UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, LETRAS E CIÊNCIAS HUMANAS
DEPARTAMENTO DE LINGUÍSTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM SEMIÓTICA E
LINGUÍSTICA GERAL

MARÍLIA SILVA VIEIRA

AÍ, DAÍ E ENTÃO EM CAMPO GRANDE E SÃO PAULO

Análise sociofuncionalista no domínio da causalidade

VERSÃO CORRIGIDA

São Paulo

2016

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, LETRAS E CIÊNCIAS HUMANAS
DEPARTAMENTO DE LINGUÍSTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM SEMIÓTICA E
LINGUÍSTICA GERAL

***AÍ, DAÍ E ENTÃO* EM CAMPO GRANDE E SÃO PAULO**

Análise sociofuncionalista no domínio da causalidade

Marília Silva Vieira

Tese apresentada ao Programa de Pós-graduação em Semiótica e Linguística Geral do Departamento de Linguística da Faculdade de Filosofia, Letras e Ciências Humanas da Universidade de São Paulo, para obtenção do título de Doutor em Letras.

Orientador: Prof. Dr. Ronald Beline Mendes

São Paulo

2016

Ao Professor e amigo Dercir Pedro de Oliveira (*in
memoriam*), idealizador dos estudos da língua falada
em Mato Grosso do Sul.
Todo o meu carinho!

Agradecimentos

- **Primeiramente:**

Aos meus pais, pelo apoio, por se fazerem presentes, mesmo quando muito distantes, pelas palavras de ânimo e por acreditarem em mim, quase sempre mais do que eu mesma.

À *Badinha*, minha tia, madrinha e segunda mãe, que faz o impossível para colaborar com a realização de cada etapa da minha vida, simplesmente porque se alegra com a minha felicidade. Sem o seu entusiasmo, teria sido muito mais difícil chegar até aqui.

Aos meus tios Cleusa e José, pela delicadeza e preciosidade de nossos encontros, que têm o dom de refinar minha alma.

Ao Saulo, meu namorado, que me transmite o equilíbrio, a paciência e a segurança que, não raro, me faltam e que foram fundamentais na reta final do Doutorado.

- **Aos professores:**

Ronald Mendes, pela orientação, desde o fim do Mestrado, por ter lapidado meu olhar para a Sociolinguística e por ter me concedido a oportunidade de viver uma experiência singular, o Grupo de Estudos e Pesquisa em Sociolinguística da USP (o GESOL).

Sebastião Gonçalves, minha admiração e minha gratidão pelas suas acuradas contribuições na ocasião do exame de qualificação – contribuições essas que foram essenciais para o bom andamento desta tese.

Esmeralda Negrão, também por suas sugestões quando do exame de qualificação e por ter sido tão marcante em minha passagem pela USP.

Olga Coelho, que acompanhou meu amadurecimento na pós-graduação e muito me inspira como professora e pesquisadora.

Norma Discini, com quem tive o privilégio de conviver durante a monitoria, no PAE (Programa de Aperfeiçoamento do Ensino). Seu exemplo, sua

sabedoria humanista e a epifania de suas aulas serão sempre um legado vivo, não só em meu exercício como docente, mas também em meu coração.

- **Aos amigos campo-grandenses, que têm me acompanhado desde o início da graduação, e com quem tudo faz mais sentido:**

Flávia de Zayas e sua mãe, Soyla, pela amizade duradoura, sincera e divertida. Irmã e mãe que tenho em Campo Grande. Agradeço especialmente pela valorosa ajuda que me prestaram com a gravação das entrevistas;

Juvenal Cezarino, pelo carinho, pelas conversas linguísticas e literárias, pela amizade que só floresce há uma década;

Leleco, parceiro de Linguística, reclamações e idealizações acadêmicas. Meu agradecimento por todo o seu empenho na busca de informantes;

Rodrigo, *Rô*, com quem divido dúvidas existenciais e que esteve presente em alguns dos momentos mais importantes da minha vida. Obrigada também por me indicar informantes para a composição do corpus!

- **Na USP:**

A todos os membros do GESOL, a quem agradeço, em especial, à *Livia Oushiro*, pela acolhida, pela paciência, por ter sido não só colega, mas também professora. À *Dayane Almeida*, por levar alegria aos encontros do grupo; à *Larissa Soriano*, pela postura engajada, que ensina a ver sempre além; à *Júlia França*, em que me reconheço, na simplicidade interiorana; à *Amanda Santana*, dona de um silêncio sensível, e ao *Fernando Gomes*, pelos momentos memoráveis que compartilhamos. Meu muito obrigada aos colegas pela troca acadêmica, que fez de mim uma pesquisadora mais madura, e, da USP, uma espécie de casa.

Ao Wendel Santos, a quem conheci em meio às angústias de quem sabe aonde quer chegar, mas que descobre o caminho ao caminhar. Obrigada por me proporcionar uma convivência calorosa e regada a muito sotaque ludovicense.

Mariane Bieler: em nossas diferenças, quanta semelhança! Minha gratidão cheia de ternura por me arrancar da melancolia no meio da semana (e sempre que necessário), pela companhia nas viagens a congressos e andanças em Sampa.

Minha gratidão ao Ricardo Marquezin, amigo que ganhei na convivência com a Mari.

- **Aos amigos paulista(no)s:**

Débora Guisso, por me mostrar a explicação filosófico-psicanalítico-política dos fatos, e por ter me ouvido sempre e com tamanha “alteridade”. Seu apoio foi fundamental para o andamento deste trabalho!

Joaquim Chaves, paulistano cujo coração não cabe em sua cidade natal. Minha gratidão pela presença acalentadora e pelo humor contagiante!

- **Na Universidade Estadual de Goiás (UEG):**

Aos amigos que me proporcionam um ambiente de trabalho leve e agradável: Eduardo Batista (pelo companheirismo, no curso de Letras e no dia a dia, pela parceria gastronômica e *nerd*, pelas dicas científicas e de formatação de texto) e Jean Vieira (pela motivação turístico-acadêmica, pelas sugestões cartográficas e pelas conversas, que não têm hora para acabar).

Aos alunos do curso de Letras, para quem tanto falei sobre *aí*, *daí* e *então* durante aulas e eventos científicos. Alguns, inclusive, pegaram gosto por essas “palavrinhas”.

- **Aos informantes:**

Campo-grandenses, que, na amada e inesquecível cidade morena, proporcionaram-me não só ótimas oportunidades profissionais e pessoais, mas também muito de sua simpatia e de seu (quase) vernáculo durante as gravações.

Paulistanos, de quem me sinto conterrânea, tanto pela influência linguística, absorvida em três anos de encantamento pela terra da garoa, quanto pela estima e identificação que tenho para com a capital paulista.

“Eu tinha vontade de fazer como dois homens que vi sentados na terra escovando osso. No começo achei que aqueles homens não batiam bem. Porque ficavam ali sentados na terra o dia inteiro escovando osso. Depois aprendi que aqueles homens eram arqueólogos. E que eles faziam o serviço de escovar osso por amor. Eles queriam encontrar nos ossos vestígios de antigas civilizações que estariam enterrados por séculos naquele chão. Logo pensei de escovar palavras. Porque eu havia lido em algum lugar que as palavras eram conchas de clamores antigos. Eu queria ir atrás dos clamores antigos que estariam guardados dentro das palavras. Eu já sabia também que as palavras possuem no corpo muitas oralidades remontadas e muitas significâncias remontadas. Eu queria então escovar as palavras para escutar o primeiro esgar de cada uma. Para escutar os primeiros sons, mesmo que ainda bígrafos ”.

Manoel de Barros

*Escova*¹

¹ BARROS, Manoel de. *Escova*. In: *Memórias inventadas: Minha infância*. São Paulo: Planeta, 2003, p. 1.

Resumo

Fundamentada na interface entre a Sociolinguística (Labov, 1978; 2008 [1972]; 2001; Lavandera, 1978) e as teorias de gramaticalização, (Halliday, 1985; Hopper & Traugott, 1991; Givón, 1995), esta tese investiga o uso variável dos itens *aí*, *daí* e *então* em contextos de causalidade (no domínio referencial - Sweetser, 1991), o que posiciona este trabalho no quadro de estudos sociofuncionalistas (Naro & Braga, 2001). Os dados foram extraídos de 48 entrevistas sociolinguísticas com nativos de Campo Grande (capital do Mato Grosso do Sul) e 48 com paulistanos (nativos da cidade de São Paulo e parte do corpus do *Projeto SP2010: Construção de uma amostra de fala paulistana*). Tais amostras foram estratificadas de acordo com as variáveis *sexo/gênero*, *faixa etária* (20-35; 35-59; 60 anos ou mais) e *nível de escolaridade*. Depois de descrever a trajetória de gramaticalização das três formas, esta tese examina os grupos de fatores que se correlacionam com seus empregos na expressão da causalidade referencial. As variáveis linguísticas são: *sequência discursiva* (narrativa, descritiva, expositiva e argumentativa), *status da informação* (velha, inferível, nova), *gradiência do tópico discursivo* (intratópico, intertópico), *aspecto lexical* (duração, telicidade, dinamismo) e *aspecto gramatical* (perfectivo, imperfectivo, resultativo e indeterminado). As análises revelam que *então* é a forma preferida em ambas as cidades (embora seja ligeiramente mais frequente em São Paulo, esta não é uma diferença estatisticamente significativa). Esta forma também é favorecida entre os falantes de nível universitário, em ambas as cidades, o que sugere que é possivelmente avaliada como a mais prestigiosa entre as três (seguindo a hipótese inicial, com base em gramáticas tradicionais, como Cunha & Cintra, 2001). No entanto, não existe uma correlação entre *sexo/gênero* e esta variável (ao contrário da tendência esperada de que as mulheres preferissem *então*, relativamente aos homens). Este fato permite um debate em torno do conceito de prestígio, quando se trata de variáveis discursivas. Além disso, não foram selecionadas exatamente as mesmas variáveis na análise de cada amostra: *nível de escolaridade* e *aspecto gramatical* são estatisticamente significativas para ambas as capitais; *faixa etária* foi selecionada para São Paulo (mas não para Campo Grande), enquanto *sequência discursiva* foi selecionada para Campo Grande (mas não para São Paulo). Na fala paulistana, *então* é a forma desfavorecida entre os informantes da segunda faixa etária, o que pode ser interpretado como um indício de uma mudança no estatuto social de *aí* e *daí*. No geral, os resultados sugerem que, embora *então* seja a forma preferida, a variável em foco é condicionada de modo diferente nas duas comunidades.

Palavras-chave: *Aí; Daí; Então*; Campo Grande; São Paulo; Sociofuncionalismo; Gramaticalização.

Abstract

Grounded at the interface between Sociolinguistics (Labov, 1978; 2008 [1972]; 2001; Lavandera, 1978) and grammaticalization theories (Halliday, 1985; Hopper & Traugott, 1991; Givón, 1995), this dissertation investigates the interchangeable use of three discourse items – *aí*, *daí* and *então* ‘so, then’ – in contexts of sentence combining to express causality (in the referential domain – Sweetser, 1991), thus placing this research within the framework of sociofunctionalism (Naro & Braga, 2001). The data were extracted from 48 sociolinguistic interviews with natives from Campo Grande (the capital of the state of Mato Grosso do Sul) and 48 with *paulistanos* (natives of the city of São Paulo, part of the project SP2010: Construction of a Sample of Paulistano Speech). These samples are stratified by the speakers’ *sex/gender*, *age group* (20-35; 35-59; 60 or older) and *level of education* (High School or College). After describing the grammaticalization trajectory of these three forms, the dissertation examines the factor groups that correlate to their variable use. These are: *discourse sequence* (narrative, descriptive, expositive, argumentative), *information status* (old, new, inferrable), *topic gradience* (intratopic, intertopic), *lexical aspect* (duration, telicity, dynamism) and *grammatical aspect* (perfective, imperfective, resultative, indeterminate). The analyses reveal that *então* is similarly preferred in both cities (while it’s slightly more frequent in São Paulo, such difference is not statistically significant). This form is also favored among college-level speakers, in both cities, which suggests that it is possibly evaluated as the most prestigious of the three forms (following the initial hypothesis, based on traditional grammars such as Cunha & Cintra, 2001). However, there is no correlation between *sex/gender* and this variable. This is contrary to the expected tendency for women to prefer *então* – the more prestigious form. This fact allows for a debate, when it comes to discourse variables, around the concept of prestige. Furthermore the variables selected in the analysis of each sample were not exactly the same: *level of education* and *grammatical aspect* are statistically significant for both, but age group was selected for São Paulo (but not for Campo Grande), while *discourse sequence* was selected for Campo Grande (but not for São Paulo). In São Paulo, *então* is disfavored by the speakers of the intermediate age group, which can be interpreted as an indication of change in the social value of *aí* and *daí*. Overall, the results show indicate that, although *então* is the preferred form, it is conditioned differently in the communities.

Keywords: *Aí; Daí; Então; Campo Grande; São Paulo; Sociofunctionalism; Grammaticalization.*

Sumário

Resumo.....	8
Abstract.....	9
Lista de tabelas	14
Lista de quadros.....	15
Lista de figuras.....	16
 Introdução	 167
 1. Um envelope de variação para <i>aí</i> , <i>daí</i> e <i>então</i>	 27
1. <i>Aí</i> , <i>daí</i> e <i>então</i> para além do nível fonético-fonológico.....	28
1.1. <i>Aí</i>	37
1.2. <i>Daí</i>	38
1.3. <i>Então</i>	38
2. A convergência das formas para um envelope de variação.....	39
 2. Funcionalismo e Sociolinguística: uma interface teórica	 58
1. Funcionalismo	58
2. Gramaticalização: uma abordagem funcionalista.....	64
2.1. Princípios de gramaticalização segundo Hopper (1991)	67
2.1.1. Estratificação (layering)	68
2.1.2. Descategorização	69
2.1.3. Divergência	70
2.1.4. Persistência	71
2.1.5. Especialização	71
2. Mecanismos de gramaticalização.....	72
2.1. Metáfora e metonímia	72
2.2. Reanálise e analogia	74
2.3. Iconicidade e marcação	75
3. Sociolinguística Quantitativa	78
4. Sociofuncionalismo.....	85

3. Córpus e metodologia.....	91
1. Três sequenciadores, duas comunidades	91
2. Para a constituição de amostras de fala.....	92
2.1. Roteiro de entrevista	93
2.2. Coleta de dados	94
2.3. Transcrição das entrevistas	95
2.4. Amostra paulistana.....	97
2.5. Amostra campo-grandense	100
2. Vizinhança, semelhanças e diferenças	103
3. A análise estatística dos dados.....	107
 4. Hipóteses e grupos de fatores.....	 112
1. Grupos de fatores.....	112
2. Grupos de fatores sociais	113
2.1. Cidade	114
2.2. Sexo/Gênero.....	116
2.3. Faixa Etária.....	117
2.4. Escolaridade	118
3. Grupos de fatores linguísticos	119
3.1. Sequência discursiva	119
3.2. <i>Status</i> da informação	126
3.3. Gradiência do tópico discursivo	128
3.4. Aspecto lexical do verbo introduzido pelo sequenciador	130
3.4.2. Telicidade (telicidade x atelicidade)	133
3.4.3. Duração (duratividade x pontualidade)	134
3.5. Aspecto gramatical do verbo introduzido pelo sequenciador	134
3.5.1. Aspecto perfectivo	136
3.5.2. Aspecto imperfectivo	136
3.5.3. Aspecto indeterminado	137
3.5.4. Aspecto resultativo.....	137
 5. Análises quantitativas	 140

1. Da perspectiva de <i>então</i>	140
2. Análise global.....	141
3. A variedade campo-grandense	150
4. A variedade paulistana	155
Conclusão	161
Referências.....	166
Índice Remissivo.....	177
Anexos	180
Anexo A – São Paulo.....	181
Anexo B – Campo Grande.....	195

Lista de Tabelas

Tabela 1. Distribuição geral dos dados em Campo Grande.....	54
Tabela 2. Distribuição geral dos dados em São Paulo	55
Tabela 3. Grupos de fatores não selecionados na amostra paulistana.....	141
Tabela 4. Tendências de uso para a variável cidade	142
Tabela 5. Distribuição dos sequenciadores por informante	143
Tabela 6. Grupos de fatores não selecionados na amostra paulistana.....	145
Tabela 7. Grupos de fatores selecionados na análise global	147
Tabela 8. Grupos de fatores não selecionados na análise global	149
Tabela 9. Grupos de fatores não selecionados na análise da amostra campo- grandense.....	151
Tabela 10. Grupos de fatores não selecionados na análise da amostra campo- grandense	153
Tabela 11. Grupos de fatores selecionados na amostra paulistana	154
Tabela 12. Grupos de fatores não selecionados na análise da amostra paulistana.....	154

Lista de Quadros

Quadro 1. Perfil dos informantes de ambas as amostras.....	98
Quadro 2. Informantes.....	103
Quadro 3. Grupos de fatores	113
Quadro 4. Aspecto gramatical	135
Quadro 5. Resumo dos grupos de fatores	139

Lista de Figuras

Figura 1. Trecho de uma transcrição da amostra paulistana no Elan	96
Figura 2. Trecho de uma entrevista da amostra campo-grandense	96
Figura 3. Divisão da cidade de São Paulo por zonas	98
Figura 4. Divisão de Campo Grande por regiões.....	101
Figura 5. Localização de São Paulo e da cidade de São Paulo no território brasileiro	103
Figura 6. Localização de Campo Grande e de Mato Grosso do Sul no território brasileiro	104
Figura 7. Campo Grande e São Paulo	105
Figura 8. Codificação dos dados no Excel.....	108
Figura 9. Interface do Rbrul	109
Figura 10. Percentuais de <i>aí + daí</i> e <i>então</i> nas amostras campo-grandense e paulistana.....	141
Figura 12. Cruzamento entre <i>escolaridade</i> e <i>sexo/gênero</i> na amostra campo-grandense	151
Figura 13. Frequência de <i>então</i> por faixa etária em Campo Grande.....	153
Figura 14. Frequência de <i>então</i> por faixa etária em São Paulo.....	155
Figura 15. Taxas de <i>então</i> no cruzamento entre <i>faixa etária</i> e <i>escolaridade</i> na amostra paulistana.....	157

Figura 16. Taxas de *então* no cruzamento entre *faixa etária* e *sexo/gênero* na amostra paulistana158

Figura 17. Cruzamento entre *escolaridade* e *sexo/gênero* na amostra paulistana.....159

Introdução

Entre o Centro-Oeste e o Sudeste do Brasil: Campo Grande e São Paulo. É à fala campo-grandense e à paulistana que se dirige a presente tese. Seria possível dizer que tais cidades vivem realidades sociolinguísticas semelhantes?

Este é o interesse geral que norteia este trabalho. Uma resposta à pergunta acima se formula aqui a partir de uma variável discursivo-pragmática que, em princípio, não parece ser tão evidente: o uso de *aí*, *daí* e *então* como sequenciadores discursivos no domínio da causalidade. A partir de uma análise qualitativa exaustiva com dados da fala de nativos dessas capitais, desenvolve-se uma pesquisa descritiva (Gil, 2002) que demonstra que *aí*, *daí* e *então* são variantes de uma mesma variável quando funcionam como sequenciadores discursivos de causalidade e, nesse contexto, seu emprego se correlaciona a outras variáveis, de natureza linguística e social.

Além da descrição dos dados e da análise variacionista que explora tais correlações, este trabalho também objetiva lançar novas luzes sobre o fenômeno estudado. A partir dos postulados de Traugott e Heine (1991), discute-se o uso desses itens: quando sequenciadores, resultam de um processo de gramaticalização, definido pelo percurso do trajeto *espaço > (tempo) > texto*. Em sua função mais antiga, *aí*, *daí* e *então* atuam como advérbios espaçotemporais, como em (1), (2) e (3):

(1) vieram pra cá... tiveram condições oportunidade...
formaram todos os filho deles aqui e [estão *aí* até hoje... em
Campo Grande]² (CG49MC-Ziraldão)³

² Em trechos com maior extensão, a fim de oferecer contextualização ao exemplo em foco, colchetes são utilizados para indicar as sequências discursivas introduzidas por *aí*, *daí* ou *então*. Quando todo o excerto transcrito está em evidência, este recurso não é utilizado.

³ Os trechos de fala de informantes paulistanos foram extraídos de entrevistas do *Projeto SP2010* (<http://projetosp2010.fflch.usp.br/>). Os de Campo Grande integram um corpus

(2) essas revista... tanto a Papéis como essas *daí* é dez reais (CG51MC-Maurício)

(3) [e a minha vó ela separou separou-se dele e... voltou pra Porto Murtinho e mora lá até *então*] e... por parte de pai a minha vó é fale/ (CG25FS-Celina)

Em um estágio posterior, tais elementos atuam intercambiavelmente como sequenciadores no domínio referencial da causalidade, tal como nos exemplos abaixo (em que cada um dos itens em negrito poderia ser substituído por qualquer um dos outros dois, sem alteração da semântica da sequenciação discursiva). A progressão de *aí*, *daí* e *então* até tal estágio é descrita e exemplificada exhaustivamente no Capítulo 1.

(4) eu não sou muito bom no português... *aí* eu... vou falar uma palavra... às vezes... eu penso (SP36SPS-Pedros)⁴

(5) [de repente a mulher está pegando mais... sendo mais tendo mais estudo que os homem hoje em dia... *daí* em algumas família está acontecendo do homem ficar desempregado] ficar em casa cuidando dos filhos e a mulher trabalhando (CG35MC-Lúcio)

(6) quando tinha é época era época da poeira assim da ventania também... muito complicado [trabalhava o dia inteiro chegava em casa aquela sujeira *então* isso me desanimava muito] (CG36FS-Renata)

Para Traugott e Heine (1991), no processo que vai dos usos dêiticos até os discursivos, ocorre transferência do contexto situacional externo ou referencial

gravado especialmente para esta tese. O código que acompanha cada excerto representa o perfil do informante: sua cidade (CG- Campo Grande; SP – São Paulo); seu sexo (M - masculino; F - feminino), sua idade (representada por dois dígitos), seu nível de escolaridade (C- colegial/médio; S - superior) e, no caso dos paulistanos, há dois códigos a mais, que indicam a região da capital paulista onde vive o informante (V – Centro velho; E – Centro expandido; P – Periferia) e a zona onde se localiza seu bairro (L – Leste; O – Oeste; N – Norte; S – Sul; C – Centro).

⁴ Para assegurar a privacidade dos informantes, foram utilizados pseudônimos que, no caso dos paulistanos, são acompanhados de uma letra maiúscula final, no sentido de diferenciar nomes repetidos (já que as entrevistas da amostra paulistana foram gravadas por vários documentadores).

para o contexto discursivo interno, resultante do conhecimento compartilhado por falante e ouvinte. A partir de tal perspectiva, este trabalho pretende demonstrar, nos termos da Sociolinguística Variacionista (Labov, 2001) e de teorias sobre gramaticalização (Givón, 1995), que tais itens funcionam como variantes de uma variável, em certos contextos.

Além de discutir os desafios da análise para fenômenos de variação que se situam além do plano fonético-fonológico, este trabalho avalia os padrões de uso de *aí*, *daí* e *então* em duas comunidades – a campo-grandense e a paulistana – cuja fala, pode-se dizer, ainda foi pouco explorada. De acordo com Mendes (2009), o português paulistano ainda foi pouco estudado (relativamente a outras variedades brasileiras) e, nas palavras de Oushiro (2015:22), “há um grande número de trabalhos *com base*⁵ no português paulistano, mas pouquíssimos *sobre* o português paulistano”. A realidade sociolinguística de Campo Grande é igualmente pouco conhecida, a ponto de ser difícil mencionar estudos sobre tal variedade e até mesmo com base na modalidade oral do português falado nesta cidade. Uma das raras obras que tratam do tema, sob um viés geossociolinguístico, é o Atlas Linguístico de Mato Grosso do Sul (Oliveira, 2007) – ainda assim, de maneira breve.

Como se trata de um estudo com dados de duas capitais, faz-se necessário ressaltar aqui a noção de *comunidade de fala*. Para Guy (2001), tal conceito pode ser definido com base em três critérios: (i) os falantes devem compartilhar traços linguísticos que sejam diferentes de outros grupos; (ii) a frequência de comunicação entre os membros dessa comunidade deve ser mais alta do que com pessoas exteriores a ela; (iii) esses indivíduos devem apresentar as mesmas normas e atitudes em relação ao uso da linguagem.

Guy (2001) argumenta ainda que os limites entre uma comunidade de fala e outra devem ser analisados com base em diferenças gramaticais e não só em diferenças na regularidade de uso de certa variável. Em síntese, o linguista postula que: (i) existem diferenças de frequência de variantes em comunidades de fala

⁵ Grifos da autora.

distintas, embora o efeito de contexto possa permanecer; (ii) há diferenças em termos do efeito de contexto entre as comunidades, o que postularia diferenças estruturais no lugar de diferenças de frequência. Labov (2008 [1972]), por sua vez, assegura existirem fatores que determinam o pertencimento de um indivíduo a certa comunidade de fala e não a outra. Desse modo, no presente trabalho, utiliza-se o termo *comunidade* para tratar de cada uma das cidades estudadas.

É evidente que São Paulo possui uma formação demográfica e social mais complexa do que Campo Grande, visto que a capital paulista pode ser considerada uma megalópole e é o centro econômico do país, ao passo que Campo Grande é a capital do estado reconhecido por um dos patrimônios ecológicos nacionais, o Pantanal. Entretanto, um olhar menos apressado pode detectar semelhanças entre São Paulo e Campo Grande: ambas são cultural e linguisticamente multifacetadas. (Capítulo 3).

Nesse sentido, tal pesquisa dialoga com outros trabalhos que enfocam a fala paulistana – especialmente Oushiro (2015), Santos (2015), Mendes (2009) – e com Oliveira (2007), sobre Campo Grande. A orientação variacionista do presente trabalho segue Labov (2008 [1972]; 2006 [1966]; 2001), ao passo que o viés variacionista é amplamente baseado em Gonçalves et al (2007), Tavares (2003), Martelotta (1999), Paiva (1995), Givón (1995), Macedo e Silva (1996), Hopper & Traugott (1993). É na interface sociofuncionalista que se busca compreender como fatores de natureza social colaboram para o uso variável dos sequenciadores discursivos *aí*, *daí* e *então* nas falas paulistana e campo-grandense.

Merecem destaque os trabalhos de Tavares (1999; 2003) sobre a variação entre *e*, *aí*, *daí* e *então* – itens em processo de gramaticalização - em cinco contextos de *sequenciação retroativo-propulsora*, na terminologia proposta pela autora. Ela analisa o emprego das formas na fala florianopolitana com base em princípios do sociofuncionalismo. Além de Tavares (2003), Martelotta (1993), sob o prisma do Funcionalismo linguístico, defende que os circunstanciadores temporais e alguns sequenciadores discursivos (entre eles, *aí*, *daí* e *então*) são

elementos da mesma natureza⁶, que se encontram em pontos diferentes de um processo de gramaticalização. Por outro lado, Braga e Paiva (2003) também abordam a multifuncionalidade de *aí* e *daí* sob a perspectiva funcionalista, mas não se atêm a uma variedade de fala específica. Pezatti (2008), por sua vez, dedica-se ao item *então* e discute se tal forma já assumiu, de fato, o estatuto de conjunção.

Conforme discussão desenvolvida no Capítulo 1, além dos contextos sugeridos por Sweetser (1991) – referencial, epistêmico e atos de fala -, dos quais se apropria este trabalho, há, evidentemente, outros contextos em que *aí*, *daí* e *então* são utilizados, mas, neles, ou a intercambiabilidade entre tais formas não existe, ou é possível, mas não atestada. Nos contextos discursivos que Koch (1998) e Tavares (2003) definem como *sequenciação temporal* e *sequenciação textual*, observam-se, respectivamente: ações eventos enunciados em função da ordem em que ocorreram no tempo, envolvendo a pressuposição de que o segundo evento ocorreu mais tarde em relação ao primeiro (7-8) e (9) e ações e eventos que se sucedem ao longo do tempo discursivo, cuja ordem de enunciação não obedece a uma sequência cronológica (9).

(7) [quando a gente/ quando eu era pequena a maioria era só paulistanos *aí* começaram a chegar pra construção] porque começou esse boom de construção (SPF65SPS-laraM)

(8) [minha mãe... a origem dela é baiana ela nasceu na Bahia diz que veio pra cá com dez anos *daí* nunca mais voltou] era da da cidade de... Xique Xique na Bahia (CG49MC-Ziraldo)

(9) eu sempre estudei no Pacaembu *então* eu estudava numa escola que chamava {C.} (SPF26SEO-VivianeC)

A exemplo dos trechos acima, observou-se, nas amostras campo-grandense e paulistana, o uso quase exclusivo de *aí* e *daí* na sequenciação temporal, por um lado, e de *então* na sequenciação textual, por outro. Ou seja, essas três formas não

⁶ O que aqui se denominam *sequenciadores discursivos*, Martelotta (1999) designa como *operadores argumentativos*. Essas nomenclaturas são discutidas no capítulo 1.

funcionam, nestes contextos, do mesmo modo como no domínio da causalidade – em termos de intercambiabilidade.

A causalidade resulta, então, da sequenciação temporal (Tavares, 1999), exemplificada em (7) e (8), quando há uma relação lógica entre dois eventos, imbricados por concepções de motivo, justificação e explicação (10). Para Tavares (1999), o traço anafórico temporal aparece reduzido quando o sequenciador, em vez de interligar ações e eventos de sucessão temporal, opera com causa e consequência:

(10) tem gente ali que fala hebraico fala um pouco de
alemão fala *aí/daí/então* ⁷ você aprende ouvindo
(CG55MC-Alberto)

No excerto acima, aprender hebraico e alemão é consequência de estar em um ambiente com pessoas nativas. Nesse exemplo, o sequenciador originalmente utilizado no enunciado pelo informante foi *então*, mas este trabalho explica, mais adiante, que *aí* e *daí* também são possíveis nesses contextos. De fato, registra-se um número expressivo de ocorrências como (10), com as três formas em foco. É este conjunto de dados que a presente tese analisa.

As amostras da fala paulistana e campo-grandense aqui utilizadas são constituídas de 48 entrevistas cada uma (totalizando, portanto, 96 gravações), estratificadas de acordo com o *sexo/gênero* dos falantes, sua *faixa etária* (de 20 a 34 anos, de 35 a 49/59 anos; 50/60 anos ou mais⁸), seu *nível de escolaridade* (Ensino Médio e Ensino Superior completo ou em curso, no caso de falantes da primeira faixa etária), além da região onde residem. As análises quantitativas dos dados são desenvolvidas com o pacote estatístico Rbrul (Johnson, 2009), que afere a correlação entre vários grupos de fatores e o emprego de uma das três formas em foco. Conforme detalha o Capítulo 4, as variáveis independentes

⁷ Nos casos em que os três itens – *aí*, *daí* e *então* – aparecem no exemplo (no sentido de indicar variabilidade possível), a forma em negrito é aquela que foi originalmente enunciada pelo informante.

⁸ Em Campo Grande, a estratificação da faixa etária é um pouco distinta daquela que norteia a compisção da amostra paulistana, conforme detalha o capítulo 3.

incluídas nas análises são: *sequência discursiva*, *status da informação*, *gradiência do tópico discursivo*, *aspecto lexical* e *aspecto gramatical*.

Aventando hipóteses

Ao demonstrar o uso variável desses sequenciadores em um dos três domínios da causalidade sugeridos por Sweetser (1991), o primeiro questionamento que aqui se levanta diz respeito a uma possível polifuncionalidade de *então*. Tal forma é atestada nos três domínios da causalidade qualitativamente analisados – referencial, epistêmico e atos de fala -, enquanto *aí* e *daí* somente são encontrados no domínio referencial ou no domínio epistêmico (neste último, sempre diante de verbos cognitivos).

Variantes discursivas, como a analisada por este trabalho, geralmente são conhecidas por sua não correlação com variáveis sociais (Macaulay, 2005). Tal premissa é norteadora deste trabalho, no sentido de verificar se a mesma realidade pode ser aplicada aos sequenciadores estudados, ou se seria possível falar na existência de prestígio ou estigma associados a *aí*, *daí* e *então*.

Antes de partir para considerações de cunho social, é preciso ressaltar que as hipóteses aqui aventadas encontram respaldo teórico na teoria laboviana (2008 [1972]; 2006 [1966]), seminal para o entendimento da variação linguística. De acordo com o linguista, as formas variantes resultam da influência de fatores sociais como *sexo/gênero*, *idade*, *escolaridade* e *estilo* (de mais casual a mais cuidadoso, no sentido da atenção dedicada à própria fala). Para Labov (1972 [2008]), há uma escala de estilos, sendo que o menos cuidadoso possível (aquele que se produz com mínimo automonitoramento) é o que ele denomina de vernáculo. Logo, em situações mais formais, os indivíduos costumam utilizar variantes de maior prestígio na comunidade de fala e a se distanciar do seu vernáculo. Assim:

- O estilo, no sentido laboviano (2008 [1972]) pode ser uma variável significativa na seleção dos sequenciadores. Conforme se discute no Capítulo 4, ao incluir o grupo de fatores *sequência discursiva* na análise, testa-se o uso de *aí*, *daí* e *então* em quatro tipos de enunciados: *narrativo*, *descritivo*, *expositivo* e *argumentativo*, caracterizados por diferentes graus de monitoramento da fala. Aqui, o respaldo teórico é

encontrado no princípio da *iconicidade*, segundo o qual, quanto maior a quantidade de informação, maior será a quantidade de material linguístico (a forma) utilizada na codificação dessa informação (Givón 1990). Desse modo, espera-se que *aí* e *daí* sejam mais frequentes nas narrativas e descrições, sequências que exigem menor planejamento e monitoramento linguístico do falante, ao passo que *então* seria preferido em exposições e argumentações.

A partir dos pressupostos labovianos, consideram-se também algumas avaliações e crenças populares acerca de ambas as capitais, tais como a de que o campo-grandense seria caipira (Amaral, 1955) (11) e (12), enquanto o paulistano seria urbanizado. Tais crenças poderiam ter alguma correlação com as diferenças verificadas no uso de *aí*, *daí* e *então* por campo-grandenses e paulistanos.

(11) eu sei que [risos] lá para São Paulo a gente é **caipira** [risos] meio atrasado [mas isso pra mim é orgulho nós já assim... nós já estamos assim taxado como **caipira**] então a gente não faz muita questão de mudar né? (CG38FS-Emília)

(12) Campo Grande tem tem um pouco mais tem mais espaços abertos tem mais árvores nas ruas tem mais aconchego... até pela população... de você estar passando na rua e estar se juntando a uma roda de tereré ali de alguém que você nem conhece direito (...) o estado é muito regionalista é mais... muito sertanejo... muito... as pessoas são mais fechadas assim na questão de geração mesmo... de serem mais **caipiras**] (CG23MS-Renan)

Com base na tradição gramatical (Cunha & Cintra, 2001), que considera *então* como *conjunção explicativa* canônica, entende-se que, comparativamente a *aí* e *daí*, essa forma goze de maior prestígio social. Hipoteticamente, portanto, *aí* e *daí* tenderiam a ser utilizados mais frequentemente por falantes com menor nível de escolaridade. Além disso:

- Considerando-se que *aí* e *daí* podem ser tomadas como formas menos prestigiosas (talvez “mais coloquiais”), é de se esperar que tais sequenciadores sejam mais comuns na fala masculina, relativamente à

feminina, de acordo com vasta literatura sobre a correlação entre *sexo/gênero* e diferentes variáveis cujas variantes diferem entre si com base no prestígio social (Chambers, 1995; Labov, 2008[1972]; Romaine, 2006, dentre muitos outros).

- Nessa mesma linha, *então* seria preferido pelos paulistanos, mais urbanizados relativamente aos campo-grandenses, conforme foi considerado acima. A identificação do campo-grandense como caipira⁹ pode estar associada ao emprego de variantes de menor prestígio social;
- No tocante ao grupo de fatores *faixa etária*, a hipótese é que *aí* e *daí*, variantes que, possivelmente, estejam associadas a um menor prestígio social, devem ser mais frequentes na fala dos mais jovens (20 a 34 anos), que estão menos inseridos no mercado de trabalho relativamente aos falantes das outras duas faixas. Por outro lado, espera-se que *então*, em virtude de seu possível status de “forma padrão”, seja mais comum na segunda e na terceira faixa etária, justamente pela exposição de tais indivíduos ao “mercado linguístico” (Bourdieu, 1991).

Estrutura do trabalho

O **Capítulo 1** discute os desafios de uma análise sociolinguística em níveis gramaticais além do nível fonético-fonológico, mais especificamente o discursivo. Atesta-se a variabilidade entre *aí*, *daí* e *então* como sequenciadores discursivos do domínio referencial da causalidade, a partir do percurso histórico de cada um dos três itens, desde seu uso dêitico espaçotemporal, até seu emprego textual (Heine et al, 1991; Traugott e Heine, 1991 e Hopper e Traugott, 1993).

O **Capítulo 2** dedica-se aos fundamentos do Sociofuncionalismo, quadro teórico em que se situa esta tese. Trata-se de uma interface com pressupostos do Funcionalismo Linguístico e da Sociolinguística Variacionista e voltado ao estudo de fenômenos em gramaticalização. Apresentam-se os princípios e mecanismos de gramaticalização relevantes à análise dos sequenciadores *aí*, *daí* e *então* no domínio da causalidade referencial.

⁹ Essa autoidentificação do campo-grandense como caipira é ilustrada no Capítulo 3.

O **Capítulo 3** trata da constituição das amostras de fala analisadas nesta tese. Nesse sentido, descrevem-se as comunidades em questão e explicitam-se as razões que possibilitam um estudo sociofuncionalista comparativo entre elas. Por fim, apresenta-se o pacote Rbrul – com o qual se desenvolvem as análises estatísticas dos dados.

No **Capítulo 4**, apresentam-se as variáveis linguísticas e sociais incluídas nas análises e reveem-se as hipóteses que as justificam. Finalmente, o **Capítulo 5** discute os resultados de tais análises, com ênfase nos grupos de fatores que se mostraram estatisticamente significativos na sua correlação com a variação dos sequenciadores. Desenvolve-se uma análise global, em que ambas as amostras são analisadas conjuntamente, a fim de verificar se, estatisticamente, Campo Grande e São Paulo se diferenciam de forma significativa no emprego de *aí*, *daí* e *então* no contexto variável que aqui se define. Em seguida, são realizadas análises de cada amostra de dados (Campo Grande e São Paulo, separadamente).

1

Um envelope de variação para *aí*, *daí* e *então*

Neste capítulo, são considerados os desafios de uma análise sociolinguística de variáveis em níveis gramaticais “mais altos” (além do nível fonético-fonológico). As formas em foco são três itens ainda pouco explorados na literatura - *aí*, *daí* e *então* – cujo funcionamento como variantes de uma variável é atestado com base na análise qualitativa de grande quantidade de dados. A partir do percurso histórico de cada um desses três itens (desde seus empregos como dêiticos espaçotemporais até aqueles em que operam como sequenciadores discursivos), verifica-se sua intercambiabilidade nos contextos do domínio referencial da causalidade. Esse trajeto (de dêiticos a sequenciadores discursivos) é explicado à luz de teorias sobre o processo de gramaticalização (Heine et al, 1991; Traugott e Heine, 1991 e Hopper e Traugott, 1993). A intercambiabilidade dessas formas nos contextos causais se dá em virtude de propriedades que estão além do valor de verdade da sequência discursiva¹⁰ em que ocorrem.

Interessa a este capítulo constatar que *aí*, *daí* e *então* constituem variantes de uma variável em contextos específicos (o chamado envelope de variação). Para isso, são analisados qualitativamente dados de uma amostra de fala que se pode considerar bastante grande (Labov, 1984): 96 entrevistas sociolinguísticas, 48 gravadas com paulistanos (parte do *Projeto SP2010: Construção de uma amostra*

¹⁰ Com base na ideia de que a função dos sequenciadores vai além da junção de sentenças, opta-se, neste trabalho, pelo termo *sequência discursiva*, em alusão às porções do discurso interligadas por *aí*, *daí* e *então*.

de fala paulistana) e 48 com campo-grandenses, um corpus inédito gravado e transcrito especialmente para esta tese.

A hipótese inicial é de que fatores sociais, além de linguísticos, correlacionam-se ao emprego dessas formas, em ambas as capitais. Dessa forma, interessa verificar o compartilhamento de normas linguísticas no que diz respeito a seu uso nessas cidades.

1. *Aí, daí e então* para além do nível fonético-fonológico

A premissa básica para um estudo variacionista é a de que duas ou mais formas configurem “maneiras diferentes de dizer uma mesma coisa”, ou, em outras palavras, que dois enunciados tenham o mesmo valor de verdade (Labov, 1978; Lavandera, 1978). No nível fonético-fonológico, esse princípio parece ser indiscutível, já que não haveria nenhuma objeção aparente em admitir-se, por exemplo, que *planta* e *pranta* (com /r/ no lugar de /l/) têm o mesmo significado referencial em qualquer contexto linguístico.

Além do nível fonético-fonológico, contudo, é necessário demonstrar que os elementos analisados funcionam, de fato, como variantes de uma variável, sob pena de que uma análise variacionista para tais formas seja invalidada. Dessa forma, não parece tão pacífica a variabilidade de “A empresa demitiu alguns funcionários” e “Alguns funcionários foram demitidos pela empresa”, em um exemplo claro de que fenômenos morfossintáticos e discursivos resultam, se comparados aos anteriormente elencados, mais complexos para o analista no que diz respeito à definição de um envelope de variação.

De acordo com Freitag (2009), em princípio, discutir a pertinência de análises variacionistas nos níveis mais altos da estrutura gramatical pode soar como um retrocesso para a constituição do paradigma teórico da sociolinguística. Com base nas pesquisas de apagamento da cópula verbal no inglês afro-americano (*African American Vernacular English*, ou *AAVE*), muitos debates foram tecidos por Weiner e Labov (1983), Lavandera (1978), Labov (1978) e Romaine (1981). A problemática, contudo, é constantemente retomada (Milroy e Gordon, 2003;

Cheshire, Kerswill e Williams, 2005) e expõe a necessidade de sua revisão, ainda que breve.

A principal questão é demonstrar que, nos níveis gramaticais mais altos, é possível verificar “diferentes meios de se dizer a mesma coisa em um mesmo contexto” (Labov, 2008 [1972]:164). De acordo com Milroy e Gordon (2003:169-179), há alguns problemas centrais que se impõem ao estudo de fenômenos de variação para além da fonologia: (a) a falta de clareza na aplicação do conceito de variável sociolinguística em relação a esses fenômenos; (b) a dificuldade de elencar, com precisão, quais fatores contribuem para delimitar as variantes que compõem uma variável linguística situada em tais planos; (c) uma análise sociolinguística de variantes sintáticas e discursivas deve, em primeiro lugar, demonstrar que há equivalência de significado entre as formas, o que não ocorre com variáveis fonético-fonológicas, em que a intercambiabilidade é patente.

Assim parece ser o caso de *aí*, *daí* e *então*, itens que, por sua origem dêitica, não configurariam, à primeira vista, um quadro de formas em variação. É o que se verifica em:

(13) eu já tive problema com droga é uma coisa difícil pra mim falar [mas é uma coisa que está **aí**] (SP31MCPS-RobsonF)

(14) sai **daí** menino você vai se machucar! (CG40MS-Ronaldo)

(15) até **então** nunca tinha visto nada parecido (SPF72SPO-DercyF)

Nos excertos acima, *aí* e *daí* funcionam como dêiticos espaciais, enquanto *então* funciona como dêitico temporal. A substituição de uma forma pelas outras duas em cada passagem não faria sentido, de modo que casos como esses não integrariam um envelope de variação.

Ocorre que, nesse caso, a variabilidade de tais itens não é uma máxima. Observa-se um contexto – o da causalidade referencial, em que a variação não só é possível, mas também constatada nos dados (maiores detalhes na seção 2 deste

capítulo). Nesses casos, *aí*, *daí* e *então* passam por uma dessemantização e funcionam como sequenciadores de causa-consequência.

(16) minha mãe nunca tinha morado em cidade grande
aí/daí/então ela se assustou muito com o assalto (CG58F5-
Valéria)

Por outro lado, a definição do nível de análise de um fenômeno variável depende do enfoque dado pelo analista. Mendes (2005), por exemplo, em sua pesquisa sobre a variação entre *ter + participípio* e *estar + gerúndio* na expressão dos aspectos durativo e resultativo (no tempo presente), leva em consideração variáveis semântico-estruturais, *como o aspecto verbal, o tipo semântico do verbo principal, adjuntos adverbiais, número do sujeito* e dos argumentos internos. Em um enfoque discursivo, como o que se leva a cabo aqui, também poderiam ser considerados o *gênero discursivo*, o *tipo de sequência discursiva*, entre outros.

A variação de *aí*, *daí* e *então* na causalidade referencial pode ser definida como morfossintática. Nesta tese, contudo, entende-se que a natureza do fenômeno é discursiva. Ressalte-se, que disso não decorre que fatores morfossintáticos não possam ser incluídos na análise, conforme se discute no Capítulo 4.

Segundo Loureiro (2004), Nebrija, em sua gramática datada de 1492, já havia destacado a existência de “partículas invariáveis do discurso”. Nesse grupo, o estudioso incluiu as classes de palavras já categorizadas desde a tradição da gramática latina, como conjunções, preposições e locuções prepositivas. Apenas séculos mais tarde são identificados elementos linguísticos que não se ajustam às classes gramaticais existentes e, pela primeira vez, observam-se suas características discursivas.

De acordo com Almeida e Marinho (2012: 171), Gili Gaya, um dos pioneiros no estudo dessas partículas, nomeou-as como “enlaces extraoracionais” e identificou alguns dos traços que os definem, tais como: a) a propriedade de veicular noções externas à relação de predicação oracional; b) seu caráter invariável - no tocante às flexões de gênero e número; c) a fluidez de categorização gramatical; d) a flexibilidade de sua distribuição sintática (ocupam a posição inicial,

medial ou final na sequência discursiva em que aparecem); d) o fato de que, ao evidenciar a coerência, realçam valores semânticos específicos, frequentemente combinados com partículas suprasegmentais, como a entonação; e) o fato de serem mais ou menos específicos de cada registro (alguns estão mais ligados à escrita e outros à oralidade).

Aqui, a opção pela nomenclatura “sequenciadores discursivos” se justifica, em um primeiro momento, pela natureza funcional de *aí*, *daí* e *então* como itens que interligam duas ou mais porções do discurso ou *sequências discursivas* (Adam, 1992). A ideia de *sequência* remete, então, a ações ou eventos interdependentes sob o ponto de vista discursivo, como em (17), em que o informante explica a origem do sotaque de Ladário (MS):

(17) devido Ladário ter a Marinha vai muito marinho lá do Rio *aí/daí/então* eles pegaram abraçaram muito o sotaque do Rio de Janeiro "de titio" "de titia" é "dois três" {pronunciando o /t/ africado e /s/ chiado} (CG35MC-Lúcio)

Sintaticamente, a sequência anterior a *daí* é independente do que vem depois do sequenciador. A relação entre a Marinha e as características fonético-fonológicas do falar de Ladário, no entanto, só é recuperada se consideradas as sequências discursivas anterior e posterior a *daí* (em negrito, para indicar que foi a forma originalmente empregada pelo informante – mas as outras duas formas poderiam ter sido igualmente empregadas, com a mesma função).

A forma *então*, comum na oralidade e no português escrito, é geralmente categorizada como conjunção conclusiva. *Aí* e *daí*, por sua vez, são típicos da fala, mas não são lembrados como conjunções que introduzem ideia de conclusão. Bechara (2003:194) afirma que conjunção “é a expressão que liga orações ou, dentro da mesma oração, palavras que tenham o mesmo valor ou função”. Para Perini (1998:333-335), os conectivos são assim chamados por serem palavras que funcionam como elementos de conexão entre constituintes. Essas definições esbarram na visão acerca desses sequenciadores no presente estudo, de cunho funcionalista (Givón, 1995; Hopper e Traugott, 1993).

Como sequenciadores discursivos de causalidade, ao mesmo tempo em que interligam as sequências discursivas, *aí*, *daí* e *então* estabelecem nexos semânticos entre elas, guiando a interpretação do ouvinte. Na verdade, a noção de continuidade veiculada pelos elementos em questão não reside apenas nos termos em si mesmos, mas na conjuntura resultante da estrutura linguística e do contexto situacional. Em outras palavras, é no discurso que se atualiza o sentido dos sequenciadores (Beaugrande & Dressler, 1981), como em:

(18) [meu avô demorou um pouquinho mais pra sair do carro *aí/daí/então* a gente ficou com medo] deles levarem ele junto (SP31MSP5-CarlosJ)

Em (18), o medo relatado pelo informante decorre do fato de seu avô ter demorado a sair do carro logo após ele e seus familiares terem sido abordados por assaltantes na estrada. Nesse excerto, uma ação desencadeia a outra e é possível utilizar indiscriminadamente qualquer um dos sequenciadores em questão.

Além de *sequenciadores*, outros trabalhos referem-se a itens como *aí*, *daí* e *então* de diferentes maneiras, como as mencionadas por Almeida & Marinho, (2012:170): *marcadores de relação textual*, *operadores discursivos*, *articuladores discursivos*, *enlaces extraoracionais*, *conectores discursivos*, *conectores pragmáticos*, *partículas pragmáticas*, *partículas discursivas*. Entretanto, considerando-se que esses itens realçam a relação discursiva existente entre duas sequências discursivas e que a sequenciação é o domínio funcional responsável pela introdução de informações no discurso (Tavares, 2003), o termo *sequenciadores discursivos* parece mais proficiente (bem como mais condizente com a perspectiva teórico-metodológica a partir da qual esses dados são aqui analisados – conforme se descreve a seguir e no Capítulo 2).

O sistema linguístico é integrado tanto por regras categóricas quanto por regras variáveis, que podem ser aplicadas com maior ou menor frequência. Todas elas são condicionadas pelo contexto linguístico e/ou social, que confere ao sistema um caráter probabilístico e abre prerrogativas para que seja quantitativamente observado, de acordo com as regularidades que lhe são naturais.

Para o modelo variacionista, uma maior taxa de utilização de uma variante, atrelada a outros fatores, pode ser um indício de mudança e, aqui, mais uma vez, o estudo de fenômenos discursivos apresenta certa resistência, já que se trata de variáveis em geral menos frequentes do que as fonológicas e certas variáveis morfossintáticas.

No plano fonológico, o falante emprega um conjunto definido, limitado e previsível de elementos, encontrados abundantemente em uma amostra de fala oriunda de entrevistas sociolinguísticas. Antes de verificar a frequência de determinado fenômeno, é preciso avaliar se a amostra sociolinguística é representativa de uma comunidade de fala, que é representada pela escolha aleatória de falantes pertencentes a ela.

Se os dados não forem suficientemente recorrentes para subsidiar uma análise estatística confiável, é porque a variável não é bem representada. Tal fato pode implicar a necessidade de nova coleta posterior, ou pode ser um sinal de que o fenômeno requer tratamento apenas qualitativo.

Garantida a relativa frequência do fenômeno em um nível gramatical mais alto na amostra analisada, ainda é preciso atentar para a possibilidade de que haja concentração de dados na fala de apenas alguns dos informantes, o que poderia caracterizar, a despeito de uma regra variável, uma marca estilística desses indivíduos.

Além da frequência, outro entrave encontrado na análise de fenômenos acima do nível fonético-fonológico é a questão da unidade semântica entre as variantes, tema do clássico debate entre Labov e Lavandera no fim da década de 1970. Lavandera (1978) questiona a regra variável de Weiner e Labov (1983), em que a passiva sem agente (*the closet was broken into*) e a ativa com sujeito genérico (*they broke into the closet*) poderiam ser consideradas variantes sintáticas. Para Lavandera (1978), haveria de se considerar não só a condição de mesmo significado, mas também o de comparabilidade funcional, de tal modo que as estruturas, apesar de terem o mesmo significado, não apresentariam necessariamente a mesma função comunicativa.

É inquestionável que o /r/ em coda silábica no português brasileiro (PB), como em *car.ne*, independentemente da forma como for pronunciado – como

retroflexo, tepe ou aspirado - , não será alvo de questionamentos em torno da significação. Contudo, de acordo com Naro e Scherre (2006), em fenômenos como a alternância entre *nós* e *a gente* na expressão pronominal de primeira pessoa do plural, e a regência do verbo *ir*, existe considerável distinção semântica entre formas variantes. Logo, sob a perspectiva da sociolinguística laboviana, não há um caso de variação canônica, caracterizado pela equivalência semântica das formas variantes em todos os contextos. O que se tem são estruturas que funcionam de modo equivalente, mas que não variam de forma incontestável.

Para tal, a definição de Nichols (1984) para o binômio função/significado tem sido ressignificada, o que resulta em uma nova definição de variantes (Labov, 1978), que passam a ser não só formas de *mesmo valor semântico*, mas também de *mesma função*, fato que realça a propriedade discursiva das variáveis. Tem-se, ao fim, função/significado atrelado ao contexto de uso, em detrimento de formas que desempenham funções ou codificam significados.

Quando se incorporam aspectos discursivos ao paradigma da sociolinguística, procede-se, na verdade uma tentativa de descrever a coexistência de duas ou mais formas de dizer a mesma coisa com o mesmo valor de verdade. Esse avanço se torna possível à medida que, na caracterização das variantes, outros tipos de significado, diferentes do referencial, e mesmo nuances de significado, são tomados como variáveis independentes (Paiva; Scherre, 1999),

As nuances de significado que caracterizam as variantes são controladas como uma variável independente, inclusive quando tais traços semânticos são muito claros. É o que acontece com os casos de ambiguidade a que se dedicou Mendes (2005). De acordo com o linguista, no domínio da expressão do aspecto verbal, as perífrases *estar + gerúndio* e *ter + particípio* constituem variantes de uma variável, no tempo presente na expressão dos aspectos durativo e resultativo, como nas ocorrências abaixo (Mendes, 2005:30):

- (i) João **tem trabalhado** pela modernização da área (durativo)
- (ii) João **está trabalhando** pela modernização da área (durativo)
- (iii) João **tem viajado** de carro (resultativo)
- (iv) João **está viajando** de carro (resultativo)

Caracterizado pela propriedade de expressar uma situação como contínua, o aspecto durativo refere-se a um intervalo dentro de uma sequência temporal. Já o aspecto resultativo exprime a repetição/intermitência de uma situação na estrutura temporal. Mendes (2005:43) explica, contudo, que há contextos em que ambas as interpretações (continuidade ou intermitência) são possíveis, o que o leva à proposta de casos de ambiguidade, como o excerto (19):

(19) atualmente pensa-se em quantidade... abre-se faculdades à vontade mas um outro campo também que eu vejo aí que seria mais um campo dos administradores e economistas... ***estão sendo invadidos*** e são pelos engenheiros (D2 SP 62¹¹)

A ambiguidade advém do processo de gramaticalização, que invisibiliza o tratamento de formas como categorias discretamente definidas. Em processos de variação que têm a ver com gramaticalização, as formas não são necessariamente especializadas para um ou outro caso. Entretanto, elas podem vir a se especializar.

No caso de *aí*, *daí* e *então*, é preciso delimitar, ao certo, os domínios funcionais considerados na análise, a fim de evitar sobreposições e no sentido de viasualizar bem os casos em que de fato funcionam como variantes de uma variável. Aqui, *domínio funcional* pode ser definido nos termos de Givón (1984), em alusão a áreas funcionais gerais (ou macrodomínios) que compõem a gramática e que podem corresponder a áreas funcionais gerais, como TAM (tempo/aspecto/modalidade), ou a áreas mais estritas (microdomínios), como o tempo futuro, o sujeito, a dêixis, etc.

Assim, *domínio* pode ser definido como diferentes domínios superordenados nas hierarquias funcionais língua: o aspecto, por exemplo, é um microdomínio dentro do domínio TAM. Desse modo, a causalidade seria um domínio funcional micro em relação à sequenciação, de que fazem parte relações conjuntivas de outra natureza.

¹¹ Excerto extraído do NURC (Projeto Norma Urbana Culta): entrevista com paulistano, diálogo entre dois informantes (ambos homens), primeira faixa etária (25 a 35 anos).

Em cada domínio funcional, há alguma forma ou formas que desempenham uma função específica, ou até mesmo um conjunto de elementos “unificados funcionalmente” (Nichols, 1984:111), que exercem papéis semelhantes. Cada uma das unidades que atuam como marca de um domínio funcional¹² representa uma camada do domínio. Aqui, “camada” (cf. Hopper, 1991) diz respeito à alternância das formas em relação de estratificação em uma etapa da língua. Como o surgimento de uma nova forma em determinada função não acarreta necessariamente o desaparecimento das mais antigas, é possível afirmar que tais vocábulos certamente remontam a épocas distintas na gramática.

Desse modo, os itens inovadores passam a conviver e a disputar com os demais, não só em relação à gramática dos indivíduos, mas também na da comunidade. É o que se verifica com os itens em foco nesta tese: embora adquiram novas funções, *aí*, *daí* e *então* continuam atuando como dêiticos. O que ocorre é que diferentes fatores, de ordem cognitiva, comunicativa, estrutural e social, podem influenciar o uso de novas funções para uma mesma forma, auxiliando em seu avanço ou em seu recuo no processo de mudança.

Definidos o domínio funcional - a causalidade - e a variável linguística no que tange à forma e à função - *aí*, *daí* e *então* como sequenciadores discursivos - procede-se ao último problema imposto na análise de um fenômeno discursivo: a necessidade de uma complementação teórica. Por ser inviável uma explicação estritamente variacionista, faz-se necessário formular um paradigma teórico que permita interpretar os resultados quantitativos e que torne possível uma análise qualitativa integrada.

A interface em que esta tese se situa é a do sociofuncionalismo (Tavares, 2003; Naro & Braga, 2001), que conjuga a teoria variacionista e processos de

¹² É importante ressaltar que, na concepção givoniana, a gramática não é apenas um mecanismo de codificação superficial dos diferentes domínios funcionais universais organizados para a comunicação. A linguista se opõe tanto à visão de Chomsky, segundo a qual a gramática seria um algoritmo, quanto à de Paul Hopper - *gramática emergente* - que define a gramática sob uma perspectiva totalmente flexível, sempre negociada e totalmente vulnerável ao contexto comunicativo. Givón (2001) assume uma posição intermediária, ao alegar que as línguas apresentam, ao mesmo tempo, *estabilidade e flexibilidade*, uma vez que essas duas características são necessárias, respectivamente, para o processamento rápido da linguagem e para a adaptação da língua aos diferentes propósitos comunicativos.

gramaticalização. O foco são relações entre funções e formas, oriundas de pressões linguísticas e sociais, com ênfase para a história e a coexistência de diferentes formas, situação de estratificação/variação (Hopper, 1991). De forma geral, a abordagem sociofuncionalista dedica-se às funções gramaticais determinadas pelo uso. Nessa perspectiva, a gramática é emergente, decorrente de constantes mudanças determinadas pelo uso dos falantes. É possível, aqui, compreender como *aí*, *daí* e *então*, aprioristicamente dêiticos e invariáveis, adquirem propriedade sequenciadora como resultado de mudança linguística por gramaticalização, constituindo uma variável linguística no âmbito da causalidade. Assim, acumulam funções diversas, a depender do contexto em que é empregado.

A seguir, apresenta-se uma breve reflexão sobre o percurso histórico desses itens (de monofuncionais dêiticos a polifuncionais), a fim de se chegar à constatação de um envelope de variação para essas formas – ou seja, a definição dos contextos em que elas funcionam como variantes de uma variável.

1.1. *Aí*

Segundo Costa (2003), a forma adverbial *aí* tem origem no latim, *ibī*. Muitas pesquisas e dicionários etimológicos atestam que os advérbios mais frequentemente empregados no latim e no português arcaico/moderno eram os locativos, anafóricos e catafóricos. Sendo assim, o uso de *aí* na esfera textual parece não ser muito antigo. De acordo com Braga et al (2003), só há registros de usos desse item como sequenciador no português contemporâneo.

Para Braga et al (2003), no latim clássico, a forma *ibī* era um advérbio espacial, com valor anafórico. Embora existisse um item adverbial específico para a exófora de segunda pessoa, *istīc*, tal forma desapareceu do português, ao passo que *i* (com as variantes *hy* e *hi*)¹³ é amplamente documentado, em português,

¹³ Margarit (1947) tenta explicar a variação entre *i*, *hy* e *hi* no português arcaico. Segundo o autor, durante a fase românica, era comum encontrar formas impessoais do verbo *habere* precedidas por *ibī*, como em *ibī habet*, justamente com o intuito de dar ênfase à noção de lugar. Cambraia (2002) também defende que a preposição *a*, no período arcaico do português, era usada para indicar lugar, o que endossa a hipótese aventada por Margarit (1947).

desde 1272” (op.cit:157), apresentando vários correspondentes em outras línguas românicas.

Foi na passagem do latim para o português que surgiu *í*, resultante de alterações fonético-fonológicas (Costa, 2003) da forma adverbial *ĩbĩ*. A forma adverbial *í* passa a ser reforçada pelo uso da preposição *a*, a partir do século XVI e, assim, dá origem à forma atual, *aí* – mais recente, portanto, no português.

1.2. Daí

Heine et al. (1991) lembram que as formas preferidas à gramaticalização são, em relação ao significado, aquelas que remetem a universais da experiência humana, que dizem respeito a aspectos concretos, básicos e gerais para a localização no espaço. Isto possibilita a tais formas a propriedade de se estenderem a conceitos menos concretos. Dentre esses elementos, destacam-se os itens lexicais que designam partes do corpo, os verbos dinâmicos, os quantificadores e os demonstrativos básicos, sobretudo os espaciais.

Este é o caso de *daí*, uma construção formada pela contração da preposição *de* com o dêitico espacial *aí*, que parece ter começado a ser usado, com as variações *dy*, *di*, *dj*, somente a partir do século XIV (Tavares, 2003).

1.3. Então

No português atual, nunca se verificou o valor espacial de *então*, embora esse valor possa ser recuperado na origem latina do elemento: a antiga forma *intunc* (*in* + *tunc*). De acordo com Ernout e Meillet (1959), *tunc* resulta da aglutinação de *tum* + *ce*. A partícula *ce* tem valor demonstrativo e é comumente encontrada nas línguas itálicas. Liga-se a pronomes demonstrativos, como *hic*(e) (este) e *illíc*(e) (aquele), ou a advérbios provenientes de temas demonstrativos, como *síc*(e) (assim) e *nunc*(e) (agora).

O elemento *tum*, segundo Faria (1975), tem valor de advérbio e pode denotar “então, naquele tempo, depois disso, donde, além disso, por outro lado”. A base demonstrativa de *tum* é registrada por Leite e Jordão (1958), que explicam que a origem desse elemento é a raiz do grego antigo *te*, também encontrada na formação de outros elementos de intensificação *talís*, *tantus*, *tot* e *tam* e dos

pronomes *iste, ista, istud*, que podem significar esse, essa, isso; tal, tamanho, semelhante. Segundo Bailly (1950), o elemento *te* do grego antigo funcionava como partícula enclítica e possuía dois sentidos básicos. No primeiro, como palavra demonstrativa, correspondia a um pronome relativo e tinha valor de “por isso, por essa razão”.

Essa origem demonstrativa (que remete a dados espaciais) do elemento *tum* e dos demais elementos de intensificação acima mencionados explica o valor anafórico que estes elementos ainda hoje apresentam. É desse valor anafórico que surgem os atuais valores argumentativos desses elementos: então (conclusivo), portanto (conclusivo), entretanto (adversativo), entre outros.

Segundo Williams (1975), a transformação de *in tunc*, do latim, para *então*, no português, deve-se, primeiramente, a um processo de apócope, que deu origem a *intun*, no latim vulgar. Por fim, a sílaba -tão, da atual forma *então*, veio por analogia com as terminações latinas *-anu, -one > -an, on, -ane > ão*, a exemplo de palavras como *cane > cão, nationes > nações* e *manu > mão*.

2. A convergência das formas para um envelope de variação

De acordo com Heine, Claude e Hünemeyer (1991), em seu estudo a respeito dos operadores argumentativos, os vocábulos com estatuto adverbial percorrem um *continuum* até se estabelecerem como mecanismos de sequenciação discursiva. Tal percurso pode ser assim caracterizado:

espaço > (tempo) > texto

Na primeira etapa, os itens têm valor dêitico-espacial; na segunda, que pode ou não ser observada nos dados (ou seja, pode ou não “emergir”), o item funciona com valor temporal; finalmente, na terceira, ele passa a operar como organizador textual. Quando não há dados que permitam exemplificar a segunda etapa, diz-se que o item passou diretamente de dêitico-espacial para um operador textual.

É possível também que, em seu percurso de gramaticalização, um advérbio seja utilizado diretamente no estágio temporal, sem que seu uso espacial seja verificado. É o caso de *então*, cujos registros com valor temporal datam desde o século XIV:

e San Gregorio disse ainda que em tempo dos godos
o ben aventurado baron don Johane, bispo desta
eigreja de Roma, i a Justiniano, o mais velho, que
enton era emperador, chegou a ti de Corinto.
(FIGUEIREDO-GOMES & PENA-FERREIRA,
2006:97).

Nessa passagem, *então* é empregado para fazer referência à época em que Justiniano era imperador, um uso dêitico temporal. Conforme Heine et al (1991:179), itens dêiticos tendem a assumir função anafórica, tendência manifestada por *aí*, *daí* e *então*. Na primeira etapa do percurso rumo à conexão de partes do texto, *aí* e *daí* atuam como dêiticos locativos (20-21) e sinalizam um lugar do mundo externo.

(20) o prefeito está fazendo melhorias pra cidade
né? [a gente vê tanta coisa bonita... e a educação
está **aí** você vê cada vez melhor] as escolas melhor
(CG30FC-Paula)

(21) é um bairro meio rico assim né? [eu tô meio
numa ponta do bairro... que depois da Rebouças
para lá é Jardins... **daí** pra cá é Pinheiros e Pinheiros]
já é um bairro um pouco mais... comercial
(SP24MSEO-Marcelo)

Em (20), *aí* indica uma localização espacial. Ao argumentar sobre a gestão do atual prefeito, a informante faz uso do advérbio para indicar que as melhorias são nítidas e “estão *aí*”. Já em (21), *daí* delimita uma porção espacial, a fim de demonstrar a extensão territorial do bairro de Pinheiros.

Como dêitico, somente é possível encontrar usos de *então* com valor temporal. O valor espacial do termo ficou retido em sua forma latina. Em (22), *então* equivale a “hoje”, “agora”, com valor nitidamente temporal.

(22) meu avô por parte de mãe já é falecido e e ele nunca saiu dessa cidade que minha mãe nasceu que é San Stanislau e [a minha vó ela separou separou-se dele e voltou pra Porto Murtinho e mora lá até *então*] e por parte de pai a minha vó é falecida (CG25FS-Celina)

Como resultado de extensão metafórica (Traugott, 1991), o valor espaço-temporal dos dêiticos relaciona-se com porções já enunciadas no texto e não com o mundo externo. Logo, *aí*, *daí* e *então* passam da dêixis para a anáfora.

Além dos processos de mudança metafórica (conforme detalha o Capítulo 2), também são comuns os que ocorrem por pressão de informatividade, típicos de contextos em que, por um processo de inferência conversacional, um determinado elemento linguístico assume um novo valor.

Traugott & König (1991) defendem que os mecanismos de metáfora espaço > (tempo) > texto e pressão de informatividade não são incompatíveis. É possível cogitar sua coocorrência no processo de mudança de alguns elementos linguísticos. A metáfora representa um sentido novo, intrinsecamente inferível do sentido original. É o que ocorre na passagem da sequencialidade temporal para a causalidade; nesta última, a ideia de consequência veiculada pelo elemento conclusivo é inferível da noção de sequencialidade temporal, já que a causa precede a consequência. No segundo processo, o sentido novo resulta de uma relação semântica estabelecida em certos contextos.

Observam-se casos de *aí* e *daí* como anafóricos locativos. Mais uma vez, não são encontradas ocorrências de *então* com valor espacial:

(23) a segunda era uma escola pequena também que [eu fiz da terceira terceira e quarta série aí eu fui pra uma maior bem maior mesmo uma escola grande que *aí* foi até o ensino médio] (CG25FS-Celina)

(24) veio pra cá com dezoito anos mais ou menos foi morar em Santos foi *daí* que ele conheceu minha mãe (SPF72SPO-DercyF)

Em seu relato sobre as escolas por que passou, a informante em (23) utiliza *aí* com valor anafórico locativo para referir-se à escola onde estudou desde a quinta até o Ensino Médio. Em (24), *daí* é utilizado para fazer referência à cidade de Santos, que já havia sido enunciada.

No caso de *aí* e *daí*, o anafórico locativo dá origem ao anafórico temporal – como em (25) e (26). Aqui, observam-se também ocorrências de *então* como anafórico temporal (25), que seriam um estágio posterior, em português, de sua função dêitica latina:

(25) eu o que eu busco é fazer um concurso onde eu adquira em que eu adquira essa autonomia pra eu poder *aí* tomar as medidas que eu entendo como corretas hoje né? (CG32MS-Raul)

(26) fiz um curso técnico de de Contabilidade... né? assim... um ano e meio fui numa escola Rui Barbosa uma escola que existe hoje (...) [eu fiz assim... um ano e meio esse curso de... de de técnico em Contabilidade na época *daí* pra cá... me relaxei] não não estudei (CG49MC-Ziraldo)

(27) e eu realmente bebi muito naquele dia e lá tinha piscina eu tomei banho de short camiseta essas coisa... [e no outro dia as pessoas comentavam... assim tipo debochando sabe? “você viu *então* a N. não sei o quê”]... todo mundo ficou sabendo (CG48FC-Nair)

Em (25), o informante almeja passar em um concurso para que tenha mais autonomia no trabalho. Ele utiliza *aí* para referir-se ao futuro, ou seja, ao dia em que for aprovado. *Daí*, em (26), substitui o termo “na época” e delimita, no discurso, o tempo em que o informante alega ter desistido dos estudos.

No excerto (27), *então* é empregado como referência anafórica que remete a “naquele dia”. Assim, parte do excerto poderia ser enunciada como “você viu a N. naquele dia?”. Este é o tipo de contexto que pressiona a mudança anafórico > sequencial. Ocorre, nesse caso, uma gramaticalização por pressão de informatividade, nos moldes do que está desenvolvido em Traugott & König

(1991), em que se pode inferir, de determinados contextos em que ocorre *então* anafórico, um valor sequencial de base temporal.

É importante ressaltar que, nos trechos de (25) a (27), não é possível substituir um dêitico pelo outro (o que só se verifica como possível no domínio funcional da sequenciação, conforme já se mencionou anteriormente e conforme se discute adiante). Para exemplificar a impossibilidade de intercambiar os dêiticos, vejam-se os exemplos (20), (21) e (22), renumerados abaixo como (28), (29) e (30):

(28) o prefeito está fazendo melhorias pra cidade né? [a gente vê tanta coisa bonita... e a educação está *aí***daí***então* você vê cada vez melhor as escolas melhor]
(CG30FC-Paula)

O uso de *daí* ou *então* descaracterizaria semanticamente o período. Com *daí*, o que se pode entender é que a Educação está perto do ouvinte e distante de quem enuncia. Como não é possível, em língua portuguesa, recuperar o sentido espacial de *então*, a substituição deste termo por *aí* daria origem a um período agramatical.

(29) é um bairro meio rico assim né? [eu tô meio numa ponta do bairro que depois da Rebouças para lá é Jardins... *aí*(?)/*daí***então* pra cá é Pinheiros]... e Pinheiros já é um bairro um pouco mais... comercial (SP24MSEO-Marcelo)

Em (29), tem-se um uso de *daí* como um dêitico locativo. De fato, o *aí* de *daí* retoma *Jardins*. Nesse caso, a troca de *daí* (de+*aí*) por *aí* não faria sentido, já que, para o sentido expresso no exemplo, a preposição é necessária à delimitação de um espaço (“dos Jardins para cá”). Quando se tenta a permuta por *então*, o excerto se torna agramatical, já que *então* não tem valor espacial.

Por outro lado, considerando-se uma realização do excerto em que houvesse uma pausa entre “Jardins” e “*daí*”, o que se teria é um uso do termo como sequenciador, de modo que a intercambiabilidade por *aí* ou *então* seria possível. Nesse caso, deve-se considerar que, como sequenciadores, os termos

estariam mais gramaticalizados e, portanto, com carga semântica reduzida (ou seja, não funcionariam especificamente como dêiticos).

(30) meu avô por parte de mãe já é falecido e e ele nunca saiu dessa cidade que minha mãe nasceu que é San Stanislau e [a minha vó ela separou separou-se dele e voltou pra Porto Murtinho e mora lá até **aí*/**daí*/***então*** e por parte de pai a minha vó é falecida] (CG25FS-Celina)

Já em (30), o uso de *aí* ou *daí* impossibilitaria a expressão de um valor temporal, equivalente a “hoje”, “agora” (como no caso de *então*), e geraria um período truncado.

O uso anafórico temporal dá origem ao sequenciador ordenativo (Vieira, 2015), em que a ordenação discursiva dos eventos narrados está alinhada com sua respectiva ordem de realização no mundo real.

A sequenciação mantém traços de indicação temporal anafórica, típicas do uso anafórico temporal (Tavares, 1999), ao indicar que o evento mais recente na cadeia da sequenciação alinha-se ao evento anterior, como mostram (31), (32) e (33):

(31) eu estou morando sozinho assim eu já morei sozinho em dois mil e dois por um ano *aí* voltei *aí* em dois mil e oito eu saí de novo (SP33MSVC-JaimeN)

(32) eu saí daqui com vinte e seis anos *daí* morei em outros dois lugares (SP45FSVC-PolianaM)

(33) quando eu vim pra cá eu já tinha meus dois filhos né? ***então*** eu morava no Sumaré (SP77FCPS-VeraD)

Nos casos acima, *aí*, *daí* e *então* estão mais gramaticalizados (relativamente àqueles casos anteriormente apresentados – 20 a 27), dessemantizados e com posição mais fixa na sequência discursiva. Quando a função sequenciadora emerge, nota-se a erosão dos valores espacial e temporal, característicos dos dêiticos. A partir disso, a intercambiabilidade dos itens passa a ser possível, como se vê em (31a), (32a) e (33a):

(31a) eu estou morando sozinho assim eu já morei sozinho em dois mil e dois por um ano *aí/daí/então* voltei *aí/daí/então* em dois mil e oito eu saí de novo

(32a) eu saí daqui com vinte e seis anos *aí/daí/então* morei em outros dois lugar

(33a) quando eu vim pra cá eu já tinha meus dois filhos né? *aí/daí/então* eu morava no Sumaré

Os exemplos originalmente encontrados na amostra (31-33) são diferentes entre si, mas têm uma estrutura e uma temática muito semelhante. Por outro lado, (31a-33a) mostram que o intercâmbio entre as formas *aí*, *daí* e *então* nesses casos não altera a relação discursivo-pragmática entre os enunciados que conectam, em cada um dos exemplos. Trata-se, portanto, de casos que atualizam o esquema abaixo, um percurso de gramaticalização que culmina com um contexto em que as formas desempenham a mesma função e podem, portanto, ser intercambiadas entre si:

dêitico > anáfora (locativa e ou temporal) > sequenciação

Essencialmente, o contexto chamado de *sequenciação* constitui-se de ações ou eventos que se sucedem na linha do tempo crônico¹⁴, de modo que os acontecimentos são enunciados em uma ordem discursiva que obedece à sua ordenação real.

No excerto (34), além do encadeamento de ações, parece haver certa dificuldade em dissociar-se uma relação puramente sequencial de uma noção de causalidade.

¹⁴ Para Benveniste (1974:70), antes de chegar ao nível da expressão linguística, é preciso “transportar” duas etapas, o tempo físico e o tempo crônico. O tempo físico do mundo é infinito, linear, segmentável à vontade do falante e tem uma duração infinitamente variável, que cada indivíduo mede pelo grau de suas emoções e pelo ritmo de sua vida interior. Do tempo físico e de seu correlato psíquico, surge a categoria do tempo crônico, que o autor define como o tempo dos acontecimentos.

(34) o meu estágio é uma iniciação científica né? **então** (tenho)... eu faço uma pesquisa ... trabalho no laboratório fazendo síntese de amostra (CG20FS-Tainá)

A informante, uma estudante de Química, participa de um projeto de iniciação científica na universidade. Ao enunciar sua experiência, ela dá margem a uma dupla interpretação: a) como aluna de iniciação científica, ela também realiza pesquisas em laboratório, ou seja, fazer iniciação científica não necessariamente induz a uma pesquisa em laboratório, mas, no caso dela, isso acontece por conta da natureza do trabalho; b) por ser aluna de iniciação científica, na área de Química, ela necessariamente precisa realizar pesquisas em laboratório.

Neste trabalho, foram desconsideradas da análise quantitativa ocorrências ambíguas como (34). É importante destacar, contudo, que exemplos como este ilustram uma fase transitória entre a sequenciação temporal (Tavares, 2003; Koch, 1987), caracterizada pelo encadeamento de ações e eventos em sua respectiva ordem cronológica, e as sequências discursivas em que uma ação ou acontecimento emerge como resultado de ações ou acontecimentos anteriores, como:

(35) você não conseguir chegar num lugar facilmente né? não conseguir conviver com a cidade **aí**a cidade se torna fechada todo mundo dentro dos carros (SPF26SEL-FabianaB)

(36) se você morasse aqui você ia conhecer tudo aqui **daí** você não ia precisar fazer a pesquisa (CG20MC-Marcelo)

(37) antigamente mulher ficava em casa ela ficava assistindo TV e os homem trabalhava e levava tudo pra elas hoje em dia não... elas saem mais o os homens se interage mais com elas acabou aquela liberdade **então** elas têm que correr mais (CG35MC-Lúcio)

Em (35), FabianaB relata que, em São Paulo, em virtude do trânsito caótico, as pessoas perdem muito tempo no trânsito. No trecho (36), Marcelo relata que

o fato de a documentadora estar realizando uma pesquisa sobre Campo Grande é um indício de que ela não seja nativa. Já em (37), o informante relata que, hoje, diferentemente do que ocorria no passado, as mulheres precisam tomar a iniciativa nos relacionamentos amorosos, uma consequência de sua independência social e financeira.

Nota-se, pois, um *continuum* entre a sequenciação temporal e as sequências discursivas de causalidade, as quais se relacionam tanto à consequência quanto à conclusão. Ao interligar dois eventos que se sucedem cronologicamente, em que o primeiro representa a causa e o segundo a consequência ou conclusão, são preservados resquícios anafóricos temporais.

A definição teórica adotada acerca da causalidade remete a Paiva (1991:108), que diz respeito à ocorrência de um ato explicativo ou consecutivo. No primeiro caso, como os analisados por esta tese, o falante expõe um fato X como origem ou motivação para um fato Y. No segundo caso, um fato é apresentado como consequência de outro, tem-se, então, uma sequência Y-X¹⁵.

Para a gramática tradicional, ao ato de explicar correspondem as orações coordenadas explicativas e subordinadas adverbiais causais; ao segundo, as coordenadas conclusivas e subordinadas consecutivas. Este trabalho dedica-se somente às ocorrências do primeiro tipo, por constituírem a estrutura prototípica de uso de *aí*, *daí* e *então* como sequenciadores discursivos. Não se verificam casos de consecução encadeados por essas formas. Portanto, aqui, a noção de causalidade relaciona-se às tradicionais orações coordenadas explicativas e subordinadas adverbiais causais.

É preciso ressaltar que a causalidade também pode ser inferida a partir da simples justaposição de sequências discursivas, tal como no excerto abaixo:

(38) meus filhos cada um já está no seu canto né? Ø
não justificava ficar naquela casa... tipo assim... já

¹⁵ “De repente o choque é tão grande *que* eu não faço é nada” (CG63MC-Bento). Nesse excerto, a passividade do informante seria resultado de um possível choque. É um tipo de causalidade diferente da analisada nesta tese. No excerto, tem-se um fato que desencadeia uma consequência, cujo sequenciador é *que*. Nos enunciados sequenciados por *aí*, *daí* e *então*, o falante apresenta um fato que motiva determinado efeito.

não usava mais aquela infraestrutura (CG59MS-Vitória)

Em (38), Vitória explica por que não mora mais em uma casa grande e, sim, em um apartamento e expõe como motivo o fato de seus filhos já serem independentes. Nesse trecho, observa-se que a relação de causa e efeito entre as sequências discursivas é não se dá por meio de *aí*, *daí*, *então* ou qualquer outro sequenciador. O uso da forma zero deixa a ideia de causalidade em aberto, exigindo do interlocutor maior esforço cognitivo para recuperar essa relação. Segundo Givón (1995), a justaposição é uma operação própria do modo pragmático, ao passo que a sequenciação é típica do modo sintático. Com base em tal assertiva, foram computadas, nesta tese, somente as ocorrências sintáticas de causalidade, em que há o uso de *aí* *daí* ou *então*, excluindo-se os casos em que a interpretação causal é possível sem o emprego de um desses sequenciadores.

O domínio funcional da causalidade, em sua amplitude, apresenta nuances de significação que se denominam domínios discursivos. Nesse universo, o uso intercambiável de *aí*, *daí* e *então* nem sempre é possível. Dessa forma, mesmo depois de constatada a variação dos itens como sequenciadores causais, ainda é preciso delimitar a natureza do domínio discursivo em que as formas efetivamente são utilizadas sem distinção de sentido.

Nos termos de Marcuschi (2008:155), domínio discursivo relaciona-se com as atividades de linguagem resultantes da diversidade de práticas:

domínio discursivo constitui muito mais uma “esfera da atividade humana” no sentido bakhtiniano do termo, do que um princípio de classificação de textos (...). Constituem práticas discursivas nas quais podemos identificar um conjunto de gêneros textuais que às vezes lhe são próprios ou específicos como rotinas comunicativas institucionalizadas.

Os domínios discursivos da causalidade representam instâncias discursivas em que os usos de *aí*, *daí* e *então* atendem a propósitos específicos. É nesse sentido que se compõe, com base em Sweetser (1991), um quadro do uso variável de *aí*, *daí* e *então* como sequenciadores de causalidade.

Em seu estudo sobre *because*, Sweetser (1991) mostra que a relação de causalidade se manifesta em três domínios: referencial, epistêmico e atos de fala. A autora considera ainda que eles representam um *continuum*, em que a noção de causalidade tende a ser mais abstrata. Para a autora, no domínio referencial, a relação causal é estabelecida entre eventos que, observáveis num mundo possível, estão mais próximos à experiência física, concreta. No domínio epistêmico, diferentemente, a relação causal se estabelece entre um evento e uma conclusão, um julgamento ou uma crença. No domínio conversacional, a relação causal se estabelece entre um ato de fala e sua explicação ou justificativa.

Os domínios, de acordo com a autora, estão ligados à noção de proximidade e afastamento conceitual. Logo, as sequências discursivas do nível do conteúdo estariam mais próximas, em termos conceituais, ao passo que as outras, no nível dos atos de fala, estariam mais afastadas. Em um estágio intermediário, ficaria o nível epistêmico, situado entre o afastamento e a proximidade conceitual das sequências discursivas.

Neves (2000) oferece um tratamento detalhado acerca da relação de causalidade. Após elencar os itens conjuncionais responsáveis pela conexão causal, descreve a relação de causalidade tomada em dois sentidos mais gerais, a saber, o estrito e o amplo. A causa estrita refere-se à relação causa-consequência, ou causa-efeito, entre dois eventos e pressupõe uma anterioridade temporal da causa em relação ao efeito. A causa ampla, por sua vez, pode se manifestar a) no domínio epistêmico, em que a relação causal se pauta em um conhecimento, julgamento ou crença do falante, tradicionalmente denominada “causa formal” e b) nas relações entre um ato de fala e a expressão do que motivou este ato. As relações causais do segundo tipo são “uma relação mais frouxa do que uma relação verdadeiramente causal (em qualquer de suas subespécies, como motivo, razão, justificativa etc.), próximo de uma explicação” (Neves, 2000:806).

De modo geral, Neves (2000) prevê que as relações causais podem ser codificadas entre predicções (estados de coisas); entre proposições (fatos possíveis) ou entre enunciados (atos de fala). A esse respeito, Amorim (2012)

elabora um *continuum*, levando em conta que a relação causal, como uma categoria cognitiva, apresenta nuances em uma escala de abstratização crescente:

**causa efetiva (entre predicções) > causa formal (entre proposições) >
explicação (entre enunciados)**

Sweetser (1991) admite haver limites tênues entre os três tipos de causalidade. Em alguns casos, nem sempre é preciso delimitar com clareza se se trata de um caso referencial ou epistêmico. A delimitação dos excertos que caracterizam o domínio dos atos de fala é, quase sempre, inequívoca. Nesse sentido, faz-se necessário descrever cada um dos domínios.

Esse tipo de ocorrência é caracterizado por verbos de evento¹⁶, que descrevem acontecimentos associados a uma mudança de estado do referente, como em (39), em que “assistir TV até tarde” é um fato da realidade, que desencadeia a reação de alguém, outro acontecimento ligado ao mundo:

(39) eu gosto de ficar assistindo tevê até tarde ficar
no telefone às vezes conversando até tarde **aí** ele
implica um pouco (SP24CPML-DalsonO)

O segundo domínio é caracterizado por estruturas argumentativas, em que o sequenciador não é empregado para encadear ações e acontecimentos do mundo. Não se trata de uma causalidade no sentido sociofísico, mas resultante da subjetividade e da argumentação, em que o sequenciador atua no fechamento de uma ideia:

(40) a gente está acostumado com o movimento de
São Paulo que é vinte e quatro horas de segunda a
segunda [se você quiser tem bar tem música tem
movimento tem restaurantes **então** ...pra quem
gosta de movimento São Paulo é melhor né?]
(SPF48CES-MeireC)

¹⁶ Caracterizados como aqueles que exprimem mudança do sujeito no tempo (Paiva, 1995).

A informante descreve a vida agitada da capital paulista e, em seguida, conclui que essa é uma rotina que agrada mais às pessoas que preferem cidades movimentadas. Nesse caso, *então* sequencia uma informação subjetiva do informante, diferentemente de (40), em que se tem um fato que não está condicionado pela opinião de uma pessoa.

O último caso trata-se de uma estrutura em que o sequenciador encadeia sequências discursivas cujos atos de fala são distintos, sendo que o segundo geralmente constitui uma sugestão ou um pedido:

(41) eu acho muita falta de respeito com a gente
dizer que a cidade é isso é aquilo... lá não é bom?
então vai pra onde veio... lá tem lugar é bom fica lá
(CG51MS-Pedro)

A relação de causalidade em (41) é ainda menos direta que em (39) e (40): neste caso, tem-se um ato de fala que motiva outro ato de fala. Ao ouvir as reclamações de pessoas de outros estados em relação a Campo Grande, mas que vivem e desfrutam dos serviços da capital, o informante sugere que elas voltem para seus lugares de origem.

No domínio dos atos de fala, o que se tem não é um fato do mundo, nem uma opinião, mas um acontecimento que precisa ser depreendido de uma ação e que não chega a se realizar efetivamente. De acordo com o que é proposto na teoria dos atos de fala (Searle, 1981; Levinson, 1983), todo ato de fala é classificado a partir de um grupo de quatro “condições de felicidade” (condição preparatória; conteúdo proposicional; condição de sinceridade; condição essencial), podendo uma delas ser posta em negociação para a realização do ato de fala. É o que se verifica em (40), excerto em que a sequenciação de causalidade, levada a cabo por *então*, caminha para a concretização de um evento.

Nesta análise, opta-se pela dicotomização dos padrões discursivo-pragmáticos da causalidade, sintetizados em causa estrita e causa alargada. Tal dicotomia baseia-se na proposta de Neves, descrita acima, segundo a qual as construções causais podem agrupar-se em dois blocos: o das causais propriamente

ditas e o das causais explicativas ou de enunciação. O primeiro grupo é composto pelos enunciados referentes aos domínios referencial e epistêmico.

A inclusão desses dois domínios em uma única categoria justifica-se pela proposta de Sweetser (1991), que alega serem estreitos os limites entre eles, fato que ajuda a compreender a existência de casos ambíguos (Sweetser, 1991:77), em que se observam nuances do domínio referencial e do domínio epistêmico. Por outro lado, as causais explicativas ou de enunciação dizem respeito aos casos em que a relação causal é estabelecida no domínio dos atos de fala. Logo, neste trabalho, causa estrita refere-se aos enunciados em que a causalidade se encontra nos domínios referencial e epistêmico, ao passo que causa alargada é utilizada para os casos em que a causalidade reside no domínio dos atos de fala.

Ainda sob essa perspectiva, um sequenciador pode ser usado para formular um conteúdo (nível do *dictum*) ou para pontuar um processamento textual (nível do *modus*). Pode-se dizer que esses dois níveis se correlacionam ao que Halliday & Hasan (1976) denominam como função ideacional e função interpessoal da linguagem. O primeiro grupo engloba eventos entendidos como a representação da realidade externa, que o falante experiencia e seria representado pelo domínio referencial; o segundo diz respeito aos eventos linguísticos, que se referem à organização que o falante imprime a seu discurso, caso em que a coesão deve ser entendida como uma relação entre significados, resultante da representação enunciada pelo falante durante a interação. A este grupo, estão ligados os domínios epistêmico e atos de fala.

A partir da proposta de domínios discursivos (Sweetser, 1991), atesta-se a variabilidade de *aí*, *daí* e *então*, com base na permuta dos sequenciadores nos excertos (42), (43) e (44):

(42) às vezes você não pode... porque você não tem condição de morar naquele naquele lugar né? perto do seu trabalho *aí/daí/então* você tem que ir pra um outro bairro (CG45MC-Zuleica)

No exemplo (42), *aí*, *daí* e *então*, ao serem intercambiados mantêm o sentido do período. Os três itens, utilizados em sua forma mais gramaticalizada,

sequeNCiam uma relação de causalidade que diz respeito a fatos do mundo (Pezatti, 2001).

Com respeito ao domínio referencial, procede-se à permuta dos itens em excertos cujos sequeNCiadores efetivamente utilizados foram *daí* e *então*:

(43) se você morasse aqui ia conhecer tudo aqui
aí//daí/então não ia precisar fazer a pesquisa
(CG20MC-Marcelo)

(44) e fica na boa quem pode tem duas empregadas
motoristas *aí//daí/então* fica fácil (SPM63CPN-
RodolfoR)

Nos trechos acima, *daí* e *então* sequeNCiam fatos da realidade e, mais uma vez, nota-se que eles podem ser substituídos um pelo outro e também por *aí*, sem alteração da significação original. Em (45), exemplo típico do domínio epistêmico, embora *aí* e *daí* sejam formas possíveis no contexto, nota-se que se trata de usos pouco expressivos e idiossincráticos de tais sequeNCiadores, o que pode sinalizar a progressiva especialização de *então* como sequeNCiador de causalidade epistêmica.

(45) a gente está acostumado com o movimento de
São Paulo que é vinte e quatro horas de segunda a
segunda [se você quiser tem bar tem música tem
movimento tem restaurantes *aí//daí/então* ...pra
quem gosta de movimento São Paulo é melhor né?]
(SPF48CES-MeireC)

No domínio epistêmico, as ocorrências com *aí* e *daí* se revelam raras e são encontradas apenas diante do verbo *acho* (*aí//daí acho*, *aí//daí eu acho*):

(46) olha não sei que eu não conversei com você
você está só perguntando mas se a gente sentasse
conversasse *aí//daí/então* eu acho que eu saberia te
responder porque eu sou daqui mas não sei
(CG28FS-Débora)

Tais usos compõem outro quadro de variação, distinto daquele que aqui se descreve, no qual as formas vêm desacompanhadas. Além disso, os usos dos

sequeenciadores com *acho* e *eu acho* dizem respeito somente ao domínio epistêmico, o que restringiria a análise apenas a esse domínio. O emprego de *então* como sequeenciador de atos de fala permite levantar a hipótese de que tal forma possa estar mais gramaticalizada que *aí* ou *daí* e, por isso, seria comumente utilizada para sequências em que a relação de causalidade tem cunho mais pragmático do que propriamente semântico (Sweetser, 1991).

Verifica-se, assim, que somente *então* é utilizado em sequências discursivas caracterizadas como atos de fala – fato que licencia a conclusão de que, quanto menos concreta a relação de causalidade, menos factível se torna o emprego de *aí* e *daí*. As tabelas 1 e 2 apresentam os números de ocorrências dos sequeenciadores nos diferentes domínios de causalidade, tanto amostra campo-grandense, quanto na paulistana:

Tabela 1 - Distribuição geral dos dados em Campo Grande

Campo Grande							
Domínio discursivo	<i>Aí</i>	%	<i>Daí</i>	%	<i>Então</i>	%	Total
Referencial	78	22	18	5	261	73	357
Ambíguos	3	22	4	28	7	50	14
Epistêmico	3	18	2	13	11	69	16
Atos de fala	0	0	0	0	3	100	3
Total	84		24		282		390

Tabela 2 - Distribuição geral dos dados em São Paulo

São Paulo							
Domínio discursivo	<i>Aí</i>	%	<i>Daí</i>	%	<i>Então</i>	%	Total
Referencial	56	16	12	4	281	80	349
Ambíguos	2	18	3	27	6	55	11
Epistêmico	3	19	4	25	9	56	16
Atos de fala	0	0	0	0	1	100	1
Total	61		19		297		377

Conforme essas tabelas permitem visualizar, em ambas capitais, *então* é a forma predominante em todos os domínios. De fato, o envelope de variação se verifica no domínio epistêmico, mas, além de serem poucos os dados de *aí* e *daí* nesse domínio, em ambas as capitais, deve-se reiterar que tais ocorrências vêm sempre acompanhadas de estruturas como *(eu) acho*. Logo, não se justificaria uma análise quantitativa.

Por outro lado, no domínio dos atos de fala, além de pouquíssimos dados (comparativamente aos outros dois domínios), não se observa variação de fato, já que todas as ocorrências encontradas são casos de *então*:

(47) [eu acho muita falta de respeito com a gente dizer que a cidade é isso é aquilo... lá não é bom? **aí/*daí/então* vai pra onde veio]... lá tem lugar é bom fica lá (CG66MS-Durval)

Dentre as escassas ocorrências de *então* em atos de fala, observa-se a impossibilidade de substituição de tal sequenciador por *aí* ou *daí*, como mostra o excerto acima. A substituição de *aí* ou *daí* por *então* geraria, em casos como (46), sequências discursivas agramaticais. Como foi demonstrado nas Tabelas 1 e 2, no domínio dos atos de fala, o emprego de *então* é categórico.

A baixa incidência de *aí* e *daí* no domínio dos atos de fala é uma evidência de que tais formas parecem se ater a relações de causalidade mais concretas, como

as do domínio referencial. Ao substituir *então* por *aí* ou *daí*, a sequência “vai pra onde veio” soaria como uma narração, não como uma sugestão decorrente de um fato relatado anteriormente, o que resultaria em uma quebra na estrutura da causalidade.

Há ainda casos em que há mais de uma interpretação possível – algo que, como visto mais acima, é típico de processos de gramaticalização. Parece ser o caso de (48):

(48) qualquer coisa às vezes eu largo a moto lá e volto de táxi depois eu pego porque é pertinho dez reais entendeu? depois eu vou de táxi *aí/daí/então* fica mais tranquilo também (SP375EO-FredericoV)

Em (48), o informante enuncia sua preferência por deixar sua moto em casa e pegar um táxi para ir para voltar do trabalho. Com isso, o trajeto fica mais tranquilo. A interpretação da relação causa-efeito, contudo, pode dar-se mediante duas perspectivas: (a) a referencial e, nesse caso, a facilidade proporcionada pelo táxi seria uma consequência concreta da escolha do informante; (b) a epistêmica – aqui, o efeito advindo do falante seria uma mera apreciação enunciada por ele. Ao optar pelo táxi, ele acredita que o percurso seja mais tranquilo (casos como esse caracterizam-se como um uso situado no plano das ideias e não dos eventos objetivos do mundo. Casos como estes, considerados como *ambíguos*, foram excluídos das análises quantitativas.

Diante das análises qualitativas desenvolvidas nesta seção, bem como diante da distribuição dos dados nas duas capitais (Tabelas 1 e 2), a análise quantitativa de dados dessa tese circunscreve-se em torno dos casos em que *aí*, *daí* e *então* são empregados na sequenciação do domínio referencial da causalidade. Assim, os casos em que esses sequenciadores operam causalidade no domínio epistêmico são rarefeitos, o que não justifica sua análise quantitativa. Já para o domínio dos atos de fala, a variação não se verifica possível e, de fato, não é atestada nos dados.

A análise quantitativa desses dados, portanto, visa a verificar se, no domínio referencial, uma das três formas acaba sendo preferida (numa das duas capitais ou em ambas). Interessa, em outras palavras, verificar se se trata de um caso de

variação estável ou se se observa uma mudança em favor de uma das formas (visto que *então* é a forma preferida no domínio epistêmico e é a única que se verifica do domínio dos atos de fala). Uma pergunta geral, dessa forma, é sobre se os usos de *aí* e *daí* estariam se expandindo, ainda que *então* pareça uma forma polifuncional (recorrente nos três os domínios da causalidade).

2

Funcionalismo e Sociolinguística: uma interface teórica

Este capítulo traça uma breve história do Funcionalismo em Linguística, a fim de demarcar qual o lugar da gramaticalização nessa escola. Em seguida, expõem-se os princípios da teoria laboviana. A partir de tais considerações, delineia-se um programa de pesquisa próprio, o Sociofuncionalismo, resultante da convergência de pressupostos do Funcionalismo e da Sociolinguística Variacionista. A heurística de tal programa é destinada ao estudo de fenômenos em gramaticalização, analisados sob uma perspectiva variável, tais como os sequenciadores *aí*, *daí* e *então* no domínio da causalidade referencial.

1. Funcionalismo

De acordo com Heine et al (1991), a noção de gramaticalização, cara ao funcionalismo, foi primeiramente reconhecida no Oriente. Os autores afirmam que desde o século X, os escritores chineses já distinguem entre símbolos linguísticos “cheios” e símbolos “vazios”.

Faraco (2004) destaca os estudos do juiz inglês William Jones, no século XVIII, que se dedicou ao sânscrito. Segundo as investigações de Jones, o grego e o latim guardavam grande semelhança com o foco de sua pesquisa sobre o sânscrito. Diante de tal constatação, extraíram-se importantes conclusões de suas pesquisas.

A gramaticalização parece estar no cerne dos estudos histórico-comparativistas, pelo menos em sua concepção mais clássica, que prevê um contínuo de mudanças ao longo do tempo.

É verdade que a gramaticalização se apresenta como um instrumento de investigação que vai além de comparações entre as línguas, entretanto não se pode

deixar de ver aí também um ponto de semelhança entre as antigas investigações filológicas e os atuais estudos funcionalistas.

Segundo Rosário (2010), ainda no século XVIII, os filósofos franceses Etienne Bonnot de Condillac e Jean Jacques Rousseau afirmavam que os lexemas concretos teriam originado vocábulos abstratos e há indícios de que tais estudiosos foram precursores ao considerar que as flexões verbais, como os sufixos, teriam vindo de palavras independentes mais antigas. Segundo Heine et al (1991), podem ter sido esses estudos a fonte de pesquisa dos gramáticos comparativistas para as pesquisas sobre o indo-europeu.

Foi Meillet (1948 [1912]) que cunhou o termo gramaticalização, primeiramente utilizado em seu trabalho denominado *L'évolution des formes grammaticales*. É necessário afirmar, contudo, que seu mérito não está somente na criação do termo, mas também no fato de ter justificado a relevância dos estudos de gramaticalização como uma das maiores atividades na ciência da linguagem.

Meillet (1948 [1912]) insistiu na ideia de *continuum*, bastante utilizada até os nossos dias, para expressar a transição de itens lexicais (*mots principaux*) para auxiliares e outros morfemas com função gramatical (*mots accessoires*), também chamadas de “palavras vazias” (*mots vides*). Também é uma contribuição do autor francês a ideia de que o aumento de frequência de uso está em correlação inversa à perda do valor expressivo das palavras.

Outra área que merece destaque é a da língua falada. Comumente se associam os estudos funcionalistas às investigações nesse âmbito. Segundo Pezatti (2004), a Análise da Conversação foi inaugurada em meados da década de 1970. Foi, contudo, entre 1856-1857 que surgiu o primeiro estudo de uma língua indo-europeia feito a partir da fala e não de textos escritos, um marco para os estudos linguísticos dos séculos XX e XXI.

Na segunda metade do século XX, por volta de 1960, surge a Sociolinguística, alguns anos antes do Funcionalismo norte-americano. O estudo de fatores sociais para explicar alternância entre formas com o mesmo significado, todavia, remonta a uma época mais remota. Os estudos de Schuchardt (1885), linguista austríaco, geraram importantes desdobramentos para o nascimento da teoria laboviana.

Outra fonte de que se beneficiou o Funcionalismo foi a Escola Linguística de Praga, de 1926, fundada por Jakobson e Nikolaj Trubetzkoy, que reconheciam sua estreita relação com o estruturalismo saussureano, ao negar a distinção entre linguística sincrônica e diacrônica e a homogeneidade do sistema linguístico.

A perspectiva funcional da sentença é um dos principais interesses do Círculo Linguístico de Praga e se traduz por um certo tipo de embate entre critérios sintáticos e enunciantivos. Para Lyons (1970), embora a ordenação das palavras seja vista como uma questão de sintaxe, a estrutura sintática dos enunciados resulta da situação comunicativa em que é pronunciada e, de modo mais específico, pelo que já é aceito pelos falantes de uma determinada comunidade e pelas informações novas apresentadas ao ouvinte.

Algumas décadas depois, Mathesius (1982) propõe um modelo para as pesquisas de Danes, Firbas e Halliday, a fim de formular uma concepção de sintaxe que pudesse descrever uma mesma sequência discursiva em vários níveis. Estabelecem-se, assim, os conceitos de marcação e não-marcação (Ilari, 2004).

Cunha et al (2003) recuperam o princípio da marcação e o associam a subprincípios de iconicidade, compreendida como a correlação entre forma e função, entre o código linguístico (expressão) e seu *designatum* (conteúdo).

Nas sequências discursivas causais, o princípio de iconicidade pode ser analisado a partir dos pressupostos atinentes à ideia de causa. Um desses pressupostos é o de sequencialidade temporal, segundo o qual os fatos obedecem a um eixo de anterioridade e posterioridade, causas precedendo efeitos. Esse pressuposto de que a causa antecede o efeito pode ser verificado, por exemplo, na refutação de argumentos a partir do raciocínio de que X não é (e não pode ser) a causa de Y, pois Y ocorreu antes de X.

Segundo Givón (1976), é a dimensão temporal dos fatos que autoriza a apreensão de relações causais, embora, por si mesma, não seja suficiente para determiná-la. A relação causal é um salto qualitativo baseado em um processo de indução de existência de razão condicional entre fatos frequentemente associados no tempo.

Um corolário natural da associação entre causa e tempo é o de que o conceito de causa se restringe à relação entre eventos, uma vez que esses se

definem em função de transformações no eixo do tempo e requerem sempre um ponto de origem. Em certo sentido, causa é um evento que provoca uma mudança de estado, ou seja, um outro evento (Van Dijk, 1984, Mann & Thompson, 1985). Uma definição prototípica de causa implica relação entre dois eventos ocorridos em tempo passado. É possível estabelecer, porém, relação entre um evento e um estado ou processo, considerando que estes também possuem um ponto de origem, uma motivação.

Aplicado aos enunciados causais, o princípio de iconicidade autoriza a hipótese de que a intrínseca relação entre causa e tempo encontra sua realização mais natural na ordenação causa-efeito, ou seja, no paradigma *A logo B*. Essa configuração sintática, conforme a conceptualização da relação causal, reproduz um pressuposto acerca da organização da realidade.

No caso dos sequenciadores discursivos estudados, a hipótese que se tem é a de que iconicidade pode ser manifestada, por exemplo, por meio da natureza da informação introduzida pelo sequenciador. Em linhas gerais, *então*, forma de maior extensão fônica, estaria mais propensa a introduzir informações novas e, portanto, de maior complexidade cognitiva, como em:

(49) já vi muita gente falar também né? de ser assaltada
tudo *aí/daí/então* a gente tem receio disso né? (SP48FCE5-
MeireC)

No excerto acima, a informante fala sobre os assaltos rotineiros que presencia e, ao introduzir o efeito de tal enunciado, utiliza *então*. O que se observa é que a informação posterior a *então* tem um conteúdo diferente daquilo que vinha sendo tratado por MeireC. Logo, considera-se como informação nova, como será descrito no Capítulo 4.

Em contrapartida, quanto menor a extensão do sequenciador, menor seria a complexidade da informação introduzida, ou seja, maior a previsibilidade de tal informação. Logo, seria mais provável que se utilizasse *aí* ou *daí*.

(50) eu fiz uma aposta com o meu filho.. se ele tirasse se
ele conseguisse tirar nota máxima em todas as matérias...
ele ganharia... um videogame... pensei que ele não ia

conseguir ele conseguiu... *aí/daí/então* eu tive que... gastar mil reais em um videogame (SP36MSPS-PedroS)

(51) se tiver dinheiro evidente né? é óbvio que a gente está falando de ter dinheiro no bolso senão São Paulo deve ser o terror também... *aí/daí/então* não tem nada pra oferecer (SP60MCEs-RômuloS)

Em (50), PedroS utiliza *aí* para relatar a consequência da aposta que fez com o filho. A sequência discursiva que sucede o sequenciador *aí* não rompe com o assunto anterior, sobre o videogame. Pelo contrário, continua a tratar do tema. Por isso, considera-se que, nesse caso, tal forma sequencia uma informação velha. Situação semelhante é verificada em (51), em que RômuloS discorre sobre as condições necessárias para que uma pessoa possa usufruir das atrações da cidade de São Paulo. Ao utilizar *daí*, o informante enuncia uma causalidade que pode ser inferida a partir de seu comentário anterior. Esse excerto, portanto, caracteriza-se como um caso intermediário entre (49) e (50), já que a informação que acompanha *daí* não é nem totalmente nova, nem recuperável no contexto.

Segundo Lyons (1970), são muito diversificadas a terminologia e a interpretação dadas aos vários tratamentos funcionalistas no âmbito da Escola de Praga. Há, porém, um ponto em comum entre elas: o de que a estrutura dos enunciados é determinada pelo uso que lhes é dado e pelo contexto comunicativo em que ocorrem. Tal asserção é a mola propulsora de todas as pesquisas desenvolvidas desde essa época até os dias de hoje, no funcionalismo linguístico, quando se inicia a introdução do caráter interacional da comunicação humana nos estudos da linguagem.

Cabe esclarecer que existem diversas facetas que revestem a abordagem funcionalista da língua. Segundo Nichols (1984) há três tipos de funcionalismos: o conservador, o extremado e o moderado. Tais orientações funcionais assinalam as grandes tendências de estudo em prol de uma gramática funcional e não de teorias funcionalistas propriamente. O que aproxima as diferentes vertentes de investigação é a visão compartilhada de que a língua sofre adaptações de suas formas para servir a funções comunicativas. O funcionalismo do tipo conservador

não propõe análises da estrutura linguística, sendo que apenas manifesta as inadequações do formalismo e do estruturalismo.

Por outro lado, o tipo moderado não apenas aponta para estas inadequações, mas também propõe uma análise funcional da estrutura. A proposta de gramática sistêmico-funcional de Halliday, e outros como Valin e Dik representam esta corrente. Além disso, Neves (1997) indica que as obras mais recentes de Givón (1984; 1995) também são de cunho moderado. Já o funcionalismo extremado nega a estrutura tal como ela é e confere que as regras se baseiam internamente na função, sem que haja restrições sintáticas.

Integram este último grupo, funcionalistas como Sandra Thompson, Paul Hopper, Givón, Traugott e Bybee, a partir da década de 1970. Nesse mesmo viés, na Alemanha, autores como Lehmann, Heine, Claude, Hünemeyer também se destacam (Gonçalves et al, 2007).

Este trabalho se insere nesta última corrente, cujas noções funcionalistas a serem apresentadas são atribuídas, principalmente, a Givón (1984; 1995). Tal postura se justifica porque, nesta direção, a língua é tomada como um instrumento de comunicação e sua estrutura não pode ser entendida sem referência ao contexto e às funções comunicativas. No que diz respeito a esta tese, a classificação de *aí*, *daí* e *então* como sequenciadores variáveis no domínio da causalidade referencial resulta da verificação dos usos de tais formas em uma amostra consideravelmente extensa, resultante de um gênero específico, a entrevista sociolinguística.

Esta compreensão advoga uma linguística baseada no uso que, segundo Martelotta & Areas (2003), relaciona o contexto e as situações extralinguísticas. Nessa perspectiva:

a sintaxe é uma estrutura em constante mutação em consequência das vicissitudes do discurso. Ou seja, a sintaxe tem a forma que tem em razão das estratégias de organização da informação empregadas pelos falantes no momento da interação discursiva. Dessa maneira, para compreender o fenômeno sintático, seria preciso estudar a língua em uso, em seus contextos discursivos específicos, pois é nesse espaço que a gramática é constituída (MARTELOTTA & AREAS, 2003:24)

Isso permite entender que os falantes não inventam livremente sequências sonoras para os novos “rótulos” linguísticos, mas que tendem a utilizar os recursos já existentes na língua (Martelotta & Areas, 2003:25). Segundo Cunha (2009), como a linguagem é uma faculdade humana, tem-se a hipótese de que a estrutura linguística revela o funcionamento da mente e as propriedades da conceitualização humana do mundo. A iconicidade é definida como a correlação natural e motivada entre forma e função, ou seja, entre o código linguístico (expressão) e seu significado (conteúdo) (Cunha, 2009).

Givón (1995) propõe que as funções linguísticas têm natureza cognitiva (representação mental da experiência). Assim, a estrutura da gramática está condicionada ao *continuum* da forma e da função no discurso. Logo, língua e gramática não podem ser descritas nem explanadas adequadamente como um sistema autônomo.

No conceito de gramática, estão envolvidos parâmetros naturais que constituem a atividade comunicativa, tais como: cognição e comunicação, processamento linguístico e cerebral, interação sociocultural, mudança e variação, aquisição e evolução. Dessa forma, a gramática e o discurso constituem uma simbiose, pois a estrutura gramatical se constrói nos arranjos e estratégias oriundas das necessidades de comunicação.

Dessa perspectiva, a visão funcionalista da linguagem pode ser resumida, segundo Givón (1995), a partir das seguintes premissas: a) a linguagem é uma atividade sociocultural; b) a estrutura serve a funções cognitivas e comunicativas; c) a estrutura é não arbitrária, motivada e icônica; d) mudança e variação são constantes; e) o sentido é contextualmente dependente; f) as categorias não são discretas; g) a estrutura é flexível, não rígida; h) as gramáticas são emergentes; i) há exceções nas regras de gramática.

Esta tese está situada em um campo em que a linguagem é entendida como um instrumento de interação social, que busca nos contextos semântico-discursivos as motivações para os fatos da língua (Cunha et al, 2003). Esta abordagem objetiva explicar as regularidades no uso linguístico a partir de seu contexto de realização.

2. Gramaticalização: uma abordagem funcionalista

O presente trabalho parte da concepção de que a gramática de uma língua natural nunca é estática e acabada: se considerada de modo sincrônico, a gramática de qualquer língua apresenta padrões regulares, rígidos e, ao mesmo tempo, padrões que não são completamente fixos, mas fluidos. Padrões novos (ou inovadores) podem se estabilizar e resultar em uma reformulação da gramática. Nesse sentido, a gramática é um “sistema adaptativo”: é parcialmente autônoma e responde a pressões externas ao sistema.

Assim, a gramática passa a ser compreendida como sistema constituído pelas regularidades que resultam das pressões do uso. Tais pressões estão relacionadas a necessidades discursivas e pragmáticas relacionadas a propósitos comunicativos do falante, no intuito de ser expressivo e informativo. Logo, o desenvolvimento de novas estruturas gramaticais decorre de pressões comunicativas e da existência de conteúdos cognitivos para os quais não há designações linguísticas pertinentes.

É o que se verifica com *aí*, *daí* e *então*, que, por meio da gramaticalização, deixam de atuar simplesmente como dêiticos espaçotemporais e passam a articular sequências discursivas com sentido causal (entre outros), constituindo um envelope de variação no domínio referencial da causalidade.

De acordo com Neves (1997), o conceito de gramaticalização começou a ser formulado no século X, mas somente no século XX Meillet o definiu como “a atribuição de um caráter gramatical a uma palavra anteriormente autônoma”.

Na década de 70, Givón lançou a máxima “a morfologia de hoje é a sintaxe de ontem”, bem como “a sintaxe de hoje é o discurso pragmático de ontem”, a fim de demonstrar que as línguas seguem um ciclo que pode ser sintetizado em três etapas: discurso > sintaxe > morfologia. Tal conclusão emerge depois que o linguista, ao analisar formas verbais africanas, depreende que os afixos de hoje remontam a arranjos de pronomes com verbos independentes.

Hopper e Traugott (1993), por sua vez, definem gramaticalização como o processo pelo qual itens e construções lexicais passam, em determinados contextos linguísticos, a desempenhar funções gramaticais, e, uma vez gramaticalizados, continuam a desenvolver novas funções gramaticais. De modo geral, o fenômeno pode ser descrito como um processo pelo qual um item sai do léxico para integrar a gramática.

Portanto, a gramaticalização pode ser considerada como resultado da rotinização. De acordo com Castilho (1997), ao se gramaticalizar, as unidades linguísticas perdem em complexidade semântica, liberdade sintática e em substância fonética:

gramaticalização é o trajeto empreendido por um item lexical, ao longo do qual ele muda de categoria sintática (=recategorização), recebe propriedades funcionais na sequência discursiva, 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. (CASTILHO, 1997:31).

Em relação às diferentes funções que um item linguístico pode exercer no decorrer do processo de gramaticalização, que vão desde as funções referenciais até as funções mais expressivas – passando pelas funções textuais – Traugott (1982) sugere um esquema de gramaticalização que recupera basicamente o trajeto que um item linguístico, situado no domínio referencial, percorre até chegar ao domínio expressivo (pragmático) da língua, que inclui os participantes:

Proposicional	>	(Textual)	>	Expressivo
(referência)		(Coesão)		(atitude do falante)

De acordo com o esquema acima, nota-se que a gramaticalização é concebida como processo dinâmico e unidirecional em que formas lexicais adquirem, ao longo do tempo, propriedades linguísticas novas, tais como atributos gramaticais. Os itens *aí*, *daí* e *então* no PB, cujos usos dêiticos aludem ao domínio proposicional (usos referenciais), ao se gramaticalizarem, passam pelo domínio textual (como sequenciadores), até chegarem ao domínio mais expressivo da língua, exercendo funções como a de marcadores discursivos.

Hopper & Traugott (1993) salientam que as palavras de “conteúdo” (palavras lexicais) distinguem-se das palavras funcionais (palavras gramaticais), e afirmam que as primeiras (substantivos, verbos e adjetivos) são usadas para informar ou descrever coisas, ações e qualidades, ao passo que as últimas indicam as relações dos nomes uns com os outros (preposições), interligam partes do discurso (sequenciadores), reportam se os participantes do discurso já foram

identificados ou não (pronomes e artigos), e mostram se os objetos estão próximos do falante ou do ouvinte (demonstrativos) (Hopper & Traugott, 1993).

O processo de gramaticalização que os autores descrevem perpassa pela ideia de que o vocábulo segue um percurso unidirecional. De acordo com esse princípio, pode-se prever um cline de gramaticalização, como sugerem Hopper & Traugott (1993:7), a saber:

[item de conteúdo] > [palavra gramatical] > [clítico] > [afixo flexional]

Esse *continuum* também pode ser representado por classes de palavras como:

[categoria maior] > [categoria medial] > [categoria menor]

Segundo os autores, a categoria maior diz respeito a nome, verbo e pronome; a categoria medial, ao adjetivo e ao advérbio; por fim, a categoria menor, à preposição, à conjunção e aos auxiliares. Bybee et al (1994:4) ressaltam que a teoria da gramaticalização começa com o pressuposto de que morfemas gramaticais se desenvolvem a partir de morfemas lexicais ou da combinação de morfemas lexicais com outros morfemas lexicais ou gramaticais. Em linhas gerais, Bybee et al (1994) observam não só a passagem de um item lexical a um item gramatical, mas também o processo diacrônico da mudança das formas linguísticas.

2.1. Princípios de gramaticalização segundo Hopper (1991)

São vários os princípios que explicam processos de gramaticalização de itens linguísticos. Entre os estudiosos que têm se dedicado a definir tais princípios, destacam-se Heine & Reh (1984) – uns dos primeiros linguistas a tentar estabelecer princípios gerais de gramaticalização para as línguas naturais –, bem como Lehmann (1995 [1982]), que se preocupa com os estágios mais avançados do item gramaticalizado, e Hopper (1991), que se ateve principalmente aos mecanismos relacionados aos estágios iniciais dos processos de gramaticalização.

Os princípios de gramaticalização de Hopper (1991) são o respaldo teórico para o envelope de variação aqui verificado – o uso variável dos sequenciadores *aí*, *daí* e *então* no domínio funcional da causalidade. Como itens intercambiáveis, eles resultam da atuação desses quatro princípios: a) *estratificação* – apresentam a mesma função sequenciadora; b) *divergência* – embora tenham origem comum, os itens desenvolvem diferentes funções ao longo do tempo; c) *descategorização* – *aí*, *daí* e *então* exibem mudança em relação à categoria original de advérbio e d) *persistência* – seriam mantidos resquícios dêiticos de *aí*, *daí* e *então* em sua função sequenciadora.

Esses princípios serão explanados com mais profundidade nas subseções a seguir.

2.1.1. Estratificação (layering)

Dentro de um domínio funcional, novas formas estão constantemente emergindo, sem que as formas antigas necessariamente desapareçam. Desse modo, itens diferentes passam a assumir vários usos e cada um deles corresponde a uma camada em domínio funcional específico. Segundo Hopper (1991), exemplos de domínios funcionais são tempo, aspecto, modalidade, caso, referência e as camadas que o constituem podem ser representadas por itens lexicais, classes particulares de construções ou registros sociolinguísticos, ou seja, podem ser expressas por diferentes significados ou por alternância de estilos.

O princípio de estratificação, segundo Braga et al (2001), Longhin-Thomazi (2003) e Gonçalves et al (2007), se refere à constatação de que a gramaticalização não implica obrigatoriamente ou necessariamente a substituição de um item por outro. Pelo contrário, em um mesmo recorte sincrônico, diferentes formas em processo de gramaticalização podem conviver, dando, assim, espaço à “variação linguística” (Longhin-Thomazi, 2003:34). Hopper exemplifica esse princípio com o caso do pretérito em inglês, no qual uma camada antiga de alternância de vogais, em verbos como *drive/drove*, *take/took*, coexiste com uma camada mais recente, em que o tempo passado é marcado morfologicamente com /t/ ou /d/, como em *notice/noticed*, *log/logged*.

2.1.2. Descategorização

Este princípio refere-se à diminuição ou perda do estatuto categorial dos itens gramaticalizados. Por exemplo, um verbo, enquanto lexical, tem propriedades sintáticas e semânticas, como o número de argumentos implicados, a categoria morfossintática e a função semântica desses argumentos, além das restrições de seleção para sua realização lexical. Quando se gramaticalizam, os verbos assumem atributos das categorias secundárias e perdem a propriedade de, por exemplo, selecionar argumentos com os quais vão se combinar.

O estudo de Omena & Braga (1996), sobre a alternância das formas *a gente* e *nós*, é um bom exemplo para ilustrar o princípio de descategorização definido por Hopper. De acordo com as autoras, no estágio mais gramaticalizado, a forma *a gente* tende a levar o verbo para a terceira pessoa do singular (ou mesmo para a primeira pessoa do plural, incluindo a variação “a gente vamos”), comportando-se, dessa forma, claramente como um pronome. Nesse caso, verifica-se a passagem do item “gente”, proveniente, em princípio, um substantivo, para a classe dos pronomes.

Neves (2002) aborda a questão dos nomes genéricos, passíveis de serem interpretados como nome ou pronome substantivo indefinido. É o caso de *coisa* em frases como *Coisa de louco*. A autora também comenta a flutuação de nomes concretos que podem indicar quantificação: um pinga de vergonha na cara, rios de dinheiro, uma senhora barriga. Com esse princípio, o que se observa é a perda gradativa de traços sintáticos e morfológicos da forma linguística, tal como ocorre com o verbo *parecer* (Gonçalves, 2003:222), que, aos poucos, vai perdendo as restrições de tempo/modo verbal e também de concordância.

Além disso, como não há nenhuma força que obrigue uma forma a se gramaticalizar, o processo de gramaticalização pode ser iniciado, finalizado ou, então, simplesmente interrompido num dado estágio de gramaticalização, sem chegar ao estágio final. Um outro aspecto importante que se pode depreender de Hopper (1991) é que, no decorrer do processo de gramaticalização de um dado item linguístico, outras formas podem conviver ou até mesmo competir com esse

item, podendo, posteriormente, se especializar na expressão de uma determinada noção ou função.

Os usos dos itens *aí*, *daí* e *então*, quando passam a exercer funções textuais (como sequenciador, principalmente), tendem a conviver com outras formas conjuncionais da língua, chegando, por vezes, até a substituir outros conectivos no Português falado. Um exemplo é uso frequente de *aí* em trechos narrativos para exercer essa função. Por fim, na proposta de Hopper (1991), os princípios de estratificação (diferentes formas podem coexistir) a existência de usos ambíguos e de sobreposição de funções dos itens linguísticos durante o processo de gramaticalização.

2.1.3. Divergência

Identifica-se tal princípio quando uma forma lexical sofre gramaticalização (por exemplo, quando um verbo se transforma em um auxiliar, um clítico ou um afixo), mas ainda é utilizada em sua função mais antiga. Em outras palavras, a forma lexical original permanece como um elemento autônomo, suscetível às mesmas mudanças de um item lexical comum. De acordo com esse princípio, é possível haver várias formas etimologicamente comuns, mas funcionalmente diferentes. A forma gramaticalizada pode ser fonologicamente idêntica à forma lexical, que permanece autônoma.

Nesse caso, a forma gramaticalizada e a(s) forma(s) lexicais correspondentes podem coexistir. Como exemplo, há o uso da partícula negativa *pas* do francês, que, paralelamente ao seu uso como partícula negativa, continua sendo empregada como nome “passo”. O resultado desse princípio aponta para a existência de várias formas que divergem funcionalmente, apesar de conterem uma etimologia comum. O mesmo ocorre com o verbo *estar*, usado como auxiliar, como cópula e como verbo intransitivo, seguido de locativo, além de integrar a perífrase de gerúndio (Mendes, 2005).

No caso de *só que*, observa-se divergência na medida em que “apesar de constituir uma perífrase, a forma original *só* permanece como elemento autônomo” (Longhin-Thomazi, 2003:178).

2.1.4. Persistência

A persistência diz respeito à permanência de traços do significado da forma original; por mais que a forma mude e assuma novos significados, permanecem alguns vestígios de usos anteriores.

Segundo Gonçalves et al (2007), esse princípio pode ser aplicado à noção de coletividade expressa pelo substantivo *gente*, que fica retida na forma gramaticalizada *a gente*, contribuindo para sua referência de indeterminação.

O princípio da persistência prevê, segundo Hopper (1991), a permanência de alguns traços de significado da referência extralinguística da categoria lexical (forma-fonte) na forma gramaticalizada (forma-alvo), o que pode influenciar as delimitações sintáticas da nova forma. De acordo com Lopes (2010: 278), é possível estender esse princípio de Hopper (1991) para além dos traços semânticos, visto que “os detalhes da história lexical do item não se restringem aos valores semânticos, mas também às suas características formais”.

2.1.5. Especialização

Ocorre quando a construção emergente deixa de ser uma escolha a mais na língua e passa a ser a forma mais utilizada. Segundo Hintze (2003), esse princípio pode ser aplicado às denominações referentes aos homens que desempenham funções sacerdotais. Mesmo dispondo de vários termos como *cura*, *monsenhor*, *vigário* e *pároco*, em que cada título revela uma função sacerdotal determinada, no uso popular, a forma *padre* é a preferida para fazer referência a todos eles, sem especificar as funções exercidas pelos religiosos.

Este princípio em geral se observa nos estágios finais da gramaticalização, quando o uso de uma forma gramaticalizada se torna obrigatório. Nesse caso, a forma ‘concorrente’ não é necessariamente eliminada ou empregada para cobrir as funções dos demais itens, mas sim empregada em certas funções ou contextos específicos ligados a um mesmo domínio. Mais uma vez, uma boa ilustração é o caso da partícula *pas* do francês, que começou a ser usada nos contextos de *ne*, passando por verbos de movimento até se tornar de fato uma partícula de negação usada com todos os tipos de verbo.

Atualmente, somente a partícula *pas* tem sido usada na língua falada para indicar negação, atingindo, portanto, o estágio de obrigatoriedade. No PB, Longhin-Thomazi (2003:178) defende que a perífrase *só que* pode ser definida como uma “forma de especialização” apenas nos casos em que tal perífrase é empregada no lugar de *mas*, em especial nos contextos de quebra de expectativa e focalização.

Partindo do princípio da especialização, acredita-se que *aí* e *daí* sejam formas específicas do domínio referencial da causalidade, ao passo que *então* atua como um sequenciador polifuncional, o que justifica seu emprego nos três domínios analisados (Capítulo 1).

2. Mecanismos de gramaticalização

Os estudos sobre gramaticalização publicados nos últimos tempos revelam que o processo de mudança linguística é sempre motivado, conforme apontam Meillet (1948 [1912]) e Heine et al (1991), por alguma força de natureza cognitiva e pragmática. Dessa forma, esta seção revê alguns processos que propiciam a ocorrência do processo de gramaticalização nas línguas naturais: a metáfora, a metonímia, a reanálise e a analogia.

A justificativa para considerar a reanálise e a analogia como mecanismos de mudança vem de Hopper & Traugott (1993), que destacam que esses dois processos estão constantemente envolvidos em casos de mudança que se encontram além do nível fonético-fonológico.

2.1. Metáfora e metonímia

Nos estudos sobre gramaticalização, a transferência metafórica e a transferência metonímica, de natureza discursivo-pragmática, são elencadas como responsáveis pelo desenvolvimento de itens lexicais em itens gramaticais (Hopper & Traugott, 1993). Segundo Heine et al (1991), a transferência metafórica implica que conceitos mais complexos e abstratos podem ser descritos ou entendidos a partir de conceitos mais concretos ou menos complexos.

Nesse caso, a transferência semântica acontece por meio da similaridade de percepções de sentido (Traugott & König, 1991). Utilizando-se dessa noção de

transferência entre espaços cognitivos, Heine et al (1991) definem a gramaticalização como um fenômeno que envolve transferência de sentidos entre categorias cognitivas básicas, como: pessoa > objeto > atividade > espaço > tempo > qualidade, que destaca a similaridade entre fontes e alvos.

Assim, a observação dessas propriedades cognitivas permite reconhecer que um conceito só é fonte em relação a outro mais abstrato, visto aqui como reflexo das experiências humanas mais elementares.

Por outro lado, a transferência metonímica é caracterizada como um contínuo de pequenas mudanças motivadas pelas relações de contiguidade entre formas-fonte e formas-alvo. Em outras palavras, diferentemente do mecanismo metafórico, a transferência metonímica envolve a especificação de um significado em termos de outro que está presente no contexto, o que implica uma transferência semântica por contiguidade. Os exemplos abaixo, são, segundo Heine et al (1991), mais bem analisados como casos de relações metonímicas e não como relações metafóricas:

(i) From Cologne to Vienna it is 10 hours by train (De Colônia a Viena são 10 horas de trem)

(ii) To get to Vienna, you travel from morning to evening
(Para chegar a Viena, você vai viajar de manhã até a noite).

Em (i), as preposições *from* e *to* possuem um significado locativo, ao passo que em (ii) elas apresentam um significado temporal. Heine et al (1991) destacam, no entanto, que na transição de espaço para tempo, pode existir um estágio intermediário ou vários estágios ambíguos. Em sua análise sobre as perífrases conjuncionais *agora que* e *já que* no Português Brasileiro, Galbiatti (2008) constata a mesma configuração de ambiguidade em alguns contextos, em específico no que tange ao funcionamento da perífrase *já que*, que apresenta usos não muito claros (evidentes), localizando-se na fronteira entre as categorias de tempo e causa.

Um mecanismo que está ligado à metonímia é a inferência por pressão de informatividade (Traugott & König, 1991), processo em que, devido à convencionalização de implicaturas conversacionais a partir do contexto de uso, o item lingüístico passa a assumir um valor novo, inferido a partir do valor original.

Assim, uma implicatura, comumente motivada por determinada forma linguística, pode ser tomada como parte do significado desta, podendo inclusive chegar a substituí-la.

Essas semelhanças permitem que as formas transcendam o domínio dêitico e atinjam o anafórico. O resultado é a atuação desses elementos no domínio textual de forma semelhante àquela no plano da dêixis. Essa transição dêixis - anáfora representa também a transferência metafórica entre os domínios, de modo que a organização espacial do mundo concreto passa a ser utilizada para organizar o universo mais complexo do texto, exatamente como ocorre com aí, daí e então, que, durante o processo de gramaticalização, passam por *bleaching*, ou dessemantização e adquirem propriedades funcionais, como a da sequenciação.

2.2. Reanálise e analogia

Além das questões apresentadas na seção anterior, Hopper & Traugott (1993), entre outros, relacionam ainda o mecanismo da metonímia ao fenômeno da reanálise, envolvido em mudanças estruturais mais locais e sintagmáticas, e o mecanismo da metáfora ao fenômeno da analogia, envolvido em mudanças paradigmáticas (mudanças superficiais).

O fenômeno da analogia, conforme se mencionou anteriormente, já tinha sido analisado por Meillet (1948 [1912]), de quem também se pode depreender uma definição de reanálise. Similarmente à metáfora e à metonímia, que são bastante caros à gramaticalização a reanálise e a analogia, definidas como mecanismos de natureza estrutural, podem ser considerados processos complementares.

A reanálise, conforme afirmam Harris & Campbell (1995), modifica a estrutura subjacente de uma expressão ou classe de expressões, envolvendo reorganização e mudança de regras lineares, sintagmáticas, frequentemente locais. O resultado disso é a alteração gradual da relação entre constituintes, que resulta na mudança da estrutura hierárquica e mudança categorial.

Por outro lado, a analogia diz respeito à atração entre formas e construções já existentes, envolvendo, assim, a reorganização paradigmática, com mudança

nas colocações de superfície (plano da expressão) e nos padrões de uso da forma linguística.

Ao rearranjar fórmulas linguísticas, a reanálise conduz a uma nova interpretação das relações entre os elementos que as constituem. Assim, a percepção de como os constituintes de língua estão inter-relacionados muda para o falante e novos “cortes” passam a existir.

Esse é um mecanismo caro a *aí*, *daí* e *então*, já que a ambiguidade observada no uso de tais formas no percurso de gramaticalização induz os usuários a novos *inputs*, que resultam em propriedades funcionais como a da sequenciação.

2.3. Iconicidade e marcação

Na Linguística Funcional, o princípio da iconicidade diz respeito à motivação na relação entre o plano da expressão e o do conteúdo. A proposição seria a de que a forma da língua deve refletir a função que exerce ou ser restringida por ela (Givón, 1990). Sendo assim, a codificação morfossintática é, na maioria dos casos, resultado do uso da língua, das experiências dos falantes. Como a linguagem é uma faculdade humana, supõe-se que a estrutura linguística revela as propriedades da conceitualização humana do mundo ou as propriedades da mente dos homens em geral.

Daí surge a concepção de língua como estrutura maleável, sujeita às pressões de uso e constituída de um código não totalmente arbitrário; e de gramática, como um conjunto de estratégias utilizadas para efetuar uma comunicação coerente. Desse modo, a estrutura da língua é entendida como o resultado de fenômenos não linguísticos, especialmente de processos cognitivos. Para Bolinger (1977), a condição natural da língua era preservar uma forma para um sentido, e vice-versa, princípio denominado isomorfismo.

Entretanto, estudos sobre os processos de variação e mudança linguística, ao constatar mais de uma forma para dizer uma mesma coisa, fizeram com que essas ideias de Bolinger (1977) fossem revistas. Na língua do cotidiano, e principalmente na língua escrita, existem muitos casos em que não há uma correspondência clara entre forma e conteúdo. Num estudo sincrônico, determinadas estruturas revelam elevado nível de opacidade em relação aos

papéis que desempenham. Nesse sentido, há, em certos momentos, correlação entre uma forma e várias funções, ou entre uma função e várias formas.

Segundo Givón (1990), em sua versão mais branda, o princípio da iconicidade se manifesta em três subprincípios: I) Subprincípio da quantidade: a) quanto maior a quantidade de informação, maior será a quantidade de forma utilizada na codificação dessa informação; b) quanto mais imprevisível (mais nova) for a informação para o interlocutor, maior será a quantidade de forma utilizada; c) quanto mais relevante for a informação em termos de continuidade temática ou discursiva, maior será a quantidade de forma requerida para codificá-la. II) Subprincípio da proximidade ou integração: os conceitos que estão mais integrados conceitual, funcional ou cognitivamente também se apresentam com maior integração morfossintática, já que serão postos mais próximos no nível da codificação – o que está mentalmente junto coloca-se sintaticamente junto.

Haiman (1983) afirma que a distância entre expressões linguísticas deve ser um índice iconicamente motivado da distância conceitual entre os termos ou eventos que elas denotam. Assim: a) A ordem das orações no discurso coerente tenderá a corresponder à ordem temporal de ocorrências dos eventos descritos; b) A informação mais importante tenderá a ser colocada em primeiro lugar (Hopper & Traugott, 1993).

Na forma e na organização, as estruturas sintáticas não devem ser muito diferentes das estruturas semânticas e cognitivas. McMahon (1994) afirma que há uma relação entre iconicidade e gramaticalização. Segundo a autora, isso ocorre pois, em um primeiro momento, certas entidades linguísticas são analogicamente aproveitadas para expressar uma determinada ideia em um contexto funcional novo. Devido a esse fato, tais entidades tendem a ser estruturalmente conjugadas, o que vai refletir o grau de união cognitiva entre elas e consequentemente, pressionar a formulação de um novo sintagma. Por via de repetição de uso, posteriormente, essa construção se regularizará, gramaticalizando-se.

De acordo com a autora, paradoxalmente, à proporção que uma construção se gramaticaliza, com maior integração entre seus elementos, estes tendem a tornar-se iconicamente menos transparentes, perdendo muito de sua

própria identidade. É o que ocorre, por exemplo, com formas que se desenvolvem para afixos ou clíticos.

Essas formas, por generalização de seu uso, tornam-se opacas, a ponto de, em certos contextos, criarem dificuldade de reconhecimento de sua função inicial. Essa opacidade é denominada *desbotamento icônico*, ou seja, um estado avançado de gramaticalização em que o elemento afixado perde em termos de liberdade sintática e conteúdo lexical.

Herdado do estruturalismo da Escola de Praga, o princípio da marcação estabelece três critérios para se distinguir binariamente entre as categorias marcadas e as não marcadas, segundo Givón (1995): (a) Complexidade estrutural: A estrutura marcada tende a ser mais complexa (ou maior) do que a correspondente não marcada; (b) Frequência de distribuição: A estrutura marcada tende a ser menos frequente, portanto, cognitivamente mais saliente, do que a categoria não marcada correspondente; (c) Complexidade cognitiva: A estrutura marcada tende a ser cognitivamente mais complexa – quando se fala de esforço mental, atenção demandada e tempo de processamento - do que a estrutura não marcada correspondente.

Segundo Martelotta (2009), a noção de marcação é particularmente importante em função daquilo que ele chama fluxo básico de mudança motivado pelo uso. Um item ou uma construção começa a ser usada com muita frequência em um determinado contexto (distribuição de frequência típica da forma não-marcada). Pela força do hábito, o falante passa a trabalhar mais automaticamente com o item ou a construção (nível de complexidade cognitiva típica de forma não-marcada).

Com isso o item ou a construção perde estrutura fonética (nível de complexidade estrutural típico de forma não-marcada). Percebe-se assim, que há uma tendência, nas línguas em geral para que esses três critérios ocorram ao mesmo tempo. A correlação entre eles é o reflexo mais geral da iconicidade na gramática, o que representa o isomorfismo entre os correlatos de natureza comunicativa e cognitiva e os formais de marcação. Desse modo, as categorias que

são estruturalmente mais marcadas tendem a ser substantivamente mais marcadas também.

Para Givón (1995) a marcação é um fenômeno dependente do contexto, pois uma mesma estrutura pode ser marcada em um e não-marcada em outro. Assim, deve ser explicada com base em fatores cognitivos, biológicos, comunicativos e socioculturais. Outro fato citado pelo autor é que a marcação não se limita apenas às categorias linguísticas, podendo se estender a outros fenômenos.

Segundo o princípio da marcação, tamanho diminuto e grande recorrência são dois fortes indícios de graus baixos de complexidade. É esperado que as formas menos marcadas adquiram mais encargos gramaticais, por serem percebidas pelos usuários da língua como pouco complexas quanto ao processamento, o que facilita seu emprego como parte de construções diversas. *Aí* e *daí* parecem se encaixar nesse caso.

3. Sociolinguística Quantitativa

O dinamismo é inerente a todas as línguas, uma vez que elas estão sujeitas à variação. Por isso, coexistem formas linguísticas alternantes em todos os níveis de análise: fonético-fonológico, morfossintático, lexical e também no nível pragmático-discursivo.

A Sociolinguística Variacionista determina, pois, que a língua seja considerada nas situações concretas de uso. Sem seus componentes extralinguísticos, ela passa a ser apenas uma abstração, o que inviabiliza o estabelecimento de qualquer ligação entre ela e o perfil social da comunidade linguística que a utiliza. É como sintetiza Calvet (2002) “as línguas não existem sem as pessoas que as falam, e a história da língua é a história de seus falantes”.

De acordo com Romaine (1994), o termo *Sociolinguística* surgiu na década de 1950 a partir das perspectivas de linguistas e sociólogos sobre a relação entre linguagem e sociedade e sobre o contexto em que se dá a variação linguística. Segundo Marcuschi (1975), a expressão foi cunhada em 1952, a partir de um

estudo que H.C. Currie havia realizado acerca do comportamento linguístico e da reação social que ele desperta.

A língua não é propriedade do indivíduo, mas da comunidade (é social). Entretanto, Labov discorda de Saussure, Chomsky e outros que insistem na homogeneidade necessária do objeto linguístico, que ignoram a heterogeneidade e que consideram a fala como caótica e desmotivada (Figuerola, 1996). Labov (2008 [1972]:259) entende que o então novo modo de fazer linguística é estudar empiricamente as comunidades de fala.

Esse argumento pode ser entendido à luz da rejeição da psicologia individual como um modelo de referência para a linguística, em como rejeição do idioleto ou gramática individual como o objeto da linguística (Figuerola, 1996: 79-80). De acordo com Labov (2001), todos os sociolinguistas concordam que produções e interpretações de um falante não são o lugar primário da investigação linguística nem as unidades finais da análise, mas os componentes usados para construir modelos de nosso objeto primário de interesse, a comunidade de fala.

Por sua vez, Saussure considerava a língua como um sistema homogêneo, regular e estático, do qual se abstraía qualquer espécie de variação. No século XX, o gerativismo de Chomsky se insurgiu contra o estruturalismo, defendendo a competência criativa do falante no uso da língua. Chomsky rechaçava a concepção mecanicista, herdada do behaviorismo de Skinner, que insistia na aquisição da língua como um processo de imitação estimulado por reforços positivos. A teoria gerativista se ateve à relação entre língua e mente, desconsiderando, a exemplo de Saussure, qualquer fator extralinguístico.

Nesse cenário, o estudo de Labov (1963) na ilha de Martha's Vineyard desvia-se da tradição de estudos estritamente linguísticos, ao investigar a variação dos ditongos /ay/ e /aw/, inaugurando uma corrente teórica que se volta à correlação entre a variação linguística e os fenômenos sociais, questionando-se, inclusive, sobre a significação social local da variante centralizada desses ditongos. A proposta mais geral de Labov consiste em depreender o sistemático no aparente caos linguístico, demonstrando que a variação é ordenada e padronizada.

A Teoria da Variação orienta-se por dois princípios: (a) identificar o caráter heterogêneo da língua; (b) compreender que as gramáticas nas quais uma mudança linguística acontece representam gramáticas de toda uma comunidade de fala (Weinreich; Labov & Herzog, 2006).

Para Labov, a mudança é uma tendência natural de todo sistema linguístico e tem, como elementos desencadeadores, os processos de variação. Toda mudança é resultado de uma variação, mas nem toda variação origina uma mudança:

os procedimentos da linguística descritiva se baseiam no entendimento de que a língua é um conjunto estruturado de normas sociais. No passado, foi útil considerar que tais normas eram invariantes e compartilhadas por todos os membros da comunidade linguística. Todavia, as análises do contexto social em que a língua é utilizada vieram demonstrar que muitos elementos da estrutura linguística estão implicados na variação sistemática que reflete tanto a mudança no tempo quanto os processos sociais extralinguísticos. (LABOV, 1968:241)¹⁷.

O termo *sociolinguística* foi aceito com relutância por Labov, já que, para ele, não seria possível considerar a língua fora do contexto social. Ao estudar a língua em uso, a Sociolinguística se volta para o tratamento da variação linguística, abordando-a sob dois eixos: o da sincronia e o da diacronia. O primeiro explica os fenômenos de variação a partir da análise de um conjunto de variáveis situadas num mesmo período do tempo. O último descreve as formas resultantes das transformações linguísticas transcorridas ao longo do tempo, baseando-se na premissa de que só se pode elucidar a estrutura de uma língua se for compreendida a história de sua formação.

A associação entre o sincrônico e o diacrônico é a base do Princípio do Uniformitarismo, segundo o qual as forças que estimulam as mudanças do presente são as mesmas que provocaram as mudanças já operadas na língua (Labov, 1994).

¹⁷ The procedures of descriptive linguistics are based upon the conception of language as a structured set of social norms. It has been useful in the past to consider these norms as invariants, shared by all members of the speech community. However, closer studies of the social context in which language is used show that many elements of linguistic structure are involved in systematic variation which reflects both temporal change and extralinguistic social processes (LABOV, 1968:241)

Para elucidar o processo de variação/mudança, a Sociolinguística estuda o sistema linguístico em correlação com a heterogeneidade social. Por isso, trabalha com variáveis como faixa etária, grau de escolaridade, classe social, gênero, dentre outros.

A variação pressupõe a coexistência de formas linguísticas diferentes que competem entre si num determinado momento e que podem originar: (a) variação estável: nesse caso, conclui-se que o quadro de variação tende a se manter por um longo período, pois não se observa a predominância de uma variante linguística sobre outra(s). Apresenta um padrão não-curvilíneo; (b) mudança em progresso: quando o processo de variação culmina com o uso de uma das variantes, que se generaliza, tornando-se praticamente categórica dentro da comunidade de fala. Desse modo, a(s) outra(s) variante(s) viria(m) a cair em desuso.

As observações em tempo aparente conectadas às observações em tempo real permitem que se verifique a mudança em progresso. A articulação entre presente e passado permite evidenciar estágios variáveis e mudanças que aconteceram (tempo real) e que estão em curso (tempo aparente). Via variação pode-se captar a direção e algumas generalizações acerca da mudança.

De acordo com Faraco (2005), a mudança não se refere à troca direta e abrupta de um elemento por outro, mas envolve sempre uma fase de concorrência. Da variação entre duas formas para a codificação de uma mesma função/significação, uma pode se fixar na função tornando a outra obsoleta, embora nem sempre seja esse o caso.

A teoria laboviana procura responder também a seguinte questão: como se explicaria a diversidade linguística dentro de uma comunidade de fala se não fosse possível sistematizar a variação linguística? Para tanto, a variação linguística na fala de cada indivíduo é tomada como consequência do contato com outros membros da comunidade em que ele está inserido. Guy (2001) define a comunidade de fala como um grupo de falantes que compartilha as mesmas normas linguísticas. Nessa perspectiva, a língua não pode ser homogênea, pois é um espelho da diversidade sociocultural da comunidade que a utiliza.

Segundo Labov (2008 [1972]), é preciso não só analisar a variação condizente a uma determinada comunidade, mas também identificar os fatores que a desencadeiam. Para descrever os fatores condicionadores da variação, é necessário sistematizar quantitativamente os dados da fala da comunidade. A partir desta análise, registra-se a probabilidade e o percentual de ocorrências de cada um dos fatores envolvidos no processo de variação.

Conforme Tarallo (2010), as variantes de uma comunidade de fala estão sempre concorrendo entre si: padrão x não-padrão, conservadoras x inovadoras; estigmatizadas x de prestígio. A variedade padrão costuma ser também conservadora e muito valorizada socialmente, embora essa associação não seja unilateral.

A mudança pressupõe um período de transição, em que estruturas linguísticas diferentes competem entre si. No início desse processo, o contato entre os membros da comunidade linguística propicia a propagação de formas inovadoras, de modo que, pouco a pouco, elas entrem em contato com as formas de mais prestígio. Finalmente, quando uma variante se sobrepõe à outra, a mudança se consolida. Labov (1994) afirma que a mudança se expande de modo proporcional ao contato entre os falantes. Quando a mudança linguística se finda, após um longo período, a nova forma linguística se implementa no sistema e pode representar baixo ou elevado *status* social. Cabe, aqui, perguntar, se, em Campo Grande e São Paulo, há mudança em progresso na direção de *então* e qual o status social dos três sequenciadores.

É importante notar que a língua opera por dois princípios aprioristicamente paradoxais: aparenta ser estável, ao mesmo tempo em que está propensa à variação e à mudança, que lhe são imanentes. A morosidade da mudança pode levar à crença de que não é possível vê-la acontecer. Em contrapartida, se seu ritmo fosse acelerado, a eficácia comunicativa poderia ficar comprometida.

Os pressupostos labovianos consideram a língua como uma forma de comportamento social e, a exemplo das demais formas de comportamento, a língua é influenciada pelo contexto da interação linguística. Elementos como a

hierarquia entre os falantes, a situação comunicativa e, principalmente, o gênero dos indivíduos atuam de forma decisiva na seleção das variantes linguísticas.

Por outro lado, a variação linguística não se motiva apenas por razões sociais. Em circunstâncias diferentes, uma mesma pessoa pode optar por variantes linguísticas igualmente distintas. Para isso, ela precisa ter conhecimento das variedades não padrão e culta para que esteja apta a selecionar qual se ajusta melhor ao seu contexto interacional.

Essa possibilidade refuta o pressuposto saussureano de que o falante teria um papel passivo diante do sistema linguístico, o qual se constituiria independentemente de sua influência e de sua vontade. Trabalhando em uma direção contrária, a Sociolinguística Variacionista confere autonomia ao falante, inserindo-o no processo de evolução linguística.

Assim, a variação e a mudança linguísticas, embora sejam condições fundamentais para a vitalidade de qualquer língua, não são fortuitas ou irrefreáveis.

Labov (2008 [1972]) resumiu em alguns pontos o percurso da mudança linguística: (a) qualquer traço linguístico relativo a uma comunidade de fala encontra-se em oposição a outro dialeto; (b) um traço utilizado pelo grupo simboliza sua identidade social e reage a pressões externas; (c) depois de finalizado o ciclo da mudança linguística, estabelece-se uma nova norma; (d) a nova norma passa a ser utilizada por grupos vizinhos, e é repassada de geração a geração.

Weinreich, Labov & Herzog (2006) afirmam que, para a mudança linguística se efetuar com sucesso, é preciso considerar cinco problemas: o dos fatores condicionantes (*constraints*); o da transição (*transition*); o do encaixamento (*embedding*); o da avaliação (*avaluation*) e o da implementação (*actuation problem*).

O problema da restrição diz respeito ao conjunto de variáveis que podem contribuir para o processo de mudança linguística, bem como as possíveis mudanças que podem ocorrer. Os linguistas anteriores à Sociolinguística tinham uma visão rígida em relação à *função*, pois consideravam que ela seria desempenhada por uma única forma linguística. A Sociolinguística, ao contrário, admite a possibilidade de que duas ou mais formas desempenhem uma mesma

função, a exemplo dos sequenciadores *aí*, *daí* e *então* no domínio referencial da causalidade.

Embora a concepção de heterogeneidade sistemática tenha sido enfatizada e comprovada pela Sociolinguística, os estruturalistas afirmavam haver assistemática no processo de transição linguística. A transição caracteriza-se como a mudança de um estado da língua a outro, típico da mudança. Nesse sentido, um mesmo falante emprega uma ou outra forma, sem se ater ao fato de que a língua está mudando. Nesse estágio, uma forma alternativa, como *então*, passa a ter usos novos, como o da causalidade referencial, até ser a forma preferida em todos os contextos, como já demonstrado no Capítulo 1 no tocante a *então*, espécie de sequenciador coringa nos três domínios da causalidade considerados.

O problema do encaixamento, operando sobre os eixos linguístico e social, considera que uma análise estritamente linguística não é suficiente para descrever a mudança. Comparando-se a frequência dos fenômenos estudados, em São Paulo e Campo Grande, será possível verificar, quantitativamente, se há diferença em relação ao ritmo do encaixamento de *aí*, *daí* e *então* nas duas capitais. Em síntese, o que se propõe é saber se o encaixamento da variante *então*, que parece ser uma forma *polifuncional*, tende a revelar maior difusão de mudança em um grande centro (Zilles, 2007), como São Paulo, e se, em uma localidade menor, como Campo Grande, seu ritmo é mais lento.

O problema da avaliação diz respeito à atitude da comunidade de fala face à mudança linguística. Para a Sociolinguística, as primeiras etapas da mudança não pertencem ao nível da consciência social, fazendo com que os falantes não se deem conta da evolução da língua. Nos estágios subsequentes, conforme afirma Labov (1982), as inovações estilísticas passam a ficar mais evidentes e, nas últimas fases da mudança, quando a sociedade já pode percebê-las, surge a avaliação inconsciente das formas linguísticas por parte dos falantes.

O problema da implementação resume-se à direção assumida pela mudança na estrutura social. Labov (1994) afirma que os grupos intermediários são caracterizados pelo uso mais recorrente de formas linguísticas padrão. A esse respeito, cabe perguntar se falantes de 35 a 49 anos, em Campo Grande e de 35

a 59, em São Paulos utilizariam com mais frequência as formas *aí* e *daí*, mais prototipicamente utilizadas no domínio referencial da causalidade.

4. Sociofuncionalismo

A natureza do fenômeno aqui estudado é mais discursiva que morfossintática, a escolha por *aí*, *daí* e *então* vai além da mera sugestão de “manutenção do significado referencial”, ou “do mesmo estado de coisas”, como sugere Labov (1978), ao final de seu debate com Lavandera (!978). A variável aqui analisada requer um enquadramento teórico que justifique “uma mesma função discursiva, o que permite a investigação dos contextos que favorecem uma forma em detrimento da outra no desempenho de uma mesma função no discurso – o da causalidade no domínio referencial. Essa perspectiva resulta em uma interface entre Funcionalismo e teoria variacionista, que tem se mostrado produtiva em pesquisas a que se vem denominado de Sociofuncionalismo.

O Sociofuncionalismo pode recobrir diferentes tipos de enfoque que envolvam pressupostos da Sociolinguística e do Funcionalismo Linguístico, daí a necessidade de se delimitar os campos dessas áreas a serem considerados numa proposta de interface. A abordagem sociofuncionalista tratada aqui articula pressupostos teórico-metodológicos da Sociolinguística Variacionista e do Funcionalismo de vertente norte-americana.

Bybee (2012) afirma que a teoria baseada no uso se desenvolveu diretamente e é, em certo sentido, apenas um novo nome para o funcionalismo norte-americano, representando uma extensão das abordagens desenvolvidas na Linguística Cognitiva e na Linguística Funcional.

Após examinar uma série de fenômenos variáveis de natureza fonológica e morfológica (e mesmo sintática), Labov (2008 [1972]) defende que “os resultados favorecem a visão (neogramática) de que a mudança linguística é [...] mecânica”. No clássico estudo da passiva sem agente, Weiner e Labov (1978) afirmam que o uso da variante passiva é condicionado sintaticamente, por um efeito mecânico (paralelismo formal). Sob essa prisma, há um problema a ser contornado: se os

condicionadores são estruturais ou mecânicos, qual o espaço para hipóteses funcionais? Na visão de Camacho (2003), “as posições assumidas por Labov impedem o alinhamento da sociolinguística variacionista com qualquer tipo de enfoque funcional.

Do lado do Funcionalismo, também existem obstáculos a serem superados. Se assumida uma perspectiva funcionalista radical, como a de Bolinger (1977) – de que a condição natural da linguagem é preservar uma forma para um significado e um significado para uma forma, numa correlação biunívoca entre forma e função (versão forte do princípio da iconicidade) -, não há conciliação teórica possível, pois simplesmente não haveria variação.

Segundo Labov (1978), o principal objetivo da teoria sociolinguística é prever a distribuição provável de informação nos níveis fonológico, prosódico, morfológico, sintático etc., de modo a se obter um retrato da estrutura gramatical da língua. Em nota, Labov (1982) reafirma que o termo gramática é usado num sentido geral, para indicar o sistema linguístico como um todo, incluindo a fonologia, o léxico e sua organização semântica. Assim, ao se estender o nível linguístico de análise, fenômenos variáveis de natureza funcional podem vir a se constituir em objeto de investigação sociolinguística.

Se fenômenos de natureza discursiva são recortados como objeto de investigação e diferentes grupos de fatores funcionais são controlados como possíveis condicionadores da variação, é natural e esperado: que se reflita sobre a viabilidade teórico-metodológica de uma interface; que se levantem pontos de divergência entre as duas teorias-mães; que se pense na possibilidade de um “terceiro programa, com núcleo e heurísticas próprias” (Borges Neto & Müller, 1987).

Neves (1999) menciona que foi no PEUL que o termo sociofuncionalismo surgiu, em referência a estudos que procuravam integrar diretrizes da sociolinguística e do funcionalismo com o objetivo de analisar tendências de uso variável como sendo reflexo da organização do processo comunicativo.

De acordo com Bybee (2012), a variabilidade inerente é um grande ponto de contato entre o Funcionalismo Linguístico e a Sociolinguística Variacionista. A

autora acredita que, em contraste com a Linguística Histórica, a Sociolinguística tende a incorporar métodos e conceitos da linguística baseada no uso, visto que se volta mais diretamente ao estudo da variabilidade inerente da língua (Bybee, 2012).

Entre tais pressupostos, destaca-se o uso linguístico, uma das pedras angulares tanto da Sociolinguística quanto do Funcionalismo. Para ambas as teorias, o que deve ser alvo das investigações é a língua em uso, em detrimento de qualquer idealização de como a língua deveria ou poderia ser usada (Hopper, 1987; Labov, 2008[1972]; Poplack, 2001). Estudos feitos em ambas as perspectivas vêm trazendo evidências de que a mudança é disseminada gradualmente ao longo do espectro linguístico e do espectro social, com incrementações contínuas em termos de frequência de uso (Labov, 2001; Hopper & Traugott, 2003; Bybee, 2012).

Apesar de uma possível complementaridade entre dados de épocas distintas para a construção de um quadro mais detalhado no que diz respeito à mudança linguística, pesquisadores afiliados às duas teorias consideram que a melhor fonte para a análise linguística são os dados atuais. A razão é que esses dados permitem uma observação mais direta e completa de um maior número de ocorrências com condições de uso mais facilmente recuperáveis. Com base em dados atuais, é possível, inclusive, a construção de hipóteses acerca de sincronias passadas.

É preciso considerar que a concepção de uniformitarismo encontrada em estudos funcionalistas foi declaradamente emprestada de Labov, que é citado, a esse respeito, por Traugott e König (1991) e Hopper e Traugott (2003), entre outros.

Na perspectiva funcionalista, a frequência de uso é tida como fundamental para o estabelecimento e a manutenção da gramática, e, além disso, acredita-se que a difusão linguística e social da mudança pode ser captada através do aumento da frequência de formas inovadoras em diferentes contextos. Para a Sociolinguística variacionista, o aumento de frequência também é compreendido como índice de difusão linguística e social, e as variantes devem ter certa recorrência para que possam ser comparadas por meio de instrumental estatístico.

Além disso, acredita-se que a difusão linguística e social da mudança pode ser captada através do aumento da frequência de formas inovadoras em diferentes contextos. Na perspectiva variacionista, o aumento de frequência também é compreendido como índice de difusão linguística e social, e as variantes devem ter certa recorrência para que possam ser comparadas por meio de instrumental estatístico.

Outra semelhança entre as teorias em apreço no que tange à mudança linguística reside nas propostas acerca da disseminação. Ambas as teorias defendem que a mudança se espalha de forma gradual ao longo do espectro social, levando-se em conta fatores como região, geração, classe social, etnia etc. Além disso, consideram que é comum haver diferença entre falantes mais velhos e mais jovens, no caso de mudança em progresso.

É justamente a relação entre o processo de gramaticalização e o fenômeno de variação morfossintática que é apontada, por diversos pesquisadores, como um dos pilares que sustenta um duplo olhar funcionalista e sociolinguístico sobre os fenômenos de variação e mudança linguística. Tagliamonte (2002), por exemplo, afirma que a gramaticalização “necessariamente produz variabilidade na gramática e essa variabilidade reflete as camadas variáveis da gramaticalização representadas por diferentes formas.

Tanto da ótica da Sociolinguística quanto do Funcionalismo, fatores de natureza interacional têm papel importante na variação e na mudança linguística. No âmbito da sociolinguística, Labov (1978) compreende a variação estilística como uma adaptação da linguagem do falante ao contexto imediato do ato de fala. No âmbito do Funcionalismo, Traugott (1991) não só defende que a mudança é motivada por práticas discursivas e sociais, como acredita que os estudos funcionalistas de gramaticalização orientados para o falante podem contribuir para o estudo sociolinguístico da variação intrafalante.

Estudos feitos em uma perspectiva sociofuncionalista costumam seguir passos de análise que podem ser, grosso modo, assim sintetizados: (i) identificação de situações de uso linguístico variável dentro de um domínio funcional; (ii) operacionalização da noção laboviana de variável, isolando formas variantes que

desempenhem uma mesma função dentro de um domínio funcional; (iii) testagem de grupos de fatores diversos para identificar os contextos (linguísticos, discursivos, estilísticos, sociais) de uso das formas; (iv) detalhamento de cada grupo de fatores buscando captar mudanças em curso ainda sutis (considerando inclusive sobreposição de funções), e posterior amalgamação de fatores em busca de generalizações; (v) interpretação da frequência das formas em determinados contextos como indício de: (a) perda de espaço de uma das variantes, (b) generalização de significado (os itens expandem seus contextos de uso), ou (c) especialização de uso (os itens adquirem significados mais específicos restritos a certos contextos dentro do domínio).

A abordagem sociofuncionalista traz, para o tratamento da variação linguística, um controle bastante refinado de grupos de fatores linguísticos, com a inclusão de restrições de natureza discursivo-pragmática (planos discursivos, status informacional dos referentes, graus de integração etc.), que podem receber, de início, tratamento analítico escalar, sujeito a posteriores amalgamações.

Por outro lado, na perspectiva funcionalista, o afrouxamento na correlação biunívoca entre forma e função – admitindo-se que a iconicidade na gramática não é absoluta mas gradiente (Givón, 1995;2001), e que essa correlação se perde por pressões diacrônicas que levam tanto ao desgaste do código como à alteração da mensagem (versão branda do princípio da iconicidade) – abre espaço para o tratamento da variação linguística, pois admite a possibilidade de mais de uma forma para a mesma função significação.

Nesta tese, tem-se três dêiticos espaçotemporais, que, com suas peculiaridades, convergem para uma mesma função. Conforme visto anteriormente, é possível recuperar o uso espacial e temporal de *aí* e *daí*, ao passo que *então* só pode ser estudado como dêitico temporal. Os dêiticos espaçotemporais adquiriram funções textuais de sequenciação e, dentro desse domínio discursivo, passaram a ser utilizados intercambiavelmente na causalidade referencial.

Do ponto de vista sociolinguístico, constata-se que *aí*, *daí* e *então* são variantes de uma mesma variável, - no domínio da causalidade. Em um mesmo contexto, o da causalidade referencial, assumem a mesma função, a de evidenciar

discursivamente a relação causa-efeito, de modo que sua substituição não causa danos ao sentido do enunciado¹⁸.

Sob a perspectiva funcionalista, desenha-se o trajeto de gramaticalização de *aí*, *daí* e *então* para definir o ponto de convergência funcional da análise variacionista levada a cabo nesta tese.

A gramaticalização e o fenômeno de variação morfossintática é comumente analisado como um dos pilares da perspectiva sociofuncionalista sobre os fenômenos de variação e mudança linguística. Para Tagliamonte (2002), por exemplo, a gramaticalização gera variabilidade na gramática e essa variabilidade pode ser verificada por meio das camadas variáveis da gramaticalização, manifestadas em diferentes formas linguísticas.

Na Sociolinguística, Labov compreende a variação estilística como uma adaptação da linguagem do falante ao contexto do ato de fala. Para o Funcionalismo (Givón, 1984; Hopper, 1991), a mudança é motivada por práticas discursivas e sociais. Desse modo, os estudos funcionalistas de gramaticalização que têm como o falante podem contribuir para o estudo da variação intrafalante.

Como variável discursiva, o uso dos sequenciadores em duas comunidades distintas revela convergências estruturais e divergências sociais atrelados à constituição das amostras de fala de cada capital.

A abordagem sociofuncionalista contribui para o refinamento da análise estatística, uma vez que inclui de fatores discursivo-pragmáticos como os que foram analisados nesta tese, tais como: domínio discursivo, sequência discursiva, ativação informacional do tópico, distribuição temporal das sequências discursivas, dentre outros (conforme detalha o Capítulo 4).

Partindo dos pressupostos do Sociofuncionalismo, o próximo capítulo tratará dos procedimentos que norteiam a constituição das amostras em cada uma

¹⁸ Aqui, a definição de enunciado é tomada de Bakhtin, que o define como unidade básica da linguagem, em relação ao signo. Para o linguista, um enunciado não é um signo pela simples razão de que para existir ele exige a presença de um enunciador (quem fala, quem escreve) e de um receptor (quem ouve, quem lê). O signo faz parte de uma construção teórica que dispensa os sujeitos reais do discurso. Um signo, num dicionário, não é e não pode ser um enunciado. Este exige uma realização histórica. Um enunciado acontece em um determinado local e em um tempo determinado, é produzido por um sujeito histórico e recebido por outro. Assim, cada enunciado é único e irrepetível.

das capitais e da análise dos dados. Nesse sentido, Campo Grande e São Paulo serão apresentadas como comunidades de fala que compartilham normas em relação ao uso de *aí*, *daí* e *então* do domínio referencial da causalidade.

3

Córpus e metodologia

Este capítulo trata da constituição das amostras de fala e da metodologia de análise dos dados. Descrevem-se as comunidades em questão - Campo Grande e São Paulo - e retomam-se as razões que justificam um estudo sociofuncionalista comparativo entre elas.

Apresenta-se o *software* Rbrul (Johnson, 2009) – o pacote de programas utilizado nas análises estatísticas, que ainda é pouco utilizado nas pesquisas sociolinguísticas brasileiras, mas oferece uma série de vantagens em relação ao Goldvarb X (Sankoff; Tagliamonte; Smith, 2005).

1. Três sequenciadores, duas comunidades

Conforme visto anteriormente, um fenômeno de variação pode se revelar estável ou como mudança em progresso. No caso dos sequenciadores aqui em foco – *aí*, *daí* e *então* em contextos de causalidade referencial –, verificar se há ou não mudança em curso significa perguntar: (i) estaria alguma das formas se

especializando nesses contextos?; (ii) a que fatores linguísticos e sociais se correlacionam os empregos dessas formas?

A fim de responder a tais perguntas, analisam-se 357 dados de *aí*, *daí* e *então* na amostra campo-grandense e 349 ocorrências desses sequenciadores no *cópus* paulistano. Os dados paulistanos foram extraídos de 48 entrevistas sociolinguísticas, dentre aquelas 60 gravadas para o Projeto SP2010 (Mendes & Oushiro, 2013). As 48 entrevistas de Campo Grande foram realizadas especificamente para esta tese¹⁹. Assim, a análise desses sequenciadores, do modo como se descreve nos capítulos anteriores, baseia-se em um *cópus* que se pode considerar robusto.

2. Para a constituição de amostras de fala

O termo *amostra* diz respeito ao grupo de pessoas selecionadas para representar a comunidade ou o universo a que pertencem, mas também costuma ser empregado para nomear o conjunto de ocorrências das variantes em foco, extraídas das entrevistas coletadas. A premissa central é de que o comportamento linguístico dos informantes reflete regularidades que resultam da sua adesão às normas dos grupos sociais a que pertencem (Labov 2008 [1972]). É nesse sentido que os resultados do estudo dos usos linguísticos por parte de um número restrito de pessoas – a amostra – estende-se para os grupos sociais que elas integram (Labov 2008 [1972]). A verificação disto se faz com métodos estatísticos – tanto descritivos quanto inferenciais (Jhonson, 2009).

Difícilmente ocorre de existir uma comunidade de falantes tão pequena a ponto de que sua “amostra” inclua todos os seus membros (o que tornaria o próprio termo “amostra” inadequado). Normalmente, o que se faz é selecionar alguns membros de uma comunidade de falantes.

Idealmente, uma amostra deve ser representativa da comunidade. Nesse sentido, é necessário considerar que:

[...] pela lei dos grandes números, sabe-se que, até certo ponto, a probabilidade de que os resultados sejam fidedignos é diretamente proporcional ao

¹⁹ Esta amostra representativa da fala campo-grandense é a primeira de que se tem notícia.

tamanho da amostra. Isso porque o possível efeito de uma variável num indivíduo será somado ao de outro indivíduo, cada um com seu efeito casual, do qual nunca se estará completamente livre. Esses efeitos tendem a se anular mutuamente, já que, sendo casuais, agirão alguns numa direção e alguns em outra (OLIVEIRA & SILVA, 2003:19)

A amostra campo-grandense aqui analisada é representativa, se avaliadas as variáveis sociais que a estratificam. Por outro lado, levando em consideração a existência de vários (i)migrantes em São Paulo, os quais não foram considerados na amostra, não se pode afirmar que o *cópus* paulistano seja suficientemente representativo.

Foram considerados, com base na literatura Sociolinguística, fatores de ordem linguística - aqueles considerados como responsáveis pela ocorrência de cada um dos fenômenos observados - atrelados a três variáveis sociais permanentes: a *escolaridade* (média e superior), *faixa etária* e *sexo/gênero* e *cidade* (Campo Grande e São Paulo).

2.1. Roteiro de entrevista

O roteiro de entrevista é constituído de 34 perguntas-guia²⁰, além da leitura de dois textos e de uma lista de palavras, o que totaliza cerca de uma hora de gravação para cada informante. Está dividido em três partes (ver Anexos). Na primeira, foram contemplados tópicos como bairro, infância, família, vida escolar, amigos, e atividades frequentes do informante em seu tempo livre.

Na segunda parte, o informante responde a perguntas relativas às suas apreciações sobre a cidade e sobre a variedade linguística campo-grandense. Na terceira parte da entrevista, faz-se a leitura de uma lista de palavras, de uma notícia de jornal e de uma narrativa com características orais. A última parte da entrevista não é utilizada como fonte de dados para esta tese. Foi gravada em virtude do

²⁰ Disponível nos Anexos: *anexo A* referente à amostra paulistana e *B*, referente ao *cópus* campo-grandense.

roteiro-modelo utilizado – o do Projeto SP2010 - e com vistas à realização de estudos vindouros.

Conforme já se mencionou, o roteiro para a gravação de entrevistas com campo-grandenses baseou-se amplamente naquele utilizado pelo Projeto SP2010. Fizeram-se necessárias, contudo, algumas adaptações. Se para a amostra da fala paulistana o roteiro foi proposto de modo a proporcionar conversas que revelassem a identificação dos informantes com a capital paulista, para que se obtivesse o mesmo tipo de material da fala campo-grandense (ver Anexo B), alguns pontos do roteiro precisaram ser alterados, sobretudo no que diz respeito à relação do informante com a cidade e à leitura de uma notícia de jornal (para a qual selecionou-se uma notícia local).

Sabe-se que a utilização do gravador na situação de entrevista é impactante e, de certa forma, age no sentido contrário ao objetivo pretendido, que é a obtenção de dados de fala os mais espontâneos possíveis - paradoxo do observador (Labov, 2006). Em virtude disso, a inclusão, no roteiro, de tópicos em que o informante discorre sobre sua infância e sobre seus momentos de lazer, é feita com o objetivo de que as amostras de fala sejam tenham um menor grau de monitoramento.

2.2. Coleta de dados

As entrevistas das amostras paulistana e campo-grandense aqui analisadas foram registradas individualmente, com um gravador digital (modelo TASCAM DR-100) e, posteriormente, salvas em um computador. As entrevistas são semiespontâneas, ou seja, os informantes desenvolveram conversas com o documentador a partir de perguntas previamente preparadas acerca de lazer, da sua infância e juventude, planos para o futuro e situações em que correram risco de vida. Entende-se que, ao relatar fatos em que se envolveu emocionalmente, ainda que em situação de “entrevista”, as pessoas tendem a utilizar variantes que empregariam nas interações mais espontâneas.

Sabe-se que a utilização do gravador na situação de entrevista é impactante e, de certa forma, age no sentido contrário ao objetivo pretendido, que é a obtenção de dados de fala os mais espontâneos possíveis - paradoxo do observador (Labov, 2006). Em virtude disso, a inclusão, no roteiro, de tópicos em que o informante discorre sobre sua infância e sobre seus momentos de lazer, é feita com o objetivo de que as amostras de fala sejam tenham um menor grau de monitoramento.

O propósito da ferramenta “entrevista sociolinguística”, então, é possibilitar o acesso, por parte do pesquisador, a diferentes “estilos” (Labov, 2006 [1966]) – desde o mais casual (as partes da conversa em que o informante tende a prestar menos atenção no modo como está falando) até o mais cuidadoso (leitura de um texto e, mais ainda, leitura de uma lista de palavras ou de pares mínimos).

Em Campo Grande, as entrevistas foram gravadas em um local onde o informante pudesse se sentir mais à vontade, de modo que pudesse responder a tudo tranquilamente. As perguntas feitas aos informantes tiveram o objetivo de gerar depoimentos inéditos, a fim de evitar qualquer tipo de pressuposição ou inferência que pudesse condicionar as respostas e, principalmente, a variante utilizada pelos entrevistados. Todos os entrevistados sabiam que seriam gravados e assinaram um termo de consentimento (Anexos).

2.3. Transcrição das entrevistas

O primeiro passo depois da gravação das entrevistas foi sua transcrição ortográfica. Todas foram transcritas integralmente, de acordo com as orientações do Projeto SP2010. A utilização de tais normas tem como objetivo ajudar o pesquisador a descrever, da maneira mais fiel possível, um corpus oral.

Quando se transcreve um texto oral, o objetivo é que a transcrição proporcione a refacção mais próxima possível do texto original. Assim, as convenções utilizadas na transcrição do corpus desta pesquisa foram estipuladas no Projeto SP2010.

Para a transcrição das entrevistas do Projeto SP2010, foi utilizado o ELAN (Hellwig & Geerts, 2013), que é um programa para anotação de arquivos de áudio e vídeo, desenvolvido pelo Instituto Max Planck de Psicolinguística. Suas principais vantagens são: (a) permite sincronizar áudio e transcrição, o que facilita enormemente a análise linguística dos dados ; (b) possibilita a criação de várias trilhas de transcrição, a fim de separar a fala de diferentes participantes, bem como anotar, de forma detalhada, outros aspectos linguísticos e contextuais (por exemplo, sobreposição de vozes, ações gestuais concomitantes às verbais); (c) viabiliza a utilização de ferramentas de buscas dentro do corpus; (d) é flexível no que diz respeito a formatos de exportação da transcrição (.txt, .textgrid), compatíveis com vários outros programas, como Word, Excel, R, Rbrul e Praat.

Figura 1 – Trecho de uma transcrição da amostra paulistana no Elan

Diferentemente das entrevistas paulistanas, retiradas do Projeto SP2010, as gravações com campo-grandenses foram transcritas manualmente em arquivos “.txt” (figura 2). Como, no Elan, o processo de transcrição tende a tomar mais tempo, houve a necessidade de agilizar a transcrição da amostra campo-grandense com vistas à extração dos dados. Em função disso, os áudios foram transcritos de forma convencional.

Figura 2 – Trecho de uma transcrição da amostra campo-grandense

2.4. Amostra paulistana

No Brasil, os primeiros estudos sobre a língua falada tiveram início com o NURC – Norma Linguística Urbana Culta – (Castilho; Pretti, 1986; Pretti; Urbano, 1988, 1990), na década de 70. O projeto reuniu dados de cinco grandes capitais brasileiras – entre elas, São Paulo. Já no Projeto Iboruna (Gonçalves, 2008),

coletou-se uma amostra constituída por falantes do interior do estado de São Paulo, na região de São José do Rio Preto²¹.

O Grupo de Estudos e Pesquisa em Sociolinguística da USP (GESOL-USP), por sua vez, tem coletado entrevistas com paulistanos desde 2006 (nascidos e criados na capital paulista). Com o Projeto SP2010, foram disponibilizadas 60 entrevistas, em site próprio. Tais entrevistas foram estratificadas de acordo com o *sexo/gênero*, a *escolaridade* e a *faixa etária dos informantes*, gerando-se 12 perfis sociolinguísticos, para cada um dos quais foram gravados 5 informantes:

Sexo/Gênero	Escolaridade	Faixa etária	Perfil do informante
Feminino	Médio	1	FC ²² 1
		2	FC2
		3	FC3
	Superior	1	FS1
		2	FS2
		3	FS3
Masculino	Médio	1	MC1
		2	MC2
		3	MC3
	Superior	1	MS1
		2	MS2
		3	MS3

Quadro 1 - Perfil dos informantes de ambas as amostras

Para que fossem diminuídos os efeitos causados pelo automonitoramento do informante, planejou-se um roteiro, que deveria ser previamente estudado pelo documentador, a fim de conduzir a entrevista da maneira mais informal possível, como se não houvesse o propósito de gravar dados para análise ulterior.

A seleção de informantes foi semialeatória, de acordo com o método “amigo do amigo” (Milroy, 2003): nesse procedimento, o contato com um

²¹ Cf: <<http://www.iboruna.ibilce.unesp.br/interna.php?Link=corpo.php&corpo=36>>.

²² A letra C faz referência a “colegial”, antiga denominação do Ensino Médio e foi adotada para que não haja confusão com o M de “masculino”.

determinado informante é feito por intermédio de uma terceira pessoa. Dessa forma, o primeiro contato do entrevistador com o sujeito a ser entrevistado fica mais natural, ao mesmo tempo que se engendra alguma aleatoriedade e abrangência, uma vez que tal método leva a grupos sociais diversificados, numa cidade cuja demografia é tão complexa como São Paulo. Assim, depois de um primeiro contato com o informante, agenda-se uma “conversa” para posterior gravação²³. A partir desse método, gravaram-se entrevistas com informantes de todas as regiões da capital paulista (Mendes & Oushiro, 2012):

Figura²⁴ 3 – Divisão da cidade de São Paulo por zonas

Dentre as 60 entrevistas do Projeto 2010, 48 foram selecionadas para este estudo. Houve o cuidado de se elegerem aquelas cuja extensão não fosse muito inferior ou muito superior a 1 hora, duração aproximada das gravações feitas em São Paulo. Outro critério adotado foi com relação à idade dos informantes. Dentro de cada faixa etária, foram escolhidos os informantes paulistanos com idades semelhantes às do campo-grandense.

²³ Ao final da entrevista, o informante assina um termo de livre consentimento, autorizando a utilização dos dados de fala para fins de pesquisa sociolinguística.

²⁴<http://projetosp2010.fflch.usp.br/sites/projetosp2010.fflch.usp.br/files/sp2010-mapas-zonas.png>

2.5. Amostra campo-grandense

As entrevistas realizadas com campo-grandenses permitiram o reforço de alguns pressupostos, como, por exemplo, o de que a fala paulistana é reconhecida nacionalmente por alguns traços característicos (52):

(52) D1²⁵: e como você acha que um **paulistano** falaria algumas dessas palavras?

S1: é aí tem mortadela Marba {retroflexo} é "quero uma mortadela" {retroflexo} "mElnta" né? essas que têm mais o *ene* "sabor de mElnta presElnte entretenimElnto" né? mais esse tipo de de sotaque eu acho que é... defender discoteca... eu não sei eu acho que mais esses com com o o *ene* com o "einto" o "ein" é alguma coisa sairia dessa forma né? imagino (CG39MS-Osmar)

Em contrapartida, a variedade campo-grandense é reconhecida por sua aparente “neutralidade”, opinião que expuseram os próprios nativos:

(53) [o **campo-grandense** em si eu acho que não as pessoas não têm não têm aquela aquela coisa que nós temos aqui de de vir um carioca você escuta o sotaque e sabe que a pessoa é de do Rio de Janeiro] **campo-grandense** eu acho que é meio difícil da pessoa identificar (CG25FS-Celina)

Quando tentam descrever as características da variedade campo-grandense, o único traço elencado é o /r/ retroflexo, que frequentemente associa a identidade linguística do campo-grandense à do interior paulista.

(54) D1: tem algum modo de falar que você acha bem campo-grandense?

²⁵ D1 diz respeito ao documentador, ao passo que S1 faz referência ao sujeito entrevistado, ou informante, termo caro à Sociolinguística.

S1: [bem **campo-grandense**? você diz o sotaque... essa questão aí? eu acho que essa esse puxado desse *erre* aí né?] um pouquinho eu acho que é característica daqui

D1: tem mais alguma coisa que você acha que caracteriza o modo de falar de Campo Grande?

S1: [não é só mesmo esse *erre* que aqui na verdade mas eu acho que fica essa esse esse sotaque característico desse *erre* puxado] que eu acho que tem outros lugares que é muito mais forte assim desse puxado desse *erre* aí eu acho que sim (CG39F5-Taís)

A despeito de fenômenos fonético-fonológicos como o uso de /r/ em coda, a variação dos sequenciadores *aí*, *daí* e *então*, fenômeno de natureza discursiva, não revela, de forma patente, características linguísticas específicas das duas variedades aqui analisadas. Ao comparar amostras dessas duas comunidades, tal tese pretende desenhar, todavia, o panorama das formas utilizadas no PB para a sequenciação causal.

Daí se projeta o caráter inovador de tal proposta, sobretudo no tocante às capitais estudadas, visto que a pesquisa linguística de maior projeção já realizada em Campo Grande é, na verdade, parte de um projeto mais amplo: o Atlas Linguístico de Mato Grosso do Sul (Oliveira, 2007). No Atlas, por exemplo, dentre as 64 gravações que compõem o Atlas, somente quatro foram realizadas com campo-grandenses²⁶.

A pequena amostra campo-grandense que compõe o Atlas Linguístico de Mato Grosso do Sul foi gravada sob uma perspectiva mais geolinguística que sociolinguística e conta com informantes com baixo nível de escolaridade. Diferentemente, o *cópus* de que parte esta tese orienta-se para o estudo de informantes que tenham cursado Ensino Médio e Superior.

Definidas nos moldes do Projeto SP2010, em Campo Grande, as faixas etárias são: 20 - 34 anos; 35 - 49 anos; 50 anos ou mais e, em São Paulo: 20-34

²⁶ Os perfis dos entrevistados eram: 2 homens, um com 31 anos e outro com 60 anos, e 2 mulheres, uma com 24 anos e outra com 58 anos, todos com baixa escolaridade (os informantes haviam concluído apenas o primeiro ciclo da educação básica).

anos; 35-59 anos; 60 ou mais. A segunda e a terceira faixas diferem das de São Paulo devido à constituição demográfica da capital sul-mato-grossense. Como se trata de uma cidade jovem, em relação a São Paulo, a média de idade dos falantes nativos é de 40 anos. A maioria das pessoas com mais de 50 anos são (i)migrantes. Também em decorrência desse fato, entre as pessoas de mais idade, foram entrevistadas não só campo-grandenses natos, mas também pessoas que moram na capital há, pelo menos 30 anos, em todas as regiões da capital:

Figura²⁷ 4 – Divisão de Campo Grande por regiões

O quadro a seguir elenca os 48 informantes campo-grandenses - cuja seleção também foi semialeatória, segundo o método “amigo do amigo” (Milroy, 2003) -. ao lado daqueles que foram selecionados na amostra do Projeto SP2010.

Quadro 2 – Informantes

Deve-se reiterar que, junto ao pseudônimo dos informantes de São Paulo, há uma inicial maiúscula, que indica o sobrenome do indivíduo.

²⁷ <http://www.skycrapercity.com/showthread.php?t=1489859>

Faixa etária	Campo-grandenses		Paulistanos	
	Feminino	Masculino	Feminino	Masculino
20-34 anos Médio	Adriana	Eduardo	CarolinaA	DalsonO
	Camila	Everton	JanaínaC	DouglasO
	Luísa	Marcelo	LaraN	EvandroO
	Paula	Raimundo	MilenaF	NicolauS
20-34 anos Superior	Celina	Fábio	RaíssaO	CarlosJ
	Débora	Júlio	SilviaB	LucasS
	Tainá	Raul	ValériaA	JaimeN
	Talita	Renan	VivianeC	RodrigoR
35-49/59 anos Médio	Joana	César	MarietaS	MaurícioB
	Regina	Lúcio	MeireC	LucianoT
	Nair	Nícolas	RenataC	NelsonF
	Zuleica	Ziraldo	RuthC	SandroS
35- 49/59 anos Superior	Emília	Francisco	ArianeG	FredericoV
	Marlene	Osvaldo	ElaineS	JoséN
	Renata	Ronaldo	LucianaN	MarcosB
	Taís	Wesley	PolianaM	PedroS
50/60 ou mais Médio	Adélia	Alberto	CarmenF	HugoF
	Eliana	Bento	MarciaM	RafaelV
	Marta	Fausto	SoraiaS	RodolfoR
	Olga	Maurício	VeraD	RômuloS
50/60 ou mais Superior	Glória	Durval	IaraM	AmadeuC
	Rosana	Mauro	RenataL	LúcioP
	Vanusa	Pedro	SueliO	NiltonR
	Vitória	Roberto	VivianT	NeimarV

2. Vizinhança, semelhanças e diferenças

Alguns dados sobre a realidade das duas capitais estudadas podem auxiliar na compreensão dos padrões sociolinguísticos encontrados na fala de tais comunidades.

Situada na região sudeste do país, foi fundada 25 de janeiro de 1554, pelos padres José de Anchieta e Manuel de Nóbrega, que haviam subido a Serra do Mar

em busca de um local seguro para se instalar e catequizar os índios. Quase cinco séculos depois, o povoado de Piratininga transformou-se na maior metrópole brasileira. São Paulo é a maior cidade da América Latina, com população estimada em mais de 11 milhões habitantes (IBGE, 2015²⁸). A capital paulista é o grande polo econômico do país. Sua economia baseia-se na prestação de serviços, seguida pela indústria. Situada na região sudeste do país, é a principal cidade da Região Metropolitana do estado de São Paulo, rodeada por outras, que formam a Grande São Paulo, com aproximadamente 20 milhões de habitantes (IBGE, 2015).

Figura²⁹ 5 – (da direita para a esquerda) Localização de São Paulo e da cidade de São Paulo no território brasileiro

Campo Grande, por sua vez, foi fundada mais de três séculos depois, em 1889, por colonizadores mineiros, que lá chegaram com o intuito de desfrutar dos campos e das águas do cerrado (Oliveira, 2006). Atualmente, a capital tem mais de 760 mil habitantes e é considerada a terceira cidade mais desenvolvida da região Centro-Oeste e a 23^a maior cidade do país (Ipea, 2006³⁰).

²⁸ <http://cidades.ibge.gov.br/xtras/perfil.php?codmun=355030>
²⁹

ftp://geofpt.ibge.gov.br/mapastematicos/mapasescolares/ensinomedio/mapas_estaduais/pdf/são-paulo.pdf

Figura³¹ 6 – (da esquerda para a direita) Localização de Mato Grosso do Sul e de Campo Grande no território brasileiro

Capitais de estados que se situam na divisa entre o Centro-Oeste e o Sudeste, o estudo comparativo de Campo Grande e São Paulo se justifica em virtude de características comuns às duas comunidades. Capitais de estados vizinhos (figura 6), Campo Grande e São Paulo têm em comum o fato de serem cidades ecumênicas, com intenso número de migrantes – principalmente do Nordeste, de Minas Gerais, de Goiás, do Rio Grande do Sul e de Santa Catarina - e imigrantes – sobretudo do Líbano, da Bolívia, da Paraguai, de Portugal e da Espanha - e com variedade linguística ainda pouco explorada sociolinguisticamente.

www.ipea.gov.br/desafios/index.php?option=com_content&view=article&id=994:catid=28&Itemid=23

31

ftp://geofp.ibge.gov.br/mapas_tematicos/mapasescolares/ensinomedio/mapas_estaduais/pdf/mato_grosso_do_sul.pdf

Figura³² 7 – Campo Grande e São Paulo

Em Campo Grande, é bastante notável, inclusive, a influência paulista nos hábitos de vida e na composição demográfica, como se verifica nos excertos abaixo, em que os informantes expõem sua opinião sobre o imaginário social sobre a capital sul-mato-grossense:

(55) [aí chega essa molecada de **São Paulo** aqui que adora um sertanejo e uma cerveja se jogam se jogam despirocam na cerveja na festa e na é uma piração aí começa a dar trabalho aí falam “não eu vim aqui para estudar”]... até a minha cabeleireira falou o seguinte... quando os pais vêm de **São Paulo** para cá para olhar os filhos e ver o dinheiro que eles estão mandando os meninos vão até cortar o cabelo (CG28F5-Débora)

³² http://atlasecolar.ibge.gov.br/images/atlas/mapas_brasil/brasil_politico.pdf

(56) a capital acho que é um lugar muito complicado mas assim se tiver oportunidade... [a gente vê mesmo o pessoal vindo de de outros e e algumas pessoas vindo de **São Paulo** e tal e encontram aqui ainda é... uma... aqui tem tem características muito próprias já de uma... tem já algumas características de uma metrópole mas também de uma cidade pequena né?] (CG49MS-Osmar)

A comparação entre essas duas comunidades – Campo Grande e São Paulo - permite verificar a existência de compartilhamento de normas entre as cidades no que tange ao uso de *aí*, *daí* e *então* como sequenciadores de causalidade referencial. Especificamente, a este estudo interessa verificar: (i) se alguns dos sequenciadores de mais prestígio social e se ele é mais frequente entre os paulistanos, tidos como mais urbanizados que os campo-grandenses, como já exposto na Introdução; (ii) quais os conjuntos de fatores selecionados para o uso das variantes *aí* e *daí* e *então* em Campo Grande e São Paulo, no sentido de verificar se tais comunidades compartilham as mesmas normas de uso em relação aos sequenciadores estudados; (iii) se o fenômeno em questão caracteriza-se como variação estável ou mudança em curso.

Aqui, tais premissas advêm da consideração, confirmada pela literatura sociolinguística, de que o controle de fatores sociais sobre o uso de variantes discursivo-pragmáticas nem sempre é significativo (Freitag, 2011). A análise variável dos sequenciadores *aí*, *daí* e *então*, revela – Capítulo 5, que, de fato, a escolha de uma forma em detrimento da outra é influenciada por fatores sociais.

3. A análise estatística dos dados

De acordo com (Sankoff, 1988), regularidades subjacentes a fenômenos variáveis podem prescindir de uma frequência alta de dados, o que conduz ao tratamento quantitativo da variável dependente (o fenômeno estudado) em função de fatores ou restrições internas e externas (grupos de fatores), que operacionalizam as hipóteses aventadas. Em termos gerais, o que se pretende em

um trabalho de orientação laboviana, como este, é fazer generalizações inferenciais sobre as tendências da comunidade que foi amostrada, por isso a possibilidade de trabalhar com amostras e não com todos os indivíduos da população.

De acordo com Scherre (2012), a ferramenta mais frequentemente utilizada nas últimas quatro décadas é o programa de regra variável (Varbrul), em suas diversas versões, como o Goldvarb X. Tal programa o efeito relativo dos fatores das variáveis independentes ou grupo de fatores, projetando pesos relativos associados a cada fator de cada variável independente em sucessivas análises (Sankoff, 1988; Guy & Zilles, 2007).

De acordo com Johnson (2009), a versão mais popular do GoldVarb carece de flexibilidade e isola seus usuários da comunidade mais ampla de linguistas quantitativos. Uma nova versão da regra variável programa, Rbrul, tenta resolver essas questões e, com efeitos mistos de modelagem, atenta para significância dos efeitos.

Nesta tese, as análises estatísticas foram realizadas por meio da plataforma R, com auxílio do pacote Rbrul (Johnson, 2009), interface utilizada para emprego das funções *glm* (do inglês “generalized linear model”, ou modelo linear generalizado) e *glmer* (do inglês “generalized linear mixed-effects models”, ou modelos de efeitos mistos lineares generalizados) (Bates; Sarkar 2008, *apud* Johnson, 2009).

O Rbrul permite considerar o papel do item lexical como uma variável aleatória (*random effect*), isto é, verificar se certos padrões se explicam não pelas variáveis propostas e incluídas na análise, mas sim em razão de certos itens lexicais.

Enquanto o GoldVarb X requer que cada fator considerado na descrição do fenômeno esteja representada por um código simples, no Rbrul, é possível utilizar vários caracteres para cada fator (como, por exemplo, o nome de um informante) – o que facilita a preparação dos dados antes da análise quantitativa.

A figura abaixo ilustra a codificação das ocorrências da presente pesquisa, para posterior análise no Rbrul, cujos resultados serão expostos no Capítulo 5,

logo após a apresentação dos grupos de fatores e das respectivas hipóteses aventadas neste estudo – Capítulo 4.

	A	B	C	D	E	F	G	H	I	J	K	L
1	V.D.	Sexo/gênero	Faixa etária	Escolaridade	Cidade	Sequência discursiva	Status da informação	Gradiência do tópico	Duração	Dinamicidade	Telicidade	Aspecto gramatical
2	então	masculino	2	superior	CG	exposição	inferível	intertópico	durativo	dinâmico	atélico	indeterminado
3	ai	masculino	2	superior	CG	argumentação	dada	intratópico	durativo	estático	atélico	indeterminado
4	então	masculino	2	superior	CG	descrição	inferível	intertópico	durativo	estático	atélico	indeterminado
5	então	masculino	2	superior	CG	argumentação	inferível	intertópico	durativo	dinâmico	atélico	indeterminado
6	então	masculino	2	superior	CG	argumentação	dada	intertópico	durativo	dinâmico	atélico	indeterminado
7	então	masculino	2	superior	CG	descrição	nova	intertópico	durativo	estático	atélico	perfectivo
8	então	masculino	2	superior	CG	narrativa	inferível	intertópico	pontual	estático	atélico	imperfectivo
9	então	masculino	2	superior	CG	narrativa	nova	intertópico	pontual	estático	atélico	iterativo
10	então	masculino	2	superior	CG	exposição	nova	intertópico	durativo	estático	atélico	imperfectivo
11	então	masculino	2	superior	CG	argumentação	nova	intratópico	durativo	estático	atélico	indeterminado
12	então	masculino	2	superior	CG	exposição	nova	intertópico	durativo	dinâmico	télico	indeterminado
13	então	masculino	2	superior	CG	argumentação	nova	intertópico	pontual	estático	atélico	iterativo
14	então	feminino	3	médio	CG	exposição	nova	intertópico	durativo	estático	atélico	indeterminado
15	então	feminino	3	médio	CG	descrição	inferível	intertópico	durativo	estático	atélico	indeterminado
16	então	feminino	3	médio	CG	argumentação	inferível	intratópico	durativo	estático	atélico	indeterminado
17	ai	feminino	3	médio	CG	argumentação	nova	intratópico	durativo	estático	atélico	indeterminado

Figura 8 – Codificação dos dados no Excel

O Rbrul também é mais indulgente em relação a nocautes, situações em que a resposta é invariante - ou 0% ou 100% - em um subconjunto do dados . Para evitar ocorrências dessa natureza, o usuário do Rbrul pode rodar os fatores juntos ou excluí-los como em GoldVarb, mas isso raramente é obrigatório, embora boas práticas continuem a exigir a sua exclusão (Guy, 2007).

Só é possível rodar o Rbrul no ambiente do Programa R. Ambos são igualmente públicos e estão disponíveis para download na internet. O pacote Rbrul, a princípio, executa as mesmas tarefas que as versões dos pacotes Varbrul para MS-DOS e Goldvarb, uma vez que é um programa para executar análise multivariada. O Rbrul tem uma série de melhorias sobre seu antecessor Goldvarb X, mas ainda produz seus resultados, da mesma forma, o que significa que sociolinguistas que começaram a usá-lo puderam preservar o modo de apresentar os resultados de suas pesquisas (Meyerhoff; Scleef; Mackenzie, 2015).

Com base em uma lista de comandos, ilustrados na figura 9, (“MODELING MENU: 1 – choose variables 2 – one-level 3 – step-up...”), escolhe-se a opção desejada. A partir disso, o Rbrul disponibiliza novos comandos (“Choose response”; “Type of response”, entre outros), a fim de que se formule o modelo de a análise mais conveniente ao conjunto de dados.

Table 5.7 *Rbrul* modeling menu.

```
MODELING MENU
1-choose variables 2-one-level 3-step-up 4-step-down 5-step-up/
step-down
8-settings 9-main menu 0-exit
10-chi-square test
1: 1
Choose response (dependent variable) by number, or Enter to keep
Dep.var (1-Dep.var 2-Main.subj 3-Verbs.3 4-Verbs.1 5-Matrix.subj
6-Add.elm 7-Sub.subj 8-Int.mat 9-Tense 10-Indiv 11-Sex 12-Age
13-Occ 14-Edu 15-Context)
1: 1
Type of response? (1-continuous Enter-binary)
1:
Choose application value(s) by number? (1-T 2-Z)
1: 2
Choose predictors (independent variables) by number, or Enter to
keep Main.subj & Verbs.1 & Matrix.subj & Add.elm & Sub.subj &
Int.mat & Tense (2-Main.subj 3-Verbs.3 4-Verbs.1 5-Matrix.subj
6-Add.elm 7-Sub.subj 8-Int.mat 9-Tense 10-Indiv 11-Sex 12-Age
13-Occ 14-Edu 15-Context)
1: 4
2: 5
3: 6
4: 7
5: 8
6: 9
7:
Are any predictors continuous? (4-Verbs.1 5-Matrix.subj 6-Add.elm
7-Sub.subj 8-Int.mat 9-Tense Enter-none)
1:
Any grouping factors (random effects)? (4-Verbs.1 5-Matrix.subj
6-Add.elm 7-Sub.subj 8-Int.mat 9-Tense Enter-none)
1:
Consider an(other) pairwise interaction between predictors? Choose
two at a time. (4-Verbs.1 5-Matrix.subj 6-Add.elm 7-Sub.subj
8-Int.mat 9-Tense Enter-done)
1:
Current variables are:
response.binary: Dep.var (Z vs. T)
fixed.factor: Verbs.1 Matrix.subj Add.elm Sub.subj Int.mat Tense
MODELING MENU
1-choose variables 2-one-level 3-step-up 4-step-down 5-step-up/
step-down
8-settings 9-main menu 0-exit
10-chi-square test
1: 5
```

Figura 9 – Interface do Rbrul (Tagliamonte, 2011:140)

No Rbrul, as variáveis são selecionadas como estatisticamente significativas somente quando as correlações são fortes o suficiente para ultrapassar a variação interfalantes (Johnson, 2009). Se, por um lado, o Varbrul tende a selecionar certas variáveis como relevantes para a variação, quando, na verdade, não o são, as análises realizadas no Rbrul podem considerar irrelevantes algumas variáveis, que são, de fato, significativas (Johnson, 2009). O Rbrul, portanto,

parece ser mais criterioso ao reportar probabilidades – o que pode ser preferível ao fato de considerar relevante uma variável que, de fato, não o é.

Este *software* dispõe de um procedimento automático que identifica a significância dos fatores. Enquanto o GoldVarb usa um limite fixo de 0,05 para determinar a significância de um grupo de fatores, no Rbrul, esse valor pode ser ajustado.

Segundo Oushiro (2015), essas diferenças são mais comuns em relação a variáveis independentes com pouca relevância na análise no Varbrul, tanto por uma pequena diferença no *range* (diferença entre maior e menor pesos relativos dentro de um grupo de fatores – Guy & Zilles (2007), quanto pela ordem de seleção no *step-up* (em geral, um dos últimos ou o último grupo de fatores a ser selecionado).

4

Hipóteses e grupos de fatores

Apresentam-se, mediante exemplos extraídos das amostras de Campo Grande e São Paulo, 12 grupos de fatores – 8 de natureza linguística e 4 de ordem social – que foram incluídos nas análises quantitativas dos dados (realizadas com o Rbrul – Johnson, 2009). Tais variáveis linguísticas e sociais advêm de hipóteses sobre o uso de *aí*, *daí* e *então* como sequenciadores variáveis no domínio referencial da causalidade.

1. Grupos de fatores

As entrevistas sociolinguísticas que compõem as amostras utilizadas nesta tese resultam em mais de 96 horas de gravação, das quais foram extraídas, no total, 706 ocorrências de *aí*, *daí* e *então* como sequenciadores da causalidade no domínio referencial.

Para Labov (2008 [1972]: 214), “[...]as pressões internas, estruturais, e as pressões sociolinguísticas agem em alternância sistemática no mecanismo da mudança linguística”. Assim, a partir da análise qualitativa dos dados, aventam-se hipóteses sobre variáveis – linguísticas e sociais – que podem se correlacionar com o uso das formas variantes. Tais hipóteses levam ao estabelecimento do que comumente se denomina *grupos de fatores* na Sociolinguística – as variáveis independentes do estudo, cuja correlação (de natureza estatística) com a variável dependente pode ou não se confirmar. O exame dessas correlações foi feito com o pacote Rbrul (Capítulo 5).

Não só com base na análise qualitativa dos dados, mas também a partir de trabalhos anteriormente desenvolvidos (Freitag et al, 2009; Tavares, 2003; Paiva, 1995; Castilho, 1968, dentre outros), foram estabelecidas doze variáveis

independentes (elencadas no Quadro 3), dentre as quais quatro são de natureza social e oito são de ordem linguística.

Grupos de fatores sociais
1. Cidade
2. Sexo/gênero
3. Faixa etária
4. Escolaridade
Grupos de fatores linguísticos
1. Sequência discursiva
2. <i>Status</i> da informação
3. Gradiência do tópico discursivo
4. Duração ³³
5. Telicidade
6. Dinamicidade
7. Aspecto gramatical

Quadro 3 – Grupos de fatores

A seguir, descrevem-se esses grupos de fatores pormenorizadamente e explicitam-se as respectivas hipóteses relativas ao uso de *aí*, *daí* e *então*.

2. Grupos de fatores sociais

De acordo com Freitag (2011), a Sociolinguística Variacionista vem contribuindo para uma ampla descrição do PB. Contudo, é preciso refletir sobre o quanto tal descrição espelha, de fato, a realidade social do Brasil em termos linguísticos, visto que os bancos de dados nem sempre são fiéis à estratificação da sociedade. Desse modo, é importante repensar os “fatores sociais clássicos” em

³³ As variáveis 4, 5 e 6 compõem o *aspecto lexical* e foram codificadas separadamente. Tal procedimento se justifica pelo fato de que cada uma dessas categorias é formada por subitens, aqui codificados como variantes.

fenômenos de variação e mudança, sobretudo naqueles atinentes a níveis mais complexos, em que nem sempre os efeitos dos fatores sociais são estatisticamente significativos.

Aqui, os grupos de fatores sociais visam a descrever e explicar tendências de uso linguístico em Campo Grande e São Paulo. Além de *sexo/gênero*, *faixa etária* e *nível de escolaridade*, inclui-se ainda a *cidade do informante*, a fim de permitir verificar sua possível correlação com a distribuição dos sequenciadores.

2.1. Cidade

Em seu trabalho sobre o emprego dos modos indicativo e subjuntivo, Santos (2012) analisou a variável *cidade de origem dos informantes*, com dados de fala das cidades do Rio de Janeiro e de São José do Rio Preto. O estudo atestou não haver diferenças significativas entre cariocas e paulistas no que concerne ao emprego do subjuntivo.

Em estudo similar, mas comparando dados de São Paulo e São Luís, divididos em três conjuntos – orações substantivas (*‘eu sempre achava que ele ia/fosse dar bronca na gente’*), relativas (*‘não tem gente que anda/ande de Porsche na Freguesia do Ó’*) e adverbiais (*‘é como se pra gente era/fosse normal’*) - Santos (2015:37) verificou diferenças nas distribuições das ocorrências de formas indicativas e subjuntivas entre paulistanos e ludovicenses. Tais diferenças são estatisticamente significativas para os três contextos de subordinação, mas, entre as orações substantivas, tal discrepância é maior. Assim, o indicativo tende a ser relativamente pouco mais usado em São Paulo do que em São Luís – nas orações adverbiais e relativas – mas bastante mais nas substantivas. Em suma, Santos (2015) afirma que ludovicenses e paulistanos mais se aproximam do que se distanciam quanto ao fenômeno analisado.

Estas pesquisas interessam aqui como inspiração para a análise comparativa entre Campo Grande (56-58) e São Paulo (59-61), no que toca aos usos de *aí*, *daí* e *então* no envelope de variação já definido:

(57) tinha horário pra chegar em casa... [não podia chegar depois das seis... um monte de coisa... *aí/daí/então* eu tive que me acostumar] (CG38F5-Emília)

(58) porque assim é tudo perto aqui lá já era um pouquinho complicado porque [é um pouquinho tudo longe as coisas *aí/daí/então* tem que ir de ônibus] (CG25FC-Adriana)

(59) minha mãe era mais flexível mas [meu pai bastante rígido *aí/daí/então* eu ...eu sempre tive muito medo dele né?] (CG35MS-Wesley)

(60) não sei como interagir porque tem muita gente *aí/daí/então* eu já começo a ficar ansiosa (SPFS28SPS-JanaínaB)

(61) mas eh elas têm que trabalhar num lugar e [vão trabalhar num lugar mais longe *aí/daí/então* tem normalmente vão correndo] (SPF24CVC-CarolinaA)

(62) minha mãe morreu quando eu tinha oito anos né? e eu não era de muita convivência com meu pai e [ele era do exército *aí/daí/então* ele é muito rígido entendeu?] (SPF20CPL-Lara)

Conforme visto anteriormente, a forma *então*, por ser ensinada na escola como *conjunção explicativa* (Cunha, 2001), pode gozar de maior prestígio social, relativamente a *aí* e *daí* (que, por sua vez, não são ensinadas na escola e são reconhecidas por serem mais coloquiais, senão vernaculares). Assim, a expectativa sobre a distribuição desses dados é a de que os campo-grandenses, por se definirem e acreditarem ser definidos, no cenário nacional, como caipiras, pudessem utilizar *então*, forma de mais prestígio social, com menos frequência que os paulistanos.

1.2. Sexo/Gênero

A fala de homens e mulheres se distingue não só no plano fisiológico, mas também no social. Anatomia e altura de voz são diferenças facilmente reconhecidas entre homens e mulheres e caracterizam de imediato o sexo do falante. Contudo, aspectos fisiológicos não constituem um ponto de partida válido para classificar indivíduos quanto ao gênero. A construção social do que é ser homem ou mulher se dá de modo muito mais complexo do que o que se verifica biologicamente.

Assim, o termo *gênero* remete a diferenças dos papéis sociais de homens e mulheres, que resultam da tradição cultural de cada comunidade. Como os papéis exercidos pelos indivíduos na sociedade são determinados não só por seus atributos fisiológicos, mas também pelos valores culturais e possibilidades sociais de uma comunidade, utiliza-se o termo *sexo/gênero*, de acordo com Paiva (2003) – ainda que, nesta tese e em muitos trabalhos sociolinguísticos, seja feita referência a homens e mulheres.

Na literatura sociolinguística, é possível atestar a relevância dessa variável em fenômenos de variação estável e de mudança linguística. Segundo Labov (2001), em casos de variação estável, as mulheres demonstram preferência pelo uso das formas de prestígio. Por outro lado, quando ocorrem casos de mudança linguística, elas seriam inovadoras e responsáveis pela difusão da variante não padrão.

É importante alertar, contudo, para a importância de avaliar a variável *sexo/gênero* com outras de natureza social (Paiva, 2003), tais como *faixa etária*. Tal fato se faz necessário pelo fato de que as mulheres mais idosas tendem a ser mais conservadoras se comparadas às mais jovens. Em contrapartida, entre os homens, tal diferenciação parece menos saliente.

Em São Paulo, o trabalho de Nascimento (2011) sobre o pronome *você/cê* revela diferenças no que tange ao uso dessas variantes por parte de homens e mulheres. Comparadas aos homens, estas empregam mais frequentemente a variante *cê* no lugar de *você*. Entretanto, a correlação entre *sexo/gênero* e variáveis

morfossintáticas parece não ser observada com frequência, pelo menos não tão comumente como no caso de variáveis fonético-fonológicas, que costumam indexar mais diretamente características sociais (Cheshire, 2005).

O estudo de Tavares (2003) revela não ser significativo este grupo de fatores, em seu estudo sobre o uso de *e*, *aí*, *daí* e *então* na sequenciação retroativo-propulsora, em Florianópolis. Similarmente, Macedo & Silva (1996) mostram que o tal grupo de fatores não se revelou relevante quanto ao uso de *aí* na fala carioca.

Cabe perguntar se *sexo/gênero* guarda correlação com a variável *aí*, *daí* e *então* em Campo Grande e São Paulo. Inicialmente, prevê-se que as mulheres tenderiam a usar *então* mais frequentemente, em relação aos homens, pelas razões anteriormente expostas: é possível que tal forma goze de prestígio, já que é aquela preconizada pela escola. Paralelamente, *aí* e *daí* seriam relativamente preferidos na fala masculina.

1.3. Faixa Etária

A *faixa etária* do falante pode indicar se a variável é estável na comunidade ou se está em processo de mudança. A estabilidade se caracteriza por um padrão em que a distribuição de dados é semelhante para mais jovens e mais velhos, enquanto a faixa intermediária exibe maior ou menor frequência da mesma variante. Já a mudança em progresso se verifica quando uma das formas (aquela chamada de inovadora) é progressivamente mais frequente, dos mais velhos aos mais jovens (Chambers & Trudgill, 1980:93).

A respeito dessa hipótese – chamada de tempo aparente – Labov (2008[1972]) afirma que o processo de aquisição do vernáculo termina por volta dos quinze anos e que, a partir disso, os usos da língua se estabilizam na fala de um indivíduo.

Além disso, é possível que, com o tempo, as formas tomadas como marcas de identidade atinjam a comunidade de fala como um todo, suplantando eventuais camadas/variantes mais antigas com as quais competiam, em um

processo que pode chegar até a extinção destas últimas (caracterizando uma situação de especialização por generalização).

A expectativa é que *aí* e *daí* sejam relativamente mais frequentes entre os falantes mais jovens (20 a 34 anos), que estão menos inseridos no mercado de trabalho comparativamente aos falantes das outras duas faixas. Por outro lado, é de se esperar que *então*, por ser a forma “escolar”, seja a preferida por pessoas da segunda e da terceira faixas etárias, justamente pela exposição de tais indivíduos ao “mercado linguístico” (Bourdieu, 1991).

1.4. Escolaridade

A *escolaridade* é uma variável frequentemente estudada em investigações sociolinguísticas por sua sempre possível relação direta com o uso de variantes padrão e não padrão. Além disso, essa variável permite avaliar, mesmo que de forma indireta, a correlação entre usos linguísticos e classes sociais (Rodrigues, 2009).

Com base nesta variável, normalmente é possível observar conceitos como prestígio e estigma, na medida em que o fato de que certos fenômenos são trabalhados na escola tem consequências para a variação na língua falada: formas preconizadas pela escola e gramática normativa costumam ser mais frequentes entre os mais escolarizados (Votre, 2003).

No caso de fenômenos discursivos, como o que se enfoca nesta pesquisa, a atuação das variáveis sociais costuma ser pequena (Macedo & Silva, 1996), mas, considerando-se o estatuto “escolar” de *então* (relativamente às outras duas formas), interessa verificar se esse panorama se confirma em relação aos sequenciadores aqui em foco, nas duas capitais.

Tavares (2003) verifica que, como o *aí*, o *daí* predomina entre indivíduos de nível primário, mas recebe um pouco mais de destaque na fala dos indivíduos de nível ginásial, comparativamente aos de formação primária ou média. A forma *então* é mais recorrente junto a informantes de níveis de escolaridade colegial e ginásial, como uma alternativa não estigmatizada de sequenciar informações, e é

bastante inibido na fala de informantes do nível de escolaridade mais baixo, o primário.

3. Grupos de fatores linguísticos

A seguir, descrevem-se as hipóteses que justificam as variáveis linguísticas incluídas nas análises do uso dos sequenciadores *aí*, *daí* e *então* no domínio referencial da causalidade.

3.1. Sequência discursiva

A variável *sequência discursiva* tem se mostrado relevante em muitas pesquisas que tratam de variação e mudança, sobretudo as de cunho discursivo. As sequências discursivas compõem os mais variados gêneros (notícia, *email*, carta, entre outros) e organizam o discurso, atuando como marcos definidores do gênero.

De acordo com Freitag et al (2009), a entrevista sociolinguística, de que fazem parte sequências discursivas como a narrativa, a expositiva, a descritiva e a argumentativa, pode ser considerada um gênero pertencente ao domínio do cotidiano.

Também segundo Freitag et al (2009), as tipologias constituídas para classificar as entrevistas sociolinguísticas não são consensuais e nem sempre partem de critérios consistentes. As linguistas consideram que os critérios de classificação das sequências discursivas são estabelecidos a partir do objeto linguístico em foco. Logo, estudos que abordam modalização, por exemplo, preocupam-se com critérios para o controle de sequências argumentativas, ao passo que trabalhos que se dedicam à sequenciação tendem a ter critérios mais refinados no que tange às sequências narrativas.

Outro fator considerado pelas autoras é que as perguntas feitas ao informante parecem desencadear sequências discursivas particulares. Assim, o controle do tipo de assunto seria um elemento importante para a classificação das sequências discursivas nas entrevistas sociolinguísticas. Essas considerações indicam

a necessidade de se delimitar com mais precisão o escopo da sequência discursiva, de modo que se confeccione uma proposta que abrigue qualquer fenômeno linguístico. Conforme proposto por Labov e Waletzky (1997), Labov (2008 [1972]), Gorski (1994) e Freitag (2009), nesta tese, opta-se pelo escopo temático como princípio de delimitação para as sequências consideradas, conforme se descreve abaixo.

3.1.1. Sequências narrativas

A sequência narrativa é frequente nas entrevistas sociolinguísticas, uma vez que o roteiro utilizado nas gravações induz o informante a relatar acontecimentos ocorridos em um tempo e em um local específicos. A sequência narrativa é constituída de fatos ocorridos no passado e que podem se prolongar por um determinado tempo.

Em (63) e (64), há dois exemplos dessas sequências: na primeira, os acontecimentos limitam-se ao passado; no segundo trecho, os fatos estendem-se aos dias atuais:

(63) porque a gente mo/ foi morar numa rua assim de trás de trás de trás de uma avenida principal lá você não via gente na rua *aí/dai/então* era muito difícil você ver gente andando ainda (SP24FCVC-CarolinaA)

(64) eu vi duas mulheres andando de mão dada... abraçada *aí/dai/então* você sempre fica assim... você sempre acaba olhando né? (SP41FCPL-RuthC)

Alguns estudiosos, como Valle (2001) e Tavares (2003), distinguem entre narrativa e descrição de vida. Nas palavras de Valle (2001:116), tais sequências são definidas, respectivamente, como “trecho constituído por relatos verbais de fatos particulares ocorridos no passado, como a conquista de uma corrida, o primeiro beijo, uma viagem etc” e “trechos em que são relatadas situações habituais ocorridas no passado, como idas do informante à escola, como passava as suas tardes, com o que costumava brincar etc”.

Nesta tese, todos os excertos em que o informante relata feitos no passado, que podem ou não se prolongar no presente, são caracterizados como *sequências narrativas*. O pressuposto é de que, seja qual for o tipo de passado em questão – perfeito, imperfeito ou mais-que-perfeito -, ambas as categorias apontadas acima relacionam-se com uma ação ou evento cuja execução envolve tempo, espaço e movimento.

O critério para a classificação dos excertos narrativos é a dinamicidade, ou seja, a existência de uma ação reportada por verbos. Independentemente da forma como o passado estiver representado por eles, ou seja, mesmo que os verbos estejam no pretérito imperfeito, a ação, o evento se sucede em determinado tempo e lugar, cuja interpretação, em certos momentos, dá-se em torno de fatos verbais plurais que são reiterados no passado.

Nos contextos narrativos, há abundância de ações decorrentes de experiências humanas no mundo concreto; assim, a despeito das sequências argumentativas, que giram em torno de discussões acerca de determinados assuntos, parecem exigir menor esforço cognitivo, tanto do ponto de vista do enunciador quanto do enunciatário (Tavares, 2003).

3.1.2 Sequências descritivas

Na entrevista sociolinguística, é comum que o entrevistado faça descrições com objetivos vários, dentre os quais estariam a demonstração, a explicação ou complementação de eventos relatados.

(65) ah tem aqueles bairros que [as pessoas são de uma classe mais... elevada né?... classe social alta... *ai/daí/então* essas pessoas têm um modo diferente de se vestir]... de... de conversar... de andar (CG21MS-Júlio)

Tavares (1999; 2003), Dal Mago (2001), Valle (2001), e Freitag (2003) controlaram, em suas pesquisas, excertos com *sequência descritiva*. Assim como em seus trabalhos, aqui se define uma sequência como descritiva quando se trata de um trecho em que se ocupa em caracterizar um evento, um indivíduo ou um

objeto. Há, no entanto, trechos em que se observa a sobreposição de sequências discursivas de caráter narrativo e descritivo, como no excerto a seguir:

(66) nos locais que não tinham a praça e as ruas não eram
asfaltada... no nas praças era só o terreno baldio onde que
prevalecia as drogas onde tinha as bocas né? *aí/daí/então*
eles aproveitavam o local e já as ruas que não eram
asfaltadas eram com barro *aí/daí/então* quando... chovia
alagava a... as ruas *aí/daí/então* entrava pra casa das
pessoas *aí/daí/então* entendeu? e quem mora na parte
baixa do bairro se prejudicava a cada dia que chovia
(SPF20CPL-Lara)

Em (66), não há somente uma sequência descritiva, mas antes um contexto narrativo, acrescido de um comentário do informante. Em casos como esse, opta-se por classificar o traço discursivo predominante na sequência – neste caso, o narrativo, visto que, a fim de relatar as atrocidades que ocorriam em seu bairro, a informante descreve como eram as ruas, as praças do local.

Assim, algumas marcas linguísticas, como uso de intensificadores e de adjetivos, podem encaminhar a uma interpretação avaliativo-referencial que, a um só tempo, transmite um conteúdo informacional e veicula uma orientação pragmática.

Nesse sentido, sequências descritivas, muitas vezes, têm função predominantemente avaliativa na construção do discurso, funcionando como um pano de fundo, o que explica e situa a ação (nos contextos narrativos) ou o que comenta e justifica os trechos argumentativos, a exemplo de (67):

(67) eu acho um bairro meio perigoso assim não/ não gosto
muito não ah [já aconteceu assalto *aí/daí/então* você
começa a ficar com medo né?] do do lugar (SP30FCPL-
AliceM)

Moreira de Sá (1999) salienta que a sequência descritiva se insere em contextos narrativos de forma natural e que isso poderia ser uma característica de descrição de ações. Em (67), a informante explica por que tem medo do bairro. Para isso, narra o assalto que já aconteceu no lugar e o qualifica como ‘perigoso’.

De acordo com Labov e Waletzky (1997), a avaliação é comumente encontrada em narrativas pessoais e apresenta-se por meio da atitude do enunciador diante do que se narra. Por outro lado, Moreira de Sá (1999) argumenta que o papel das sequências descritivas é de grande relevância, pois elas veiculam informações importantes para a compreensão dos acontecimentos e, por vezes, trazem também comentários avaliativos.

Assim sendo, as sequências descritivas encontradas nas entrevistas caracterizam-se linguisticamente por meio de intensificadores, adjetivos e tempos verbais e expressam não só conteúdo referencial, mas também avaliativo.

3.1.3. Sequências expositivas

É comum comparar a *sequência expositiva* à narrativa. A esse respeito, Moss (2004) afirma que os textos narrativos são caracterizados por contar uma história, ao passo que o texto expositivo visa ao relato de informação. Colomer & Camps (2002:53), por sua vez, argumentam que “os textos expositivos apresentam uma organização textual muito distinta daquela dos narrativos”.

Desse modo, a sequência expositiva é caracterizada pela ordenação de conceitos lógicos, de modo imparcial, e, nas palavras de Marcuschi (2002:29), marcada por “sequências analíticas ou explicitamente explicativas”. Diferentemente da argumentação, nesse tipo de texto, não há a adesão a um ponto de vista. O objetivo é fazer saber, a exemplo de (68):

(68) minha patroa trabalha é na área cível né? *aí/daí/então*
ela lida muito com divórcio com inventário sabe?
(SP48FCES-MeireC)

No trecho acima, MeireC expõe a área de atuação de sua chefe, advogada. Como se nota, é apenas uma explicação, isenta de posicionamento ou adesão ideológica.

Na superfície linguística, notam-se tempos verbais do mundo comentado - presente, pretérito perfeito composto, futuro do presente, futuro do presente composto, além das locuções verbais formadas com esses tempos, todos no

indicativo -, que atuam como um alerta para advertir o ouvinte de que se trata de algo que o afeta diretamente e de que o discurso requer sua resposta, mesmo que de modo não verbal (Koch, 1998).

Outro elemento linguístico da sequência expositiva é a presença do interdiscurso, definido em função da interpelação ideológica – o “pré- construído”, no nível da memória, a imaterialidade dos saberes, as formulações imaginárias necessárias para suscitar as estratégias do discurso.

Nesse tipo de sequência discursiva, o objetivo central do emissor é discorrer sobre determinado assunto, com base em recursos como a definição, a comparação, a informação e a enumeração.

3.1.4. Sequências argumentativas

Em virtude da natureza da entrevista sociolinguística, também é comum encontrar contextos argumentativos. Givón (1993) destaca, acerca dessa sequência, a exposição de opiniões do falante sobre determinado fato ou ideia, correlacionando-se com o tempo presente e o aspecto imperfectivo, durativo e incompleto.

Em (69), a informante defende uma postura menos tradicional dos professores face aos alunos do ensino básico, de modo que seja possível adequar o ensino ao perfil dos discentes.

(69) [sexto ano por exemplo dificilmente ele vai conseguir ficar sentado na cadeira prestando atenção na explicação quadro e lousa ali entende? *ai/daí/então* esbarra com os professores tradicionais] que são resistentes ao uso da tecnologia (CG39F5-Renata)

De acordo com Tavares (1999), a sequência discursiva de caráter argumentativo costuma ser definida como exposição ou justificativa do ponto de vista do falante sobre determinado tema. Já Valle (2001:117) salienta que as sequências argumentativas costumam ocorrer no tempo presente, acompanhadas de modalizadores, como *eu acho* e *para mim*.

Segundo Guy et al (1986), as sequências argumentativas estariam divididas em opiniões e explicações. Sequências explanativas são aquelas em que o falante revela o motivo ou a razão de determinados assuntos, enquanto as opinativas são se caracterizam pela expressão de um posicionamento.

Nesta tese, não se utiliza a tipologia de Guy et al (1986) para classificação das sequências argumentativas. Aqui, o critério definidor é o escopo da avaliação. Em linhas gerais, são classificadas nessa categoria, todas as ocorrências em que se observa uma tomada de atitude do narrador.

Com base no princípio da marcação, o esquema a seguir (Tavares, 2003) exhibe um contínuo que parte da sequenciação de eventos passados, típicos da narrativa; passa pela ordenação de características típicas de um ser, um objeto ou um acontecimento, caros à descrição; segue para a exposição de fatos e, por fim, termina com a ordenação de argumentos e opiniões, típicas da argumentação.

Desse modo, considera-se a narração como um polo e a argumentação como outro. Em posições intermediárias, estão a descrição e a exposição. Esta, por envolver a manifestação de opiniões, é considerada mais complexa que aquela, constituída pela mera exposição de características de um elemento.

narrativa	descrição	exposição	argumentação
- marcado			+ marcado

Diante de tal raciocínio, é possível aventar, de acordo com o princípio da iconicidade, as seguintes hipóteses: a narrativa, definida como menos complexa, tenderia a ser sequenciada por formas menos marcadas (Givón, 1995), ou seja, formas estruturalmente menores como *aí* e *daí*. Por outro lado, a argumentação, sequência que exige maior dispêndio cognitivo, deve favorecer *então*, forma de maior substância fônica, portanto, mais marcada.

3.2. *Status* da informação

A análise do *status* informacional foi feita considerando a relação de *status* informacional entre a sequência discursiva causal e aquela que traz a ideia de consequência. Esse método pressupõe a progressão da informação no discurso (de *velha*, passando por *inferível*, até a *nova*).

Na classificação do *status* informacional das sequências discursivas causais, adotou-se um critério estritamente textual. O discurso do falante é o ponto de partida para a definição do grau de informatividade, classificada como *nova*, *velha* e *inferível*.

Informação nova traz um conteúdo não mencionado e não inferível da sequência discursiva anterior à que leva a noção de causalidade (70); já *informação velha* se define como uma sequência discursiva que carrega informação já mencionada (71):

(70) tenho outro [irmão] aqui no em Dois Irmãos do Buriti tem um que mora com a minha irmã que ele é ele não é paralítico né? mas ele tinha tem problema epilético e eu não sei como que foi que os dedo dele engrossou os dedo do pé ele não firma quase no pé e a mão quebrou o braço ficou torto e os dedo a/ atrofiou o dele dele n/ aí/dai/então vive com a minha irmã (CG55MC-Alberto)

(71) agora a cidade tá crescendo está aumentando... criminalidade essas coisas mas... eu acho que... tudo isso depende do... do... do poder público... de como ele... controla a situação... aqui... [não digo que eles façam certo... mas eles investem... uma quantia legal em segurança... a gente se sente seguro aí/dai/então... dá aquela sensação de segurança] não dá... não é igual você andar no Rio de Janeiro... e ter medo de ser assaltado por... dentro de um túnel essas coisas (CG21MS-Júlio)

Em (70), o informante fala de seus irmãos e conclui o assunto introduzindo uma relação de causa-consequência, ao expor que, devido a um problema no pé, um de seus irmãos vive com uma irmã, que cuida dele. Ao enunciar a sequência

discursiva *vive com a minha irmã*, o informante expõe, portanto, informação nova. Em (71), tanto a causa como a consequência já foram mencionadas pelo falante e somente são retomadas. A informação pode também ser inferida do discurso anterior, embora não seja diretamente mencionada. O trecho (72) ilustra esse caso:

(72) não vou dizer assim que a mulher está cem por cento porque não está... porque tem muita mulher ainda que apanha do marido mulher que cuida de filho e tem muita mulher também que não pensa né? que [vai arruma um filho e não tem marido não tem sustento *aí/daí/então* fica batendo cabeça] eu acho isso tem muito muito em Cam/ quer dizer no Brasil inteiro eu acho que tem né? (CG66FC-Eliana)

No trecho acima, Eliana relata, em tom crítico, a postura de mulheres solteiras que engravidam. Com essas informações, prepara o interlocutor para a consequência que pode ser inferida a partir desse comportamento: a situação financeira complicada com que se deparam depois de algum tempo.

Segundo Paiva (1997), a ativação informacional pode ser bem compreendida com mais clareza se estudada à luz da tradição que alude à Escola de Praga. Nessa perspectiva, a língua é um conjunto de meios empregados com vistas a objetivos pré-determinados, dentre os quais se destaca a transmissão de informação.

Paiva (1997), por sua vez ressalta que o princípio de dinamismo comunicativo é reforçado na organização linear da sequência discursiva. Desse modo, há, inicialmente, uma sequência discursiva que carrega elementos de baixo grau na hierarquia comunicativa, passando, de modo gradual, elementos de grau mais alto.

O falante organiza seu discurso em consonância com expectativas acerca das representações mentais já ativadas, não ativadas ou semiativadas na mente do ouvinte. Chafe (1998) reconhece que os esquemas mentais pré-existentes ao discurso e a capacidade inferencial do ouvinte são fatores preponderantes na estruturação do fluxo discursivo.

É preciso ressaltar que as abordagens cognitiva e textual se equivalem somente de modo parcial. Logo, uma entidade nova, sob o prisma do ouvinte, é também nova no texto. Em contrapartida, uma entidade nova no texto pode ser velha se tomada pelo ponto de vista daquilo que já está ativado na mente do ouvinte.

Tais fatos não tornam incompatíveis as duas perspectivas. As propostas que se prestam a definir os limites da ativação informacional das sequências discursivas consolidam um princípio funcional que prevê que, em situações não marcadas, informação velha precede informação nova (Halliday, 1985; Chafe, 1987). Desse modo, a distribuição de informação está atrelada à posição das unidades linguísticas: em primeiro lugar, estão as que trazem uma informação velha e em seguida, aquelas cuja informação é nova.

A hipótese que está por trás desse grupo de fatores inspira-se no princípio da iconicidade (Givón, 1990), segundo o qual quanto mais imprevisível (mais nova) for a informação, maior será a quantidade de material linguístico utilizado. Desse modo, *então* poderia ser a forma sequenciadora preferida nesses casos. Por outro lado, *aí* e *daí*, relativamente mais diminutas, seriam as formas mais propensas ao emprego no sequenciar de informações velhas ou inferíveis.

3.3. Gradiência do tópico discursivo

De acordo com Risso (1990), a organização do tópico pode ser observada em dois níveis: dentro do próprio tópico – *intratópico* (73) – ou entre dois tópicos – *intertópico* (74). No primeiro nível, o elemento discursivo estabelece conexões que se limitam a um tópico em especial, o que resulta na organização de sua estrutura interna. Essa posição interna decorre da propriedade de prosseguir com seu desenvolvimento. Por outro lado, quando há intertópico, nota-se que o sequenciador parte de uma informação já enunciada para instaurar um pronunciamento novo.

(73) o avô pai do meu pai morreu muito cedo... num acidente de carro *aí/daí/então* ficou a minha vó cuidando de todo mundo ali (CG28FS-Débora)

Em (73), antes de falar sobre o tópico “avó”, a informante discorre sobre outro assunto: a morte do avô. Nota-se que *então* sequencia uma relação intertópica. Em (74), Joana relata como era a rotina da família durante a infância de sua filha. A relação de causalidade é intratópica e gira em torno do crescimento da filha, com a consequente mudança de hábitos do casal. Logo, o mesmo tópico conversacional é mantido.

(74) mas antes era mais a questão do cerrado as frutas que a gente saía procurava [agora não agora minha filha cresceu *aí/daí/então* a gente já nem sai quase] (CG45FC-Joana)

Por ser o intertópico, relativamente ao intratópico, uma sequência discursiva que exige maior esforço cognitivo, já que se caracteriza pelo encaminhamento de um novo conteúdo referencial, aventa-se, com base no princípio da iconicidade, a hipótese de que *então*, seja a forma preferida nesse contexto, por ser a forma com mais material linguístico. Por outro lado, *aí* e *daí* (com menos material fônico) podem ser mais comumente utilizados na sequenciação da causalidade intratópica.

Segundo Marcuschi (1991), as conversações iniciam-se com o tópico que motivou o encontro. Sendo assim, o tópico é a mola propulsora de uma interação. Definido como aquilo de que trata o texto, é o princípio central da organização discursiva. Essa organização se dá por meio da coerência conversacional. Portanto, existem regras no interior da conversa. Elas constituem um consenso entre os falantes para que a comunicação seja funcional e coerente.

O processo colaborativo entre os interlocutores, com a articulação de um turno após o outro, faz com que a conversa seja mantida de forma coesa e coerente, em torno de um conjunto de referências que se projeta como foco da interação. Assim, o tópico discursivo passa a ser entendido como o elemento que compõe, organiza e fundamenta o texto falado, realizando recortes de segmentos textuais, a fim de que o analista identifique o que se fala e como isso está disposto no texto. Ao demarcar a estrutura de um texto, o tópico torna-se responsável pela organização discursiva, instaurando seu fio condutor.

3.4. Aspecto lexical do verbo introduzido pelo sequenciador

Comrie (1976) entende que o aspecto se define em função dos diferentes modos de observar a constituição temporal interna de uma situação. Desse modo, a expressão “constituição temporal interna”, utilizada para a definição de aspecto, pode ser compreendida em termos da oposição entre “tempo interno da situação” e “tempo externo da situação”. Para Comrie (1976), o “tempo interno da situação” diz respeito a aspecto, enquanto o “tempo externo” diz respeito a tempo, no sentido mais clássico em que é estudado pela gramática tradicional.

O termo *aspecto* pode referir-se a traços de natureza muito distinta. O *aspecto lexical*, que Comrie (1976) nomeou como *aspecto semântico*, refere-se às propriedades aspectuais atinentes às raízes verbais e a outros itens lexicais utilizados pelo enunciador, a fim de descrever uma dada situação. As distinções de aspecto nessa categoria não seriam expressas por marcas gramaticais visíveis.

Por outro lado, o *aspecto gramatical* atua nas distinções aspectuais que são marcadas explicitamente na morfologia, geralmente por auxiliares e/ou morfemas flexionais e derivacionais.

De acordo com Comrie (1976), o aspecto semântico – aqui chamado de *aspecto lexical* – realiza-se independentemente do tempo de referência e de qualquer marca morfológica. Entretanto, essa questão gera controvérsias entre os pesquisadores, já que alguns sugerem que há certa interdependência entre *aspecto lexical* e alguns afixos que codificam distinções condizentes com *aspecto gramatical*.

O modelo de classificação formulado por Vendler (1967) e expandido por Smith, 1991) parte de três pares básicos de valores para o *aspecto lexical*: *estatividade X dinamicidade*, que diz respeito ao fato de descrever, respectivamente, um estado fixo em um período de tempo ou uma sucessão de estados ou estágios de um processo ao longo do tempo; *telicidade X atelicidade*, relativo à possibilidade de um predicado apresentar, respectivamente, um término ou não; e *pontualidade X duratividade*, que se liga à possibilidade de um predicado apresentar um evento que não se prolonga no tempo ou um evento ou estado cuja duração se prolonga por um determinado período de tempo. Assim,

foram consideradas três variáveis na análise dos dados: *dinamicidade*, *telicidade* e *duração*.

Nas análises realizadas na presente tese, o verbo codificado é aquele sobre o qual recai a noção de causalidade, no caso, o aquele que é introduzido pelo sequenciador. Assim, a análise de tal variável se faz necessária para avaliar se a escolha do sequenciador se correlaciona com propriedades semânticas do verbo introduzido pelo sequenciador – o que levaria à seleção das categorias de aspecto lexical aqui elencadas - ou com propriedades de natureza morfossintática, o que seria equivalente à seleção de *aspecto gramatical*. Por outro lado, considerando-se que esses dois tipos de aspecto sejam interdependentes, seria possível esperar que as duas variáveis fossem selecionadas.

A respeito desses grupos de fatores, que compõem o *aspecto lexical* do verbo, podem-se aventar algumas hipóteses, com base no princípio da persistência (Hopper, 1991). Desse modo, o que se considera em tais hipóteses, é que a escolha dos verbos, no que diz respeito aos seus traços semânticos, teria relação com resquícios dêiticos de *aí*, *daí* e *então* em sua função sequenciadora:

- (i) *aí* e *daí* seriam utilizados com mais frequência diante de verbos *durativos*. Se comparados a *então*, são, em sua origem, formas espaçotemporais, diferentemente de *então*, cujo uso no português é marcadamente temporal (Capítulo 2). Logo, *aí* e *daí* teriam maior ligação com a noção de espaço, o que poderia levar à crença de que seriam formas com apelo mais estativo que *então*;
- (ii) a frequência de *então* seria maior diante de verbos *dinâmicos* e *télicos*, cujo uso estritamente temporal na língua portuguesa influenciaria o uso de verbos cujo aspecto carreguem a noção de intermitência temporal.

A seguir, serão detalhadas as variantes que compõem cada uma das três categorias constituintes do *aspecto lexical*, de acordo com a proposta de Vendler (1967), adaptada por Smith (1991).

3.4.1. Dinamicidade (estatividade X dinamicidade)

Em português, os verbos de eventos são associados à ideia de movimento e se caracteriza por denotarem um deslocamento, não necessariamente obrigatório. Nesse grupo, o fato que os une é o de, lexicalmente, não denotarem uma direção intrínseca, mas sim uma forma ou modo de movimento, em que a delimitação de uma direção é obrigatória e se dá por meio de um sintagma direcional.

Segundo Leal e Oliveira (2007), uma análise mais atenta dos verbos de movimento em português revela, de fato, a necessidade de uma subdivisão. Esses linguistas defendem que o verbo *caminhar*, por exemplo, comporta-se de forma distinta de *vaguear*, no que tange à especificação da direção através de um SP, como atesta o trecho a seguir (Leal & Oliveira, 2007:287):

- (i) a. O João caminhou para casa.
- b. * O João vagueou para casa.

Nesta tese, contudo, os verbos de movimento são denominados *dinâmicos*, serão inclusos, para efeito de análise, em um único grupo geral (75), que se opõe aos verbos de estado, ou *estativos*, como rotulados aqui.

(75) são bem limitados têm problema de diabete
[minha mãe meu pai tem problema de pressão alta
já já sofreu derrame no no... passado né? muito
tempo atrás mas... não ficou com sequela não mas...
(tosse) já anda com dificuldade tem prolema de
pressão aí/dáí/então minha irmã cozinha cuida³⁴ da
casa pra eles] (CG39MC-Ziraldo)

Basso & Ilari (2004) afirmam que, desde a classificação aristotélica, a classe dos verbos estativos sempre foi vista como inteiramente homogênea e distinta das

³⁴ Os verbos introduzidos pelo sequenciador, sob os quais recai o foco da análise, foram sublinhados para melhor visualização.

demaís. Os verbos estativos têm sido caracterizados como “verbos que não indicam uma ação”, e, de modo geral, são acompanhados por um sujeito experienciador, e não agente. Basso e Ilari (2004:16) citam como exemplos dessa grupo os verbos *parecer(-se) com, possuir, localizar(-se), saber, amar e ter fome*, em que não há controle por parte do sujeito sobre o estado de coisas referido pelo predicado, tal como em (76):

(76) quando a minha mãe/ a gente foi morar nesse lugar [a minha mãe... recebia pouco a gente mal comia minha [minha irmã saiu de casa *aí/daí/então* eu fiquei triste triste triste triste triste] chegou uma hora que eu chorava muito mas eu falava ‘nossa que vida’ (SPF24CVC-CarolinaA)

No excerto acima, *ficar triste* é um predicado não agentivo, experienciado por CarolinaA, diferentemente do exemplo anterior, em que *cuida* expressa uma ação em que a pessoa, no caso, a irmã do informante não é experienciadora, e sim, agente.

3.4.2. Telicidade (telicidade x atelicidade)

Segundo Comrie (1976), a *atelicidade* é comum nos estados e eventos que não se caracterizam por um limite temporal bem definido. Já a *telicidade* é expressa pelos eventos em que um processo se desenrola até um ponto além do qual não será possível haver continuidade.

Em (77), o verbo *conheço*, atélico, apresenta um ponto de encerramento.

(77) realmente ali onde eu moro né? não sei se é porque [eu fui criada ali né? já.. já... fui criada *aí/daí/então* eu conheço a maioria do pessoal] (CG52MC-Adélia)

Já em (78), o verbo *fechou*, télico, expressa um fim inerente, isto é, um ponto em que a ação chega a seu fim natural: JoséN relata que financiou seu apartamento devido às facilidades que encontrou na época.

(78) na época estava barato tal [o cara facilitava e tinha financiamento tinha não sei o quê ah

aí/daí/então a gente fechou] e eu fiquei meio com medo porque assim questão de um ano ou dois atrás aconteceram duas coisas com o prédio (SPM46SPN-JoséN)

3.4.3. Duração (duratividade x pontualidade)

Duratividade e *pontualidade* dizem respeito à presença ou ausência de intervalos internos. Os verbos podem conter semanticamente o traço de momentaneidade ou duratividade, porém essa característica original pode ser alterada pela contextualização.

De acordo com Travaglia (1994), é possível questionar a existência de situações pontuais pelo fato de que qualquer evento ou ação tem uma duração, mesmo que pequena. No trecho abaixo, o verbo *brincava*, definido como *durativo*, alude a uma trajetória que pode ser percebida como tendo uma certa duração:

(79) é uma rua bem calma assim *aí/daí/então* a gente brincava bastante na rua assim (SPM27SEL-SérgioA)

Por outro lado, o verbo *acordar*, no trecho a seguir, denota uma ação com início e fim, definida, portanto, como *pontual*:

(80) ela trabalha longe *aí/daí/então* ela acorda muito cedo né? (SPM31SPS-CarlosJ)

Travaglia (1994), no entanto, assume que o que importa não é a medida de tempo em termos absolutos, mas o sentimento linguístico do falante, o qual concebe a situação como pontual.

3.5. Aspecto gramatical do verbo introduzido pelo sequenciador

Para Castilho (1968), o *aspecto verbal* – em sua acepção gramatical - é a visão objetiva da relação entre o processo e o estado expressos pelo verbo e a

ideia de duração ou desenvolvimento. Em suma, é a representação espacial do processo.

Se ação verbal indica uma duração, o aspecto é imperfeito; se a ação é cumprida, o aspecto é perfeito. Em caso de ação repetida, há o aspecto resultativo. Em último caso, se o verbo assume uma característica virtual, imune à atualização por qualquer categoria, o aspecto é indeterminado.

De acordo com a proposta de Castilho (1968), na língua portuguesa, é possível delimitar quatro aspectos principais, devidamente expressos no quadro a seguir:

Valor	Aspecto
Duração	Imperfeito
Completamento	Perfeito
Repetição	Resultativo
Negação da duração e do completamento	Indeterminado

Quadro 4 – Aspecto gramatical (Castilho, 1968)

No interior de cada modalidade de aspecto, há noções que não podem ser desprezadas. Algumas delas não são necessariamente aspectuais (representam, antes de mais nada, "modos da ação"), uma vez que escapam à oposição duração/completamento. Por aderir aos aspectos, acabam por caracterizá-los. Dizem respeito aos conceitos de mudança de estado, progressão, resultado e cessamento, que não serão consideradas na análise desta tese. A codificação de cada uma dessas ocorrências resultaria em um número pouco expressivo de ocorrências por variante, o que impossibilitaria uma análise estatística confiável.

A exemplo das hipóteses levantadas para a variável anterior, aqui, também se considera o princípio da persistência (Hopper, 1991):

- (i) *aí* e *daí*, formas que têm, relativamente a *então*, maior ligação com a noção de espaço, conforme já explorado no item acima, seriam

utilizados com mais frequência diante de verbos em aspecto perfectivo e indeterminado.

- (ii) a frequência de *então* seria maior diante de imperfectivo e resultativo, em virtude do uso exclusivamente temporal, se considerada sua função de advérbio.

3.5.1. Aspecto perfectivo

O processo é apresentado em sua totalidade, com começo, meio e fim. Para Lyons (1997), a propriedade básica do aspecto perfectivo é o fato de apresentar a situação como um todo em si mesmo, o que faz com que sua análise seja incompatível com a proposta de fases internas particulares de um evento.

Comrie (1976: 4) avalia que “o perfectivo olha para a situação de fora, sem necessariamente distinguir quaisquer das estruturas internas da situação, enquanto que o imperfectivo olha para a situação de dentro e, como tal, preocupa-se crucialmente com a estrutura interna da situação”. Essa definição é verificada a seguir:

(81) ele tinha u/ uma desconfiança total de todo mundo achou que eu estava roubando dele
aí/daí/então me mandou embora (CG52MC-Fausto)

No trecho acima, Fausto diz por que foi demitido. O verbo utilizado pelo informante - *mandou* - enuncia um processo pontual, finito.

3.5.2. Aspecto imperfectivo

Está relacionado à ação incompleta, de cuja duração não é possível delimitar princípio e fim, apresentando, desse modo, o processo em seu curso (Castilho, 1968). Desse modo, o enfoque não reside sobre a ação como um todo, mas sobre diferentes fases dela:

(82) o meu pai era... ele tinha uma vida... ele era marceneiro ele tinha uma marcenaria tal...
aí/daí/então ele... *dava* estudo dava educação assim para nós né?] (CG66MC-Durval)

Durval relata o ofício do pai e diz que, a partir de renda obtida na marcenaria, ele e os irmãos tinham boas condições de vida. Nota-se que o verbo utilizado na sequência discursiva causal - *dava* - expressa um processo ininterrupto, não sendo, passível, portanto, de segmentação temporal.

3.5.3. Aspecto indeterminado

Caracteriza-se por não ser nem imperfeito nem perfeito. É um aspecto neutro, revestido de certa virtualidade. Intitulado por Said Ali (1964:161) e Cunha (1975:430) de presente “durativo” e, por Dias (1970:184), de “presente Universal”, designando a ação que se desenrola em todo o tempo. De acordo com Travaglia (2006:35), caracteriza-se por não ser nem perfeito nem imperfeito e foge às expressões de aspecto e de tempo.

(83) realmente ali onde eu moro né? não sei se é porque [eu fui criada ali né? já.. já... fui criada *aí/daí/então* eu conheço a maioria do pessoal] (CG52FC-Adélia)

No excerto acima, relata que o fato de conhecer a maioria do pessoal do bairro é consequência de morar no local desde que nasceu. Nesse caso, *conhecer* é atemporal.

3.5.4. Aspecto resultativo

Caracteriza-se pela repetição da ação e é considerado um coletivo de ações (Travaglia, 2006:35), que podem ser durativas ou pontuais. De acordo com Castilho (1968), seria um aspecto intermediário entre o perfeito e o imperfeito, normalmente expresso por perífrases (Vendryes, 1921:117).

No excerto abaixo, por exemplo, a informante relata a reação das pessoas diante de um tipo de acontecimento rotineiro, como o vivenciado por ela:

(84) o tempo vai passando cada um vai... indo pra um canto né? *aí/daí/então* você acaba deixando é

mais... bom dia boa tarde boa noite... né?
(SP30FCPL-AliceM)

Ao relatar o afastamento de seu grupo de amigos, AliceM usa o aspecto resultativo para reportar a que se limitou a convivência entre eles. O Quadro a seguir resume os grupos de fatores, com suas respectivas variantes:

Grupos de fatores estratificados na amostra	
Cidade	Campo Grande São Paulo
Sexo/Gênero	Feminino Masculino
Faixa etária	20-34; 35-49/59; 50/60 ou mais
Escolaridade	Média Superior
Outros grupos de fatores	
Sequência discursiva	Narração Descrição Exposição Argumentação
Status da informação	Velha Inferível Nova
Gradiência do tópico discursivo	Intratópico Intertópico
Duração do verbo	Durativo Pontual
Telicidade do verbo	Télico Atélico
Dinamicidade do verbo	Dinâmico Estativo
Aspecto gramatical	Perfectivo Imperfectivo Indeterminado Resultativo

Quadro 5 – Resumo dos grupos de fatores

5

Análises quantitativas

Reportam-se, neste capítulo, os resultados das análises quantitativas de *aí*, *daí* e *então* em contexto de causalidade no domínio referencial, extraídas das amostras campo-grandense e paulistana.

No intuito de testar as hipóteses aventadas e descritas no Capítulo 4, foram realizadas três análises: uma *global*, com dados de ambas as capitais, a fim de verificar se é significativa a diferença na distribuição dos dados entre as duas cidades; uma apenas com os dados de São Paulo; e outra apenas com os de Campo Grande. No geral, os resultados revelam que as diferenças entre campo-grandenses e paulistanos não são significativas e abre a discussão em torno da questão do prestígio social no que tange a variáveis discursivas.

1. Da perspectiva de *então*...

No Capítulo 1, definiu-se o envelope de variação, ou seja, os contextos em que *aí*, *daí* e *então* funcionam como variantes de uma variável – a sequenciação no domínio referencial da causalidade (Sweetser, 1991; Paiva, 1995). No Capítulo 4, foram estabelecidos os grupos de fatores que hipoteticamente se correlacionam ao emprego variável dessas formas nesse envelope. O presente capítulo apresenta os resultados das análises quantitativas que testam tais hipóteses, com as ocorrências extraídas das 96 entrevistas sociolinguísticas (48 com campo-grandenses, 48 com paulistanos).

A análise quantitativa é feita da perspectiva de *então* (ou seja, estabeleceu-se tal forma como valor de aplicação nas análises estatísticas), uma vez que a hipótese inicial é a de que *então* seria, em princípio, uma forma que goza de maior prestígio, na medida em que tem um estatuto mais “escolar”, como visto em capítulos anteriores. Nesse sentido, se um fator de um grupo favorece *então*,

significa que tal fator desfavorece *aí/daí*. Da mesma forma, se um fator desfavorece *então*, significa que favorece *aí/daí*.

Para a realização das análises quantitativas, procedeu-se ao amálgama de *aí* e *daí*, sob o rótulo de *aí*, analisando tais ocorrências em contraposição a *então*. Tal procedimento justifica-se por três razões: (i) Em primeiro lugar, porque somente seria possível realizar, com o Rbrul, análises com variáveis binárias; (ii) do ponto de vista morfológico, *daí* se assemelha mais a *aí* que a *então*; (iii) são poucos os dados de *daí* (30), relativamente aos outros dois sequenciadores (134 casos de *aí* e 542 de *então*). Embora sua ocorrência não seja numericamente expressiva, não é recomendável descartá-las já que, na fala de alguns informantes, seu uso é predominante (ver tabelas 5 e 6).

2. Análise global

A Tabela 3 abaixo mostra a distribuição geral dos dados, nas amostras de cada cidade. São 357 as ocorrências nas entrevistas campo-grandenses e 349 nas paulistanas. A forma *então* ocorre em 73% dos dados na primeira amostra e em 80% no caso da segunda.

Tabela 3- Distribuição geral dos dados

Cidade	<i>Aí</i>		<i>Daí</i>		<i>Então</i>		Total
	N	%	N	%	N	%	
Campo Grande	78	22	18	5	261	73	357
São Paulo	56	16	12	4	281	80	349

Como os totais gerais de dados para cada cidade são bastante semelhantes entre si, pode-se já visualizar (especialmente no gráfico da Figura 10) que há mais casos de *então* na amostra paulistana do que na campo-grandense. Interessa verificar se tal diferença é estatisticamente significativa. Nesse sentido, o histograma abaixo permite uma melhor visualização das ocorrências dos sequenciadores em cada cidade:

Figura 10 – Percentuais de *aí + daí* e *então* nas amostras campo-grandense e paulistana

Todos os grupos de fatores estabelecidos foram incluídos nesta análise global (mais adiante, discutem-se os resultados para aqueles que foram selecionados como estatisticamente significativos). Aqui, interessa destacar o fato de que a diferença percentual visualizada no gráfico não é significativa. Com pesos obtidos de uma análise *one-level*, a tabela abaixo revela que as tendências (.46 para Campo Grande e .53 para São Paulo) são mais próximas do que distantes entre si.

Tabela 4 – Tendências de uso para a variável *cidade* (análise *one-level*)

Cidade	% <i>então</i>	P.R.	Total
Campo Grande	49	.46	357
São Paulo	51	.53	349

Input: 0,384. p < 0,05

Conforme a tradição variacionista, quando se desenvolvem análises sociolinguísticas, o intuito é descrever e explicar padrões de uso (Guy, 2001). Vale questionar, contudo, se tais padrões seriam uma média de padrões individuais (que

poderiam divergir entre si), ou se se trata, de fato, de um padrão observado na “gramática de cada falante”.

Dentre as premissas que embasam a análise da regra variável (Labov, 1969; Cedergren & Sankoff, 1974; Sankoff & Labov, 1979), pode-se ressaltar, de acordo com Guy (1991), que, no nível individual, pode haver divergência em relação ao *input*, ou seja, às taxas de aplicação de uma regra variável. Assim, de acordo com Oushiro (2015:201), é possível que, na fala de dois falantes de uma mesma comunidade, encontrem-se frequências de uso de uma variante de 5% ou 75%:

Tabela 5 – Distribuição dos sequenciadores por informante³⁵ – Campo Grande

Informante	<i>Então</i>	%	<i>Aí + Daí</i>	%	Total
Adriana(1) ³⁶	4	1	4	1	8
Camila(1)	1	0,5	1	0,5	2
Celina(1)	6	1,7	2	0,5	8
Débora(1)	6	1,7	-----	----	6
Eduardo(1)	3	0,8	2	0,5	5
Everton(1)	5	1,4	-----	----	5
Fábio(1)	3	0,8	-----	----	3
Júlio(1)	5	1,4	5	1,4	10
Luísa(1)	3	0,8	1	0,2	4
Marcelo(1)	4	1,1	2	0,5	6
Paula(1)	2	0,5	1	0,2	3
Raimundo(1)	-----		3	0,8	3
Raul(1)	17	4,7	1	0,2	18
Renan(1)	4	1,1	-----	----	4
Tainá(1)	6	1,7	2	0,5	8
Talita(1)	6	1,7	2	0,5	8
César(2)	3	0,8	2	0,5	5
Emília(2)	11	3	5	1,4	16

³⁵ Embora o Rbrul permita a inclusão do informante como variável aleatória num modelo de efeitos mistos, problemas técnicos com o *software* não possibilitaram tal análise. Por isso, a distribuição dos dados por informante é aqui apresentada no intuito de ilustrar a grande variabilidade no emprego das formas.

³⁶ O número que acompanha o nome indica a faixa etária do informante.

Frederico(2)	8	2,2	1	0,2	9
Joana(2)	9	2,5	7	1,9	16
Lúcio(2)	7	1,9	3	0,8	10
Marlene(2)	4	1,1	2	0,5	6
Nair(2)	6	1,6	-----	----	6
Nícolás(2)	4	1,1	4	1,1	8
Osmar(2)	14	3,9	4	----	18
Regina(2)	6	1,6	3	0,8	9
Renata(2)	7	1,9	6	1,6	13
Ronaldo(2)	3	0,8	-----	----	3
Taís(2)	11	3	1	0,2	12
Wesley(2)	11	3	1	0,2	12
Ziraldo(2)	5	1,4	-----	----	5
Zuleica(2)	2	0,5	4	1,1	6
Adélia(3)	3	0,8	1	0,2	4
Alberto(3)	1	0,2	1	0,2	2
Bento(3)	2	0,5	2	0,5	4
Durval(3)	13	3,6	2	0,5	15
Eliana(3)	4	1,1	-----		4
Fausto(3)	3	0,8	1	0,2	4
Glória(3)	-----	----	1	0,2	1
Marta(3)	2	0,5	5	1,4	7
Maurício(3)	4	1,1	1	0,2	5
Mauro(3)	2	0,5	5	1,4	7
Olga(3)	-----	----	-----	----	-----
Pedro(3)	-----	----	-----	----	-----
Roberto(3)	-----	----	-----	----	-----
Rosana(3)	4	1,1		----	4
Vitória(3)	1	0,2	10	2,8	11
Vanusa(3)	5	1,4	13	3,6	18
Total					357

Em Campo Grande, catorze informantes concentram suas ocorrências em apenas um dos sequenciadores – *então*, o que permite afirmar que, embora haja muita variação idiossincrática, no final, a maioria usa mais este sequenciador. Somente quatro indivíduos, de faixa etária intermediária e mais velhos, – Renata, Zuleica, Mauro e Marta – empregam categoricamente *aí* e *daí*.

Tabela 6 – Distribuição dos sequenciadores por informante – São Paulo

Informante	<i>Então</i>	%	<i>Aí + Daí</i>	%	Total
AliceM(1)	13	3,7	1	0,2	14
CarlosJ(1)	6	1,7	1	0,2	7
CarolinaA(1)	12	3,4	3	0,8	15
DalsonO(1)	6	1,7	3	0,8	9
DouglasS(1)	5	1,4	3	0,8	8
EvandroO(1)	-----	-----	-----	-----	-----
JaimeN(1)	8	2,2	-----	-----	8
LaraN(1)	4	1,1	1	0,2	5
LucasS(1)	7	2	-----	-----	7
NicolauS(1)	-----	-----	-----	-----	-----
RaíssaO(1)	11	3,1	2	0,5	13
RodrigoR(1)	8	2,2	2	0,5	10
SérgioA(1)	13	3,7	1	0,2	14
SílviaB(1)	9	2,5	1	0,2	10
ValériaA(1)	-----	-----	-----	-----	-----
VivianeO(1)	6	1,7	-----	-----	6
AlbertoM(2)	4	1,1	1	0,2	5
ArianeG(2)	9	2,5	-----	-----	9
ElaineS(2)	5	1,4	-----	-----	5
FredericoV(2)	5	1,4	2	0,5	7
JoséN(2)	8	2,2	1	0,2	9
LucianaN(2)	11	3,1	5	1,4	16
LucianoT(2)	3	0,8	2	0,5	5
MarietaS(2)	1	0,2	1	0,2	2
MaurícioB(2)	4	1,1	1	0,2	9

MuriloM(2)	6	1,7	1	0,2	7
PedroS(2)	7	2	2	0,5	9
NelsonF(2)	7	2	3	0,8	10
PolianaM(2)	4	1,1	2	0,5	6
RenataC(2)	6	1,7	-----	-----	6
RuthC(2)	1	0,2	5	1,4	6
SandroS(2)	1	0,2	-----	-----	1
AnselmoM(3)	1	0,2	2	0,5	3
BenjaminT(3)	5	1,4	5	1,4	10
DarcyM(3)	6	1,7	-----	-----	6
EdnaC(3)	5	1,4	-----	-----	5
EthelM(3)	4	1,1	1	0,2	5
GilvanS(3)	4	1,1	1	0,2	5
HugoF(3)	2	0,5	-----	-----	2
IaraM(3)	6	1,7	-----	-----	6
JorgeV(3)	11	3,1	1	0,2	12
MarcosB(3)	1	0,2	-----	-----	1
PedroS(3)	-----	-----	-----	-----	-----
NiltonR(3)	2	0,5	-----	-----	2
RodolfoR(3)	4	1,1	3	0,8	7
RômuloS(3)	1	0,2	1	0,2	2
VeraD(3)	7	2	1	0,2	8
VivianT(3)	3	0,8	-----	-----	3
Total					349

Em São Paulo, parece haver mais variação entre os falantes. Há três indivíduos em cuja fala não se revela nenhum caso de sequenciação no domínio da causalidade referencial. Já para treze deles, *então* é a forma categoricamente empregada. Diferentemente do que se verifica em Campo Grande, contudo, nenhum paulistano utiliza categoricamente *aí* e *daí*.

Na análise global, os grupos de fatores que foram selecionados pela análise estatística como relevantes são: (a) *aspecto gramatical*; (b) *sequência discursiva* e (c) *escolaridade*, conforme indica a Tabela 7.

Tabela 7 - Grupos de fatores selecionados na análise global

Análise Global					
	Fatores	Total	% <i>então</i>	P.R.	Range
Aspecto gramatical	Imperfectivo	350	49	.70	.34
	Resultativo	146	20	.63	
	Indeterminado	128	19	.38	
	Perfectivo	82	12	.36	
Sequência discursiva	Argumentativa	298	42	.61	.28
	Expositiva	245	34	.60	
	Descritiva	103	15	.47	
	Narrativa	60	9	.33	
Escolaridade	Ensino Superior	400	56	.60	.19
	Ensino Médio	306	44	.41	
Total		706			

Input: 0,201; p < 0,05

Sobre o *aspecto gramatical*, a distribuição dos dados confirma as hipóteses aventadas para este grupo de fatores:

- (i) *aí* e *daí*, formas que têm, relativamente a *então*, mais ligação com a noção de espaço do que com a ideia de tempo, são utilizadas com mais frequência diante de verbos em *aspecto perfectivo* e *indeterminado*;
- (ii) *então* é favorecido diante de verbos com *aspecto imperfectivo* ou resultativo (caracterizados pela ideia de repetição de uma ação ao longo do tempo): uma possível explicação para tal fato pode ser atribuída à sua característica preponderantemente temporal quando em função dêitica. Desse modo, *então* seria o sequenciador preferido diante de aspectos que trazem, de forma mais saliente, a noção de tempo. Em outras palavras, há, aqui, outro exemplo de persistência (Hopper, 1991).

Quanto ao grupo *sequência discursiva*, também se confirma a hipótese inicial, aventada com base no princípio da marcação: como exposto no Capítulo 4, considerou-se um contínuo em que o ponto de partida é a narrativa, passa pela descrição, segue para a exposição e termina com a argumentação, de modo que, em relação ao grau de complexidade das informações, a narração está em um polo e a argumentação, em outro.

Conforme o previsto pelo princípio da *iconicidade* (Hopper, 1991), a narrativa, por ser cognitivamente menos complexa, tende a ser sequenciada por formas menos marcadas, como *aí* e *daí*. Por outro lado, a argumentação, sequência de maior esforço cognitivo, favorece *então*, forma de maior complexidade estrutural e, portanto, mais marcada, nos termos do Funcionalismo linguístico.

Pode-se cogitar o fato de que *então* esteja se estendendo, aos poucos, a tipos de discurso que, em princípio, não lhe eram muito característicos. De acordo com esse raciocínio, seria possível compreender o emprego de tal sequenciador em enunciados como a narrativa e a descrição, em que se espera que o falante opte por *aí* ou *daí*, formas menos marcadas.

Os resultados para a variável *escolaridade* também confirmam a hipótese inicial: falantes com ensino superior tendem a utilizar *então* com mais frequência do que aqueles que têm somente ensino médio. Diante de tal resultado, é possível entender que, embora variável em foco seja de natureza discursiva (que, em princípio, não teria correlação com fatores sociais), o uso variável de *aí*, *daí* e *então* se revela como um fenômeno em que emerge algum tipo de avaliação social.

A tabela abaixo expõe os grupos de fatores não selecionados estatisticamente, (a partir das análises *one-level*).

Tabela 8 – Grupos de fatores não selecionados na análise global

	Fatores	Total	% <i>então</i>	P.R.
Status da informação ^(a)	Inferível	473	67	[.56]
	Nova	194	27	[.47]
	Dada	39	6	[.45]
Duração ^(b)	Durativo	548	78	[.53]
	Pontual	158	22	[.48]
Dinamicidade ^(c)	Estativo	498	79	[.52]
	Dinâmico	208	30	[.47]
Gradiência do tópico discursivo ^(d)	Intratópico	440	62	[.52]
	Intertópico	266	38	[.48]
Telicidade ^(e)	Télico	556	79	[.50]
	Atélico	150	21	[.49]
Faixa etária ^(f)	20-34	268	38	[.55]
	50 ou mais	251	35	[.50]
	35-49	187	27	[.44]
Sexo/gênero ^(g)	Feminino	373	53	[.52]
	Masculino	333	47	[.47]
Total		706		

^(a)Input: 0,087. ^(b)Input: 0,082. ^(c)Input: 0,079. ^(d)Input: 0,076. ^(e)Input: 0,070. ^(f)Input: 0,068.

^(g)Input: 0,062.

Sobre o *status da informação*, a hipótese não é confirmada. Acredita-se que *então* venha sendo utilizado para sequenciar não só informações novas, que dispendem maior esforço cognitivo, mas também as inferíveis e dadas, cujo teor de informatividade é mais baixo. O que se pode cogitar é que, à semelhança de *aí* e *daí*, *então* esteja se tornando, progressivamente, uma forma menos marcada, talvez em face de outras, como *por isso*, ou mesmo pelo emprego de perífrases conjuncionais, como *aí então* e *daí então*.

A respeito dos grupos de fatores *duração*, *dinaminicidade* e *telicidade*, é importante reiterar que esses três elementos são, na verdade, um trinômio que compõe a proposta de *aspecto lexical*, formulada por Smith (1991), a partir dos

postulados de Vendler (1967). O *aspecto lexical* diferencia-se do *aspecto gramatical* por sua natureza semântica; diz respeito à classificação dos verbos em grupos de acordo com sua carga semântica, diferentemente de uma distribuição guiada por elementos de natureza morfológica, como o tempo, e que caracteriza o *aspecto gramatical*. Nesse caso, quem sabe se possa pensar que o uso de *aí*, *daí* e *então* como sequenciadores da referencialidade seja influenciado, no tocante à seleção dos verbos, mais pelo *aspecto gramatical* do que pelo *aspecto lexical*.

Por outro lado, a compreensão de tal cenário mostra-se um pouco complexa, dado que a *escolaridade* é um grupo de fatores significativo para a escolha de *aí*, *daí* e *então*, fato que revela a relevância de questões sociais no emprego de uma ou outra variante. A variável *sexo/gênero*, contudo, não foi selecionada. Uma possível explicação para tais fatos é de que haveria certo prestígio de uma das formas (na análise global), de modo que, em princípio, causa estranhamento o fato de que *sexo/gênero* não tenha correlação com a variável dependente. Mais adiante, tal discussão será desenvolvida, assim que forem apresentados os resultados obtidos cada uma das amostras.

Os resultados que se apresentam a seguir, em conjunto com estes, vão permitir uma conclusão mais geral a respeito dos padrões de uso dos sequenciadores em ambas as comunidades analisadas.

3. A variedade campo-grandense

Os grupos de fatores selecionados na análise da amostra campo-grandense coincidem exatamente com os da análise global:

Tabela 9 – Grupos de fatores selecionados na análise da amostra campo-grandense

Variedade campo-grandense					
	Fatores	Total	% <i>então</i>	P.R.	Range
Escolaridade (a)	Ensino Superior	218	61	.63	.27
	Ensino Médio	139	39	.36	.
Sequência discursiva (b)	Argumentativa	158	44	.66	.38
	Expositiva	113	31	.53	
	Descritiva	57	16	.52	
	Narrativa	29	9	.28	
Aspecto gramatical (c)	Imperfectivo	181	50	.74	.38
	Resultativo	78	22	.48	
	Indeterminado	63	18	.38	
	Perfectivo	35	10	.36	
Total		357			

Input: 0,044.; p < 0,05

Deve-se atentar para o fato de que, diferentemente do que se observa na análise global, em que a *escolaridade* é o terceiro grupo de fatores selecionado, aqui esta é a primeira variável selecionada pelo Rbrul. Este é mais um indício da relevância deste grupo de fatores nos usos desses sequenciadores - *então*, comparativamente a *aí* e *daí*, é uma forma que pode ser de fato entendida como mais “escolar”. É importante observar que se, por um lado, *escolaridade* é um grupo de fatores significativo, *sexo/gênero* não é relevante, como mostra a tabela 10. Nesse sentido, o cruzamento entre tais variáveis se faz interessante como ponto de partida para pesquisas posteriores.

Figura 11 - Cruzamento entre *escolaridade* e *sexo/gênero* na amostra campo-grandense

Para esta amostra, procedeu-se ao cruzamento entre tais variáveis (figura 11), que permite visualizar que as taxas de *então* por homens e mulheres são praticamente idênticas, independentemente do nível de escolaridade dos falantes, discussão que será retomada adiante.

De forma semelhante aos resultados obtidos na análise global, *sequência discursiva* e *aspecto gramatical* são relevantes, tendo sido selecionados na mesma ordem pela análise estatística.

Os resultados para os grupos de fatores não selecionados na análise da amostra campo-grandense (obtidos de análises *one-level*) estão na Tabela 10:

Tabela 10 – Grupos de fatores não selecionados na análise da amostra campo-grandense

	Fatores	Total	% <i>então</i>	P.R.
Status da informação ^(a)	Inferível	219	61	[.57]
	Dada	117	33	[.51]
	Nova	21	6	[.38]
Duração ^(b)	Durativo	282	79	[.52]
	Pontual	75	21	[.47]
Telicidade ^(c)	Télico	283	80	[.52]
	Atélico	74	20	[.47]
Dinamicidade ^(d)	Estativo	262	73	[.51]
	Dinâmico	95	26	[.48]
Gradiência do tópico discursivo ^(e)	Intratópico	240	67	[.50]
	Intertópico	117	32	[.49]
Sexo/gênero ^(f)	Feminino	189	53	[.53]
	Masculino	168	47	[.47]
Faixa etária ^(g)	35-49	158	44	[.51]
	20-34	102	29	[.50]
	50 ou mais	97	27	[.48]
Total		357		

^(a)Input: 0,085; ^(b)Input: 0,08; ^(c)Input: 0,080; ^(d)Input: 0,076; ^(e)Input: 0,074; ^(f)Input: 0,070;

^(g)Input: 0,067

Como apresentado na tabela, a variável *faixa etária* não foi selecionada para a amostra de Campo Grande. Há, contudo, um indício de que, no “mercado linguístico” (Bourdieu, 1991), *então* seja uma forma mais bem avaliada socialmente (apesar do fato de que *sexo/gênero* tampouco tenha sido selecionada aqui).

Figura 12 - Frequência de *então* por faixa etária em Campo Grande

Conforme tem sido apresentado, *faixa etária* e *sexo/gênero* não são variáveis relevantes – nem para a amostra global, nem para a campo-grandense. Aqui, o cruzamento entre tais grupos de fatores evidencia que, independentemente da faixa etária, homens e mulheres campo-grandenses não se diferenciam muito quanto à frequência de *então* (as taxas para mulheres são sempre um pouco maiores, mas essa diferença não é significativa).

Figura 13 - Taxas de *então* no cruzamento entre *faixa etária* e *sexo/gênero* na amostra campo-grandense

A fim de comparar padrões, seguem as análises realizadas com dados da amostra paulistana. A seguir, os resultados são apresentados de acordo com a ordem de seleção do programa.

4. A variedade paulistana

Na análise dos dados da amostra paulistana, também são três os grupos de fatores selecionados. Entretanto, são dois de ordem social e um de natureza linguística. Na tabela abaixo, pode-se verificar como as ocorrências das formas em foco se distribuem, na amostra paulistana.

Tabela 11 – Grupos de fatores selecionados na análise da amostra paulistana

Variedade paulistana					
	Fatores	Total	% <i>então</i>	P.R.	Range
Faixa etária (a)	60 ou mais	110	31	.55	.24
	35-59	89	26	.36	
	20-34	150	43	.60	
Escolaridade (b)	Ensino Superior	182	52	.59	.16
	Ensino Médio	167	47	.43	
Aspecto gramatical (c)	Imperfectivo	169	48	.64	.31
	Resultativo	68	20	.62	
	Indeterminado	65	18	.43	
	Perfectivo	47	14	.32	
Total		706			

Input: 0,201; p < 0,05.

À semelhança dos resultados obtidos na análise global, a hipótese acerca do *aspecto gramatical* se confirma aqui também, com os dados da amostra paulistana analisados separadamente: os aspectos indeterminado e perfectivo desfavorecem *então* (portanto, favorecem *aí* e *daí*). Em contrapartida, os aspectos imperfectivo e resultativo, caracterizados pela ideia de repetição de uma ação ao longo do tempo, favorecem o uso de *então*, forma caracterizada pelo uso

preponderantemente temporal em sua forma adverbial, a despeito de *aí* e *daí*, vocábulos com uso dêitico espacial. Em outras palavras aspectos com carga temporal mais saliente favorecem o sequenciador com a mesma característica.

Para *faixa etária*, o peso relativo para os falantes da faixa etária intermediária encontra-se abaixo do ponto neutro – de modo que desfavorecem *então*, ao passo que pessoas jovens e mais velhas favorecem tal forma. Em outras palavras, pessoas de faixa etária intermediária assemelham-se mais com as de terceira faixa etária do que com as mais jovens em relação ao uso dos sequenciadores. Desse modo, o que se tem é variação estável, com pessoas de faixa etária intermediária desfavorecendo *então*, um dado surpreendente, se levada em conta a questão do mercado linguístico (Bourdieu, 1991). Uma possível interpretação para tal fato é que, entre paulistanos e campo-grandenses, poderia haver diferentes concepções de prestígio associadas a *aí*, *daí* e *então*, de modo que, em São Paulo, paulatinamente, *aí* e *daí* estivessem sendo mais valorizadas socialmente.

É possível, portanto, comparar esse resultado com aquele obtido para a amostra campo-grandense. Na capital de Mato Grosso do Sul, *então* é mais frequente entre os informantes da segunda faixa etária (ainda que não significativamente, do ponto de vista estatístico). Assim, evidentemente não se pode falar em qualquer sinal de mudança na direção de qualquer forma, em nenhuma das comunidades.

Figura 14 – Frequência de *então* por faixa etária em São Paulo

O que se verifica para o grupo de fatores *escolaridade* é bem semelhante aos resultados reportados para a análise global: *então* é o sequenciador favorecido pelos falantes com formação universitária – o que se revela como uma semelhança entre as falas campo-grandense e paulistana.

O gráfico a seguir mostra a distribuição de *então* na amostra paulistana, a partir do cruzamento entre *faixa etária* e *escolaridade*. Os resultados reiteram aqueles já expostos na Tabela 11:

Figura 15 - Taxas de *então* no cruzamento entre *faixa etária* e *escolaridade* na amostra paulistana

A Figura 15 reitera que, embora *então* seja o sequenciador preferido, falantes de faixa etária intermediária utilizam-no com menos frequência do que que jovens e pessoas mais velhas.

Com base nos resultados obtidos a partir de rodadas one-level, que mostram os grupos de fatores não selecionados nas análises da amostra paulistana, o que se observa é que o conjunto de variáveis que se correlacionam com o emprego de *aí*, *daí* e *então* não se assemelha àquele verificado para Campo Grande. Para São Paulo, destaca-se o fato de que *faixa etária* é uma variável significativa (algo que não se observa para a amostra campo-grandense). Por outro

lado, para aquela amostra, *sequência discursiva* é uma variável estatisticamente significativa – o que não se observa aqui, para os dados paulistanos.

Tabela 12 – Grupos de fatores não selecionados na análise da amostra paulistana

	Fatores	Total	% <i>então</i>	P.R.
Sequência discursiva ^(a)	Argumentativa	140	40	[.52]
	Expositiva	132	37	[.51]
	Descritiva	46	14	[.45]
	Narrativa	31	9	[.37]
Status da informação ^(b)	Inferível	254	72	[.53]
	Dada	77	22	[.53]
	Nova	18	6	[.42]
Gradiência do tópico discursivo ^(c)	Intratópico	200	57	[.52]
	Intertópico	149	43	[.45]
Duração ^(d)	Durativo	266	76	[.52]
	Pontual	83	24	[.48]
Dinamicidade ^(e)	Estativo	236	67	[.51]
	Dinâmico	113	33	[.44]
Telicidade ^(f)	Télico	273	78	[.51]
	Atélico	76	22	[.48]
Sexo/gênero ^(g)	Feminino	184	52	[.55]
	Masculino	165	48	[.45]
Total		349		

^(a)Input: 0,080; ^(b)Input: 0,079; ^(c)Input: 0,07; ^(d)Input: 0,074; ^(e)Input: 0,071; ^(f)Input: 0,072;

^(g)Input: 0,061.

Ainda sobre os grupos de fatores não selecionados, a frequência de uso da forma *então* entre homens e mulheres é praticamente igual para todas as faixas etárias entre paulistanos, conforme ilustra a Figura 16:

Figura 16 – Taxas de *então* no cruzamento entre *faixa etária* e *sexo/gênero* na amostra paulistana

Desse modo, embora *escolaridade* seja uma variável relevante para o emprego dos sequenciadores nas duas amostras, a noção de prestígio não parece operar, em nenhuma das duas, do modo como se vem atestando numa multiplicidade de trabalhos – nos quais se verifica que a forma prestigiosa tende a ser favorecida pelas mulheres. Quando se cruzam *sexo/gênero* e *escolaridade*, identifica-se o mesmo padrão tanto para homens quanto para mulheres: *então* é a forma mais frequente entre os mais escolarizados.

Figura 17 - Cruzamento entre *escolaridade* e *sexo/gênero* na amostra paulistana

No seu conjunto, as análises reportadas neste capítulo permitem concluir que: (i) paulistanos e campo-grandenses se assemelham no que diz respeito a taxas gerais de ocorrências das variantes: *então* é, no geral, a forma favorecida em ambas as comunidades (ou seja, a variável *cidade* não é selecionada como relevante na análise global); (ii) por outro lado, os grupos de fatores estatisticamente significativos são diferentes para cada amostra, quando analisadas separadamente (o que aponta para um encaixamento social e linguístico diferente entre elas) (iii) a seleção da variável *escolaridade* para ambas as amostras, com *então* como forma favorecida pelos mais escolarizados, indica que há algum prestígio associado a tal forma, mas não a ponto de levar a uma diferenciação significativa entre homens e mulheres.

Conclusão

Esta pesquisa centrou-se sobre o uso intercambiável de *aí*, *daí* e *então* – como sequenciadores da causalidade no domínio referencial (Sweetser, 1991) – em duas comunidades brasileiras, Campo Grande e São Paulo, a partir de amostras robustas da fala de cada capital. Além de contribuir para o entendimento sociolinguístico dessas comunidades de fala, bem como das semelhanças e diferenças entre elas, este estudo comparativo objetivou investigar os padrões de emprego das variantes de uma variável discursiva.

Com base na teoria laboviana e nos pressupostos da gramaticalização, discutiu-se o *continuum* no uso das formas em foco – de dêiticos a sequenciadores discursivos. A análise se define como sociofuncionalista, já que tem como objetivo a integração de pressupostos teórico-metodológicos da sociolinguística variacionista e do funcionalismo.

Conforme visto no início do Capítulo 5, a variável *cidade* não foi selecionada na análise global: *então* é a forma preferida tanto em Campo Grande quanto em São Paulo. O fato de que as taxas dessa forma não são significativamente diferentes entre as duas cidades não permite seguir com a hipótese de que, devido a um possível “orgulho caipira”, campo-grandenses pudessem empregar menos frequentemente tal forma e preferir as formas *aí* e *daí* (hipoteticamente mais coloquiais, com base, inclusive, no fato de que não são preconizadas pela escola como sequenciadores no domínio causal).

Por outro lado, o fato de que Campo Grande e São Paulo se assemelham no que diz respeito ao favorecimento de *então* poderia ser tomado como algo não surpreendente, já que, em geral, variáveis além do nível fonético-fonológico, sobretudo as discursivas, são tidas como fenômenos que não se correlacionam fortemente com fatos de natureza social (Macaulay, 2005).

Entretanto, pode-se verificar que não apenas que diferentes variáveis se correlacionam aos empregos dessas formas nas duas cidades, mas também que fatores têm papel importante nos dois casos:

a) *faixa etária* não é um grupo de fatores selecionado para os dados de Campo Grande, mas é para os de São Paulo (onde *aí* e *daí* são as formas favorecidas entre os informantes da faixa etária intermediária);

b) o fato de que *escolaridade* tenha sido um grupo selecionado na análise de ambas as amostras sugere que a variante em foco nas análises estatísticas – *então* – pode, de fato, ser entendida como aquela que tem um estatuto mais “escolar”, pois tende a ser a forma mais empregada por informantes que têm nível superior;

c) o fato de que não se verifica correlação entre *sexo/gênero* e a variável dependente é interessante na medida em que se observa a correlação com escolaridade. Se há prestígio associado a *então*, por que tal forma não seria preferida pelas mulheres, relativamente aos homens? Seria a relação entre prestígio e “norma escolar” uma relação necessária, mesmo quando se trata de uma variável no nível discursivo? Os resultados aqui apresentados sugerem que não. Seria então necessário rediscutir a questão do prestígio para variáveis discursivas, através da comparação entre os presentes resultados e aqueles de estudos futuros nessa mesma linha.

Além disso, é preciso reiterar que há diferenciação significativa entre as *faixas etárias* dos paulistanos – um quadro que não se verifica em Campo Grande. Por um lado, uma possível explicação para a relevância desse grupo de fatores entre os paulistanos e não entre os campo-grandenses pode estar relacionada à estratificação da amostra. Conforme explica o capítulo 3, em São Paulo, considera-se como segunda e terceira faixas etárias, respectivamente, falantes de 35-59 anos e 60 anos ou mais. Em Campo Grande, essas faixas são menores (35 a 49 anos e 50 ou mais), devido à dificuldade de se encontrarem informantes nativos com idade superior a 40 anos e escolaridade média ou superior. A adoção de tal critério pode ter influenciado os resultados obtidos para o grupo de fatores em questão. Por outro lado, o fato de que os falantes da faixa intermediária desfavorecem *então* (e, portanto, favorecem *aí* e *daí*) pode ser um indicativo de

que o valor social dessas duas formas venha apresentando diferenças entre campo-grandenses e paulistanos, de modo que, somente entre os primeiros, parece haver uma progressiva valorização das formas *aí* e *daí*.

No que diz respeito às variáveis linguísticas, a análise dos dados também revela diferenças e semelhanças em relação às capitais. Em Campo Grande, tem-se *sequência discursiva* e *aspecto gramatical* como grupos de fatores significativos para o uso de *então*, enquanto que, em São Paulo, a única variável interna selecionada é *aspecto gramatical*.

Deve-se considerar que o fato de a amostra campo-grandense ter sido toda coletada por uma mesma documentadora pode ter conferido às entrevistas um caráter discursivamente mais homogêneo. Desse modo, a relação entrevistadora-entrevistado seria um modo de explicar uma consistente coloquialidade por parte dos entrevistados campo-grandenses nos trechos narrativos (em que as formas *aí* e *daí* são favorecidas).

Outro ponto a ser destacado é a seleção do grupo de fatores *aspecto gramatical*, em detrimento do *aspecto lexical*. Cogita-se que tal resultado seja fruto da natureza do fenômeno estudado - a função de sequenciação discursiva resulta de um ganho de propriedade gramatical de *aí*, *daí* e *então* e da respectiva perda de propriedades lexicais dessas formas enquanto itens dêiticos. Em outras palavras, o *aspecto lexical* não parece influenciar a produção de uma variável de natureza discursiva, diferentemente do *aspecto gramatical* (ou seja, não têm peso as propriedades semânticas do verbo, diferentemente de suas características estruturais e da relação que elas mantêm com os sequenciadores).

A variação entre os sequenciadores analisados encontra respaldo em alguns dos princípios de gramaticalização (Hopper, 1991) expostos no Capítulo 2:

- a) *Descategorização*: perda categorial por parte dos itens gramaticalizados. Quando gramaticalizados, *aí*, *daí* e *então* passam por um enfraquecimento de suas propriedades dêiticas, em função do ganho de propriedades funcionais, próprias dos sequenciadores discursivos.
- b) *Estratificação*: uma vez gramaticalizados os itens, novas funções passam a coexistir com funções antigas no sistema linguístico. Esse princípio ilustra

não só o uso de *aí* e *daí* como sequenciadores causais no domínio referencial, ao lado de formas canônicas, mas também o uso maciço de *então* no domínio referencial.

- c) *Persistência*: Resquícios do uso dêitico de *então* mantêm-se ao longo do processo de gramaticalização e contribuem para que tal forma seja mais comum diante de verbos com aspecto resultativo e imperfectivo, ou seja, aqueles cujos traços de tempo são mais evidentes, em contraposição ao aspecto indeterminado e perfectivo.

Por fim, deve-se atentar para a relevância do mecanismo da *iconicidade* e da *marcação* (Hopper, 1991). As hipóteses para o grupo de fatores *sequência discursiva* foram confirmadas. Com base no princípio da iconicidade, a narrativa, entendida como menos complexa cognitivamente (citação), tende a ser sequenciada por formas menos marcadas, como *aí* e *daí*. A argumentação, por sua vez, sequência que exige maior dispêndio cognitivo, favorece o uso de *então*, forma mais marcada.

O presente trabalho visou à descrição de fenômenos discursivos no português brasileiro. Para isso, foram analisadas duas variedades de fala ainda pouco exploradas no cenário nacional, com interesse nas suas semelhanças e divergências no tocante ao uso variável de três itens gramaticalizados como sequenciadores.

Os resultados obtidos neste estudo permitirão observar se os padrões de uso descritos para a fala campo-grandense e paulistana aplicam-se também a outras comunidades do Brasil. Pesquisas futuras poderão permitir avançar nos seguintes pontos:

- (i) A discussão em torno do “valor escolar” de *então*, que não reflete numa diferenciação de sexo/gênero. Seriam tais resultados observados também em outras comunidades? Seriam tais resultados observados para outras variáveis de natureza discursiva?
- (ii) Quais seriam os padrões de uso de tais formas entre informantes de escolaridade rudimentar?

- (iii) Como poderia ser desenvolvida uma pesquisa sobre a percepção sociolinguística acerca dessas formas, no sentido de cotejar os resultados com os que se apresentam aqui, sobre sua produção?

As análises aqui desenvolvidas também promovem um questionamento de natureza linguística: considerando-se que *aí* e *daí* são formas muito mais frequentes no domínio referencial, haveria outros possíveis sequenciadores constituindo, com *então*, um envelope de variação nos domínios epistêmico e de atos de fala?

Esta tese contribui, enfim, para o refinamento da descrição dos contextos em que os sequenciadores analisados constituem um envelope de variação: primeiramente, no que diz respeito ao amplo domínio da sequenciação, e, em seguida, no que se refere à causalidade. *Aí*, *daí* e *então* já não passam despercebidos como variantes de uma variável: seu ponto de encontro é o domínio referencial.

Referências

- ADAM, Jean-Michel. *Les textes: Types et prototypes - Récit, description, argumentation et dialogue*. Paris: Nathan, 1992.
- ALMEIDA, Daniel M. V. de; MARINHO, Janice H. C. *Dos Marcadores Discursivos e Conectores: Conceituação e Teorias Subjacentes*. Gláuks, v. 12. n. 1, 2012, 169 - 203.
http://www.revistaglauks.ufv.br/arearestrita/arquivos_internos/artigos/Artigo_7_Daniel_e_Janice_diagramado.pdf
- AMARAL, Amadeu. *O dialeto caipira*. São Paulo: Anhembi, 1955.
- AMORIM, F. S. *Alguns indícios sincrônicos da gramaticalização do conector por causa (de) que*. Domínios de Linguagem, v. 5, n. 1. 2011.
- BAILLY, A. *Dictionnaire grec français*. Paris: Hachette, 1950.
- BASSO, Renato; Ilari, Rodolfo. *Estativos e suas características*. Rev. Brasileira de Linguística Aplicada, v. 4, n. 1, 2004. Disponível em:
<http://www.scielo.br/pdf/rbla/v4n1/02.pdf>
- BATES, D.; D. SARKAR. *lme4: linear mixed-effects models using Eigen and R*, 2008.
<http://cran.r-project.org>. R package, version 0.999375-8
- BECHARA, Evanildo. *Moderna gramática portuguesa*. Rio de Janeiro: Nova Fronteira, 2009.
- BRAGA, M. L. et al. *Aí e Então e a hipótese da trajetória universal*. In: NEVES, M. H. M. (org.). *Descrição do português*. Araraquara: Acadêmica, 2001, p. 13-23.
- BOLINGER, D. *Meaning and Form*. London: Longman, 1977.
- BORGES NETO, J.; MÜLLER, A. L. P. *Linguistas ou camaleões? Uma resposta a Tarallo*. D.E.L.T.A., vol. 3, n. 1, 1987.
- BOURDIEU, P. *Language and Symbolic Power*. Cambridge: Polity Press, 1991.
- BYBEE, J. *Language, usage and cognition*. Cambridge: Cambridge University Press, 2010.
- _____. *Usage-based models in linguistics*. Entrevista de Tiago Timponi Torrent. Revista Linguística, v. 8, n. 1, p. 1-6. 2012.
- CALVET, Louis-Jean. *Sociolinguística, uma introdução crítica*. Trad. Marcos Marcionilo. São Paulo: Parábola, 2002.

CAMBRAIA, César Nardelli. *Sobre as origens do “aonde” na língua portuguesa*. In: COHEN, M.A.; RAMOS, Jânio (org.). *Dialeto mineiro e outras falas; estudo de variação e mudança linguística*. Belo Horizonte: UFMG, 2002.

CASTILHO, A. T. *Introdução ao estudo do aspecto verbal na língua portuguesa*. Tese (Doutorado em Linguística), Universidade de São Paulo, Marília, 1968.
<http://seer.fclar.unesp.br/alfa/article/viewFile/3311/3038>

_____; *A gramaticalização*. Estudos linguísticos e literários, 19: 25-64. Salvador: Programa de Pós-graduação em Letras e Linguística da UFBA, 1997.

_____; PRETI, D. *A linguagem falada culta na cidade de São Paulo: materiais para seu estudo*. vol. I – Elocuções Formais. São Paulo: T.A. Queiroz, 1986.

_____; *Aspecto verbal no português falado*. In: ABAURRE, M. B.; RODRIGUES, A. (Ed.). *Gramática do português falado – novos rumos*. Campinas: Editora da Unicamp, 2003. v. VIII. p. 83-121.

CEDERGREN, Henrietta J. & SANKOFF, David. *Variable rules: Performance as a statistical reflection of competence*. *Language*, vol. 50(2), 333–355, 1974.
<http://www.ling.upenn.edu/courses/cogs501/CedergrenSankoff1974.pdf>

CHAFE, W. *Significado e estrutura linguística*. Rio de Janeiro: Livros Técnicos e Científicos. 1979.

_____; DANIELEWICZ, Jane. *Properties of speaking and written language*. In: HOROWITZ, Rosalind; SAMUELS, S. Jay (eds.). *Comprehending oral and written language*. New York: Academic Press, 1987, p. 83-113.

_____. *Linking intonation in spoken English*. In: HAIMAN, J.; THOMPSON, S.A. (eds.) *Clause combining in grammar and discourse*. Amsterdam: John Benjamins Publishing, 1988, p. 1-27

CHAMBERS, J. K.; TRUDGILL, Peter. *Dialectology*. Cambridge: Cambridge University Press, 1980.

_____. *Sociolinguistic Theory. Linguistic variation and its social significance*. Oxford: Blackwell, 1995.

CHESHIRE, J. *Syntactic variation and beyond: gender and social class variation in the use of discourse-new markers*. *Journal of Sociolinguistics*, Hoboken, v.9, n.4, p.479-508, 2005.

COLOMER, T. & CAMPS, A. *Ensinar a ler, ensinar a compreender*. Porto Alegre: Artmed, 2002.

COMRIE, B. *Aspect*. Cambridge: Cambridge University Press, 1976.

COSTA, S. B. B. *Adverbiais espaciais e temporais do Português: indícios diacrônicos de gramaticalização*. Tese de Doutorado. Universidade Federal da Bahia - Programa de Pós-Graduação em Letras e Linguística. Salvador: UFBA, 2003.

CUNHA, Celso & CINTRA, Luís L. *Nova Gramática do Português Contemporâneo*. Rio de Janeiro: Nova Fronteira, 2001.

DAL MAGO, D. *Quer dizer: percurso de mudança via gramaticalização e discursivização*. Dissertação (Mestrado em Linguística). Florianópolis, UFSC, 2001.

ERNOUT, A. e MEILLET, A. *Dictionnaire étymologique de la langue latine: histoire des mots*. Paris: Librairie C. Klincksieck, 1959.

FARACO, Carlos Alberto. *Estudos pré-saussurianos*. In: MUSSALIM, Fernanda; BENTES, Anna Christina (Org.). *Introdução à Linguística*. Vol. 3: Fundamentos epistemológicos. São Paulo: Cortez, 2004.

FARIA, E. *Dicionário escolar latino-português*. Rio de Janeiro. FENAME, 1975.

FIGUEREDO GOMES; PENA FERREIRA (orgs.) *Corpus mínimo de textos da língua portuguesa – COMTELPO*. Lisboa: Faculdade de Letras da Universidade de Lisboa/PDEECAPES, 2006.

FREITAG, Raquel Meister Ko. “*Problemas teórico-metodológicos para o estudo da variação linguística nos níveis gramaticais mais altos*”. *Matraga*, v. 16, p. 115-132, 2009.

www.pglettras.uerj.br/matraga/matraga24/arqs/matraga24a05.pdf

_____; et al. *O controle do gênero textual/sequências discursivas na motivação da variação linguística: apontamentos metodológicos*. Em *Odisseia*, 3, 1-15, 2009.

www.periodicos.ufrn.br/odisseia/article/download/2051/1485

_____. *O social da sociolinguística: o controle de fatores sociais*. Diadorim, Rio de Janeiro, v.8, 2011, p.43-58.

<file:///C:/Users/Mar%C3%ADlia/Local%20Settings/Downloads/Diadorim8%20O%20social%20da%20sociolingu%C3%ADstica.pdf>

GONÇALVES, S.C.L. et al. *Introdução à gramaticalização: princípios teóricos e aplicação*. São Paulo: Parábola Editorial, 2007.

_____. “*Banco de dados Iboruna: amostras eletrônicas do português falado no interior paulista*”, s/d. Último acesso em 25 set./2014. www.alip.ibilce.unesp.br/iboruna

GIL, Antônio Carlos. *Como elaborar projetos de pesquisa*. 4. Ed. São Paulo: Atlas, 2002.

GIVÓN, T. *Topic, Pronoun and Grammatical Agreement*. In: Charles N. Li (ed.) *Subject and Topic*, p. 149-88. New York: Academic Press, 1976.

_____. *A functional-typological introduction*. Vol.1. Amsterdam/Philadelphia: John Benjamins Publishing Co, 1984.

_____. *Verbal Inflections: Tense, Aspect, Modality and Negation*. In: *English Grammar: a functional-based introduction*. Vol I. Amsterdam/Philadelphia: John Benjamins Publishing Co, 1993.

_____. *Functionalism and Grammar*. Amsterdam/Philadelphia: John Benjamins, 1995.

_____. *Syntax*. v. 1 e 2. Amsterdam: John Benjamins, 2001.

GUY, Gregory, HORVATH, Barbara, VONWILLER, Julia, DAISLEY, Elaine, ROGERS, Inge. *An intonational change in progress in Australian English*. In: *Language*, n. 15, p. 23-52, 1986.

_____. *Varbrul: análise avançada*. In: Guy, Gregory R. & Zilles, Ana M. (Eds.), *Sociolinguística quantitativa: instrumental de análise*, p. 47–70. São Paulo: Parábola, 2007 [1988].

_____. *Explanation in variable phonology: An exponential model of morphological constraints*. *Language Variation and Change*, vol. 3, 1–22, 1991.

HALLIDAY, M. A. K. *An introduction to functional grammar*. Great Britain: Edward Arnold, 1985.

_____; H. HASAN. *Cohesion in English*. London: Longman, 1976.

HAIMAN, John & THOMPSON (1988), *Introduction*. In: J. Haiman & Thompson (ed). *Clause Combining in Grammar and Discourse*, Philadelphia, John Benjamins, p. X-XIII.

HEINE, Bernd. ULRIKE, Claudi & HUNNEMEYER. *Grammaticalization: a conceptual framework*. Chicago: University of Chicago Press, 1991.

HELLWIG, B.; GEERTS, J. ELAN – *Linguistic Annotator*. Versão 4.4.0.
<http://www.mpi.nl/corpus/manuals/manual-elan.pdf>.

HINTZE, Ana Cristina Jaeger. *Gramaticalização e as variâncias das locuções de base nominal*. In: *Anais do 5º Encontro do Celsul*, Curitiba-PR, 2003, p. 186-192.

HOPPER, P. J. *On Some principles of grammaticization*. In: TRAUGOTT, E.C.,

HEINE, N. (ed.). *Approaches to grammaticalization*. Amsterdam: John Benjamin, 1991, p. 17-35.

_____; TRAUGOTT, Elizabeth. *Grammaticalization*. Cambridge: Cambridge University Press, 1993.

ILARI, Rodolfo. *O Estruturalismo Linguístico: alguns caminhos*. In: MUSSALIM, Fernanda; BENTES, Anna Christina (Org.). *Introdução à Linguística*. Vol. 3: Fundamentos epistemológicos. São Paulo: Cortez, 2004.

JOHNSON, Daniel Ezra. "Getting off the GoldVarb standard: introducing Rbrul for mixed-effects variable rule analysis." *Language and Linguistics Compass*, vol. 3(1), 359–383, 2009.

KOCH, Ingedore G. V. *Dificuldade na leitura/produção de textos: os conectores interfrásticos*. In: Marta Kist & Elvo Clemente (org.). *Linguística Aplicada ao Ensino de Português*. Porto Alegre: Mercado Aberto. 1987.

_____ et al. *Aspectos do processamento do fluxo de informação no discurso oral dialogado*. Em: Ataliba Teixeira de Castilho (Org.). *Gramática do português falado*. Campinas: Editora da UNICAMP; São Paulo: FAPESP, vol. I – A ordem, 1990, p. 143-84.

LEAL, António; OLIVEIRA, Fátima. *Subtipos de verbo de movimento e classes aspectuais*. Disponível em: http://www.apl.org.pt/docs/actas-23-encontro-apl-2007_2.pdf

LABOV, William. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press, 1972.

_____. *Where does the linguistic variable stop? A response to Beatriz Lavandera*. Working Papers in Sociolinguistics 44, 1978.

_____. *Field methods of the project on linguistic change and variation*. In John Baugh and Joel Sherzer (eds.), *Language in Use: Readings in Sociolinguistics*. 28- 66. Englewood Cliffs, NJ: Prentice Hall, 1984.

_____; WALETZKY, Joshua. *Narrative analysis. Essays on the Verbal and Visual Arts*, ed. J. Helm, 12 - 44. Seattle: U. of Washington Press. Reprinted in *Journal of Narrative and Life History* 7: 3 - 38, 1997.

_____. *Principles of Linguistic Change: Social Factors*. Oxford: Blackwell, 2001.

_____. *Ordinary events*. In: THORNBORROW, J.; COATES, J. (Eds.) *The sociolinguistics of narrative*. Amsterdam: John Benjamins, 2004. p. 31-43.

_____. *The social stratification of english in New York city*. 2.ed. New York: Cambridge University Press, 2006.

LAVANDERA, Beatriz. *Where does the sociolinguistic variable stop?* *Language in Society* 7, 1978; p.171-82.

LEHMANN, C. Towards a Typology of Clause Linkage . In: THOMPSON, S. & HAIMAN, J (orgs.). *Clause combining in grammar and discourse* . Amsterdam: John Benjamins, 1988, p. 151-225

LEITE, J.F. Marques e JORDÃO, A.J. Novaes. *Dicionário latino vernáculo*. Editora Lux Ltda. 1958.

LEVIN, Beth. *English Verb Classes and Alternations: a Preliminary Investigation*. Chicago: The University of Chicago Press, 1933.

LEVINSON, S. *Pragmatics*, Cambridge, Cambridge Univ.Press, 1983.

LOPES, M. H. C. *Aspectos sintáticos, semânticos e pragmáticos das construções causais: contributo para uma reflexão sobre o ensino da gramática*. 2004. Tese (Doutorado) – Faculdade de Letras da Universidade do Porto, Portugal, 2004.

LOPES, Célia R. dos S. *A persistência e a decategorização nos processos de gramaticalização*. In.: VITRAL, L.; COELHO, S. M. (orgs.) *Estudos de processos de gramaticalização em português: metodologias e aplicações*. Campinas, SP: Mercado de Letras, 2010, p. 275-314.

LOUREIRO, V. J. S. *Los marcadores del discurso dentro de la perspectiva del análisis gramatical*. In: BARROS, L. G.; COSTA, M^a J. D.; VIEIRA, V. R. A. (Orgs.). Congresso Brasileiro de Hispanistas: Hispanismo 2004: língua espanhola. Florianópolis: ABH, 2004, p. 471-480.

LYONS, John. *Linguagem e Linguística: uma introdução*. Rio de Janeiro: Zahar, 1970.

_____. *Semantics*. Cambridge: Cambridge University Press, 1977.

MACEDO, Alzira Tavares; SILVA, Gisele Machline de Oliveira e. *Análise sociolinguística de alguns marcadores conversacionais*. In: MACEDO, Alzira Tavares de; RONCARATI, Cláudia Nívea; MOLLICA, Maria Cecília. *Variação e discurso*. Rio de Janeiro: Tempo Brasileiro, 1996.

MACAULAY, R. *Discourse variation*. In: CHAMBERS, J. K., TRUDGILL, P. & SCHILLING-ESTES, N. (eds.). *The handbook of language, variation and change*. Oxford: Blackwell, 2005, p. 283-305.

MANN, W. C.; THOMPSON, S. A. *Assertions from discourse structure*. California: University of Southern California, 1985, 14 p.

MARCUSCHI, Luiz Antônio. *Análise da conversação*. São Paulo: Ática, 1991.

_____. *Gêneros Textuais & Ensino*. 5^o ed. São Paulo: Lucerna, 2002.

_____. *Produção de texto, análise de gêneros e compreensão*. São Paulo: Parábola Editorial, 2008.

MARGARIT, A. M. B. *Los complementos pronominalo-adverbiales derivados de ibi e inde en la Península Ibérica*. Revista de Filología Española, Anejo XXXVIII, Madrid, 1947.

MARTELOTTA, M.E.; VOTRE, S.; CEZARIO, M. M. (orgs.). *Gramaticalização no português: uma abordagem funcional*. Rio de Janeiro: Tempo Brasileiro: UFRJ. Departamento de Linguística e Filologia, 1996.

_____. (org.) *Linguística funcional teoria e prática*. Rio de Janeiro, DP&A, 2003.

_____; AREAS, E. K. *A visão funcionalista da linguagem no século XX*. In: FURTADO DA CUNHA, M. A.; OLIVEIRA, M.R.; MARTELOTTA, M. E (Orgs.) *Linguística Funcional: Teoria e Prática*. Rio de Janeiro: DP & A, 2003.

MATHESIUS, V. *Jazyk, Kultura, a slovesnost [língua, cultura e literatura]*. Praga: Odeon, 1982.

MEILLET, Antoine. *Linguistique historique et linguistique générale*. Paris: Edouard Champion, 1948 [1912].

MENDES, Ronald Beline. *Estar + gerúndio e ter + participio: aspecto verbal e variação no português*. Tese (Doutorado em Linguística) – Programa de Pós-graduação em estudos da linguagem da Universidade Estadual de Campinas, 2005. <http://www.bibliotecadigital.unicamp.br/document/?code=vtls000347284>

_____; OUSHIRO, Livia. *Documentação do Projeto SP2010 – Construção de uma amostra da fala paulistana*, 2013. Último acesso em 25 dez./2015. <http://projetosp2010.fflch.usp.br/producao-bibliografica>

_____; *O paulistano no mapa sociolinguístico brasileiro*. In: Alfa, São Paulo, 56 (3): 973-1001, 2012.

MEYERHOFF, Miriam. SCLEEF, Erick. MACKENZIE, Laurel. *Doing Sociolinguistics*. New York: Routledge, 2015.

MILROY, L.; GORDON, M. *Sociolinguistics – Method and Interpretation*. Oxford: Blackwell, 2003.

MOREIRA DE SÁ, Maria da Piedade. *A descrição na narrativa oral*. Revista do GELNE, n. 1, 1999, 61-65.

MOSS, B. *Teaching expository structures through information trade book retellings*. *The reading teacher*, 57 (8), 2004, 710 – 718.

NARO, Anthony. J. *Modelos quantitativos e tratamento estatístico*. In: MOLLICA, M. (Org.). *Cadernos didáticos UFRJ: Introdução à sociolinguística variacionista*. Rio de Janeiro: UFRJ, 1992, p. 17-25.

_____; & BRAGA, M. L. *A interface sociolinguística/gramaticalização*. Gragoatá, Niterói, n.9, p.125-134, 2001.

_____; SCHERRE, M. M. P. *Variação linguística, expressividade e tradição gramatical*. In: GORSKI, E. e COELHO, I. L. (orgs.). *Sociolinguística e ensino: contribuições para a formação do professor de língua*. Florianópolis: EdUFSC, 2006. p. 235-266.

NASCIMENTO, Ivanete B. *O uso variável de você(s)/cê(s) na cidade de São Paulo*. Dissertação de Mestrado. Universidade de São Paulo. São Paulo, 2011.

NEBRIJA, A. *Gramática de la lengua castellana*. Edición preparada por Antonio Quilis. Madrid: Editora Nacional, 1980.

NEVES, M. H. de M. *Estudos funcionalistas no Brasil*. D.E.L.T.A., v. 15, n. esp., p. 71-104. 1999.

_____. *Gramática de usos do português*. São Paulo: Unesp, 2000.

_____. *A gramática funcional*. São Paulo: Martins Fontes, 2004.

NICHOLS, J. *Functional theories of grammar*. Annual Review of Anthropology, Stanford, v. 13, p. 7-117, 1984.

OLIVEIRA, Dercir Pedro. *O Atlas Linguístico de Mato Grosso do Sul*. SIGNUM: Estudos Linguísticos, Londrina, n.9/2, dez. 2006. p.169-183.

<http://www.uel.br/revistas/uel/index.php/signum/article/viewFile/3944/3153>

_____. (Org). *ALMS – Atlas Linguístico de Mato Grosso do Sul*. 1. ed. Campo Grande: Editora UFMS, 2007.

OUSHIRO, Livia. *Identidade na pluralidade: avaliação, produção e percepção linguística na cidade de São Paulo*. Tese (Doutorado em Linguística). São Paulo, Universidade de São Paulo, FFLCH. 2015.

_____. MENDES, Ronald Beline. *A pronúncia do (-r) em coda silábica no português paulistano*. Revista do GEL, vol. 8(2), 66–95, 2013. URL <http://revistadogel.gel.org.br/rg/article/view/39/21>

PAIVA, Maria da Conceição de. *Ordenação de cláusulas causais: forma e função*. 1991. Tese (Doutorado em Linguística). 238 f. Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1991.

_____. *Isomorfismo informacional entre período simples e período composto*. Alfa, São Paulo, 41(n.esp.), 127-140, 1997.

_____; & SCHERRE, M. M. P. *Retrospectiva sociolinguística: contribuições do PEUL*. DELTA, vol. 15, n.º especial, 1999, p. 201-232.

_____; BRAGA, M. L. *Cláusulas causais introduzidas por porque: da sintaxe ao discurso*. In: MOLLICA, Maria Cecília. (Org.). *Usos da linguagem e sua relação com a mente humana*. Rio de Janeiro: Editora Tempo Brasileiro, v. 1, p. 55-71, 2010.

PEZATTI, Erotilde Goreti. *O advérbio então já se gramaticalizou como conjunção?* DELTA., São Paulo, v. 17, n. 1, 2001.

_____. *O Funcionalismo em linguística*. In: MUSSALIM, Fernanda e BENTES, Anna Christina (orgs.). *Introdução à Linguística: fundamentos epistemológicos*. Vol. 3. São Paulo: Cortez, 2004.

PERINI, Mário A. *Níveis de detalhamento na descrição gramatical – Uma perspectiva pedagógica*. Trabalhos em Linguística Aplicada nº 12. Campinas: Unicamp/IEL – setor de publicações, 1988.

POPLACK, S. *Grammaticalization and linguistic variation*. In: NARROG, Heiko; HEINE, Bernd (Eds.). *The Oxford handbook of grammaticalization*. Oxford: Oxford University Press, 2011. p. 209-224.

RODRIGUES, C. de J. M. *Gênero e aprendizagem participativa orientada para a ação em educação sexual: um estudo com alunos(as) do 7º ano de escolaridade*. 2009. Dissertação (Mestrado em Educação) - Universidade do Minho, Braga, 2009.

ROSÁRIO, Ivo da Costa do. *Funcionalismo em linguística: raízes passadas e irradiações futuras*. Via Litterae, Anápolis, v. 2, n. 2, p. 435-450, jul./dez. 2010. <http://www2.unucseh.ueg.br/vialitterae/index.php?id=70>

SAID ALI, Manuel. *Gramática histórica da língua portuguesa*. 3ª ed., São Paulo: Edições Melhoramentos, 1964.

SANKOFF, D.; TAGLIAMONTE, S.; SMITH, E. *Goldvarb X: A variable rule application for Macintosh and Windows*. Department of Linguistics, University of Toronto, 2005.

SANTOS, W. S. *A morfologia do indicativo na expressão do subjuntivo em São Paulo e São Luís*. Dissertação de Mestrado. Universidade de São Paulo. São Paulo, 2015.

SCHERRE, Maria Marta Pereira. *Padrões sociolinguísticos do português brasileiro: a importância da pesquisa variacionista*. In: Tabuleiro de Letras, n. 04, junho de 2012.

http://www.tabuleirodeletras.uneb.br/secun/numero_04/pdf/no04_artigo09.pdf

SCHUCHARDT, Hugo E.M. *Über die Lautgesetze. Gegen die Junggrammatiker*, Berlin, Oppenheim, 1885.

SEARLE, John R. *Os Atos de fala*. Coimbra: Almedina, 1981.

SMITH, C. *The Parameter of Aspect*. Dordrecht: Kluwer, 1991.

SOUZA, Edson Rosa Francisco. *Gramaticalização dos itens linguísticos assim, já e aí no Português Brasileiro: um estudo sob a perspectiva da Gramática Discursivo-Funcional*. -- Campinas, SP: [s.n.], 2009.

SWEETSER, Eve. *From etymology to pragmatics*. Cambridge: Cambridge University, 1991. p. 23-48.

TAGLIAMONTE, Sali. *Variationist sociolinguistics: change, observation, interpretation*. Malden, MA: Wiley-Blackwell, 2011.

_____.; *Comparative linguistics*. In Peter Trudgill, Jack Chambers & Natalie Schilling-Estes (eds.), *The handbook of language variation and change*. Oxford: Blackwell, 2002, p. 729-763.

TARALLO, Fernando. *Pesquisa Sociolinguística*. São Paulo: Ática, 2010.

TAVARES, M. A. *A gramaticalização de e, aí, daí e então: estratificação/variação e mudança no domínio funcional da sequenciação retroativo-propulsora de informações – um estudo sociofuncionalista*. Florianópolis, 2003. Tese (Doutorado em Linguística) – Universidade Federal de Santa Catarina.

_____. *A gramaticalização do aí como conector – indícios sincrônicos*. Working Papers in Linguística. UFSC, n.3, 1999.

<https://periodicos.ufsc.br/index.php/workingpapers/article/download/2316/2008>

_____. *Sociofuncionalismo: um duplo olhar sobre a variação e a mudança linguística*. Edição Especial ABRALIN/SE, Itabaiana/SE, Ano VIII, v.17, jan./jun. 2013.

TRAUGOTT, Elizabeth & Bernd Heine (eds.). *Approaches to Grammaticalization*. Vol.1: Focus on Theoretical and Methodological Issues. Amsterdam/ Philadelphia: John Benjamins, 1991, p. 189-218.

_____.; KÖNIG, *The semantic-pragmatics of grammaticalization revisited*. TRAUOGOTT, E., HEINE, B. (orgs.) *Approaches to grammaticalization*. Vol. 1. John Benjamins Publishing Company, 1991.

_____.; C. *Constructions in grammaticalization*. In: Joseph, B. D. & R. D. Janda (eds.) *The Handbook of Historical Linguistics*. Cambridge USA / Oxford UK: Blackwell, 2003, pp. 624 – 647.

TRAVAGLIA, Luiz Carlos. *O aspecto verbal no português: a categoria e sua expressão*. ed. rev. Uberlândia: Universidade Federal de Uberlândia, 2006, p.49-54.

VALLE, C. R. M. *Sabe?~ não tem? ~ entende?: itens de origem verbal em variação como requisitos de apoio discursivo*. Dissertação (Mestrado em Linguística), Florianópolis, UFSC, 2001.

VAN, DJIK. *Prejudice in discourse: an analysis of ethnic prejudice in cognition and conversation*. Amsterdam: Benjamins, 1984.

VENDLER, Z. *Linguistics in Philosophy*. Ithaca, NY: Cornell, 1967.

VENDRYES, J. *Le langage*. Paris: La Renaissance du Livre, 1921, p.117.

VIEIRA, Marília S. *Aí, daí e então: contextos discursivos e variação sociolinguística*. Revista De Letra em Letra. Vol. 2, 2015, p.121-130.
http://issuu.com/revistadeletraemletra/docs/revistadeletraemletra_vol2_2015

VOTRE, Sebastião. *Cognitive verbs in Portuguese and Latin – unidirectionality revisited*. University of California, Santa Barbara, 1999.

WILLIAMS, Edwin B. *Do latim para o português*. Rio de Janeiro, Tempo Brasileiro, 1975.

WEINER, E. J e LABOV, W. 1983. *Constraints on the agentless passive*. Journal of Linguistics, 19: 29-58.

WEINREICH, U., LABOV, W.; HERZOG, M. I. *Fundamentos empíricos para uma teoria da mudança linguística*. Tradução: Marcos Bagno. São Paulo: Parábola Editorial, 2006.

ZILLES, Ana Maria S. *O que a fala e a escrita nos dizem sobre a avaliação social do uso de a gente?* In: Letras de Hoje, Porto Alegre, v. 42, n. 2, p.27-44, junho de 2007.

Índice remissivo

A

Adam (1992), 31
Almeida & Marinho (2012), 30, 32
Amorim (2012), 50

B

Bailly (1950), 39
Basso & Ilari (2004), 133
Beaugrande & Dressler (1981), 31
Bechara (2003), 31
Benveniste (1974), 45
Bolinger (1977), 75, 86
Borges Neto & Müller (1987), 86
Bourdieu (1991), 25, 118, 153
Braga e Paiva (2003), 21
Bybee (2012), 85, 86
Bybee et al (1994), 67

C

Calvet (2002), 78
Camacho (2003), 86
Cambraia (2002), 38
Castilho (1968), 112, 135, 137
Castilho (1997), 66
Cedergren & Sankoff (1974), 143
Chafe (1987), 128
Chafe (1998), 128
Chambers & Trudgill (1980), 117
Chambers (1995), 25
Cheshire (2005), 117
Cheshire, Kerswill e Williams (2005), 28
Colomer & Camps (2002), 123
Comrie (1976), 130, 133, 136
Costa (2003), 38
Cunha & Cintra (2001), 24
Cunha (1975), 137
Cunha (2001), 115

Cunha (2009), 64
Cunha et al (2003), 60, 64
Cunha & Cintra (2001), 24

D

Dal Mago (2001), 121

E

Ernout e Meillet (1959), 38

F

Faraco (2004), 58
Faraco (2005), 81
Faria (1975), 39
Figueroa (1996), 79
Figueiredo-Gomes & Pena-Ferreira (2006), 40
Freitag (2003), 121
Freitag (2009), 120
Freitag (2011), 107, 113
Freitag et al (2009), 112
Freitag et al (2009), 119

G

Galbiatti (2008), 73
Gil (2002), 17
Givón (1976), 60
Givón (1984), 35, 63, 90
Givón (1990), 75, 76, 128
Givón (1993), 124
Givón (1995), 19, 31, 48, 64, 77, 78, 89, 124, 125, 128
Givón (2001), 116, 120, 121, 124, 166, 168, 169, 170, 172, 174, 175
Gonçalves (2003), 69
Gonçalves (2008), 98
Gonçalves et al (2007), 20, 63, 68, 71

Gorski (1994), 120
 Guy (1991), 143
 Guy (2001), 19, 20, 81, 142
 Guy (2007), 109
 Guy et al (1986), 125
 Guy & Zilles (2007), 108, 111

H

Halliday & Hasan (1976), 52
 Halliday (1985), 128
 Harris & Campbell (1995), 74
 Heine et al (1991), 38, 40, 58, 59, 72, 73
 Hellwig & Geerts (2013), 96
 Hintze (2003), 71
 Hopper & Traugott (1993), 31, 65, 66, 72, 74
 Hopper & Traugott, 2003, 87
 Hopper (1991), 10, 36, 67, 68, 69, 70, 71, 90, 131, 136, 147, 148
 Hopper e Traugott (2003), 87
 Hünemeyer, Claude e Heine (1991), 39

I

IBGE (2015), 104
 Ilari (2004), 60
 Ipea (2006), 105

J

Jhonson (2009), 93, 112
 Johnson (2009), 92, 108, 111, 112

K

Koch (1987), 46
 Koch (1998), 21, 124

L

Labov ([1972] 2008), 20, 23, 24, 29, 79, 81, 83, 85, 93, 112, 120

Labov (1963), 79
 Labov (1969), 143
 Labov (1978), 28, 34, 85, 86, 88, 89
 Labov (1982), 84, 86, 89
 Labov (1984), 27
 Labov (1994), 80, 82, 84
 Labov (2001), 79, 87
 LABOV (2006), 95
 Labov e Waletzky (1997), 120, 123
 Lavandera (1978), 28, 33
 Leal e Oliveira (2007), 132
 Levinson (1983), 51
 Longhin-Thomazi (2003), 68, 72
 Lopes (2010), 71
 Loureiro (2004), 30
 Lyons (1970), 60, 62
 Lyons (1997), 136

M

Macedo & Silva (1996), 20, 117, 118
 Meyerhohh, Scleef & Smith (2005), 109
 Mann & Thompson (1985), 61
 Marcuschi (1975), 78
 Marcuschi (1991), 129
 Marcuschi (2002), 123
 Margarit (1947), 38
 Martelotta & Areas (2003), 63
 Martelotta (1993), 21
 Martelotta (1999), 20, 21
 Martelotta (2009), 77
 Mathesius (1982), 60
 McMahon (1994), 76
 Meillet (1948 [1912]), 59, 72, 74
 Mendes (2005), 30, 34, 70
 Milroy (2003), 99, 102
 Milroy e Gordon (2003), 28, 29
 Moreira de Sá (1999), 122
 Moss (2004), 123

N

Naro & Braga (2001), 36

Naro (1992), 117
 Naro & Scherre (2006), 33
 Nascimento (2011), 116
 Neves (1997), 63, 65
 Neves (1999), 86
 Neves (2000), 49, 50
 Nichols (1984), 34, 35, 62

O

Oliveira (2006), 106
 Oliveira (2007), 19, 20
 Omena & Braga (1996), 69
 Oushiro (2014), 111, 142, 143
 Oushiro (2015), 19, 20, 143

P

Paiva (1991), 47
 Paiva (1995), 20, 112, 140
 Paiva (1997), 127
 Paiva (2003), 116
 Perini (1998), 31
 Pezatti (2001), 53
 Pezatti (2004), 59
 Pezatti (2008), 21

R

Rodrigues (2009), 118
 Romaine (1981), 28
 Romaine (2006), 25
 Rosário (2010), 59

S

Said Ali (1964), 137
 Sankoff & Labov (1979), 143
 Sankoff (1988), 108
 Santos (2012), 114
 Santos (2015), 20, 114
 Bates, 108
 Scherre (2012), 108
 Searle (1981), 51

Macedo & Silva (1996), 118
 Smith (1991), 132, 149
 Sweetser (1991), 21, 23, 49, 50, 52, 140

T

Tagliamonte (2002), 88, 90
 Tagliamonte (2011), 108, 110
 Tarallo (2010), 82
 Tavares (1999), 20, 121, 124
 Tavares (2003), 20, 21, 32, 36, 38, 46, 112, 115, 116, 117, 118, 120, 121, 125, 167, 168, 172, 175
 Tavares (2013), 90
 Traugott & König (1991), 41, 43, 72, 73
 Traugott (1982), 66
 Traugott (1991), 88
 Traugott e Heine (1991), 17, 19
 Traugott e König (1991), 87
 Travaglia (1994), 134
 Travaglia (2006), 134, 135, 137, 138 170

V

Valle (2001), 120
 Van Dijk (1984), 61
 Vendler (1957), 130, 150
 Vendler (1967), 132
 Votre (2003), 118, 148

W

Weiner e Labov (1983), 28, 33
 Weinrech, Labov & Herzog (2006), 80, 83
 Williams (1975), 39

Z

Zilles (2007), 84

Anexos

Anexo A – São Paulo

Entrevistas com paulistanos

SP2010: Construção de uma amostra da fala paulistana” (Processo Fapesp nº 2011/09278-6)

I. Primeira parte BAIRRO (aproximadamente 5 minutos)

Objetivos: descobrir o “grau de enraizamento” do informante no bairro onde vive/outros bairros; descobrir padrões de sociabilidade nos diferentes bairros; descobrir o “grau de mobilidade” da pessoa.

1. Há quanto tempo você mora na (Mooca, Pinheiros, Bexiga...)?
2. Você gosta de morar aqui? 3. Por que você escolheu morar neste bairro? (manter em mente que o informante pode não ter “escolhido” morar ali: pode ter sido porque a família já morava lá etc.) a. (Se o informante mora há bastante tempo) Como era o bairro antes/quando você se mudou pra cá? Mudou muito? b. (Se o informante não mora lá há muito tempo, ou se nem sempre viveu ali)
3. Em que outros lugares você já morou? Como era lá comparado com esse bairro aqui? Onde você preferia morar?
4. Você conhece seus vizinhos?
5. O que esse bairro tem de diferente de outros bairros na cidade?
6. Aqui costuma haver festas do bairro? Existe algum lugar no bairro em que as pessoas se reúnem?
7. As pessoas se ajudam por aqui? a. Se você precisa de ajuda, a quem você recorre? Se você ficar doente, a quem você pode pedir para tomar conta de sua família?
8. Com quais pessoas você tem mais contato?
9. Já aconteceu alguma coisa aqui que te fez pensar em se mudar?
10. Hoje, tem algum outro bairro em que você gostaria de morar?

INFÂNCIA (aproximadamente 5 minutos)

Objetivos: relaxar o informante (em geral, as pessoas gostam de falar de sua infância); obter informações sobre mudanças no bairro/cidade de São Paulo; grau de mobilidade do informante; obter informações sobre escolaridade

11. E como foi a sua infância (no bairro X)? Você pode contar um pouco de como foi, o que você fazia...? a. brincava na rua/dentro de casa? Do que vocês brincavam? b. Como eram os seus pais? Eram rígidos...? Você tinha horário para estar em casa? c. Vocês tinham alguma tradição de família?

12. Você foi pra escola no mesmo bairro? Como era a escola? Você sempre estudou na mesma escola? Você gostava de ir pra escola? Tem algum professor que te marcou? Até que série você estudou? Você acha que a escola fornece aquilo que uma pessoa precisa para encontrar um emprego?

13. Enquanto ainda era criança/adolescente, você ia pra outros lugares dentro da cidade de São Paulo? (pra onde, pra fazer o quê...)

14. Que roupas você usava? Que tipo de corte de cabelo as pessoas usavam?

FAMÍLIA (aproximadamente 5 minutos)

Objetivos: obter informações sobre rede social do informante, grau de enraizamento no bairro/cidade

15. Você tem irmãos? (é possível que esta informação já tenha aparecido na parte sobre infância; neste caso, falar algo como: “Você disse que tem uma irmã... Você tem outros irmãos?”) Quantos anos eles têm?

16. Onde seus pais nasceram? (Se não são paulistanos, perguntar também sobre avós, bisavós... até encontrar a primeira geração da família que veio pra cá). Quantos anos eles (pais) têm? Quando seus (pais/avós/bisavós) vieram pra São Paulo? Você sabe por que eles vieram?

17. E o resto da família, tios, primos, também vivem aqui em São Paulo? (Se sim, em que bairros? Moram próximo? Se não, onde? Têm contato sempre?)

18. Você é casado? Você tem filhos? Quantos anos eles têm?

19. Quais são as pessoas que vivem na mesma casa que você?

20. Como é a vida em família hoje em São Paulo? a. É muito diferente de quando você era criança? (Para os mais velhos) b. É muito diferente do que seus pais contam para você? (Para os mais jovens) c. Muitas pessoas falam que as crianças hoje são mal-educadas. O que você acha disso? Você acha que as crianças hoje têm menos respeito pelos adultos?

21. No passado, esperava-se que as mulheres ficassem em casa para tomar conta das crianças. E hoje, como é isso? Mudou? a. Em sua casa, os homens ajudam nos afazeres domésticos? O que você acha de um homem ficar em casa e cuidar dos filhos?

22. O que você acha da lei sobre casamento gay, que foi recentemente aprovada?

TRABALHO/OCUPAÇÃO (aproximadamente 5 minutos)

Objetivos: obter informações sobre rede social do informante; características socioeconômicas

23. Você trabalha aqui por perto? (se não souber onde a pessoa trabalha)

24. Como você faz para chegar até o seu trabalho? Quais meios de transporte você utiliza?

25. O que você faz? Faz tempo que você trabalha nesse serviço?

26. Você gosta do seu serviço? (Se não, o que a pessoa preferiria fazer?)
27. Você se sente reconhecido no seu trabalho?
28. Você quer alcançar outro cargo? O que você almeja no seu trabalho?
29. Qual é a profissão dos seus sonhos?
30. Se você ganhasse na Mega Sena, o que você faria?
31. As pessoas devem continuar trabalhando, mesmo se elas têm muito dinheiro?

LAZER (aproximadamente 5 minutos)

Objetivos: obter informações sobre rede social do informante; mobilidade na cidade; características socioeconômicas

32. E nas horas de lazer, o que você e sua família gostam de fazer? (Se saem, vão pra que lugares?) Você acha que a cidade de São Paulo tem boas opções de lazer? Quais?
33. A maioria dos seus amigos mora aqui nesse bairro mesmo? (Se não, onde?)
34. Você tem alguma atividade de recreação em grupo, algum clube...?
35. Quais são seus amigos mais antigos? Você mantém contato com os amigos de colégio?
36. Você costuma viajar? Pra que lugares já viajou? Que lugares gostaria de conhecer?

II. Segunda parte A CIDADE DE SÃO PAULO (aproximadamente 5 minutos)

37. Uma pesquisa recente mostrou que 57% das pessoas em São Paulo deixariam a cidade caso pudessem. O que você acha disso? (se o informante perguntar, a pesquisa é do Ibope e foi publicada em janeiro/2010³⁷)
38. Você gosta de morar em São Paulo? (Se não, em que lugar preferiria morar?) Por quê?
39. O que você acha que caracteriza a cidade (tanto as coisas boas quanto ruins)?
40. O que você acha que caracteriza o paulistano (tanto as coisas boas quanto ruins)?
41. Olhando pra mim, você diria que eu sou paulistano(a)? Por quê?
42. O que você mais gosta em São Paulo?
43. O que você não gosta em São Paulo? (a depender do tópico mencionado pelo informante - violência, trânsito, poluição etc. - procurar explorar mais o assunto. P.ex.: você já foi assaltado? O que aconteceu? (para obter narrativa pessoal) O

³⁷http://www.nossasaopaulo.org.br/portal/arquivos/Pesquisa_IRBEM_Ibope_2010_completa.pdf. Último acesso em dezembro de 2015.

que o governo deve fazer pra solucionar esse problema? (para obter uma fala mais distanciada)) (explorar o subtópico por aproximadamente 10 min.)

44. Pras pessoas que não vivem em São Paulo, como você acha que elas imaginam que seja a cidade? Qual é a imagem que as pessoas de fora de São Paulo têm da cidade?

45. Você falou que já foi pra (X, Y, Z). Quando você foi pra esses lugares, as pessoas percebiam que você era paulistano? (Se sim) como elas percebiam?

46. Quando você conhece alguém, você percebe se a pessoa não é daqui de São Paulo? a. (Se sim) como você percebe? (Se o informante mencionar o modo de falar, seguir nessa linha e tentar conseguir informações mais precisas. É normal as pessoas não saberem definir o porquê, mas devemos tentar tirar mais informações). b. (Se não) Quando você ouviu uma pessoa falando, por exemplo, você percebe que ela é de fora pelo sotaque?

PRODUÇÃO/PERCEPÇÃO/AVALIAÇÃO LINGUÍSTICA (aproximadamente 5 minutos)

47. Já que a gente está falando de sotaque... qual o sotaque do Brasil que você mais gosta? E tem algum que te irrita? Como é que (o gaúcho/o carioca/o mineiro/o caipira etc. – a depender dos sotaques mencionados) fala(m)?

48. Aqui em São Paulo tem muito (i)migrante, né? De onde vêm a maioria das pessoas? Tem algum bairro específico em que eles (italianos, coreanos, nordestinos etc.) se concentram? Hoje em dia, existem bairros étnicos em São Paulo?

49. E como é que as pessoas falam na cidade de São Paulo? (evitar usar a palavra “sotaque”)

50. Dentro da cidade de São Paulo, você consegue identificar se a pessoa é de alguma região ou bairro específico da cidade? (Se sim) Você poderia dar alguns exemplos? (Imprimir a lista de palavras, a notícia e o trecho, todos abaixo, para mostrar ao informante) LISTA de PALAVRAS:

Apresentar ao informante a lista de palavras e dizer: “Agora eu queria te pedir pra ler algumas coisas. Eu tenho aqui uma lista de palavras, e eu queria que você lesse cada uma delas. Você pode dar uma olhada primeiro, antes de começar a ler.” Esperar que o informante leia a lista de palavras. Depois perguntar:

51. A gente estava falando de sotaques agora há pouco... Como você acha que uma pessoa do interior falaria algumas dessas palavras? Tem diferenças?

52. E como um carioca falaria algumas dessas palavras? 53. Tem mais algum sotaque no Brasil que você conhece?

LEITURA DE NOTÍCIA: Dizer ao informante: “Agora eu queria que você lesse esse texto. Você pode ler em voz baixa antes de falar em voz alta.” Esperar que o informante leia a notícia.

LEITURA DE TRECHO: Dizer ao informante: “Por último, queria que você lesse esse trechinho.” Esperar que o informante leia o trecho.

54. (Sobre /e/ nasal) O que você acha desse modo de falar: “Você tá entendendo o que eu tô dizendo?” (com “en” ditongado e exagerado). Se a pessoa manifestar uma atitude negativa, perguntar: Como você acha que deveria ser? O que tem de errado aí?

55. Quem você acha que fala assim? (se falar “paulistanos”, perguntar: “você acha que todos os paulistanos falam assim ou é uma coisa de uma região mais específica na cidade?”)

56. Você fala desse modo?

57. (Sobre concordância nominal) E o que você acha de “Me vê dois pastel e um chopps?” (Repetir as perguntas 52-53)

58. (Sobre /r/ retroflexo) E o que você acha de “A porta tá aberta.” (com retroflexo exagerado). (Repetir as perguntas 52-53).

59. E tem algum modo de falar que você acha assim bem paulistano?

Palavras

alma
amargo
animação
argola
atitude
barqueiro
biscoito
cacique
carteiro
cerca
chácara
circo
cisne
curto
defender
discoteca
elefante
enchente
entender
entretenimento
erguer
fazenda
felicidade
firme
fornalha
furgão
fusquinha
geleira
gérmen de trigo
gordo
gula
Hércules
hilário

irmã
justiça
lento
mortadela
Marba
mosca
necessidade
orca
Ordem e progresso
orgânico
órgão
ostracismo
penteado
pertencimento
perto
porto
presente
rapidez
riqueza
sabor de menta
soberba
trabalho
turco
urgente
utilidade
vulto
zebra

Notícia

08/02/2010

Moradores cobram solução para área alagada em SP e aguardam encontro com prefeito Moradores da região do Jardim Pantanal protestaram nesta segunda-feira em frente à Prefeitura de São Paulo contra a inundação em bairros da zona leste, que completa dois meses hoje. Uma comissão de manifestantes foi recebida pela Secretaria de Relações Institucionais e apresentou reivindicações, mas o prefeito Gilberto Kassab (DEM) não participou do encontro.

Representantes dos manifestantes esperam que o governador José Serra (PSDB) também participe da reunião de sexta. De acordo com os integrantes da comissão, não havia integrantes do governo estadual no encontro realizado hoje na prefeitura. Além de cobrar a limpeza imediata das águas nos bairros, os moradores cobram uma solução para a falta de moradia. A maioria dos desalojados continua em escolas municipais e, com o início do ano letivo, temem não ter para onde ir.

Trecho

Tá chovendo muito! Choveu tanto, tanto na semana passada que ficou uma piscina na minha casa. Ó, pra você ver: Molharam todos os armários, a cama, os colchões, tudo... Foi um sacrifício... O que a gente fez? Nós tivemos que erguer os móveis pra limpar tudo: a geladeira, o forno... minha irmã até veio me ajudar, sabe? E meus filhos compraram umas cadeiras novas, mas é aquela coisa, assim... quando chover de novo, vai molhar tudo outra vez. Você fica sem ter o que fazer.

E tem um rio lá perto que sempre alaga... quer dizer, é água dentro e fora de casa! Daí, o que acontece? Fica aquele trânsito, os carros todos parados, a gente demora um tempão pra chegar em casa... Não aguento mais enchente nessa cidade... Agora que eu vou fazer? Os políticos falam, falam, mas eles tinham que fazer alguma coisa urgente. Você tá entendendo o que eu tô dizendo?

Anexo II

Declaração de livre consentimento

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, LETRAS E CIÊNCIAS HUMANAS
Av. Prof. Luciano Gualberto, 403 - Cx. Postal: 26097 - CEP: 01060-970 - São Paulo
Fone: (011) 3091.4298 - e-mail: flil@usp.br

Projeto de Pesquisa: SP2010 – Construção de uma amostra da fala paulistana

Docente responsável: Prof. Dr. Ronald Beline Mendes

Contato: rbeline@usp.br

Declaração

Declaro ter consentido em ter gravada minha conversa com o aluno universitário _____, como parte do projeto acima identificado, em execução pelo Grupo de Estudos e Pesquisa em Sociolinguística da USP. Estou ciente de que a gravação será transcrita e disponibilizada na Internet para o desenvolvimento de estudos sociolinguísticos na cidade de São Paulo. Entendo que essa gravação é idônea e que meus dados pessoais não serão tornados públicos na divulgação de resultados da pesquisa.

São Paulo, ____ de _____ de ____.

Nome do informante: _____

Assinatura: _____

Anexo III

Ficha do Informante e da Gravação

Nome do Informante:

Pseudônimo:

Data de Nascimento:

Idade:

Endereço:

Telefones para contato:

E-mail:

Ocupação:

Pessoas com quem mora:

Documentador:

Arquivo de gravação e transcrição:

Bairros em que já morou na cidade de São Paulo, em ordem cronológica, e tempo aproximado (em anos) de quanto tempo morou em cada bairro:

P.ex.: Tatuapé – 10 anos

Número de irmãos:

() mais velhos

() mais novos

Nome, idade e escolaridade dos irmãos:

Geração na cidade: () 0 – pais não nascidos na cidade de São Paulo

() 1 – pai ou mãe nascido na cidade de São Paulo

- () 2 – pai e mãe nascidos na cidade de São Paulo
() 3 – um avô ou uma avó nascido na cidade de São Paulo
() 4 – dois ou mais avós nascidos na cidade de São Paulo

Nome do Pai:

Idade do Pai:

Grau de Escolaridade do Pai:

Ocupação do pai:

Naturalidade: () São Paulo (capital) () outra: _____

Nome da Mãe:

Idade da Mãe:

Grau de Escolaridade da Mãe:

Ocupação da mãe:

Naturalidade de: () São Paulo (capital) () outra: _____

Escolas em que estudou e tipo de escola (pública municipal (M), pública estadual (E) ou particular (P)):

P.ex.: Fundamental 1: Escola Estadual President Roosevelt (E)

Fundamental 1: _____

Fundamental 2: _____

Ensino Médio: _____

Faculdade: _____

(concluído ou não)

Data da Entrevista:

Horário:

Local da Entrevista e breve descrição do local:

Presença e atuação de terceiros:

Como o documentador conheceu o informante:

Grau de relação entre informante e documentador:

() 1 – Bastante próximo. O entrevistado faz parte do meu círculo imediato de amigos/familiares e conversamos frequentemente.

- () 2 – Próximo. Conversamos frequentemente, mas o entrevistado não faz parte do meu círculo imediato de amigos/familiares.
- () 3 – Próximo, mas não conversamos frequentemente.
- () 4 – Neutro. Ele é meu conhecido, mas não nos falamos com frequência.
- () 5 – Distante. Não o conhecia anteriormente e praticamente só conversamos na ocasião da entrevista.

Outras informações:

Anexo III

Questionário Socioeconômico

Por favor, assinale a opção que melhor descreve sua situação:

1) Em que tipo de imóvel você vive?

☐ Casa

☐ Apartamento

☐ Outro: _____

☐ 1

☐ 2

☐ 3 ou mais

2) Há quantos anos o imóvel foi construído?

☐ Há menos de 10 anos

☐ De 10 a 20 anos

☐ Há mais de 20 anos

☐ 1

☐ 2

☐ 3

☐ 4

3) Você é proprietário do imóvel em que vive?

☐ Sim, é um imóvel próprio e quitado.

☐ Sim, é um imóvel financiado.

☐ Não, é um imóvel alugado.

☐ 5

☐ 6

4) Quantos quartos tem a sua casa?

☐ 0

☐ 1

☐ 2

☐ 3

5) Quantas pessoas moram em sua casa?

6) Quantas pessoas que vivem em sua casa contribuem para a renda familiar?

☐ 4

☐ 5

7) Qual é a sua renda média mensal?

☐ Até R\$ 600,00

☐ De R\$ 600,00 a R\$ 1.200,00

☐ De R\$ 1.200,00 a R\$ 2.400,00

☐ De R\$ 2.400,00 a R\$ 6.000,00

☐ De R\$ 6.000,00 a R\$ 12.000,00

☐ Acima de R\$ 12.000,00

8) Qual é a renda média mensal de sua família?

☐ Até R\$ 600,00

☐ De R\$ 600,00 a R\$ 1.200,00

☐ De R\$ 1.200,00 a R\$ 2.400,00

☐ De R\$ 2.400,00 a R\$ 6.000,00

☐ De R\$ 6.000,00 a R\$ 12.000,00

☐ Acima de R\$ 12.000,00

Anexo B – Campo Grande

Roteiro de entrevistas com campo-grandenses

I. Primeira parte BAIRRO (aproximadamente 5 minutos)

Objetivos: descobrir o “grau de enraizamento” do informante no bairro onde vive/outros bairros; descobrir padrões de sociabilidade nos diferentes bairros; descobrir o “grau de mobilidade” da pessoa.

1. Há quanto tempo você mora nesse bairro?
2. Você gosta de morar aqui? 3. Por que você escolheu morar neste bairro? (manter em mente que o informante pode não ter “escolhido” morar ali: pode ter sido porque a família já morava lá etc.) a. (Se o informante mora há bastante tempo) Como era o bairro antes/quando você se mudou pra cá? Mudou muito? b. (Se o informante não mora lá há muito tempo, ou se nem sempre viveu ali)
3. Em que outros lugares você já morou? Como era lá comparado com esse bairro aqui? Onde você preferia morar?
4. Você conhece seus vizinhos?
5. O que esse bairro tem de diferente de outros bairros na cidade?
6. Aqui costuma haver festas do bairro? Existe algum lugar no bairro em que as pessoas se reúnem?
7. As pessoas se ajudam por aqui? a. Se você precisa de ajuda, a quem você recorre? Se você ficar doente, a quem você pode pedir para tomar conta de sua família?
8. Com quais pessoas você tem mais contato?
9. Já aconteceu alguma coisa aqui que te fez pensar em se mudar?
10. Hoje, tem algum outro bairro em que você gostaria de morar?

INFÂNCIA ((aproximadamente 5 minutos)

Objetivos: relaxar o informante (em geral, as pessoas gostam de falar de sua infância); obter informações sobre mudanças no bairro/cidade de Campo Grande; grau de mobilidade do informante; obter informações sobre escolaridade

11. E como foi a sua infância (no bairro X)? Você pode contar um pouco de como foi, o que você fazia...? a. brincava na rua/dentro de casa? Do que vocês brincavam? b. Como eram os seus pais? Eram rígidos...? Você tinha horário para estar em casa? c. Vocês tinham alguma tradição de família?

12. Você foi pra escola no mesmo bairro? Como era a escola? Você sempre estudou na mesma escola? Você gostava de ir pra escola? Tem algum professor que te marcou? Até que série você estudou? Você acha que a escola fornece aquilo que uma pessoa precisa para encontrar um emprego?

13. Enquanto ainda era criança/adolescente, você ia pra outros lugares dentro de Campo Grande? (pra onde, pra fazer o quê...)

14. Que roupas você usava? Que tipo de corte de cabelo as pessoas usavam?

FAMÍLIA ((aproximadamente 5 minutos))

Objetivos: obter informações sobre rede social do informante, grau de enraizamento no bairro/cidade

15. Você tem irmãos? (é possível que esta informação já tenha aparecido na parte sobre infância; neste caso, falar algo como: “Você disse que tem uma irmã... Você tem outros irmãos?”) Quantos anos eles têm?

16. Onde seus pais nasceram? (Se não são paulistanos, perguntar também sobre avós, bisavós... até encontrar a primeira geração da família que veio pra cá). Quantos anos eles (pais) têm? Quando seus (pais/avós/bisavós) vieram pra Campo Grande? Você sabe por que eles vieram?

17. E o resto da família, tios, primos, também vivem aqui em Campo Grande? (Se sim, em que bairros? Moram próximo? Se não, onde? Têm contato sempre?)

18. Você é casado? Você tem filhos? Quantos anos eles têm?

19. Quais são as pessoas que vivem na mesma casa que você?

20. Como é a vida em família hoje em Campo Grande? a. É muito diferente de quando você era criança? (Para os mais velhos) b. É muito diferente do que seus pais contam para você? (Para os mais jovens) c. Muitas pessoas falam que as crianças hoje são mal-educadas. O que você acha disso? Você acha que as crianças hoje têm menos respeito pelos adultos?

21. No passado, esperava-se que as mulheres ficassem em casa para tomar conta das crianças. E hoje, como é isso? Mudou? a. Em sua casa, os homens ajudam nos afazeres domésticos? O que você acha de um homem ficar em casa e cuidar dos filhos?

22. O que você acha da lei sobre casamento gay, que foi recentemente aprovada?

TRABALHO/OCUPAÇÃO (aproximadamente 5 minutos)

Objetivos: obter informações sobre rede social do informante; características socioeconômicas

23. Você trabalha aqui por perto? (se não souber onde a pessoa trabalha)

24. Como você faz para chegar até o seu trabalho? Quais meios de transporte você utiliza?

25. O que você faz? Faz tempo que você trabalha nesse serviço?

26. Você gosta do seu serviço? (Se não, o que a pessoa preferiria fazer?)
27. Você se sente reconhecido no seu trabalho?
28. Você quer alcançar outro cargo? O que você almeja no seu trabalho?
29. Qual é a profissão dos seus sonhos?
30. Se você ganhasse na Mega Sena, o que você faria?
31. As pessoas devem continuar trabalhando, mesmo se elas têm muito dinheiro?

LAZER (aproximadamente 5 minutos)

Objetivos: obter informações sobre rede social do informante; mobilidade na cidade; características socioeconômicas

32. E nas horas de lazer, o que você e sua família gostam de fazer? (Se saem, vão pra que lugares?) Você acha que Campo Grande tem boas opções de lazer? Quais?
33. A maioria dos seus amigos mora aqui nesse bairro mesmo? (Se não, onde?)
34. Você tem alguma atividade de recreação em grupo, algum clube...?
35. Quais são seus amigos mais antigos? Você mantém contato com os amigos de colégio?
36. Você costuma viajar? Pra que lugares já viajou? Que lugares gostaria de conhecer?

II. Segunda parte CAMPO GRANDE (aproximadamente 5 minutos)

37. Uma pesquisa recente mostrou que 57% das pessoas em Campo Grande deixariam a cidade caso pudessem. O que você acha disso? (se o informante perguntar, a pesquisa é do Ibope e foi publicada em janeiro/2010³⁸) Você acha que isso acontece em Campo Grande também? Por quê?
38. Você gosta de morar em Campo Grande? (Se não, em que lugar preferiria morar?) Por quê?
39. O que você acha que caracteriza a cidade (tanto as coisas boas quanto ruins)?
40. O que você acha que caracteriza o campo-grandense (tanto as coisas boas quanto ruins)?
41. Olhando pra mim, você diria que eu sou campo-grandense(a)? Por quê?
42. O que você mais gosta em Campo Grande?
43. O que você não gosta em Campo Grande? (a depender do tópico mencionado pelo informante - violência, trânsito, poluição etc. - procurar explorar mais o assunto. P.ex.: você já foi assaltado? O que aconteceu? (para obter narrativa

³⁸http://www.nossasaopaulo.org.br/portal/arquivos/Pesquisa_IRBEM_Ibope_2010_completa.pdf>. Último acesso em dezembro de 2014.

peçoal) O que o governo deve fazer pra solucionar esse problema? (para obter uma fala mais distanciada)) (explorar o subtópico por aproximadamente 10 min.)

44. Pras pessoas que não vivem em Campo Grande, como você acha que elas imaginam que seja a cidade? Qual é a imagem que as pessoas de fora de Campo Grande têm da cidade?

45. Você falou que já foi pra (X, Y, Z). Quando você foi pra esses lugares, as pessoas percebiam que você era paulistano? (Se sim) como elas percebiam?

46. Quando você conhece alguém, você percebe se a pessoa não é daqui de Campo Grande? a. (Se sim) como você percebe? (Se o informante mencionar o modo de falar, seguir nessa linha e tentar conseguir informações mais precisas. É normal as pessoas não saberem definir o porquê, mas devemos tentar tirar mais informações). b. (Se não) Quando você ouve uma pessoa falando, por exemplo, você percebe que ela é de fora pelo sotaque?

PRODUÇÃO/PERCEPÇÃO/AVALIAÇÃO LINGUÍSTICA (aproximadamente 5 minutos)

47. Já que a gente está falando de sotaque... qual o sotaque do Brasil que você mais gosta? E tem algum que te irrita? Como é que (o gaúcho/o carioca/o mineiro/o caipira etc. – a depender dos sotaques mencionados) fala(m)?

48. Aqui em Campo Grande tem muito (i)migrante, né? De onde vêm a maioria das pessoas? Tem algum bairro específico em que eles (italianos, coreanos, nordestinos etc.) se concentram? Hoje em dia, existem bairros étnicos em Campo Grande?

49. E como é que as pessoas falam em Campo Grande? (evitar usar a palavra “sotaque”)

50. Dentro de Campo Grande, você consegue identificar se a pessoa é de alguma região ou bairro específico da cidade? (Se sim) Você poderia dar alguns exemplos? (Imprimir a lista de palavras, a notícia e o trecho, todos abaixo, para mostrar ao informante)

LISTA DE PALAVRAS

Apresentar ao informante a lista de palavras e dizer: “Agora eu queria te pedir pra ler algumas coisas. Eu tenho aqui uma lista de palavras, e eu queria que você lesse cada uma delas. Você pode dar uma olhada primeiro, antes de começar a ler”. Esperar que o informante leia a lista de palavras. Depois perguntar:

51. A gente estava falando de sotaques agora há pouco... Como você acha que uma pessoa do interior falaria algumas dessas palavras? Tem diferenças?

52. E como um carioca falaria algumas dessas palavras?

53. Tem mais algum sotaque no Brasil que você conhece?

LEITURA DE NOTÍCIA: Dizer ao informante: “Agora eu queria que você lesse esse texto. Você pode ler em voz baixa antes de falar em voz alta”. Esperar que o informante leia a notícia.

LEITURA DE TRECHO: Dizer ao informante: “Por último, queria que você lesse esse trequinho.” Esperar que o informante leia o trecho.

54. (Sobre /e/ nasal) O que você acha desse modo de falar: “Você tá entendendo o que eu tô dizendo?” (com “en” ditongado e exagerado). Se a pessoa manifestar uma atitude negativa, perguntar: Como você acha que deveria ser? O que tem de errado aí?

55. Quem você acha que fala assim? (se falar “campo-grandenses, perguntar: “você acha que todos os campo-grandense falam assim ou é uma coisa de uma região mais específica na cidade?”)

56. Você fala desse modo?

57. (Sobre concordância nominal) E o que você acha de “Me vê dois pastel e um chopps?” (Repetir as perguntas 52-53)

58. (Sobre /r/ retroflexo) E o que você acha de “A porta tá aberta.” (com retroflexo exagerado). (Repetir as perguntas 52-53).

59. E tem algum modo de falar que você acha assim bem campo-grandense?

Palavras

alma
amargo
animação
argola
atitude
barqueiro
biscoito
cacique
carteiro
cerca
chácara
circo
cisne
curto
defender
discoteca
elefante
enchente
entender
entretenimento
erguer
fazenda
felicidade
firme
fornalha
furgão
fusquinha
geleira
gérmen de trigo
gordo
gula

Hércules
hilário
irmã
justiça
lento
mortadela
Marba
mosca
necessidade
orca
Ordem e progresso
orgânico
órgão
ostracismo
penteado
pertencimento
perto
porto
presente
rapidez
riqueza
sabor de menta
soberba
trabalho
turco
urgente
utilidade
vulto
zebra

Notícia³⁹

10/03/2012

Com um calor acima de 30 °C, Mato Grosso do Sul registra pior índice de umidade do ano

Com um calor acima de cinquenta graus Celsius, Mato Grosso do Sul registra o pior índice de umidade do ar. Com um fim de verão marcado por pouca chuva e altas temperaturas, Mato Grosso do Sul registrou ontem a menor umidade relativa do ar deste ano. A umidade caiu dezesseis por cento, o índice é comparável ao clima de deserto, onde, durante o dia, a umidade chega a treze por cento, com temperatura na casa dos cinquenta graus.

Ontem, os índices abaixo do sessenta por cento recomendados pela OMS (Organização Mundial de Saúde) foram registrados em boa parte dos municípios. Essa situação foi verificada em toda a região central de Mato Grosso do Sul, incluindo Campo Grande, Maracaju, Sidrolândia e Terenos, explica o meteorologista da Anhanguera Uniderp, Natálio Abrão: “neste período, também aumenta a intensidade de raios ultravioleta, por isso devemos tomar alguns cuidados como uso de filtro solar aumento da ingestão de água e umidificação de ambientes com aparelhos e toalhas molhadas”.

³⁹ Adaptada de: <http://www.campograndenews.com.br/meio-ambiente/com-calor-acima-de-30c-ms-registrou-pior-indice-de-umidade-do-ano> (último acesso em janeiro de 2016).

Trecho

Está fazendo muito calor! Sexta-feira estava tão quente, que dormi na área de serviço lá de casa. Ó para você ver: mesmo usando filtro solar, minha pele está toda queimada e manchada, nossa está parecendo um deserto aqui. No meu trabalho, nós tivemos que deixar uma garrafa de água bem gelada na mesa para tomar tereré durante a tarde.

Nossa, que clima insuportável! Você escancara as janelas e continua um forno, fica sem ter o que fazer. A gente liga o ar-condicionado para refrescar, mas fica uma sensação estranha, parece que resseca tudo. Não aguento mais. Quando tenho que pegar o ônibus, aí sim piora, sem falar que essa tarifa do transporte público só aumenta, sem melhorar nada. Temos que andar nesses ônibus lotados e quentes. Os políticos falam e falam, mas eles tinham que fazer alguma coisa urgente. Você está entendendo o que eu estou dizendo?

Anexo II

Declaração de livre consentimento

UNIVERSIDADE DE SÃO PAULO
FACULDADE DE FILOSOFIA, LETRAS E CIÊNCIAS HUMANAS
Av. Prof. Luciano Gualberto, 403 - Cx. Postal: 26097 - CEP: 01060-970 - São Paulo
Fone: (011) 3091.4298 - e-mail: fl@usp.br

Projeto de Pesquisa: *Aí, daí e então* em Campo Grande e São Paulo: análise sociofuncionalista no domínio da causalidade

Contato: marilia.vieira@usp.br

Declaração

Declaro ter consentido em ter gravada minha conversa com a aluna universitária Marília Vieira, como parte do projeto acima identificado, em execução pelo Grupo de Estudos e Pesquisa em Sociolinguística da USP. Estou ciente de que a gravação será transcrita e disponibilizada na Internet para o desenvolvimento de estudos sociolinguísticos nas cidades de Campo Grande e de São Paulo. Entendo que essa gravação é idônea e que meus dados pessoais não serão tornados públicos na divulgação de resultados da pesquisa.

Campo Grande, _____ de _____ de _____.

Nome do informante: _____

Assinatura: _____

Anexo III

Ficha do informante e da gravação

Nome do Informante:

Pseudônimo:

Data de Nascimento:

Idade:

Endereço:

Telefones para contato:

E-mail:

Ocupação:

Pessoas com quem mora:

Documentador:

Arquivo de gravação e transcrição:

Bairros em que já morou em Campo Grande, em ordem cronológica, e tempo aproximado (em anos) de quanto tempo morou em cada bairro:

P.ex.: Moreninhas – 10 anos

Número de irmãos:

() mais velhos

() mais novos

Nome, idade e escolaridade dos irmãos:

Geração na cidade: () 0 – pais não nascidos em Campo Grande

() 1 – pai ou mãe nascido em Campo Grande

() 2 – pai e mãe nascidos em Campo Grande

() 3 – um avô ou uma avó nascido em Campo Grande

() 4 – dois ou mais avós nascidos em Campo Grande

Nome do Pai:

Idade do Pai:

Grau de Escolaridade do Pai:

Ocupação do pai:

Naturalidade: () Campo Grande () outra: _____

Nome da Mãe:

Idade da Mãe:

Grau de Escolaridade da Mãe:

Ocupação da mãe:

Naturalidade de: () em Campo Grande () outra: _____

Escolas em que estudou e tipo de escola (pública municipal (M), pública estadual (E) ou particular (P)):

P.ex.: Fundamental 1: Escola Estadual President Roosevelt (E)

Fundamental 1: _____

Fundamental 2: _____

Ensino Médio: _____

Faculdade:
(concluída ou não) _____

Data da Entrevista:

Horário:

Local da Entrevista e breve descrição do local:

Presença e atuação de terceiros:

Como o documentador conheceu o informante:

Grau de relação entre informante e documentador:

() 1 – Bastante próximo. O entrevistado faz parte do meu círculo imediato de amigos/familiares e conversamos frequentemente.

() 2 – Próximo. Conversamos frequentemente, mas o entrevistado não faz parte do meu círculo imediato de amigos/familiares.

- () 3 – Próximo, mas não conversamos frequentemente.
- () 4 – Neutro. Ele é meu conhecido, mas não nos falamos com frequência.
- () 5 – Distante. Não o conhecia anteriormente e praticamente só conversamos na ocasião da entrevista.

Outras informações:

Anexo III

Questionário Socioeconômico

Por favor, assinale a opção que melhor descreve sua situação:

**1) Em que tipo de imóvel
você vive?**

☐ Casa

☐ Apartamento

☐ Outro: _____

☐ 1

☐ 2

☐ 3 ou mais

**5) Quantas pessoas moram em
sua casa?**

**2) Há quantos anos o imóvel
foi construído?**

☐ Há menos de 10 anos

☐ De 10 a 20 anos

☐ Há mais de 20 anos

☐ 1

☐ 2

☐ 3

☐ 4

**3) Você é proprietário do imóvel
em que vive?**

☐ Sim, é um imóvel próprio e
quitado.

☐ Sim, é um imóvel financiado.

☐ Não, é um imóvel alugado.

☐ 5

☐ 6

**6) Quantas pessoas que vivem
em sua casa contribuem para
a renda familiar?**

☐ 1

☐ 2

☐ 3

**4) Quantos quartos tem a
sua casa?**

☐ 0

☐ 4

☐ 5

7) Qual é a sua renda média mensal?

☐ Até R\$ 600,00

☐ De R\$ 600,00 a R\$ 1.200,00

☐ De R\$ 1.200,00 a R\$ 2.400,00

☐ De R\$ 2.400,00 a R\$ 6.000,00

☐ De R\$ 6.000,00 a R\$ 12.000,00

☐ Acima de R\$ 12.000,00

8) Qual é a renda média mensal de sua família?

☐ Até R\$ 600,00

☐ De R\$ 600,00 a R\$ 1.200,00

☐ De R\$ 1.200,00 a R\$ 2.400,00

☐ De R\$ 2.400,00 a R\$ 6.000,00

☐ De R\$ 6.000,00 a R\$ 12.000,00

☐ Acima de R\$ 12.000,00