UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE LETRAS
CURSO DE PÓS-GRADUAÇÃO EM LETRAS
MESTRADO EM AQUISIÇÃO DA LINGUAGEM

A CONCORDÂNCIA VERBAL NA FALA DE CRIANÇAS DE PORTO ALEGRE

SIMONE MENDONÇA SOARES

Dissertação apresentada como requisito parcial à obtenção do título de Mestre em Aquisição da Linguagem.

Orientadora: Profª Dra. Luciene Juliano Simões

Porto Alegre, janeiro de 2006.

AGRADECIMENTOS

O incentivo, a colaboração e a amizade de algumas pessoas foram fundamentais para a realização deste trabalho.

Agradeço, inicialmente, à minha orientadora, Luciene Juliano Simões, a quem devo boa parte de minha formação acadêmica, pela qualidade da orientação, tanto em relação à pesquisa quanto a outros processos envolvidos na realização de uma dissertação, e também pela estimada amizade.

Ao professor Luís Amaral pela disponibilidade e valiosa ajuda com as rodadas estatísticas.

À professora Ana Maria Zilles, por suas contribuições em várias etapas do trabalho.

Às minhas colegas Elaine Capellari, Cristiane dos Santos e Hires Batista, pela ajuda e amizade.

Ao Artur e seus pais, que me abriram sua casa e sua vida durante um tempo tão longo e com tanto carinho.

Ao Programa de Pós-Graduação em Letras da UFRGS, pela oportunidade de realizar um curso qualificado, e à CAPES pela bolsa que me foi concedida.

Por fim, aos meus três homens: Thomas, pelo amor imenso e compreensão ainda maior; Vinicius, pelo amor incondicional; e ao nenê que vai nascer, pela alegria de sua companhia irrequieta na fase mais dura do trabalho.

SUMÁRIO

Lista de Tabelas..... 5
Lista de quadros gráficos 6
Resumo 7
Résumé 8

Introdução 9

1. A Concordância Verbal 13
1.1 O mecanismo da CV 13
1.2 CV e tópico 15

2. A Socioligüística e a CV 32
2.1 A Sociolingüística laboviana 32
2.2 A CV no PB 34
2.3 A CV em Porto Alegre 41

3. A Aquisição da Linguagem e a CV 54
3.1 Os modelos de aquisição da linguagem 54
3.2 As etapas do desenvolvimento da linguagem 57
3.3 A natureza da gramática inicial 61
3.4 Linguagem e socialização 71

4. Metodologia 77
4.1 Os dados 77
4.2 O envelope de variação 84
4.2.1 A variável dependente 85
4.2.2 As variáveis lingüísticas 86
4.2.3 As variáveis sociais 96
4.3 O modelo analítico 97

5. Análise e Discussão dos Resultados 99

5.1 Análise preliminar 99
5.2 Resultados gerais de presença/ausência de marca 104
5.2.1 Tipo de sujeito 110
5.2.2 Tempo verbal 113
5.2.3 Posição do sujeito em relação ao verbo 115
5.2.4 Tonicidade 116
5.2.5 Tipo de verbo 118
5.2.6 Faixa etária 119
5.2.7 Variáveis não selecionadas 123
5.3 A concordância de primeira pessoa do plural 125
5.3.1 A forma *a gente* 125
5.3.2 Presença versus ausência da DNP-P4 127
5.3.2.1 Tipo de verbo 130
5.3.2.2 Tipo de sujeito 131
5.3.2.3 Tipo de coleta 132
5.3.3 Padrão *versus* não padrão: o apagamento do /s/ 133
5.3.3.1 Faixa etária 134
5.3.3.2 Tipo de coleta 138
5.4 A concordância de terceira pessoa do plural 138
5.4.1 Posição do sujeito em relação ao verbo 140
5.4.2 Tonicidade 141
5.4.3 Saliência fônica 142
5.4.4 Tipo de verbo 146
5.4.5 Tipo de sujeito 147
5.4.6 Gênero 150
5.4.7 Faixa etária 151

Considerações Finais 155

Referências Bibliográficas 160

LISTA DE TABELAS

Tabela 1 – Resultados gerais da rodada preliminar 101
Tabela 2 – Resultados gerais do total de dados relativamente às cinco pessoas gramaticais consideradas 105
Tabela 3 – Índices de aplicação da marca de concordância em relação à variável tipo de sujeito 111
Tabela 4 – Índices de aplicação da marca de concordância em relação à variável tempo verbal 114
Tabela 5 – Índices de aplicação da marca de concordância e relação à variável posição do sujeito em relação ao verbo 116
Tabela 6 – Índices de aplicação de marca de concordância em relação à variável tonicidade 118
Tabela 7 – Índices de aplicação de marca de concordância em relação à variável tipo de verbo 119
Tabela 8 – Índices de aplicação de marca de concordância em relação à variável faixa etária 119
Tabela 9 – Relação entre a aplicação da DNP-P4 e a variável tipo de verbo 130
Tabela 10 – Relação entre a aplicação da DNP-P4 e a variável tipo de sujeito 132
Tabela 11 – Relação entre a aplicação da DNP-P4 e a variável tipo de coleta 133
Tabela 12 – Relação entre a aplicação da DNP-P4 padrão e a faixa etária 134
Tabela 13 – Relação entre a aplicação da DNP-P4 padrão e o tipo de coleta 138
Tabela 14 – Relação entre a aplicação da DNP-P6 e a variável posição do sujeito 140
Tabela 15 – Relação entre a aplicação da DNP-P6 e a variável tonicidade 141
Tabela 16 – Relação entre a aplicação da DNP-P6 e a variável saliência fônica 142
Tabela 17 – Relação entre a aplicação da DNP-P6 e a variável tipo de verbo 146
Tabela 18 – Relação entre a aplicação da DNP-P6 e a variável tipo de sujeito 147
Tabela 19 – Relação entre a aplicação da DNP-P6 e a variável gênero 150
Tabela 20 – Relação entre a aplicação da DNP-P6 e a variável faixa etária 152

LISTA DE QUADROS E GRÁFICOS

Quadro 1 – Perfil sócio-econômico dos informantes do projeto DELICRI 79
Quadro 2 – Informantes da coleta longitudinal do projeto DELICRI, idade e categoria sócio-econômica 81
Quadro 3 – Composição da amostra relativamente a idades e faixa etária 84
Gráfico 1 – Aplicação da marca de concordância em relação à faixa etária 121
Gráfico 2 – Distribuição geral das formas da DNP-P4 126
Gráfico 3 – Presença *versus* ausência da DNP-P4 128
Gráfico 4 – Presença *versus* ausência da DNP-P6 139
Gráfico 5 – Atuação dos níveis de saliência fônica na aplicação da DNP-P6 144
Gráfico 6 – Atuação em peso relativo dos níveis de saliência fônica na aplicação da DNP-P6 em diferentes pesquisas 145
Gráfico 7 – Atuação da faixa etária na aplicação da DNP-P6 em termos de peso relativo 153

RESUMO

Este trabalho se propõe investigar o sistema de concordância verbal na fala de nove crianças, quatro meninos e cinco meninas, entre 2;5 e 8 anos de idade, da comunidade de Porto Alegre. A análise é alicerçada metodologicamente na Sociolingüística Quantitativa, tendo descrições da fala adulta como parâmetro interpretativo dos dados, tanto em relação a aspectos lingüísticos quanto sociais.

O objetivo é estabelecer um quadro descritivo geral do sistema flexional de verbos referente a primeira, segunda e terceira pessoas do singular e primeira e terceira do plural por um lado, e por outro, examinar mais detidamente a concordância variável de primeira e terceira pessoas do plural. Esse exame mais detalhado contou com uma análise da desinência de primeira pessoa do plural dividida em dois momentos. O primeiro opondo a presença à ausência da marca, mesmo procedimento aplicado à desinência de terceira pessoa do plural. O segundo opondo a marca padrão, com manutenção do morfema de plural /s/, à marca não padrão, com apagamento do /s/.

De modo geral, os resultados obtidos mostram que o comportamento lingüístico da criança é precoce em relação a tal sistema. Além disso, está sujeito às mesmas regras estabelecidas pelas diversas forças simultâneas que atuam na fala do adulto. Tal panorama levanta a possibilidade de se compreender o sistema de concordância variável como operando tanto sob domínio sintático quanto social.

RÉSUMÉ

Ce travail se propose d'étudier le système de concordance verbale dans la parole de neuf enfants de la communauté de Porto Alegre: quatre petits garçons et cinq petites filles, de 2 ans et demi à 8 ans. Du point de vue méthodologique, l'analyse est basée sur la Sociolinguistique Quantitative. Les descriptions de la parole adulte sont le paramètre interprétatif des donnés, soit par rapport à des aspects linguistiques, soit par rapport à des aspects sociaux.

L'objectif est d'établir, d'un côté, un tableau descriptif général du système de flexion des verbes de la première, la deuxième et la troisième personne du singulier, ainsi que de la première et la troisième du pluriel, et d'un autre côté, d'examiner la concordance variable de la première et de la troisième personne du pluriel. Cette analyse se déroule en deux moments : d'abord, en opposant la présence à l'absence de marque du pluriel ; le même procédé appliqué à la désinence de la troisième personne du pluriel. Ensuite, en opposant la marque modèle avec le maintien du morphème de pluriel /s/ à la marque non-modèle, avec l'effacement du /s/.

De façon générale, les résultats obtenus montrent que le comportement de l'enfant est précoce par rapport à un tel système. En plus, il est soumis aux mêmes règles établies par les diverses forces simultanées qui agissent dans la parole de l'adulte. Un telle évidence peut permettre de comprendre le système de concordance variable opérant dans le domaine syntaxique aussi bien que dans le domaine social.

INTRODUÇÃO

Este trabalho tem como objetivo estabelecer um quadro descritivo do sistema de concordância verbal (daqui em diante CV) variável na fala de crianças da comunidade de Porto Alegre.

Nesta perspectiva, dou continuidade a um estudo anterior (Soares, 2001), no qual, sob orientação da professora Luciene Simões, examinei a concordância verbal variável de primeira pessoa do singular e primeira pessoa do plural em dados de duas crianças adquirindo o Português do Brasil como língua materna, um menino, entre 2;9 e 5;4, e uma menina entre 4;3 e 8;5[1] de idade.

No referido estudo os resultados mostraram, em relação a P1[2], no menino, num total de 640 dados, 633 casos de presença de marca de concordância (98,9%), sendo

[1] O primeiro número indica a idade em anos, e o segundo, em meses. Assim, 8;5 indica a idade de 8 anos e 5 meses.
[2] Daqui em diante, P1, P2 e P3 para as três pessoas do singular respectivamente, e P4, P5 e P6 para as três pessoas do plural respectivamente.

que, dentre esses, 17 ocorrências eram de supergeneralização (como *eu sabo*, por exemplo), e 7 eram de desinência não marcada. Na menina, 171/172 casos de marca de concordância e 1 caso de não marcação. Ocorrências de supergeneralização e verbo com a desinência não marcada de terceira pessoa do singular relativos à primeira pessoa do singular estão descritos na literatura psicolingüística (Perroni Simões e Stoel-Gammon, 1977; Perroni Simões, 1976), sobretudo para crianças em torno de dois anos de idade.

Naquele trabalho ainda, em relação a P4, foram observadas três formas variáveis: a desinência padrão (*–mos*), a desinência não padrão (*-mo*) e a ausência de marca (*Ø*) em relação a três tipos de sujeito: o pronome reto (*nós*), seu nulo e o SN pleno, além de considerada a forma inovadora *a gente* e seu nulo. Não foi possível observar a concordância variável nos dados da menina devido à ausência de casos do pronome reto. Assim sendo, decidi pela continuidade do estudo da concordância verbal variável em dados de crianças a fim de obter um quadro descritivo mais detalhado desse sistema lingüístico. Além disso, o sistema de CV é bastante contemplado no âmbito da pesquisa sociolingüística no Brasil no que concerne a dados de adultos. Entretanto, sobre dados de criança, ainda pouco se sabe, o que justifica um estudo tal como empreendo aqui. Por outro lado, no âmbito acadêmico, existe uma carência de estudos que promovam a associação entre as áreas de Psicolingüística e de Sociolingüística, associação essa que diz respeito não só a cada uma das áreas individualmente, mas também, de forma mais ampla, à discussão que promove a construção de uma teoria lingüística no âmbito da pesquisa brasileira.

Assim, o objetivo aqui é mapear as desinências de CV nos termos da análise sociolingüística quantitativa de modo a indicar caminhos em direção a uma nova abordagem da fala da criança.

O trabalho é composto por sete partes. A Introdução traz as justificativas e motivações, além da estrutura organizacional geral.

O primeiro capítulo trata do sistema de CV. Inicialmente, apresento uma rápida descrição do mecanismo de CV em termos estruturais. Na segunda sessão, discorro sobre a teoria givoniana de CV, relacionada ao tópico do discurso, discutindo alguns estudos sobre CV no português brasileiro que têm por base a teoria lingüística de Givón. Além disso, abordo a hipótese inacusativa em sua relação com o sistema de CV.

O segundo capítulo é dedicado à Sociolingüística laboviana. Na primeira seção trato de seus principais fundamentos. Na segunda, apresento uma seleção de trabalhos, todos de cunho laboviano, examinando o sistema de CV em diferentes comunidades do português do Brasil. Na seção seguinte, faço uma compilação similar de trabalhos sobre a comunidade de Porto Alegre.

O terceiro capítulo é dedicado à pesquisa aquisicionista, com ênfase no sistema flexional de CV. A primeira seção trata de dois modelos teóricos de tratamento da aquisição de língua materna. A segunda seção trata das etapas do desenvolvimento, sobretudo aquelas relativas ao sistema de concordância. A terceira aborda a natureza da gramática no período inicial de aquisição da linguagem. A quarta considera as relações existentes entre linguagem e socialização.

No quarto capítulo descrevo o procedimento analítico de tratamento dos dados, bem como a caracterização da amostra submetida à análise, e apresento também as variáveis consideradas no estudo.

O quinto capítulo é dedicado à análise e discussão dos resultados.

Por fim, na última parte, apresento uma síntese dos principais resultados alcançados, traçando algumas considerações finais sobre a pesquisa.

1. A CONCORDÂNCIA VERBAL

Nesta seção abordarei, inicialmente, o mecanismo de CV na gramática da língua portuguesa. A seguir, apresentarei e discutirei a concepção givoniana de concordância, relacionada à função de tópico na sentença.

1.1 O mecanismo da CV

A concordância é um princípio que estabelece a harmonização de um termo flexionável a outro ao qual se refere no discurso. Segundo Said Ali (2001:205),

> *"Consiste a concordância em dar a certas palavras as formas de gênero, número ou pessoa correspondentes à palavra a que no discurso se referem. (...) Desde que de um vocábulo se oferecem várias formas à escolha, e o dito vocábulo vem determinar, esclarecer ou informar alguma coisa a respeito do outro, escolhemos naturalmente aquela forma que se harmonizar com estoutro termo".*

A concordância verbal, então, opera nos elementos flexionáveis do verbo em relação ao sujeito da oração. No tratamento normativo que a gramática tradicional dá à CV, trata-se do sistema que estabelece as relações número-pessoais entre o sujeito e o verbo, ou seja, o sujeito que se faz presente, em número e pessoa, na forma verbal. Cunha (1972: 466) diz que

> *"A solidariedade entre o verbo e o sujeito, que ele faz viver no tempo, exterioriza-se na concordância, isto é, na variabilidade do verbo para conformar-se ao número e à pessoa do sujeito"*.

Em português, como em outras línguas, a estrutura morfológica básica do vocábulo verbal contém um *locus* relativo à CV. Na primeira posição da estrutura temos um radical, responsável pela carga semântica, associado a uma vogal temática. Dois morfemas compõem, na seqüência, as desinências flexionais (DF). Elas distinguem duas noções gramaticais distintas, uma estritamente verbal – modo-temporal (DMT) – e outra nominal – número-pessoal (DNP). Temos, assim, a regra geral da estrutura do vocábulo verbal em português tal como nos apresenta Câmara Jr. (1982): o tema, formado por radical e vogal temática, associado à desinência flexional, formada pelas desinências modo-temporal e número-pessoal.

T (R+VT) + DF (DMT + DNP)[3]

A desinência de número e pessoa, que responde pela concordância com o sujeito, ocupa, como vemos, a última posição nessa estrutura. Tal como qualquer dos outros constituintes, excetuando-se o radical, pode assumir a forma zero (∅). É importante

[3] Optei pelas siglas DF, DMT e DNP ao invés de, respectivamente, SF, SMT e SNP, tal como aparece em Câmara Jr, em função da maior freqüência daquelas na literatura.

salientar ainda que a relação com o sujeito na forma da desinência faz com que o vocábulo verbal contenha em si as distinções discursivas que opõem falante (primeira pessoa) e ouvinte (segunda pessoa), e estes a uma terceira pessoa, tanto no singular como no plural, conforme aponta Mattoso Câmara (1988).

Segundo as gramáticas de cunho normativo, como a de Cunha (1972), por exemplo, a aplicação da regra de CV é categórica, isto é, obrigatória e invariável, e qualquer construção que foge a esse uso é considerada erro. Associada a esse quadro, como veremos mais detidamente no próximo capítulo, existe a forte estigmatização das formas não marcadas.

1.2 CV e tópico

A teoria lingüística de Talmy Givón contempla de modo particular o sistema de CV. Para esse autor (Givón, 1976, 1984), é o tópico que rege a concordância. Além disso, aborda esse sistema relativamente ao processamento do discurso, sua continuidade/descontinuidade e os elementos aí envolvidos. Apresentarei e discutirei, a seguir, essa perspectiva.

Tratando sobre a concordância gramatical, tanto em termos sincrônicos quanto diacrônicos, Givón (1976) propõe a derrubada de dois "mitos" sobre a concordância[4]. O primeiro identificaria concordância gramatical com concordância verbo-

[4] O termo "mito" é empregado pelo próprio autor: "*In this paper I would like to lay to rest two myths about grammatical agreement*", não refletindo minha adesão a ele.

sujeito. Para o autor, é o tópico, e não o sujeito, o elemento controlador da concordância. O segundo mito seria a aceitação de que concordância e pronominalização são dois processos distintos: para Givón, são fundamentalmente um único e mesmo fenômeno.

Partindo de uma perspectiva diacrônica, o autor afirma que a concordância gramatical surge[5] via construções de mudança de tópico em que o SN topicalizado é correferencial a um dos argumentos do verbo. O resultado disso é que o nome correferente dentro da sentença é substituído por um pronome anafórico, sendo que este pronome atestaria um caso de concordância de tópico. Entretanto, quando acontece a reanálise desta estrutura, sendo o constituinte tópico identificado como sujeito ou objeto de uma sentença neutra, não topicalizada, forçosamente a concordância também é reanalisada como sendo de sujeito ou de objeto (e não mais de tópico). O fato de línguas que não usam pronome anafórico em construções de mudança de tópico, como Mandarin e Japonês, não desenvolverem concordância verbo-sujeito ou verbo-objeto reforçaria a teoria.

Em termos sincrônicos, os antigos pronomes seriam reanalisados como morfemas de concordância, continuando a desempenhar sua função anafórica. Dessa forma, línguas com um paradigma viável de concordância verbo-sujeito podem apagar o sintagma nominal sujeito sem a necessidade de substituí-lo por um pronome independente.

Discursivamente, os pronomes tônicos tenderiam a ser enfáticos, contrastivos, enquanto que os átonos e a concordância tenderiam a ser não enfáticos,

[5] Sobre esse ponto, Sérgio Menuzzi me chamou atenção para uma contradição na afirmação de Givón, já que se o sistema de concordância "surge" (o trecho do texto de Givón é o seguinte: *"I will show that agreement arises via topic-shifting constructions (...)"*) a partir das construções de mudança de tópico, é porque não são a mesma coisa, e portanto, não pode se tratar de um único e mesmo fenômeno. Essa é uma discusão, entretanto, que foge ao âmbito deste trabalho.

marcadores de continuidade e anafóricos (Givón, 1984). Assim, a concordância gramatical estaria, nessa teoria, relacionada ao domínio funcional dos pronomes na cadeia de referência, ou seja, atuando na *identificação* ou *continuidade* de tópicos.

O autor afirma também que a probabilidade de concordância com o verbo é governada pela hierarquia universal da topicalidade, que prediz, em relações hierárquicas binárias, que tipo de sintagmas nominais argumentos seriam tópico em sentenças ou em construções de mudança de tópico. As relações hierárquicas são as seguintes:

(1) a) humano > não-humano
b) definido > indefinido
c) participante mais envolvido > participante menos envolvido
d) 1ª pessoa > 2ª pessoa > 3ª pessoa

Em (1), a) diz respeito à tendência de seres humanos falarem mais sobre seres humanos; b) mostra que os SNs definidos, veiculando informação dada, tendem a ser constituídos como tópico, enquanto que os indefinidos, que veiculam informação nova, tendem a ser comentário; c) traduz a hierarquia da semântica de caso, em que agente > dativo > acusativo, indicando o altíssimo pareamento do *tópico* com a função semântica *agente*, agregados na função sintática *sujeito*; e, finalmente, d) reflete o caráter egocêntrico do discurso, no qual o falante tende a ser o ponto de referência universal e o argumento mais altamente pressuposto.

Sobre essa predição tipológica de SNs, Givón (1984) afirma que ela pode ser dividida em dois tipos relacionados: sincronicamente uma língua vai mostrar provavelmente mais concordância obrigatória com um sintagma nominal de tipo mais alto

que mais baixo na escala; diacronicamente a concordância gramatical vai se desenvolver primeiro em SNs de tipo mais alto na escala.

Nesse contexto teórico, discutirei o seguinte aspecto: se no PB a concordância verbo-sujeito é regida por um elemento funcional – o tópico da oração. Para tanto, examinarei dois trabalhos funcionalistas – Decat (1983) e Camacho (1993) – que se debruçam sobre a assertiva de que é o tópico, e não o sujeito, o controlador da CV em determinados tipos de construções no PB.

No primeiro estudo, Decat (1983) dedica-se à analise da CV em estruturas de sujeito posposto, como em Decat (1981), mas também relacionadas a processos de topicalização. Antes de apresentar o artigo destaco uma crítica feita por Camacho (1993) relativamente a esse trabalho, que o autor chama de excesso de generalização: o *corpus* em análise é formado por uma variedade falada e uma variedade escrita e o tratamento analítico e explicativo é o mesmo para ambos. Camacho diz que: "*O que é criticável, nesse aspecto da questão, é o tratamento igual dado às variedades culta e popular e às modalidades oral e escrita. Parece não ser possível englobar todos esses fatores variáveis numa mesma e única gramática*" (Camacho, 1993:109). De fato, do ponto de vista sociolingüístico, é problemático tratar diferentes variedades da língua sem controlar fatores que possivelmente provoquem essas diferenças. Contudo, não se tratam de diferentes gramáticas, mas de uma única gramática, variável[6].

Quanto ao artigo, a hipótese da autora é de que nas sentenças de sujeito posposto a concordância é ausente, isto é, não apresenta marcas, por se tratarem de

[6] Agradeço à minha orientadora por ter me chamado atenção a esse aspecto.

estruturas tópico-comentário formadas unicamente pelo comentário, ou seja, é o tópico o controlador da concordância e sua ausência afeta a marcação da concordância.

Para fundamentar seu ponto de vista, a autora parte do tratamento que a gramática tradicional dá à relação entre a CV e o SN sujeito, já que a CV seria usada como critério para a identificação do sujeito da sentença. Sua constatação, entretanto, é de que nos casos de posposição do sujeito a concordância não ocorre. Ela discute sentenças tais como:

(2) Saiu as notas de matemática
(3) Foi feito duas viagens
(4) Essas roupas não saem tinta[7]

Segundo ela, "*a aplicação da regra nem sempre parece ser determinada ou controlada pelo SN sujeito, mas por outros SNs que estejam próximos ao verbo"* (Decat, 1983:15).

É preciso salientar aqui que em (2) e (3) há sujeito posposto, mas em (4), não há razão para não considerar "*Essas roupas"* o sujeito da sentença. A análise que ela propõe para (4) é a seguinte:

> *"o verbo não está concordando com o SN tinta que, nos termos da Gramática Tradicional, seria o sujeito da sentença, o qual, por sua vez, iria determinar a CV, já que o sintagma essas roupas seria tratado como adverbial, originado de um sintagma dessas roupas, através da supressão da preposição.* (Decat, 1983:15)

[7] Extraídos de Decat (1983:15), exemplos (11) a (13) no original.

Para o exame de outros tipos de sentenças, são apresentados quatro grupos de exemplos com sujeito posposto. O Grupo I é composto por sentenças com verbos existenciais e de movimento, todos intransitivos, como vemos em (5) e (6) a seguir:

(5) Não está aqui os melhores exemplos [8]
(6) ... ficou as camadas mais fortes

O Grupo II constitui-se ou de sentenças na forma passiva, como em (7) e (8) abaixo, ou de sentenças em que o verbo (*ser*, *estar*, *ficar*, etc.) é seguido de um adjetivo ou de um particípio passivo, como em (9) e (10) a seguir.

(7) seja tirado uma comissão
(8) Deve ser inserido... uma frase...
(9) tá claro nessa assembléia duas posições
(10) tá previsto cinco meses

O Grupo III, segundo a autora, apresenta exemplos que contrariam o que é postulado pela gramática tradicional para sujeito posposto e composto:

(11) ... entrou eu, o Sérgio e o Túlio...
(12) ... nós fomos no Rio Vermelho (...) foi eu, Túlio e César...

E, finalmente, o Grupo IV mostra o fenômeno de não concordância em sentenças com verbos transitivos, o que seria bastante interessante para o argumento dela, dadas as restrições de VS para o PB. Entretanto, os exemplos não são de transitivos típicos, mas sim, de transitivos não agentivos, tais como:

[8] Os exemplos (2) a (15) foram extraídos de Decat (1983:18). Todos os exemplos utilizados neste trabalho foram copiados tal como estão no texto original.

(13) me chateou certos votos
(14) Dá muito trabalho essas coisas
(15) "Começou a fazer o mesmo efeito os remédios"

A autora salienta que há casos, como (15), em que o sujeito posposto é do tipo "*afterthought*", ou seja, o falante repete o sujeito no final da sentença lembrando que faltou informação [9].

Em relação a esses grupos de sentenças, a autora afirma que eles "*constituem evidência de que a CV nem sempre se dá com o SN sujeito posposto*" (Decat, 1983:19).

Na seqüência da argumentação, a autora considera outros dois grupos de sentenças, apresentados abaixo:

Grupo A:

(16) Essa cerveja eu acho muito boa
(17) Esses brinquedos os meninos gostam muito
(18) Meus móveis esse ano eu vou arrumar

Grupo B:

(19) Minhas gavetas não cabem mais nada
(20) As minhas canetas acabaram a tinta
(21) Os meninos cresceram a barba[10]

Conforme a autora, os SNs sublinhados são tópicos das sentenças e a diferença entre os dois grupos é que nas sentenças do primeiro houve somente topicalização e, no segundo, além de topicalização, foi aplicada posposição do sujeito. No Grupo A, a

[9] A autora cita Hyman na explicação do termo; ele aparece também em Givón (1976:154).
[10] Exemplos extraídos de Decat (1983:20-1).

concordância foi aplicada com os SNs que exercem a função de sujeito. No Grupo B, entretanto, a concordância se deu não com o sujeito, mas com o SN topicalizado. A análise seria aquela que apresento em (22).

(22) a. Essa cerveja eu acho muito boa.
tópico *suj.* *CV com suj.*

b. Os meninos cresceram a barba.
tópico *CV com tópico* *suj. posposto*

A autora conclui que *"estamos, assim, diante de um impasse: ora a CV se dá com o sujeito, ora se dá com o tópico"* (Decat, 1983:21).

Partindo do argumento de Givón (1976) apresentado acima, a autora conclui que a concordância verbal em português é controlada, normalmente, pelo tópico. Para as sentenças do Grupo A (exemplos (16) a (18) acima), em que a gramática tradicional considera como de "duplo sujeito", Decat (1983) sugere que uma maior generalização é alcançada se for postulado que as sentenças são construídas com dois tópicos, em níveis diferentes. Nesses casos a concordância se dará com o tópico imediatamente precedente ao verbo.

Cabe aqui ressaltar que a posição de Pontes (1987) sobre o assunto é distinta daquelas de Camacho (1993) e Decat (1983). Segundo a autora, um dos argumentos que indicariam ser o português uma língua de tópico (em que a sentença é mais bem descrita em termos de tópico e comentário que de sujeito e predicado) é justamente o fato de o tópico prescindir de relações selecionais com o verbo, isto é, o sujeito tem relação direta com o

verbo através da seleção de concordância, enquanto que o tópico não apresenta este tipo de seleção. Ela diz que "*não há caso de concordância de verbo com o tópico, mas apenas de verbo com o sujeito"* (Pontes, 1987:19) e aponta exemplos que, segundo ela, mostram isso:

(23) Isso aí eu tenho dúvida.
(24) Banana ouro – é a única banana que eu gosto.
(25) Qualquer elemento você pode fazer isso. (com...)
(26) Essas regras, sejam da base, sejam da ES, elas são construídas...
(27) Essa competência ela é de natureza mental.[11]

Em relação às sentenças acima, pode-se dizer que (24), (26) e (27) são irrelevantes para o exame da regência da concordância, já que, em (24), o tópico é o sujeito da oração principal, e em (26) e (27) o tópico é co-referente com o sujeito. Entretanto, (23) e (25) mostram perfeitamente o ponto. Elas podem ser adaptadas, respectivamente, para (28) abaixo:

(28) a. Isso aí nós temos dúvida.
b. (Com) qualquer elemento vocês podem fazer isso.

Voltando ao artigo de Decat, o exame dos exemplos discutidos pela autora leva a pensar em dois aspectos relacionados. O primeiro e mais aparente diz respeito ao tipo de estrutura e conseqüentemente ao tipo de verbo que ela está considerando. O segundo traduz as propriedades dos SNs que ela considera o sujeito das respectivas sentenças, já que se apóia na gramática tradicional.

[11] Exemplos extraídos de Pontes (1987:19).

Todas as sentenças consideradas no estudo, devido ao fato de serem estruturas de sujeito posposto, contêm verbos que podem, em alguma medida, ser tratados como inacusativos; ou seja, são construções cujos verbos possuem baixa agentividade.

A hipótese inacusativa, proposta inicialmente por Perlmutter e posteriormente incorporada ao contexto da gramática gerativa por Burzio, estabelece que há verbos monoargumentais cujo único argumento é interno. Tais verbos, contrariamente à maioria dos verbos que selecionam complementos, não atribuem caso acusativo a seu argumento (Mioto, 1995). Duarte (1993) assim explica a Generalização de Burzio:

> "*um verbo só atribui papel temático a seu Sujeito se atribuir Caso a seu Objeto. Conseqüentemente, a Estrutura-P de verbos ergativos apresenta a posição Sujeito vazia, detematizada, dado que um verbo ergativo*[12] *não tem a capacidade de atribuir Caso a seu Objeto*" (Duarte, 1993:32).

O fato de não haver argumento externo selecionado pelo verbo possibilita ao argumente interno ser alçado à posição de sujeito. Conforme Mioto (1995:10), "*em termos gerativos, aquele sintagma torna-se o sujeito por necessidade de conseguir seu caso*".

Menuzzi (2001), em nota de rodapé, explica que os inacusativos são verbos cujos sujeitos, na estrurura de superfície, possuem propriedades semânticas e gramaticais que são características do *objeto*, sendo uma delas a de ter papel temático não agentivo.

Mioto (1995) destaca três motivações em favor da hipótese inacusativa. A primeira é a possibilidade de distinção dos verbos monoargumentais em duas classes, de

[12] O termo *ergativo* é tomado como sinônimo de *inacusativo*.

acordo com suas restrições selecionais. Uma classe englobando os verbos que selecionam um argumento [+ animado], que seriam os intransitivos típicos, como em (24) abaixo:

(29) a. O gato miou.
b. João telefonou.[13]

E outra classe abarcando os verbos cujo argumento único não é necessariamente agente, como *chegar*, em (30) abaixo, que seria a classe dos ergativos.

(30) A encomenda chegou.[14]

A segunda motivação explica o fato de sujeitos pospostos se adaptarem melhor aos verbos da segunda classe. Segundo o autor, a posição pós-verbal desses sujeitos se compatibiliza com a do argumento interno.

A terceira e de maior peso refere-se aos verbos cujo complemento não é nominal, como o verbo *parecer*. Para discutir esse verbo, Mioto apresenta três construções sintáticas distintas nas quais o verbo figura:

(31) a. Parece que Maria é feliz.
b. Maria parece ser feliz.
c. Maria parece feliz.

Pela Gramática Tradicional, segundo o autor, em (31.a), *parece* seria um verbo intransitivo; em (31.b), seria um auxiliar e, em (31.c), um verbo de ligação. Segundo a hipótese inacusativa, o verbo teria a mesma classificação nas três estruturas, já que as

[13] Exemplos extraídos de Mioto (1995:12).
[14] Exemplo extraído de Mioto (1995:11).

diferentes estruturas derivariam da natureza do complemento (em (31.a) uma sentença encaixada, em (31.b) um constituinte cujo núcleo é o verbo *ser* e em (31.c) uma pequena oração[15]). Como explica o autor, *"Em todos os casos, fica assegurado que o papel temático de* Maria *é-lhe atribuído em sua relação local com (ser) feliz"* (Mioto, 1995:14).

Assim sendo, na hipótese inacusativa, os verbos são classificados em transitivos, intransitivos e inacusativos.

Um trabalho que lança luz à questão é o de Menuzzi (2001), que, discutindo a estrutura VS nas análises formalista e funcionalista, aborda, entre outras coisas, as relações entre sujeito invertido e concordância. Para compreender as restrições a que a estrutura VS está submetida no PB, um estudo que o autor cita é o de Figueiredo Silva (1996). A autora afirma que a restrição de definitude do sujeito (apenas sujeitos indefinidos podem ser pospostos (Nascimento, 1984 *apud* Menuzzi, 2001)) aponta para outra restrição que diz que a ordem VS só é possível com os verbos inacusativos, e não com os intransitivos. De acordo com a autora:

> *"Em resumo, vimos que o PB não tem a possibilidade de atribuir Caso Nominativo sob regência aos sintagmas nominais que ocupam uma posição à direita do VP. Entretanto, a opção de atribuir Caso Partitivo aos sintagmas nominais que ocupam a posição de objeto, isto é, uma posição interna ao VP, está sempre aberta, o que explica por que objetos profundos de verbos ergativos e de uma certa classe de verbos intransitivos podem permanecer nesta posição interna"* (Figueiredo Silva, 1996:103).

[15] Em inglês *small clause*. Trata-se de um constituinte com significado proposicional, como nas estruturas plenas, mas desprovido de forma verbal. Para maiores detalhes ver Haegeman, L. (1991) *Introduction to Government and Binding Theory.* Oxford: Basil Blackwell.

Os sujeitos pospostos devem ser indefinidos porque recebem Caso partitivo e, como esse é um Caso que só pode ser atribuído a um argumento interno, isto é, a um objeto direto, isto "*explicaria por que somente 'sujeitos' de verbos inacusativos poderiam ser pospostos"* (Menuzzi, 2001:5). Assim, a hipótese é de que só sujeitos que possam ser, de alguma forma, analisados estruturalmente como objetos podem ser pospostos.

Citando Kato e Tarallo, o autor explica que na estrutura inacusativa do sujeito invertido, este ocupa a posição de objeto e o verdadeiro sujeito gramatical da sentença é um pronome nulo expletivo, como em (32) abaixo. Uma evidência para essa análise seria a tendência atual do PB de evitar concordância entre verbo e sujeito na ordem VS.

(32) Apareceu um monte de gente.[16]

O autor discute ainda duas outras estruturas nas quais a ordem VS ocorre com verbos não ergativos. Numa delas, o sujeito ocuparia sua posição canônica, mas o verbo seria deslocado para uma posição mais alta, por exemplo, pela presença de um constituinte interrogativo na sentença, como em (33):

(33) Onde dormem os meninos?

A outra possibilidade seria um SV com verbo não inacusativo numa estrutura de anti-tópico, cujo sujeito invertido é um *afterthought*, como em (34):

(34) a. Ele está pronto, o vestido.

[16] Os exemplos 32, 33 e 34 foram extraídos de Menuzzi (2001).

b. Tá pronto, o vestido.

Na superfície, somente (34.b) configuraria uma estrutura VS, mas (34.a) mostra a correferência com um pronome que ocupa a posição de sujeito e o sujeito invertido estaria numa posição não argumental adjunta à oração.

Menuzzi apresenta, do ponto de vista gerativista, predições feitas em relação às estruturas VS com respeito à CV:

(i) não há concordância entre forma verbal e sujeito invertido em (32);
(ii) há concordância entre forma verbal e sujeito invertido em (33) e (34).

E predições que correlacionam tipo de verbo e concordância:

(i) verbos inacusativos favorecem a ausência de concordância;
(ii) outros verbos favorecem a presença da concordância.

Do ponto de vista da abordagem funcionalista, o autor mostra que uma restrição e um princípio (*Restrição da Ambigüidade* e *Princípio da Baixa Tensão*) dão conta de explicar a relação entre concordância e tipo de verbo; isto é, VS ocorre principalmente com verbos monoargumentais. Mas essa restrição não explica a diferença entre verbos inacusativos e não ergativos, que pode ser explicada pelo *Princípio da Baixa Tensão*, em que verbos inacusativos pertenceriam à figura, dentro da distinção figura/fundo. Isso porque, entre outras características, têm baixa transitividade. Esse quadro teórico oferece instrumentos para a discussão dos argumentos de Decat (1983).

A questão sobre ser o tópico e não o sujeito o elemento controlador da concordância (Decat, 1983 e Camacho, 1993) pode ser discutida com relação à ordem VS, tal como ocorre nos dois trabalhos. Entretanto, vários autores (ver a discussão em Menuzzi, 2001) salientam as fortes restrições sofridas pela ordem VS em PB. Na ordem canônica (SV), tópico e sujeito confundem-se, na posição e, conseqüentemente, na condução da concordância. Assim, não parece produtivo considerar o tópico e não o sujeito o controlador da concordância, já que esta não é uma hipótese verificável para a maior parte das construções produzidas no PB.

Essa análise estende-se também para o trabalho de Camacho (1993). Apoiado nos mesmos pontos da teoria givoniana expostos anteriormente, afirma que, mesmo não havendo indicações seguras sobre o desenvolvimento do fenômeno de concordância para o PB nos termos sugeridos por Givón, encontramos, no PB, indícios do processo de pronominalização anafórica em construções de mudança de tópico, como em (35) abaixo:

(35) Agora o ladrão ele desapareceu correndo.[17]

A partir desse exemplo, o autor afirma o seguinte:

> "*É pertinente, neste caso, a idéia de que o tópico, e não o sujeito gramatical, é o principal elemento no controle da concordância gramatical, não como uma relação entre processos de mudança de tópico e pronominalização anafórica, mas em relação a um tipo específico de concordância gramatical – a que se dá entre o verbo e seu sujeito*" (Camacho, 1993:107).

[17] Exemplo extraído de Camacho (1993:107).

A sentença em (35) é outro caso de co-referência entre o pronome sujeito e o tópico, portanto, não é possível saber qual dos dois elementos está na regência da concordância.

Por outro lado, se no PB não há indícios de relação entre topicalização e pronominalização anafórica, este apresenta processos de redução pronominal e gramaticalização de pronomes que alteram o quadro do sistema pronominal, não em termos de cliticização, mas de enfraquecimento de pronomes. Talvez, e isso é mera suposição, a teoria givoniana, nos termos acima descritos, possa lançar luz sobre estudos desse tipo. O que quero sugerir é que a teoria givoniana relacionada a CV para o PB pode ser produtiva do ponto de vista diacrônico. Um estudo dessa natureza, no entanto, não pode ser realizado no âmbito deste trabalho, ficando como sugestão para projeto futuro.

O outro aspecto da teoria givoniana diz respeito ao fato de que a CV (e os elementos e estruturas a ela relacionados, como pronomes, tipo de verbo, a posição do verbo em relação ao sujeito, entre outros) vai estar operando de forma cruzada com elementos relacionados ao processamento cognitivo da linguagem. Barbisan e Machado (2000:72) oferecem uma boa compreensão dos fundamentos da teoria de Givón:

> *"Para esse autor, a linguagem, em sua organização nocional, funcional e estrutural, é estreitamente controlada e, ao mesmo tempo, motivada pelo sistema cognitivo. Os elementos referenciais tópicos são usados para codificar a coerência referencial. Esses elementos têm a tarefa de acionar operações específicas na mente do receptor. Tais operações envolvem os domínios cognitivos da atenção e da procura do referente na memória. É a informação velha que age como suporte para a informação nova. Acrescenta-se a essas ponderações o fato de o discurso humano ser verbalmente codificado, primordialmente, de nomes."*

Assim, essa é uma teoria que pode lançar luz sobre a relação entre estruturas lingüísticas e estruturas cognitivas.

As questões tratadas nessa sessão mostram vários aspectos, de ordem funcionalista, a serem examinados na fala das crianças em estudo. O objetivo é que tal abordagem sirva de suporte na elaboração de variáveis, tal como pode ser visto em 4.2.2. Serão levados em conta o comportamento da ordem SV/VS, o comportamento dos verbos (considerando a distinção entre transitivos, intransitivos e ergativos) e a estrutura de figura/fundo, todos em relação à marcação da CV.

Os capítulos seguintes tratarão de questões do sistema de CV relacionados às teorias sociolingüística e aquisicionista respectivamente.

2. A SOCIOLINGÜÍSTICA E A CV

Neste capítulo, abordo a teoria lingüística laboviana apresentando, inicialmente, suas principais concepções de base. Na seqüência, traço um quadro do sistema de CV variável i) em comunidades diferentes do PB; e ii) na fala de informantes de Porto Alegre. Vale salientar que em virtude da amostra aqui em estudo ser composta de crianças da comunidade de Porto Alegre, darei mais ênfase à seção que relata pesquisas sobre dados de Porto Alegre, para posterior comparação. Os estudos comentados a seguir servirão de suporte teórico, por um lado, para ajudar a tecer as hipóteses sobre essa investigação e, por outro, para a comparação com os resultados obtidos, permitindo que se investigue a CV na fala da criança nos termos da língua do adulto.

2.1 A Sociolingüística Laboviana

A Sociolingüística é uma área de estudos fundamentada no vínculo inerente entre linguagem e sociedade, ou seja, na linguagem como comportamento social (Camacho, 2003). Seu objeto de investigação é a língua usada na interação social, falada por indivíduos que se relacionam e orientam seu comportamento verbal por um mesmo conjunto de regras (Alkmin, 2003). Dessa forma, tem como projeto descrever e sistematizar a língua falada tendo em vista o cruzamento de aspectos lingüísticos e sociais responsáveis pela coocorrência de formas variantes. A base da teoria está na concepção de língua como sistema heterogêneo, tendo a variação não como conseqüência do uso arbitrário e aleatório de formas, mas sim como sendo sistemática e regular.

O sistema de concordância verbal, no português brasileiro falado, é um exemplo de fenômeno sujeito à variação, tal como ilustra o trecho abaixo:

> *ah@i nós chegamo-0s fazemo-0s uma rodinha <e daí> [/] e daí nós fizemos que nem o ajudante e depois nós brincamo-0s e daí depois <nós fizemo-0s um traba> [//] daí nós fazemo-0s um trabalho daí lavamo-0s as mão-0s lanchemo-0s e vamo-0s pra pracinha brincar quando nós voltamo-0s nós pegamo-0s a mochila <e sen> [//] e senta na rodinha* . (Mat 01'40-46)

Nessa passagem, a criança usa, no mesmo turno de fala, não só diferentes formas de realização da desinência de primeira pessoa do plural como também a ausência da marca. O conjunto de opções possíveis para um mesmo ambiente lingüístico em relação a um determinado fenômeno da língua é tratado como variável dependente (ou variável em estudo), e as formas que a compõem constituem as variantes lingüísticas. A preferência pelo uso de cada uma das formas é resultado da atuação simultânea de diferentes forças, lingüísticas e sociais. Essas forças são estudadas através das variáveis independentes. Dessa

forma, estabelece-se, na relação da variável dependente com as independentes, a análise dos condicionadores operantes no fenômeno variável.

Os estudos sobre o sistema de CV referentes ao uso da regra variável no português falado no Brasil, que trabalham, na maioria das vezes, com a estratificação social em sexo, faixa etária, escolarização e classe social, têm demonstrado que a atuação das forças variáveis é conduzida, dentre outras, pela noção de prestígio. Assim, tem-se o papel da escola fortemente associado à ratificação das formas prestigiadas na comunidade, formas essas que são, normalmente, variantes da língua padrão.

Assim sendo, com este trabalho busco conhecer os condicionadores sócio e psicolingüísticos que atuam no uso variável do sistema de CV na fala das crianças.

2.2 A CV no PB

Os trabalhos que elencarei a seguir tratam da CV variável em diferentes comunidades de falantes do PB.

A primeira pesquisa que referirei é a de Lemle e Naro (1977), que examina a concordância variável de terceira pessoa do plural em dados de falantes analfabetos participantes do projeto MOBRAL. A amostra foi composta para o projeto Competências Básicas de Português e contou com 20 informantes (9 mulheres e 11 homens), com idades entre 17 e 50 anos, naturais do Rio de Janeiro ou adjacências (Grande Rio).

Esse estudo discute a atuação da variável morfológica saliência fônica, que tem mostrado forte influência no sistema de concordância variável em várias pesquisas sociolingüísticas sobre o português falado no Brasil. A hipótese testada pela variável é de que quanto maior a quantidade de material fônico presente na distinção entre a forma de terceira pessoa do singular e a de terceira pessoa do plural, mais saliente é esse dado lingüístico, acarretando maior aplicação da regra de concordância. Os fatores foram estabelecidos dentro de uma escala que vai da menor para a maior saliência. Assim, nas formas verbais em que o plural se dá unicamente pelo acréscimo da nasalização na forma plural (*come/comem; fala/falam*), a manutenção da marca seria desfavorecida, ao passo que nas formas verbais com maior distinção, quando o plural apresenta mudança na posição do acento (*foi/foram; teve/tiveram*) ou quando as formas são inteiramente distintas (*é*/*são*), haveria favorecimento da aplicação da desinência.

Os resultados do estudo, em termos de percentuais e pesos relativos, ratificaram a hipótese: para a distinção *come/comem*, 13,6% - 0,06; para *fala/falam,* 29,6% - 0,17; para *teve/tiveram*, 72,5% - 0,69; para *é*/*são*, 82,4% - 0,81.

Outra variável testada foi a de saliência posicional, medindo a influência da posição do sujeito em relação ao verbo. Os resultados percentuais e de pesos relativos de cada um dos fatores são os seguintes: sujeito posposto, 23,1% - 0,22; sujeito separado (posposto distante), 35,3% - 0,44; sujeito preposto: 49% - 0,70; sujeito oculto, 54,3% - 0,65. Na análise desses valores, os autores concluem que (p.43-44):

> *"A categoria menos saliente é aquela em que o sujeito segue o verbo. Neste caso, o elemento determinante da concordância segue ao elemento determinado, fazendo com que a falta de concordância seja*

> *menos óbvia. No caso do sujeito separado há uma distância física, com a correspondente demora temporal, entre determinante e determinado. O caso mais saliente é o do sujeito pré-posto, quando o sujeito plural antecede imediatamente ao verbo candidato à concordância".*

Sobre a variável morfológica, penso que cabe uma reflexão, já que o princípio da saliência fônica tem sido tomado, em algumas pesquisas, como um princípio funcional que balizaria fortemente a concordância variável. A generalização funciona de forma ampla, mas não dá conta das minúcias que distinguem cada nível da hierarquia. Como saliência é um objeto da psicolingüística, ou seja, diz respeito à percepção, os níveis de diferença fônica deveriam ser cruzados com outro tipo de construto (como teste de percepção) para então se chegar a conhecer as diferenças categoriais. Há circularidade no momento em que os resultados dos testes de significância estatística mostram quais categorias são mais salientes e posteriormente elas são testadas como mais salientes. Do ponto de vista científico, as categorias que discernem os níveis não foram criadas aprioristicamente, embasadas em um construto consistente, e posteriormente testadas; elas foram criadas a partir dos resultados, o que torna a testagem redundante[18].

Além disso, Lemle e Naro concentram-se nas distinções fônicas, sem considerar as forças que geram estas distinções. Loregian, 1996 e 2001 e Amaral, 2002 e 2003 salientam que há cruzamento da saliência fônica com tempo verbal, tonicidade e vogal temática do verbo. Portanto, é uma variável que precisa ser examinada com maior profundidade.

O segundo estudo que mencionarei é o de Rodrigues (1992), que investiga a concordância de primeira e terceira pessoas do plural na fala de quarenta informantes

[18] Essa discussão foi levantada pela professora Luciene Simões.

analfabetos ou semi-escolarizados, pertencentes a uma comunidade de favelados da periferia de São Paulo (Zona Oeste – Carombé). A população é constituída por pessoas de baixa renda, em sua maioria migrantes das zonas rurais sem qualificação profissional. A autora postula, para a variedade falada por essas pessoas, a noção de dialeto social popular, vinculado a *"um estrato social que utiliza uma variedade de língua de menor prestígio, não codificada e não normalizada, eminentemente de cunho oral, rotulada de língua popular"* (p.155-56). Os resultados gerais mostram aplicação da marca de concordância em 59% dos casos de DNP-P4 e 29% nos de DNP-P6[19].

A pesquisa se concentra em dois grupos de fatores. O primeiro é a presença/ausência de sujeito pronominal, que pretende verificar a hipótese de que a elipse do sujeito formal favoreceria a manutenção das marcas de concordância, compensando uma possível perda de informação de pessoa e número a respeito do sujeito, ao passo que as formas verbais não marcadas seriam favorecidas pelo sujeito pronominal explícito. Os fatores considerados para o teste dessa hipótese foram: sujeito não pronominal, explícito e não explícito. Em relação à DNP-P1, os resultados mostraram índices bastante robustos de aplicação da desinência relacionada ao sujeito nulo (74% e 0,81), ao passo que os sujeitos pronominal explícito e não pronominal mostraram-se desfavorecedores da aplicação da marca com índices percentuais e probabilísticos de 47% - 0,45 e 32% - 0,22[20] respectivamente. Em relação à DNP-P6, a diferença entre o sujeito elíptico e os sujeitos explícitos é menor, mas a freqüência e a probabilidade também apontam o primeiro como

[19] Esses resultados foram extraídos de Almeida (2005:29), que calculou os índices gerais baseados na Tabela 1 de Rodrigues (1992:167) sobre a variável sexo.

[20] Esses índices foram extraídos de Almeida (2005:29), que converteu os resultados contidos em Rodrigues (1992:159), apresentados em termos de não aplicação da marca, para índices de aplicação da marca. Tal como no trabalho daquela, os percentuais e pesos relativos são mais úteis aqui se apresentados como aplicação, para fins de comparação.

maior favorecedor de manutenção da marca: 52% e 0,62 (elíptico); 28% e 0,55; (pronominal); 18% e 0,33 (não-pronominal).

O segundo grupo de fatores examinado pela autora é sexo. Segundo ela (p.166), *"nas sociedades urbanizadas do mundo ocidental, falantes do sexo feminino tendem a usar formas de prestígio mais freqüentemente que falantes do sexo masculino"*. Diferentemente dos resultados da variável anterior, a distinção de sexos mostrou probabilidades tanto para DNP-P4 como para DNP-P6 bastante próximas do ponto neutro: 0,57 (masculino) e 0,43 (feminino) e 0,47 (masculino) e 0,53 (feminino) respectivamente. É interessante notar que são índices contrários entre si, no sentido de que os homens favorecem a aplicação da marca na DNP-P4 e desfavorecem na DNP-P6, ao passo que as mulheres desfavorecem em P4 e favorecem em P6. Ou seja, a distinção entre sexos é um fator complexo, atuando, provavelmente, de forma cruzada com outros fatores.

O terceiro estudo que referirei é o de Monguilhott e Coelho (2002). A pesquisa se concentra no exame da terceira pessoa do plural em dados de fala de 24 informantes da comunidade de Florianópolis, estratificados por sexo (masculino e feminino), idade (15 a 24 anos, 25 a 45 anos e 52 a 76 anos) e escolaridade (4 anos de escolarização e 11 anos de escolarização), a partir de entrevistas extraídas do Banco de Dados do projeto VARSUL. O estudo segue a perspectiva da sociolingüística paramétrica, cujo objetivo é conciliar os pressupostos metodológicos da teoria variacionista laboviana com o modelo chomskiano de Princípios e Parâmetros.

Do total de 1583 dados obtidos, 1251 (79%) apresentaram aplicação da desinência e 332 (21%) mostraram a variante zero. As variáveis selecionadas pelo

programa VARBRUL estatisticamente relevantes para a amostra foram: saliência fônica, posição do sujeito em relação ao verbo, paralelismo formal, traço humano no sujeito, tipo de verbo e tipo de sujeito.

A variável saliência fônica, tida como a mais relevante, foi controlada pelas autoras em dois níveis: o nível 1 de oposição não acentuada e o nível 2 de oposição acentuada. Os resultados em percentuais e pesos relativos foram os seguintes: no nível 1, 25% e 0,02 em *conhece/conhecem*; 80% e 0,46 em *ganha/ganham*; 66% e 0,13 em *diz/dizem*; no nível 2, 96% e 0,88 em *vai/vão*; 83% e 0,65 em *bateu/bateram*; 90% e 0,75 em *é/são*. As autoras afirmam que os resultados corroboram os de outras pesquisas já que a oposição acentuada favorece a marca enquanto que a não acentuada desfavorece. Entretanto, os índices mostram que não há atuação da hierarquia de saliência dentro de cada nível.

A variável posição do sujeito em relação ao verbo mostrou que o sujeito anteposto favorece a desinência em 84% dos casos, com peso relativo de 0,58, ao passo que o sujeito posposto a desfavorece (52% e peso relativo de 0,17). A análise das autoras sobre esses resultados se pauta, por um lado, sobre um viés sociolingüístico de trabalhos empíricos e, por outro, sobre o viés teórico da hipótese inacusativa, que sustenta que esses verbos aceitam argumentos à sua direita, sem que o verbo apresente marcas de concordância. As autoras salientam, com isso, o cruzamento existente entre tipo de verbo e posição do sujeito em relação ao verbo.

Em relação ao tipo de verbo, a cópula apresentou maior retenção da marca, com freqüência de 78% e peso relativo de 0,63. Os transitivos e intransitivos ficaram numa

posição intermediária, pouco abaixo do ponto de indiferença estatística, com resultados de 82% e 0,49 e 82% e 0,46 respectivamente. Os inacusativos se mostraram os mais influentes na ausência da marca, com 71% de freqüência e peso relativo de 0,37.

Os trabalhos aqui expostos, de pesquisas empíricas da fala de adultos, mostram a atuação de alguns grupos de fatores que servirão de termos de comparação para o estudo da fala das crianças.

Em relação à DNP-P6, a variável saliência fônica mostrou que atua no sentido da retenção da marca nas formas verbais em que há maior distinção fônica entre a forma alvo e a forma não marcada (Lemle e Naro, 1977). Monguilhott e Coelho (2002) encontraram atuação da variável em termos de oposição não acentuada *versus* oposição acentuada, mas o segundo não mostrou atuação da hierarquia de saliência entre os fatores.

A variável posição do sujeito apresentou resultados comparáveis nos três estudos, com o sujeito anteposto favorecendo a aplicação da desinência e o posposto a desfavorecendo.

Em relação tanto à DNP-P4 quanto à DNP-P6, as variáveis presença/ausência de sujeito pronominal e sexo apresentaram desempenhos parelhos. O sujeito elíptico mostrou-se atuante na manutenção da desinência, seguido pelo sujeito pronominal, enquanto que o sujeito não-pronominal mostrou-se favorecedor da ausência de marca. Quanto ao sexo (Rodrigues, 1992), homens e mulheres apresentaram desempenho lingüístico invertido em relação à marca de CV.

2.3 A CV em Porto Alegre

Meu objetivo nesta seção é apresentar resultados de trabalhos de modo a formar um quadro representativo do uso de flexão de CV por informantes de Porto Alegre. Essa abordagem será útil para, posteriormente, traçar um perfil comparativo com os dados das crianças em análise. Organizarei a exposição contemplando distintamente P2, P4 e P6, já que são ambientes lingüísticos, em relação ao sistema de CV, conduzidos por regra variável.

Em relação a P2, apresento a análise referente às variáveis sociais do trabalho de Loregian (1996). O estudo examina o fenômeno de concordância verbal com o pronome *tu* em dados de fala de informantes de três localidades da região sul do Brasil: Porto Alegre, Florianópolis e Ribeirão da Ilha. As duas primeiras são capitais e a última, um bairro rural de Florianópolis. A base teórica e metodológica é a da Teoria da Variação.

As variáveis lingüísticas independentes examinadas no estudo foram: paralelismo formal no nível discursivo, interação emissor/receptor, explicitação do pronome, tempo verbal, saliência fônica, tonicidade do verbo, número de sílabas do verbo, contexto fonológico seguinte. As variáveis sociais foram: região, grau de escolarização, faixa etária e sexo.

Dessas, a primeira a ser escolhida pelo programa VARBRUL como estatisticamente significativa foi paralelismo formal, seguida, na ordem, por região, tempo

verbal, explicitação do pronome, interação emissor/receptor, tonicidade, número de sílabas, grau de escolarização, faixa etária, contexto fonológico seguinte, saliência fônica e sexo.

As variáveis sociais relevantes para o programa estatístico foram região, grau de escolarização e faixa etária. Os resultados destes fatores foram analisados primeiro individualmente, e posteriormente através de cruzamentos entre as variáveis selecionadas.

Os resultados gerais por região mostram que o índice de presença de marca de concordância em Porto Alegre contrasta fortemente com o das outras localidades: entre os informantes de Porto Alegre é predominante o uso do pronome *tu* com verbo sem marcação de concordância, com freqüência de 4% e peso relativo de 0,12, em Florianópolis 40% e 0,71 e em Ribeirão da Ilha, 57% e 0,81.

No cruzamento das variáveis região e grau de escolarização, os resultados mostram que em Porto Alegre o grau de escolarização não interfere no aumento da presença da marca de concordância com o pronome *tu*, diferentemente do que ocorre nas outras localidades em que a escolaridade é diretamente proporcional ao uso da flexão.

Outro cruzamento feito na pesquisa é entre faixa etária e região. A autora salienta que, mesmo com pesos relativos baixos (2% e 0,13 no grupo entre 25 e 49 anos e 6% e 0,29 no grupo com mais de 50 anos), os informantes com mais de 50 anos apresentam mais marcas de concordância em sua fala que os da faixa etária anterior.

O fator sexo não foi selecionado dentre as variáveis no âmbito geral do trabalho, já que os resultados mostram uma diferença probabilística praticamente

inexistente entre os falantes do sexo masculino (34% de aplicação da marca e peso relativo de 0,52) e do sexo feminino (29% de aplicação e peso de 0,49). Entretanto, foi selecionado para Porto Alegre, mas esses resultados a autora não especifica.

Diante dos resultados encontrados por Loregian, com um índice tão baixo de manutenção da marca de P2 em adultos, não espero encontrar tal marcação em crianças, já que elas teriam um contato muito pequeno com esta forma verbal.

Passo agora à marcação de concordância verbal de primeira pessoa do plural. O estudo de Zilles e Batista (2005) examina a marca de DNP-P4 em estruturas de sujeito explícito *nós* ou seu nulo, considerando três variantes: a marca padrão (*-mos*), a marca não padrão com apagamento parcial da desinência *(-mo*) e a marca não padrão com apagamento total da desinência (desinência zero). Trata-se de um estudo de tendência, sob aporte da Teoria da Variação, sobre duas amostras de informantes cultos (com grau de escolaridade de nível superior) da cidade de Porto Alegre, uma dos anos 1970 e outra dos anos 1990, tendo interesse em examinar um possível processo de mudança na desinência de CV de primeira pessoa do plural nesse segmento social. O fato de se tratar de falantes cultos é importante para o presente trabalho, uma vez que as crianças cujos dados estão sob análise, conforme pode ser visto no capítulo 5 a seguir, são, na maioria, filhas de pais com terceiro grau de escolaridade, da mesma cidade que os informantes de Zilles e Batista, o que torna relevante a comparação entre os resultados das crianças aqui tratadas e os adultos do citado estudo.

Na primeira etapa do estudo, foi feita uma análise ternária considerando as três variantes distintamente. O resultado geral mostra a seguinte distribuição percentual:

desinência zero, 3%; desinência não-padrão (*-mo*), 15%; e desinência padrão (*-mos*), 82%, que mostra, por um lado, a ocorrência da desinência padrão muito robusta e, por outro, a desinência zero muito baixa.

A segunda etapa da análise contrapôs a presença à ausência de marca de concordância, reunindo as formas padrão e não padrão em uma única variante. Os resultados mostraram 33 ocorrências de desinência zero (sendo 3 ocorrências, ou 1%, nos anos 70, e 30, ou 5%, nos anos 90) de um total de 1195 ocorrências, ou seja, 3% (índice geral relativo às duas décadas) de omissão da DNP-P4. Vê-se, assim, que a omissão da marca de desinência é muito pequena nos dados de falantes cultos de Porto Alegre.

O resultado relacionado à omissão total da desinência refere-se aos tempos verbais nos quais a forma alvo produziria uma palavra proparoxítona, forma prosodicamente marcada no PB. Segundo as autoras, mesmo o apagamento sendo bastante estigmatizado, este contexto seria propício para sua ocorrência. Para testar tal hipótese foi feita uma rodada amalgamando os tempos verbais de forma a reunir, numa variante, as paroxítonas e, noutra, as proparoxítonas. Os resultados mostraram que, dos 33 casos de apagamento da desinência, 27 são de tempos verbais que formariam proparoxítonas, perfazendo um total de 82% dos casos. E ainda, em relação às outras 6 ocorrências de ausência de marca, estão em contextos nos quais o padrão prevê infinitivo flexionado. Sobre isso, Zilles e Batista (2005:14) afirmam que:

> *"Houve apagamento da desinência em apenas 2% das ocorrências de tempos em que a forma alvo não seria proparoxítona. Cabe acrescentar que tivemos 318 ocorrências de verbos em que a forma alvo*

> *seria uma palavra proparoxítona, mas destas, apenas 27 não apresentaram concordância, num percentual de 8%. Quando há a concordância (92% dos casos), deve-se notar ainda que o apagamento do* –s *é altamente desfavorecido, podendo-se mesmo pensar em uma restrição ou bloqueio da regra de redução da desinência em favor da regra que gera a forma sem desinência".*

Em relação ao apagamento do *–s* (redução da desinência), os resultados apresentados referem-se à variável faixa etária, testando a hipótese de que os falantes com menos de 50 anos apagariam mais do que aqueles com mais de 50. Os valores são de 697 ocorrências de presença da desinência nos falantes mais velhos, sendo que, dentre essas, 105 ocorrências (15%) sofrem apagamento do *–s*. Dentre os mais jovens, há 316 ocorrências, com apagamento em 62 delas (19%), com pesos relativos de, respectivamente, 0,45 e 0,60. A análise é feita, ainda, em relação às duas décadas: nos anos 70 a diferença de índice de apagamento entre as idades (isto é, entre os mais jovens e os mais velhos) é de 7% e de 10% nos anos 90.

Na análise quantitativa, o programa VARBRUL selecionou também como significativa para a redução da desinência a variável tempo e modo verbais, num grupo composto por quatro variantes: uma que reunia os tempos cuja forma alvo era uma palavra proparoxítona; uma reunindo presente a pretérito perfeito, tempos que mostram a mesma forma; uma que isolava a perífrase de futuro: verbo *ir* + infinitivo; e, finalmente, uma que considerava uma perífrase modal: verbo *ir* + infinitivo com usos injuntivo e exortativo. Os resultados foram os seguintes para ocorrências de apagamento parcial da desinência:

- formas proparoxítonas: 6/260 (2%), com peso relativo de 0,15;

- presente e perfeito do Indicativo: 83/556 (15%) e peso relativo de 0,57;

- perífrase de futuro: 6/29 (17%) e peso relativo de 0,61;

- perífrase modal: 72/141 (51%) e peso relativo de 0,89%.

Os índices mostram a forte restrição à redução da desinência quando a forma alvo é proparoxítona. Os demais casos favorecem o apagamento do *–s*, enfatizando o comportamento do verbo *ir* (nas duas perífrases consideradas) como os ambientes que mais conduzem à redução.

A pesquisa de Zilles, Maya e Silva (2000) investiga a variação na concordância verbal de primeira pessoa do plural em dados de língua falada de duas comunidades, Panambi (cidade gaúcha de colonização alemã) e Porto Alegre, em dados provenientes do projeto VARSUL. O estudo trata a variação segundo dois processos distintos, um fonológico – de apagamento do /s/ – e um morfossintático – de aplicação da regra de concordância. Assim, a variável dependente é composta de três variantes: a forma padrão (*-mos*), o apagamento do /s/ (*-mo*) e a desinência zero (∅). A análise estatística foi feita em três etapas, sendo a primeira ternária, relacionada à distribuição das três variantes, e as duas outras, binárias, considerando respectivamente omissão de DNP versus *-mo* e *-mos* somados, e alternância de *-mo* e *-mos*.

Na primeira rodada, num total de 1035 ocorrências, os percentuais distribuem-se da seguinte forma: 53% de aplicação da forma padrão (*-mos*), 34% de apagamento do /s/(*-mo*) e 13% de desinência zero (∅). Como salientam os autores, o índice de manutenção da desinência, 87%, é bastante robusto.

As outras duas rodadas relacionam a variável em estudo a 13 fatores significativos, sendo 9 lingüísticos – conjugação do verbo, tempo e modo verbal, realização do sujeito, estrutura verbal, tipo de discurso, contexto seguinte, posição e distância do sujeito em relação ao verbo, alternância da vogal temática e posição do acento na forma verbal alvo – e 4 sociais – escolaridade, sexo, idade e comunidade (Panambi e Porto Alegre). Os resultados, considerando omissão da desinência, foram os seguintes:

a) posição do acento: as proparoxítonas favorecem a omissão da DNP-P4 em 43% dos casos (0,97 de peso relativo) contra 2% (0,28) nas paroxítonas.

b) posição do sujeito: a posposição favorece a omissão da desinência em 80% dos casos.

Os autores citam, à parte, os casos de sujeito nulo, que correspondem a 4% de desinência zero.

Nos resultados relativos ao apagamento do /s/, os números relativos a tempo verbal são os seguintes: favorecem o apagamento *Vamos* + infinitivo, com 73% e peso relativo de 0,81; presente do indicativo, com 38% e peso relativo de 0,57; perfeito, com 40% e peso de 0,55; imperativo, com 37% e peso de 0,41; futuro do subjuntivo + infinitivo flexionado, com 13% e peso de 0,24; e imperfeito, em 3% dos casos (e peso relativo de 0,09). Sobre esses resultados, os autores sugerem que as formas verbais mais raras no uso da língua oral favorecem a forma padrão *-mos*, enquanto que as de uso mais freqüente na fala favorecem a forma não padrão, além de reforçar as formas proparoxítonas como inibidoras da marca.

O estudo de Barden (2004) examina a concordância variável de terceira pessoa do plural sob base teórica e metodológica da Teoria da Variação, utilizando o Pacote VARBRUL de análise estatística. A amostra provém do projeto VARSUL e é composta por informantes de Porto Alegre, estratificados por nível de instrução (de 1 a 4 anos de escolaridade, de 5 a 8 anos, e de 9 a 11 anos), sexo (masculino e feminino) e idade (menos de 50 anos e mais de 50 anos). Os outros grupos de fatores considerados foram saliência fônica, paralelismo formal, tipo de sujeito explícito (opondo pronomes a nomes quantificados) e posição do sujeito (imediatamente anteposto ao verbo, anteposto com material interveniente de 1 a 3 sílabas, anteposto com mais de 3 sílabas entre sujeito e verbo, e posposto ao verbo). A variável dependente examina a) a presença total da marca flexão, b) a forma reduzida, desnasalizada, que ela chama de *presença parcial de flexão*, e c) a ausência total da marca.

A distribuição percentual a que chega a análise é a seguinte: concordância total, 57%; concordância parcial, 22%; e sem concordância, 21%. Em termos de percentagem, o índice de aplicação da regra é bastante robusto, como salienta a autora, perfazendo um total de 79% (1039/1321 dados). Para verificar o comportamento diferenciado das duas formas de marcação da desinência, foram feitas duas rodadas distintas, uma considerando apenas a aplicação total, e outra, a aplicação total e a parcial somadas. Na rodada relativa apenas à aplicação total, o programa selecionou como estatisticamente significativos os seguintes fatores: 1) saliência fônica, 2) escolaridade, 3) paralelismo formal, e 4) posição do sujeito. Os fatores sexo, idade e tipo de sujeito foram excluídos. Na rodada que considerava as duas formas de aplicação da marca, as variáveis selecionadas foram: 1) saliência fônica, 2) posição do sujeito, 3) paralelismo formal, 4)

escolaridade, e 5) tipo de sujeito, excluindo sexo e idade. Diante deste quadro, a autora optou pela união das formas relativas à manutenção da marca.

Os resultados (em número de ocorrências e peso relativo) referentes à aplicação da regra de concordância em função da posição do sujeito são os seguintes:

- imediatamente anteposto: 414/502 (82%) e 0,63;
- anteposto com 1 a 3 sílabas: 202/251 (80%) e 0,60;
- anteposto distante (mais de 3 sílabas): 200/267 (75%) e 0,39;
- sujeito posposto: 40/116 (34%) e 0,11.

Os índices revelam um comportamento semelhante entre o sujeito imediatamente anteposto e o sujeito com pouco (até 3 sílabas) material interveniente, ou seja, a adjacência ou proximidade do sujeito anteposto conduz à aplicação da regra. Aumentando essa distância (mais de 3 sílabas), a tendência em relação à marcação inverte, passando a um contexto de condução ao apagamento da marca. O sujeito posposto mostra-se como um contexto robusto de não aplicação da regra.

Em relação ao paralelismo formal, os resultados revelam, segundo a autora, "*forte correlação entre o aparecimento de um verbo marcado (flexionado), total ou parcialmente, e a presença de marcas explícitas no verbo subseqüente"*. Os índices são os seguintes: verbo precedido de verbo marcado mostra 86% (254/218 ocorrências) e peso relativo de 0,65 de manutenção da marca, e verbo precedido de verbo parcialmente marcado apresenta 89% (90/101 ocorrências) e peso relativo de 0,64. Ainda, *"o verbo não-marcado (não-flexionado) conduz à ausência de marca na ocorrência verbal seguinte"*

(55% de ocorrências e 0,27 de peso relativo), e *"o sintagma verbal que ocorre isolado apresenta um comportamento neutro: ele tem 0,46 de probabilidade de receber as marcas de concordância"* (73% de ocorrências).

Os resultados do tipo de sujeito explícito revelaram, segundo a autora, que sujeitos pronominais promovem maior tendência de uso da desinência (0,57 de peso relativo) do que sujeitos com nomes quantificados (0,44 de probabilidade). Pode-se dizer, em relação a esses índices, que o primeiro tipo de sujeito conduz à manutenção da marca enquanto o segundo conduz à não marcação.

Em relação à escolaridade, os índices mostraram que à medida que aumenta o número de anos de escolarização do indivíduo cresce a tendência à manutenção da marca, sendo os resultados polarizados no sentido de não favorecimento da marca pelos falantes com Primário e Ensino Fundamental, e favorecimento pelos falantes com Ensino Médio: Primário, 70% e peso relativo 0,39; Ensino Fundamental, 75% e peso relativo 0,47; e Ensino Médio, 82% e peso relativo 0,64.

As variáveis sexo e idade, como dito anteriormente, mostraram-se estatisticamente irrelevantes. Em relação a sexo, tanto o resultado de homens quanto de mulheres mostrou peso relativo de 0,50, o que indica que homens e mulheres têm comportamento lingüístico semelhante no que tange à marcação da flexão verbal. Em relação à idade acontece o mesmo, os mais novos e os mais velhos apresentam índices de 0,50 de probabilidade estatística de concordância, ou seja, o ponto neutro.

Ainda em relação à terceira pessoa do plural, o trabalho de Batista e Zilles (2005) apresenta um *estudo de painel*[21] em tempo real, tendo como base de dados informantes de Porto Alegre que nos anos de 1970 compuseram parte do Projeto NURC e nos anos de 1990 foram recontatados pelo Projeto VARSUL. A amostra consta de 3 homens e 4 mulheres que, na década de 70, com exceção de uma, estavam na faixa etária de 25-44 anos, e na década de 90 estavam todos na faixa de mais de 45 anos. O objetivo do trabalho é verificar se houve mudança estatisticamente significativa na concordância da DNP6.

Sob a variável dependente foram consideradas três formas: a presença da marca padrão (/aw/ nasalizado, /ey/ nasalizado, aN, eN), a presença da marca não padrão (/u/ nasal, /u/, /i/ nasal, /i/) e a ausência da marca. As variáveis lingüísticas eram: posição do sujeito; tipo de sujeito; animacidade do sujeito; traço humano; saliência fônica; conjugação verbal; tempo e modo verbais. As variáveis sociais eram idade, sexo, década, informante.

A análise estatística, amparada no pacote VARBRUL, foi feita em três etapas. Na primeira, uma análise ternária, foram consideradas as três variantes, resultando na seguinte distribuição percentual dos 1525 dados computados: forma padrão: 89%; forma não padrão: 6%; desinência zero: 5%. A segunda rodada, visando contrapor a presença à ausência de marca, teve reunidas as ocorrências padrão e não padrão. Nesta etapa, o programa selecionou como estatisticamente relevantes as variáveis posição do sujeito, tempo/modo verbais e saliência fônica. A primeira delas mostra que a desinência zero ocorre preferencialmente em contextos de sujeito posposto.

[21] Segundo as autoras, "*o estudo de painel tem como objetivo investigar a estabilidade individual da fala durante a vida adulta, comparando entrevistas dos mesmos indivíduos em dois momentos de sua vida*". PÁGINA

Os pesos relativos, considerando o apagamento da DNP-P6, são os seguintes: sujeito contíguo (ordem SV): 0,43; sujeito não contíguo (ordem SV): 0,74; sujeito posposto: 0,89. Em relação a tempo/modo verbais, também considerando o apagamento da desinência, os resultados mostram que o perfeito (peso relativo 0,06) está fortemente relacionado à presença da marca, enquanto as outras formas verbais consideradas favorecem o apagamento: perífrase de futuro: 0,54; presente do indicativo: 0,58; imperfeito do indicativo: 0,70; presente do subjuntivo: 0,73; e finalmente, o maior favorecedor de ausência de marca, o infinitivo flexionado: 0,85. Os resultados referentes à saliência fônica mostram maior apagamento nos três primeiros níveis; entretanto, por falta de dados de ocorrência zero em quatro níveis, as autoras concluíram que a escala funciona nos níveis menos salientes.

Em relação à DNP-P4, os estudos de Zilles e Batista (2005) apontam como inibidores da desinência as formas proparoxítonas e a posposição do sujeito. Para a DNP-P6, a posposição do sujeito também foi atestada tanto por Barden (2004) quanto por Batista e Zilles (2005). O grupo de fatores paralelismo formal se mostrou atuante no sentido de correlação entre um verbo marcado e a presença da marca no verbo seguinte. Além disso, o tipo de sujeito explícito mostrou que o sujeito pronominal favorece a aplicação da desinência enquanto o nome quantificado a desfavorece. Das variáveis sociais, escolaridade se mostrou atuante no sentido de aumento da tendência à aplicação da marca com o aumento do tempo de escolarização. As variáveis sexo e idade se mostraram irrelevantes.

Os trabalhos apresentados neste capítulo, nas duas sessões anteriores, têm como objetivo traçar um quadro da atuação de algumas variáveis na fala de adultos de

diferentes comunidades. Tal quadro servirá, por um lado, para estabelecer as hipóteses de pesquisa; por outro, como parâmetro de comparação para que se compreenda a fala das crianças. Tendo por base um perfil desse tipo, espero que seja possível interpretar os resultados de modo a distinguir os casos de atuação da variação sociolingüística dos casos de atuação da variação aquisicional; ou seja, espero poder perceber quando a concordância variável na fala da criança está respondendo ao fato de ela ser integrante da comunidade de fala e quando está respondendo ao fato de o sujeito se tratar de uma criança.

3. A AQUISIÇÃO DA LINGUAGEM E A CV

Neste capítulo tratarei das teorias psicolingüísticas, abordando certas questões relevantes à área de aquisição da linguagem. Inicialmente, examinarei algumas descrições das etapas do desenvolvimento da gramática inicial da criança. Apresentarei, na seqüência, uma discussão em torno da comparação estrutural da fala inicial com a língua do adulto e, por fim, uma reflexão sobre a relação entre a aquisição da linguagem e a socialização da criança.

3.1 Os modelos de aquisição da linguagem

As questões centrais da área de aquisição giram em torno da emergência da faculdade da linguagem no ser humano: como, quando e por que, num determinado momento, a criança irrompe em direção à linguagem? Qual a natureza da linguagem inicial da criança? Que tipo de conhecimento lingüístico a criança tem no início e ao longo do

processo de aquisição da língua materna? O que é inato e o que é aprendido no processo de aquisição da língua pela criança?

Uma teoria fortemente direcionada para a elaboração de instrumentos teóricos que dêem conta dessas e de outras questões relativas à aquisição da linguagem é a gramática gerativa, proposta por Chomsky e seus seguidores. De acordo com essa teoria, o exame extensivo de diferentes línguas (numa concepção adultocêntrica) é capaz de mostrar que a gramática das línguas naturais compartilha um núcleo comum de propriedades universais. O conjunto de tais propriedades é conhecido como *Gramática Universal* (GU), cujo estudo se concentra na busca da teoria universal da gramática das línguas naturais. O argumento gerativista é de que se as diferentes línguas compartilham propriedades da GU (propriedades, essas, abstratas), elas não são simplesmente aprendidas pela criança através da exposição à experiência lingüística (*input*). A conclusão a que chega a teoria é de que os seres humanos são inatamente dotados de um sistema ricamente estruturado de conhecimento lingüístico. O sistema opera de modo a guiar a criança na análise do estímulo da língua que está sendo adquirida.

Para os gerativistas, todas as gramáticas de língua adulta compreendem um complexo sistema de *módulos* interconectados. Cada módulo é responsável por um aspecto específico de propriedades da gramática, por exemplo, o módulo *categorial* é responsável pelos princípios que regulam a estrutura categorial dos constituintes das sentenças. As línguas possuem também um sistema modular paralelo, que opera com possibilidades de variação intramodular (variação essa possível dentro de limites fixados). O modelo de Princípios e Parâmetros dá conta desses dois sistemas: os *princípios* codificam as propriedades invariáveis das línguas, sendo, portanto, universais, similares para todas as

línguas adultas (embora alguns autores afirmem que nem todos os princípios operam em todas as línguas). Os *parâmetros* codificam as propriedades que variam de uma língua para outra.

A tarefa da GU, nessa proposta, é determinar quais são os Princípios universais, comuns a todas as línguas, e quais são os Parâmetros de variação possíveis ao longo das línguas. A tarefa da criança que está adquirindo sua língua materna é selecionar os valores apropriados dos parâmetros especificados pela GU, ajustando-os ao valor expresso pela língua do ambiente.

Outra abordagem que oferece modelos para a busca de respostas às questões de desenvolvimento da gramática inicial da criança é a de cunho cognitivista. Essa perspectiva explica o desenvolvimento da linguagem em termos de compartilhamento, ao menos em parte, por uma base de processamento psicológico com outros domínios cognitivos. Nesse sentido, essa proposta opõe-se ao gerativismo em relação aos módulos. Aqui, o desenvolvimento da linguagem se dá em termos não modulares, isto é, os módulos não são específicos para o processamento da linguagem.

Outra diferença em relação ao gerativismo diz respeito ao tratamento das categorias lingüísticas. Mesmo reconhecendo-as como instrumentos descritivos úteis, os estudiosos cognitivistas não tomam essas categorias como analíticas, inatas e de domínio específico. Antes, sugerem a possibilidade de que habilidades cognitivas mais gerais, de diferentes tipos, atuando tanto na linguagem quanto em outros processos de aprendizagem contribuem para o desenvolvimento da linguagem, tanto em relação às categorias gramaticais específicas quanto em relação às habilidades cognitivas da linguagem. Alguns

desses modelos também assumem que as categorias lingüísticas podem ser adquiridas por processos de aprendizagem de natureza probabilística, e não categórica.

Dentro dessa abordagem, vários autores enfatizam que as diferenças quantitativas e qualitativas que caracterizam a aquisição da linguagem, tanto numa criança individualmente, quanto numa criança em relação à outra, numa mesma língua ou em línguas diferentes, podem ser usadas como argumento contrário à explicação de aquisição da linguagem em termos de um bioprograma inato e universal (Pizzuto e Caselli, 1992).

A seguir, apresento um quadro das etapas iniciais do desenvolvimento da linguagem.

3.2 As etapas do desenvolvimento da linguagem

Antes de entrar no assunto, preciso fazer um esclarecimento a respeito da relação entre os temas abordados neste capítulo e o (des)compromisso ontológico de minha parte com alguns termos e conceitos de que me valerei na seqüência do estudo. Uma observação que se faz necessária advém do fato de que uma abordagem variacionista, tal como a que conduz o tratamento dos dados aqui em análise, e, portanto, conduz minha leitura sobre o sistema da CV em relação aos dados, não condiz com o determinismo subjacente à abordagem gerativista. A inclusão de trabalhos gerativistas na base teórica que construo não significa minha adesão a essa escola. Entretanto, dada a importância de pesquisas nessa linha para a descrição do sistema de CV em crianças, lanço mão de tais estudos, cujos termos e conceitos serão devidamente explicados à medida que aparecerem.

Passo agora às fases do desenvolvimento da linguagem. Brown (1973), utilizando-se do método de diários naturalísticos baseados em observações longitudinais de três crianças adquirindo inglês como língua materna, descreve o surgimento de relações sintáticas e semânticas no que diz respeito às construções das sentenças e aos significados atribuídos a essas construções. O autor sugere que o padrão da linguagem inicial segue uma ordem de desenvolvimento em cinco estágios. Nessa obra, Brown trata dos dois primeiros, sendo que os estágios foram estabelecidos em função do número médio de morfemas por enunciados (MLU – *Mean Length of Utterance*), medida largamente usada nos estudos posteriores na área aquisicionista. Segundo o autor, o modelo de linguagem em ordem desenvolvimental pode ser aplicado a outras línguas que não o inglês, desde que sejam levadas em conta certas diferenças entre as línguas.

O Estágio I é caracterizado pela fala telegráfica. As características da fala telegráfica foram inicialmente descritas a partir de experimentos em que crianças recebiam a tarefa de repetir (por imitação) um enunciado da fala adulta. Nas produções das crianças mais jovens, foram observadas as seguintes propriedades: limite em relação ao comprimento do enunciado (essa propriedade traduz-se numa limitação do número de morfemas por enunciado da criança, manifestando-se na omissão de flexões verbais e de plural, de cópula e verbos auxiliares, de artigos, preposições e conjunções), manutenção das palavras de conteúdo (nomes, verbos e adjetivos) e preservação da ordem de palavras. Estudos de fala espontânea confirmaram essas características.

Em torno dos 20 meses de idade, ainda no Estágio I, a criança começa a combinação de duas palavras, em relações que são definidas antes semântica que

gramaticalmente. É importante ressaltar que é nesse ponto que tem início o desenvolvimento da gramática da língua que está sendo adquirida e, por ter esse caráter particularmente importante no desenvolvimento lingüístico das crianças, é um período muito caro para as teorias em aquisição da linguagem. Por um lado, qualquer teoria em aquisição deve dar conta desse momento. Por outro, estudos detalhados da gramática inicial lançam luz sobre questões cruciais para a compreensão do processo de aquisição da linguagem, tais como, qual a extensão da gramática inicial da criança regida por princípios lingüísticos inatos, o ponto no qual os diversos princípios passam a ser operativos e o modo pelo qual os princípios relevantes interagem com a experiência lingüística da criança (Radford, 1990).

Voltando ao estudo de Brown, o Estágio II caracteriza-se pelo "primeiro brotar" de morfemas gramaticais na produção da criança. Surgem, nesse período, algumas preposições (especialmente *in* e *on*), ocasionalmente artigos e o verbo cópula, flexões nominais de plural e possessivo e flexões verbais de progressivo, passado e terceira pessoa do presente do indicativo. Nem todos os morfemas, entretanto, apresentam um desenvolvimento completo dentro do estágio. Em alguns casos, seu uso pleno só se dará no estágio V.

Perroni Simões e Stoel-Gammon (1977) estudaram o desenvolvimento da flexão verbal de pessoa em dados de fala de uma menina adquirindo português brasileiro, Carina, observada entre 2;1 e 2;8 de idade. A evolução do uso das desinências é analisada através de regras que apontam diferentes estágios de desenvolvimento do sistema flexional. Estes estágios partem do uso exclusivo pela criança da forma de terceira pessoa do singular para todas as pessoas do discurso, chegando à emergência da flexão de plural. Os dados do

estudo não mostram produção de marcas de plural a não ser pela forma *"vamos"*, sendo, então, o último estágio proposto, mas não atestado. O paradigma de desinência verbal é composto por três formas do indicativo: presente, perfeito e imperfeito, nas três conjugações; este é constituído de quatro pessoas, opondo primeira a segunda/terceira, com a distinção de número. Para as autoras, P2 e P3 formam, no padrão infantil nesta faixa etária, uma categoria única.

Os resultados do trabalho mostram quatro estágios de desenvolvimento para os verbos regulares. No primeiro, a criança utiliza a marca de segunda/terceira pessoa para todas as pessoas do discurso, sendo a interpretação de P1 feita a partir do contexto e do uso do pronome. Os estágios seguintes apontam a distinção da marca de primeira pessoa, que se desenvolve diferentemente para o presente (segundo estágio) e o perfeito (terceiro estágio). As autoras chamam atenção para o fato de que a distinção da primeira pessoa do perfeito não é regular como o do presente, já que o paradigma varia com a conjugação (*-ei [ey]* para primeira e *-i* para segunda e terceira). Finalmente, o quarto estágio marca a aquisição da marca de plural.

O processo de aquisição dos verbos irregulares se dá de outra forma, já que, diversamente dos verbos regulares, cuja produção têm por base um conjunto limitado de regras, devem ser memorizados. Seu desenvolvimento, então, se dá em cinco estágios:

1) a repetição da forma de terceira pessoa conforme o padrão adulto ("*fez*")[22];
2) aplicação de forma regular de terceira pessoa ("*fazeu*");
3) aplicação da regra de distinção de marca de P1 no presente ("*fazei*");
4) aplicação da regra de distinção de marca de P1 no perfeito ("*fazi*");

[22] Os exemplos são extraídos do texto.

5) memorização da forma irregular (“*fiz*”).

3.3 A natureza da gramática inicial

Uma das questões nucleares na área da aquisição da linguagem gira em torno de o quanto da linguagem inicial da criança é inatamente regida e o quanto emerge, em termos desenvolvimentais. Um estudo que aponta um interessante meio-termo no debate é o de Givón (1999).

O autor aborda o tema discutindo duas perspectivas diametralmente opostas em relação à questão. De um lado está a gramática gerativa chomskiana, uma visão algorítmica da linguagem em que regras da gramática operam como regras lógicas ou matemáticas. De outro está o conceito de gramática emergente, proposto por Hopper (1997 *apud* Givón, 1999), no qual as regras são inteiramente dependentes do contexto comunicativo. Enquanto naquela as regras são 100% rígidas, sem exceções e governadas por critérios de necessidade e suficiência; nesta, as regras são 100% flexíveis, sempre negociadas na ocasião, e sistematicidade e rigidez são vistas como uma “*miragem dos hábitos analíticos do lingüista*”. Givón examina os antecedentes lógicos e psicológicos das duas abordagens do ponto de vista da teoria das categorias. Em Chomsky, cuja posição é traçada a partir da doutrina platônica essencialista do significado, os indivíduos são membros ou não-membros de uma categoria, jamais situados parcialmente ao longo de um *continuum* de similaridade, sendo as categorias, portanto, discretas. O antecedente filosófico em Hopper é expresso por Wittgenstein, em *Investigações Filosóficas*, que

argumenta que o significado é profundamente dependente do contexto e dirigido pelo uso, sendo não discreto e envolvendo um *continuum* graduado.

Para Givón, as duas posições são extremadas e redutoras, e ele argumenta em favor de uma perspectiva de meio termo, reconhecendo que tanto Chosmsky quanto Hopper têm razão, embora parcialmente. O autor conclui que:

> *"Ambas as perspectivas são reducionistas num domínio complexo, que justamente clama por uma abordagem não-reducionista, uma abordagem que reconheça ambas como faces extremas de um todo complexo, interativo. A mente-cérebro tanto é modular como interconectada. Depende de ambas as capacidades, tanto de um domínio específico (sensório-motor) quanto de um domínio geral (atenção executiva, memória episódica). Parte dela está concluída no nascimento (portanto, 'inata'), a outra amadurece durante – e na interação com – a experiência de vida (emergente). Tem módulos altamente automatizados (invariáveis, rígidos, discretos, independentes do* input*, livres do contexto). Mas eles interagem com mecanismos mais flexíveis (sensíveis ao contexto, dependentes do* input*, escalares). Insistir em uma para a exclusão da outra é atalhar a enorme complexidade desse sistema quintessencialmente* ***híbrido****, a mente-cérebro"*. (Givón, 1999:107-8)[23]

Voltando à natureza da sintaxe, qualquer tentativa de descrição da gramática inicial deve dar conta do fato de que no período de surgimento das primeiras combinações de palavras, em torno dos 20 a 24 meses, descrito como o estágio I de Brown, a criança produz sentenças que, em geral, não possuem marca de flexão, nem verbos auxiliares, modais e cópula. Examinado sob a ótica de várias línguas (inglês, italiano, alemão, francês, português, entre outras), o debate aquisicionista em torno do assunto opõe duas hipóteses

[23] *"They are both reductionist perspectives on a complex domain that fairly cries for a non-reductionist approach that would recognize both extremes as facets of a complex, interactive whole. The mind-brain is* both *modular and interconnected. It depends on* both *domai-specific (sensory, motor) and domain-general (executive attention, episodic memory) capacities. Parts of it are mature at birth (thus 'innate'), others mature during – and interaction with – lifetime experience (emergent). It has highly automated (invariant, rigid, discrete, input-independent, context-free) modules. But they interact whit more flexible (context-sensitive, input-dependent, scalar) mechanisms. To insist on one to the exclusion of the other is to short-change the enormous complexity of this quintessentially* ***hibrid*** *system, the mind-brain"* [Tradução minha].

principais: a das pequenas orações e a da competência plena. De acordo com a primeira (Radford, 1990, 1997; Nicolau, 1998), as estruturas sintáticas na gramática inicial da criança têm um caráter estritamente pré-funcional ou léxico-temático. Essa perspectiva parece ser apoiada, sobretudo, por línguas de morfologia verbal pobre. A segunda sustenta que as categorias funcionais (seja em parte ou no todo) estão presentes desde cedo na gramática inicial da criança (Kato, 1995; Clahsen et al., 1993/1994; Guasti, 1993/1994; Guasti, 2002;).

Radford (1990, 1997) examinou uma série de dados disponíveis de crianças adquirindo inglês como língua materna, com idade entre 20 e 30 meses. O objetivo é descrever a gramática inicial em termos de: i) o ponto em que tem início o desenvolvimento da gramática; ii) a transição da fala que ele chama de *agramatical* para *gramatical*; iii) quais são e o que define cada estágio do desenvolvimento; e iv) a extensão do que é produto de conhecimento inato e do que é produto da experiência lingüística na gramática inicial.

Segundo o autor, é razoável supor que uma determinada estrutura particular reflete conhecimento sintático por parte da criança quando a estrutura é usada *produtivamente*, ou seja, quando a estrutura é aplicada a uma variedade de diferentes itens lexicais em cada posição da estrutura. Dessa forma, o que surge como pré-requisito para a aquisição da gramática é o desenvolvimento de um sistema de *categorias*. Pode-se pensar, então, o desenvolvimento da gramática em termos de aquisição de um conjunto formalmente diferenciado de categorias gramaticais (Radford, 1990:19). Afirma o autor que:

> *"Usando evidência morfológica, podemos dizer que temos clara evidência de que uma dada criança adquiriu uma dada categoria assim que a criança começa a adquirir uma ou mais flexões, as quais afixa a um item pertencente àquela categoria – contanto que ela sempre use a flexão relevante* produtivamente *(i.e. ligando-a a uma larga extensão de raízes, e não apenas usando-a em expressões fixas, aprendidas por memorização),* seletivamente *(e.g. ligando flexões nominais somente a raízes nominais e nunca a qualquer outro tipo de raiz),* contrastivamente *(e.g. afixando uma flexão de plural somente a formas plurais e não a formas singulares), e* apropriadamente *(i.e. usando uma flexão somente em contextos em que ela é requerida)"* (Radford, 1990:24)[24].

A hipótese de Radford é de que as estruturas sintáticas existentes na produção da criança são significativamente diferentes daquelas encontradas no inglês adulto. Ele assume, com base na Teoria dos Princípios e Parâmetros, que as lacunas presentes na fala das crianças nesse período do desenvolvimento se referem aos elementos funcionais das sentenças. Assim, as sentenças das crianças com menos de dois anos de idade seriam projeções dos núcleos lexicais principais (substantivo, verbo, adjetivo e preposição), não possuindo os núcleos funcionais (auxiliares, complementizadores, determinantes e partículas de caso) existentes na gramática do adulto.

O que interessa, nesse ponto, é que o autor faz algumas previsões sobre a produção infantil: a fala das crianças não apresentará os modais do inglês, as crianças não farão uso produtivo dos auxiliares não-modais *do*/*have*/*be* nesse estágio. Há ainda outras predições propostas pelo estudo, contudo, a principal, para o presente trabalho, é que as formas verbais não apresentarão marca de concordância.

[24] *"Using morphological evidence, we can say that we have clear evidence that a given child has acquired a given category once the child starts to acquire one or more of the inflections which attach to items belonging to that category – provided always that he is using the relevant inflections* productively *(i.e. attaching them to a wide range of stems, and not just using them in rote-learned 'set expressions'),* selectively *(e.g. attaching noun inflections only to noun stems and never to any other kind stem),* contrastively *(e.g. attaching a plural inflection only to plural forms and not singular forms), and* appropriately *(i.e. using an inflection only in contexts where it is required)"* [Tradução minha].

Esse estudo de Radford interessa aqui, sobretudo, porque ele foi replicado para o PB por Nicolau (1998). A autora quis verificar se dados do PB fornecem evidências para a hipótese de Radford (1990) de que a criança entre 20 e 24 meses de idade não domina ainda o sistema flexional da língua em aquisição. Para tanto, os dados de uma criança, na faixa etária analisada por Radford, foram contrastados quantitativamente com os de um adulto, a mãe da criança. Um dos argumentos de Radford considerado pela autora é de que a ausência do sistema flexional seria evidenciada: (1) pelo fato de as crianças produzirem sentenças que contêm formas verbais não-finitas (um núcleo composto por uma forma básica de infinitivo, ou gerúndio ou particípio), tanto em contextos de fala espontânea quanto de respostas a estruturas de pergunta com verbo finito, marcado com Tempo e Concordância; e (2) pelo fato de as formas finitas produzidas pela criança integrarem "fórmulas" ou enunciados "semi-prontos", ou também serem fragmentos de sentenças da fala do adulto, não podendo, por isso, ser analisadas como formas flexionadas.

A autora salienta que num estudo sobre aquisição do português é importante o tratamento detalhado das marcas flexionais, já que, diferentemente do inglês, é uma língua morfologicamente rica em desinências verbais. Além disso, esse detalhamento é imprescindível na discussão sobre o português estar ou não num processo de empobrecimento do sistema de flexões verbais.

As variáveis se concentram em examinar os contextos de formas finitas e não-finitas em relação ao tipo de forma verbal, à estrutura na qual a forma verbal está contida, ao contexto de produção das formas verbais e à representação da flexão.

Os resultados, num primeiro momento, mostram um certo pareamento em relação aos índices de formas com e sem flexão nos dados da criança (88% de formas com flexão) com os da mãe (97,4% de formas com flexão). No entanto, como salienta a autora, esses valores incluem as formas verbais consideradas por Radford como “aparentemente” flexionadas (fórmulas e fragmentos da fala da mãe). Excluindo esses dados, o índice de formas flexionadas na fala da criança cai substancialmente em relação ao da mãe: nas formas que apresentam flexão, aquelas consideradas aparentemente flexionadas, nos dados da mãe, têm índice de 44,8%, e da criança, 83,9%; as realmente flexionadas computam, na mãe, 52,6% do total, e na criança, 4,1%. Além disso, no total de formas “flexionadas” na fala da criança está computado um grande número de repetições das formas que a criança usa uma primeira vez. Isso faz com que o uso de formas produtivas caia ainda mais.

Com esses resultados, a autora conclui que a fala da criança apresenta (1) com relativa freqüência, formas sem flexão (infinitivo e gerúndio) no lugar daquelas que incluiriam verbos auxiliares flexionados (complexas); e (2) com “altíssima freqüência”, formas verbais apenas “aparentemente flexionadas”. Esses resultados, segundo a autora, mostram que a criança em questão ainda não adquiriu o Sistema de Flexão do PB, fato que evidencia, tal como propôs Radford, que o período entre 20 e 24 meses de idade da criança se caracteriza pela ausência de flexão.

Apesar da importância de tais achados para a descrição da fala inicial no PB, é preciso pensar que a teoria de Radford foi elaborada a partir do inglês, uma língua que difere substancialmente do PB em termos de morfologia flexional dos verbos. Além disso, o PB está passando por um processo de mudanças que alteram as flexões verbais na fala. A autora salienta esses fatos no artigo. Entretanto, essas duas condições fazem com que haja a

necessidade de discussões que permitam que se compreenda e descreva o sistema de flexão no PB para então se abordar o PB na perspectiva de Radford.

A segunda hipótese a que fiz alusão anteriormente é a da competência plena. Ela se opõe à alegação de que a estrutura da fala inicial no período de primeiras palavras seja desprovida do sistema funcional.

A evidência para esta abordagem se apóia, sobretudo, na distinção feita pelas crianças entre verbos finitos e infinitivos. Não entrarei em detalhes com relação à análise gerativista, já que é possível compreender que, se a hipótese da pequena oração é operante, espera-se que as crianças não façam distinção entre as duas formas e mostrem a mesma distribuição e posição estrutural para ambas.

Guasti (2002), tratando da perspectiva da evidência plena, comenta o argumento da distribuição das formas verbais em relação à negação. Em dinamarquês, holandês, alemão e francês, os verbos finitos precedem a partícula de negação, ao passo que os verbos infinitivos seguem-na. Guasti cita alguns trabalhos que, com suporte de análise quantitativa, mostram evidências robustas sobre o uso adequado dessas estruturas por crianças entre 1;8 e 2;3 de idade.

Guasti (1993/1994) é um dos estudos que contribuem com evidências quantitativas sobre o uso de verbos finitos e não finitos e uso de morfologia verbal de concordância por crianças pequenas. Baseada numa análise de dados naturalísticos de três informantes italianos entre 1;8 e 2;7 de idade, argumenta em favor de a estrutura da gramática inicial incluir categorias funcionais. Para tanto, apóia-se na percepção da criança

para diferentes contextos de uso dos dois tipos de verbos, na distribuição de pronomes clíticos e no conhecimento do sistema de concordância verbal pelas crianças nesse período etário.

Em relação à distribuição dos verbos finitos e não finitos, ponto crucial para determinar se as crianças fazem distinção entre os dois tipos de verbos, a autora comenta que, tal como na fala dos adultos, as raízes infinitivas na fala inicial das crianças são encontradas depois de verbos regentes, tais como modais, volitivos e aspectuais, e depois de preposições. As preposições tendem a ser omitidas na fala das crianças. Entretanto, em contextos que requerem formas infinitivas, as crianças não usam formas finitas, ou seja, *modais + V finito* ou *preposição + V finito* não são formas encontradas na amostra. Esses fatos, segundo a autora, atestam que as crianças são sensíveis às diferenças entre os tipos de verbos e também alocam os verbos corretamente em relação à gramática do adulto.

Sobre a colocação dos clíticos em relação às raízes finitas e infinitivas, a autora afirma que, na língua adulta, os clíticos recebem posições diferentes de acordo com a natureza do verbo: à esquerda dos finitos e à direita dos infinitivos. Os resultados mostram que, quando as crianças usam pronomes clíticos, eles aparecem em posição de acordo com a gramática do adulto.

O último argumento do estudo tem relação com o paradigma de concordância. Segundo a autora, as crianças italianas não só distinguem entre verbos finitos e infinitivos como também conhecem o paradigma de concordância bastante cedo. Os resultados mostram o uso de flexão de pessoa, especialmente no singular, sem erros sistemáticos de concordância. Além disso, mostram que as flexões de singular emergem

mais cedo que as de plural, fato relacionado não particularmente à flexão verbal, mas a uma manifestação mais tardia do fenômeno da pluralidade na fala das crianças.

Ainda sobre a natureza da gramática inicial da criança, apresentarei, a seguir, um debate, com base no inglês e no italiano, sobre a extensão das propriedades inatas e das propriedades aprendidas dessa gramática. A discussão gira em torno de dois modelos divergentes de descrição de aquisição da linguagem: de um lado, o modelo gerativista, parametrizado; de outro, o cognitivista, de processamento de informação.

O primeiro estudo considerado é o de Pizzuto e Caselli (1992), cujo objetivo é avaliar a plausibilidade do modelo inatista de fixação paramétrica assumido em Hyams (1986) em pesquisa feita com dados de onze crianças falantes de italiano com idades entre 1;10 e 2;4. As autoras analisam quantitativamente dados longitudinais do italiano de três crianças entre 1;4 e 3;0 de idade, focando três aspectos da fala dessas crianças: o paradigma morfológico verbal, os pronomes e os artigos. Examinaremos aqui somente o primeiro aspecto, que pretende analisar os padrões de emergência e desenvolvimento das flexões verbais no italiano.

A partir de uma abordagem empiricista, as autoras assumem que o processo de aquisição da morfologia verbal se dá de forma gradual. Para tanto, empregam no estudo um critério que estabelece o ponto de aquisição de cada morfema examinado. Esse critério, utilizado em outros trabalhos empiricistas, como o de Brown (1973), faz a distinção entre a primeira aparição do morfema e sua aquisição. O ponto de aquisição se dá no primeiro de três inquéritos consecutivos no qual o morfema foi produzido corretamente em ao menos 90% dos contextos em que é obrigatório.

Os resultados mostraram que apenas um pequeno conjunto das formas observadas atendia o critério de aquisição, ou seja, nenhum dos paradigmas flexionais examinados estava plenamente dominado até os três anos de idade dos sujeitos. Em duas das crianças as formas adquiridas eram as mesmas: primeira e terceira pessoas do indicativo, segunda pessoa do indicativo e do imperativo e terceira pessoa do indicativo da cópula; na terceira criança, somente a terceira pessoa do indicativo. Diante dos resultados, as autoras concluíram que o controle incompleto do sistema de flexão verbal constitui evidência contrária ao modelo inatista de relações gramaticais, um dos quais, examinado por Hyams (1986), sendo o de aquisição de concordância verbal.

Sobre os achados de Pizzuto e Caselli, Hyams (1992) argumenta que os resultados não são relevantes para o postulado inatista de um processo formal de aquisição, já que o fato de não haver marcas de flexão no plural pode indicar que as crianças não falem sobre coisas no plural, além disso, o fato de as crianças usarem sujeito no plural associado a uma forma verbal no singular indica evidência de que as crianças não dominam a morfologia de plural.

Rubino e Pine (1998) abordam o debate salientando que se por um lado ele explicita as dificuldades de diferenciar empiricamente entre as abordagens empiricista e nativista o desenvolvimento sintático ou morfológico das crianças, por outro, levanta a questão da validade de argumentos empíricos para explicar o conhecimento ou habilidade sintática baseados unicamente em baixas proporções de erros (Rubino e Pine, 1998:37).

O objetivo do estudo é refletir sobre o significado da proporção de erros produzidos por uma criança, examinando a validade dos argumentos apresentados para a regra de concordância entre verbo e sujeito. Os dados são de uma criança falante de PB com idade entre 3;02 e 3;04, obtidos em interações com a mãe, além do investigador.

As seqüências sujeito-verbo produzidas foram classificadas em corretas ou incorretas em relação à concordância. Os primeiros resultados mostraram percentuais de seqüências corretas bastante robustos (1420/1464 - 97%) e percentuais muito diferentes de seqüências no singular (1414/1464 – 96,6%) e no plural (50/1464 – 3,4%)[25].

3.4 Linguagem e socialização

Para as abordagens desenvolvimentistas, é muito cara a relação entre a linguagem e os processos sociais nos quais a criança está imersa, sobretudo no período inicial da aquisição. Uma linha de pesquisa que lança luz sobre essas questões é a da socialização da linguagem. Um dos argumentos é que a experiência social da criança funciona como suporte para a aquisição da língua. Nesta seção discuto alguns trabalhos centrados nessa área. Apresento também estudos que atestam que as crianças são sensíveis a fatores sociais.

Snow (1999) reflete sobre um duplo sentido para o termo *social* em relação à emergência da linguagem: por um lado, refere-se ao argumento de que as capacidades

[25] Mais detalhes do trabalho são apresentados na página 120.

sociais da criança são a fonte de emergência da linguagem; por outro, identifica o contexto no qual se dá a aquisição.

Em relação ao primeiro sentido, que diz respeito, sobretudo, ao período de emergência da linguagem, ou seja, o primeiro ano de vida da criança, a autora sustenta que, se a linguagem não se estabelece como um fenômeno desenvolvimental separado de outros domínios cognitivos, são necessárias certas condições (pré/não-lingüísticas) que sirvam de suporte para que o sistema da língua possa emergir. É preciso saber, então, em que domínio a criança possui habilidades precoces que sustentem esse emergir. O argumento da autora é de que esse domínio é social, relativo à capacidade pragmática precoce da criança de compreender a natureza intencional e interpessoal da comunicação.

A chave para a base pragmática da linguagem, e que dá sentido à afirmação de que a comunicação ocorre antes da linguagem convencional, é a *intersubjetividade*. Esse construto marca a capacidade (específica de nossa espécie) de compreender que o outro pensa como nós próprios, e interpretar o comportamento do outro como intencional. A realização da intersubjetividade pela criança abre caminho para a comunicação verdadeira, ou seja, a produção de sinais mediados pela expectativa de que eles serão interpretados tal como pretendido. A autora acrescenta que essa realização é responsável pelo interesse da criança nos significados convencionalizados da língua, um começo potencial da linguagem. Todo esse processo pragmático torna-se saliente a partir da necessidade das crianças pequenas de participarem de jogos interacionais, do engajamento em expressões comunicativas convencionalizadas, relativamente sem conteúdo (do ponto de vista semântico) e previsível, dentro de formatos sociais altamente restritos.

Em relação ao outro sentido do termo *social*, evidências sugerem que algumas características do ambiente lingüístico têm efeito sobre o desenvolvimento da linguagem, o que leva a pensar que esse desenvolvimento é fortemente dependente do possível lugar que a criança ocupa como agente comunicativo intencional. Dessa forma, no período inicial de aquisição, o desenvolvimento lingüístico é fortemente apoiado na ação do adulto de obter uma comunicação efetiva com a criança. Como diz a autora, "e*vidências consideráveis sugerem que o reconhecimento do adulto e a resposta à pretensão comunicativa da criança são demonstravelmente úteis para a criança adquirir a linguagem*" (Snow, 1999:267)[26].

Um trabalho que examina o processo de aquisição da língua pela criança do ponto de vista da socialização da linguagem é o de Ochs e Schieffelin (1997). Esse estudo trata, inicialmente, do modo como uma abordagem de socialização da linguagem pode colaborar para a compreensão mais ampla do desenvolvimento lingüístico.

O objetivo do modelo é explicar o desenvolvimento gramatical das crianças em termos de significados *indexicais* que as formas gramaticais possuem. Para tanto, baseia-se no pressuposto de que tais formas, em relação a cada comunidade, "*indexam situações culturalmente organizadas de uso"*, além disso, que "*os significados indexicais das formas gramaticais influenciam a produção e compreensão destas formas pela criança"* (*op. cit.*, p.60).

[26] *"Considerable evidence suggests that adult recognition of and responsiveness to children's communicative intents is demonstrably helpful to children in acquiring language"* [Tradução minha].

O primeiro ponto que as autoras discutem é se o desenvolvimento gramatical da criança depende ou não de fala simplificada. Essa questão, bastante discutida pelas diversas teorias de aquisição da linguagem, coloca uma interrogação sobre se a competência gramatical das crianças está relacionada com a participação delas em intercâmbios de fala simplificada, que teriam como objetivo facilitar a compreensão por parte das crianças. As autoras exploram esse ponto examinando de que forma culturas diferentes organizam a comunicação para e pelas crianças pequenas. Partindo do pressuposto de que os membros, em todas as sociedades, desejam transmitir suas intenções às crianças, o que as autoras percebem é que o que varia nas comunidades é o ponto em que a criança assume o papel de destinatário na interação. Segundo as autoras, em comunidades tais como as de classe média, brancas, dos Estados Unidos e Canadá, as crianças recebem desde o nascimento essa atribuição, quando as mães já começam a conversar com seus bebês. Em outras comunidades, as crianças pequenas não fazem o papel de destinatário até serem capazes de produzir palavras reconhecíveis na língua, como é o caso dos maias falantes de *k'iché,* de comunidades rurais e urbanas de Java, de comunidades tradicionais de Samoa Ocidental e dos kaluli de Papua-Nova Guiné. Ao contrário daquelas, em que os bebês são escolhidos desde cedo como destinatários preferenciais, nestas eles participam das interações como *ouvintes casuais.* O que acontece é que, mesmo assim, as crianças tornam-se falantes-ouvintes gramaticalmente competentes.

Outro ponto de interesse do estudo está relacionado à seguinte questão: *"Os sistemas culturais de crença, conhecimento e ordem social explicam, em parte, a aquisição de determinadas construções gramaticais por crianças pequenas?"*. Segundo as autoras, a organização cultural da linguagem pode influenciar profundamente a aquisição de algumas construções gramaticais. Isso por que as construções gramaticais estão intimamente ligadas

"*a normas, preferências e expectativas que organizam o modo como os membros devem agir, pensar e sentir em situações sociais"* (*op. cit.*, p.78).

As autoras mostram que a aquisição de uma determinada forma verbal não pode ser explicada simplesmente em termos de freqüência da forma no meio verbal da criança. Um exemplo comentado ocorre em comunidades samoanas tradicionais em relação ao verbo dêitico *sau* (vir), que não é produzido por crianças pequenas. Nessas comunidades há associação entre o movimento físico de ir e vir e as pessoas de *status* relativamente inferior. Assim, as pessoas de status superior impõem ações que exijam movimento às de *status* inferior. Dessa forma, as crianças pequenas são bombardeadas com formas imperativas do verbo *sau* (vir), sem, contudo, produzi-lo. Por outro lado, as crianças são encorajadas a pedir comida e outras coisas, e o fazem através do verbo *aumai* (dar/trazer). Como explicam Ochs e Schieffelin:

> *"Embora as crianças samoanas demonstrem compreender o verbo* sau *"vir" no início de seu desenvolvimento (por volta dos 19 meses), elas tendem a produzir o verbo dêitico* 'aumai *antes de produzir sau, e elas produzem* 'aumai *com uma freqüência muito maior que* sau" (*op. cit.*, p.78).

O trabalho de Andersen (1990), de cunho sociolingüístico, aborda outro ângulo da sensibilidade lingüística das crianças ao arcabouço social. A autora examina a variação de registro em relação a diferentes contextos de produção em 18 crianças entre 4 e 7 anos.

O instrumento de coleta era composto de três situações distintas em que as crianças deveriam manipular bonecos que caracterizavam diferentes personagens dentro da

mesma cena. As crianças eram estimuladas pelo entrevistador a encenar os diferentes personagens. Além disso, o entrevistador sugeria algumas situações de interação que proporcionavam uma contextualização maior da situação.

Os resultados demonstraram, de modo geral, que as crianças ajustam-se lingüisticamente a dimensões como sexo, idade e hierarquia social, manipulando recursos lingüísticos como indicativos de identidades sociais. Foram relatados ajustes de registro na fala das crianças em relação a aspectos fonológicos (entonação e volume), morfológicos (omissão de verbos, artigos e preposições marcando a fala dirigida à criança) e de léxico (relativamente a cada situação distinta).

Na literatura psicolingüística, o sistema de flexão verbal está amplamente descrito e explicado (não sem divergências teóricas). Os achados dos trabalhos relatados neste capítulo são de valia para outras áreas de estudo. Assim, o objetivo de abordar aqui questões teóricas da psicolingüística é ter suporte descritivo e analítico para os dados em estudo, promovendo uma compreensão mais ampla do fenômeno variável na fala da criança, em relação a outras abordagens. Tendo como suporte do trabalho tanto a sociolingüística quanto a psicolingüística é possível discernir, como afirmei anteriormente, entre variação sociolingüística e variação psicolingüística na fala das crianças.

4. METODOLOGIA

Este capítulo será dedicado à descrição dos procedimentos metodológicos implicados na pesquisa. Inicialmente, tratarei dos dados que compõem a amostra, características dos informantes, das coletas e o modo como foi montado o *corpus*. A seção seguinte apresenta os grupos de fatores envolvidos na análise, tanto os que foram relevantes estatisticamente quanto os que não foram, com vistas a expor as reflexões que nortearam a pesquisa desde seu início. Por fim, na última seção, comentarei o modelo de análise estatística adotado.

4.1 Os dados

Os dados que integram a amostra provêm de três fontes distintas. Os das crianças maiores, que compõem as faixas etárias de 4 a 8 anos, foram retirados do Banco de Dados do Projeto Desenvolvimento da Criança em Fase de Letramento (DELICRI). Os

dados do menino menor (Art) foram coletados por mim e os dados da menina menor (Eri) fazem parte de outro momento do banco de dados do DELICRI. Todas as coletas são longitudinais, mas com basicamente duas rotinas diferentes de entrevistas[27]. Preciso salientar que nenhuma das coletas foi produzida especificamente para análise do fenômeno de concordância verbal variável. Entretanto, como o objetivo geral do trabalho era a descrição das formas verbais variáveis produzidas por crianças (e verbos são uma constante na fala das crianças desde cedo), a utilização desses dados era pertinente. Além disso, em trabalho anterior (Soares, 2001) já havia examinado uma parte dos dados do menino menor e também os de uma das crianças maiores. Assim, decidi por estender o *corpus* nos mesmos conjuntos de dados.

A amostra que compõe as faixas etárias de 4 a 8 anos é constituída pela totalidade da amostra longitudinal do banco de dados do Projeto DELICRI. As informações aqui prestadas foram extraídas de Guimarães (1995). O projeto foi constituído com o objetivo de permitir uma larga abrangência de estudos do desenvolvimento da linguagem da criança nas faixas etárias alcançadas, em áreas distintas tais como fonologia, sintaxe, semântica, pragmática e discurso. A amplitude desses fins fez com que o banco fosse montado a partir de dois tipos de coleta, uma transversal e uma longitudinal. Como já dito anteriormente, no presente trabalho me vali desta última, já que, de um modo geral, a coleta longitudinal pode propiciar a observação de estágios de desenvolvimento das crianças.

A coleta longitudinal prolongou-se de 1992 a 1996, contando inicialmente com doze crianças das quais sete atingiram o término do trabalho. As entrevistas tinham

[27] As entrevistas não seguem o modelo de entrevista sociolingüística proposto por Labov.

intervalos de 2 a 3 meses e uma duração média de 10 minutos, totalizando aproximadamente 16 horas de coletas. Eram organizadas seguindo quatro rotinas:

- interação entrevistador/informante;
- relato pessoal oral e escrito;
- relato de história em seqüência.
- narrativa ficcional

A primeira das rotinas acima é de conversa informal entre o entrevistador e o informante; na segunda, a criança é solicitada a narrar um fato ocorrido com ela e, posteriormente, a escrever e ler o relato; na terceira, o entrevistador apresenta um pequeno livro de gravuras em seqüência e sem texto, pedindo à criança que conte a história que há ali; na última, é pedido à criança que ela conte uma história (ficcional) qualquer.

É preciso salientar que essa coleta se deu a partir de uma escola particular de Porto Alegre, fato que faz com que a maioria das crianças seja pertencente à classe econômica caracterizada como A pelo projeto. Um dos informantes é filho de pai uruguaio e freqüenta a casa do pai nas férias, daí ser considerado bilíngüe. O perfil dos informantes é apresentado no quadro abaixo:

Quadro 1 – Perfil sócio-econômico dos informantes do projeto DELICRI

Origem	**Contato L2**	**Escola**	**Categoria Social**
Porto Alegre: 85,7%	Nenhum: 57,1%	Particular: 100%	A: 57,1%
Interior: 14,3%	Casa assistemático: 14,3%		B: 26,6%
	Escola: 14,3%		C: 14,3%
	Bilíngüe: 14,3%		

A categorização sócio-econômica estabelecida para o projeto é a seguinte:

> Classe A: nível superior – escolaridade superior de pelo menos um dos pais; profissional liberal, professores de 3º grau, empresários, dirigentes, bairro de moradia considerado A na cidade, escola particular.
>
> Classe B: nível médio – escolaridade 1º grau completo ou 2º grau; funcionário público, comerciante, técnico, bancário, escriturário, professor de escola secundária ou primária, bairro de moradia considerado de classe média, escola pública.
>
> Classe C: nível baixo – pais com 1º grau incompleto ou sem escolaridade; empregado doméstico, pedreiro, faxineiro, motorista, serviços gerais, bairro considerado pobre, escola pública. (cf. Guimarães, 1995:19)

Os dados foram transcritos ortograficamente com algumas notações fonológicas. Relativamente à CV, há identificações tais como:

nós chegamo-0s

na qual o número zero associado à marca de plural indica que ela não foi produzida pela criança, e

lavamo-0s as mão-0s lanchemo-0s [: lanchamos] (…) *e senta [: sentamos]*

em que a forma preconizada pela gramática é colocada entre colchetes na seqüência da forma produzida pela criança.

A amostra é composta por quatro meninas e três meninos com idades que variam de 4;03 a 9 anos de idade. Como somente uma das crianças tem inquérito com nove anos de idade, esta faixa etária não foi considerada na análise. Apresento abaixo um quadro geral dos informantes desse *corpus* em relação a idade e categoria sócio-econômica.

Quadro 2 – Informantes da coleta longitudinal do projeto DELICRI, idade e categoria sócio-econômica

Nome	**Idade**	**Categoria sócio-econômica**
Alexandra	4;08 a 8;05	A
Camila	4;11 a 8;09	C
Carmela	4;03 a 8;05	B
Gabriel	5;05 a 8;06	A
Matheus	6;02 a 9;00	A
Natália	5;04 a 8;10	A
Rodrigo	5;05 a 7;07	B

A segunda coleta a que me referi anteriormente foi toda feita por mim entre outubro de 1998, estando o menino na época com 2;9 de idade, e maio de 2005, quando ele contava com 9;3 meses de idade, embora a última coleta tenha um intervalo de nove meses da coleta anterior em razão de um problema de saúde da criança. Os intervalos entre as entrevistas eram quinzenais até os quatro anos da criança e mensais dos quatro aos cinco, época prevista inicialmente para término do trabalho de coleta. Contudo, como havia a possibilidade de as entrevistas serem retomadas, pela disponibilidade tanto da criança e seus pais quanto minha, decidimos (eu e a professora Luciene Simões, então coordenadora do Projeto DELICRI) retomar as visitas à casa do menino, em intervalos quadrimensais, até que ele completasse nove anos.

A rotina das entrevistas não previa qualquer tipo de elicitação de formas. Eu passava um turno inteiro (normalmente o da manhã) brincando com a criança e gravava em áudio uma hora desta interação. Posteriormente, procedi à transcrição ortográfica das coletas. Em relação a informações sócio-econômicas, o menino tem ambos os pais com

escolaridade superior, estuda desde a pré-escola em colégio particular e mora em um bairro de classe média de Porto Alegre. O pai é professor em escola pública e a mãe trabalha com artesanato em papel, ministrando cursos e produzindo para empresas.

Após ter em mãos as transcrições das coletas, me detive em selecionar e conferir acusticamente nas fitas de áudio todos os dados considerados, descartando os que não eram passíveis de confirmação, por falta de qualidade da gravação. Considerando as seis pessoas gramaticais e oito faixas etárias, tal como era o objetivo no início do trabalho, a quantidade de ocorrências se mostrou absurdamente grande (a média era de 300 a 500 por inquérito), além de boa parte serem dados provavelmente não variáveis (os casos de primeira e terceira pessoas do singular).

Em vista disso, codifiquei inicialmente o primeiro inquérito de Artur (aos 2;9) e o total dos inquéritos de Camila e rodei as percentagens para definir sobre o restante dos dados. Os resultados desta primeira análise podem ser vistos em 6.1 adiante. A partir deles, decidi por codificar os dados das crianças de 4 a 8 anos somente em relação a primeira e terceira pessoas do plural, já que não havia variação que justificasse uma análise estatística nas outras pessoas gramaticais.

Como a idéia inicial era tratar também os dados da menina menor, cujas coletas tiveram início numa idade anterior à do menino, fiz um recorte nos inquéritos que já estavam transcritos, conferi acusticamente, selecionei os dados e codifiquei um inquérito relativo à faixa etária de 2;0 a 2;5 anos (o único disponível nessa faixa etária) e três

relativos à faixa etária seguinte, de 2;6 a 2;11. Todos esses dados foram codificados relativamente às seis pessoas gramaticais e às quinze variáveis independentes.

As coletas dessa menina foram feitas em casa, em sessões de brincadeira, sendo a interação gravada em áudio e posteriormente transcrita ortograficamente. O pai tem curso superior e exerce a profissão de professor de educação física. A mãe tem escolaridade média. A criança freqüenta escola de educação infantil, vivendo em ambiente letrado.

Para compor a faixa etária de 3 anos, comecei por codificar o primeiro inquérito do menino nas seis pessoas gramaticais. Como o volume de dados era muito grande relativamente a um único inquérito, codifiquei a entrevista seguinte considerando somente P4 e P6. Um consulta[28] por pessoa gramatical nessa faixa etária nesse momento da pesquisa mostrou que não havia dados suficientes sequer para analisar P4 e P6.

Como até esse ponto já contava com 2589 dados codificados, já não havia tempo para recompor as três faixas etárias iniciais. Assim sendo, decidi por fazer uma análise consistente sobre as DNPs P4 e P6 nas faixas etárias de 4 a 8 anos e relatar os resultados das faixas anteriores como indicadores de caminhos e auxiliares em tomadas de decisões para pesquisas futuras. Assim sendo, a amostra foi composta do seguinte modo:

[28] A consulta é aquela disponível no programa Acces. Este disponibiliza os dados de qualquer cruzamento que se queira entre variáveis previamente escolhidas.

Quadro 3 – Composição da amostra relativamente a idades e faixa etária

Informante	Inquéritos	Idades	Faixas Etárias
Alexandra	Ale01 a Ale18	4;08 a 8;05	4 a 8
Camila	Cam01 a Cam14	4;11 a 8;09	4 a 8
Carmela	Car01 a Car18	4;03 a 8;05	4 a 8
Gabriel	Gab01 a Gab13	5;05 a 8;06	5 a 8
Matheus	Mat01 a Mat13	6;02 a 9;00	6 a 9
Natália	Nat01 a Nat15	5;04 a 8;10	5 a 8
Rodrigo	Rod01 a Rod08	5;04 a 7;07	5 a 4
Artur	Art01, Art06 e Art07	2;9, 3;0 e 3;0	2, 3 e 3
Érica	Eri33, Eri36, Eri41 e Eri 46	2;5, 2;6, 2;7 e 2;9	d, 2, 2 e 2

Em termos lingüísticos, alguns dados precisaram ser descartados:

i) verbos *ter*/*ver* no presente do indicativo, já que as formas singular e plural são homófonas, sendo distinguidas somente por acento na escrita;

ii) formas de gerúndio;

iii) casos em que a forma verbal produzida pela criança é uma resposta equivalente a uma pergunta sim/não em outras línguas[29];

iv) estruturas formulaicas invariáveis: *tá bom, era uma vez*;

v) casos de sujeito nulo em que não era possível atribuir a desinência, sem dúvida, ao pronome *nós* ou à forma inovadora *a gente*;

vi) casos de repetição idêntica pela criança sem a interposição de outra fala (da criança ou do entrevistador).

A seguir apresentarei os grupos de fatores que compuseram a análise.

4.2 O envelope de variação

[29] Sobre isso ver Simões (1997).

4.2.1 A variável dependente

Em razão do grande número de variantes que compunham inicialmente a variável dependente (as desinências relativas a cada pessoa gramatical), resolvi desdobrá-la em uma variável dependente com seis variantes, considerando as seis pessoas gramaticais, e uma variável independente, que chamei de *realização da forma verbal*, abrangendo todos os casos de presença e ausência de marca de CV. Como cada código da variável independente estava associado a uma única variante, produzindo o que o pacote VARBRUL chama de *knockout*, esta variável independente foi posteriormente reorganizada em relação à variável dependente para as rodadas individuais das DNPs P4 e P6. Tais variáveis foram compostas do seguinte modo:

Variável dependente

P1 – 1
P2 – 2
P3 – 3
P4 – 4
P5 – 5
P6 – 6

Realização da forma verbal

P1 →	marca (categórica) ..... c – (eu sei) ausência de marca ..... a – (eu aprendeu) supergeneralização ..... g – (eu sabo)
P2 →	marca não padrão ..... v – (tu recebesse, fizesse) marca padrão ..... p – (tu recebeste, fazes) ausência de marca ..... n – (tu fez, recebe)
P3 →	marca ..... h – (ele sabe, faz) marca de P6 ..... i – (a bundona delas eram)
P4 →	-mo (não padrão) ..... u – (nós andamo) -mos (padrão) ..... s – (nós andamos) ausência de marca ..... z – (nós anda/ a gente anda)

P5 → marca ..... k – (vocês recebem)
ausência de marca ..... y – (vocês recebe)

P6 → marca ..... m – (eles foram, fizeram)
ausência de marca ..... x – (eles foi, fez, canta)

Esse grupo de fatores visava a contemplar todas as formas verbais produzidas pelas crianças. Por isso, além de abranger a variação descrita pela literatura sobre dados de adultos, alguns códigos foram acrescidos ao longo da codificação, quando uma desinência produzida não estava prevista na variável.

A codificação da primeira pessoa do singular leva em consideração os casos de ausência de marca (a forma neutra de terceira pessoa) e os de supergeneralização (verbos irregulares flexionados tal qual o paradigma regular). A da primeira pessoa do plural prevê análise distinguindo as formas padrão e não padrão. Elas serão reunidas na rodada que opõe a presença à ausência da marca. Na terceira pessoa do plural, mesmo que a literatura aponte variação fonológica na aplicação da marca (presença ou não da nasalidade e/ou mudança na qualidade da vogal temática), não há previsão de distinção dessa natureza no presente trabalho, já que não me senti segura em confirmar vários casos, com os aparelhos de que dispunha e com a qualidade das fitas de áudio a partir das quais compus o *corpus*. As demais pessoas gramaticais serão examinadas levando em conta presença *versus* ausência de marca.

4.2.2 As variáveis lingüísticas

O envelope de variação continha, inicialmente, além daquela já citada, quatorze variáveis independentes. Comento a seguir as variáveis internas que foram consideradas na pesquisa.

Tempo verbal

O objetivo desta variável é medir a influência dos tempos verbais na manutenção ou não das desinências número-pessoais na fala das crianças. Zilles et al. (2000) mostram, em relação à DNP-P4, que os tempos verbais cujas formas alvo produziriam proparoxítonas desfavorecem fortemente a manutenção da desinência. Por outro lado, as formas verbais mais freqüentes na fala – a perífrase *Vamos+infinitivo*, presente e pretérito perfeito – favorecem o apagamento do /s/. Em relação à DNP-P6, Batista e Zilles (2005) apontam o perfeito como favorecedor da marca enquanto que as outras formas verbais consideradas no estudos a desfavorecem.

Sigo, aqui, a codificação proposta por Amaral (2003), acrescentando a forma *ir + infinitivo* e modificando o código "@" de "outros verbos" para "locuções verbais". A codificação adotada foi a seguinte:

P – presente do indicativo
M – pretérito imperfeito do indicativo
T – pretérito perfeito do indicativo
F – futuro do pretérito do indicativo
E – futuro do presente do indicativo
S – presente do subjuntivo
R – pretérito imperfeito do subjuntivo
U – futuro do subjuntivo
V – infinitivo
I – imperativo
Y – ir + infinitivo
@ – locução

Paralelismo Formal

Esta é uma variável freqüentemente examinada em trabalhos sobre concordância, tanto nominal quanto verbal (Naro, 1981; Scherre e Naro, 1993; Loregian, 1996; Vieira, 1997; Amaral, 2003, entre outros). Ela verifica a tendência, em seqüências discursivas, à ocorrência de repetição de formas ou marcas usadas anteriormente pelo falante. Conforme Scherre e Naro (1993), o princípio do paralelismo formal traduz "*uma tendência de marcas levarem a marcas e zeros levarem a zeros*" (cf. Poplack, 1980 *apud* Scherre e Naro, 1993), ou ainda, "*em termos mais gerais, uma tendência de formas gramaticais semelhantes ocorrerem juntas*" (cf. Schiffrin, 1981 *apud* Scherre e Naro, 1993).

A atuação dessa variável, segundo os autores, se dá tanto no nível clausal, com as marcas do sujeito influenciando as marcas no verbo, quanto no nível do discurso, em que a marca de flexão (ou ausência dela) em uma forma verbal atua sobre a presença ou ausência da marca na forma verbal seguinte numa série discursiva. Para os autores, a definição de construção em série se atém a dois critérios: o primeiro é a necessária co-referência entre a construção analisada e a construção anterior; o segundo é que elas não podem estar separadas por mais de dez sentenças ou pelo discurso de um interlocutor.

Neste trabalho me deterei no paralelismo formal de nível discursivo, seguindo a mesma codificação de Loregian (1996) e Amaral (2003), que examinaram, ambos, o impacto dessa variável na concordância verbal de segunda pessoa do singular e que aqui se estenderá às outras pessoas do discurso. Há, contudo, uma pequena diferença na codificação da mesma nos dois trabalhos citados. Amaral utiliza a codificação proposta por Loregian introduzindo uma modificação nos dois últimos fatores. Para o autor, T e S representarão a opção, respectivamente, por 100% de aplicação e 100% de não aplicação da

regra, sendo, posteriormente, ambos os fatores, retirados da análise (já que o VARBRUL só trabalha dados variáveis). Essa codificação se faz importante aqui, principalmente pelos casos de regra categórica, na primeira pessoa do singular de presença da marca e na terceira do singular de ausência da marca. Utilizarei aqui a codificação de Amaral (2003), composta de seis fatores:

I – verbo em construção isolada
P – primeiro verbo de uma série
M – mistura de marcas em que o elemento anterior ao elemento analisado é marcado
N – mistura de marcas em que o elemento anterior ao elemento analisado é não marcado
T – verbo em uma seqüência com todas as marcas de concordância
S – verbo de uma seqüência sem marcas de concordância

Discurso Reportado

Com esta variável minha intenção é examinar se o discurso reportado direto inserido na fala, principalmente nas narrativas das crianças, teria algum efeito sobre o uso variável das DNPs. Zilles e Faraco (2002:15) assinalam que:

> *"não se pode tomar os enunciados dos informantes simplesmente como dados homogênceos, isto é, não se pode desconsiderar que, na composição do dizer dos informantes, pode estar explicitamente presente a voz de outros (o chamado discurso reportado ou citado) e que isso pode motivar o uso diferenciado de variantes".*

Em razão dos tipos de coleta, estruturei a variável pensando nas possibilidades de discursos reportados que seriam encontrados na fala das crianças. Assim sendo, ela abrange os seguintes fatores:

H – discurso reportado direto de histórias
D – discurso reportado direto de outra pessoa
S – discurso reportado de si mesmo
N – discurso não reportado

Contexto Seguinte

Este grupo de fatores quer examinar se o contexto fônico seguinte à DNP exerce influência na manutenção ou não da marca, já que a terminação da desinência em vogal ou consoante pode provocar um processo de ressilabificação dependendo do contexto seguinte. Assim, é uma variável cujo domínio deverá ser testado especialmente na manutenção ou não do /s/ nas formas de primeira pessoa do plural em função do possível processo de ressilabificação. Os códigos adotados distinguem as vogais átonas, as vogais tônicas, as consoantes e as pausas.

vogais	consoantes
a – /a/ átono	b
1 – /a/ tônico	c
e – /e/ átono	d
2 – tônico	f
i – átono	g
3 – tônico	j
o – átono	l
4 – tônico	m
u – átono	n
5 – tônico	p
	q
	r
	s
	t
	v
	x
	z

0 – pausa

Saliência Fônica

Esta variável, proposta inicialmente por Lemle e Naro (1977), quer examinar a influência da saliência fônica na manutenção da desinência relativa à CV, tendo por pressuposto que uma maior quantidade de material fônico na oposição entre a forma de terceira pessoa do singular e a forma alvo do plural torna o dado lingüístico mais saliente,

conduzindo a uma maior aplicação da marca de concordância. O exame se dá a partir de uma escala, indo dos primeiros níveis, menos salientes, aos mais salientes.

A codificação aqui adotada é aquela contida em Almeida (2005), tanto para o exame da DNP-P4 quanto da DNP-P6.

DNP-P4

1 – oposição -V/-Vmos não é tônica nas duas formas (*falava/falávamos*)
2 – oposição -V/-Vmos é tônica em uma das formas (*fala/falamos, trouxe/trouxemos*)
3 – oposição -V/-Vmos é tônica nas duas formas (*está/estamos, tem/temos*)
4 – oposição -V/-Vmos é tônica nas duas formas e a terceira pessoa do singular tem uma semivogal que não aparece na primeira pessoa do plural (*comeu/comemos, partiu/partimos, vai/vamos, foi/fomos*)
5 – oposição -V/-Vmos é tônica nas duas formas e a vogal acentuada é diferente *(falou/falamos)*
6 – oposição -V/-Vmos é tônica nas duas formas e a vogal acentuada é diferente do caso único: *é/somos*

DNP-P6

1 – As formas verbais apresentam a mesma posição da sílaba tônica (paroxítona) e a diferença entre singular e plural existe devido ao processo de nasalização da vogal final /e/ na forma verbal singular que passa a ser um ditongo nasalizado /ey/ na forma verbal plural. Ex.: *come/comem, fale/falem*.
2 – As formas verbais apresentam a mesma posição da sílaba tônica e a diferença entre singular e plural existe devido ao processo de nasalização da vogal final /a/ na forma verbal singular que passa a ser um ditongo nasalizado /ãw/ na forma verbal plural. Ex.: *fala/falam, ia/iam.*
3 – A diferença entre plural e singular reside no acréscimo de uma vogal final átona, possivelmente nasalizada. Ex.: *faz/fazem, quer/querem*.
4 – A sobreposição de raiz e desinência, com acento, marca a diferença entre singular e plural nesse caso. Ex.: *dá/dão, está/estão, falará/falarão.*
5 – A diferença entre as duas formas não está na posição do acento na palavra, mas há uma diferença fonológica maior do que a simples nasalização: a troca do ditongo da sílaba final da forma verbal do singular para a permanência de uma das vogais do ditongo ou a substituição dele por uma nova vogal, seguida do acréscimo de uma sílaba. Ex.: *sumiu/sumiram, comeu/comeram, falou/falaram.*
6 – Existe a mudança da posição da sílaba tônica e da raiz da forma verbal, seguida do acréscimo de duas sílabas. Ex.: *fez/fizeram, teve/tiveram*.
Casos únicos que não se enquadram nos níveis já descritos acima:
7 – Formas totalmente distintas para singular e plural): *é/são.*
8 – Caso único: *foi/foram.*
9 – Caso único: *veio/vieram.*

Os códigos foram os seguintes:

0 – não se aplica (casos de P1, P2 e P3)

A – (*falava/falávamos*)

B – (*fala/falamos, trouxe/trouxemos*)
C – (*está/estamos, tem/temos*)
D – (*comeu/comemos, partiu/partimos, vai/vamos, foi/fomos*)
E – *(falou/falamos)*
F – caso único: *é/somos*

1 – *come/comem, fale/falem.*
2 – *fala/falam, ia/iam.*
3 – *faz/fazem, quer/querem.*
4 – *dá/dão, está/estão, falará/falarão.*
5 – *sumiu/sumiram, comeu/comeram, falou/falaram.*
6 – *fez/fizeram, teve/tiveram.*
7 – caso único: *é/são.*
8 – caso único: *foi/foram.*
9 – caso único: *veio/vieram.*

Tonicidade

Este grupo de fatores examina a relação entre a sílaba tônica do verbo e a presença ou não da marca de concordância, sobretudo pensando na tendência do PB falado de evitar palavras proparoxítonas (ver Bortoni, 1984; Zilles et al., 2000 e Batista e Zilles, 2004). Além disso, pensando em termos de saliência fônica, minha expectativa é de que as formas verbais que contém a desinência na sílaba tônica, ou seja, monossílabos tônicos e oxítonas, favoreçam a retenção da marca de concordância. A codificação foi feita do seguinte modo:

M – monossílabo tônico
O – oxítona
P – paroxítona
R – proparoxítona

Tipo de Verbo

A variável tipo de verbo foi proposta com o intuito de averiguar a relação entre características de maior ou menor transitividade do verbo e a aplicação da marca de CV.

Monguilhott e Coelho (2002) relacionam verbos que selecionam argumentos com traço [+] humano como argumento externo com a posição que o sujeito ocupa em relação ao verbo. Os transitivos e intransitivos constituem preferencialmente construções de sujeito anteposto, favorecendo, portanto, a marca.

A expectiva em relação aos verbos ergativos e aos verbos *Ter/Haver* (aqui considerados em estruturas apresentativas), ao contrário, é que desfavoreçam a aplicação da desinência, por constituírem, com maior freqüência, sentenças com sujeito posposto, dado o caráter de inacusatividade desses verbos[30].

Em relação ao verbo cópula, as autoras o distinguiram dos outros verbos ergativos, porque aquele seleciona um argumento diferente destes, qual seja, uma *pequena oração*. Seus resultados mostraram esse verbo como o maior responsável pela manutenção da marca de concordância.

O verbo *Vamos* foi considerado separadamente em virtude de compor um tipo de estrutura com peculiaridades em relação à CV (Ver Zilles et al., 2000; Zilles e Batista, 2005), quais sejam, a de promover aplicação de marca, na distinção entre presença *versus* ausência da DNP-P4, e o apagamento do –s na distinção padrão *versus* não padrão, em virtude do processo de gramaticalização do verbo *Ir*.

A hipótese a ser testada é a de que os verbos ergativos e *Ter/Haver* promovam a não aplicação da marca, enquanto que os verbos transitivos, intransitivos e a cópula provoquem a aplicação da marca.

[30] Ver seção 1.2 deste trabalho.

A codificação considerada foi a seguinte:

T – transitivo
I – intransitivo e reflexivo
E – ergativo
C – cópula
H – *Ter*/*Haver*
V – *Vamos*

Tipo de Sujeito

Esta variável avalia se a forma como o pronome é expresso ou o fato de ser elíptico influencia na produção da desinência verbal das crianças. Zilles (2000) afirma que relativamente ao sujeito do tipo SN pleno ou pronome indefinido existe maior tendência de este estar colocado à direita do verbo, construção que favorece a não aplicação da marca, ao contrário do que ocorre com sujeitos do tipo pronome pessoal e demonstrativo. Rodrigues (1992) mostra que o sujeito nulo favorece a aplicação da marca, enquanto o sujeito não pronominal (SN pleno) desfavorece. A codificação vislumbra as formas de sujeito nas seis pessoas gramaticais:

s – SN pleno
p – pronome reto
i – pronome indefinido
n – nulo expletivo
z – nulo de pronome
g – “a gente”
y – nulo de “a gente”
d – pronome demonstrativo
j – uso genérico

Posição do sujeito em relação ao verbo

O objetivo desta variável é averiguar de que modo se dá a atuação da posição que o sujeito tem em relação ao verbo sobre a manutenção ou não da marca de CV. Vários trabalhos na área (Lemle e Naro,1977; Bardem, 2004; Batista e Zilles, 2005, entre outros) atestam que o sujeito anteposto ao verbo favorece a aplicação da marca, ao passo que o sujeito posposto desfavorece.

Além disso, a hipótese inacusativa sugere que os inacusativos são verbos cujos sujeitos possuem, na estrutura de superfície, características de objeto, ou seja, semanticamente a não agentividade, e gramaticalmente a não regência de concordância[31]. Dessa forma, os verbos dessa classe se adaptam à posição pós-verbal.

Assim, a hipótese a ser testada aqui em relação à variável posição do sujeito é de que os sujeitos antepostos favoreçam a aplicação da marca de concordância ao passo que os pospostos a desfavoreçam. Os fatores que compõem o grupo dão conta também da possível atuação de material interveniente entre o sujeito e o verbo. Dessa forma, o grupo conta com os seguintes fatores:

A – anterior adjacente
D – anterior distante
F – posterior distante
P – posterior adjacente
Z – não se aplica (casos de sujeito nulo)

Animacidade do Sujeito

Com essa variável quero verificar se o *status* de animacidade do sujeito influencia a produção de desinência verbal.

[31] Ver seção 1.2 deste trabalho.

A – sujeito [+] animado
M – sujeito [–] animado
N – não se aplica (casos de sujeito expletivo)

Focalização

A variável focalização quer averiguar se o modo como os eventos se dispõem na estrutura narrativa em termos de foco discursivo influencia a presença ou não de marca de concordância no verbo. As categorias consideradas são:

f – foreground
b – background

4.2.3 As variáveis sociais

As variáveis sociais que compuseram a análise são as seguintes:

- *Gênero*: feminino e masculino
- *Faixa etária*: 2;0 a 2;5; 2;6 a 2;11; 3;0 a 3;11; 4 anos; 5 anos; 6 anos; 7 anos e 8 anos
- *Tipo de coleta*: relato pessoal, narrativa, história em seqüência, leitura, conversa informal, leitura de desenho, música.

Em relação ao grupo de fatores gênero o objetivo é observar se há distinção de comportamento lingüístico relativamente à aplicação de marcas de CV entre os meninos e meninas observados, já que a literatura variacionista aponta, por vezes, tal distinção.

A variável faixa etária quer averiguar o efeito do desenvolvimento lingüístico da criança, lembrando que, mesmo que a escolarização não tenha sido uma variável considerada aqui, ela está cruzada com a faixa etária, considerando que o período de educação formal na escola tem início por volta dos seis anos de idade. Por outro lado, entre dois e três anos, a criança ainda está focada na aquisição de marca de CV, o que pode influir no favorecimento do uso da desinência.

A variável tipo de coleta foi inserida com o propósito de medir a influência de alterações estilísticas no uso das DNPs. Os diferentes fatores dizem respeito às diferentes situações de interação e elicitação dos dados nas entrevistas. A hipótese a ser testada é de que os fatores relacionados à coleta de narrativas – narrativa ficcional, história em seqüência e leitura –, em virtude de um caráter mais formal desses fatores, relacionado ao modo como se deu a entrevista, favoreceriam mais a manutenção da marca que a conversa informal e o relato pessoal.

4.3 O modelo analítico

O arcabouço teórico-metodológico que orienta essa pesquisa é o da Sociolingüística Quantitativa, a qual tem como projeto a descrição do uso da linguagem nos termos da variabilidade inerente ao sistema lingüístico. Para tanto, utiliza métodos estatísticos que avaliam a influência probabilística dos diversos grupos de fatores evocados na análise sobre a variável dependente. Scherre (1998:43) explica:

> *"uma das questões centrais da metodologia variacionista consiste em desenvolver ou definir modelos matemáticos que sejam capazes de associar adequadamente pesos relativos ou probabilidades aos diversos*

> *fatores de cada variável independente ou grupo de fatores, a fim de que se possa medir a influência que cada um destes fatores exerce sobre a presença de uma ou outra variante de um determinado fenômeno lingüístico ou, em outras palavras, de uma dada variável dependente”*.

A análise estatística utiliza o pacote de programas VARBRUL (Pintzuk, 1988), que estabelece, através de pesos relativos, os condicionantes relativos ao fenômeno lingüístico em estudo. A compreensão interpretativa de tais pesos relativos se dá em termos de um ponto de neutralidade (0,50), sendo os pesos abaixo desse índice indicativos de desfavorecimento da variável independente em análise e os pesos acima de 0,50 indicativos de favorecimento.

5. ANÁLISE E DISCUSSÃO DOS RESULTADOS

Meu objetivo neste capítulo é descrever, analisar e discutir os resultados obtidos. Por uma questão organizacional, ele será dividido em três blocos. No primeiro, tratarei dos resultados gerais, iniciando pela análise preliminar, seguida pelos resultados gerais que englobam as três pessoas do singular e duas do plural consideradas na pesquisa. As duas sessões seguintes serão dedicadas aos resultados referentes às DNPs P4 e P6, respectivamente.

5.1 Análise Preliminar

A rodada preliminar tinha o objetivo de me ajudar a tomar algumas decisões a respeito das rodadas futuras. Isso porque, como já mencionado no capítulo 5, a quantidade de dados por entrevista, levando em conta as seis pessoas gramaticais para cada criança, era muito grande para o escopo deste trabalho. Ademais, não fazia sentido

codificar uma quantidade tão grande de dados em sua maioria não variáveis. Assim sendo, codifiquei todos os inquéritos (um total de 13) de uma das crianças da faixa etária a partir de 4 anos (Camila) e o primeiro inquérito do menino menor (Artur), o que perfez um total de 1047 ocorrências.

O envelope de variação continha quinze variáveis, tal como descritas no capítulo 5:

1) *Realização da forma verbal*:
   para P1: (c) marca; (a) ausência de marca; (g) supergeneralização;
   para P2: (v) marca não padrão; (p) marca padrão; (n) ausência de marca;
   para P3: (h) marca; (i) marca de P6;
   para P4: (u) *–mo*; (s) *–mos*; (z) ausência de marca;
   para P5: (k) marca; (y) ausência de marca;
   para P6: (m) marca; (x) ausência de marca.
2) *Tempo verbal*: (P) presente do indicativo; (M) pretérito imperfeito do indicativo; (T) pretérito perfeito do indicativo; (F) futuro do pretérito do indicativo; (E) futuro do presente do indicativo; (S) presente do subjuntivo; (R) pretérito imperfeito do subjuntivo; (U) futuro do subjuntivo; (V) infinitivo; (I) imperativo; (Y) ir + infinitivo; (@)locução.
3) *Paralelismo formal:* (I) verbo em construção isolada; (P) primeiro verbo de uma série; (M) mistura de marcas em que o elemento anterior ao elemento analisado é marcado; (N) mistura de marcas em que o elemento anterior ao elemento analisado é não marcado; (T) verbo em uma seqüência com todas as marcas de concordância; (S) verbo de uma seqüência sem marcas de concordância.
4) *Discurso reportado*: (H) discurso reportado direto de histórias; (D) discurso reportado direto de outra pessoa; (S) discurso reportado de si mesmo; (N) discurso não reportado.
5) *Contexto seguinte:* vogais: (a) /a/ átono; (1) /a/ tônico; (e) /e/ átono; (2) /e/ tônico; (i) /i/ átono; (3) /i/ tônico; (o) /o/ átono; (4) /o/ tônico; (u) /u/ átono; (5) /u/ tônico; e as consoantes correspondentes aos códigos: (b); (c); (d); (f); (g); (j); (l); (m); (n); (o); (p); (q); (r); (s); (t); (v); (x); (z), e também (0) pausa.
6) *Saliência fônica*: (0) não se aplica; 0 – não se aplica; (A) (*falava/falávamos*); (B) (*fala/falamos, trouxe/trouxemos*); (C) (*está/estamos, tem/temos*); (D) (*comeu/comemos, partiu/partimos, vai/vamos, foi/fomos*); (E) *(falou/falamos)*; (F) caso único: *é/somos;*(1) *come/comem, fale/falem*; (2) *fala/falam, ia/iam*; (3) *faz/fazem, quer/querem*; (4) *dá/dão, está/estão, falará/falarão*; (5) *sumiu/sumiram, comeu/comeram, falou/falaram*; (6) *fez/fizeram, teve/tiveram*; (7) caso único: *é/são*; (8) caso único: *foi/foram*; (9) caso único: *veio/vieram.*
7) *Tonicidade*: (M) monossílabo tônico; (O) oxítona, (P) paroxítona; (R) proparoxítona.
8) *Tipo de verbo:* (T) transitivo; (I) intransitivo e reflexivo; (E) ergativo; (C) cópula; (H) *ter/haver* com função apresentativa; (V) *vamos.*
9) *Tipo de sujeito*: (s) SN pleno; (p) pronome reto; (i) pronome indefinido; (n) nulo expletivo; (z) nulo de pronome; (g) *a gente*; (y) nulo de *a gente*; (d) pronome demonstrativo; (j) pronome de uso genérico.
10) *Posição do sujeito em relação ao verbo:* (A) anterior adjacente; (D) anterior distante; (F) posterior distante; (P) posterior adjacente; (Z) não se aplica (casos de sujeito nulo).
11) *Tipo de coleta*: (r) relato pessoal; (n) narrativa; (h) história em seqüência; (l) leitura; (c) conversa informal; (d) leitura de desenho; (m) música.
12) *Animacidade do sujeito*: (A) sujeito [+] animado; (M) sujeito [-] animado; (N) não se aplica.
13) *Focalização:* (f) *foreground*; (b) *background.*

14) *Gênero:* (F) feminino; (M) masculino.
15) *Faixa etária*: (d) 2;0 a 2;5; (2) 2;6 a 2;11; (3) 3 anos; (4) 4 anos; (5) 5 anos; (6) 6 anos; (7) 7 anos; (8) 8 anos.

Os resultados percentuais gerais desta primeira etapa podem ser vistos na Tabela 1 a seguir:

Tabela 1 – Resultados gerais da rodada preliminar

		Ocorrências	**Total**	**%**
P1	presença	233	238/1047	**23**
	ausência	4		
	supergeneralização	1		
P2	presença	-	85/1047	**8**
	ausência	85		
P3	desinência não marcada	623	624/1047	**60**
	marca de DNP-P6	1		
P4	marca padrão	14	38/1047	**4**
	marca não padrão	15		
	ausência	9		
P5	presença	1	1/1047	**0**
	ausência	-		
P6	presença	55	61/1047	**6**
	ausência	6		

Tal como esperado, o maior percentual de ocorrências é relativo à terceira pessoa do singular (60%), enquanto que o menor (um único caso), refere-se à segunda pessoa do plural. Esses índices estão relacionados às especificidades das coletas em relação aos tipos de discurso produzidos. Nas conversas informais entre o entrevistador e a criança, assim como nos relatos pessoais, o foco de interesse é a criança, levada a falar muito mais dela mesma que de ou para seu interlocutor (e nesse ponto é interessante ressaltar que há um alto percentual de imperativos, 92%, relativamente à segunda pessoa do singular, o que faz pensar sobre a atribuição de papel que o interlocutor assume nestas interações, qual seja, o de ser exortado pela criança a executar ações). As narrativas e histórias em seqüência, por terem fundamentalmente estrutura narrativa, conduzem à produção de

terceira pessoa, singular (na maioria dos casos, já que as narrativas infantis em geral são muito mais acerca de um único personagem que de vários) ou plural.

Outro aspecto a ser levantado é que alguns trabalhos psicolingüísticos, concentrados, sobretudo em idades até três anos (Hyams, 1992; Pizzuto e Caselli, 1992), relatam a baixa produção de formas de plural pelas crianças. Perroni-Simões e Stoel-Gammon (1977) estabelecem que as formas de plural são tardias na produção de crianças pequenas. Não tenho condições de aprofundar aqui questões dessa natureza, até porque extrapolam o escopo do trabalho, mas é interessante pensar que enquanto a produção de sujeitos no singular soma 90% dos casos, a de sujeitos no plural contabiliza apenas 10%. Isso significa que para um estudo mais consistente sobre regra variável nas formas de plural (DNPs P4, P5 e P6) seriam importantes coletas específicas com a devida elicitação dessas formas.

Voltando aos percentuais, relativamente a P1, os números de ocorrências de ausência de marca (4) e de supergeneralização (1) são bastante baixos. Esses casos estão descritos na literatura psicolingüística como parte do desenvolvimento da linguagem infantil, seja em termos de estágios de desenvolvimento (Perroni-Simões e Stoel-Gammon, 1977), seja em termos de desvios (ou erros) em relação à fala do adulto (Hyams, 1992; Pizzuto e Caselli, 1992; Rubino e Pine, 1998; Nicolau, 1998). Dessa forma, pensei que talvez o aumento da quantidade de dados, sobretudo nas faixas etárias até 3 anos, revelasse resultados diferentes desses, com índices maiores das formas não previstas na fala do adulto. Assim sendo, na seqüência do trabalho, decidi por manter a codificação de P1 nas faixas etárias iniciais.

Em relação a P2, o percentual de 100% de ausência de marca poderia já indicar que não seria produtivo considerar a segunda pessoa do plural na amostra. Contudo, os resultados da variável tempo verbal mostraram que 92% (48/52) dos casos de imperativo são relativo à P2, contra 2% de perfeito, 7% de presente, 4% de *ir + infinitivo* e 12% de infinitivo. Ressalto aqui que no imperativo, sendo a desinência de segunda pessoa idêntica à forma não marcada do presente do indicativo, decidi por codificar a forma atestada pela gramática para segunda pessoa do singular como ausência de marca. Dessa forma, novamente o aumento na quantidade de dados, com a possibilidade de haver mais dados nos tempos verbais que não o imperativo, poderia apresentar outros resultados, mesmo que estudos sobre concordância variável na DNP-P2 indiquem baixos percentuais de aplicação da marca[32].

Na terceira pessoa do singular, o único caso de ausência de marca nessa rodada foi produzido por Camila. Contudo, como já havia lido alguns dos outros inquéritos da criança menor, sabia que havia poucos, mas mais casos desse tipo. Além disso, codificar as outras pessoas gramaticais deixando de lado a terceira do singular poderia enviesar o *corpus*.

Note-se ainda que nos percentuais de P4 vistos na Tabela 1 estão incluídos os casos de sujeito *a gente*. Ressalte-se que 8 dos 9 casos de ausência de marca são relativos a essa forma, que deverá ser excluída das rodadas específicas à DNP-P4.

[32] Amaral (2003) atesta 7,4% de manutenção da marca na comunidade de Pelotas e Loregian (1996), 4% na comunidade de Porto Alegre.

Assim sendo, a partir desses resultados, tomei as seguintes decisões: (i) eliminar a segunda pessoa do plural da análise; (ii) compor o *corpus* das faixas etárias a partir de 4 anos por DNP-P4 e DNP-P6 somente; e (iii) compor o *corpus* das faixas etárias entre 2;0 a 3;11 considerando as cinco pessoas gramaticais.

5.2 Resultados gerais de presença/ausência de marca

Os resultados presentes nesta seção são relativos ao total de crianças da amostra e às cinco pessoas gramaticais.

Primeiramente, apresento o resultado geral em termos percentuais sobre o total de dados (2588 ocorrências). É preciso ressaltar, neste ponto, que há problemas nas três faixas iniciais e, portanto, esses resultados precisam ser considerados com cautela; também, que os índices referentes às três pessoas do singular são relativos às faixas etárias d, 2 e 3 e aos dados de uma das crianças maiores. Os demais casos, relativos à primeira e terceira pessoas do plural, referem-se ao total de informantes. Decidi expor essa rodada para poder discutir, mesmo que rapidamente, as três pessoas do plural. Além disso, mesmo que a amostra apresente problemas nas primeiras três faixas etárias, os resultados podem apontar certos caminhos que venham a auxiliar em algumas decisões para pesquisas futuras sobre concordância variável em dados de crianças. Os números são apresentados na Tabela 2 a seguir:

Tabela 2 – Resultados gerais do total de dados relativamente às cinco pessoas gramaticais consideradas

		Ocorrências	Total	%
P1	presença	433	446/2588	**17**
	ausência	9		
	supergeneralização	4		
P2	presença	3	249/2588	**10**
	ausência	246		
P3	desinência não marcada	997	1000/2588	**39**
	marca de DNP-P6	3		
P4	marca padrão	44	363/2588	**14**
	marca não padrão	99		
	ausência	218		
P6	presença	461	529/2588	**20**
	ausência	68		

É necessário salientar que, no tratamento analítico empreendido nessa rodada, não há cruzamento possível de algumas das pessoas gramaticais com alguns fatores das variáveis consideradas. Um caso bastante óbvio é o de P1, P2, P3 e P6 em relação ao uso da forma pronominal *a gente* e da forma verbal *Vamos*; ou, no âmbito discursivo, sujeito [-] animado associado a P1, P2 e P4; ou, ainda, a variável saliência fônica, que além de não se aplicar a P1, P2 (este trabalho não prevê análise de saliência fônica para P2), e P3, tem tratamento diferente para P4 e P6. Assim, esses resultados devem ser vistos com cautela.

Em relação à primeira pessoa do singular, do total de 446 ocorrências, apenas 13 (3%)[33] não apresentam a DNP-P1. Os quatro casos de supergeneralização são relativos à fala de Érica, aos 2 anos e 7 meses de idade; já as ocorrências de ausência de marca, listados abaixo, são distribuídas da seguinte forma:

[33] Fiz esse cálculo com base no número de ocorrências contidas na Tabela.

a) 4 episódios de fala de Camila, listados abaixo. Vale atentar para as idades em que ocorrem, indicadas dentro do parêntese, após o sinal de travessão:

36) *depois eu vou te [/] te dizer as música-0s que* ***eu aprendeu*** *[:aprendi] no colégio.* (Cam04-6; 01)

37) *eu andava de bicicleta depois foi a minha prima depois* ***foi*** *[: fui]* ***eu*** *depois foi a minha prima daí depois <a gente foi tomar sor> [//] a gente foi comer pipoca tomar um Sprite@s ["] andar mais um pouco e aí voltamo-0s pra casa .* (Cam12-8;3)

38) *a história do Parcão a história do Parcão foi muito legal primeiro* ***foi*** *[: fui]* ***eu*** *depois e a minha prima eu e a minha prima tomamos refri@g e comemos pipoca andamos mais um pouco e fomos embora.* (Cam12-8;3)

39) *mas* ***foi*** *[: fui]* ***eu*** *que trouxe pra casa .* (Cam13-8;6)

b) 4 episódios relativos à fala de Érica:

40) *MIL: machucou o dedinho?*
ERI: ***machucou****.* (Eri33-2;5)

41) *MIL: e que que é isso aqui?*
ERI: ***é eu****.* (Eri41-2;7)

42) *MIL: e o que que foi isso aqui Érica?*
ERI: ***foi eu****.* (Eri41-2;7)

43) *MIL: quem é esse rapazinho aqui?*
ERI: ***é eu****.* (Eri46-2;9)

c) 1 caso da fala de Artur:

44) *INV: e tu sabe quem sou eu?*
ART: eu sabe. (Art06-3;0)

Vistas em sua totalidade, as ocorrências em (36-44) atestam que há dois tipos de variação na fala das crianças. Por um lado, há variação aquisicional, descrita e discutida pela literatura psicolingüística, relacionada, nesses casos, aos processos aquisicionais de supergeneralização e uso da forma não marcada relativamente à primeira pessoa do discurso. Por outro, há variação sociolingüística, descrita e discutida pela literatura sociolingüística sobre dados de adultos, em termos de forças conduzindo ora a uma variante, ora a outra.

A literatura aquisicionista (cf. Perroni Simões, 1976 e Faria, 1993) atesta que há uma fase da aquisição da linguagem em que a criança usa a forma de terceira pessoa para referir-se a si mesma. Esse processo perdura ainda mesmo em fases posteriores, nas quais a criança já adquiriu o contraste de primeira pessoa[34]. Simões (1997), em estudo sobre a aquisição de sujeito nulo por um menino entre 2;4 e 3;0 de idade, afirma, a respeito das ocorrências de primeira pessoa com verbo com desinência não marcada, tal como a atestada em (36), serem provavelmente sobrevivência dessa estratégia. O interessante de notar aqui é que, se minha interpretação é adequada, trata-se de uma estratégia que perdura até os 6 anos de idade. Os dados em (36) e também em (44) podem estar relacionados a tal processo.

Nos casos em (37), (38) e (39) vê-se a atuação de parâmetros sociolingüísticos, já que se tratam de ocorrências de sujeito pós-verbal, ambiente descrito na literatura sociolingüística como favorável à ausência de marca. Além disso, em (37) e (38) há uma situação sintática de enumeração com alternância entre sujeitos de primeira e

[34] Brown (1973), entre outros, estabelece como ponto de aquisição o momento em que a criança passa a usar determinado morfema numa freqüência acima de 90% dos casos em que aquele morfema é requerido.

terceira pessoas do discurso, com verbos na mesma forma, ou seja, a disposição em série pode ter levado a criança a usar a mesma forma verbal para os dois sujeitos.

A ocorrência vista em (40) pode ser um caso em que a criança analisou o SN *o dedinho*, da fala da entrevistadora, como sujeito.

As ocorrências em (41) a (43), e talvez também em (44), podem estar influenciadas pala fala da entrevistadora, já que as respectivas formas verbais aparecem no turno imediatamente anterior ao da criança. O estudo de Nicolau (1998) mostrou que quando extraídas dos dados as formas "aparentemente" flexionadas, que constituem, em parte, fragmentos da fala da mãe, a freqüência de flexão na fala da criança cai de 83,9% para 4,1%[35].

Relativamente à segunda pessoa do singular, como já comentei acima, nos casos de verbo no imperativo, considerei a desinência prescrita pela gramática como ausência de marca e a forma de DNP-P3 como presença de marca. Reforço isso porque os três casos de manutenção da desinência são relativos a essa estrutura. Além disso, as três situações discursivas estão relacionadas a músicas, ou que estavam sendo cantadas pelas crianças (45) e (46) ou era repetição do trecho da música, retomado pela criança. Trata-se de um tipo de discurso em que a criança repete a forma verbal como forma cristalizada, como em (46) e (47). Entretanto, no caso de (45), a música em questão tem outra estrofe (ambas reproduzidas abaixo do exemplo) com a mesma estrutura – um vocativo reduplicado seguido por um verbo na forma imperativa –, que é produzida com a forma considerada aqui não marcada; esse caso, portanto, não se trata de uma forma cristalizada,

[35] Ver seção 3.3 deste trabalho.

mas ou a associação com o pronome de segunda pessoa *você*, que recebe a marca de terceira pessoa, ou um caso de hipercorreção.

(45) *motorista, motorista* ***olhe*** *o sinal* (Cam04-6;01)

> *["]motorista motorista* ***olha*** *o sinal olha o sinal ele (es)tá vermelho ele (es)tá vermelho pare então pare então motorista motorista* ***olhe*** *o sinal olhe o sinal ele está amarelo ele está amarelo atenção atenção*

(46) ***tenha*** *mais educação* (Art06-3;0)
(47) ***tenha*** *mais educação a senhora* (Art06-3;0)

Os três casos de terceira pessoa do plural que recebem DNP-P6 compartilham da mesma estrutura, na qual entre o sujeito (singular) e o verbo (plural) há um pronome possessivo plural. Nesses casos, a concordância parece ser regida pelo elemento anterior adjacente, e não pelo sujeito sintático.

(48) *tava aqui o vestido né@i e* ***a bunda delas eram*** *lá.* (Cam10-7:10)
(49) *ó como* ***o pescoço delas são*** *comprido.* (Art06-3;0)
(50) *é que* ***a casa deles são*** *iguais!* (Art06-3;0)[36]

Os resultados relativos às DNPs P4 e P6 serão discutidos na análise a seguir.

Passo agora aos resultados gerais com peso relativo. Nesta rodada, o objetivo era obter um quadro geral de aplicação das DNPs na fala das crianças consideradas, opondo os casos de presença aos de ausência das marcas de CV. Para tanto, os dados receberam uma recodificação considerando os dois casos, e a variável dependente foi reorganizada do seguinte modo:

1 – Presença de marca: c, g, v, p, i, u, s, k, m;

0 – Ausência de marca: a,n, z, y, x.

[36] Conferi os dados em (49) e (50) para me certificar de que é possível distinguir a produção do morfema de plural no pronome possessivo.

Do total de 1592 ocorrências, 1049 (66%) foram de presença da marca ao passo que 543 (34%) foram de ausência.

Dos quinze grupos de fatores considerados inicialmente, seis foram selecionados como estatisticamente relevantes pelo programa, na seguinte ordem: tipo de sujeito, tempo verbal, posição do sujeito em relação ao verbo, tonicidade, tipo de verbo e faixa etária, os quais passo a discutir a seguir.

5.2.1 Tipo de Sujeito

A variável tipo de sujeito, tal como composta originalmente, teve de ser reorganizada em virtude de alguns fatores apresentarem um número muito reduzido de ocorrências. Assim sendo, os casos de elipse de *a gente* foram amalgamados com a respectiva forma; o fator *nulo expletivo*, consistindo em um único caso, foi amalgamado com os casos de *nulo de pronome*; o fator *sujeito genérico*, além de contar com um pequeno número de casos (14), apresentou 100% de ocorrências de manutenção da marca, fato que não pode ser computado pelo programa. Por isso, esse fator foi amalgamado com *pronome demonstrativo*, que também apresentava um único caso. Os resultados podem ser vistos na Tabela 3 a seguir.

Tabela 3 – Índices de aplicação da marca de concordância em relação à variável tipo de sujeito

	Ocorrências	**/**	**Total**	**%**	**Peso Relativo**
pronome indefinido	7	/	8	87	0,86
pronome reto	435	/	57	88	0,76
nulo de pronome; nulo expletivo	376	/	530	70	0,75
SN pleno	124	/	179	69	0,70
pronome demonstrativo; genérico	14	/	15	93	0,65
“a gente”; nulo de “a gente”	2	/	214	0	0,01
TOTAL	**958**	**/**	**1438**	**66**	

Input: 0,74

A tabela mostra que a presença da desinência é favorecida por todos os tipos de sujeito, com exceção da forma inovadora *a gente* e seu respectivo nulo, que têm um índice praticamente categórico (0% - 0,01) de ausência de marca. Os dois únicos casos em que a forma pronominal *a gente* está associada à presença da marca (ou seja, o emprego da DNP-P4) são listados a seguir:

> (51) *daí o filho perguntou: mas* ***a gente não comemos*** *antílopes?*(Mat10-8;3)
>
> (52) ***a gente saímos*** *[: saiu] junto com as professora-0s no passeio a gente nunca fica # longe das professora-0s # a gente sempre fica perto delas porque senão a gente se perde* . (Nat01-5;4)

A estrutura em (51) pode estar ligada ao fato de tratar-se de discurso reportado direto dentro de uma narrativa. A criança está narrando o filme *Rei Leão*. Quem fala, nesse discurso, é o personagem principal, Simba, que conversa com seu pai, o então rei dos animais. Na cena, o rei mostra o reino para o filho, que virá a ser rei, dando uma espécie de lição de sabedoria sobre o “ciclo da vida”. Portanto, está inserido em uma conversa entre dois personagens hierarquicamente superiores aos outros do filme, numa situação mais formal que a de brincadeiras, por exemplo, entre os dois. Ademais, a

estrutura narrativa aqui propicia um tipo de discurso com um grau de formalidade maior que a simples conversa.

A ocorrência em (52) pode ser um caso em que a criança monitora mais o seu discurso. A estrutura é produzida no início do primeiro inquérito, em que ela se depara com a solicitação, pela entrevistadora, de um relato pessoal na situação formal de entrevista de um projeto de pesquisa acadêmico. Além disso, o assunto do relato escolhido pela criança é algo ocorrido numa situação escolar. Cabe salientar que a combinação de sujeito *a gente* e verbo com DNP-P4 não existe nos dados das crianças pequenas, ou seja, nas faixas de idades pré-escolares. Li os inquéritos (a totalidade das entrevistas do menino e as entrevistas da menina até três anos que já estão codificados) e não encontrei nenhuma ocorrência desse tipo. Pode-se pensar com isso que a presença da marca de concordância em verbos cujo sujeito é a forma *a gente* está ligada à pressão exercida pela escola em favor da presença da DNP[37]. Talvez o estigma que envolve a desinência zero agregada ao pronome de primeira pessoa do plural seja fruto da ratificação que a escola faz pela presença da marca.

Voltando aos índices da Tabela 3, pode-se ver que os pronomes indefinidos lideram a motivação da manutenção da marca (87%, 0,86). É preciso salientar que o número de ocorrências de pronome indefinido (e também de demonstrativo e genérico) é bastante baixo, por isso esses resultados devem ser vistos com cautela. Relativamente aos indefinidos, fiz uma consulta ao banco de dados para localizar as ocorrências e saber de que pessoas gramaticais se tratavam, já que, se estivessem relacionadas às três pessoas do singular, que mostraram índices quase categóricos de presença da DNP, essa situação seria

[37] Almeida (2005) encontra maior aplicação de marca de concordância nos dados da comunidade quilombola por ela examinada que nos de outras comunidades rurbanas (de transição da fala rural para a fala urbana). A diferença é atribuída à pressão da escolarização nessa comunidade.

esperada. O resultado apontou um número maior de ocorrências que têm como sujeito um pronome indefinido que os que aparecem computados pelo programa. Provavelmente isso se deve a dados descartados em razão da metodologia estatística (a necessidade de eliminação de *Knockouts*). O total de casos que apareceram na consulta aponta 9 ocorrências de pronome indefinido relacionado a P3 (como em ***um*** *é preto e* ***o outro*** *é bege*) e 7 ocorrências associadas a P6 (como em ***uns*** *vão ali* e ***todos*** *caíram na cabeça*). O único episódio de ausência de marca é em uma estrutura com sujeito pós-verbal:

(53) * *ó # é* ***todas*** *# fechada toda encaixada.*

Assim sendo, o efeito desse fator não está relacionado unicamente às pessoas gramaticais do singular. Os pesos relativos bastante próximos dos fatores pronome reto, nulo de pronome e nulo expletivo e SN pleno (0,76, 0,75 e 0,70 respectivamente) mostram que seus efeitos são similares de favorecimento da aplicação da desinência.

Em relação aos achados de Rodrigues (1992), os resultados aqui corroboram que o sujeito nulo favorece a presença da marca. Entretanto, contradizem os resultados da autora em relação ao sujeito pronominal explícito que, naquele trabalho, propiciam a perda da marca de flexão verbal tanto em relação à DNP-P4 (47% e 0,45) quanto à DNP-P6 (32% e 0,22).

5.2.2 Tempo verbal

Os ajustes feitos na variável tempo verbal foram devidos ao baixo número de ocorrências de três fatores: futuro do pretérito do indicativo (sem ocorrências), presente do subjuntivo, que contava com um único caso (de aplicação de marca), e imperfeito do subjuntivo (3 casos de aplicação e 2 de ausência de marca, totalizando 5 casos). Esses fatores foram retirados da análise. Os resultados da variável podem ser vistos na Tabela 4 a seguir.

Tabela 4 – Índices de aplicação da marca de concordância em relação à variável tempo verbal

	Ocorrências	**/**	**Total**	**%**	**Peso Relativo**
ir + infinitivo	149	/	167	89	0,76
pretérito perfeito do indicativo	406	/	565	71	0,74
infinitivo	18	/	26	69	0,47
presente do indicativo	209	/	285	73	0,48
locução	92	/	142	64	0,43
pretérito imperfeito do indicativo	71	/	114	62	0,41
imperativo	10	/	130	7	0,01
TOTAL	**955**	**/**	**1429**	**66**	

Input: 0,74

Os resultados apontam a perífrase *ir + infinitivo* e o pretérito perfeito do indicativo como os tempos verbais que favorecem a aplicação das DNPs (0,76 e 0,74 de peso relativo respectivamente), ao passo que os demais (infinitivo, presente do indicativo, as locuções, imperfeito do indicativo e imperativo) desfavorecem a aplicação. Vale chamar atenção para o forte efeito de desfavorecimento da presença da marca relacionado ao imperativo. Isso certamente se deve ao tratamento que dei à codificação da segunda pessoa do singular para essa forma verbal, já que do total de 130 ocorrências de imperativo, 123 são de segunda pessoa do singular. Os demais casos de imperativo são de DNP-P4 do verbo *ir*, ou seja, a forma *Vamos*, como em *Vamo lá no meu quarto?*. Em relação à perífrase, julguei que o alto índice de futuro perifrástico (*vamos + infinitivo*) estaria

influenciando muito o resultado. Entretanto, dos 167 casos da perífrase, apenas 65 referem-se a P4. Simões (1997) mostra, para os dados de um menino entre 2;4 e 3;0 de idade, que, junto com o presente e o passado, o futuro perifrástico está, em termos de freqüência de enunciados, presente na amostra desde a primeira coleta. E como o exemplo citado pela autora é referente à primeira pessoa do plural, não parece ser um tempo verbal produtivo somente para P4.

5.2.3 Posição do sujeito em relação ao verbo

Em relação a esse grupo de fatores, as ocorrências de posterior distante (2 casos de aplicação da marca e 4 de não aplicação) foram amalgamadas à posterior adjacente em função de seu pequeno número. Os casos que haviam recebido código Z (não se aplica), relativos aos casos de sujeito nulo, foram desconsiderados na análise. Os resultados, conforme apontados na Tabela 5 a seguir, indicam que as estruturas com sujeito posposto ao verbo favorecem enormemente (peso relativo de 0,06) a ausência de aplicação da marca, conforme esperado, ao passo que o sujeito em posição anterior adjacente ao verbo favorece a presença (0,57 de peso relativo). A posição anterior distante, embora desfavoreça a marca (peso relativo abaixo de 0,50), tem um índice próximo ao anterior distante. Isso faz pensar que na fala das crianças em análise a atuação desse fator diz respeito à oposição anterior *versus* posterior, mais do que ao material interveniente que haja entre o sujeito e o verbo. A Tabela 5 mostra os resultados dessa variável.

Tabela 5 – Índices de aplicação da marca de concordância em relação à variável posição do sujeito em relação ao verbo

	Ocorrências	**/**	**Total**	**%**	**Peso Relativo**
posterior	18	/	68	26	0,06
anterior adjacente	453	/	675	67	0,57
anterior distante	95	/	132	71	0,48
TOTAL	**566**	**/**	**875**	**64**	

Input: 0,74

5.2.4 Tonicidade

A variável tonicidade teve o fator proparoxítonas eliminado da análise em razão de apresentar uma única ocorrência. Sobre essa variável, percebi que cometi um erro ao considerar, na codificação, a forma produzida pela criança, ao invés da forma alvo. Com isso, deixaria de obter informações probabilísticas a respeito do cruzamento entre proparoxítonas e ausência de marca, atestadas na literatura sobre primeira pessoa do plural (Bortoni, 1984; Zilles et al., 2000 e Batista e Zilles, 2004) como fortemente relacionados. Fiz então uma consulta ao banco de dados para ter uma idéia geral desse cruzamento. Para tanto, cruzei os tempos verbais cuja forma alvo é proparoxítona (nesse caso, somente imperfeito do indicativo, já que o futuro do pretérito do indicativo e o imperfeito do subjuntivo não entraram na rodada). O resultado da consulta mostrou que há somente quatro casos de imperfeito do indicativo relativos à primeira pessoa do plural, sendo um deles marcado (o que aparece na rodada) e três não marcados. Esse número mostra que mesmo que eu tivesse codificado segundo a forma alvo, o fator seria igualmente eliminado da rodada. As quatro ocorrências podem ser vistas a seguir:

(54) *nós três era namorada dele.* (Cam06’219)
(55) *nós tava XXX* (Mat05’22)

(56) *nós tava brincando de lutar* (Mat05’23)
(57) *estávamos em casa, e a minha mãe estava grávida* (Gab13)

O único caso de proparoxítona (57) em toda a amostra (o que significa em todos os informantes) foi na fala de um menino de 8 anos que demonstra ter estabelecido uma clara correlação entre a escrita e a forma padrão. Quando solicitado a contar algo que tinha acontecido com ele (relato pessoal), usou a forma inovadora *a gente* (em (58) abaixo); quando solicitado a escrever, e a seguir, ler o mesmo relato (em (59) abaixo), ele usou o pronome reto com a desinência padrão, produzindo a proparoxítona, deixando bastante claro o caráter formal da segunda produção em relação à primeira.

(58) *é assim né@i <**a gente (es)tava** ãhn@i na> [//] a minha tia # ela tem um vídeo # daí quando <xxx> [//] (es)tava estragado o nosso vídeo daí ela mandou o nosso vídeo pra gente a minha mãe (es)tava no hospital lá e o meu pai (es)tava junto # não meu pai (es)tava com a gente # não # (es)tava lá no hospital é # e foi [/] daí alugaram um filme do Bambi ["] # eu não gostava muito né@i # daí [/] daí ligaram pra gente dizendo que [/] # que o nenê nasceu # já .*

(59) *tá@i # <**estávamos** em casa e a minha mãe estava grávida> [= lendo] (es)pera aí que errei aqui (...) <e de repente o telefone # tocou e nós atendemos e disseram que o nenê nasceu .> [= lendo].*

Assim sendo, não seria possível avaliar com consistência a influência das proparoxítonas na presença/ausência das DNPs na amostra, mas a indicação é de que a produção das crianças corresponde àquela dos adultos atestada por Bortoni (1984), Zilles et al. (2000) e Batista e Zilles (2004) de evitação de formas proparoxítonas no português falado no Brasil. Os resultados obtidos na rodada, portanto, dizem respeito às paroxítonas, oxítonas e monossílabos tônicos, e podem ser vistos na Tabela 6 a seguir.

Tabela 6 – Índices de aplicação de marca de concordância em relação à variável tonicidade

	Ocorrências	/	Total	%	Peso Relativo
paroxítona	655	/	858	76	0,63
monossílabo tônico	216	/	385	56	0,38
oxítona	86	/	190	45	0,19
TOTAL	**957**	**/**	**1433**	**66**	

Input: 0,74

Percebe-se certa polaridade na qual tem-se as paroxítonas de um lado, favorecendo a presença das DNPs (0,63), e os monossílabos tônicos e oxítonas de outro, desfavorecendo a aplicação da marca (0,38 e 0,19 respectivamente), o que contraria minha expectativa em relação a essa variável. A hipótese era de que as formas que apresentassem a desinência na sílaba tônica, ou seja, oxítonas e monossílabos tônicos, favoreceriam a retenção da marca.

5.2.5 Tipo de verbo

Os resultados obtidos para a variável tipo de verbo mostram a forma *Vamos* como fortemente favorecedora da manutenção da marca, com percentual de 98% e peso relativo de 0,98. É preciso salientar que nesse fator estão incluídas tanto as formas com a presença do –s quanto as de seu apagamento. No outro extremo, vê-se os verbos ergativos e as construções apresentativas com *Ter/Haver* desfavorecendo a aplicação (pesos de 0,37 e 0,25, respectivamente). Esses resultados dão suporte ao que já era esperado. Zilles e Batista (2005) atestam o forte favorecimento do uso da desinência que a forma *Vamos* representa. Em relação a verbos ergativos e estruturas apresentativas, estão ambos relacionados a estruturas VS. Os resultados podem ser vistos na Tabela 7 a seguir.

Tabela 7 – Índices de aplicação de marca de concordância em relação à variável tipo de verbo

	Ocorrências	/	Total	%	Peso Relativo
Vamos	74	/	75	98	0,98
intransitivo e reflexivo	146	/	216	67	0,52
função copulativa	32	/	54	59	0,50
transitivo	626	/	951	65	0,44
ergativo	76	/	120	63	0,37
Ter/*Haver*	4	/	18	22	0,25
TOTAL	**958**	**/**	**1434**	**66**	

Input: 0,74

5.2.6 Faixa etária

A última variável escolhida pelo programa foi faixa etária, cujos resultados são apresentados na Tabela 8 a seguir.

Tabela 8 – Índices de aplicação de marca de concordância em relação à variável faixa etária

	Ocorrências	/	Total	%	Peso Relativo
4 – 4 anos	10	/	18	55	0,74
3 – 3;0 a 3;11	235	/	326	72	0,59
5 – 5 anos	54	/	123	43	0,54
6 – 6 anos	146	/	228	64	0,53
7 – 7 anos	220	/	305	72	0,52
8 – 8 anos	196	/	266	73	0,49
d – 2;0 a 2;5	13	/	19	68	0,41
2 – 2;6 a 2;11	175	/	307	57	0,35
TOTAL	**1049**	**/**	**1592**	**65**	

Input: 0,74

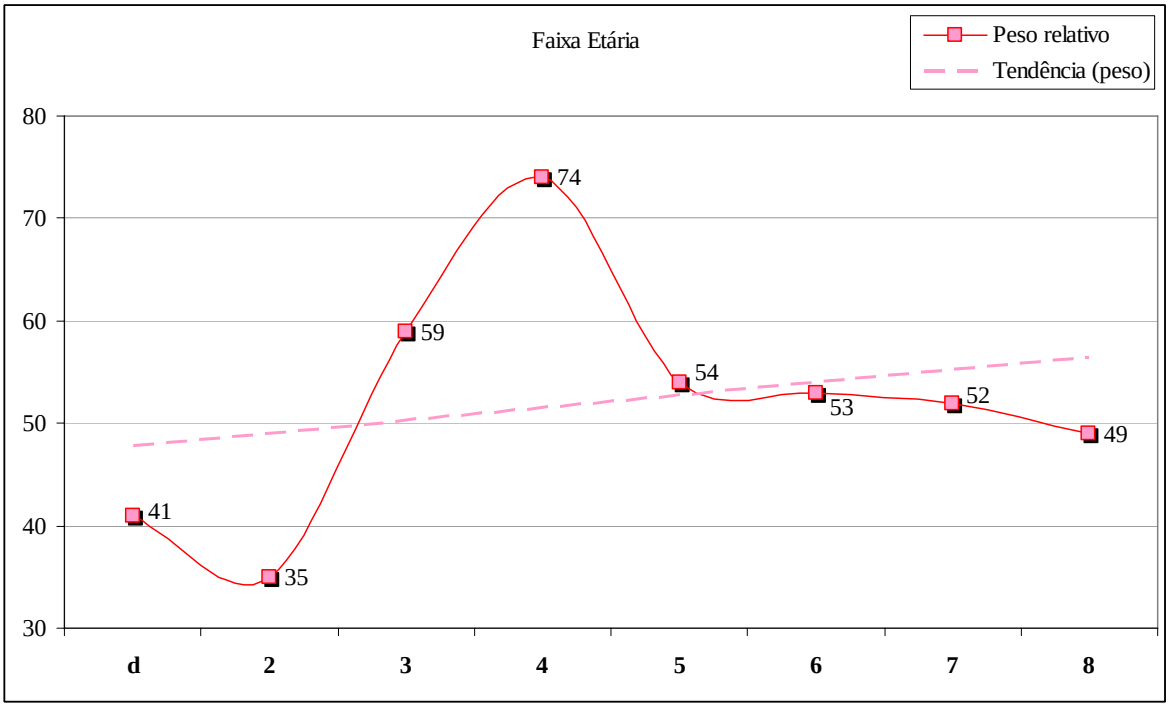
O modo como foi montada a amostra nas três faixas etárias iniciais, associado ao fato de a rodada considerar as cinco pessoas gramaticais juntas, dificulta enormemente a interpretação desses resultados. Além disso, nas faixas de 4 e 2;0 a 2;5 anos, o número de ocorrências é significativamente baixo. Examinando os índices de uma

forma geral (Gráfico 1), poder-se-ia dizer que a tendência é a de que com o aumento da idade haja maior número de marcas, com um pico de aquisição aos 4 anos e um processo de perda de foco (e da relevância estatística) a partir dos 5 anos[38]. Os resultados parecem coerentes, a não ser pela faixa de 3 anos, que apresenta índices muito altos de produção de marca. Esses índices certamente são relativos a características da fala de Artur. Em trabalho anterior (Soares, 2003) sobre concordância nominal padrão *versus* não padrão, Artur mostrou resultados surpreendentemente altos de produção da forma padrão (114/127 – 89,7%). Com isso, pode-se pensar que o fato de a faixa etária conter somente dados desse menino enviesa o resultado.

Assim, uma análise mais conclusiva só seria possível mediante uma reestruturação da amostra, preferencialmente, inclusive, com um número maior de crianças nas faixas iniciais.

[38] Essa análise, bem como o Gráfico 1, foram propostos, e este também executado, pelo prof. Luís Amaral, a quem agradeço a colaboração.

Gráfico 1 – Aplicação da marca de concordância em relação à faixa etária



Devido ao caráter exploratório que tem essa rodada geral, vale comentar o trabalho de Rubino e Pine (1998). As autoras examinaram a proporção dos erros de flexão verbal em uma criança adquirindo PB e revelaram contrastes entre a freqüência de produção de diferentes verbos e a proporção de erros de concordância verbal associados a eles. Na Tabela 5 (p. 44) elas relacionam o número de produções incorretas de flexão (de desinência de número ou pessoa) sobre o total de ocorrências relativamente às seis pessoas gramaticais, incluindo a forma *a gente*. Os resultados apontaram a terceira pessoa do plural como o ambiente com maior freqüência de erros (10/23 ou 43,5%), seguida pela forma *a gente* (4/17 ou 23,5%), pela primeira pessoa do plural (25/297 ou 8,4%) e, finalmente, pela terceira pessoa do singular (5/929 ou 0,5%). Esses resultados são reveladores no sentido de que corroboram os ambientes lingüísticos relativamente às pessoas gramaticais em que se

dão os processos de variação na fala das crianças. Associados aos resultados obtidos na presente pesquisa, vê-se que a variação sociolingüística na fala das crianças desde cedo não difere em grandes termos daquela encontrada na fala dos adultos. Pode-se pensar com isso, que o tratamento dado em pesquisas psicolingüísticas (Rubino e Pine, 1998; Pizzuto e Caselli, 1992, 1993) para as formas de flexão que não atendem ao modelo padrão, sendo consideradas erro só faz sentido se ignorar por completo os achados da sociolingüística. Os "erros" cometidos pelas crianças são explicáveis em termos de variação sociolingüística, fato que atesta o quanto, desde cedo, a fala das crianças é sensível a tais forças.

Por outro lado, em relação aos resultados gerais apresentados até aqui, algumas considerações são necessárias. De um ponto de vista estatístico, eles poderiam ter sido refinados, seja com a amalgamação de alguns fatores que mostraram pesos relativos próximos, seja eliminando um ou outro fator que estaria atuando de forma maciça na rodada. Entretanto, não julguei necessário fazê-lo, em virtude, sobretudo, do tratamento igualitário dado a todas as pessoas gramaticais no modo como a análise foi feita. Tanto do ponto de vista morfossintático quanto discursivo, as pessoas gramaticais têm especificidades em relação à regra variável (ou categórica) de CV, de aplicação ou não da marca. Por exemplo, o grupo de fatores tipo de verbo, escolhido como estatisticamente significativo pelo programa, possui um fator (*Ter/Haver)* que só diz respeito à oposição singular/plural na terceira pessoa, não tendo relação com primeira e segunda pessoas. Ainda na mesma variável, o verbo *Vamos* só se insere na reflexão sobre DNP-P4. Além disso, na terceira pessoa do singular, a ausência da marca é categórica[39]. Generalizando, os casos de P1 são categóricos na aplicação da marca, ao passo que os casos de P2 e P3 são categóricos

[39] Com pouquíssimos casos, como os atestados aqui, de palavra interpolente no plural entre um sujeito singular e um verbo com DNP de plural, mesmo na língua adulta.

em termos de não aplicação da marca. Disso resulta a possibilidade de que seja gerada uma análise estatística (como de fato o foi), entretanto, sobre objetos lingüisticamente distintos, com regras de marcação de concordância opostas em termos de presença/ausência da desinência.

Assim sendo, os resultados apresentados anteriormente têm muito mais um caráter descritivo geral de aplicação de marca de CV na fala das crianças em estudo que um caráter propriamente analítico.

5.2.7 Variáveis não selecionadas

Ainda em relação a essa rodada geral, vale tecer alguns comentários sobre algumas das variáveis que não foram selecionadas como estatisticamente relevantes.

A variável paralelismo formal, tal como codificada inicialmente, deveria ser reorganizada com a eliminação dos fatores categóricos (T – verbo de uma seqüência com todas as marcas de concordância, e S – verbos de uma seqüência sem marcas de concordância). Ela previa ainda a codificação de verbo em construção isolada, fator maciço na fala das crianças menores e nas conversas informais em geral e que, no entanto, não servia para medir a influência da variável na aplicação ou não de marca de CV, e também deveria ser retirado da rodada. Os fatores eliminados na análise somaram um total de 1404/1592. Do número de casos que restaram – 188 –, 166 deles são ocorrência de primeiro verbo de uma série, cujo exame só faz sentido em relação aos verbos que estão em série.

Dessa forma, em relação à amostra da pesquisa, não há possibilidade de examinar a variável paralelismo formal.

Em relação à variável discurso reportado, os fatores referentes a discurso reportado de si mesmo e discurso reportado direto de outra pessoa foram amalgamados a discurso reportado direto de histórias. Os três fatores somados contabilizaram 51 ocorrências, em oposição a discurso não reportado, que contava com 1370 casos. A variável não foi selecionada e, de qualquer modo, não serviria para examinar a influência dos tipos de discurso reportado, já que sua manifestação na fala das crianças é muito baixa. Para examinar tal variável na fala de crianças, seria necessário elicitar tais tipos de discurso na coleta.

O objetivo inicial da variável contexto seguinte era examinar seu efeito, sobretudo, na distinção das formas padrão e não padrão na primeira pessoa do plural em função do processo de ressilabificação. Como a inclusão dela na rodada gerou um número muito elevado de *knockouts*, foi recodificada levando em conta o modo de articulação, sendo, então, organizada da seguinte forma[40]:

Vogais:	A – arredondadas: a, o, u (átonas e tônicas)
	N – não arredondadas: e, i (átonas e tônicas)
Consoantes:	P – plosivas (p, b, t, d, c, k, g)
	F – fricativas (f, v, s, z, x, j)
	L – laterais (r, l)
	N – nasais (n, m)

Mesmo com tal organização, a variável não se mostrou produtiva.

[40] A reorganização dessa variável também foi proposta e executada pelo prof. Luís Amaral.

A variável saliência fônica não fez parte dessa rodada porque, tal como foi prevista (com códigos para a primeira e a terceira pessoas do plural), só fazia sentido nas análises específicas de DNP-P4 e DNP-P6.

Passo, na seqüência, à apresentação e discussão dos resultados referentes às DNPs P4 e P6.

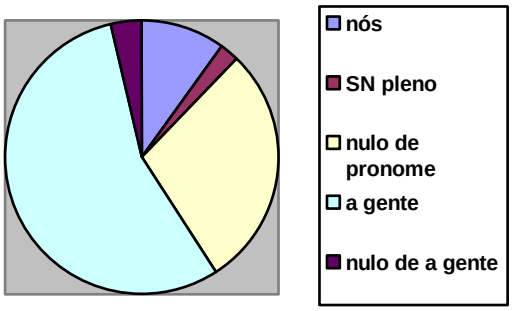
5.3 A concordância de primeira pessoa do plural

Em relação à DNP-P4 foram feitas duas análises distintas. Uma examinando aplicação *versus* não aplicação da desinência, e outra focalizando o apagamento ou não do morfema de plural (-mo *versus* -mos), ambas desconsiderando a forma pronominal *a gente*. Além disso, cada rodada foi feita, inicialmente, com o total de faixas etárias e, nos casos em que a variável faixa etária foi selecionada, foi realizada uma segunda rodada abordando somente as faixas de 4 a 8 anos. Antes de apresentar as análises, vou expor os resultados relativos à forma inovadora *a gente*.

5.3.1 A forma *a gente*

No total das faixas etárias, as 363 ocorrências de DNP-P4 foram distribuídas da seguinte forma: 36 ocorrências (9,9%) do pronome *nós*, 9 ocorrências (2,4%) de SN pleno, 104 ocorrências (28,6%) de elipse do pronome, 202 ocorrências (55,6%) de *a gente* e 13 ocorrências (3,5%) de elipse de *a gente*. O gráfico 2 abaixo mostra a distribuição:

Gráfico 2 – Distribuição geral das formas da DNP-P4



Seara (2000) apresenta um estudo sobre a distribuição do pronome *nós* e da forma *a gente* em dados extraídos de entrevistas do VARSUL. O *corpus* é composto por doze informantes florianopolitanos com nível de escolaridade primário e colegial, sendo seis do sexo feminino e seis do masculino, divididos em três faixas etárias: 15 a 24 anos; 25 a 50 anos e mais de 50 anos. Nos resultados, contabilizou um percentual da forma inovadora de 72% (contra 28% do pronome *nós*).

Zilles (2005) aborda o processo de gramaticalização de *a gente* comparando o uso dessa forma nas décadas de 1970 e 1990. Na análise sobre tempo aparente, relativamente à década de 1990, os dados foram extraídos de entrevistas do projeto VARSUL de falantes de Porto Alegre, divididos por gênero (masculino e feminino), faixa etária (25-49 anos e mais que 50 anos) e escolaridade (ensino fundamental, primário, médio e superior). Os resultados mostraram 69% de uso da forma inovadora, com uso categórico do verbo com a desinência não marcada (a desinência de terceira pessoa do singular). Como a presente pesquisa não aprofundará a análise sobre o uso de *a gente*, me limitarei à comparação com os índices gerais nos dois estudos.

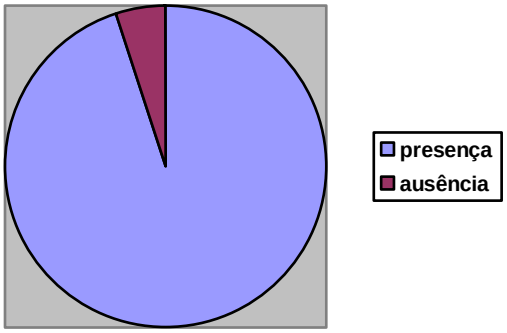
A soma de ocorrências de *a gente* e sua elipse na amostra das crianças totalizou 59%, um índice menor tanto em relação aos falantes de Florianópolis, quanto aos de Porto Alegre. Esse resultado pode estar relacionado ao número de ocorrências da forma verbal *Vamos*, que nas crianças chega a 75/363 no total de ocorrências, correspondendo a 75/104 ocorrências da elipse do sujeito pronominal. Isso quer dizer que o percentual mais baixo de *a gente* na fala das crianças em relação aos adultos pode ser reflexo do percentual de pronomes nulos, ampliado em razão do uso de *Vamos.*

Do total de casos de *a gente*, apenas dois deles foram realizados com DNP-P4 (ver 6.3.1 acima). Os demais foram produzidos com a desinência não marcada (DNP-P3). Essa distribuição não justifica uma análise estatística da forma pronominal inovadora.

5.3.2 Presença *versus* ausência da DNP-P4

Na análise sobre a presença da DNP-P4, do total de 146 ocorrências, 140 (95%) foram de presença da desinência e 6 (5%) foram de ausência. O gráfico 3 exibe os resultados.

Gráfico 3 – Presença *versus* ausência da DNP-P4



Nesta rodada, três variáveis foram selecionadas pelo VARBRUL como estatisticamente significativas: tipo de verbo, tipo de sujeito e tipo de coleta, nessa ordem.

A análise estatística foi feita em termos de presença da DNP-P4. Entretanto, como o número de ocorrências de ausência de marca é pequeno, vou apresentá-los e discutir os resultados também em termos de ausência de marca. Os seis casos estão listados a seguir.

(60) ***nós três era*** *namorada dele.* (Cam06'219)
(61) ***nós tava*** *XXX* (Mat05'22)
(62) ***nós tava*** *brincando de lutar* (Mat05'23)
(63) *daí* ***eu e o meu primo tava*** *vendo o filme do Aladin* (Car15'79)
(64) ***foi eu a Beth e a Inês****.* (Ale03'126)
(65) *quando nós voltamo-0s nós pegamo-0s a mochila <e sen> [///] e* ***senta*** *na rodinha.* (Mat 01'40-46)

Batista e Zilles (2004) estudaram dois grupos de falantes, todos com formação universitária, de Porto Alegre, em duas décadas distintas (1970 e 1990), compondo um Estudo de Painel sobre a CV de primeira pessoa do plural. A pesquisa examina tanto a presença da DNP-P4 (padrão ou não padrão) em oposição à sua ausência quanto o apagamento do –s (–mo *versus* –mos). Os resultados da primeira análise atestam

um índice bastante baixo de ausência da desinência: 33/1195, sendo 3 ocorrências na década de 70 (1%) e 30 (5%) nos anos 90. É interessante ressaltar que a comunidade de fala a que se relaciona o estudo é comparável à das crianças em análise aqui. Os contextos de ausência de marca são discutidos em termos de tonicidade da forma alvo: dos 33 casos de ausência da desinência, 27 (82%) são de tempos verbais que produziriam uma forma proparoxítona. As autoras afirmam que mesmo o apagamento da DNP-P4 sendo estigmatizado (em relação ao sujeito expresso pelo pronome), o fato de a forma alvo da produção da fala ser uma proparoxítona propicia o apagamento da marca.

Como é possível perceber, dos seis casos de ausência da desinência, quatro (66%) são de formas que produziriam uma proparoxítona caso a DNP-P4 fosse realizada. Pensando no percentual de Batista e Zilles de 82% de ausência de marca em relação ao total de casos em que a forma alvo seria proparoxítona, o resultado das crianças é bastante próximo: 4/5 ocorrências ou 80% (as ocorrências são as quatro listadas acima e aquela de presença de DNP discutida em 6.3.4[41]). Certamente a variável tonicidade não foi escolhida como significativa pelo programa por causa do modo como procedi à codificação, ou seja, em função da forma produzida e não da forma alvo, como já comentei antes. Assim, a explicação do baixo índice de ausência da marca nos dados das crianças em termos de tonicidade da forma alvo não pode ser apreciada aqui em termos estatísticos, mas o exame das ocorrências dá suporte para uma análise desse tipo.

Apresento a seguir os resultados referentes às variáveis selecionadas.

[41] Na seção 6.3.4 aparecem somente três ocorrências de ausência de marca porque uma foi codificada como locução no tempo verbal, o que não permitiu que ela aparecesse quando cruzei ausência de marca com o pretérito imperfeito do indicativo.

5.3.2.1 Tipo de Verbo

Como já mencionado, a primeira variável selecionada foi tipo de verbo. Nesta rodada foram descartados os seguintes fatores: a forma verbal *Vamos*, que tal como esperado não apresenta nenhum caso de ausência de marca, assim como os verbos intransitivos e reflexivos, e também o verbo cópula, por conter apenas duas ocorrências. A variável, dessa forma, contou com dois fatores: verbos transitivos e verbos ergativos. Os resultados podem ser vistos na Tabela 9.

Tabela 9 – Relação entre a aplicação da DNP-P4 e a variável tipo de verbo

	Ocorrências	**/ Total**	**%**	**Peso Relativo**
transitivo	49	51	96	0,61
ergativo	4	7	57	0,03
TOTAL	**53**	**58**	**91**	

A distribuição dos percentuais e pesos relativos da variável revela que os verbos transitivos favorecem a manutenção da marca (96%, 0,61), enquanto que os ergativos a desfavorecem fortemente (57%, 0,03).

Esse resultado corrobora as expectativas. Os verbos transitivos, por sua característica de selecionar sujeito [+ agentivo], em grande parte das vezes engajam-se em estruturas SVO, favorecedoras da aplicação da desinência. Por outro lado, os verbos ergativos, de forma geral, são ambientes propícios a estruturas em que o verbo está anteposto ao sujeito. Entretanto, analisando os resultados de não aplicação da marca ((60) a (65)), tem-se que as construções com verbos ergativos dizem respeito a uma única forma

verbal, que justamente coincidem com três das ocorrências de verbo cuja forma alvo com a desinência seria proparoxítona ((61), (62) e (63)).

Dos outros dois casos acima ((64) e (65)), um diz respeito a uma forma VS, ambiente em que a ausência da desinência é esperada. Vale ressaltar que a variável posição do sujeito em relação ao verbo foi eliminada da rodada pela seguinte situação dos fatores: anterior adjacente: 39/43 ocorrências distribuídas entre presença (35) e ausência (4) de marca (90%); anterior distante: 3 casos de presença e nenhum de ausência; e posterior distante: 1 caso de ausência e nenhum de presença da desinência, ou seja, o único caso de verbo posposto ao sujeito é de ausência de marca, e entre os verbos antepostos a percentagem de verbos sem marca é baixa (10%). Digo isso porque talvez em relação à ocorrência (64) o que esteja pesando mais para ausência da desinência seja a influência da estrutura VS, mesmo que não tenha podido ser examinada pelo programa.

5. 3.2.2 Tipo de Sujeito

A variável tipo de sujeito foi rodada com os seguintes fatores: SN pleno, o pronome reto *nós* e a elipse do pronome. A distribuição dos resultados pode ser vista na Tabela 10 a seguir, mostrando o pronome (tanto em sua forma explícita quanto elíptica) como favorecedor da manutenção da DNP-P4 e o SN pleno como altamente desfavorecedor da aplicação da marca. É preciso salientar o problema de falta de ortogonalidade nessa variável, já que a distribuição do total de ocorrências entre os fatores é bastante desigual.

Tabela 10 – Relação entre a aplicação da DNP-P4 e a variável tipo de sujeito

	Ocorrências /	Total	%	Peso Relativo
pronome nulo	102	103	99	0,58
pronome reto	32	34	94	0,57
SN pleno	6	9	66	0,01
TOTAL	**140**	**146**	**95**	

Voltando à lista de ocorrências, os casos de SN pleno já foram analisados como tendo outra motivação para a ausência da marca do que, propriamente, o tipo de sujeito em questão. O número de ocorrências deste fator é muito baixo para uma análise estatística consistente, que possa refletir matematicamente a sua influência (e do grupo de fatores) na amostra. O que quero dizer com isso é que, olhando as produções em (63) e (64), não me parece que a motivação da ausência da marca seja o sujeito SN pleno.

Quanto aos outros fatores, os pesos relativos são bastante próximos, sugerindo uma influência quase coincidente sobre a manutenção da desinência.

5.3.2.3 Tipo de Coleta

A variável tipo de coleta queria medir a influência de diferenças estilísticas na presença/ausência da marca de concordância. A hipótese era que os fatores relacionados à coleta de narrativas (ou seja, narrativa ficcional, história em seqüência e leitura), em virtude do caráter mais formal relacionado à estrutura da entrevista, favoreceriam mais a manutenção da marca que a conversa informal e o relato pessoal. A distribuição dos resultados é apresenta na Tabela 11.

Tabela 11 – Relação entre a aplicação da DNP-P4 e a variável tipo de coleta

	Ocorrências	**/**	**Total**	**%**	**Peso Relativo**
conversa informal	86		87	98	0,75
relato pessoal	35		40	87	0,09
TOTAL	**121**	**/**	**127**	**95**	

A interpretação desses índices é bastante delicada. A rodada contou somente com os fatores conversa informal e relato pessoal, justamente os dois que poderiam ser amalgamados no rearranjo da variável. Os outros fatores foram retirados por não terem dados de ausência de desinência (fato que, de certa forma, corrobora a expectativa em relação a tais fatores). Os índices atestam uma diferença percentual média entre a conversa e o relato (98% e 87% respectivamente), mas uma diferença de peso relativo muito robusta, com a primeira favorecendo em 0,75 e o segundo desfavorecendo em 0,09. Assim, pode-se dizer que os índices do relato corroboram a expectativa, mas os da conversa, não. Desse modo, talvez seja mais adequado avaliar que não foi possível averiguar a hipótese de forma satisfatória.

5.3.3 Padrão *versus* não padrão: o apagamento do /s/

Na rodada que examina o apagamento do morfema de plural, duas foram as variáveis selecionadas pelo programa VARBRUL, na seguinte ordem: faixa etária e tipo de coleta. O grupo etário tendo sido selecionado, foi realizada outra rodada, desconsiderando as três primeiras faixas (d, 2 e 3). Entretanto, não ouve alteração dos resultados a não ser pelo índice do Input (0,22 e 0,34 respectivamente). Assim sendo, vou apresentar os

resultados da primeira rodada (relativa ao total de faixas etárias), para traçar alguns comentários a respeito das faixas de crianças menores.

5.3.3.1 Faixa Etária

A distribuição percentual e de peso relativo dos fatores revela que as crianças de 7 e 8 anos empregam mais a forma padrão que as outras, respectivamente, 60% - 0,79 e 68% - 0,83. Já as faixas etárias intermediárias, apontam pesos relativos próximos ao ponto neutro, com as crianças de 5 anos favorecendo a aplicação do padrão (0,51) e as de 6 anos desfavorecendo (0,46). Os resultados podem ser vistos na Tabela 12.

Tabela 12 – Relação entre a aplicação da DNP-P4 padrão e a faixa etária

	Ocorrências	**/ Total**	**%**	**Peso Relativo**
2;6 a 2;11	4	15	26	0,52
3;0 a 3;11	7	53	13	0,32
5 anos	2	9	22	0,51
6 anos	9	37	24	0,46
7 anos	9	15	60	0,79
8 anos	11	16	68	0,83
TOTAL	**42**	**146**	**28**	

Saliento que a faixa 4 não foi considerada na análise por não apresentar nenhum dado, assim como a faixa d, que computava apenas um. Nessa análise é preciso levar em conta que os números de ocorrências em cada faixa etária é baixo, sobretudo aos 5 anos. Entretanto, eles corroboram a expectativa de que com o aumento da faixa etária, correspondendo a uma maior experiência na escola, aumente a produção da criança em

direção ao uso da forma padrão. A faixa 6, que corresponde ao início da educação formal, mostra uma probabilidade maior de apagamento do /s/.

Mas é preciso ainda buscar uma explicação para a diminuição nos pesos relativos de 5 para 6 anos. Para tanto, fiz uma consulta à base de dados, cruzando as formas verbais em análise com os respectivos sujeitos. O resultado mostrou que as ocorrências correspondentes à faixa 5 são de três meninas, Camila, Natália e Alexandra, ao passo que na faixa 6, do total de 37 ocorrências, 26 são de um único menino, Matheus. Quando fazia a seleção dos dados, me chamou a atenção o fato de ele ser a única criança a produzir o pronome *nós* (elíptico) com o verbo na forma não marcada (de terceira pessoa do singular), e também a produzir o único caso da amostra de mudança na qualidade da vogal temática (que não está em estudo aqui) associada à DNP-P4. Cito, abaixo, o trecho de fala da criança que contém esses episódios:

> *ah@i nós chegamo-0s fazemo-0s uma rodinha <e daí> [/] e daí nós fizemos que nem o ajudante e depois nós brincamo-0s e daí depois <nós fizemo-0s um traba> [//] daí nós fazemo-0s um trabalho daí lavamo-0s as mão-0s* ***lanchemo****-0s e vamo-0s pra pracinha brincar quando nós voltamo-0s nós pegamo-0s a mochila <e sen> [//] e* ***senta*** *na rodinha* (Mat 01’40-46).

Na fala acima, de 12 formas produzidas de DNP-P4, apenas uma está na forma padrão, ou seja, provavelmente na faixa 6 são os dados desse menino que explicam o favorecimento da forma não padrão. Examinei o relatório do projeto DELICRI (Guimarães, 1995) em busca do perfil social dessa criança. É filho de pais com nível superior de escolaridade, ele, engenheiro civil, ela, odontóloga, estudou na Escola Projeto e no Colégio Americano, morava no bairro Rio Branco e a categoria sócio-econômica a que pertencia a

família era a considerada A pelo projeto. Com esses dados, não encontro explicação social para essas particularidades na fala da criança, já que essas formas são relacionadas, na literatura, a classes sociais mais baixas, com baixa escolarização[42].

Essa situação aponta duas questões. A primeira diz respeito a uma falha minha na pesquisa, qual seja, não ter tratado os informantes enquanto variável. Se houvesse codificação por informante seria possível avaliar a influência de individualidades das crianças no sistema de CV de sua fala.

A segunda questão a que me refiro aponta a necessidade de um controle mais rigoroso das variáveis sociais, desde a coleta dos dados. Em se tratando de informantes crianças, talvez além de informações sócio-econômicas consistentes, seja necessário mapear, por exemplo, a situação de outras pessoas que têm um convívio intenso com a criança, seja babá, avó, etc. Sabe-se que no processo de aquisição da linguagem é possível que apareçam na fala da criança formas que fazem parte da fala das pessoas que convivem diariamente com ela. Assim seria possível compreender o sistema de CV na fala de crianças de modo mais abrangente.

Voltando aos resultados, os números relativos à faixa 3 apontam uma robusta preferência pela forma com apagamento do /s/. Sobre isso, a consulta ao banco de dados mostrou que das 53 ocorrências relativas a essa faixa etária, 44 (83%) delas correspondem à forma *Vamos*. Esse resultado é bastante próximo àquele encontrado por Batista e Zilles (2004), segundo o qual essa forma verbal é responsável por 88% do total de apagamentos do /s/. As autoras atribuem o resultado ao processo de auxiliarização do verbo

[42] Fenômeno atestado, por exemplo, na comunidade de São Miguel dos Pretos, cf. Almeida (2005).

ir, em construções de perífrase de futuro e perífrase modal. Não investiguei aqui a forma verbal *Vamos* nesses termos; contudo, dos 75 casos do verbo na amostra, apenas 3 deles não estão funcionando como auxiliar nos tipos de construções tratadas pelas autoras. Assim sendo, parece bastante plausível atribuir a preferência pela forma não padrão na faixa 3 à presença maciça de *Vamo*.

Em relação à faixa 2, a situação é praticamente a mesma, porém com um total de dados menor. A totalidade das 15 ocorrências são construções de *Vamos*, 4 das quais mostrando a retenção do –s. Provavelmente a diferença estatística apontada pelo programa seja reflexo da diferença entre os números de ocorrências nas duas faixas, em relação ao Input de 0,22.

Vale ressaltar que a produção de *Vamos* por crianças pequenas está atestada na literatura psicolingüística como uma forma precoce em relação às demais formas verbais de plural: Perroni Simões e Stoell-Gammon (1977) não encontram flexão de plural entre 2;1 e 2;8 de idade, a não ser por esta forma. Rubino e Pine (1998) não discutem especificamente a forma *Vamos*, mas seus resultados apontam que aos 3 anos de idade a criança não produziu erros de concordância de flexão na produção relativa à primeira pessoa do plural, e os dois exemplos discutidos pelas autoras são construções com a forma *Vamos.* Assim, talvez seja interessante, em pesquisas futuras, examinar melhor a produção de *Vamos* por crianças pequenas, tanto do ponto de vista psicolingüístico quanto sociolingüístico.

5.3.3.2 Tipo de Coleta

Nessa rodada, devido aos problemas de distribuição dos dados, os fatores narrativa e história em seqüência foram amalgamados a relato pessoal, e os fatores leitura, leitura de desenho e música foram retirados. Assim sendo, os resultados opõem dois fatores: conversa informal e relato pessoal. Os números podem ser conferidos na Tabela 13.

Tabela 13 – Relação entre a aplicação da DNP-P4 padrão e o tipo de coleta

	Ocorrências	**/ Total**	**%**	**Peso Relativo**
conversa informal	15	87	17	0,53
relato pessoal	13	45	28	0,45
TOTAL	**28**	**132**	**21**	

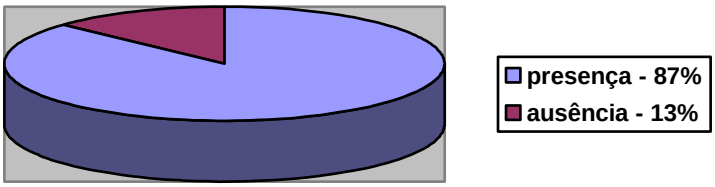
O problema analítico é o mesmo que aquele discutido em 6.3.2.3. Aqui, a diferença é que os dados dos fatores que foram amalgamados ao relato elevaram o peso relativo deste para próximo do ponto estatisticamente neutro. Os resultados apontam um índice fraco de favorecimento das formas padrão (17% e peso de 0,53) nas conversas e um índice igualmente fraco de favorecimento das formas não padrão (28% e peso de 0,45) nos relatos. Novamente, os problemas distribucionais impedem que se verifique a atuação da variável na amostra.

Passo, na próxima seção, à análise da DNP-P6.

5.4 A concordância de terceira pessoa do plural

A análise da concordância de terceira pessoa do plural será feita em termos de presença da DNP-P6. Do total de 529 ocorrências, 461 (87%) apresentaram a aplicação da marca, enquanto que 68 (13%) não a apresentaram, conforme mostra o Gráfico 4.

Gráfico 4 – Presença *versus* ausência da DNP-P6



O alto índice de aplicação da desinência na fala das crianças aqui observadas corresponde ao comportamento lingüístico observado em dados de adultos com maior escolarização e classe social mais alta. Os resultados apresentados no estudo de Rodrigues (1992) sobre falantes de classe baixa, menos escolarizados, apresentam índices de aplicação de 29%, ao passo que os falantes mais escolarizados (da comunidade de Porto Alegre, tal como as crianças em análise), representados na pesquisa de Batista e Zilles (2005), aplicam a desinência em 95% das ocorrências. Assim, provavelmente o ambiente de letramento em que vivem as crianças da amostra influencia o percentual de aplicação da marca de DNP-P6 em sua fala.

Dentre as variáveis testadas, sete foram selecionadas pelo programa VARBRUL, na seguinte ordem: posição do sujeito em relação ao verbo, tonicidade, saliência fônica, tipo de verbo, tipo de sujeito, gênero e faixa etária. A seguir, apresento os

resultados examinando a influência dessas variáveis na aplicação da marca de concordância.

5.4.1 Posição do Sujeito em relação ao Verbo

Essa variável teve os fatores posterior distante e posterior adjacente amalgamados, assim como os fatores anterior adjacente e anterior distante. Assim, a análise se deteve na oposição entre sujeito anterior e posterior ao verbo. Os resultados, vistos na Tabela 14, apontam uma forte tendência de o sujeito posposto favorecer a ausência da desinência e de o sujeito anteposto favorecer sua presença, corroborando a hipótese inicial.

Tabela 14 – Relação entre a aplicação da DNP-P6 e a variável posição do sujeito

	Ocorrências	**/**	**Total**	**%**	**Peso Relativo**
anterior	297	/	307	96	**0,63**
posterior	18	/	61	29	**0,07**
TOTAL	**315**	**/**	**368**	**85**	

Os estudos de Lemle e Naro (1977) atestam percentuais e pesos relativos de 49% - 0,70 de presença da desinência em sujeitos pré-postos e 23,1% - 0,22 de desfavorecimento da desinência em sujeitos pospostos. Bardem (2004) encontra índices ainda mais polarizados: 82% - 0,63 para sujeitos antepostos e 34% - 0,11 para sujeitos pospostos em relação à aplicação. Batista e Zilles (2005) descrevem os percentuais e pesos em relação à ausência da desinência: 3% - 0,43 para anteposto e 23% - 0,89 para posposto. Assim, os resultados mostram que o comportamento lingüístico das crianças no tocante à posição do sujeito em relação ao verbo é comparável ao dos adultos.

5.4.2 Tonicidade

Contrariamente à minha expectativa, a distribuição dos resultados em relação à tonicidade aponta as paroxítonas como favorecedoras da manutenção da marca e os monossílabos tônicos e as oxítonas como fortemente desfavorecedores, tal como mostra a Tabela 15.

Tabela 15 – Relação entre a aplicação da DNP-P6 e a variável tonicidade

	Ocorrências	**/**	**Total**	**%**	**Peso Relativo**
paroxítona	401	/	433	92	0,77
monossílabo tônico	58	/	80	72	0,01
oxítona	2	/	16	12	0,00
TOTAL	**461**	**/**	**529**	**87**	

Os pesos relativos dos dois últimos fatores são impressionantemente baixos (0,01 e 0,00, respectivamente). Para entender os resultados recorri novamente ao banco de dados e me deparei mais uma vez com o problema de ter codificado sobre a forma produzida, e não sobre a forma alvo. Em relação às oxítonas, se tivesse procedido da outra maneira, o total de ocorrências teria sido codificado como paroxítona, não havendo nenhum caso de oxítona na amostra em relação à DNP-P6. No tocante aos monossílabos tônicos, três casos somente teriam sido recodificados como paroxítonas. Entretanto, das 22 ocorrências de ausência de marca, 15 são casos de sujeito posposto. Como foi visto no item anterior, pesquisas sobre adultos (Lemle e Naro, 1977; Rodrigues, 1992; Monguilhott e Coelho 2002) atestam que o sujeito anteposto favorece a marca, enquanto que o posposto desfavorece. Além disso, a variável Posição do Sujeito foi selecionada em primeiro lugar

como influenciadora do percentual de aplicação da DNP-P6 para esta amostra. Assim, atribuo o baixo percentual dos monossílabos à atuação das ocorrências de sujeito posposto.

Com essa análise, sobretudo pelo procedimento de codificação, creio ser difícil estabelecer aqui de modo consistente a atuação da variável tonicidade.

5.4.3 Saliência Fônica

A terceira variável escolhida pelo programa foi saliência fônica.

Tabela 16 – Relação entre a aplicação da DNP-P6 e a variável saliência fônica

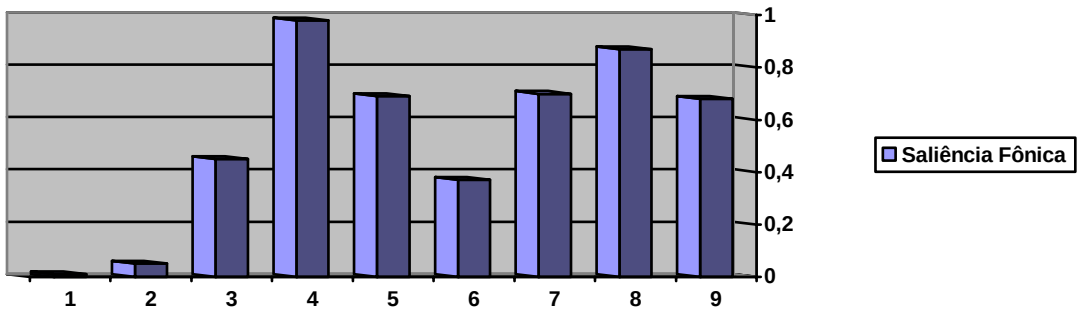
	Ocorrências	/	Total	%	Peso Relativo
4 – *dá/dão, está/estão, falará/falarão*	46	/	55	83	0,98
8 – caso único: *foi/foram*	11	/	19	57	0,87
7 – caso único: *é/são*	34	/	36	94	0,70
5 – *sumiu/sumiram, comeu/comeram, falou/falaram*	216	/	229	94	0,69
9 – caso único: *veio/vieram*	5	/	6	83	0,68
3 – *faz/fazem, quer/querem.*	12	/	15	80	0,45
6 – *fez/fizeram, teve/tiveram*	19	/	20	95	0,37
2 – *fala/falam, ia/iam.*	113	/	140	80	0,05
1 – *come/comem, fale/falem*	5	/	8	62	0,01
TOTAL	**461**	**/**	**528**	**87**	

Os resultados, vistos na Tabela 16, mostram, de forma geral, que não houve aumento progressivo de aplicação de marca em relação à escala, ou seja, que a hierarquia de saliência fônica não atuou como esperado[43].

Os índices dos primeiros níveis (1 e 2) se mostram em concordância com a hipótese, com os pesos relativos (0,01 e 0,05) respectivamente mais baixos dentre todos os níveis, indicando um desfavorecimento robusto da presença da marca. O peso relativo do terceiro nível (0,45) sobe bastante em relação aos dois primeiros, mas ainda se encontra na faixa de desfavorecimento da marca. No quarto nível, entretanto, o peso dispara para 0,98, mostrando ser o nível com maior índice de favorecimento de aplicação da desinência. É preciso notar que a rodada tem problemas de ortogonalidade, já que os números de ocorrências em cada nível são muito díspares e isso pode estar afetando os resultados. O peso do nível 5 cai para 0,69, ainda elegendo a aplicação da desinência. A seguir, há uma queda brusca no sexto nível (0,37), estimulando a ausência da marca. O sétimo e oitavo níveis mostram pesos crescentes de emprego da desinência (0,70 e 0,87), enquanto que no último nível o peso diminui um pouco (0,69), mas ainda prevalece a marca. O gráfico 5 expõe os resultados ordenados por nível.

[43] Ver reflexão sobre a variável saliência fônica na seção 2.2.

Gráfico 5 – Atuação dos níveis de saliência fônica na aplicação da DNP-P6

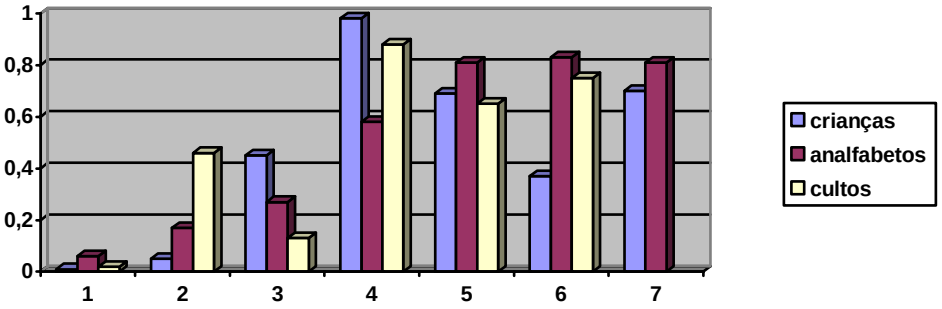


O gráfico mostra que o aumento da marca de concordância se deu progressivamente nos quatro primeiros níveis. Além disso, os níveis 5, 7 e 9 manifestaram pesos relativos muito próximos. Entre os níveis 5 e 6 há uma queda brusca, provocada, talvez, pela diferença numérica de ocorrências entre eles, que no nível 5 é de 229 e no 6, apenas 20.

Lemle e Naro (1977) e Naro (1981) apontam o fato de haver dois graus de atuação da saliência fônica em relação ao acento da forma verbal. No primeiro, o acento do verbo é na raiz; no outro, é na desinência. Assim, nos três primeiros níveis tem-se o contraste entre singular e plural somente na desinência, não coincidindo com o acento. O quarto nível é o primeiro cujo contraste está na sobreposição da desinência com o acento. Nos resultados obtidos por Lemle e Naro (1977), tem-se que o quarto nível é o primeiro a apresentar favorecimento da aplicação de concordância em relação aos outros três primeiros. Monguilhott e Coelho (2002) examinaram a fala de informantes cultos de Florianópolis e, em relação à saliência fônica, encontraram o maior peso relativo de ocorrências com marca no nível quatro, tal como nos dados das crianças aqui consideradas. O gráfico 5 representa um panorama da aplicação da marca da DNP-P6 relativamente à

saliência fônica da amostra, no grupo de analfabetos do Rio de Janeiro (Lemle e Naro, 1977) e no grupo de falantes cultos de Florianópolis (Monguilhott e Coelho, 2002).

Gráfico 6 – Atuação em peso relativo dos níveis de saliência fônica na aplicação da DNP-P6 em diferentes pesquisas



Nos três grupos, os três primeiros níveis se mostram favorecedores da ausência da desinência ao passo que os outros níveis apontam favorecimento da aplicação da marca. O único caso destoante com esse panorama é o resultado das crianças no nível 6, fato atribuído ao baixo número de ocorrências. Além disso, não há dados no nível 7 de Monguilhott e Coelho, não permitindo comparação. Em relação a tal panorama, não é possível dizer que a produção das crianças se aproxima mais da fala de uma ou de outra comunidade em relação à saliência fônica. Além disso, como salientaram Loregian (1996 e 2001) e Amaral (2002), há outros fatores atuando nessa variável que precisariam ser examinados juntamente com uma análise mais refinada dos dados.

Por fim, de forma geral, pode-se dizer que a hipótese da saliência fônica de proporcionalidade entre o nível de saliência e o peso de favorecimento da marca atuou nos quatro primeiros níveis e, sobretudo, em relação aos dois graus referentes à saliência no

acento, fazendo com que os três primeiros níveis influenciassem na ausência da marca e os outros (com exceção do nível 6) influenciassem no sentido de manutenção da marca.

5.4.4 Tipo de Verbo

A variável Tipo de Verbo apresenta resultados condizentes com a hipótese inicial. A Tabela 17 expõe os índices percentuais e os pesos relativos referentes à aplicação da DNP-P6.

Tabela 17 – Relação entre a aplicação da DNP-P6 e a variável tipo de verbo

	Ocorrências	**/**	**Total**	**%**	**Peso Relativo**
intransitivo e reflexivo	82	/	90	91	**0,82**
transitivo	305	/	319	95	**0,52**
cópula	22	/	33	66	**0,51**
ergativo	49	/	70	70	**0,16**
Ter/*Haver*	3	/	17	17	**0,07**
TOTAL	**461**	**/**	**529**	**87**	

Os verbos intransitivos e reflexivos se mostram altamente favorecedores à aplicação da desinência com percentual de 91% e peso relativo de 0,82. Já os transitivos e o verbo cópula manifestam pesos muito próximos entre si (0,52 e 0,51), e próximos do ponto neutro, indicando fraco favorecimento à presença da marca. Já os ergativos e os verbos *Ter*/*Haver* em construção apresentativa, ambos com tendência a construções VS, desfavorecem a marca (pesos de 0,16 e 0,07 respectivamente), como esperado.

Os resultados das crianças mostraram-se diferentes daqueles encontrados por Monguilhott e Coelho (2002). As autoras exibem os seguintes índices de percentuais e pesos relativos em relação à marcação da concordância: cópula, 78% e 0,63; transitivos, 82% e 0,49; intransitivos, 82% e 0,46; e inacusativos, 71% e 0,37.

5.4.5 Tipo de sujeito

Os tipos de sujeitos considerados nesta rodada foram: pronome reto, indefinido, SN pleno, pronome nulo, pronome demonstrativo e genérico. Os genéricos (14 ocorrências de presença de marca) foram amalgamados aos demonstrativos (1 única ocorrência de ausência de marca). Os resultados corroboram a hipótese inicial de que os sujeitos expressos por pronome reto favorecem a manutenção da marca de concordância, ao passo que pronomes indefinidos e SNs plenos a desfavorecem. Os índices de percentuais e pesos relativos podem ser vistos na Tabela 18 a seguir.

Tabela 18 – Relação entre a aplicação da DNP-P6 e a variável tipo de sujeito

	Ocorrências	**/**	**Total**	**%**	**Peso Relativo**
pronome reto	193	/	194	99	**0,90**
pronome indefinido	7	/	8	87	**0,35**
SN pleno	116	/	166	69	**0,25**
pronome nulo	131	/	146	89	**0,21**
pronome demonstrativo + genérico	14	/	15	93	**0,05**
TOTAL	**461**	**/**	**529**	**87**	

Os pesos relativos mostram uma forte polaridade entre as ocorrências de pronome reto, favorecendo enormemente a aplicação da marca (0,90), e os demais tipos de

sujeitos, que a desfavorecem (indefinidos, 0,35; SNs plenos, 0,25; pronomes nulos, 0,21; e demonstrativos somados a nulos genéricos, 0,05). Monguilhott e Coellho (2002) encontraram o maior percentual de aplicação de marca também entre os sujeitos expressos por pronome pessoal (amalgamados com os demonstrativos), indicando relação entre o tipo de sujeito e a posição ocupada em relação ao verbo.

Uma consulta ao banco de dados relativamente a demonstrativos e genéricos mostrou que o único caso de ausência de marca é referente ao demonstrativo, enquanto que os genéricos, todos apresentando elipse de sujeito, apontaram 100% de presença de marca. As ocorrências estão listadas a seguir:

(66) *essas daqui é só uma.*
(67) *a historinha que me contaram*
(68) *daí alugaram um filme do Bambi*
(69) *daí ligaram pra gente*
(70) *daí ligaram dizendo*
(71) *e disseram que o nenê nasceu .*
(72) *roubaram o carro da minha mãe*
(73) *roubaram ali na frente só tiraram né@i a tarde*
(74) *e acharam ontem o carro num terreno baldio .*
(75) *tinham roubado o rádio*
(76) *disseram que não .*
(77) *ele falou assim que depenaram o carro do pai dele #*
(78) *daí ligaram pra casa sete horas da manhã*
(79) *dizem que se ele olhar pro próprio pé ele morre*

A ocorrência em (66) mostra que entre o núcleo do SN (o pronome demonstrativo) e o verbo há um advérbio no singular, que pode ser o responsável pela desinência não marcada do verbo, que estaria sendo regido pelo elemento imediatamente anterior ao invés de sê-lo pelo núcleo do sujeito. Por se tratar de uma única ocorrência de pronome demonstrativo, e com a particularidade de conter material interveniente entre o pronome e o verbo, penso que não é possível tirar qualquer conclusão acerca da atuação dos

pronomes demonstrativos sobre a marcação de concordância de terceira pessoa do plural na amostra aqui considerada.

Sobre o nulo genérico, tem-se que o traço de especificidade que opõe um referente anafórico a um genérico reflete a possibilidade ou não de o referente ser recuperável por uma anáfora (Callou et al., 2000). Assim sendo, se o falante opta pelo uso genérico de plural com sujeito nulo, não parece haver a possibilidade de equivalência semântica na forma cuja DNP-P6 está ausente, já que a informação de plural não pode ser recuperada. Assim, em construções de sujeito nulo e genérico ([-] específico) parece esperável a presença da marca de forma categórica, o que acontece nos dados. Essas ocorrências deveriam ter sido amalgamadas aos pronomes nulos. Assim, a análise apresentada na tabela em relação ao fator demonstrativo + genérico fica sem sentido, com um índice extremamente baixo de aplicação da marca.

A respeito do sujeito pronominal nulo, Rodrigues (1992) argumenta que, funcionalmente, a elipse do sujeito deve favorecer a aplicação da DNP, em virtude da possível perda de informação de plural se a desinência não for marcada. Essa hipótese foi corroborada em seu trabalho (52% e peso relativo de 0,62), e também pelos percentuais e pesos relativos de outras pesquisas: Lemle e Naro (1977), 54% e 0,65; Naro (1981), 50% e 0,65; Bortoni-Ricardo (1985), 48%. A produção das crianças se mostrou contrária a essa hipótese, com os nulos desfavorecendo a aplicação da marca. Examinei as ocorrências de ausência de marca com sujeito nulo e vi que 9/15 (60%) são casos de categoria vazia produzidas pela retomada do referente por pronome relativo. Monguilhott e Coelho (2002) atestam um percentual de 83% e peso relativo de 0,47 no que tange a esse tipo de sujeito,

marcando desfavorecimento da marca. Talvez a preferência das crianças por sujeito nulo associado à não marcação, seja explicada pelos casos de atuação do pronome relativo.

5.4.6 Gênero

A sexta variável selecionada pelo programa foi gênero. O construto teórico variacionista afirma que as mulheres são mais sensíveis às formas de maior prestígio na comunidade. Os resultados obtidos a partir da fala das crianças, apresentados na Tabela 19, mostram que as meninas usam a DNP-P6 em 90% dos casos, com peso relativo de 0,68, conduzindo, portanto, à preferência pela marca, ao passo que os meninos usam em 83% dos casos, com peso relativo de 0,30, indicando preferência pela ausência da marca. Simões (2003) examinou a concordância nominal de número, opondo as formas padrão e não padrão em SNs de dados de cinco das sete crianças do projeto DELICRI examinadas aqui, relativamente a três situações discursivas, opondo relato pessoal a história ficcional e história em seqüência. Os resultados mostraram percentuais distintos entre meninos e meninas, com as meninas produzindo mais formas padrão que os meninos em ambas as situações discursivas. As meninas apresentaram freqüência de 48,5% (99/204) nos relatos e 75% (123/164) contando histórias, e os meninos, 23,5% (33/104) e 51,4% (72/140), respectivamente.

Tabela 19 – Relação entre a aplicação da DNP-P6 e a variável gênero

	Ocorrências	**/**	**Total**	**%**	**Peso Relativo**
Feminino	253	/	279	90	**0,68**
Masculino	208	/	250	83	**0,30**
TOTAL	**461**	**/**	**529**	**87**	

Com esses índices, pode-se dizer que as meninas da amostra seguem a tendência geral que ocorre nos fenômenos lingüísticos variáveis de opção pelo padrão. Entretanto, alguns trabalhos relativos à CV variável atestam não haver diferenças relevantes entre a produção de homens e mulheres no tocante ao uso da DNP-P6 (Naro, 1981; Rodrigues, 1992; Almeida, 2005).

Segundo Paiva (2003:35), a tendência atribuída às diferenças lingüísticas entre os gêneros aponta para "*uma maior consciência feminina do* status *social das formas lingüísticas*", havendo necessidade de a pesquisa variacionista contemplar, além do prestígio atribuído a certas variáveis, também o papel ocupado por cada gênero na organização social da comunidade de fala. Andersen (1990) mostrou, num experimento sobre representação de papéis sociais pelas crianças, que os meninos davam preferência a papéis de maior *status*, enquanto que as meninas se dispunham a qualquer papel. Além disso, as meninas apresentaram uma identificação maior com o papel da professora. Segundo a autora, isso mostra que as crianças são sensíveis desde cedo aos processos de socialização relativos aos gêneros. Assim, apesar de os resultados obtidos corroborarem a hipótese de que as meninas apresentariam mais marcas de concordância, para uma compreensão mais acurada do fenômeno em relação aos gêneros é importante uma análise mais sofisticada do papel social de homens e mulheres na comunidade de fala na qual essas crianças se inserem.

5.4.7 Faixa Etária

A última variável selecionada nessa rodada foi faixa etária. O grupo 4 foi amalgamado ao 5 por conter apenas 5 ocorrências (todas de aplicação da marca). A Tabela 20 expõe os resultados.

Tabela 20 – Relação entre a aplicação da DNP-P6 e a variável faixa etária

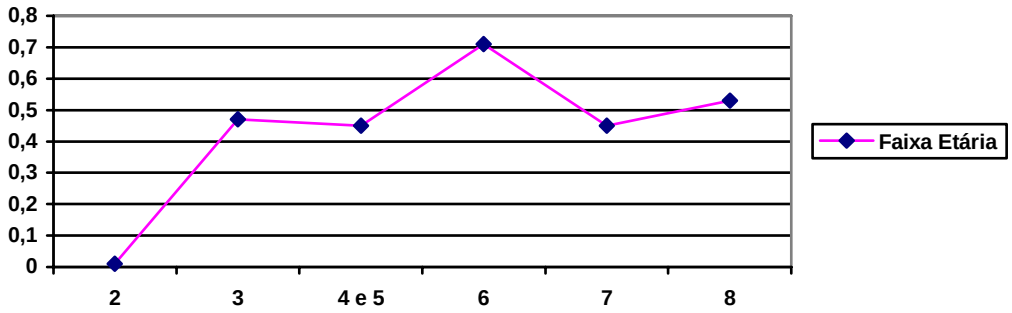
	Ocorrências	/	Total	%	Peso Relativo
2 – 2;6 a 2;11	2	/	6	33	**0,01**
3 – 3;0 a 3;11	91	/	100	91	**0,47**
5 – 4 e 5 anos	35	/	41	85	**0,45**
6 – 6 anos	65	/	80	81	**0,71**
7 – 7 anos	169	/	191	88	**0,45**
8 – 8 anos	99	/	111	89	**0,53**
TOTAL	**461**	**/**	**529**	**87**	

A faixa 2, mostrando um percentual de 33% com um peso relativo de 0,01, indica forte desfavorecimento de aplicação da marca. Isso pode indicar a aquisição tardia da DNP-P6 em relação às outras desinências de CV. Além disso, a faixa contém um número bastante baixo de ocorrências (apenas 6), o que poderia também explicar tais resultados.

O grupo etário 3 foi constituído somente por dados de Artur, que apresentou o maior percentual de aplicação da marca (91%), embora o peso relativo indique desfavorecimento pela marcação. Aqui vale comentar os achados de Rubino e Pine (1998) sobre a proporção de erros de flexão verbal em uma criança entre 3;02 e 3;04 de idade. Os resultados revelaram (Tabela 5, p. 44) que a maior proporção de erros (de desinência de número ou pessoa) se dá na terceira pessoa do plural (10/23 – 43,5%). Isso novamente indica que esse pode ser tardiamente dominado pela criança. Os resultados aqui obtidos possivelmente refletem a fala particular de Artur.

O gráfico 7 exibe um panorama geral de aplicação da DNP-P6 nos grupos etários em termos de peso relativo.

Gráfico 7 – Atuação da faixa etária na aplicação da DNP-P6 em termos de peso relativo



O grupo 6 apresentou a maior influência na aplicação da DNP-P6, com percentual de 81% e peso relativo de 0,71. Uma explicação possível para o fato é o início do contato com a língua escrita, gerando implemento no uso da forma padrão. Os grupos 4/5 e 7, com percentuais de 85% e 88%, respectivamente, mostram o mesmo valor de peso relativo (0,45), enquanto que no grupo 8 os percentuais sobem levemente (89% e 0,53 de peso relativo), dessa vez em favor da desinência.

No gráfico 6 (p. 145) vê-se que há pouca diferença entre as faixas 3, 4/5, 7 e 8 (embora a 8 favoreça e as outras desfavoreçam a marca), todas próximas do ponto neutro. Ao que parece, com exceção da faixa 6, as crianças não estão focadas na DNP-P6.

Por outro lado, o Gráfico 7 (p. 153) confirma, de certa forma, o Gráfico 1, relativo aos resultados gerais por faixa etária, de uma curva em U, com um pico de aquisição, e uma certa estabilidade girando em torno do ponto neutro.

CONSIDERAÇÕES FINAIS

A pesquisa aqui desenvolvida debruçou-se sobre o sistema de concordância verbal variável na produção oral infantil. Inicialmente, tracei um quadro geral do sistema flexional na fala das crianças para, a partir daí, conduzir a investigação sobre a regra variável de desinência verbal. Os parâmetros levados em conta tanto para a construção das hipóteses quanto para a avaliação dos resultados obtidos foram extraídos, na maior parte, da literatura sociolingüística quantitativa sobre o fenômeno de CV na fala adulta, em função da metodologia analítica aqui empregada, bem como da literatura psicolingüística, quando possível.

A pesquisa foi executada a partir de dados de nove crianças, cinco meninas e quatro meninos, com idades entre 2;5 e 8 anos, todas da comunidade de Porto Alegre. Os dados foram tratados numa perspectiva quantitativa seguindo rodadas do pacote VARBRUL de análise estatística sob quatro focos: o primeiro, geral, de aplicação *versus* não aplicação da marca, considerando o sistema flexional de primeira, segunda e terceira

pessoas do singular e primeira e terceira do plural; dois relativos à DNP-P4 – um opondo presença e ausência de marca e outro opondo as formas padrão e não padrão –, e ainda um relativo a presença *versus* ausência da DNP-P6.

Na rodada geral, os percentuais apontaram 66% de presença da marca de concordância e 34 % de ausência. Os grupos de fatores selecionados como relevantes pelo programa estatístico foram, na ordem, tipo de sujeito, tempo verbal, posição do sujeito em relação ao verbo, tonicidade, tipo de verbo e faixa etária. A interpretação dos resultados nessa rodada se deu em termos quantitativos e também qualitativos em relação à amostra, já que o tratamento reunindo as cinco pessoas gramaticais não consegue ser claro no tocante a cada pessoa distintamente. Os resultados dessa rodada tiveram, acima de tudo, um caráter descritivo geral de aplicação da marca de concordância, com o intuito de apontar alguns caminhos em pesquisas futuras. Eles mostram, por exemplo, que as formas de variação aquisicional na primeira pessoa do singular perduram até quase a idade escolar; que o exame da regra variável da DNP-P2 só é possível em dados coletados com elicitação dessa forma verbal; que a forma *a gente* com emprego da DNP-P4 é pouco usual (nessa amostra) e não é encontrada na fala das crianças em idade pré-escolar; que a partir de 2;5 de idade a fala das crianças já está sujeita à atuação de regras variáveis e, de modo geral, mostra um comportamento lingüístico, em relação a tais regras, próximo do adulto.

No tocante ao tipo de sujeito, todos os fatores se mostraram favoráveis à aplicação da marca, com exceção da forma pronominal *a gente*. Os tempos verbais que implicaram em manutenção da marca foram o futuro perifrástico (*ir + infinitivo*), com 89% e 0,76 de peso relativo, e perfeito do indicativo, com 71% e peso relativo de 0,74, formas atestadas por Simões (1997) (juntamente com o presente do indicativo) como presentes na

fala do sujeito de sua pesquisa desde os 2;4 de idade. Os demais tempos verbais testados (infinitivo, presente do indicativo, locução, imperfeito do indicativo e imperativo) mostraram influenciar o sistema de CV no sentido de não aplicação da desinência. Os resultados da variável posição do sujeito em relação ao verbo indicaram que sua influência se dá mais em termos de anterioridade e posterioridade do que desses fatores associados a material interveniente. Os índices referentes à tonicidade apontaram as paroxítonas como influenciadoras da presença da desinência (76% e 0,63 de peso relativo) e os monossílabos tônicos e oxítonas conduzindo à ausência. O único caso de proparoxítona com desinência é discutido em termos de maior formalidade da estrutura. A variável tipo de verbo apontou a forma *Vamos* como a que mais favorece a aplicação da marca, com percentual de 98% e peso relativo de 0,98, em conformidade com o atestado por Zilles e Batista (2005). Os dois fatores que consideravam verbos inacusativos (ergativos e *Ter/Haver*) apresentaram, conforme esperado, desfavorecimento da marca. O verbo cópula manifestou imparcialidade, com percentual de 59% e peso relativo de 0,50, enquanto que os transitivos influíram levemente na ausência e os intransitivos/reflexivos, na presença da desinência. A única variável social selecionada foi faixa etária, que, devido a problemas na constituição da amostra, manifestou resultados que talvez não sejam condizentes com a realidade lingüística dos informantes. O mais concreto em termos de análise parece ser os resultados das faixas 5 a 8 anos, que mostraram baixa relevância estatística, com índices próximos ao ponto neutro de aplicação das DNPs.

No cômputo geral da análise das DNPs P4 e P6, que tiveram seus resultados comparados aos da fala adulta, o quadro que se mostra é de que as crianças aqui examinadas apresentam um comportamento lingüístico comparável ao dos adultos. Isso indica que na fala das crianças em estudo a variação relacionada ao sistema de

concordância verbal é tanto de ordem sintática quanto social. O fato de os resultados das crianças corroborarem a maioria das hipóteses formuladas sobre dados de adultos mostra que a ação das variáveis analisadas, bem como a simultaneidade de forças entre elas, dizem respeito antes ao sistema lingüístico do PB como um todo, que à fala dos adultos e das crianças especificamente.

O trabalho teve limitações das seguintes ordens:

1. Os dados não foram coletados com vistas ao estudo específico da concordância verbal variável. Tal fato imputou limitações no tocante às formas de plural, devido ao baixo número de ocorrências, sobretudo nas entrevistas das crianças mais novas;

2. A amostra apresenta problemas de estruturação nas faixas etárias de 2 e 3 anos devido à disparidade de entrevistas consideradas. Além dos dados das duas crianças pequenas, o ideal seria ainda a inclusão de outras crianças nesses grupos etários para compor uma amostra mais representativa;

3. A não codificação por informante impediu que a análise estatística contemplasse particularidades das crianças. Além disso, um controle maior das variáveis sociais seria importante para uma maior compreensão da dimensão social da variação na fala das crianças;

4. A variável escolaridade (que nas crianças deveria incluir informações também da pré-escola) não foi contemplada aqui e possivelmente revelaria nuances de influência na produção das crianças;

5. Algumas variáveis mostraram que operam juntamente com outras, numa conjunção de forças. Para melhor delimitar a influência de cada uma, seriam necessárias análises mais refinadas cruzando tais variáveis (por exemplo, saliência fônica, tempo verbal e tonicidade, ou tipo de verbo e posição do sujeito em relação ao verbo).

REFERÊNCIAS BIBIOGRÁFICAS

Alkmim, T. (2003) Sociolingüística: Parte I. In: Mussalin, F. e Bentes, A. C. (orgs.) *Introdução à lingüística: domínios e fronteiras,* vol.I. São Paulo: Cortez.

Almeida, A. P. de (2005) *A concordância verbal na comunidade de São Miguel dos Pretos, Restinga Seca, RS*. Porto Alegre: UFRGS. Dissertação de Mestrado.

Amaral, L. I. C. (2002) A importância de variáveis estilístico-discursivas para a análise de fenômenos lingüísticos. In: Vandresen, P. (org.) *Variação e mudança no português falado na Região Sul*. Pelotas: EDUCAT. p.47-67.

_____ (2003) *A concordância verbal de segunda pessoa do singular em Pelotas e suas implicações lingüísticas e sociais*. Porto Alegre: UFRGS. Tese de Doutorado.

Barbisan, L. B. e Machado, R. F. (2000) O tópico no texto argumentativo. *Letras de Hoje*. Porto Alegre. vol.35, n. 3, p.69-106.

Barden, L. T. V. (2004) *A variação na concordância verbal de terceira pessoa do plural*. Porto Alegre: PUCRS. Dissertação de Mestrado.

Batista, H. H. R. B. e Zilles, A. M. S. (2004) A concordância verbal de primeira pessoa do plural na fala culta de Porto Alegre. Trabalho apresentado no 6º Encontro do CELSUL, mimeo.

Batista, H. H. R. B. e Zilles, A. M. S. (2005) A concordância verbal de terceira pessoa do plural em dados de Porto Alegre: um Estudo de Painel. Trabalho apresentado no XVII Salão de Iniciação Científica da UFRGS, Porto Alegre, mimeo.

Bortoni-Ricardo, S. M. (1984) Problemas de comunicação interdialetal. *Tempo Brasileiro*, Rio de Janeiro, n.78, p. 9-32.

Bortoni-Ricardo (1985) *The urbanization of rural dialect speakers. A sociolinguistic study in Brazil*. Cambridge: Cambridge University Press.

Brown, R. (1973) *A first language: the early stages*. London: George Allen e Unwin.

Callou, D., Portela, K., Avelar, J. e Serra, C. (2000). Dinâmica do específico e do genérico: artigo definido e construções existenciais. *Veredas*, vol. 4, n. 2, p.80-88.

Camacho, R. G. (1993) Aspectos funcionais e estruturais da concordância verbal no português falado. *Alfa*, vol.37, p.101-116.

_____. (2003) Sociolingüística: Parte II. In: Mussalin, F. e Bentes, A. C. (orgs.) *Introdução à lingüística: domínios e fronteiras,* vol.I. São Paulo: Cortez.

Câmara Jr., J. M. (1982) *Estrutura da Língua Portuguesa*. Petrópolis: Vozes, 11ª ed.

_____. (1988) *Problemas de Lingüística Descritiva*. Petrópolis: Vozes.

Clahsen, H., Penke, M. e Parodi, T. (1993/1994) Functional Categories in Early Child German. *Language Acquisition*, vol.3, n.4, p.395-429.

Cunha, C. (1972) *Gramática da Língua Portuguesa.* Rio de Janeiro: FENAME/ Ministério da Educação e Cultura.

Decat, M. B. (1981) Verbal agreement differences in spoken and written Brasilian Portuguese and their consequences for the teaching of composition. *Ensaios de Lingüística*, vol.5, p. 25-39.

_____. (1983) Concordância verbal, topicalização e posposição do sujeito. *Ensaios de Lingüística*, vol.9, p. 91-148.

Duarte, Y. (1993) A hipótese inacusativa e as evidências do português. *D.E.L.T.A.*, vol.9, n.1, p. 31-58.

Faria, I. H. (1993) A aquisição da noção de “agente” e a produção de sujeitos sintáticos por crianças portuguesas até aos dois anos e meio. *Revista Internacional de Língua Portuguesa*, Lisboa, n. 10, p.16-50.

Figueiredo Silva, M. C. (1996) *A posição do sujeito no português brasileiro: frases finitas e infinitivas*. Campinas: Editora da UNICAMP.

Givón, T. (1976) Topic, pronoun and grammatical agreement. In: Charles Li Ed. *Subject and Topic*. New York: Academic Press.

_____ . (1984) *Sintax: a functinal-typological introduction.* Vol.I. Amsterdam: John Benjamins.

_____ **.** (1999) Generativity and Variation: The Notion 'Rule of Grammar' Revisited. In: MacWhinney, B. (org) *The Emergence of Language.* Mahwah, NJ: Lawrence Erlbaum.

Guasti, M. T. (1993/1994) Verb syntax in Italian child grammar: finite and nonfinite verbs. *Language Acquisition,* vol.3, n.1, p.1-40.

_____. (2002) *Language acquisition: the growth of grammar*. Cambridge, MA: The MIT Press.

Guimarães, A. (1995) *Relatório: Projeto Desenvolvimento da Linguagem da Criança em Fase de Letramento*. Porto Alegre, mimeo.

_____ (1996) Conferência intitulada "*A questão da crioulização no Português do Brasil*". I Encontro de Variação Lingüística do Cone Sul, Projeto VARSUL/UFRGS.

Heath, S. B. (2001) What no Bedtime Story Means: Narrative Skills at Home and School. In: Duranti, A. (ed.) *Linguistic Antropology: A Reader*. Blackwell Publishing.

Hyams, N. (1986) *Language Acquisition and the Theory of Parameters.* Dordrecht: Reidel.

_____ (1992) Morphosyntatic development in Italian and its relevance to parameter-setting models: comments on the paper by Pizzuto e Caselli. *Journal of Child Language*, Cambridge, vol.19, p. 695-709.

Kato, M. (1995) Raízes não finitas na criança e a construção do sujeito. Campinas: *Cadernos de Estudos Lingüísticos*, vol.29, p.119-136.

Lemle, M. e Naro, A. (1977). *Competências Básicas do Português*. Rio de Janeiro, MOBRAL.

Loregian, L. (1996). *Concordância verbal com o pronome tu na fala do sul do Brasil.* Dissertação de Mestrado. Florianópolis, UFSC.

_____ (2001) Sobreposição de variáveis lingüísticas: saliência fônica versus tempo verbal. *Anais do 4º Encontro do CelSul*, Curitiba.

Menuzzi, S. (2001) *A ordem verbo-sujeito no português do Brasil: para uma comparação das abordagens formalistas e funcionalistas*. (mimeo)

Mioto, C. (1995) Lingüística e Ensino da Gramática. *Anais do Seminário Lingüística e Ensino de Língua Portuguesa*. Porto Alegre: EDIPUCRS.

Monguilhott, I. O. e Coelho, I. L. (2002) Um estudo da concordância verbal de terceira pessoa em Florianópolis: *In*: Vandresen, P. (org.) *Variação e mudança no português falado na Região Sul.* Pelotas: EDUCAT. p.189-216.

Naro, A. J. (1981) The social and the structural dimensions of a syntatic change. *Language*, vol.57, n.1, p.63-98.

Nicolau, E. M. D. (1998) Sobre a aquisição do sistema de flexão do português do Brasil: a representação morfo-fonológica dos traços de pessoa e número nas gramáticas iniciais de crianças brasileiras. *Letras de Hoje*. Porto Alegre. vol.33, n.2, p.191-201.

Ochs, E. e Schieffelin, B. (1997) O Impacto da Socialização da Linguagem no Desenvolvimento Gramatical. In: Fletcher, P. e MacWhinney, B. *Compêndio da Linguagem da Criança*. Porto Alegre: Artes Médicas. p.69-84.

Perroni-Simões, M. C. (1976) *Aspectos da gramática portuguesa aos 2 anos de idade*. Dissertação de Mestrado, Campinas: UNICAMP.

Perroni-Simões, M. C. e Stoel-Gammon, C. (1977) The acquisition of inflections in Portuguese: a study of the development of person markers on verbs. *Journal of Child Language*, Cambridge, vol.6: 53-67.

Pintzuk, S. (1988) *VARBRUL programs*. Mimeo.

Pizzuto, E. e Caselli, M. C. (1992) The acquisition of Italian morphology: implications for models of language development. *Journal of Child Language*, Cambridge, vol.19, n.3, p. 401-557.

_____ (1993) The acquisition of Italian morphology: a reply to Hyams. *Journal of Child Language*, Cambridge, vol.20, n.3, p.707-712.

Pontes, E. (1987) *O tópico no português do Brasil*. Campinas: Pontes.

Radford, A. (1990) Syntatic theory and the acquisition of English syntax. Oxford: Basil Blackwell.

_____. (1997) Estrutura Frasal a Categorias Funcionais. In: Fletcher, P. e MacWhinney, B. *Compêndio da Linguagem da Criança*. Porto Alegre: Artes Médicas. p.395-415.

Rodrigues, A. (1992). Língua e contexto sociolingüístico: concordância verbal no português popular em São Paulo. Publicação do curso de Pós-Graduação em Lingüística e Língua Portuguesa. Araraquara, UNESP – Campus de Araraquara, n.2, p.153-171.

_____. 1994. Concordância verbal no português falado no Brasil: uma abordagem funcionalista. *Boletim ABRALIN*, vol.15, p.115-121.

Rubino, R. J. e Pine, J. M. (1998) Subject-verb agreement in Brazilian Portuguese: what low error rates hide. *Journal of Child Language* vol.25**,** p.35-59.

Said Ali, M. (2001). Gramática histórica da língua portuguesa. 8. ed. São Paulo: Melhoramentos/Brasília, DF: Editora Universidade de Brasília.

Scherre, M. M. P. e Naro, A. J. (1993) Duas dimensões do paralelismo formal na concordância verbal no português popular do Brasil. *D.E.L.T.A.*, vol.9, n.1.

Scherre, M. M. P. (1998) Pressupostos teóricos e suporte quantitativo. In: Scherre e Silva (org.). *Padrões Sociolingüísticos: análise de fenômenos variáveis do português falado na cidade do Rio de Janeiro*. Rio de Janeiro: Tempo Brasileiro.

Seara, I. C. (2000) A variação do sujeito *nós* e *a gente* na fala florianopolitana. *Organon*, Porto Alegre, UFRGS, vol.14, n. 28 e 29.

Simões, L. J. (1997) *Sujeito nulo na aquisição do português brasileiro: um estudo de caso.* Tese de Doutorado. Porto Alegre: PUCRS.

_____ (2003) Concordância nominal de número e a aquisição de regras variáveis. *Cadernos de Pesquisa em Lingüística*. Porto Alegre, vol. 1, n. 1, p. 39-42.

Soares, S. (2001). A concordância verbal de primeira pessoa: um estudo em aquisição. Monografia de final de curso. Porto Alegre: Instituto de Letras, UFRGS.

_____ (2003) Marcação de concordância nominal e verbal nos dados de uma criança observada longitudinalmente em sessões de fala espontânea. *Anais do 5º Encontro do CELSUL*: Curitiba, UFPR.

Snow, C. (1999) Social Perspective on the Emergence of Language. In: MacWhinney, B. (ed.) *The Emergence of Language.* Mahwah, NJ: Laurence Erlbaum.

Zilles, A. M. S. (1992) A ordenação de sujeito, verbo e objeto no discurso narrativo de crianças de 4 a 6 anos. Porto Alegre: PUCRS. Tese de Doutorado.

_____ (2000) A posposição do sujeito ao verbo no português falado no Rio Grande do Sul. *Letras de Hoje*, Porto Alegre, vol. 35, n.1, p. 75-96.

_____ (2005) The development of a new pronoun: The linguistic and social embedding of a *a gente* in Brazilian Portuguese. *Language Variation and Change*. Cambridge, vol. 17, p. 19-53.

Zilles, A. M., Maya, L. e Silva, K. (2000) A concordância verbal com a primeira pessoa do plural em Panambi e Porto Alegre. *Organon*, vol. 14, n. 28 e 29, p.195-219.

Zilles, A. M. e Faraco, C. A. (2002) Considerações sobre o discurso reportado em *corpus* de língua oral. In: Vandresen, P. (org.) *Variação e mudança no português falado na Região Sul*. Pelotas: EDUCAT. p.15-46.

Zilles, A. M. e Batista, H. H. R. B. (2005). A concordância verbal de primeira pessoa do plural na *fala culta* de Porto Alegre. In: *Variação, mudança e contato lingüístico no Português da Região Sul*. Pelotas: EdUCPel.