etárias mais jovens, sendo que o oposto ocorria com *você*. Já os dados de 2002 são curiosos. Até a faixa III (de 16 a 30 anos), vê-se mudança em progresso, com o aumento gradativo da forma *cê* sobre *você*. Entretanto, os resultados das crianças e, principalmente, dos adolescentes levam a pensar numa reversão dessa tendência.

Contudo, essa hipótese foi descartada, tendo em vista que, nos testes – que consistiram na primeira parte das entrevistas, quando os informantes estavam menos acostumados a essa situação –, o uso de *você* cresceu acentuadamente: 45,2%, contra 52,7% de *cê*. Já na segunda parte das entrevistas, o uso de *você* diminuiu e o de *cê* aumentou, com relação aos testes. Assim, a frequência mais baixa de uso de *cê* nessas duas faixas etárias foi analisada como sendo o fruto da formalidade gerada pelas entrevistas, mesmo na segunda parte, quando esses informantes já estavam um pouco mais à vontade. Se pensarmos no que dizem Chambers (1995) e, principalmente, Piaget (2003) a respeito das crianças e

adolescentes, essa hipótese se justifica. Portanto, pode-se dizer que as crianças e os adolescentes não estão retardando a mudança de *você* para *cê* e que os dados dos dois *corpora* indicam mudança em curso. Se compararmos o peso relativo da faixa mais jovem de 1982 (.57) com o de 2002 (.65), comprovaremos, então, este fato.

Enfim, os dados relativos à idade confirmam que a mudança de *você* para *cê*, em Belo Horizonte, é geracional.

Finda essa discussão, será analisado, agora, o processo lingüístico que envolve essas formas. Viu-se, no capítulo 1, que uma mudança não é um fato isolado, mas que ocorre dentro de um contexto lingüístico e social, e a mudança sob estudo não é diferente. Faraco (1996), Lopes (2003; 2003a) e Lopes e Duarte (2003; 2004) já descreveram parte dela e, no capítulo 2 desta tese, foi feito um resumo acerca do assunto. Cabe agora, então, correlacionar as origens dessa mudança aos resultados obtidos neste estudo, a fim de analisá-la mais amplamente, respondendo, assim, às questões apresentadas por Weinreich, Labov e Herzog (1968).

A mudança de *você* para *cê* originou-se há muito tempo, exatamente no século XIV, com a introdução, na língua portuguesa, de expressões do tipo *Vossa + Nome*, mais especificamente, neste caso, *Vossa Mercê*. Essa expressão, que inicialmente era de uso exclusivo ao rei, passa a se generalizar, servindo, posteriormente, para o tratamento aos nobres e, por fim, a pessoas de qualquer classe social a quem se quisesse tratar com deferência ou simplesmente com respeito.

A generalização do uso trouxe o aumento da frequência dessa expressão, o que, de acordo com os estudos de Fidelholtz, Phillips e Bybee, apresentados no item 1.3, gera mudanças sonoras. No caso específico de *Vossa Mercê*, o uso frequente levou-a a se reduzir, gerando diversas variantes. Aliado à expansão e à variação de *Vossa Mercê* está o fato de que o pronome *vós* caiu em desuso, o que abriu espaço para a entrada na língua da variante *vocês*.

Essa era a situação lingüística que foi trazida da metrópole para o Brasil, à época da colonização. Aqui, o uso de *você* ocorre concomitantemente ao de *Vossa Mercê*, *vós* e *tu*. Entretanto, aos poucos as formas *Vossa Mercê* e *vós* caem completamente em desuso. Quanto à forma *você*, na primeira metade do século XX, suplanta o uso de *tu* em grande parte do país. Assim, *você*, que se originou nas classes sociais mais baixas, em Portugal, passou a ser de uso geral no Brasil.

A entrada de *você* na língua portuguesa trouxe como consequência uma série de rearranjos em sua estrutura: o processo de gramaticalização de *Vossa Mercê* para *você* restringiu o contexto de atuação desta forma, que ocupava, inicialmente, somente a função de sujeito pré-verbal, ao contrário de *Vossa Mercê*, que tinha maior mobilidade na frase. Além disso, sendo *você* uma forma originada de uma expressão nominal, a qual, em realidade, era uma exaltação às qualidades do rei, essa forma herdou de *Vossa Mercê* a correspondência à segunda pessoa, mas a concordância com a terceira.

O resultado do quadro acima é que, numa língua de sujeito nulo, como era o PB, formas como “amava”, “partia” etc. serviam tanto à primeira pessoa quanto à segunda. Por fim, para desfazer essa ambigüidade, o falante sentiu a necessidade de preencher o sujeito de suas frases.¹⁵³ Tem-se aí o quadro onde se insere a mudança de *Vossa Mercê* para *você*.

O que se disse até aqui foi discutido pelos autores citados no capítulo 2 desta tese. Entretanto, este objeto de estudo refere-se à continuação desse processo de redução fonética e o surgimento de *cê*. Assim, para se analisar essa mudança, é preciso continuar a descrevê-la.

O processo de gramaticalização pelo qual passou *você*, que o transformou em pronome – uma palavra funcional, mais susceptível às mudanças sonoras, conforme apontou Phillips (2001, p. 128-9) – e a crescente obrigatoriedade do preenchimento do sujeito culminaram no aumento do uso da forma *você*, e essa freqüência levou-a a continuar seu processo de redução fonética, originando a forma *cê*.

¹⁵³ A hipótese acima, com a qual concordo, aparece em Faraco (1996, p. 68)

Como foi mostrado no item 1.1 deste trabalho, em alguns *corpora* analisados sobre a variação de uso de *você*, *ocê* e *cê*, constatou-se que, ao contrário da forma *você*, a forma *cê* somente ocorria como sujeito pré-verbal, era favorecida pela contigüidade ao verbo, não aparecia como foco contrastivo etc. Esse fato levou alguns dos autores citados no capítulo 1 a lançar a hipótese de que a forma *cê* estaria continuando o processo de gramaticalização iniciado pela forma nominal *Vossa Mercê*: *cê* estaria tornando-se um clítico, sendo talvez, por ora, um pronome fraco.

Essa hipótese é interessante e justificável, tendo em vista os resultados dos *corpora* analisados por esses autores. Mas, por outro lado, outros dados, obtidos por meio de outros *corpora*, demonstram que *cê*, mesmo timidamente, aparece em contextos onde clíticos não apareceriam. Os dados desta tese e os de Coelho (1999), analisados por ela própria e por Barbosa (2005), demonstram que *cê* exerce a função de objeto de verbo e de preposição, além de poder vir muito distante do verbo e receber foco ou ênfase. Então, como analisar o comportamento de *cê*? As ocorrências desta forma nessas situações, tão escassas, seriam expressões pronunciadas por equívoco, saídas ao acaso?

Em primeiro lugar, é preciso lembrar que as mudanças “*from below*”, que são as mudanças lingüísticas propriamente ditas, ocorrem exatamente desse modo, ou seja, abaixo do nível de consciência dos falantes. Imagina-se, então, que os processos de mudança se iniciem dessa maneira, com esse tipo de inovações por parte dos que os iniciam.

Também é importante ter-se em mente que, no princípio, toda mudança é lenta. Assim, esse processo, que, em seu início, restringe a atuação de uma forma a determinados contextos, aos poucos permite o espraiamento de seu uso. Isso aconteceu com a forma *você*: inicialmente restrita à função de sujeito pré-verbal, foi aos poucos ampliando sua atuação para outros contextos. O mesmo sucedeu com os pronomes “eu” e “ele” (cf. Barbosa, 2005, p. 53) e “you”, no inglês (cf. Nevalainen, 2000, p. 262): no início, exerciam somente a função de sujeito, passando depois a exercer outras funções.

Há um outro ponto a ser discutido sobre o comportamento da forma *cê*. Vê-se que os contextos em que ela é mais desfavorecida são aqueles em que exerce as funções de objeto de verbo e de objeto de preposição e em que recebe foco. E esses contextos têm um ponto em comum: são pouco freqüentes, com uma pequeníssima porcentagem de ocorrência, ou seja, são marcados. Omena (2003, p. 70) observa que os fatores não-marcados favorecem as mudanças, ao passo que os marcados tendem a restringi-las. Pode-se, então, aqui, estipular uma hipótese análoga: contextos marcados tendem a frear as mudanças, como já apontaram alguns estudos sobre mudança lingüística. Portanto, pode-se pensar que a focalização e as funções sintáticas de objeto de verbo e de preposição, por serem marcadas, retardam a atuação da forma inovadora *cê*, mesmo ela já sendo a grande preferida da comunidade de Belo Horizonte, com mais de 70% do total das ocorrências.

Com relação a isso, é importante observar novamente o que diz Trudgill (1988, p. 44-5), a respeito do espraiamento das formas inovadoras:¹⁵⁴

This graph [que apresenta o resultado de [t]] thus provides an interesting example of the way in which a change, having gone almost to completion in casual speech, continues to spread from style to style. We can take this as a vivid example of the way in which linguistic innovations spread not only from person to person, area to area, class to class, and linguistic environment to linguistic environment, but also from contextual style to contextual style. Here is confirmation that most linguistic changes begin in unmonitored, vernacular informal styles and only later spread to more formal varieties of speech.

Assim, vamos voltar aos poucos casos de uso de *cê* pelos informantes nesses contextos marcados. Está claro que não se pode afirmar com certeza que eles sejam um princípio de mudança. Como apontam Hopper e Traugott (1993, p. 95): “... there is nothing deterministic about grammaticalization and unidirectionality. Changes do not have to occur.” (grifo meu). Além disso, uma mudança, para acontecer, passa pela avaliação da

¹⁵⁴ O autor, nesse trecho, refere-se à expansão do [t] como oclusiva glotal, a forma inovadora analisada, chamando a atenção mais especificamente para o seu espraiamento dos contextos informais para os mais formais. Entretanto, está-se apontando aqui para o espraiamento de um modo geral, e não somente para o estilo de fala.

comunidade. E, conforme estes autores apontam, no início, toda mudança é estranha; assim, os falantes podem rejeitar certas inovações. É a comunidade que determina o avanço ou o bloqueio das mudanças lingüísticas.

Entretanto, o aumento – mesmo muito modesto – das ocorrências de *cê* nesses ambientes, atestado nas amostras de 1982 e de 2002, e, também, esse uso sendo encontrado nas interações lingüísticas cotidianas, em Belo Horizonte, podem indicar uma melhor avaliação e uma maior aceitação dos falantes para esse uso, indicando, assim, igualmente, um início de mudança, em que *cê* avança – lentamente, é verdade – para outros contextos, antes de uso exclusivo de *você* e *ocê*.

Neste ponto, deve-se atentar de novo para o fato de que, nesta tese, trabalhou-se com entrevistas sociolingüísticas, o que implica uma maior formalidade da linguagem dos informantes. Assim, o Paradoxo do Observador, se é um problema para o pesquisador que analisa fenômenos fonológicos, por exemplo, torna-se um verdadeiro desafio quando o fenômeno analisado envolve itens com níveis distintos de formalidade. Dessa forma, deve-se analisar os dados desta tese tendo-se em mente que eles retratam apenas uma parte da realidade, ou seja, eles não mostram o que realmente ocorre nas situações naturais de fala, dentro da comunidade estudada.

Não resta dúvidas de que um “bom” *corpus*, no caso de estudos deste tipo, deveria ser constituído de gravações da fala natural, ou seja, de falantes em situação normal de interação. Entretanto, adviriam, para o pesquisador, problemas para a coleta de dados, de forma que se satisfizessem as exigências metodológicas da Teoria da Variação, principalmente quando se lida com tempo real.

Em resumo, os dados das duas amostras desta tese indicam que a forma *cê* está implementada na comunidade de Belo Horizonte, sendo usada preferencialmente na função de sujeito, e que apresentou um avanço bastante pequeno nas funções marcadas, que oferecem maior resistência a esse tipo de mudança. Dentre estas, o maior avanço do peso

relativo de *cê* foi quanto à função de objeto de verbo¹⁵⁵, e o menor, na de objeto de preposição.

Quanto aos fatores extralingüísticos, viu-se que *cê* é a forma preferencialmente usada pelas duas classes sociais e pelas mulheres, o que indica que não existem estereótipos nem estigmatização relacionados a ela, ou seja, sua avaliação pela comunidade não é negativa. Por outro lado, os informantes da faixa etária mais alta a desfavorecem, atestando a mudança em curso e também o fato de que os falantes adultos têm uma tendência maior a formalizar sua fala, o mesmo ocorrendo com os informantes homens do *corpus* de 2002.

Até agora se discutiu a variação de uso e da mudança de *cê* – e, automaticamente, de *você*. Entretanto, é importante discutir também a forma *ocê*. Nos resultados das duas amostras, viu-se que o percentual de seu uso caiu um pouco: de 6,6% das ocorrências, em 1982, passou para 3,9%, em 2002, computando-se os dados das cinco faixas etárias. Também pelos dados, vê-se que o maior favorecimento de *ocê* ocorre nas funções de objeto de verbo e de preposição, e isso se deve ao ambiente fonético em que *ocê* se encontra, ao exercer essas duas funções. Na função de sujeito, o uso de *ocê* ocorre maciçamente nas orações subordinadas ou coordenadas, em que o ambiente fonético lhe é igualmente propício.

Assim sendo, cabe perguntar por que *ocê* tem ocorrência restrita como sujeito de oração principal. Pode-se pensar em duas razões, uma de natureza lingüística e outra de natureza extralingüística. Sabe-se que, normalmente, a oração principal é precedida de pausa, mesmo curta. E, em situação de pausa, existe uma tendência natural de redução ou de apagamento da sílaba pretônica, especialmente se ela for constituída por uma única vogal. É este o caso de *ocê*. Entretanto, veja-se que esse fator não impede sua ocorrência, como se verifica nos dados dos dois *corpora* deste estudo.

Já a causa extralingüística não é tão óbvia, mas é possível ser inferida através de outros dados. Embora em nenhum dos estudos citados no primeiro capítulo desta tese se

¹⁵⁵ Verificou-se um aumento do peso relativo de *cê* como objeto de verbo analisando-se os dados no tempo real, conforme apresentados no capítulo 6.

encontrem testes de reação subjetiva, sabe-se que a forma *ocê* é reconhecida como própria das pessoas do interior (cf. Ramos, 2000, p. 183). Esse dado pode ser comprovado, embora indiretamente, pelos resultados de Coelho (1999), que atestam que, na zona urbana de São Francisco, a forma *ocê* tem 16% do total das ocorrências; na zona rural, essa porcentagem sobe para 35%, o que realmente relaciona *ocê* à fala das pessoas do interior. Assim, essa correlação pode gerar uma avaliação negativa pela comunidade de Belo Horizonte e, conseqüentemente, retardar ou impedir o avanço de *ocê* para exercer a função de sujeito da oração principal, tendo em vista que, aí, a forma apareceria destacada.

Concluindo, esta pesquisa tratou da variação do uso e da mudança das formas *você*, *ocê* e *cê*, observando os aspectos lingüísticos – fonético/fonológicos, sintáticos e semânticos – e extralingüísticos relacionados a ela. Entretanto, por falta de tempo e também por fugir aos objetivos mais imediatos deste estudo, alguns pontos importantes deixaram de ser analisados, mas merecem estudos posteriores que os aprofundem.

O primeiro é uma análise mais completa do ambiente fonético precedente às três formas. Viu-se que ele exerce influência sobre as ocorrências de *ocê*, e é importante que esse fato seja mais detalhado. Coelho (1999) e Barbosa (2005) utilizaram uma classificação mais geral desse ambiente onde as três formas se encontram: depois de pausa, de vogal e de consoante. Entretanto, é necessário especificar esses contextos, buscando determinar sua influência quanto ao uso de *você*, *ocê* e *cê*.

Um outro dado que seria interessante ter é a avaliação real da comunidade de Belo Horizonte sobre o uso dessas formas, através de um teste de reação subjetiva. Com isso, poder-se-ia saber com um pouco mais de segurança as razões da baixa ocorrência de *ocê* nos *corpora*, bem como se ter uma idéia da avaliação atual dos belo-horizontinos quanto ao uso de *cê* como objeto de verbo e de preposição, e, conseqüentemente, das chances de essa forma espalhar-se, no momento atual.

Um terceiro estudo bastante proveitoso seria a coleta de dados que permitissem rastrear a história, ou seja, as origens das formas *ocê* e *cê*, através da literatura¹⁵⁶, de gravações antigas de programas de rádio e de TV, de discos e de filmes brasileiros. Esses dados são importantes para se traçar a trajetória do processo que se originou em *Vossa Mercê*.

Por último, seria de extrema valia um estudo que enfocasse a variação de *você*, *ocê* e *cê* do ponto de vista pragmático-discursivo. Se as reduções estão relacionadas à predizibilidade e à automação dos itens lexicais (Bybee, 2002b, p. 09), qual(is) é(são) a(s) restrição(ões) discursiva(s) ao uso de uma forma reduzida como *cê*? E qual(is) é(são) a(s) implicação(ões) do uso de um termo informal, como *cê*, numa situação mais formal, como a de entrevista? Analisar esses aspectos torna-se essencial, quando se tem em mente as diferenças entre a fala gravada e o vernáculo, e se estuda aquela correlacionando-a com este.

Por fim, esta tese termina com um pensamento de Lass (1980), citado por Hopper e Traugott (1993, p. 63), que resume o que se disse neste último capítulo:

“... the phenomena that give rise to language change are so complex that they will perhaps never be understood in enough detail for us to state precisely why a specific change occurred in the past or to predict when one will occur and if it does what it will be.”

¹⁵⁶ Como foi visto, as peças teatrais citadas no Capítulo II não trazem nem as formas *ocê*, nem *cê*. Por outro lado, o uso de *ocê* e *cê* ocorreu, por exemplo, no conto “A terceira margem do rio”, do livro “Primeiras Estórias”, de Guimarães Rosa, escrito em 1962. Assim, pode-se esperar que esse uso seja citado por outros autores que registrem a fala do povo do interior ou das camadas populares das zonas urbanas.

REFERÊNCIAS BIBLIOGRÁFICAS

ALI, Saïd M. *Gramática histórica da língua portuguesa*. 5. ed. São Paulo: Melhoramentos, 1965.

_____. *Investigações filológicas*. 2. ed. Rio de Janeiro: Grifo, 1976.

AMARAL, A. *O dialeto caipira*. São Paulo: Hucitec, 1976.

ALMEIDA, Napoleão M. de. *Gramática metódica da língua portuguesa*. 44. ed. São Paulo: Saraiva, 1999.

ALVES, Nilton A. *As formas você e cê e a indeterminação do sujeito no português brasileiro*. 1998. 93 f. Dissertação (Mestrado em Lingüística) - Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 1998.

ANDRADE, Adriana L. V. S. *A variação de 'você', 'cê' e 'ocê' no português brasileiro falado*. 2004. 136 f. Dissertação (Mestrado em Lingüística) – Faculdade de Letras, Universidade de Brasília, Brasília, 2004.

ANDRÉ, Hildebrando A. *Gramática ilustrada*. 5. ed. São Paulo: Moderna, 1999.

BAILEY, Charles N. *Variation and linguistic theory*. Arlington, Virgínia: Center for Applied Linguistics, 1973.

BAILEY, Guy, WIKLE, Tom, TILLER, Jan e SAND, Lori. The apparent time construct. *Language Variation and Change*, New York, 3, p. 241-264, 1991.

BARBOSA, Liliâne P. *Estatuto da forma cê: clítico ou palavra?* 2005. 137 f. Dissertação (Mestrado em Lingüística) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 2005.

BARLOW, Jessica e GUIERUT, Judith. Optimality theory in phonological acquisition. In: *Journal of Speech, Language and Hearing Research*, v. 42, p. 1482-1498, dec. 1999

BASTO, Cláudio. Formas de tratamento, em português. *Revista Lusitana*, XXIX, p. 182-202, 1931.

BECHARA, Evanildo. *Moderna gramática portuguesa*. 37. ed. Rio de Janeiro: Lucerna, 2003.

BISOL, Leda. Sândi vocálico externo. In: ILARI, Rodolfo (Org.). *Gramática do português falado*. Vol. II: Níveis de Análise Lingüística. 2. ed. Campinas: Ed. UNICAMP, 1993.

_____. Os constituintes prosódicos. In: BISOL, L. *Introdução a estudos de fonologia do português brasileiro*. 3. ed. Porto Alegre: EDIPUCRS, 2001. p. 229-241.

BOBERG, Charles. Real and apparent-time in language change: late adoption of changes in Montreal English. *American Speech*, v. 79, n. 3, 2004

BOT, Kees de; STOESSEL, Saskia. Introduction: language change and social networks. In: *International Journal of Social Language*, Walter de Gruyter, v. 153, p. 1-7, 2002.

BRAGA, Maria Luíza. E aí se passaram 19 anos. In: PAIVA, Maria da Conceição; DUARTE, Maria Eugênia L. (Org.) *Mudança lingüística em tempo real*. Rio de Janeiro: Contra Capa Livraria, 2003. p. 159-174.

BYBEE, Joan. Main clauses are innovative, subordinate clauses are conservative: consequences for the nature of constructions. In: BYBEE, J.; NOONAN, M. (Eds.) *Complex sentences in grammar discourse: essays in honor of Sandra A. Thompson*. Amsterdam: John Benjamins, 2001. p. 1-17.

_____. Phonological evidence for exemplar storage of multiword sequences. *SSLA*, Cambridge, n. 24, p. 215-221, 2002.

_____. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language variation and change*, v. 14, p. 261-290, 2002b.

_____. Mechanisms of change in grammaticization: the role of frequency. In: RICHARD, Janda; BRIAN, Joseph (Ed.) *Handbook of historical linguistics*. Oxford: Blackwell, 2003. p. 602-23.

CALLOU, Dinah & LEITE, Yonne. *Iniciação à fonética e à fonologia do português*. 8. ed. RJ: Jorge Zahar Ed., 2001.

CALLOU, Dinah et al. Topicalização e deslocamento à esquerda: sintaxe e prosódia. In: CASTILHO, Ataliba (Org.) *Gramática do português falado*. 2. ed. Campinas, SP: Ed. UNICAMP, 1996. v. 3, p. 315-357.

CÂMARA Jr, Joaquim M. *História e estrutura da língua portuguesa*. 2. ed. Rio de Janeiro: Padrão, 1976.

CÂMARA Jr., Joaquim M. *Estrutura da língua portuguesa*. 14. ed. Petrópolis: Vozes, 1884

CASTILHO, Ataliba (Org.) *Gramática do português falado*. 2. ed., v. 3. Campinas, SP: Unicamp, 1996.

CASTILHO, Ataliba T. A gramaticalização. In: *Estudos Lingüísticos e Literários*. Salvador, BA: UFBA, n. 19, p. 11-24, mar. 1997.

CEGALLA, Domingos P. *Novíssima gramática da língua portuguesa*. 43. ed. São Paulo: Cia. Editora Nacional, 2000.

- CHAMBERS, J.K. *Sociolinguistic theory*. Oxford, Cambridge: Blackwell, 1995.
- CINTRA, Luís F. Lindley. *Sobre Formas de Tratamento na língua portuguesa: ensaios*. Lisboa: Horizonte, 1972.
- CIRÍACO, Larissa, VITRAL, Lorenzo e REIS, César. Intensidade e duração de formas reduzidas no português brasileiro. *Revista de Estudos da Linguagem*, v. 12, n.2, p. 143-147, 2004.
- COELHO, Maria do Socorro V. *Uma abordagem variacionista do uso de formas de tratamento no norte de Minas*. 1999. 85 f. Dissertação (Mestrado em Linguística) - Faculdade de Letras, Belo Horizonte, 1999.
- COLLISCHONN, Gisela. A sílaba em português. In: BISOL, L. *Introdução a estudos de fonologia do português brasileiro*. 3.ed. Porto Alegre: EDIPUCRS, 2001. p. 91-123.
- _____. O acento em português. In: BISOL, L. *Introdução a estudos de fonologia do português brasileiro*. 3.ed. Porto Alegre: EDIPUCRS, 2001. p. 125-158.
- COOK, Manuela. Uma teoria de interpretação das formas de tratamento na língua portuguesa. *Hispania*, n. 80, p. 451-464, set. 1997.
- CRISTÓFARO-SILVA, Thaís. *Fonética e fonologia do português: roteiro de estudos e guia de exercícios*. SP: Contexto, 1999.
- _____. Branching onsets in Brazilian Portuguese. *Revista de Estudos Lingüísticos*. Belo Horizonte: UFMG, v. 10, n.1, p. 91-107, jan./jun. 2002
- CRYSTAL, David. *Dicionário de lingüística e fonética*. Tradução Maria Carmelita Pádua Dias. Rio de Janeiro: Jorge Zahar Editor, 2000.
- CUNHA, Celso F. da. *Gramática da língua portuguesa*. 11.ed. Rio de Janeiro: FAE, 1985.
- CUNHA, Celso; CINTRA, Luís F. Lindley. *Nova gramática do português contemporâneo*. 3.ed. Rio de Janeiro: Nova Fronteira, 2001.
- CYRINO, Sônia M. L.; BRITO, Onilda R. M. de. *Perda (do uso do tu/te) e aquisição (de você/te)*. Marília, SP: Fundação Eurípides, [199_]
- HORA, Demerval da. Teoria da variação: uma retrospectiva. In: _____ (Org.). *Diversidade lingüística no Brasil*. João Pessoa, PB: Idéia, 1997.
- PAULA, Graciane de. *O clítico 'te' no dialeto mineiro: um fenômeno de manutenção*. Manuscrito, 2004.

DUARTE, M. Eugênia L. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. In: ROBERTS, Ian; KATO, Mary A. *Português brasileiro: uma viagem diacrônica*. 2. ed., Campinas, SP: Ed. da UNICAMP, 1996.

_____. A evolução na representação do sujeito pronominal em dois tempos. In: PAIVA, Maria da Conceição; DUARTE, Maria Eugênia L. (Org.) *Mudança lingüística em tempo real*. Rio de Janeiro: Contra Capa Livraria, 2003, p. 115-128.

DOMINGOS, Tânia R. E. *Pronomes de tratamento do português do século XVI: uma gramática de uso*. São Paulo: Annablume; Rondônia: Unir, 2001.

ELIA, Sílvio. *El portugués en Brasil: historia cultural*. Madrid: Editorial Mapfre, 1992.

FARACO, Carlos Alberto. O tratamento você em português; uma abordagem histórica. In: *Fragmenta*, Curitiba: Ed. da UFPR, n. 13, p. 51-82, 1996.

_____. *Lingüística histórica*. São Paulo: Ática, 1998.

FARACO & MOURA. *Gramática*. 10. ed. São Paulo: Ática, 1997.

FIDELHOTZ, James. Word frequency and vowel reduction in English. *Chicago Linguistic Society*, v. 11, p. 200-213.

GORDON, Elizabeth. Sex, speech, and stereotypes: why women use prestige speech forms more than men. *Language in society*, Cambridge, n.26, p.47-63, 1997.

GUY, Gregory. Language and social class. In: NEWMAYER, Frederik (Ed.). *Linguistics: the Cambridge survey*. V. 4: Language: the socio-cultural context. Cambridge: Cambridge University Press, 1988. p. 37-63.

HERNANDORENA, Carmen L. Introdução à teoria fonológica. IN: BISOL, L. *Introdução a estudos de fonologia do português brasileiro*. 3.ed. Porto Alegre: EDIPUCRS, 2001. p. 11-89.

HINSKENS, F. ;HOUT, R.v.; WETZELS, Leo. Um balanço de dados e teoria no estudo da variação e da mudança fonológica. *Letras de Hoje*. Porto Alegre, v. 35, n.1, p. 7-46, mar. 2000.

HUBACK, Ana Paula. *Cancelamento do (r) final em nominais na cidade de Belo Horizonte: uma abordagem difusionista*. 2003 – 130 f. Dissertação (Mestrado em Lingüística) - Belo Horizonte, Universidade Federal de Minas Gerais, 2003.

KROCH, Anthony. Mudança sintática. IN: “Syntactic Change”, BALTIN, Mark and COLLINS, Chris. (Ed.) *The handbook of contemporary syntactic theory*. Blackwell, 2001. Tradução de Silvia Regina Cavalcante, 2003.

LABOV, William. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press, 1972.

_____. *Language in the inner city; studies in the Black English Vernacular*. Philadelphia: University of Pennsylvania Press, 1972a.

_____. *Principles of linguistic change; internal factors*. Cambridge: Blackwell, 1994.

_____. *Principles of linguistic change; social factors*. Cambridge: Blackwell, 2001.

_____. Driving forces in linguistic change. International Conference on Korean Linguistics, Seoul National University, August 2, 2002.

LEE, Seung-Hwa. Acento secundário do PB. IN: *Letras de Hoje*. Porto Alegre, v. 37, n.1, p. 149-162, mar. 2002.

LIBERATO, Yara G. *Sobre a oposição dado/novo*. 1980. Dissertação (Mestrado em Lingüística) - Belo Horizonte, Universidade Federal de Minas Gerais, 1980.

LOPES, Célia R. S. *O quadro dos pronomes pessoais*. Rio de Janeiro, manuscrito. 2003.

_____. Vossa Mercê > você e Vuestra Merced > Usted: o percurso evolutivo ibérico. *Lingüística*. ALFAL, v. 14, 2003a.

LOPES, Célia R. S. e DUARTE, M. Eugênia L. De *Vossa Mercê* a *você*: análise da pronominalização de nominais em peças brasileiras e portuguesas setecentistas e oitocentistas. In: BRANDÃO, Silvia F. & MOTA, M. Antónia (Org.). *Análise contrastiva de variedades do Português: primeiros estudos*. RJ: In-fólio, 2003, p. 61-76.

_____. *Notícias sobre o tratamento em cartas escritas no Brasil dos séculos XVIII e XIX*. Comunicação apresentada no V Seminário do PHPB. Ouro Preto, UFOP/UFMG, 2004.

LUFT, Celso P. Tratamento depreciativo. *Revista Brasileira de Filologia*. Rio de Janeiro: Livraria Acadêmica, v. 3, T. II, dez. 1957.

LUZ, Marilina dos Santos. Fórmulas de tratamento do português. *Revista Portuguesa de Filologia*. Coimbra, v. II. T. I, II, p. 256-363, 1956.

MACAULAY, Ronald. Social class and language in Glasgow. *Language in Society*, v. 5, p. 173-188, 1976.

MAGALHÃES, José Olímpio. *Prosódia e significado em sentenças declarativas do português (e de outras línguas)*. Projeto de Estágio Pós-doutoral. Belo Horizonte: FALE/UFMG, 2005.

MAIA, Marcus; FINGER, Ingrid. *Processamento da linguagem*. Porto Alegre: Educarte, 2005.

MARCHELLO-NIZIA, Christiane. Changes in the structure of grammatical systems: the evolution of French. *Court of the University of St. Andrews*, v. xxxix, n. 4, 2003.

MENÓN, Odete P. S. O sistema pronominal do português do Brasil. *Letras*, Curitiba: Ed. da UFPR, n. 44, p. 91-106, 1995.

_____. Pronome de segunda pessoa no sul do Brasil: tu/você/o senhor em Vinhas da Ira. In: *Letras de Hoje*. Porto Alegre/RS, v. 35, n.1, p. 121-164, mar. 2000.

MESQUITA, Roberto M. *Gramática da língua portuguesa*. 8. ed. São Paulo: Saraiva, 1999.

MILROY, James. *Linguistic variation & change*. Oxford: Blackwell, 1992.

MILROY, Lesley. *Language and social network*. 2nd. Ed. Oxford (UK): Blackwell, 1987.

MILROY, L. & MILROY, J. Social network and social class: toward an integrated sociolinguistic model. *Language in Society*, Cambridge University Press, v. 21, p. 1 – 26, 1992.

MOLLICA, M. Cecília. Relevância das variáveis não lingüísticas. IN: MOLLICA, M.C.; BRAGA, M. L. (Org.). *Introdução à sociolingüística: o tratamento da variação*. São Paulo: Contexto, 2003. p. 27-32,

MORAIS, João A. e LEITE, Yone. Ritmo e velocidade da fala na estratégia do discurso: uma proposta de trabalho. IN: ILARI, Rodolfo (Org.). *Gramática do português falado*. V. 2: Níveis de Análise Lingüística. 2. ed. Campinas: Ed. UNICAMP, 1993.

NARO, Anthony J. O dinamismo das línguas. In: MOLLICA, M. Cecília & BRAGA, M. Luíza (Org.) *Introdução à Sociolingüística: o tratamento da variação*. São Paulo: Contexto, 2003.

NARO, Anthony & SCHERRE, M. Marta P. Variação e mudança lingüística: fluxos e contrafluxos na comunidade de fala. *Cadernos de Estudos Lingüísticos*. Campinas, v. 20, p. 9-16, jan./jun. 1991.

_____. Estabilidade e mudança lingüística em tempo real: a concordância de número. In: PAIVA, Maria da Conceição; DUARTE, Maria Eugênia L. (Org.) *Mudança lingüística em tempo real*. Rio de Janeiro: Contra Capa Livraria, 2003, p. 47-62.

NASCENTES, Antenor. O tratamento de “você” no Brasil. In: *Letras*. Curitiba/PR: Ed. UFPR, v.6, n.05, p. 114-122, 1956.

NEVALAINEN, Terttu. Mobility, social network and language change in Early Modern England. In: *European Journal of English Studies*. Swets & Zeitlinger, v. 4, n. 3, p. 253-264, 2000.

OLIVEIRA, Gilván Muller de. Última fronteira: história da língua portuguesa no Brasil meridional (1680-1830). In: GROBE, Sybille; ZIMMERMANN, Klaus (Ed.). *O português brasileiro: pesquisas e projetos*. Frankfurt am Main: TFM, 2000, p. 345-376.

OLIVEIRA, Marilza & RAMOS, Jânia. *O estatuto de 'você' no preenchimento do sujeito*. Comunicação apresentada no Encontro da Alfal, Costa Rica, 2002.

OLIVEIRA, Sândi M. de. Contribuição para um estudo comparativo de formas de tratamento em Espanha e Portugal. *Actas del Congreso Internacional Luso-Español de Lengua y Cultura en la Frontera*. Cáceres: Universidad de Extremadura, 1996, p. 123-139.

OMENA, Nelize Pires de. A referência à primeira pessoa do plural: variação ou mudança? In: PAIVA, Maria da Conceição; DUARTE, Maria Eugênia L. (Org.) *Mudança lingüística em tempo real*. Rio de Janeiro: Contra Capa Livraria, 2003, p. 63-80.

PAIVA, Maria da Conceição. A variável gênero/sexo. In: MOLLICA, M. Cecília; BRAGA, M. Luíza (Org.) *Introdução à Sociolingüística; o tratamento da variação*. São Paulo: Contexto, 2003.

PAREDES SILVA, Vera L. Variação e funcionalidade no uso de pronomes de segunda pessoa do singular no português carioca. *Revista de Estudos da Linguagem*. Belo Horizonte: UFMG, v. 7, n. 2, p. 121-138, 1998.

_____. Variação e funcionalidade. In: MOLLICA, M. Cecília & BRAGA, M. Luíza (org.) *Introdução à Sociolingüística: o tratamento da variação*. São Paulo: Contexto, 2003, p. 67 - 72.

PASCHOALIN & SPADOTO. *Gramática; teoria e exercícios*. São Paulo: FTD, 1989.

PASQUALE & ULISSES. *Gramática da língua portuguesa*. São Paulo: Scipione, 2003.

PHILLIPS, Betty. Word frequency and the actuation of sound change. *Language*, v. 60, n. 2, p. 320-342, 1984.

_____. Lexical diffusion is *not* lexical analogy. *Word*, v. 49, n. 3, dec. 1998. p. 369-381.

_____. The mental lexicon: evidence from Lexical Diffusion. *Brain and Language*, v. 68, p. 104-109, 1999.

_____. Lexical Diffusion and competing analyses of sound change. Paper presented at SHEL-1 (Studies in the History of the English Language), Los Angeles, UCLA, may, 2000.

_____. Lexical diffusion, lexical frequency and lexical analysis. In: BYBEE, J.; HOPPER, P. *Frequency and the emergence of linguistic structure*. Amsterdam: Benjamins, 2001. p.123-136.

PIAGET, Jean. O desenvolvimento mental da criança. In: PIAGET, Jean. *Seis estudos de psicologia*. 24. ed. Rio de Janeiro: Forense Universitária, 2003.

PONTES, Eunice. *O tópicio no português do Brasil*. Campinas: Pontes, 1976.

PREFEITURA MUNICIPAL DE BELO HORIZONTE. *Caracterização socioeconômica da área piloto do Programa BH Cidadania: Vila Santa Rosa*. Belo Horizonte: PMBH, 2001.

RAMOS, Jânia. O uso das formas você, ocê e cê no dialeto mineiro. In: DA HORA, Demerval (Org.). *Diversidade lingüística no Brasil*. João Pessoa, PB: Idéia, 1997, p. 43-60.

_____. O surgimento de um novo clítico no português brasileiro: análise quantitativa e qualitativa da forma *cê*. In: GÄRTNER, Eberhard, HUNDT, Christine; SCHÖNBERGER, Axel (Ed.). *Estudos de sociolingüística brasileira e portuguesa*. Frankfurt am Main: TFM, 2000. p. 181-189.

RAMOS, Jânia & OLIVEIRA, Marilza. *Pronomes de segunda pessoa: uma abordagem diacrônica*. Comunicação apresentada na Reunião da ANPOLL, Gramado, RS, 2002.

ROCHA LIMA. *Gramática normativa da língua portuguesa*. 38.ed. Rio de Janeiro: José Olympio, 2000.

RONCARATI, Cláudia. Os mostrativos na variedade carioca falada. In: PAIVA, Maria da Conceição; DUARTE, Maria Eugênia L. (Org.). *Mudança lingüística em tempo real*. Rio de Janeiro: Contra Capa Livraria, 2003. p. 139-158

RUSSO, Ieda; BEHLAU, Mara. *Percepção da fala: análise acústica do português brasileiro*. São Paulo: Ed. Lovise Científica, 1993.

SACCONI, Luiz A. *Nossa gramática: teoria e prática*. 25. ed. São Paulo: Atual, 1999.

SALLES, Miguel. *Pronomes de tratamento do interlocutor no português brasileiro: um estudo de pragmática histórica*. 2001. - 246 f. Tese (Doutorado em Lingüística) – Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo, São Paulo, 2001.

SANTOS, Armindo dos. O tratamento por “você” e por “tu” nas relações de parentesco: o exemplo da Beira Baixa. *Jornal de Letras, Artes e Idéias*. Lisboa: Publicações Projornal, ano V, p. 168, 1985.

SCHANE, Sanford. *Generative phonology*. Englewood Cliffs, New Jersey: Prentice-Hall, 1973.

SUNDGREN, Eva. Men and women in language change: a Swedish case study. *NORA*, v. 9, n. 2, 2001, p. 113-122.

TARALLO, Fernando. *A pesquisa sociolingüística*. São Paulo: Ática, 1994.

_____. Diagnosticando uma gramática brasileira: o português d'aquém e d'além mar ao final do século XIX. In: ROBERTS, I.; KATO, M. A. (Ed.). *Português brasileiro: uma viagem diacrônica*. 2. ed. Campinas: Editora da UNICAMP, 1996. p. 69-105.

TERRA, Ernani. *Curso prático de gramática*. 3. ed. São Paulo: Scipione, 2000.

TEYSSIER, Paul. *História da língua portuguesa*. Tradução Celso Cunha. 2. ed. São Paulo: Martins Fontes, 2001.

TRUDGILL, Peter. Norwich revisited: recent linguistic changes in the English urban dialect. *English World-Wide*. Amsterdam: John Benjamins, v. 9, n. 1, p. 33-49, 1988.

VIEGAS, Maria do Carmo. Alçamento de vogais médias pretônicas: uma abordagem sociolingüística. 1987. Dissertação (Mestrado em Lingüística) – Faculdade de Letras, Universidade Federal de Minas Gerais, Belo Horizonte, 1987.

VITRAL, Lorenzo. A forma CÊ e a noção de gramaticalização. *Revista de Estudos da Linguagem*. Belo Horizonte, FALE, jan./jun., 1996

WEINREICH, Uriel; LABOV, William & HERZOG, Marvin. *Empirical foundation for a theory of language change*. Austin: University of Texas Press, 1968.

WEISS, Curtis E., GORDON, Mary E. e LILLYWHITE, Herold S. Phonologic theory, development, and disorders. In: _____. *Clinical management of articulatory and phonologic disorders*. 2nd ed. Baltimore: Williams & Wilkins, 1987. Chapter 4, p. 63-83.

ANEXO 1

**ROTEIRO DE PERGUNTAS PARA O TESTE
COM CRIANÇAS E ADOLESCENTES - 2002****Auto-retrato**

1. Nome completo
2. Apelido(s)
3. Aniversário
4. Família
5. Suas predileções (respostas rápidas, com justificativas curtas)
 - a) cor
 - b) flor
 - c) fruta
 - d) cidade
 - e) animal
 - f) estação do ano
 - g) mês do ano
 - h) dia da semana
 - i) parte do dia
 - j) matéria
 - k) cantor(a)
 - l) música
 - m) instrumento
 - n) ator/atriz
 - o) filme
 - p) escritor/escritora
 - q) livro
 - r) esporte
 - s) passatempo

6. Do que você mais gosta em você? Por quê?
7. Seu melhores amigos
8. Uma palavra para eles
9. Os melhores lugares:
 - a) para se divertir
 - b) para descansar
 - c) para estudar
 - d) para refletir

10. Adoro
11. Detesto
12. Suas mais preciosas qualidades
13. Seus piores defeitos
14. Qualidades que admira nas pessoas
15. Defeitos com os quais não consegue conviver
16. Planos para o futuro
17. Profissão desejada
18. Como você desejaria ser lembrado?

Obs.: As perguntas feitas aos entrevistados, com ou sem modificações, são as que aparecem sublinhadas.

ANEXO 2

CASOS DE OBJETO DE PREPOSIÇÃO – 1982

- (E 13) Você não ganha nem PRA VOCÊ, vai querer resolver os problemas dos outros!
- (E 17) Eles escrevem assim cartinha PRA VOCÊ cada coisa mais bonitinha.
- (E 64) Se o cara tá disposto a casar COM VOCÊ é porque...
- (E 66) Lógico que se ele gosta DE VOCÊ ele nunca vai abusar.
- (E 102) E se ela quiser estar num boteco e não quiser estar lá na sua casa COM VOCÊ?
- (E 161) Eu vou lá olhar P'OCÊ.
- (E 240) E se cê não dá valor PR'OCÊ, cê não tem condição de prosseguir.
- (E 269) Aí eu posso tentar arrumar isso P'OCÊ.
- (E 271) Eu vou olhar isso P'OCÊ.
- (E 285) É mais suficiente PR'OCÊ.
- (E 346) Você pensa EM VOCE.
- (E 354) Cê faz a pergunta PRA VOCÊ mesmo
- (E 381) Eu mando isso PR'OCÊ.
- (E 382) Se quiser ficar com esse PR'OCÊ pode levar.
- (E 410) Na casa dessa moça que estava COM VOCÊ, já roubaram na casa lá.
- (E 456) Se eu fiz eu chego e falo COM VOCÊ: ...
- (E 460) ...acabou o filme PRA VOCÊ ...
- (E 485) ...pra não dar problema PRA VOCÊ.
- (E 518) ... não estava fazendo entrevista COM VOCÊ aqui não.
- (E 519) Não estava fazendo entrevista COM VOCÊ.
- (E 529) ... só pra ficar olhando PR'OCÊ.
- (E 532) ... que tá doida pra encontrar COM VOCÊ.
- (E 544) Se eu saio COM VOCÊ de braço dado na rua ...
- (E 549) Eu falei C'OCÊ que num dava.
- (E 563) Pô, tô a fim de transar COM VOCÊ.

CASOS DE OBJETO DE PREPOSIÇÃO – 2002

- (E 35) “Olha, vai faltar isso PRA VOCÊS daqui dez anos”
- (E 144) Isso não vai dar certo PRA VOCÊ.
- (E 158) ... então PRA VOCÊ vai ser um orgulho.
- (E 178) A pessoa fala assim COM VOCÊ: ...
- (E 182) ... guardar aquilo ali P'OCÊ dentro de si.
- (E 186) Ela nunca vai pisar na bola COM VOCÊ.
- (E 191) ... e leva COM VOCÊ, ...
- (E 200) Vou mandar PR'OCÊ.
- (E 231) Não vou falar nada C'OCÊ não.
- (E 233) Depois eu conversei C'OCÊ, tá?
- (E 262) Mais bonito depois DE VOCÊ, é claro.
- (E 266) ... que ela gostasse DE VOCÊ, não?
- (E 270) Eu vou abrir o portão PR'OCÊ.
- (E 310) Vou pegar mais um pouquinho P'OCÊ.
- (E 439) ... que eu vou lá COM VOCÊ.
- (E 454) Tá muito difícil PR'OCÊ, né?
- (E 541) O que que tá acontecendo COM VOCÊ?
- (E 683) Eu vou COM CÊ.
- (E 704) ... baixar um trem N'OCÊ aí, ...
- (E 706) ... que tem um troço aí C'OCÊ.
- (E 710) Se chega perto DE VOCÊ é porque ...
- (E 717) É, eu também concordo C'OCÊ.
- (E 724) Cê tá falando que eu menti PR'OCÊ?
- (E 734) Xo [Deixe eu] perguntar P'CÊ.
- (E 738) De repente eu vou fazer C'OCÊ lá mesmo (...)
- (E 743) Não sei se assim aconteceu COM VOCÊ.
- (E 759) A minha carteira não foi assinada nem POR VOCÊ nem pelo seu pai.

- (E 764) Eles ficam até um pouco dependentes DE VOCÊ.
- (E 848) Isso aqui ia ser ideal PRA VOCÊ.
- (E 860) Se eu pego uma arma e atiro EM VOCÊ ...
- (E 861) ... ou se eu atiro EM VOCÊ de costa, ...
- (E 864) ... como é que eu tava me defendendo DE VOCÊ?
- (E 910) ... pra ter conhecimento A VOCÊ, eu teria que ...
- (E 959) Sair C'OCÊS dá peso.
- (E 961) ... deixa eu ir C'OCÊ, pr'eu ficar lá na biblioteca.
- (E 967) Não falei C'OCÊ que nós ia ser interrompida?
- (E 1137) ... Ele chega PRA VOCÊ com um bilhete ...
- (E 1197) "Eu não vou deixar nada PRA VOCÊ."
- (E 1285) ...ficam com raiva DE VOCÊ.
- (E 1291) Aquela pessoa que chega perto DE VOCÊ e...
- (E 1399) ... a pegar PR'OCÊ.
- (E 1410) ... me vingar DE VOCÊ se cê ligar.
- (E 1435) ... eu tô a fim de abrir COM VOCÊ.
- (E 1448) ... vou conversar com ela e eu ligo PR'OCÊ
- (E 1449) No final de semana não dá PRA VOCÊ não, né?

CASOS DE OBJETO DE VERBO - 1982

- (E 18) "Tia, gostei muito de ter VOCÊ de novo."
 (E 117) ... cê cê confunde muito OCÊ com a outra pessoa e...
 (E 548) Eu tô querendo VOCÊ mesmo.

CASOS DE OBJETO DE VERBO - 2002

- (E 97) ... e nunca larguei VOCÊS por isso.
 (E 98) Nunca larguei OCÊ.
 (E 117) Eu vou matar CÊ, fdp!
 (E 142) Dinheiro sujo te leva CÊ à morte (...)
 (E 167) ... mas às vezes te joga OCÊ pro buraco.
 (E 172) ... ela acaba te jogando OCÊ pro buraco.
 (E 188) ... te joga OCÊ pra trás.
 (E 437) Eles podem te... levar VOCÊS lá na zona rural.
 (E 614) Cê vai assistir o futebol cê...nego te arrasta CÊ lá dentro.
 (E 1288) Eu passo e nem soudo [saúdo] VOCÊ..
 (E 1418) "Eu vou matar sua mãe e VOCÊ".