

UNIVERSIDADE FEDERAL DE JUIZ DE FORA
FACULDADE DE LETRAS
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA

Cristina Azalim

**Concordância nominal variável na produção infantil: dados naturalísticos,
experimentais e caracterização formal**

JUIZ DE FORA

2021

Cristina Azalim

**Concordância nominal variável na produção infantil: dados naturalísticos,
experimentais e caracterização formal**

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Linguística da Faculdade de Letras da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Doutora em Linguística. Área de concentração: Linguística e Cognição.

Orientadora: Profa. Dra. Mercedes Marcilese

Coorientadora: Profa. Dra. Paula Roberta Gabbai Armelin

Juiz de Fora

2021

Ficha catalográfica elaborada através do programa de geração automática da Biblioteca Universitária da UFJF, com os dados fornecidos pelo(a) autor(a)

Azalim, Cristina.

Concordância nominal variável na produção infantil: : dados naturalísticos, experimentais e caracterização formal / Cristina Azalim. -- 2021.

247 p. : il.

Orientadora: Mercedes Marcilese

Coorientadora: Paula Roberta Gabbai Armelin

Tese (doutorado) - Universidade Federal de Juiz de Fora, Faculdade de Letras. Programa de Pós-Graduação em Linguística, 2021.

1. Concordância nominal variável. 2. Produção infantil. 3. Saliência fônica. 4. Programa Minimalista. 5. Tipos de morfemas. I. Marcilese, Mercedes, orient. II. Armelin, Paula Roberta Gabbai, coorient. III. Título.

Cristina Azalim

Concordância nominal variável na produção infantil: dados naturalísticos, experimentais e caracterização formal

Tese apresentada ao Programa de Pós-Graduação em Linguística da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do título de Doutora em Linguística.

Área de concentração:
Linguística

Aprovada em 13/12/2021.

BANCA EXAMINADORA

Profa. Dra. Mercedes Marcilese - Orientadora

Universidade Federal de Juiz de Fora

Profa. Dra. Paula Roberta Gabbai Armelin - Coorientadora

Universidade Federal de Juiz de Fora

Profa. Dra. Maria Cristina Lobo Name

Universidade Federal de Juiz de Fora

Profa. Dra. Marina Rosa Ana Augusto

Universidade do Estado do Rio de Janeiro

Profa. Dra. Ruth Elisabeth Vasconcellos Lopes

Universidade Estadual de Campinas

Profa. Dra. Livia Oushiro

Universidade Estadual de Campinas

Juiz de Fora, 06/12/2021.



Documento assinado eletronicamente por **Mercedes Marcilese, Professor(a)**, em 13/12/2021, às 15:48, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Paula Roberta Gabbai Armelin, Professor(a)**, em 13/12/2021, às 15:51, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Maria Cristina Lobo Name, Professor(a)**, em 13/12/2021, às 15:53, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Livia Oushiro, Usuário Externo**, em 14/12/2021, às 10:58, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Marina Rosa Ana Augusto, Usuário Externo**, em 14/12/2021, às 11:11, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Ruth Elisabeth Vasconcellos Lopes, Usuário Externo**, em 14/12/2021, às 13:13, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no Portal do SEI-Ufjf (www2.ufjf.br/SEI) através do ícone Conferência de Documentos, informando o código verificador **0604447** e o código CRC **01CAEE9B**.

Aos meus pais, Maria das Graças e Getulio,

e à minha filha, Isabelle.

AGRADECIMENTOS

A Deus, por me conceder a vida e estar ao meu lado em todos os momentos. Por me conduzir, abençoar e proteger sempre. Por me amparar e me dar forças nos momentos em que fraquejava. A Ti, Senhor, devo e agradeço por mais essa conquista!

Aos meus pais, Getulio e Maria das Graças, minha irmã, Simone, minha filha, Isabelle e meu sobrinho, Felipe, por me incentivarem, apoiarem e, principalmente, compreenderem, todas as vezes que eu estava “ausente” em virtude dos estudos. Sem vocês, essa conquista não teria sido possível. Por isso, a vocês, todo o meu amor, carinho, admiração e gratidão. Amo vocês!

À minha orientadora, Mercedes Marcilese, por todo seu carinho, dedicação e paciência. Por acreditar que eu era capaz de trilhar caminhos que nem eu mesma achava que conseguiria. Se hoje tenho a felicidade de ver esse trabalho concluído, devo muito a você! Por isso, procurei sempre dar o meu melhor pra fazer jus à orientação que você tem me dado desde o mestrado. Sempre levarei comigo as lembranças de tudo que você fez por mim ao longo desses 7 anos. Como já te disse, você sempre foi muito mais que orientadora: você tem sido uma verdadeira mãe acadêmica. Por isso, a você, todo o meu carinho, respeito, admiração e, é claro, minha gratidão!!!

À minha orientadora, Paula Armelin, por ter aceitado se juntar a nós nessa caminhada. Te agradeço, também pela paciência, disponibilidade, didática e carinho que você tem tido pra esclarecer as minhas dúvidas sobre nosso querido, porém difícil Minimalismo (rs). Bendita a hora que Mercedes me sugeriu cursar a disciplina de morfo com você. Afinal, foi ali que comecei a ver sua dedicação, seu amor à profissão e respeito pelos seus alunos. E hoje, estou aqui, feliz demais da conta e muito grata por esse trabalho tão bonito que você me ajudou a concretizar.

Então, a vocês duas, o meu carinhoso Muito obrigada!!!

Às Professoras Cristina Name e Marina Augusto, por acompanharem minha trajetória acadêmica com tanto carinho e pelas sugestões valiosas para o desenvolvimento de meus trabalhos.

Às Professoras Livia Oushiro e Ruth Lopes por aceitarem a compor a banca examinadora e pelas valiosas contribuições para a versão final do trabalho.

À Olívia e Nilton, os “filhos” que o doutorado me deu. Muito obrigada por todos os momentos que compartilhamos, nossas conversas, risadas e corridas à beira mar de Maceió. (Ehhh, saudade...)

À Dani Molina, pela companhia nos eventos, pelas conversas, por sua amizade e pela paz que você transmite pra gente.

Aos professores do PPG Linguística pela sua dedicação, em especial aos Professores Fábio Fortes e Ana Paula El Jaick.

Ao meu querido “Tio Wayne”, por toda a ajuda que você deu para que eu conseguisse aplicar meu experimento, além, é claro, pelas conversas, trocas de experiências e risadas. Você é sensacional!!! Adoro vc, querido!!!

Ao Igor Costa, faço minhas, as palavras de Marina Reis, uma vez que na presente pesquisa, sua ajuda também foi estatisticamente significativa! Obrigada, meu amigo!

A todos os participantes dos experimentos conduzidos no âmbito desta tese.

À FAPEMIG, pelo apoio financeiro que contribuiu para a realização deste sonho.

RESUMO

Esta tese investiga a realização variável da concordância nominal (redundante X não redundante) no português brasileiro (PB) e visa a propor uma formalização para o fenômeno a partir de dados de produção de adultos e de crianças. Em termos empíricos, com base na articulação de dados longitudinais espontâneos e metodologia experimental, nosso estudo explora: (i) possíveis efeitos da variável saliência fônica (SF) na produção adulta e infantil dos padrões de concordância nominal; (ii) a atuação de aspectos de natureza extralinguística, tais como grau de escolarização, faixa etária e origem do falante na alternância entre tais padrões e (iii) a relação entre a alternância dos padrões de marcação e um conjunto de aspectos estruturais e morfofonológicos tais como a configuração da estrutura, a natureza do núcleo do sintagma, a morfofonologia dos numerais presentes no sintagma, etc. Em termos teóricos, buscamos fornecer uma caracterização formal para a alternância de padrões de marcação de plural identificados nos dados coletados, aliando o arcabouço teórico fornecido pelo Programa Minimalista (CHOMSKY, 1993, 1995, 1999) e a Teoria dos 4M (MYERS-SCOTTON; JAKE, 2000a). Em relação aos dados naturalísticos, são relatados os resultados obtidos a partir da análise de três *corpora*, compreendendo a produção espontânea de crianças em contextos de interação com seus cuidadores primários ou com a professora. Em relação aos dados experimentais, são reportados os resultados de um estudo conduzido com crianças e adultos falantes do PB. Quanto à saliência fônica dos nomes, os resultados naturalísticos e experimentais não se mostram compatíveis com uma atuação significativa dessa variável. No que tange à produção de sintagmas plurais redundantes e não redundantes por crianças e adultos, observou-se que nos dados de produção espontânea, o comportamento das crianças parece refletir o padrão registrado na fala adulta. Os resultados do estudo experimental, por sua vez, indicam diferenças importantes entre a produção adulta e infantil. Questões associadas à natureza da situação comunicativa em cada caso podem explicar as diferenças observadas. Em termos teóricos, os resultados obtidos tanto por meio de metodologia naturalística quanto experimental e a interpretação deles decorrentes são compatíveis com uma abordagem formulada no contexto da Teoria dos 4M para explicar o fenômeno da concordância variável de número no PB. Propomos que a concordância não redundante é licenciada pelo tipo de morfema envolvido, sua constituição de traços, bem como pelas relações que eles estabelecem

no sistema de *Agree* (CHOMSKY, 1999, 2004). Mais especificamente, propomos que o morfema de número no determinante seja do tipo sistêmico precoce e o morfema de número no nome e adjetivo do tipo sistêmico tardio. Nos moldes do sistema de *Agree*, propomos que os primeiros sejam valorados e interpretáveis e que os últimos correspondam a traços não valorados e não interpretáveis. A partir desse raciocínio, defendemos – com base no sistema de Frampton e Guttman (2000) e na proposta de Pesetsky e Torrego (2007) – que o mecanismo de concordância envolve, na verdade, compartilhamento de traços.

Palavras-chave: Concordância nominal variável. Produção infantil. Saliência fônica. Programa Minimalista. Tipos de morfemas.

ABSTRACT

This thesis investigates the variable realization of nominal agreement (redundant X non-redundant) in Brazilian Portuguese (BP) and aims to propose a formalization for the phenomenon based on data produced by adults and children. In empirical terms, based on the articulation between spontaneous longitudinal data and experimental methodology, our study explores: (i) the possible effects of phonic salience on adult and child production of the nominal agreement patterns; (ii) the role of extralinguistic aspects, such as level of education, age and origin of the speaker in the alternation between such patterns and (iii) the relationship between the alternation of the patterns and a set of structural and morphological aspects such as, the configuration of the structure, the nature of the phrase head, the morphology of numerals, etc. In theoretical terms, we aim at providing a formal account for the alternation of plural marking patterns identified in the data, combining the theoretical framework provided by the Minimalist Program (CHOMSKY, 1993, 1995, 1999) and the 4M Model (MYERS-SCOTTON; JAKE, 2000a). Regarding naturalistic data, the results obtained from three *corpora* are reported, comprising the spontaneous production of children in contexts of interaction with their primary caregivers or with their teacher. Regarding experimental data, the results of a study conducted with BP-speaking children and adults are reported. Concerning the phonic salience of the nouns, the naturalistic and experimental results are not compatible with a significant effect of this variable. Regarding the production of redundant and non-redundant plural phrases by children and adults, it was observed that in the spontaneous production data, the behavior of children seems to reflect the pattern registered in adult speech. The results of the experimental study, in turn, indicate crucial differences between adult and child production. Issues associated with the nature of the communicative situation in each case may explain the differences. In theoretical terms, the results obtained both through naturalistic and experimental methodology and the resulting interpretation are compatible with the an approach formulated in the context of the 4-M Theory to explain the phenomenon of variable number agreement in BP. We claim that non-redundant agreement is licensed by the type of morpheme involved and its features, as well as the relationships they establish in the *Agree* system (CHOMSKY, 1999, 2001). More specifically, we propose that the number morpheme in the determiner is an early system morpheme, and the number morpheme in the noun and adjective is a late

system morpheme. In line with the *Agree* system, we propose that the former are valued and interpretable and that the later correspond to unvalued and uninterpretable features. Based on this reasoning, we propose – according to Frampton and Guttman’s (2000) system, and the proposal of Pesetsky and Torrego (2007) – that the agreement mechanism actually involves features sharing.

Keywords: Variable nominal agreement. Child production. Phonic salience. Minimalist Program. Types of morphemes.

LISTA DE ILUSTRAÇÕES

LISTA DE FIGURAS

- Figura 1** – Exemplo de cartão alvo utilizado no experimento – Sintagma para identificação do cartão (nome alvo gambá): Quatro gambás pretos grandes108
- Figura 2** – Exemplo de prancha utilizada no experimento109
- Figura 3** – Exemplo de prancha utilizada no Experimento 2.....136
- Figura 4** – Exemplo de cartão alvo utilizado no Experimento 2 – Sintagma-alvo para identificação do cartão: Duas flores rosas pequenas136
- Figura 5** – Traços e categorias funcionais vinculados à concordância no âmbito do DP, segundo Carstens (2000).171
- Figura 6** – Estrutura interna do DP proposta por Magalhães (2004).....172
- Figura 7** – Estrutura básica para a relação sonda-alvo em DPs contendo Adjs assumida por Magalhães (2004, p[13]), com base em Abney (1987) e173
- Figura 8** – Configuração de DPs no PB segundo Menuzzi (1994, p.131).....175
- Figura 9** – Configuração de DPs no PB segundo Menuzzi (1994, p.132).....176
- Figura 10** – Explicando a operação de Agree no sistema de Magalhães (passo 1)177
- Figura 11** – Explicando a operação de Agree no sistema de Magalhães (passo 2)177
- Figura 12** – Explicando a operação de Agree no sistema de Magalhães (passo 3)178
- Figura 13** – Explicando a operação de Agree no sistema de Magalhães (passo 4)178
- Figura 14** – Explicando a operação de Agree no sistema de Magalhães (passo 5)179
- Figura 15** – Explicando a operação de Agree no sistema de Magalhães (passo 6)179
- Figura 16** – Explicando a operação de Agree no sistema de Magalhães (passo 7)180
- Figura 17** – Classificação de morfemas187
- Figura 18** – Diagrama do processo de produção linguística: ativação do lema e estrutura lexical abstrata, extraída de Namba (2004).....193
- Figura 19** – Estrutura de DP proposta na presente pesquisa.....207
- Figura 20** – Explicando a operação de Agree no nosso sistema.....211
- Figura 21** – Estrutura de DP com Numeral proposta na presente pesquisa.....215
- Figura 22** – Explicando a operação de Agree no nosso sistema nas estruturas com Numeral.....216

LISTA DE GRÁFICOS

- Gráfico 1** – Ocorrências de nomes em sintagmas plurais em função da SF nos dados naturalísticos (adultos e crianças)81
- Gráfico 2** – Gráfico de efeitos das condições experimentais na distribuição dos sintagmas não redundantes no grupo controle (adultos) –118
- Gráfico 3** – Gráfico de efeitos da produção de nomes não marcados em sintagmas não redundantes em função do grupo analisado (Crianças X Adultos) - Experimento 1 – nomes alvo120
- Gráfico 4** – Percentual de nomes não marcados em sintagmas não redundantes em função de padrão de acentuação – no grupo experimental (crianças) e controle (adultos) – Experimento 1 – nomes alvo.....122
- Gráfico 5** – Percentual de nomes não marcados em sintagmas não redundantes em função de número de sílabas - no grupo experimental (crianças) e controle (adultos) – Experimento 1 - nomes alvo123
- Gráfico 6** – Gráfico de efeitos das condições experimentais na distribuição de nomes não marcados no grupo experimental (Crianças) – Experimento 1 - nomes alvo.....125
- Gráfico 7** – Gráfico de efeitos da produção de sintagmas plurais não redundantes em função do grupo analisado (Crianças X Adultos) –139
- Gráfico 8** – Distribuição de padrões de marcação internos registrados nas estruturas Num+N+Adj nas duas atividades experimentais– dados brutos.....150
- Gráfico 9** – Resultado do Teste de Intervalo de Confiança conduzido em função dos padrões de marcação internos na estrutura (Crianças e Adultos) – dados brutos – Experimentos 1 e 2.....151
- Gráfico 10**– Gráfico de efeitos em função da morfofonologia dos numerais na distribuição de sintagmas plurais não redundantes – grupo experimental - (crianças) – Experimentos 1 e 2.....155
- Gráfico 11** – Panorama geral da produção de sintagmas plurais redundantes nos dados naturalísticos e experimentais156
- Gráfico 12** – Percentual de sintagmas plurais produzidos por crianças e adultos na primeira tarefa experimental – Experimento 1 - dados brutos243

Gráfico 13	– Gráfico de efeitos da variável padrão de acentuação na produção de sintagmas plurais não redundantes pelos Adultos (Grupo controle) – Experimento 1 – dados brutos.....	244
Gráfico 14	– Gráfico de efeitos da produção de sintagmas plurais não redundantes em função da variável padrão de acentuação (Grupo de crianças) – Experimento 1 – nomes alvo.....	245
Gráfico 15	– Gráfico de efeitos da produção de sintagmas plurais não redundantes em função das condições experimentais (Grupo de crianças) – Experimento 1 – nomes alvo.....	246
Gráfico 16	– Gráfico de efeitos da produção de nomes não marcados em sintagmas plurais não redundantes em função das condições experimentais (Grupo de Adultos) – Experimento 1 – nomes alvo	247

LISTA DE QUADROS

Quadro 1	– Síntese dos estudos sobre o papel da SF na realização da marcação de número no âmbito do DP	58
Quadro 2	– Síntese da composição dos <i>corpora</i> analisados.....	71
Quadro 3	– Condições experimentais do Experimento 1.	103
Quadro 4	– Síntese dos resultados do Experimento 1 obtidos considerando as variáveis independentes separadamente (padrão de acentuação e número de sílabas)	126
Quadro 5	– Síntese dos resultados do Experimento 1 obtidos considerando as condições experimentais (cruzamento das variáveis padrão de acentuação e número de sílabas).....	127
Quadro 6	– Condições experimentais do Experimento 2	131
Quadro 7	– Síntese com os principais achados experimentais	159
Quadro 8	– Quadro contrastivo em função dos mecanismos propostos para concordância e do locus da interpretabilidade do traço de número apresentados nos trabalhos resenhados.....	183
Quadro 9	– Classificação de morfemas segundo modelo dos 4M (MYERS-SCOTTON; JAKE, 2000a) e suas principais propriedades.....	190
Quadro 10	– Algumas hipóteses para explicar a ocorrência de concordância não redundante no PB.....	198
Quadro 11	– Padrões de marcação internos ao sintagma encontrados nas estruturas formadas por três elementos nos dados experimentais.....	218
Quadro 12	– Quadro síntese com as interpretações teóricas para nossos resultados	221

LISTA DE TABELAS

- Tabela 1** – Distribuição dos sintagmas plurais em função da forma de concordância aplicada (redundante X não redundante) em cada *corpus*.....75
- Tabela 2** – Distribuição dos sintagmas plurais em função do padrão de concordância aplicada (redundante X não redundante) e da SF do nome em cada *corpus*80
- Tabela 3** – Distribuição dos sintagmas plurais em função do tipo de estrutura e concordância em cada *corpus*83
- Tabela 4** – Distribuição dos sintagmas plurais em função dos processos de formação morfológica dos numerais (Num + N X Num-S + N) por crianças e adultos – dados naturalísticos.....85
- Tabela 5** – Distribuição dos sintagmas plurais em função dos padrões de marcação internos ao sintagma e dos tipos de estrutura analisados em cada *corpus*87
- Tabela 6** – Distribuição dos sintagmas plurais em função do padrão de concordância aplicado (redundante X não redundante) no grupo experimental (crianças) e controle (adultos) – Experimento 1 – nomes alvo112
- Tabela 7** – Tabela de coeficientes extraída do modelo de regressão logística para ocorrências de sintagmas plurais não redundantes em função do grupo analisado (Crianças X Adultos) – Experimento 1 – nomes alvo.....113
- Tabela 8** – Distribuição dos sintagmas plurais em função do padrão de acentuação (oxítonos X paroxítonos) no grupo experimental (crianças) e controle (adultos) – Experimento 1 - nomes alvo114
- Tabela 9** – Tabela de coeficientes extraída do modelo de regressão logística para ocorrências de sintagmas plurais não redundantes em função da variável padrão de acentuação (Adultos) – Experimento 1 - nomes alvo115
- Tabela 10** – Distribuição dos sintagmas plurais em função do número de sílabas (dissílabos X trissílabos) no grupo experimental (crianças) e controle (adultos) – Experimento 1 – nomes alvo.....115
- Tabela 11**– Distribuição dos sintagmas plurais em função das condições experimentais no grupo experimental (crianças) e controle (adultos) – Experimento 1 - nomes alvo.....117

Tabela 12	– Distribuição de sintagmas plurais não redundantes em função da marcação de plural no nome (nome marcado X nome não marcado) no grupo experimental (crianças) e controle (adultos) – Experimento 1 – nomes alvo	120
Tabela 13	– Distribuição de marcação de plural no nome em sintagmas não redundantes em função do padrão de acentuação (oxítonos X paroxítonos) no grupo experimental (crianças) e controle (adultos) – Experimento 1	121
Tabela 14	– Distribuição de marcação de plural no nome em sintagmas não redundantes em função do número de sílabas (dissílabos x trissílabos) no grupo experimental (crianças) e controle (adultos) – Experimento 1 - nomes alvo	123
Tabela 15	– Distribuição de marcação de plural no nome em sintagmas não redundantes em função das condições experimentais no grupo experimental (crianças) e controle (adultos) – Experimento 1 - nomes alvo	124
Tabela 16	– Tabela de coeficientes extraída do modelo de regressão logística para ocorrências de não marcação no nome em função das condições experimentais (Crianças) – Experimento 1 – nomes alvo	125
Tabela A	– Distribuição dos sintagmas plurais em função do padrão de concordância aplicado (redundante X não redundante) no grupo experimental (Crianças) e controle (Adultos_ - Experimento 2 - dados brutos	137
Tabela 17	– Distribuição dos sintagmas plurais em função do padrão de concordância aplicado (redundante X não redundante) no grupo experimental (crianças) e controle (adultos) – Experimento 2 – nomes alvo	138
Tabela 18	– Distribuição dos sintagmas plurais em função dos tipos de processos de formação de plural (-R X -L X plural regular) no grupo experimental (crianças) e controle (adultos) Experimento 2 – nomes alvo	139
Tabela 19	– Tabela de coeficientes extraída do modelo de regressão logística para ocorrências de sintagmas plurais não redundantes em função dos tipos de processos de formação de plural (Crianças e Adultos) –	140
Tabela 20	– Distribuição dos sintagmas plurais não redundantes em função da marcação ou não marcação de plural no nome (nome marcado X nome	

	não marcado) no grupo experimental (Crianças) e controle (Adultos)	141
Tabela 21	– Distribuição de marcação de plural no nome em sintagmas não redundantes em função dos tipos de processos de formação de plural (-R X -L X plural regular) no grupo experimental (Crianças) e controle (Adultos).....	142
Tabela 22	– Distribuição dos sintagmas plurais em função do tipo de estrutura analisada (Num + N X Num + N + A X Num + N + A1 + A2) no grupo experimental (crianças) e controle (adultos) – dados brutos .	146
Tabela 23	– Tabela de coeficientes extraída do modelo de regressão logística para ocorrências de sintagmas plurais não redundantes em função dos tipos de estrutura - Experimento 2 – nomes alvo	147
Tabela 24	– Distribuição dos sintagmas plurais em função dos padrões de marcação internos ao sintagma nas duas tarefas experimentais – dados brutos.	149
Tabela 25	– Distribuição dos sintagmas plurais em função dos processos de formação morfológica dos numerais (Num + N X Num-S + N) no grupo experimental (crianças) e controle (adultos) – dados brutos .	153
Tabela 26	– Ocorrências de marcação explícita de plural considerando a variável posição dos elementos dentro do DP – dados naturalísticos.....	201
Tabela 27	– Ocorrências de marcação explícita de plural considerando a variável posição dos elementos dentro do sintagma – dados experimentais .	202
Tabela 28	– Percentual de marcação explícita de plural considerando o tipo de morfema associado aos afixos de plural no PB – crianças Corpus B	204
Tabela 29	– Distribuição dos sintagmas plurais em função do padrão de concordância aplicada (redundante X não redundante) e faixa etária em cada corpus– dados naturalísticos	242
Tabela 30	– Distribuição dos sintagmas plurais em função do padrão de concordância aplicada (redundante X não redundante) e SF em cada faixa etária– dados naturalísticos	242
Tabela 31	– Tabela de coeficientes extraída do modelo de regressão logística para ocorrências de sintagmas plurais não redundantes em função do grupo analisado (Crianças X Adultos) – Experimento 1 – dados brutos.....	243

Tabela 32 – Distribuição dos sintagmas plurais em função do padrão de acentuação (oxítonos X paroxítonos) em ambos os grupos (Crianças e adultos) – Experimento 1 – dados brutos.....	243
Tabela 33 – Distribuição dos sintagmas plurais em função do número de sílabas (dissílabos X trissílabos) no grupo experimental (crianças) e controle (adultos) – Experimento 1 – dados brutos.....	244
Tabela 34 – Tabela de coeficientes extraída do modelo de regressão logística para ocorrências de sintagmas plurais não redundantes em função das variáveis padrão de acentuação e número de sílabas (Modelo sem interação) – Crianças e Adultos - Experimento 1	245
Tabela 35 – Tabela de coeficientes extraída do modelo de regressão logística para ocorrências de sintagmas plurais não redundantes em função das variáveis padrão de acentuação e número de sílabas (Modelo com interação) – Crianças e Adultos – Experimento.....	246

SUMÁRIO

1	INTRODUÇÃO.....	22
1.1	HIPÓTESES INICIAIS E OBJETIVOS DA PESQUISA.....	26
1.2	JUSTIFICATIVAS DA PROPOSTA.....	30
1.3	ORGANIZAÇÃO DO TRABALHO.....	31
2	CONCORDÂNCIA NOMINAL VARIÁVEL NO PB E O PAPEL DA SALIÊNCIA FÔNICA	33
2.1	CONCORDÂNCIA NOMINAL VARIÁVEL: CONTRAPONTO ENTRE A GRAMÁTICA TRADICIONAL E A SOCIOLINGUÍSTICA VARIACIONISTA	33
2.2	CONCORDÂNCIA VARIÁVEL NO PROCESSAMENTO ADULTO: A VISÃO DA PSICOLINGUÍSTICA EXPERIMENTAL	45
2.3	CONCORDÂNCIA NOMINAL VARIÁVEL NA AQUISIÇÃO DO PB....	49
3	CONCORDÂNCIA NOMINAL VARIÁVEL NA PRODUÇÃO INFANTIL: DADOS NATURALÍSTICOS	63
3.1	ASPECTOS METODOLÓGICOS NO ESTUDO DA PRODUÇÃO LINGUÍSTICA INFANTIL E ADULTA.....	65
3.2	A PRODUÇÃO DA CONCORDÂNCIA NOMINAL VARIÁVEL POR CRIANÇAS ADQUIRINDO O PB.....	69
3.2.1	Constituição das amostras e metodologia de coleta	70
3.2.2	Caracterização dos dados.....	72
3.3.3	Análise da produção	75
4	CONCORDÂNCIA NOMINAL NA PRODUÇÃO INFANTIL: ABORDAGEM EXPERIMENTAL.....	92
4.1.1	Produção eliciada	96
4.2	EXPERIMENTO 1: INVESTIGANDO A SALIÊNCIA FÔNICA NA PRODUÇÃO DA CONCORDÂNCIA NOMINAL VARIÁVEL – PADRÃO DE ACENTUAÇÃO E NÚMERO DE SÍLABAS.....	100

4.2.1	Método	102
	Participantes	105
	Procedimento	107
4.2.2	Resultados e discussão	109
4.2.2.1	<i>Análise 1 – SF dos nomes e marcação de plural – Dados brutos</i>	111
4.2.2.2	<i>Análise 2 – SF dos nomes e marcação de plural – Análise não atomística – nomes alvo</i>	112
4.2.2.3	<i>Análise 3 – SF dos nomes e marcação de plural – Análise atomística – nomes alvo</i>	119
4.2.2.4	<i>Síntese dos principais resultados do Experimento 1</i>	126
4.3	EXPERIMENTO 2: INVESTIGANDO A SF NA PRODUÇÃO DA CONCORDÂNCIA NOMINAL VARIÁVEL – PROCESSOS MORFOFONOLÓGICOS DE FORMAÇÃO DE PLURAL	129
4.3.1	Método	131
	Materiais	133
	Procedimento	135
4.3.2	Resultados e discussão	137
4.3.2.1	<i>Análise 1 – SF dos nomes e marcação de plural – Análise não atomística – nomes alvo</i>	137
4.3.2.2	<i>Análise 2 – SF dos nomes e marcação de plural – Análise atomística – nomes alvo</i>	141
4.4	ANÁLISES COMPLEMENTARES DOS DADOS EXPERIMENTAIS: INVESTIGANDO A CONFIGURAÇÃO ESTRUTURAL DO SINTAGMA E A FONOLOGIA DOS NUMERAIS.....	145
4.4.1	Configuração dos sintagmas e marcação interna de plural.....	145
4.4.2	Fonologia dos numerais e marcação de plural no sintagma	152
4.5	DIFERENÇAS ENTRE A PRODUÇÃO ADULTA E INFANTIL: DADOS NATURALÍSTICOS X EXPERIMENTAIS	155
4.6	SÍNTESE DOS PRINCIPAIS ACHADOS EXPERIMENTAIS.....	158

5	INTERPRETAÇÕES TEÓRICAS DOS FATOS EMPÍRICOS: A CONCORDÂNCIA NOMINAL NO PB SOB UMA ABORDAGEM FORMAL.....	161
5.1	PROGRAMA MINIMALISTA: NOÇÕES FUNDAMENTAIS	162
5.2	CONCORDÂNCIA NA TEORIA GERATIVA	166
5.3	PROPOSTAS PARA O ESTUDO DA CONCORDÂNCIA NO DP.....	170
5.3.1	Uma alternativa de análise para a concordância nominal variável: a Teoria dos 4M	184
5.3.1.1	<i>A Teoria dos 4M.....</i>	184
5.3.1.1.1	<i>A Teoria dos 4M e a concordância de número variável no PB</i>	195
5.3.1.2	<i>Nova proposta de análise para a concordância nominal variável.....</i>	199
5.3.2	Interfaces do modelo de classificação de morfemas com Minimalismo e desdobramentos para a presente pesquisa	204
5.3.2.1	<i>Interpretando a variação na marcação morfofonológica observada nos nossos dados</i>	218
6	CONSIDERAÇÕES FINAIS	223
	REFERÊNCIAS BIBLIOGRÁFICAS.....	231
	APÊNDICES	242

1 INTRODUÇÃO

Esta tese investiga a alternância de padrões de marcação de plural na realização da concordância nominal no português brasileiro (PB) e visa a propor uma formalização para esse fenômeno a partir de dados de produção de adultos e de crianças.

Dois padrões gerais de concordância nominal podem ser identificados na língua, quais sejam, marcação morfológicamente redundante *versus* marcação não redundante (Ex. *As pecinhas pequenas X Os pés vermelhoØ/ Seis bonéØ verdeØ*)¹. Para abordar o fenômeno escolhido, buscamos estabelecer um diálogo entre a sociolinguística variacionista, a psicolinguística experimental – em particular, no que tange à aquisição da linguagem – e a teoria linguística de base formalista.

Em termos empíricos, com base na articulação de dados longitudinais espontâneos e metodologia experimental, nosso estudo explora: (i) possíveis efeitos da variável saliência fônica (SF) na produção adulta e infantil dos padrões de concordância nominal; (ii) a atuação de aspectos de natureza extralinguística, tais como grau de escolarização, faixa etária e origem do falante (rural X urbano) na alternância entre tais padrões; e (iii) a relação entre a alternância dos padrões de marcação e um conjunto de aspectos estruturais e morfofonológicos tais como, configuração da estrutura, natureza do núcleo do sintagma, morfofonologia dos numerais, etc. Em termos teóricos, buscamos fornecer uma caracterização formal para a alternância de padrões de marcação de plural identificados nos dados coletados, aliando o arcabouço teórico fornecido pelo Programa Minimalista (doravante, PM) (CHOMSKY, 1993, 1995, 1999) e a Teoria dos 4M (MYERS-SCOTTON; JAKE, 2000a).

Inicialmente, a escolha pelo estudo de tais aspectos justifica-se em virtude de nosso interesse por aprofundar e refinar a pesquisa desenvolvida em Azalim (2016), acerca do papel que o fator SF desempenharia na realização dos padrões de concordância nominal. Isso porque os resultados experimentais obtidos em função de tal variável em Azalim (2016) e Azalim et al (2018), bem como os obtidos a partir de uma abordagem descritiva (CAMPOS, 2015, dentre outros), parecem apontar para a

¹ Todos os exemplos do PB fornecidos ao longo da tese correspondem a ocorrências reais coletadas, seja no âmbito do estudo naturalístico conduzido ou a partir dos nossos dados experimentais.

necessidade de uma revisão da própria caracterização da SF quando a dimensão cognitiva e perceptual da variação é levada em consideração. Nesse sentido, a presente tese amplia os estudos previamente conduzidos ao incluir a produção infantil como uma rica fonte de informações para a caracterização do fenômeno de variação investigado.

A variação linguística tem recebido papel de destaque nos estudos linguísticos, principalmente a partir de pesquisas pioneiras lideradas por Labov(2008 [1972]), dentre outros. No entanto, vale destacar, como apontado por Fortes (2012, p. 198), que “reflexões em torno das línguas enquanto construções variáveis, suscetíveis de serem normatizadas segundo determinados usos, já estavam presentes nos primórdios da tradição gramatical ocidental.” A variação é um fenômeno inerente aos sistemas linguísticos e consiste na ocorrência de formas que se alternam em um mesmo contexto linguístico, expressando *grosso modo* o mesmo significado, pelo menos no que tange ao valor de verdade dos enunciados. Contudo, vale lembrar que essa alternância não ocorre de forma aleatória. Ao contrário, estaria condicionada à atuação de fatores linguísticos – tais como, contexto fonológico, categoria sintática, posição do elemento na sentença – e extralinguísticos – dentre os quais, classe social, faixa etária, grau de escolarização, origem e sexo dos falantes – que podem favorecer ou desfavorecer a ocorrência de um ou outro padrão.

Entretanto, o tratamento dado ao tema da variação a depender da perspectiva adotada – no contexto da gramática normativa ou no âmbito de teorias linguísticas – pode apresentar posições divergentes. Para a gramática normativa, boa parte dos casos de variação costumam ser vistos como “erro”, “desvio” da norma culta², visão já difundida na tradição gramatical antiga em textos de autores como Quintiliano (*apud*

²Segundo Pagotto (1998, p.50), “as formas linguísticas da chamada norma culta estão impressas de sentidos que se ligam tanto à ancestralidade de uma cultura superior quanto à inserção social do sujeito no domínio simbólico das diferenças. A constituição de uma nova norma culta no Brasil foi um longo processo executado por eminentes gramáticos, jornalistas, escritores, etc. que construíram o significado que hoje atribuímos às formas da escrita”. Além disso, Pagotto argumenta que foi o “discurso científico que deu o último suporte para a manutenção da norma culta codificada na segunda metade do século XIX”. O referido autor destaca ainda que “o estabelecimento da norma culta se dá por meio de associações mais ou menos explícitas com determinados grupos dentro de uma sociedade, bem como é fruto de assentamento de certos usos nas camadas mais escolarizadas da população, e ainda se trata de um processo que se integra historicamente ao modelo de sociedade que uma certa elite implanta, calcado na manutenção da estrutura de dominação.”

FORTES, 2012, p.201), para o qual “a língua é uma realidade regulada por princípios a ela intrínsecos, e organiza-se segundo certa lógica das coisas.” Nesse sentido, o caráter estigmatizante geralmente atribuído à variação já se fazia presente na antiguidade, como pode ser observado na visão apresentada por Quintiliano (*apud* FORTES, 2012, p.201) segundo o qual, o uso da linguagem estaria associado à modalidade de fala observada na produção de pessoas escolarizadas, conhecedoras da tradição letrada, e não às formas vulgares ou populares da fala.

Em contrapartida, para a sociolinguística, campo dos estudos da linguagem que considera a língua como um veículo de comunicação, de informação e de expressão entre os indivíduos da espécie humana, a variação é considerada um fenômeno natural e inerente às línguas naturais. E é por meio da observação da variação que constatamos o caráter heterogêneo e diversificado e, até mesmo, “rico” das línguas. Nesse contexto, Mollica (2003, p.11) destaca:

[...]a Sociolinguística enquanto ciência tem como um de seus objetivos investigar o grau de estabilidade ou de mutabilidade da variação linguística, diagnosticar as variáveis que têm efeito positivo ou negativo sobre a emergência dos usos linguísticos alternativos e prever seu comportamento regular e sistemático.

A psicolinguística, por sua vez, investiga as relações existentes entre questões associadas ao conhecimento e ao uso de uma língua, tais como a do processo de aquisição de linguagem e a do processamento linguístico, bem como os processos mentais que se supõe estarem a elas associados. Nesse sentido, segundo Field (2004), a psicolinguística provê *insights* acerca (i) da maneira como formulamos nosso discurso oral e escrito, bem como compreendemos o dos outros, (ii) como estocamos e usamos vocabulário, (iii) e como adquirimos uma língua. Para tal, a psicolinguística se relaciona a muitos outros domínios, tais como a fonética, a análise do discurso, o estudo das patologias linguísticas, a neurociência e a aquisição da linguagem, dentre outros.

Recentemente, o estudo das dimensões cognitiva e perceptual da variação linguística tem recebido destaque na literatura acerca do processamento linguístico e tem se tornado um objeto de pesquisa de crescente interesse (CAMPBELL-KIBLER, 2009, 2010; LABOV et al, 2011; SQUIRES, 2013, 2014, e outros). O objetivo principal dessa abordagem consiste em uma compreensão mais aprofundada sobre a

maneira como os falantes lidam com a variação linguística em termos de processamento. Sendo assim, estudos voltados para o aspecto cognitivo da variação buscam compreender de que maneira os falantes não só adquirem como também armazenam o conhecimento sobre variação linguística e em que medida tal conhecimento seria semelhante em todos os níveis linguísticos, a saber, fonética/fonologia, morfologia, sintaxe, semântica e pragmática.

Tradicionalmente, nas pesquisas sobre variação linguística, prevalece o foco na produção de falantes adultos, com raras exceções de trabalhos sobre a variação na língua materna observada na produção infantil conduzidos no campo da sociolinguística (ROBERTS, 2002). Afinal, Roberts (2002) menciona que o interesse pela variação na linguagem infantil é um tipo de estudo relativamente novo dentro do campo da sociolinguística. Questões metodológicas têm sido apontadas como um dos principais motivos para essa lacuna na literatura. Em particular, são destacadas dificuldades relativas ao tamanho reduzido das amostras de fala espontânea infantil, fato que limita as possibilidades de tratamento estatístico dos dados. Por outro lado, nos estudos sociolinguísticos é frequentemente assumido que o vernáculo apenas alcançaria sua forma estável na idade adulta e que a produção infantil passa por períodos de instabilidade que não necessariamente refletem as características da comunidade de fala. Logo, segundo Lorandi (2013, p.134), “durante muito tempo, a criança era vista como alguém que está adquirindo o vernáculo³ e não, necessariamente, como alguém que contribui para a sua manutenção ou mudança.” Pesquisas mais recentes, entretanto, têm começado a colocar em destaque a relevância da fala infantil no contexto da sua comunidade linguística para a o estudo da variação. Estudos sobre as línguas crioulas, que passam a ser internalizadas como língua da comunidade apenas por meio da aquisição pelas gerações mais novas, têm sido paradigmáticos nesse sentido (LORANDI, 2013).

No âmbito dessa questão, Lorandi (2013) estabelece uma distinção que se mostra relevante para a nossa pesquisa. Trata-se da diferença entre a variação linguística presente na fala das crianças e a aquisição da variação que, por vezes,

³Importante destacar que por vernáculo de uma comunidade de fala entende-se o estilo em que o mínimo de atenção é dado ao monitoramento da fala, ou seja, “é a língua que se produz em contextos tais como uma conversa entre amigos, ou que se conta uma piada, ou seja, em que se caracteriza o uso linguístico espontâneo, ou pelo menos, com o menor monitoramento possível”. (FREITAG, 2010, p.85)

podem parecer um único assunto. Segundo Lorandi (2013), a variação linguística dentro de uma comunidade pode ser observada em dados de falantes de diferentes idades, o que faz com que a variável idade seja frequentemente considerada nesse tipo de estudo.

Estudos que investigam a aquisição da variação, por sua vez, buscam compreender o processo pelo qual determinada regra variável é adquirida pela criança, ou seja, como e quando tal regra passa a fazer parte do seu sistema linguístico. Os estudos de Simioni (2006), a partir de dados longitudinais e de Reis (2020), combinando dados naturalísticos e experimentais, por exemplo, correspondem a essa segunda abordagem. Tendo-se em vista tal distinção, cumpre esclarecer que, apesar de em alguns momentos – principalmente na análise dos dados naturalísticos – estabelecermos uma comparação entre o *input* que a criança recebe e o que ela efetivamente produz, não é nosso objetivo neste trabalho propor uma discussão aprofundada acerca de como de fato se dá o processo de aquisição do fenômeno de variação linguística aqui investigado por parte das crianças brasileiras. Diferentemente, como veremos a seguir, a presente tese traz uma análise de dados de produção adulta e infantil – com especial atenção para o papel da SF – a partir da qual é construída uma formalização do fenômeno da concordância nominal variável para o PB.

Assim sendo, embora o processo de aquisição *per se* não seja o foco da tese, consideramos que os dados da produção infantil podem fornecer valiosos *insights* para nossa pesquisa. A fala infantil se mostra especialmente relevante para os estudos sobre variação pela influência que a mesma tem, tanto na manutenção quanto na mudança de aspectos linguísticos no contexto da comunidade de fala na qual a criança está inserida. Tais dados podem revelar em que medida o fenômeno variável aqui investigado está se consolidando na direção de uma efetiva mudança. Assim sendo, procuramos obter o maior número de dados atuais e representativos (aliando dados adultos e infantis) que possam nos responder a diferentes questões de pesquisa.

1.1 HIPÓTESES INICIAIS E OBJETIVOS DA PESQUISA

Tendo-se em vista que o objeto empírico investigado corresponde a um fenômeno de variação linguística, boa parte das nossas hipóteses iniciais foram

calcadas nos estudos da sociolinguística variacionista, bem como em trabalhos sobre o assunto desenvolvidos no âmbito da linguística formalista. Assim sendo, as hipóteses de trabalho que nortearam nosso estudo foram as seguintes:

- (i) Os dados de produção infantil são compatíveis com o *input* fornecido pelos cuidadores primários da criança, ou seja, a gramática infantil – pelo menos, nas faixas etárias investigadas – , espelha aquilo que é encontrado na produção adulta (CAPELLARI; ZILLES, 2002; LAMPRECHT, 1997; MILLER; SCHMITT, 2010, 2012)
- (ii) Partindo do pressuposto de que a SF seria um preditor para fenômenos de variação tal como a realização dos padrões de marcação de número (OUSHIRO; GUY, 2015), mas considerando também resultados prévios sobre a atuação dessa variável (AZALIM, 2016; AZALIM et al, 2018), assumimos que a SF configura-se como uma propriedade dos itens lexicais que dependeria de um conjunto de aspectos ou dimensões – de acordo com a literatura, número de sílabas, padrão de acentuação e processos morfofonológicos de formação do plural. Tais dimensões, por sua vez, não têm a mesma relevância ou efeito quando da classificação de um item em função do realce fônico do mesmo.
- (iii) A concordância é um processo sintático guiado pela configuração de traços nos moldes minimalistas (CHOMSKY, 1999, 2001) e a natureza dos morfemas envolvidos pode dar conta dos diferentes padrões de marcação identificados na língua.

Para o desenvolvimento da pesquisa, partimos ainda da ideia de que o PB exhibe um padrão variável de concordância de número e que na concordância nominal. Especificamente, é possível encontrar marcação redundante em todos os itens do sintagma determinante (DP⁴, do inglês *Determiner Phrase*) ou, alternativamente,

⁴ No presente trabalho, adotamos a caracterização do sintagma determinante (DP) formulada por Abney (1987), segundo a qual, uma categoria funcional (D) faria parte da projeção estendida do NP (*Nominal Phrase*, sintagma nominal). Nesse sentido, o DP seria um sintagma composto por um elemento da categoria D na posição de núcleo e um NP como seu complemento. Tal caracterização é amplamente difundida no âmbito da Teoria Gerativa, arcabouço que norteia os pressupostos teóricos de nossa pesquisa no que tange ao modelo de língua assumido. No entanto, não é muito comum encontrarmos o termo DP em trabalhos de orientação sociolinguística – com os quais a nossa pesquisa dialoga – nos quais se utiliza mais frequentemente o termo NP para fazer referência a estruturas de natureza nominal. Nesse

marcação não-redundante em que o morfema de plural seria obrigatório apenas no D (determinante)⁵, podendo ou não ser realizado no nome e no adjetivo, tal como exemplificado em (1-2) abaixo:

- (1) Os⁶ piratas pequenininhos – concordância redundante
 (2) Os pés vermelho∅/Os “homi”∅grande∅ – concordância não redundante

Diante dessa caracterização inicial, nossa tese tem como objetivo geral contribuir para a discussão mais ampla sobre a variação linguística presente na produção infantil e adulta a partir da investigação de uma série de questões relacionadas ao fenômeno da concordância nominal variável de número no DP no PB.

Para tal, como fundamentação teórica e metodológica, a presente pesquisa procura – como já mencionado – estabelecer um diálogo entre a sociolinguística variacionista, a psicolinguística experimental e teoria linguística de cunho formalista. Mais especificamente, nos moldes da interface entre uma teoria de classificação de morfemas (Teoria dos 4M, MYERS-SCOTTON; JAKE, 2000) e o PM (CHOMSKY, 1993, 1995, 1999;) para a caracterização das relações de concordância em termos formais (FRAMPTON; GUTTMAN, 2000).

Acreditamos que o domínio empírico escolhido – a concordância variável de número no DP – seja um ambiente bastante rico para correlacionarmos as três áreas supracitadas, uma vez que (i) a sociolinguística tem elucidado os fatores tidos como relevantes para a alternância entre os padrões redundante e não redundante de concordância no PB, (ii) no contexto da psicolinguística, embora as pesquisas ainda sejam escassas, têm sido reportadas diferenças significativas no processamento de ambos os padrões de marcação e pouco se sabe sobre a ocorrência desse fenômeno na

sentido, ao longo do presente trabalho, o termo NP será utilizado quando forem discutidos trabalhos que utilizam preferencialmente tal denominação para fazer menção às estruturas nominais. Sendo assim, em alguns momentos, ambos os termos podem ser utilizados como sendo intercambiáveis.

⁵ Vale destacar a possível exceção de estruturas do tipo D-Poss-N (determinante-possessivo-nome), em que a marca ocorreria em Poss (COSTA; FIGUEIREDO SILVA, 2006), tal como em *A∅ minhas coleguinha∅*.

⁶Importante esclarecer que na presente tese estão sendo discutidos sintagmas nos quais há a ocorrência de um elemento foneticamente à esquerda do nome (D, Num, Poss...). Assim, exemplos como “meninas bonitas” em que poderia ocorrer um D foneticamente nulo ou se tratar de uma estrutura do tipo *bare noun*, não estão sendo considerados.

fala infantil e (iii) a teoria linguística de base formalista, por sua vez, vem tentando mapear as relações estruturais e caracterizar a natureza dos traços que estariam envolvidos nos mecanismos de concordância de número, sendo que os padrões de variação observados no PB constituem um desafio teórico instigador nessa perspectiva.

A partir dessas considerações, como objetivos específicos primários desta tese pretendemos:

- (i) Verificar em que medida a produção da criança durante a aquisição segue os mesmos padrões da fala adulta e se mostra compatível com *input* recebido a partir da fala dos seus cuidadores, quando consideradas amostras de diferentes grupos socioeconômicos;
- (ii) Investigar possíveis efeitos da variável SF na alternância dos padrões de concordância nominal na produção espontânea adulta e infantil, bem como possíveis correlações entre SF e origem do falante na alternância de padrões de concordância nominal (zona urbana X zona rural); e investigar experimentalmente a sensibilidade de crianças e adultos a dimensões específicas relativas à SF na realização da concordância variável no interior do sintagma nominal, quais sejam: o padrão de acentuação, o número de sílabas e os processos morfofonológicos de formação de plural nos nomes;
- (iii) Articular os resultados obtidos a partir das análises dos dados naturalísticos e experimentais à discussão teórica dos modelos formais propostos na literatura para propor uma formalização que dê conta do sistema de concordância nominal variável identificado no PB.

Importante lembrar que, como mencionado previamente, a SF foi tomada como ponto de partida para nossas investigações empíricas no âmbito desta tese. Sendo assim, tal fator consiste no tópico que exploramos de forma mais extensa e sistemática. Todavia, com o avanço da investigação, surgiram uma série de novas questões relacionadas à marcação da concordância, bem como acerca das diferenças que existem entre as produções adulta e infantil que se mostraram muito relevantes para proporcionar subsídios para embasarmos a discussão teórica conduzida a partir desses dados. Diante disso, foram traçados os seguintes objetivos secundários para nossa investigação:

- (i) Coletar dados que permitissem avaliar outros pontos relevantes – tais como a configuração da estrutura na qual os nomes ocorrem, os padrões de marcação internos ao sintagma, a presença de numeral na estrutura –, no intuito obtermos subsídios adicionais para embasar nossa formalização do fenômeno da concordância nominal variável no PB;
- (ii) Investigar possíveis efeitos da constituição morfofonológica dos numerais na realização dos padrões de concordância, no intuito de verificar se a terminação em -S, encontrada em numerais como *dois*, *três* e *seis*, por exemplo, poderia ser vista como resultado de uma possível falsa analogia entre tal marca e o morfema de plural -S dos nomes.

1.2 JUSTIFICATIVAS DA PROPOSTA

De modo geral, a pesquisa conduzida se mostra relevante em virtude da necessidade de se ampliar o conjunto – ainda reduzido – de estudos sobre a variação na fala infantil no PB e reforçar a relevância da articulação de dados naturalísticos e experimentais para uma melhor compreensão dos processos de aquisição, variação e mudança que ocorrem no interior de uma comunidade de fala. Afinal, embora a literatura disponibilize numerosas pesquisas sobre o fenômeno da concordância variável na perspectiva da sociolinguística, raros são os estudos conduzidos na área da psicolinguística experimental e, com exceção de poucos trabalhos recentes (JAKUBÓW, 2018; MOLINA, 2018; REIS, 2020), não se tem conhecimento de investigações que abordem o tema a partir da análise de dados infantis sob o prisma da referida interface no PB.

Especificamente no tocante à SF, nosso estudo mostra-se relevante no sentido de apresentar uma investigação minuciosa desse fator – até onde sabemos, inédita na área da psicolinguística com foco na aquisição – tanto no que se refere às dimensões da SF (número de sílabas, padrão de acentuação e processos de formação de plural), como também no controle de tais variáveis. Ainda em relação à SF, é importante ressaltar que a pesquisa experimental desenvolvida no âmbito desta tese dialoga com nossos estudos anteriores (AZALIM, 2016; AZALIM et al, 2018) que, por sua vez, trazem evidências empíricas que sugerem que a SF não seria um fator com efeitos tão nítidos quanto poderia parecer inicialmente a partir de pesquisas desenvolvidas na

área da sociolinguística variacionista.

No que diz respeito aos fatores sociais considerados – grau de escolarização e origem do falante (rural X urbano) –, nossa escolha justifica-se em virtude de estes serem tidos como condicionadores relevantes na alternância da realização dos padrões de concordância.

Além disso, a tese fornece um conjunto de dados que se configuram como uma importante fonte de informações acerca do assunto mais amplo de nossa pesquisa, i.e. a concordância nominal variável no PB, no sentido de permitir levantar e analisar outras questões associadas à marcação de concordância, tais como configuração estrutural dos sintagmas, a presença de numeral e os padrões de marcação internos aos sintagma, no intuito de embasarmos a discussão teórica sobre o tema.

Do ponto de vista da teoria formalista, a presente tese se justifica ao avançar em dois níveis diferentes. O primeiro deles é que a tese configura-se com uma das poucas pesquisas a propor uma perspectiva dialógica entre teoria formalista de base gerativista, a sociolinguística e a psicolinguística. O segundo, por sua vez, concretiza-se na proposta de uma formalização para o fenômeno da concordância nominal de número de maneira ampla e também para os padrões de variação encontrados no PB, buscando propor um sistema que permita superar algumas limitações encontradas em análises anteriores.

1.3 ORGANIZAÇÃO DO TRABALHO

Esse trabalho está organizado da seguinte forma: no segundo capítulo, apresentamos uma resenha de algumas das principais pesquisas com base em dados provenientes de diferentes regiões do Brasil sobre o fenômeno da concordância variável no PB na perspectiva da sociolinguística variacionista, enfatizando os fatores linguísticos e extralinguísticos que têm se mostrado relevantes como favorecedores ou desfavorecedores da presença de marcações morfológicas explícitas de plural no âmbito do DP, com destaque para o papel desempenhado pela SF. Ainda no capítulo dois, apresentamos uma resenha de trabalhos que investigam o fenômeno da concordância nominal variável no PB tanto do ponto de vista do processamento quanto da aquisição da linguagem. No Capítulo 3, apresentamos a análise conduzida a

partir de *corpora* longitudinais e as considerações tecidas a partir de dados de crianças e adultos no que tange à alternância entre os padrões de concordância nominal de número. No Capítulo 4, apresentamos o estudo experimental conduzido no âmbito desta pesquisa, no qual utilizou-se a técnica de produção semieliciada a partir de imagens, no intuito de propiciar a produção de sintagmas plurais e investigar as principais dimensões comumente associadas à SF na realização da concordância variável no interior do sintagma nominal. No Capítulo 5, dedicado às reflexões formalistas, apresentamos de forma sucinta alguns conceitos importantes dentro do PM. Além disso, fazemos uma breve apresentação de algumas propostas de análise formal que visam a dar conta da caracterização da concordância no âmbito do DP no PB, formuladas dentro do arcabouço gerativista, bem como da nossa proposta de análise para a concordância nominal variável. Na construção do nosso sistema propomos uma associação entre o modelo de classificação de morfemas, o *4M* (MYERS-SCOTTON; JAKE, 2000a) e o sistema de *Agree* (FRAMPTON; GUTTMAN, 2000; PESETSKY; TORREGO, 2007). Por fim, no Capítulo 6, tecemos as considerações finais desta tese e indicamos futuros encaminhamentos para a pesquisa.

2 CONCORDÂNCIA NOMINAL VARIÁVEL NO PB E O PAPEL DA SALIÊNCIA FÔNICA

A concordância variável no PB é, ao mesmo tempo, um dos fenômenos mais pesquisados no cenário sociolinguístico brasileiro (LUCCHESI, 2019) e um dos tópicos do ensino formal de gramática que mais estigmatizam aqueles que se distanciam da norma tida como “cultura” (RODRIGUES, 2006). No entanto, de acordo com Scherre (2005, p.137), a concordância variável de número no PB, caracterizada como “erro” ou “desvio da norma” pela gramática normativa, “obedece a padrões estruturais e sociais de riqueza ímpar que precisamos enxergar, entender e conservar”. Nesse sentido, num contexto de variação como o que caracteriza o PB, destacamos a relevância de estudos que busquem investigar, do ponto de vista cognitivo, de que modo aspectos linguísticos variáveis são processados pelos falantes, no intuito de compreender melhor os fatores – linguísticos e extralinguísticos – comumente apontados como condicionadores de tal variação.

Em virtude do anteriormente exposto, o presente capítulo é dedicado à revisão de pesquisas que versam sobre o tema da concordância variável no âmbito do DP, no intuito de apresentar como tais fatores têm se mostrado relevantes para a compreensão do fenômeno sob a perspectiva de diferentes abordagens. Especificamente, daremos destaque aqui à visão da sociolinguística variacionista, bem como a estudos que levam em consideração o processamento adulto e a aquisição da linguagem. Antes, porém, pensamos ser necessário fazer uma breve apresentação de alguns conceitos associados à gramática tradicional que servem de ponto de partida para nosso estudo da concordância nominal variável no PB.

2.1 CONCORDÂNCIA NOMINAL VARIÁVEL: CONTRAPONTO ENTRE A GRAMÁTICA TRADICIONAL E A SOCIOLINGUÍSTICA VARIACIONISTA

A gramática tradicional, também chamada de gramática normativa ou prescritiva, representa uma das concepções do termo “gramática” e é a abordagem usualmente adotada nas práticas escolares. Segundo Bechara (2015, p.54):

Cabe à gramática normativa, que não é uma disciplina com finalidade científica e sim pedagógica, elencar os fatos

recomendados como modelares da exemplaridade idiomática para serem utilizados em circunstâncias especiais do convívio social.

Para atingir tal objetivo, a gramática normativa estabelece regras fundamentadas no emprego da linguagem identificado nas obras de um grupo seletivo de escritores – “àqueles de linguagem corrente, estilizada dentro dos padrões da norma culta” (ROCHA LIMA, 2017, p.38-39). Ainda de acordo com Bechara (2015, p.54): “a gramática normativa tem por objetivo recomendar como se deve falar e escrever segundo o uso e a autoridade dos escritores corretos e dos gramáticos e dicionaristas esclarecidos.”

No que diz respeito especificamente ao fenômeno da concordância, no ensino formal de português é apresentada apenas a forma redundante de concordância como sendo “a correta” na norma culta (MARTELOTTA, 2011, p.45). Nesse contexto, fica claro porque a gramática normativa trata fenômenos de variação linguística – tais como o da concordância variável de número – como desvios da norma ou erros, ou seja, como aspectos que fogem e ferem as regras daquilo estabelecido como “norma padrão” da língua. Pelo fato de serem tidas como “falhas”, esses tipos de ocorrência devem, na visão tradicional, ser evitadas em prol da manutenção da “boa formação” da língua.

A sociolinguística variacionista, por sua vez, desenvolvida a partir das pesquisas de Weinreich, Labov e Herzog (1968), Labov (2008 [1972]), dentre outros, se distancia da visão prescritiva e propõe o estudo da gramática das comunidades de fala no intuito de sustentar a hipótese da heterogeneidade das línguas, baseada nas diferenças sistemáticas observadas tanto no nível linguístico quanto social, como podemos ver na citação extraída de Weinreich, Labov e Herzog (1968, p.187-8):

The association between structure and homogeneity is an illusion. Linguistic structure includes the orderly differentiation of speakers and styles through rules, which govern variation in the speech community; native command of the language includes the control of such heterogeneous structures.

Nesse sentido, para Weinreich, Labov e Herzog (1968), a língua, longe de ser um fenômeno invariável, exhibe estruturas heterogêneas que, segundo eles,

configuram-se como parte da competência linguística⁷ dos falantes. Nesse contexto, fenômenos tais como o da concordância variável de número, “deixam de ser considerados falhas, desvios da norma padrão, e passam a fazer parte da lista de construções permitidas e chanceladas pela língua” (TARALLO, 2002, p.8). Sendo assim, para a sociolinguística, a língua apresenta um dinamismo próprio e, por essa razão, é reconhecido que os falantes fazem uso de “variantes linguísticas”, ou seja, formas diferentes para expressar sentidos semelhantes, ou como expresso nas palavras de Tarallo (2002, p.8): “diversas maneiras de dizer a mesma coisa, com o mesmo valor de verdade”. Importante lembrar que, para Weinreich, Labov e Herzog (1968), tal variação pode ser observada tanto na fala de uma comunidade linguística quanto na produção de um mesmo indivíduo.

Vale destacar que a concordância variável no PB – tanto no âmbito do DP quanto no nível sentencial – configura-se como um dos fenômenos mais amplamente estudados no contexto da sociolinguística no Brasil. Tais estudos concentram-se, principalmente, em explorar os fatores linguísticos e sociais condicionadores da ocorrência das variantes linguísticas, e partem da premissa de que a alternância de padrões é inerente ao conhecimento de um falante sobre sua língua, e que seu estudo pode promover uma melhor compreensão acerca dos processos de variação e mudança linguística.

⁷Aqui, faz-se necessário mencionar a distinção entre as diferentes acepções para o termo “competência linguística”. Do ponto de vista da teoria gerativa, por competência linguística entende-se o conhecimento (mental e inato) que o falante-ouvinte possui da sua língua. Na visão de Chomsky, a competência pode ser estudada independentemente do uso concreto e efetivo da linguagem (i.e. do desempenho). Em contrapartida, Labov critica o fato de a teoria linguística estar preocupada principalmente com “um locutor-ouvinte ideal, em uma comunidade de fala homogênea, que conhece sua língua (da comunidade de fala) perfeitamente e não é afetado por condições gramaticalmente irrelevantes como limitações de memória, distrações, mudanças de atenção e interesse, e erros (aleatórios ou característicos) na aplicação de seu conhecimento dessa linguagem no desempenho real” (CHOMSKY, 1965:3-4 *apud* WEINREICH; LABOV; HERZOG (1968, p.125). Afinal, para Labov, a diversidade observada na fala não deveria ser vista como decorrente de erros de desempenho, e sim, como parte de uma descrição realista da competência de um membro de uma comunidade de fala (WEINREICH; LABOV; HERZOG (1968, p.125).

No que tange à concordância nominal, numerosas pesquisas têm investigado o tema, iniciando pelos estudos conduzidos por Lemle e Naro (1976,1977), Braga (1977) e Scherre (1978). Dada a relevância desses trabalhos para a presente investigação, esta seção traz uma revisão de trabalhos conduzidos a partir de dados da produção de falantes residentes em diferentes regiões do Brasil, no intuito de fornecer um panorama sociolinguístico do fenômeno. Em particular, são destacados aqui estudos que abordam o possível papel da SF na alternância de padrões na marcação da concordância.

Como vimos, estudos conduzidos no âmbito da sociolinguística variacionista têm defendido amplamente a ideia de que fenômenos de variação linguística, tais como da concordância de número, não são aleatórios. Ao contrário, estariam condicionados a fatores tanto linguísticos quanto extralinguísticos. Dentre tais fatores, destacamos a SF, cujo princípio – formulado inicialmente por Lemle e Naro (1976, 1977) – estabelece que formas fonologicamente mais salientes, considerando a diferença de material fônico entre a forma singular e plural de um determinado item lexical como em *coração* [ko.ra.'sẽw̃] – *corações* [ko.ra.s'õjs], seriam mais perceptíveis acusticamente e, portanto, mais suscetíveis a serem marcadas morfologicamente no plural do que as formas menos salientes – tais como *coisa* ['koj.zẽ] – *coisas* ['koj.zɛs].

As pesquisas sobre o possível papel das características fônicas dos itens nominais na alternância dos padrões de concordância tiveram início com o estudo de Braga e Scherre (1976), conduzido a partir de dados de 7 falantes residentes no Rio de Janeiro, de classes sociais (baixa-média-alta), faixa etária e procedência geográfica distintas. As autoras propuseram uma escala hierárquica com cinco níveis de diferenciação no material fônico dos nomes, partindo do que consideravam ser os itens mais salientes em direção aos menos salientes, a saber:

- 1) Plural metafônico: inserção de -S e mudança silábica (olho, olhos);
- 2) Inserção de -S e mudança silábica (botão/botões, quartel/quarteis);
- 3) Inserção de -ES em palavras terminadas em -R (dor/dores);
- 4) Inserção de -S em palavras de plural regular (coisa/coisas).
- 5) Inserção de -ES em palavras terminadas em -S (vez/vezes).

Os resultados obtidos por Braga e Scherre (1976) são compatíveis com o

princípio da SF anteriormente mencionado e parecem sustentar a visão mais difundida na literatura de que formas mais salientes favorecem a marcação explícita de plural quando comparadas a formas menos salientes. Segundo as autoras, embora os dados tenham se mostrado compatíveis com a escala de saliência postulada em todos os grupos considerados, a hierarquia de saliência tem um papel mais expressivo nos dados das classes média e média alta.

Braga (1977), por sua vez, investigou a relação entre SF e nível social a partir de dados de 7 falantes de classe média e baixa do Triângulo Mineiro. Os resultados dessa pesquisa são semelhantes aos informados por Braga e Scherre (1976) e sugerem uma correlação positiva entre marcação redundante da concordância e grau de SF dos itens para os dados da classe média. Entretanto, essa mesma situação não foi verificada nos dados dos falantes da classe baixa que, por sua vez, apresentam apenas uma oposição binária: itens terminados em -S, de maneira geral, desfavorecendo a marcação de plural, versus os demais itens da escala hierárquica de SF, que favoreceram a marcação redundante.

Ponte (1979 *apud* SCHERRE, 1988) investigou a relação entre SF e grau de escolarização a partir de dados de falantes semi-analfabetos de Porto Alegre (RS). Adotando uma metodologia similar à de Scherre (1978), conclui que a hierarquia da saliência encontrada para os falantes gaúchos semi-analfabetos é bastante semelhante à dos falantes semi-escolarizados cariocas. Apenas uma diferença é reportada: a condição que menos favorece a aplicação da regra é a constituída por nomes regulares (livro/livros) e não pelos itens terminados em -S (mês/meses). Segundo Ponte (1979, p.180), os falantes semi-analfabetos gaúchos apresentam uma diminuição dos níveis da escala de saliência.

Carvalho Nina (1980 *apud* SCHERRE, 1988), por sua vez, aborda o mesmo tema mas, neste caso, considerando falantes analfabetos da micro- região Bragantina (PA). Em sua pesquisa, Carvalho Nina adota quatro níveis na escala da saliência – agrupando todos os casos de inserção de -ES em um único nível, em virtude de seu comportamento semelhante. A partir das análises conduzidas, a autora conclui que os processos de formação de plural não se mostram relevantes para a realização da concordância de número na fala dos seus informantes.

Mais recentemente, Oushiro (2015), investigou a relação entre SF e variáveis

sociais – Escolaridade, Classe Social e Mobilidade⁸ – na realização da concordância verbal de 1PP e 3PP⁹. A partir dos cruzamentos estabelecidos entre a SF e cada uma das variáveis sociais mencionadas, a autora observou que a hierarquia proposta para a saliência foi seguida por todos os grupos, praticamente sem exceção. Sendo assim, a omissão de marca explícita no verbo apresentou um percentual maior de ocorrências no Grau A – com menor grau de saliência – passando a diminuir gradualmente nos graus B, C e D, à medida que aumentava o grau de saliência. Em outros termos, a referida autora conclui que quanto maior o nível de escolaridade, maior *status* socioeconômico e maior a mobilidade geográfica do falante, maior será a realização da marca explícita de plural no verbo com alto grau de saliência fônica.

Considerados em conjunto, os achados das pesquisas citadas anteriormente (BRAGA; SCHERRE, 1976; PONTE, 1979; CARVALHO NINA, 1980; OUSHIRO, 2015) parecem corroborar a relevância de se incluir fatores sociais na investigação dos possíveis efeitos da saliência fônica. Em particular, escolaridade e procedência geográfica dos falantes – bem como a mobilidade dos mesmos – parecem apresentar algum tipo de correlação com os efeitos de saliência. Vale destacar, no entanto, que o estudo de Oushiro (2015), apesar de trazer dados bastante recentes, considera a saliência associada a itens verbais e não a nomes quando propõe essa investigação. Assim sendo, nossa própria investigação – que leva em consideração alguns dos fatores sociais mencionados e fornece dados contemporâneos – se mostra particularmente interessante para uma visão atualizada sobre o papel da saliência fônica no que tange à concordância nominal.

A partir dessas pesquisas pioneiras conduzidas nas décadas de 70 e 80, a questão da SF tem sido retomada em diversos estudos. Contudo, como apontado por Scherre (1988) e reforçado por Chaves (2014), o papel que a mesma desempenha na realização da concordância de número ainda é um tema de debate, afinal, os estudos conduzidos nem sempre apresentam resultados convergentes. Há inclusive autores que rejeitam parcial ou totalmente a relevância dessa variável (cf. CAMPOS, 2015; NICOLAU, 1984, 1995).

⁸ Para a variável Escolaridade, os participantes foram estratificados em três níveis: Fundamental, Médio e Superior. Para Classe Social, foram considerados quatro níveis: A-B1, B2, C1 e C2-D. E no que tange à Mobilidade, os participantes foram estratificados em três níveis: Baixa, Média e Alta mobilidade geográfica.

⁹ Por 1PP entende-se primeira pessoa do plural e por 3PP entende-se terceira pessoa do plural.

As diferenças que podem ser identificadas entre os estudos já desenvolvidos recobrem aspectos distintos, mas uma questão fundamental diz respeito ao número de dimensões consideradas na caracterização da SF, isto é, quais seriam as propriedades que definem um item como mais ou menos saliente. Afinal, o que faz um item ser percebido – e por quem – como mais ou menos saliente? Para o estudo da SF, três dimensões¹⁰ têm sido destacadas até o momento:

- (i) processos morfofonológicos de formação de plural, tendo sido propostas escalas de saliência a partir de tal dimensão, como a encontrada em Scherre (1988) e a já mencionada, postulada por Braga e Scherre (1976), dentre outras;
- (ii) tonicidade da sílaba dos itens lexicais singulares (itens monossílabos átonos e tônicos, oxítonos, paroxítonos e proparoxítonos); e
- (iii) número de sílabas dos itens lexicais quando considerada sua forma singular (monossílabos, dissílabos e itens com mais de duas sílabas).

Scherre (1988), por exemplo, investigou a concordância nominal variável a partir de dados extraídos do Corpus Censo do Programa de Estudos sobre o Uso da Língua (PEUL), cujos informantes foram distribuídos em dois grupos: (i) 48 adultos e (ii) 16 crianças. Cada grupo, por sua vez, foi subdividido em função de: (i) anos de escolarização; (ii) sexo e (iii) faixa etária. Na análise, foram consideradas as três dimensões da saliência anteriormente mencionadas (processos morfofonológicos, tonicidade da sílaba e número de sílabas dos itens lexicais singulares). Os resultados sugerem que as duas primeiras (processos e tonicidade) exercem influência na realização da concordância nominal de número. Mais especificamente, em relação aos processos, os dados indicam que quanto maior for a diferenciação do material fônico na oposição singular/plural, maior será o número de ocorrências com marcação explícita de plural. A autora observou ainda um maior número de ocorrências de marcação redundante quando se tratava de itens oxítonos, resultado semelhante ao

¹⁰ As dimensões da SF das quais tratamos ao longo da presente tese são relativas aos itens nominais. No entanto, vale lembrar que a literatura disponibiliza estudos investigando a SF dos verbos, tais como Naro (1981), Chaves (2014), Scher (2021), dentre outros. A diferença de material fônico quando comparadas as formas singular e plural e a acentuação dos itens são aspectos que têm sido apontados como dimensões relevantes para a caracterização da SF dos verbos.

reportado previamente por Guy (1981), sugerindo uma influência da tonicidade da sílaba na realização da concordância redundante. Por outro lado, Scherre (1988) reporta que, diferentemente das duas primeiras, a terceira dimensão avaliada – número de sílabas dos itens lexicais singulares – não apresentou diferenças significativas.

Carvalho (1997), por sua vez, analisou a concordância nominal na fala urbana de Rio Branco (AC), correlacionando a alternância nos padrões a um conjunto de variáveis linguísticas e extralinguísticas, dentre as quais, destacamos a SF, a posição linear dos elementos no sintagma e grau de escolarização.

Nesse estudo, a SF foi analisada com base nas mesmas três dimensões acima destacadas: (i) processos morfofonológicos de formação de plural, (ii) tonicidade e (iii) número de sílabas dos itens lexicais singulares. A partir dos dados obtidos, a autora concluiu que a SF apresenta correlação positiva com a concordância variável, principalmente no que se refere aos processos morfofonológicos de formação de plural.

Além da SF, a posição linear dos elementos no sintagma também se mostrou relevante no estudo de Carvalho (1997), de forma que os resultados obtidos pela autora sustentam a visão mais difundida na literatura, de acordo com a qual a primeira posição do sintagma é mais frequentemente marcada morfologicamente no plural.

No concernente à variável grau de escolarização, os resultados encontrados por Carvalho (1997) são compatíveis com a hipótese mais frequentemente aventada na literatura acerca da correlação entre tal fator e a marcação/omissão da marcação explícita de plural no âmbito do NP (BRANDÃO; VIEIRA, 2012; DIAS; FERNANDES, 2000; SCHERRE, 1988; dentre outros). Em outros termos, os resultados reportados sugerem uma variação em escala progressiva do menos para o mais escolarizado, sustentando a hipótese de que quanto mais escolarizado for o informante, mais propenso ele se mostra a pluralizar os itens flexionáveis que compõem o NP.

Nessa mesma linha de investigação, Andrade (2003) analisou a concordância nominal a partir de dados de 24 informantes, sendo 12 de Tubarão (SC) e 12 de São Borja (RS), ambas cidades da região sul do Brasil. Considerando a SF como uma das variáveis linguísticas, os dados apresentam indícios sustentando uma correlação entre tal fator e a alternância nos padrões de concordância – pelo menos no que concerne à dimensão processos de formação de plural –, uma vez que alguns plurais irregulares,

tais como os terminados ortograficamente¹¹ em -R (o mar – os mares) e em -L (fácil – fáceis), favoreceram mais a aplicação do padrão redundante do que os plurais regulares. No entanto, a dimensão tonicidade dos itens lexicais não se mostrou relevante para a alternância entre os padrões de concordância nominal.

No que diz respeito às variáveis extralinguísticas consideradas, Andrade (2003) conclui que escolaridade se revelou bastante relevante no referido estudo, no sentido de que a realização do padrão redundante de concordância se mostrou diretamente proporcional aos anos de escolarização dos informantes.

Mais recentemente, Ribeiro (2013) investigou a variação linguística na concordância de número nos sintagmas nominais e verbais entre os falantes do município de Oliveira Fortes (MG). Para a análise da concordância nominal, foram selecionados como fatores linguísticos: (i) a SF na oposição singular-plural – considerando a dimensão processos de formação de plural, (ii) a posição linear dos itens no sintagma, (iii) a classe gramatical, e (iv) as marcas precedentes dos elementos que compõem o sintagma. No concernente aos fatores sociais, foram selecionados: (i) zona de residência, (ii) sexo, (iii) faixa etária e (iv) escolaridade. Foram registradas 1407 ocorrências de sintagmas nominais plurais e 810 ocorrências de sintagmas verbais plurais, dentre as quais registrou-se uma forte tendência pela variante não redundante, representando 89,6% das produções.

Quanto à SF do núcleo do sintagma nominal, foram registrados resultados compatíveis com pesquisas anteriores de acordo com as quais a presença de um item com traço [-saliente] favoreceria a omissão de marcas explícitas de plural, ao passo que o traço [+saliente] tenderia a favorecer a presença de concordância redundante no sintagma. Apesar disso, Ribeiro (2013) chama a atenção para o fato de que em sua pesquisa, o papel da SF não possibilita uma interpretação mais precisa da variação, uma vez que se observou um uso quase categórico de padrão não redundante de concordância, independente da SF dos itens produzidos. Ademais, a autora destaca que as ocorrências de nomes com o traço [+saliente] representam um número significativamente menor do que as de itens com traço [-saliente]. A escassa produção de itens nominais considerados [+salientes] também é reportada por Martins (2010)

¹¹ Vale ressaltar que apesar de estarmos tratando de saliência fônica e, portanto termos consciência de que uma representação fonológica dos itens analisados seria mais apropriada, optamos por manter a caracterização nos moldes da proposta original de Scherre (1988) espelhada na morfologia, no intuito de manter a unidade ao longo da presente tese.

na sua pesquisa sobre a concordância nominal por moradores do município amazonense de Benjamin Constant e, como veremos no Capítulo 3 desta tese, também foi identificada nos dados naturalísticos considerados na presente investigação.

Dentre os trabalhos conduzidos a partir de dados de falantes residentes na região nordeste do Brasil, a pesquisa de Tabosa (2016) investigou a concordância nominal com base na análise de dados extraídos de 24 entrevistas do *corpus O Português falado no Ceará* inserido no PROFALA¹², cujos informantes foram estratificados em função de: (i) sexo/gênero; (ii) faixa etária, e (iii) nível de escolarização. Dentre os fatores linguísticos considerados nesse estudo, interessam aqui, em particular, a posição e classe gramatical dos itens em relação ao núcleo do sintagma, os processos morfofonológicos de formação de plural, a tonicidade das sílabas dos itens lexicais singulares e as marcas precedentes de plural no âmbito do sintagma nominal.

No que tange aos resultados reportados, vale destacar que, as variáveis tonicidade da sílaba dos itens lexicais e escolarização não se mostraram relevantes nos dados analisados. Tal resultado configura-se como, no mínimo, intrigante, visto que o grau de escolarização tem sido comumente apontado como um dos fatores extralinguísticos que mais influenciam a alternância dos padrões de concordância no PB (BRAGA, 1977; OUSHIRO, 2015; SALOMÃO, 2010; SCHERRE, 1991; dentre muitos outros). Não podemos deixar de mencionar também a variável tonicidade das sílabas, que vem ganhando destaque nas pesquisas tanto na área da sociolinguística (CAMPOS, 2015; SCHERRE, 1988; dentre outros), quanto no âmbito da psicolinguística (AZALIM, 2016; AZALIM et al, 2018), no tocante às investigações sobre o papel da SF na alternância dos padrões de concordância.

Por outro lado, os resultados obtidos em Tabosa (2016) para a variável posição linear do constituinte no sintagma se mostraram compatíveis com o reportado em estudos anteriores (SCHERRE, 1978, 1988), que sugerem que a primeira posição é a aquela que mais favorece à presença da marcação de plural. Os resultados relativos às

¹² O Projeto Variação e Processamento da Fala e do Discurso: Análises e Aplicações. Desenvolvido em parceria pelos Programas de Pós-Graduação em Linguística e Teleinformática, ambos da Universidade Federal do Ceará (UFC) como objetivo de implantação de um sistema baseado em tecnologia da informação para análises e aplicações à língua falada e ao discurso. Para tal, o projeto conta com um banco de dados com *corpora* já existentes na referida instituição, a saber, o Português Não-Padrão do Ceará, o Português Oral Culto de Fortaleza, o Projeto AliB-CE, o *Corpus* de Língua Inglesa Falada.

variáveis posição e classe gramatical em relação ao núcleo do sintagma, também dialogam com pesquisas prévias que indicam que os constituintes antepostos ao núcleo se apresentam favoráveis à presença de marcação de plural no NP, enquanto que os constituintes alocados depois do núcleo tendem a concentrar menos marcas de plural nos elementos que compõem tal tipo de sintagma.

Ainda sobre a investigação de Tabosa (2016), os resultados para a variável processos morfofonológicos de formação de plural, reforçam a relevância do fator SF no fenômeno da concordância variável, uma vez que foi registrada uma tendência à marcação explícita de plural nos itens de plural duplo ou metafônico¹³ com alternância vocálica, podendo haver ou não inserção de –S, bem como nos casos de plural irregular, tal como ocorre nos itens terminados em –L ou em –R, em oposição aos constituintes com plural regular e aos itens terminados em –S, com inserção de –ES, que exibiram forte tendência à omissão de marca de plural.

Com isso, Tabosa (2016) conclui que os dados analisados na denominada região do CRAJUBAR, compreendendo os municípios de Crato, Juazeiro e Barbalha, do Cariri cearense, fornecem indícios, não só de que o comportamento linguístico de seus falantes não diverge muito dos falantes das outras regiões do Brasil no que se refere à concordância de número, como também que o fenômeno em questão se mostra como um processo de variação estável que, por sua vez, é regulada por fatores de natureza linguística e social.

Com base na literatura apresentada até aqui, pode-se constatar que diversos estudos têm apontado uma correlação entre SF e variação nos padrões de marcação da concordância plural no âmbito nominal. Ainda assim, vários pontos relevantes permanecem em aberto, dentre os quais destacamos os seguintes:

- (i) as três dimensões da SF postuladas até o momento (processos, tonicidade e número de sílabas) não têm sido investigadas com a mesma extensão, uma vez que nem todas as pesquisas que examinam o papel da SF, o fazem com base na investigação das três dimensões, geralmente focando apenas nos processos de formação de plural;

¹³ Por plural duplo, entendem-se os casos em que há duas possibilidades de marcas de plural: (i) inserção do morfema –S e (ii) plural metafônico – que consiste na abertura vocálica, ou seja, na alteração da vogal tônica fechada (ô) pela tônica aberta (ó), em um item ao ser pluralizado, tal como acontece em *olho* [o]λ[u] - *olhos* [ɔ]λ[us].

- (ii) têm sido reportadas diferenças nos efeitos da variável SF em função do grupo social, sendo que os possíveis efeitos parecem ser menos evidentes nos dados de falantes de classe socioeconômica mais baixa;
- (iii) nos dados considerados, itens menos salientes parecem ocorrer de forma significativamente mais frequente do que itens mais salientes, o que muitas vezes dificulta uma análise mais refinada do papel da variável em questão.

Com isso, levantam-se algumas questões sobre as quais julgamos importante refletir:

- (i) Em que medida as três dimensões comumente associadas à SF se mostram suficientes para se propor uma caracterização refinada para tal variável? Até que ponto o contexto sintático ou morfossintático em que os nomes aparecem também pode contribuir para definir o grau de saliência ou realce fônico de um item? Por exemplo, o realce fônico de um N numa configuração D+N (em que o D é átono) seria distinto do atribuído a um N numa configuração do tipo Numeral + N (com um numeral tônico)? Ou ainda sobre a configuração Numeral + N, será que a constituição morfofonológica dos numerais (terminados ou não em -S) teria algum efeito no grau de SF do N?
- (ii) Embora a posição linear que um item ocupa no sintagma não costuma ser apontada na literatura como sendo uma dimensão diretamente relacionada e/ou relevante para a SF, consideramos que é possível que exista alguma correlação entre tais variáveis (NARO; GÖRSKI; FERNANDES, 1999).

Além dos pontos acima colocados, é importante ressaltar que a literatura apresenta lacunas no que diz respeito ao modo como a SF afetaria o processamento, bem como a aquisição da concordância de número no âmbito do DP no PB, questões que abordaremos mais adiante. Antes, porém, nas duas próximas seções, sintetizamos alguns dos principais resultados obtidos por estudos com foco no processamento adulto e na aquisição da concordância variável no PB, em particular, no âmbito nominal.

2.2 CONCORDÂNCIA VARIÁVEL NO PROCESSAMENTO ADULTO: A VISÃO DA PSICOLINGUÍSTICA EXPERIMENTAL

Embora a concordância variável tenha sido vastamente explorada no campo da sociolinguística, o número de pesquisas sobre o assunto do ponto de vista da aquisição da linguagem e do processamento linguístico é relativamente pequeno, seja com base no processamento adulto (AZALIM 2016; AZALIM et al, 2018; HENRIQUE, 2016; MARCILESE et al, 2015; MARCILESE et al, 2017; SCHER, 2021), seja focalizando especificamente dados da aquisição infantil (CORRÊA; AUGUSTO; FERRARI NETO, 2005; JAKUBÓW, 2018; MOLINA, 2018; REIS, 2020; ROZA; CASAGRANDE, 2015). Dado que nem todos esses trabalhos abordam a concordância no âmbito nominal e/ou consideram de alguma forma o papel da SF na alternância de padrões, abordaremos aqui apenas aqueles que trazem uma contribuição mais direta para a melhor compreensão de pelo menos um desses dois pontos.

Desenvolvido no âmbito da psicolinguística experimental, o estudo conduzido por Marcilese et al (2015) investigou o processamento da concordância variável no DP e no domínio sentencial (relação entre sujeito-verbo) por falantes com nível superior de escolaridade. O estudo buscou verificar experimentalmente em que medida os falantes com alto nível de escolaridade tratam os dois padrões mais gerais de concordância – redundante e não redundante – como igualmente funcionais, bem como investigar se a informação de número manifesta exclusivamente em D seria suficiente para tomar um referente como sendo plural. Para tal, foi concebida uma tarefa experimental baseada na técnica da escuta auto-monitorada, conduzida com 48 participantes adultos (estudantes de cursos universitários ou que já possuíam ensino superior completo). Para fins de análise, foram consideradas como variáveis independentes: *concordância* (redundante vs. não redundante) e *congruência na retomada pronominal* (retomada congruente vs. incongruente). De acordo com esse *design*, no que diz respeito à variável *concordância*, foram contrastadas sentenças com concordância morfologicamente redundante (Ex. *Os esquilos desceram da árvore*) e não redundante (Ex. *Os esquiloØ desceuØ da árvore*). No que tange à variável *congruência*, foi comparada a retomada pronominal congruente com o número expresso pelo antecedente/sujeito previamente introduzido (=prônimo plural

Ex. *Os esquilo*Ø *desceu da árvore. Eles recolheram nozes no quintal* OU *Os esquilos desceram da árvore. Eles recolheram nozes no quintal*) e a retomada incongruente com o número do antecedente (=pronomes singular Ex. *Os esquilo*Ø *desceu*Ø *da árvore. Ele recolheu nozes no quintal* OU *Os esquilos desceram da árvore. Ele recolheu nozes no quintal*). Note-se que o sujeito da primeira sentença era sempre plural, ora apresentando marcação redundante, ora não redundante. Como variáveis dependentes foram considerados: o *tempo de escuta para cada segmento crítico*, bem como o *tempo de reação* e a *média de respostas-alvo* para as perguntas de compreensão que eram exibidas na tela do computador após a escuta de cada estímulo experimental (Ex. *Os esquilos procuravam nozes?* SIM).

Os resultados obtidos são compatíveis com a ideia de que ambas as regras de concordância seriam gramaticais para os falantes escolarizados de PB, além de que a presença da informação de número apenas em D viabiliza uma leitura plural para todo o sintagma, o que pode ser sustentado em virtude de não terem sido registradas diferenças entre as condições *redundante* e *não redundante*, no que tange à compreensão da pergunta de compreensão (o número de respostas-alvo foi equivalente em ambas as condições) e nem ao tempo de reação para fornecer a resposta à pergunta de compreensão. No entanto, os tempos médios de escuta (aferidos no verbo e no segmento imediatamente posterior) revelaram que a condição *redundante* parece ser processada mais rapidamente pelos universitários do que a *não redundante*, o que poderia ser interpretado como indício de que a variante não redundante demanda um maior esforço cognitivo para o seu processamento.

Nessa mesma abordagem, Azalim (2016) e Azalim et al (2018), investigaram o processamento adulto da concordância variável de número no âmbito do DP. Especificamente, buscou-se investigar em que medida a variável SF, influencia o processamento da concordância nominal variável no PB, bem como uma possível interação entre esse fator e a variável extralinguística *nível de escolaridade*. Além disso, buscou-se avaliar em que medida dados relativos à percepção e à produção se mostram compatíveis com os achados reportados previamente com base em análise de *corpora*. Para tal, foram desenvolvidos dois experimentos de produção eliciada por repetição. O primeiro experimento foi conduzido com um grupo de falantes adultos com ensino superior completo ou em andamento, bem como por um grupo com

Ensino Fundamental completo (cursando a EJA¹⁴). Foram utilizados nomes reais do PB tidos como mais e menos salientes, quando considerados os *processos morfofonológicos de formação de plural*. Os dados obtidos foram analisados com base na consideração das seguintes variáveis independentes: *saliência* (itens +/- salientes), *concordância* (redundante e não redundante) e *nível de escolaridade*. No que tange à variável *saliência*, nomes tidos como [+salientes], representados por itens terminados ortograficamente em -R ou -L, como por exemplo, (*dever/deveres*), (*quartel/quarteis*) foram contrastados a nomes [-salientes] cuja forma plural se dá apenas pela inserção de -S tais como, (*maiô/maiôs*). Vale ressaltar que a escolha de tais itens como [+/- salientes] toma como base a escala hierárquica de SF proposta por Scherre (1988). As variáveis dependentes consideradas foram: o *número de repetições-alvo* e o *tempo de reação* para a repetição, que, neste caso, consistiu-se no intervalo entre o fim do estímulo auditivo e o *onset* da repetição. Para o segundo experimento, foram criados pseudonomes (Ex: *nabé, mecal*), respeitando o padrão fonotático da língua e também dos nomes reais utilizados no primeiro experimento (Ex. *boné, varal*), visando a manter a distinção entre itens mais e menos salientes, com o objetivo de isolar a influência da dimensão morfofonológica da SF de outros aspectos relevantes que possam ter afetado a resolução da tarefa no Experimento 1, por exemplo, a frequência de uso dos itens lexicais nas suas formas plurais e singulares.

Em conjunto, os resultados parecem contrariar a visão mais difundida na literatura com relação ao papel da SF na concordância nominal variável, mas se mostram compatíveis com o reportado por alguns estudos prévios (CAMPOS, 2015; NICOLAU, 1984; 1995;) e sugerem que formas pluralizáveis tendem a ser marcadas morfológicamente na forma redundante por falantes com nível superior completo, independentemente do grau de *saliência* – definida nesse estudo em termos da dimensão *processos morfofonológicos de formação de plural*. No concernente à flexão de número dos elementos no interior do DP, Azalim et al (2018) compartilham com Campos (2015) a ideia de que a SF não parece exercer uma influência significativa sobre o comportamento linguístico dos falantes do PB com alto nível de

¹⁴Educação de Jovens e Adultos (EJA): modalidade de ensino que trabalha com o mesmo conteúdo e pretende desenvolver as mesmas competências do ensino fundamental e médio, sendo, no entanto, voltada especificamente para o público adulto (maiores de 18 anos).

escolaridade.

Por outro lado, faz-se necessário ressaltar que os resultados do Grupo 2 (alunos da EJA) no Experimento 1 sugerem que tal variável poderia ter alguma influência – em termos perceptuais – no comportamento linguístico de falantes com menor grau de escolaridade. No entanto, os resultados de Azalim et al (2018) sugerem que um eventual efeito estaria mais vinculado à dimensão relativa à *tonicidade* do que a *processos morfofonológicos de formação de plural* propriamente ditos – sendo que essa última – como visto anteriormente - é a dimensão a partir da qual a variável SF é mais comumente caracterizada e analisada. Essa interpretação se mostra compatível ainda com o contraste identificado – nos grupos avaliados – no que diz respeito às formas redundantes e não redundantes quando associadas aos denominados “plurais regulares”, tipicamente tidos como [-salientes] fonologicamente. Contrariando as expectativas iniciais, nos grupos analisados, foram registrados TRs significativamente maiores para itens [-salientes] na condição não redundante, quando comparados com os itens [-salientes] na condição redundante (ex. *os nabé* x *os nabés*), mas o mesmo não foi observado para os nomes tidos inicialmente como [+salientes] (*os mecal* x *os mecais*), contrariando, mais uma vez, as previsões iniciais.

Recentemente, Scher (2021) investigou a influência da SF no processamento da concordância verbal variável no PB por parte de falantes adultos universitários. Nesse estudo, a autora conduziu uma análise de *corpus* e um experimento de produção eliciada por repetição com pseudo-verbos. Os estímulos experimentais utilizados foram elaborados com base na proposta de classificação da SF nos verbos apresentada por Naro (1981)¹⁵ e os resultados reportados por Scher não revelaram efeitos estatisticamente significativos da variável saliência na marcação dos verbos. No que diz respeito à análise de *corpus*, a autora informa que não foi possível conduzir uma análise detalhada acerca da influência da SF na alternância de padrões

¹⁵ Para Naro (1981), a definição de SF associada aos verbos compreende duas dimensões: a diferença de material fonológico entre as formas singular e plural e a acentuação dos segmentos fonéticos referentes à oposição (acentuado X não acentuado). No entanto, importante destacar que, assim, como acontece no âmbito nominal, a relevância da SF na marcação de concordância verbal, bem como a formulação conceitual de tal fator ou ainda a disposição escalar das formas verbais em virtude da saliência têm sido objeto de debate na literatura. Não entanto, os níveis das dimensões comumente associadas à SF dos nomes nos parece seguir critérios mais bem definidos. Em virtude de nossa pesquisa focalizar a concordância no sintagma nominal, recomendamos a leitura de Chaves (2014) e Scher (2021) para uma revisão detalhada de estudos sobre a SF nos verbos.

de concordância verbal em virtude do número insuficiente de ocorrências de padrão não redundante nos dados. Das 16 ocorrências registradas, foram identificados verbos que não constavam da classificação inicialmente adotada pela autora, não foram registradas ocorrências de todas as classes de saliência presentes na classificação utilizada e algumas classes apresentaram uma representatividade muito reduzida.

Os resultados experimentais obtidos em função da variável SF para os nomes por Azalim et al (2018), para os verbos por Scher (2021), bem como aqueles obtidos a partir de uma abordagem descritiva (CAMPOS, 2015; CHAVES, 2017) parecem apontar para a necessidade de uma revisão da própria caracterização de SF quando a dimensão cognitiva e perceptual da variação é levada em consideração, revisão esta que representa um dos objetivos da presente tese e será retomada mais adiante no capítulo 4, que trata do estudo experimental desenvolvido na presente pesquisa.

Em virtude dos resultados conflitantes mencionados anteriormente no que tange à produção e percepção adulta numa abordagem psicolinguística, realizamos uma revisão da literatura em busca de resultados acerca da concordância variável na produção infantil do PB em função da relevância que a fala infantil possui no contexto de sua comunidade linguística. Dessa forma, na próxima seção, apresentamos os estudos que julgamos serem mais diretamente relevantes para a nossa pesquisa.

2.3 CONCORDÂNCIA NOMINAL VARIÁVEL NA AQUISIÇÃO DO PB

Um primeiro ponto a ser ressaltado é que a literatura na área da aquisição da linguagem é notadamente escassa no que diz respeito aos estudos que investigam os efeitos dos padrões variáveis na aquisição de uma língua. Dentre os estudos disponíveis conduzidos com base na aquisição do PB, a pesquisa de Corrêa, Augusto e Ferrari Neto (2005) investiga o processamento da concordância de número no DP. Para tal, os autores utilizaram uma atividade experimental baseada na técnica de identificação de imagens com crianças de dois anos de idade, cujo objetivo era verificar em que medida a informação fornecida pelo morfema de número –S seria levada em consideração pelas crianças para a identificação do referente de um DP contendo um pseudônimo (Ex. *Os dabo* X *Os dabos*). Os resultados sugerem que crianças na faixa etária investigada (18 a 30 meses de idade (média – 25 meses) são sensíveis à expressão morfológica de número e que a informação crucial seria

extraída de elementos da categoria D, uma vez que as crianças não parecem fazer uma diferenciação entre os dois padrões de concordância coexistentes no PB (redundante e não redundante). Nessa pesquisa – com foco na compreensão de enunciados – a SF não foi avaliada. É interessante destacar que, em estudo posterior, Castro e Ferrarino (2007) reportam diferenças entre crianças adquirindo o PB e crianças expostas ao Português Europeu (PE), sendo que as crianças adquirindo o PE demonstram preferência pelo padrão redundante de marcação da concordância nominal que seria categórico no *input* que as mesmas recebem.

No que diz respeito à investigação do fenômeno a partir de dados de produção, tem-se, por exemplo, o estudo conduzido por, Capellari e Zilles (2002), que investigaram em que medida o comportamento linguístico da criança se assemelha ao do adulto no tocante à marcação de plural. Além disso, ainda que de maneira indireta, buscou-se examinar a relação entre o comportamento linguístico da criança e o processo de alfabetização, tendo em vista que a escolarização poderia proporcionar às crianças uma maior exposição à variante padrão da língua.

Para atingirem tais objetivos, as referidas autoras analisaram dados longitudinais coletados¹⁶ por meio de entrevistas realizadas com uma criança por um período de aproximadamente quatro anos (entre os 4;3 e os 8;5 anos de idade do informante). Feitas as análises, as autoras concluíram que, semelhante ao reportado previamente por Lamprecht (1997), os dados ressaltam a importância do *input* na aquisição da concordância no PB.

No que concerne à possível influência da escola no uso progressivo de marcação redundante na produção das crianças, Capellari e Zilles (2002) argumentam que não foram registradas evidências empíricas sustentando tal ideia. Ao contrário, os dados obtidos sugerem o uso dos dois sistemas que as autoras supõem estarem presentes no *input*: o da modalidade falada e o padrão, que, por sua vez, segundo as autoras, parecem estar relacionados a contextos discursivos específicos. Vale ressaltar que no referido estudo, os dados foram coletados por meio de três contextos: (i) relato pessoal, (ii) narrativa de historinha e (iii) elaboração de narrativa a partir de imagens. Na análise, observou-se, por exemplo, que quando se tratava de uma situação de fala mais espontânea, ou seja, quando a criança fazia um relato pessoal, era comum

¹⁶ Os dados analisados na referida pesquisa pertencem ao Banco de Dados do projeto Desenvolvimento da Linguagem da Criança em Fase de letramento – DELICRI da Universidade Federal do Rio Grande do Sul- UFRGS.

observar a omissão de marcas explícitas de plural. Em contrapartida, quando a criança baseava seu discurso em textos escritos, isto é, fazia a narrativa oral de uma historinha para a entrevistadora, seu comportamento tendia ao uso de formas redundantes – ainda de acordo com as autoras – possivelmente em virtude da memorização das formas encontradas nas histórias, reforçando assim a ideia de que o *input* tem um impacto importante na produção das crianças.

Já em Simioni (2006), são reportados resultados obtidos de um estudo de caso, conduzido com uma criança adquirindo o PB e realizado por meio de uma coleta longitudinal, no intuito de se explorar o processo de aquisição da concordância de número no DP tomando como base teórica o arcabouço do PM (CHOMSKY, 1993, 1995). A partir dos dados de fala espontânea registrados, Simioni (2006) descreve a ordem de emergência dos padrões de concordância de número e marcação de plural observados:

- (i) Primeiramente se observa, por volta dos 21 meses, a emergência das regras que a autora chama de *formas de plural*¹⁷ *não padrão, esperado*¹⁸, que constituem sintagmas formados por numerais ou PPs indicando quantidade, como em *dois palhaço, um monte de bala*;
- (ii) Num segundo momento, em torno dos 25 meses, a criança começa a produzir as *formas de plural não padrão, não esperado*, que consistem

¹⁷*Formas de plural*, segundo Simioni (2006) seriam aquelas formas em que não se observa evidência morfológica do estabelecimento de concordância entre os elementos que compõem o sintagma, tais como aquelas em que a marcação explícita de plural é manifesta somente em N, ou ainda os sintagmas formados por Numeral+Ns ou Numeral+N. Por *formas de concordância*, a autora entende aquelas formas em que a marcação morfológica de plural é realizada de forma redundante, ou seja, em todos os itens flexionáveis do sintagma, ou apresenta-se na forma não redundante, podendo ser expressa no item à esquerda de N, preferencialmente, D.

¹⁸De acordo com Simioni (2006), por *marcação não padrão, não esperada* entende-se a manifestação de marca morfológica de plural apenas em N. Em contrapartida, por *marcação não padrão, esperada*, entende-se a marcação apenas em D ou no elemento mais à esquerda do sintagma, bem como expressões quantificadas com N no singular, tais como em *um monte de bala*. Nesse sentido, podemos entender que a noção de *marcação não padrão esperada* se refere às ocorrências de marcação de plural que, apesar de não seguirem o padrão culto do PB, são amplamente presentes na gramática do adulto. Já a *não esperada* representa a marcação considerada agramatical na fala adulta.

nos sintagmas nominais em que a marca morfológica de plural é expressa apenas em N, como em *o meninos, a florzinhas*¹⁹;

(iii) Um pouco mais tarde, por volta dos 30 meses, já é possível notar a emergência tanto de *formas de plural padrão*, como em *dois meninos* quanto de *formas de concordância não padrão*²⁰, *esperada*, como em *as cadeirinha*;

(iv) E aos 35 meses, observa-se a emergência das formas de concordância padrão, como em *os copos, as patas*.

A autora discute três hipóteses aventadas na literatura, visando a identificar a que (i) melhor possa explicar a variação observada na produção da criança analisada, e a variação observada entre os registros adulto e infantil, bem como (ii) seja capaz de prever como os padrões de concordância de número observados na fala da criança irão se desenvolver até atingirem o estágio da gramática adulta. A primeira hipótese apresentada por Simioni (2006) consiste na ideia defendida em Chomsky (1995) e Lopes (2004) de que o traço de número seria interpretável em N. A segunda consiste na proposta de Magalhães (2004) que argumenta que o *locus* da interpretabilidade do traço de número seria D. Por fim, a terceira hipótese compreende a proposta de Carstens (2000), para a qual o traço de número não seria interpretável nem em N, tampouco em D, e sim em uma terceira categoria Num, uma categoria funcional – localizada entre N e D na estrutura sintática – que, funcionaria como uma fronteira para a distribuição do morfema de plural na estrutura do DP. Por ser um núcleo funcional, Num nos remete à ideia de que o traço de número seja atribuído pela sintaxe e não pelo léxico. Segundo Simioni, a última hipótese apresentada, ou seja, aquela defendida por Carstens (2000) seria a única capaz de sustentar os padrões observados nos dados registrados, em particular no que diz respeito à variação observada entre as regras de concordância e marcação de plural utilizadas pelos adultos vs. a ocorrência, na produção da criança, de marcação explícita de plural

¹⁹Vale ressaltar ainda que, como a própria autora aponta, tal análise considerando a ordem de emergência dos padrões de concordância de número e marcação de plural observados, tem um caráter exploratório. Em função de contemplar os dados de apenas um falante, julgamos que seja, no mínimo, questionável considerar a emergência de *formas de plural padrão*, como em *dois meninos* quanto de *formas de concordância não padrão, esperada*, como em *a florzinhas*, como uma etapa propriamente dita.

²⁰O estudo não fornece, no entanto, informações relativas à frequência de ocorrência de cada um desses tipos de formas nos dados.

expressa exclusivamente nos nomes – tida como agramatical no português. Aproveitamos para apontar que discussões dessa natureza serão aprofundadas nos próximos capítulos da tese, quando da apresentação e análise de nossos dados.

Importante ressaltar que os estudos disponíveis sobre a aquisição dos padrões de concordância não se restringem à análise de *corpora*. Há estudos que têm abordado o tema com base em dados obtidos por meio de metodologia experimental. Miller e Schmitt (2010), por exemplo, investigam o papel do *input* variável na aquisição de morfologia gramatical. Em particular, as autoras questionam que tipo de impacto a variação sociolinguística teria na produção e na compreensão infantil em relação à aquisição da morfologia de plural. Para tal, são consideradas duas variedades do espanhol: (i) o espanhol mexicano (neste caso, especificamente, a variante falada na Cidade do México), no qual se observa um uso categórico da marca /-S/ de plural, e (ii) o espanhol chileno, no qual a marcação de plural se dá de maneira variável já que tal marca pode ser omitida em virtude de um processo fonológico de lenição do -S em sílaba final, processo este que pode reduzir o -S neste ambiente a uma aspiração ou à omissão do mesmo, como reportado por Lipski (1999).

As autoras levantaram duas importantes questões de pesquisa, a saber (i) o *input* variável concernente à morfologia de plural afeta a habilidade das crianças para associar a marca de plural a uma interpretação de “mais de um”? e (ii) no contexto de *input* variável, qual seria o papel desempenhado pela frequência? No intuito de responder a estas questões, foi conduzida uma série de três experimentos, utilizando as tarefas de produção eliciada a partir de imagens no Experimento 1, *Act-Out* no segundo, e seleção de imagens no terceiro experimento. Em conjunto, os resultados sustentam a ideia de que o *input* variável afeta o processo de aquisição, visto que as crianças chilenas – expostas a um *input* menos consistente – levam um tempo maior para adquirirem a marcação de plural quando comparadas às crianças mexicanas – expostas à regra categórica de marcação de plural.

No tocante à segunda pergunta de pesquisa, as autoras defendem que, apesar de inicialmente o fator *frequência* da presença de um determinado item no *input* ser considerada importante, talvez o tipo de *input* (variável ou consistente) seja mais relevante para determinar os caminhos da aquisição.

De volta para os estudos acerca da concordância nominal variável na aquisição no PB, citamos Roza e Casagrande (2015), que desenvolveram uma pesquisa a partir de dados de 19 crianças de 4 e 5 anos de idade frequentando uma escola particular do

município de Francisco Beltrão (PR) e de 10 adultos (de distintos níveis de escolaridade e classe econômica), coletados por meio de tarefas de produção e de imitação eliciada. Os resultados mostram que as crianças avaliadas empregam, de modo significativo, a marcação redundante, ainda que apresentem comportamento compatível com a existência de diferentes padrões de marcação de plural em sua gramática. As autoras chamam a atenção para um dado curioso registrado na tarefa de produção eliciada: as crianças registraram um maior percentual de ocorrências de plural redundante do que o registrado nos dados dos adultos (63% dos sintagmas plurais apresentou marcação redundante); na produção adulta, o percentual de ocorrências de plural redundante foi de 52,5%. Já na tarefa de imitação eliciada tal situação se inverte, com adultos apresentando um uso semicategórico do padrão redundante de concordância (89,7% dos sintagmas plurais produzidos), percentual bastante superior ao observado nos dados das crianças (59,5%).

Apesar das diferenças observadas entre os participantes, os resultados adultos e infantis se assemelham quanto aos aspectos que mais favorecem a alternância na marcação: os dados fornecem evidências compatíveis com a ideia de que a SF e a posição do elemento no âmbito do NP influenciam a alternância entre os padrões redundante e não redundante de concordância nominal. No que concerne à SF, os resultados de Roza e Casagrande sustentam a visão mais difundida na literatura (SCHERRE, 1988, 1991, 1994; SILVA, 2011): quanto maior a diferenciação entre a forma singular e plural de um item, maior a possibilidade de ocorrer marcação redundante. No caso dos adultos, itens mais salientes foram produzidos em 100% das instâncias com o padrão redundante (em ambas as tarefas). Já no caso das crianças, na tarefa de produção, itens mais salientes foram produzidos com marcação de plural 100% das vezes, mas apenas 40% das vezes na tarefa de imitação. Vale destacar que, como vimos anteriormente na revisão de estudos sociolinguísticos (MARTINS, 2010; RIBEIRO, 2013), nos dados de Roza e Casagrande (2015) o número de itens mais salientes produzidos também é proporcionalmente pequeno quando comparado como o número de itens menos salientes, em ambas as tarefas utilizadas: 16 itens mais salientes e 62 menos salientes nos dados dos adultos e 31 itens mais salientes e 94 menos salientes nos dados das crianças.

Jakubów (2018), por sua vez, investigou como crianças que vivem na cidade do Rio de Janeiro lidam com o *input* variável no que se refere às marcas morfofonológicas de concordância de número no PB. Partindo da premissa de que o

grau de uso dos padrões de concordância seria diretamente proporcional ao nível de escolaridade e pelo nível socioeconômico (NSE) do falante, tal como é comumente defendido na literatura de cunho sociolinguístico (ANDRADE, 2003; MARTINS, 2010; OUSHIRO, 2015; SCHERRE, 1978, 1988), Jakubów (2018) assume que a cidade do Rio de Janeiro configura-se como um ambiente de contato entre tais padrões, em virtude da interação constante entre pessoas de diferentes NSEs. Para testar suas hipóteses, Jakubów (2018) conduziu um experimento com base na técnica de produção eliciada por repetição com crianças em idade pré-escolar e com alunos do sexto ano do ensino fundamental de escolas das redes pública e particular do Rio de Janeiro.

Os dados coletados são compatíveis com os achados reportados por pesquisas sociolinguísticas (BRANDÃO, 2013) que consideram a variável *nível socioeconômico* como fator condicionador da produção variável de concordância de número. Os resultados do experimento de produção eliciada sugerem que tanto crianças em idade pré-escolar, quanto os alunos do 6º ano do Ensino Fundamental demonstram preferência pela variante padrão de concordância, embora seja possível perceber uma influência do fator *nível socioeconômico* nas taxas de produção da variante não padrão, em virtude de um número significativamente maior de produção da variante não padrão por alunos da rede pública de ensino. Jakubów (2018) argumenta ainda que os seus resultados são compatíveis com a hipótese de que o NSE do ambiente em que a criança está inserida exerce influência na alternância dos padrões de concordância.

Ainda no âmbito da aquisição da linguagem, mas desta vez com foco na concordância verbal, Molina (2018) investigou a produção²¹ de morfemas verbais de terceira pessoa do plural a partir da análise de dados naturalísticos de crianças residentes na região urbana de Juiz de Fora (MG) e na área rural. No concernente especificamente à SF, considerada aqui a partir da dimensão processos de formação de plural, os resultados apontam que a produção dos cuidadores adultos residentes na zona urbana apresenta uma forte tendência pela marcação redundante tanto para

²¹Vale destacar que Molina (2018) também investigou a compreensão dos morfemas verbais de terceira pessoa do plural por meio de metodologia experimental. No entanto, nesse estudo foi considerado apenas o padrão redundante de marcação da concordância (contrastando sentenças com sujeito nulo e preenchido). Por esse motivo, embora a pesquisa de Molina – assim como a presente tese – articule dados naturalísticos e experimentais, a concordância verbal variável somente é abordada nesse trabalho a partir de dados de fala espontânea.

formas verbais mais salientes (95% de marcação redundante) quanto para itens menos salientes (85% de padrão redundante). No caso das crianças da zona urbana, itens salientes foram marcados de forma redundante de forma mais consistente do que formas verbais menos salientes (72% para os primeiros e 53% para os segundos). Para as crianças residentes na zona rural, por sua vez, a marcação não redundante foi o padrão predominante na produção (independente da SF do verbo). No entanto, até mesmo nesse grupo é possível identificar diferenças em função de SF: quando o verbo era um item mais saliente foi registrado 45% de ocorrências de concordância redundante. Já para os itens menos salientes, 92% das instâncias corresponde à concordância do tipo não redundante. Considerados em conjunto, os resultados sugerem que a SF teria uma influência mais clara na produção das crianças do que na dos adultos, quando considerada a concordância verbal.

Por fim, o trabalho de Reis (2020) tem em comum com a nossa própria pesquisa o fato de investigar a concordância variável a partir de dados naturalísticos e experimentais. Nesse caso, a autora explora a produção e a compreensão da concordância variável tanto no âmbito nominal quanto verbal. Quanto à produção, de modo geral, os resultados reportados atestam a presença de variação nos padrões de marcação da concordância em ambos os domínios e nos dois tipos de dados de produção infantil considerados. Especificamente quanto à saliência fônica dos nomes, Reis (2020) não encontra efeitos estatisticamente significativos dessa variável na produção adulta e identifica um efeito na direção oposta daquela inicialmente prevista no caso das crianças. Em outras palavras, os resultados sugerem que nomes salientes ocorreram mais frequentemente sem a marca morfofonológica de plural na fala infantil do que nomes não salientes. Além do observado quanto à saliência fônica, outros dois achados reportados por Reis dialogam – como veremos nos próximos capítulos – com os resultados da nossa pesquisa empírica. O primeiro diz respeito à ideia defendida pela autora, segundo a qual a variação registrada nos sintagmas nominais e verbais refletiria processos distintos: enquanto no SN haveria de fato a aquisição de duas variantes, no caso da concordância verbal a alternância encontrada nos dados estaria associada ao processo de aquisição da concordância de número pela criança. O segundo achado relevante diz respeito às diferenças encontradas na produção de adultos e crianças no que diz respeito à proporção de marcação redundante e não redundante: os adultos parecem preferir o padrão redundante

enquanto as crianças produziram mais marcação não redundante, em ambos os tipos de sintagmas investigados.

No quadro (1) a seguir, apresentamos uma síntese dos trabalhos resenhados até aqui, no intuito de viabilizar ao leitor um panorama dos resultados reportados nas pesquisas sobre a atuação do fator SF na alternância de padrões de concordância variável no PB.

Quadro 1 – Síntese dos estudos sobre o papel da SF na realização da marcação de número no âmbito do DP

Pesquisa	Perspectiva de análise do fenômeno do fenômeno	Corpus / grupo experimental analisado	Resultados reportados	Parecer a respeito da relevância da SF	Justificativa
Braga e Scherre (1976)	Foi considerada uma variável: aspectos morfológicos de formação do plural	Dados de 7 falantes residentes no Rio de Janeiro, mas de classe social e origem geográfica distintas.	Itens terminados em -ÃO e em -L apresentam comportamento distinto dos regulares.	Favorável.	Formais mais salientes parecem favorecer mais marcas de plural do que as formas menos salientes.
Braga (1977)	Idem anterior.	7 falantes de classe média e baixa do Triângulo Mineiro.	Classe média mineira realiza a marcação explícita de plural, em função do grau de SF do item pluralizável.	Favorável, em relação aos falantes de classe média.	Dados da classe baixa mineira sugerem que os graus de SF não parecem exercer papel relevante quando da aplicação das regras de concordância de número
Scherre (1978)	Idem anterior.	Dados de 10 falantes da área urbana do Rio de Janeiro, estratificados em função de escolaridade: 6 semi- escolarizados, 3 universitários, e 1 com 11 anos de escolarização.	- Endossando Braga e Scherre (1976), itens em -L configuram o 2º fator que mais favorece a marcação explícita de plural. - Itens regulares se assemelham aos itens em -S, e não como os itens em -ÃO ou em -L.	Favorável no que diz respeito aos falantes escolarizados.	Falantes semiescolarizados também apresentam uma gradação na variável SF, porém, não tão marcada como nos falantes escolarizados. A escala observada nos dados daqueles falantes, apresenta apenas 3 níveis, e não 5.
Scherre (1988)	Foram consideradas 3 dimensões: processos	Entrevistas com duração de aproximadamente 60 minutos com 48 adultos e	-O plural duplo mais favorece a marcação formal de plural, e o plural regular é o que menos	Favorável.	Processos e tonicidade dos itens no singular foram consideradas relevantes no condicionamento da realização da marcação

	morfológicos, tonicidade do item no singular e número de sílabas do item no singular	16 crianças estratificados em função: (i) gênero; (ii) anos de escolarização, faixa etária.	a favorece. - os itens em -R favorecem mais a marcação de plural do que os itens em -ÃO.		explícita de plural nos itens analisados.
Carvalho (1997)	Idem Scherre (1988)	Dados de 24 informantes estratificados em função de (i) Sexo e (ii) grau de escolarização	Formas mais salientes em termos de processos são mais suscetíveis à marcação	Favorável	Os dados sugerem uma correlação positiva entre SF e marcação redundante, com destaque para a dimensão Processos
Andrade (2003)	Foram consideradas as dimensões Processos e Tonicidade	Dados de 24 informantes estratificados em função de: (i) idade, (ii) sexo e (iii) escolaridade	Itens com plural irregular (terminados em -L e -R) parecem favorecer mais a marcação redundante do que itens com plural regular	Favorável com ressalva	SF parece ser relevante somente para a dimensão Processos
Ribeiro (2013)	Foi considerada a SF da oposição singular/plural	Dados de informantes divididos em função de : (i) sexo, (ii) faixa etária e (iii) escolaridade	O traço [-saliente] se mostra favorecedor da ausência de concordância	Favorável com ressalva	A variante ausência de concordância prevalece nas produções analisadas, independente do grau de saliência
Roza e Casagrande (2015)	Foi considerada a maior semelhança ou diferença na oposição	Dados de 19 crianças (4 e 5anos) e 10 adultos (distintos níveis de escolaridade e classe	Quanto maior a diferenciação de material fônico entre singular e plural, maior a possibilidade de ocorrer	Favorável	No entanto, número de itens [+salientes] produzidos também é proporcionalmente pequeno quando comparado à produção

	singular/plural	econômica) coletados por meio de tarefas de produção e imitação eliciada.	marcação redundante		de itens [-salientes]
Campos (2015)	Idem Scherre (1988)	15 inquéritos que compõem o <i>corpus</i> mínimo do projeto NURC.	Tendência de falantes ao uso da marcação explícita de plural nos itens pluralizáveis, independentemente do seu grau de SF. Das dimensões da SF consideradas, os resultados evidenciam apenas a relevância do subfator Tonicidade.	Desfavorável	A influência que a variável SF desempenha no comportamento linguístico dos falantes brasileiros com alto nível de escolaridade se se mostra insignificante.
Tabosa (2016)	Foram consideradas as dimensões: Processos e tonicidade	Dados de 24 informantes estratificados por: (i) sexo, (ii) faixa etária e (iii) nível de escolarização	Maioria dos itens com formação de plural irregular apresenta-se com os maiores índices de marcação de plural	Favorável	Os dados sustentam o princípio da SF, pelo menos no que diz respeito a Processos. Tonicidade não se mostrou significativa
Azalim (2016) e Azalim et al (2018)	Foi considerada a dimensão processos morfofonológicos de formação de plural	Dados de um grupo formado por: 49 estudantes universitários e 12 alunos de EJA coletados por meio de dois experimentos em que aplicou a técnica de produção eliciada ²² .	Formas pluralizáveis tendem a ser marcadas morfológicamente na forma redundante por falantes com nível superior, independente do grau de SF	Desfavorável	Falantes com ensino superior parecem contrariar a visão mais difundida na literatura com relação à SF. No entanto, os dados sugerem que a SF poderia ter alguma influência para falantes com menor grau de escolaridade. Um eventual efeito estaria mais vinculado à tonicidade do que à SF em termos de processos.

²²A técnica de produção eliciada será apresentada com mais detalhes na seção 4.1.1.

Reis (2020)	Foi considerada a dimensão Processos a partir de um contraste binário: itens [+salientes] X [-salientes]	Dados de um grupo formado por: 15 adultos e 75 crianças coletados por meio de um experimento de produção eliciada.	A variável não se mostrou significativa para os adultos.	Desfavorável	Os dados das crianças apresentaram diferença estatisticamente significativa, porém, na direção oposta à esperada: itens mais salientes favorecendo omissão das marcas de plural.
Scher (2021)²³	Foram consideradas 4 classes de saliência nos verbos dentre aquelas propostas por Naro (1981)	Dados de um grupo formado por 35 adultos com ensino superior completo ou em andamento por meio de um experimento de produção eliciada por repetição de pseudoverbos.	Não se observou diferença significativa em relação aos dois níveis de saliência da escala de Naro (1981) para os itens verbais. Tampouco se observou uma correspondência hierárquica entre as classes pertencentes à classificação.	Desfavorável	Os resultados sustentam a visão defendida em trabalhos anteriores (AZALIM et al, 2018; CAMPOS, 2015, dentre outros) de que a saliência, da forma como é definida e classificada atualmente na literatura, não se mostra relevante na alternância entre os padrões de concordância.

Fonte: Adaptado de Azalim (2016, p. 38-41)

²³Como vimos, o estudo de Scher (2021) não aborda a concordância variável no âmbito nominal, mas investiga o papel da SF na marcação da concordância de número no verbo. No entanto, pelo fato de se tratar de um estudo pioneiro que emprega uma metodologia semelhante à adotada na presente pesquisa, consideramos pertinente incluir esse trabalho na síntese de trabalhos relevantes aqui apresentada.

Em síntese, embora a literatura disponibilize um conjunto de estudos investigando os efeitos da SF na produção dos padrões de concordância nominal a partir de análise de *corpora*, bem como de dados coletados por meio de metodologia experimental, o exato papel de tal fator no fenômeno em questão ainda está longe de ser bem compreendido, seja sob um viés descritivo, seja a partir do ponto de vista do processamento ou, ainda, da aquisição da linguagem. Inserindo-se nesse debate, a nossa pesquisa tem como um de seus objetivos investigar a possível influência da SF na produção adulta e infantil dos padrões variáveis de concordância nominal no PB, com base em um diálogo entre investigações desenvolvidas pela sociolinguística e pela psicolinguística, articulando dados obtidos por meio da análise de corpora de produção espontânea e também a partir de duas tarefas experimentais de produção eliciada, que serão apresentados nos capítulos 3 e 4, respectivamente. Segundo Squires (2014), trabalhos recentes têm sugerido que o processamento de sentenças se mostra, pelo menos em parte, sensível a fatores extralinguísticos, ou sociais. Nesse sentido, julgamos que a articulação de dados longitudinais e de metodologia experimental, seja capaz de prover evidências mais robustas acerca do fenômeno da concordância nominal variável observado no PB. Passemos, então, ao capítulo que apresenta os dados naturalísticos analisados na presente pesquisa.

3 CONCORDÂNCIA NOMINAL VARIÁVEL NA PRODUÇÃO INFANTIL: DADOS NATURALÍSTICOS

Como vimos na introdução da presente tese, embora a aquisição da linguagem esteja estreitamente relacionada aos processos de mudança e variação linguística (LIGHTFOOT; WESTERGAARD, 2007), nota-se, na literatura, uma lacuna considerável no que se refere a pesquisas que investiguem os efeitos de regras variáveis na aquisição. Sendo assim, julgamos oportuno propor um diálogo entre os campos da sociolinguística, da psicolinguística experimental e dos estudos sobre aquisição da linguagem, por acreditarmos que, ao aliar dados naturalísticos – no caso, provenientes de coletas longitudinais – à metodologia experimental, seja possível obter evidências mais robustas acerca da produção infantil da concordância nominal variável observada no PB.

Neste capítulo, reportamos uma análise de dados naturalísticos elaborada com base em dois *corpora* coletados por Molina (2018) e um terceiro coletado por nós especificamente para esta pesquisa. O objetivo principal dessa análise consiste em investigar em que medida a variação linguística presente no *input* afetaria a realização dos padrões de concordância nominal por crianças adquirindo o PB. Sendo assim, julgamos pertinente conduzir a análise não somente da produção de crianças, como também de seus cuidadores – em especial, suas mães. A nossa escolha justifica-se em virtude de compartilharmos a ideia – bastante difundida na literatura – de que as mães possivelmente representam a principal fonte de *input* recebido pelas crianças²⁴, por serem, geralmente, as principais cuidadoras das crianças em fase inicial de aquisição linguística (ELY; GLEASON, 1996, p.256).

Como objetivos específicos, buscamos:

- (i) Verificar em que medida os padrões de concordância nominal variável no PB (redundante e não redundante) identificados pela sociolinguística, a

²⁴No entanto, tem-se, por exemplo, pesquisas tais como desenvolvida em Lieven (1994, p.58-9), que reportam estudos que têm investigado outras situações em que a linguagem é adquirida: há culturas em que as crianças passam bastante tempo com a mãe e outros irmãos/crianças; outras crianças ou pessoas que atuam como seus cuidadores, etc. Lieven reporta ainda que há estudos acerca de algumas culturas em que adultos falam com as crianças, mas de formas normalmente consideradas inúteis pelos pesquisadores da linguagem infantil ou aquelas em que há pouco ou nenhuma interação com crianças até que estas comecem a falar. E ainda assim, sabe-se que todas essas crianças aprendem a falar.

- partir de dados de fala adulta, são atestados na produção de crianças adquirindo o PB;
- (ii) Verificar em que medida a produção da criança se mostra compatível com o *input* recebido a partir da fala de seus cuidadores, quando consideradas amostras de diferentes grupos socioeconômicos;
 - (iii) Investigar possíveis efeitos da variável SF na produção espontânea dos padrões de concordância nominal na fala adulta e infantil;
 - (iv) Investigar possíveis correlações entre SF e origem do falante na alternância das formas de concordância nominal (zona urbana X zona rural);
 - (v) Fazer um levantamento das ocorrências de padrões relacionados ao sintagma nominal – Ex. D+N *As meninas*; D+N+Adj*As bonecas vermelhas*; Num+N *Dois alunos* – que possa alimentar nossa discussão teórica no intuito de fornecer uma formalização para o fenômeno da concordância variável no PB.

Partindo-se da premissa de que a concordância variável não é um fenômeno aleatório, estando condicionada à atuação de fatores de natureza linguística e/ou extralinguística (ALMEIDA, 1997; SCHERRE, 1988, 1991, 1994, 2014; dentre outros), a análise conduzida levou em consideração o possível papel na alternância de padrões de marcação da concordância de fatores tais como: (i) o nível de SF dos itens nominais; (ii) o tipo de estrutura sintática em que aparecem os nomes; (iii) os padrões de marcação de plural internos a cada tipo de sintagma e (iv) o grau de escolarização dos adultos que fornecem *input* linguístico (cuidadores primários da criança ou professor, a depender da amostra considerada).

Além disso, visamos, posteriormente, comparar tais resultados aos encontrados por nós em estudo anterior (AZALIM, 2016), no qual contrastamos o processamento adulto das regras de concordância nominal variável no PB a partir de dados de grupos com graus de escolarização distintos obtidos por meio de metodologia experimental.

Porém, antes da descrição do estudo e da análise dos resultados obtidos, julgamos necessária a descrição de alguns aspectos metodológicos importantes para a abordagem descritiva, fundamentalmente, a partir da perspectiva

sociolinguística variacionista, mas também relativos à coleta de dados naturalísticos da produção infantil.

3.1 ASPECTOS METODOLÓGICOS NO ESTUDO DA PRODUÇÃO LINGUÍSTICA INFANTIL E ADULTA

A sociolinguística variacionista é um ramo da linguística que se dedica ao estudo – identificação, descrição e explicação – da linguagem em uso, considerando-se as relações existentes entre a estrutura linguística e os aspectos sociais e culturais da produção linguística. Um aspecto fundamental desta abordagem diz respeito aos seus pressupostos teóricos, orientados à investigação acerca da regularidade e sistematicidade subjacentes a fenômenos que, considerados superficialmente, podem ser vistos como caóticos e desordenados. Segundo Tarallo (2002), a pesquisa variacionista investiga o padrão que subjaz à variação linguística, por considerá-lo parte importante do sistema linguístico de cada falante e também da comunidade de fala à qual pertence. Para tal, no contexto dessa abordagem, costumam ser conduzidos distintos tipos de análises. Numa perspectiva quantitativa, por exemplo, a metodologia utilizada costuma seguir três etapas. Na primeira, dá-se a coleta de dados, com base em critérios bem definidos. Num segundo momento, é feita a análise estatística e apresentação dos dados e, finalmente, faz-se a interpretação e explicação dos resultados. Em outros termos, nesta etapa, aliam-se os dados obtidos na análise estatística aos pressupostos teóricos assumidos na pesquisa, viabilizando, dessa maneira, a análise do fenômeno em questão, bem como a identificação e explicação do papel de fatores linguísticos e/ou sociais envolvidos na comunicação humana. Nesse tipo de abordagem, a coleta de dados é comumente feita por meio de entrevistas sociolinguísticas e costuma envolver, como já mencionado, informantes adultos.

Tendo-se em vista que o interesse da sociolinguística é analisar o vernáculo de uma comunidade de fala, Freitag (2010, p.88) defende que as entrevistas sociolinguísticas sejam a melhor estratégia para obtê-lo. Existem diversas possibilidades para a realização da coleta via entrevista, ou seja, diferentes técnicas podem ser empregadas, a depender do fenômeno investigado. Tem-se, por exemplo, as entrevistas estruturadas, nas quais todos os entrevistados respondem a perguntas padronizadas que foram formalmente organizadas num roteiro (KAJORNBOON,

2005). Além disso, há também as entrevistas semiestruturadas que se caracterizam por serem mais flexíveis, quando comparadas às anteriores. Isso porque durante as entrevistas semiestruturadas, em vez de perguntas padronizadas, o roteiro elaborado para as entrevistas pode consistir em temas específicos que sugerem tópicos que o entrevistado deseja explorar, embora as perguntas a serem feitas possam variar de acordo com os informantes. Essa forma de entrevista tem a vantagem de dar ao entrevistador a oportunidade de aprofundar em qualquer assunto. Ademais, (KAJORNBOON, 2005) menciona métodos de entrevistas sociolinguísticas ainda mais flexíveis, tais como as entrevistas não estruturadas em que não há a necessidade de se seguir um roteiro detalhado, ou seja, cada entrevista é diferente da outra, e os entrevistados são encorajados a responder de maneira bem à vontade as perguntas que vão surgindo de acordo com o desenvolvimento da conversa. Há também métodos de coleta naturalísticos que consistem em coletas rápidas, com perguntas pré-definidas que, por sua vez, exigem respostas pré-definidas, tais como as utilizadas por Labov (2008 [1972]) para seu estudo acerca a estratificação social do /r/ em Nova Iorque, com base em respostas de funcionários de três lojas de departamento. Para tal, o entrevistador perguntava aos informantes, – funcionários do local – onde ficava uma determinada seção da loja, a fim de obter como resposta a expressão *fourth floor* (quarto andar). Tal tipo de abordagem é bastante útil quando se pretende – tal como no estudo de Labov – conseguir um número significativo de dados em um curto espaço de tempo. Como vimos, existem diferentes formas de coleta de dados naturalísticos. No entanto, um aspecto comum a todos os tipos de coleta acima mencionados, diz respeito a alguns cuidados aos quais o pesquisador precisa estar atento no intuito de garantir a maior espontaneidade dos participantes. Dentre tais cuidados, destaca-se a necessidade de minimizar ao máximo o denominado “Paradoxo do observador” (LABOV, 2008 [1972]). Afinal, por menor que seja a intervenção do pesquisador na coleta de dados, a sua simples presença pode inibir os falantes e comprometer a naturalidade da produção.

Além disso, segundo Labov (2008 [1972]), os dados mais interessantes provêm de narrativas de experiências pessoais. Isso porque, ao envolver o falante em assuntos que remetem a emoções vividas por ele no passado, o informante deixa de prestar atenção em sua produção, gerando, então, uma fala mais natural, espontânea. Outra característica importante das entrevistas sociolinguísticas é que elas podem não só envolver diálogos entre entrevistador e informante, como também interações entre

pares de informantes, buscando minimizar a artificialidade de situação - o que nos remete ao fato de que cada tipo de entrevista pressupõe níveis de monitoramento diferentes.

Já no contexto dos estudos sobre a aquisição da linguagem, a coleta de dados produzidos por crianças costuma se dar de maneira um pouco distinta do que com adultos. Podemos citar como exemplo, a coleta de dados anedóticos, como são chamados os registros de fala espontânea, ou seja, simples anotações que os pais ou outros cuidadores costumam fazer na forma de diários. Tais registros são relevantes para as pesquisas na área de aquisição em virtude de viabilizarem o levantamento de várias questões sobre o desenvolvimento linguístico da criança.

Por outro lado, um dos tipos de coleta de dados de produção infantil mais utilizados é a coleta de caráter longitudinal. Em tal tipo de coleta, fazem-se registros da produção espontânea da criança durante um determinado período de tempo, geralmente em intervalos regulares. Tais registros podem ser feitos por meio da gravação de áudio ou de vídeo/áudio pelo próprio pesquisador ou por um adulto do círculo social da criança.

Na coleta de dados infantis, geralmente é necessário um tempo maior de gravação para que se consiga obter uma amostra significativa e variada de dados de fala das crianças que possibilite uma análise estatística mais robusta, principalmente no caso de crianças mais novas. Há também de se adaptar os métodos de coleta de dados, buscando proporcionar condições favoráveis à coleta. Por essa razão, Eisenbeiss (2010, p.15) aponta que muitos *corpora* de produção espontânea compreendem desde brincadeiras livres até outras atividades não estruturadas, tais como conversas na hora das refeições e atividades semiestruturadas, como por exemplo, jogos de tabuleiro, ou rituais associados à hora de dormir etc. Para encorajar as crianças a falarem, alguns pesquisadores lançam mão de algumas estratégias. Dentre elas, o uso de brinquedos contendo peças diversas que, além de permitirem uma variedade de atividades envolventes, encorajam a fala em situações que as crianças precisam dela para coordenar suas ações às de outras crianças.

Eisenbeiss (2010) defende ainda que os métodos naturalísticos são versáteis e oferecem um excelente ponto de partida para a pesquisa em uma ampla gama de fenômenos, incluindo o estudo do *input* que a criança recebe. Outra característica relevante de tais métodos consiste no fato de registrarem não só as formas e estruturas que as crianças ouvem e produzem, como também fornecer evidências de variação

linguística que permitem investigar os mecanismos usados pelas crianças no processo de aquisição da linguagem.

Além disso, como mencionamos anteriormente, pesquisas acerca da aquisição da variação, podem incluir nas análises não somente os dados de produção das crianças, como também os de seus cuidadores – em especial das mães – para verificar os possíveis efeitos que o *input* primário teria sobre as formas produzidas durante o processo de aquisição.

No que diz respeito aos registros, os pesquisadores devem estar atentos a seu envolvimento – nos mesmos moldes seguidos na coleta de dados de adultos -, no intuito de minimizar ao máximo sua interferência durante o processo de coleta de dados. Afinal, registros da interação entre crianças e pesquisadores não provêm dados de *input* representativos, bem como podem gerar efeitos significativos em decorrência da presença do observador. Na tentativa de minimizar tais efeitos, dá-se a oportunidade de a criança se familiarizar com o pesquisador antes de serem feitos os registros. Há, ainda, a possibilidade de os registros serem feitos pelos próprios cuidadores das crianças – como é o caso dos corpora aqui analisados. Já em relação aos ambientes propícios para os registros, Eisenbeiss defende que além da residência das crianças, as gravações possam ocorrer em instituições com as quais a criança esteja familiarizada, como escolas, creches, etc. No entanto, a referida autora argumenta que a residência da criança costuma ser o ambiente mais adequado em virtude de proporcionar a observação das crianças em suas atividades de rotina.

Na pesquisa aqui reportada, a metodologia escolhida para a coleta de dados foi pautada na abordagem naturalística e buscou neutralizar a influência do investigador. Assim sendo, todas as gravações foram realizadas pelos próprios cuidadores das crianças, sem a intervenção ou presença do pesquisador, no intuito de se preservar ao máximo o caráter espontâneo das interações entre adulto e criança. Somos cientes de que a presença de um dispositivo de gravação pode gerar efeitos no nível de monitoramento da fala – principalmente no caso dos cuidadores adultos. Todavia, consideramos que o fato de se tratar de interações cotidianas, no contexto doméstico, bem como a própria presença das crianças na cena comunicativa, são fatores que podem favorecer uma maior naturalidade (e um menor nível de monitoramento) das produções.

A partir do exposto até aqui, passemos à descrição dos dados coletados e das análises conduzidas nesta etapa de nossa pesquisa.

3.2 A PRODUÇÃO DA CONCORDÂNCIA NOMINAL VARIÁVEL POR CRIANÇAS ADQUIRINDO O PB

Como visto anteriormente, a alternância na realização dos padrões de concordância no PB constitui um fenômeno largamente investigado na literatura sociolinguística variacionista. Contudo, a maioria das pesquisas conduzidas nesta linha baseia-se em *corpora* formados pela fala de adultos e poucos trabalhos consideram a produção infantil (GOMES et al, 2011). Nesse sentido, o estudo conduzido e descrito nesta seção teve como objetivo principal coletar, descrever e analisar dados da produção oral relativa à concordância nominal variável de número por crianças em fase de aquisição do PB, bem como investigar o possível papel de fatores linguísticos – tais como SF, tipo de sintagma, padrões de marcação de plural internos ao sintagma–, e extralinguísticos – como grau de escolarização dos cuidadores (Ensino Superior X Ensino Médio) e origem do falante (zona urbana X zona rural) – na alternância na marcação morfofonológica da concordância nominal no PB.

Embora, como vimos no capítulo anterior, o papel da SF tenha sido amplamente estudado na literatura sociolinguística, não há um ainda consenso entre os pesquisadores no que tange ao tema. De fato, nossos próprios trabalhos anteriores (AZALIM, 2016; AZALIM et al, 2018) trazem resultados que vão na contramão da visão bastante difundida na literatura de que tal fator seria um dos principais condicionadores da alternância na realização dos padrões de concordância nominal no PB. Assim sendo, no intuito de refinar e aprofundar nossa pesquisa anterior, consideramos relevante investigar o papel da SF na marcação de número no âmbito do DP na produção de crianças adquirindo o PB, bem como comparar suas produções às de seus cuidadores. Buscamos assim caracterizar a produção infantil e verificar em que medida o *input* recebido poderia estar refletido no comportamento linguístico das crianças cujos dados compõem as amostras. Além disso, buscamos avaliar se a atuação de tal fator pode ser influenciada por algum outro, tal como grau de escolarização dos cuidadores primários (no caso, as mães), que representam uma das principais fontes de *input* linguístico para as crianças.

3.2.1 Constituição das amostras e metodologia de coleta

O material analisado foi extraído de três *corpora* longitudinais (identificados aqui como *corpus* A, B e C), totalizando 25 horas de fala gravada de adultos e crianças em situações de interação espontânea²⁵. As amostras foram coletadas em momentos distintos e cada uma delas teve objetivos e critérios específicos de estratificação dos participantes. O *corpus* A (MOLINA, 2018) é formado por produções em contextos de interação mãe/criança, em que todas as mães possuem grau de escolarização superior e pertencem à classe média, residentes da zona urbana de Juiz de Fora (MG) ou de Paraíba do Sul (RJ). Foram coletados dados de quatro crianças, de 3, 4, 5 e 6 anos de idade, respectivamente, sendo gravadas seis sessões com cada criança participante, com uma periodicidade mensal. O *corpus* B (MOLINA, 2018) é formado pela produção espontânea de crianças com idades entre 4 e 6 anos – divididas em três turmas: 1º e 2º períodos da Educação Infantil e 1º ano do Ensino Fundamental I – compreendendo uma sessão por turma, em contexto de interação com a professora em ambiente escolar, residentes na área rural de Juiz de Fora. Importante esclarecer que o *corpus* B não fornece dados de crianças com 3 anos de idade, visto que na escola em que foram conduzidas as coletas, são admitidas crianças a partir de 4 anos.

Finalmente, o *corpus* C é formado por duas crianças residentes na área urbana de Juiz de Fora (de 3 e 5 anos de idade, respectivamente), totalizando 8 gravações com a criança de 3 anos e 7 com a de cinco anos, realizadas a cada quinze dias e, neste caso, as mães possuem ensino médio. A escolha por grupos de mães com nível de escolaridade diferentes justifica-se em virtude de resultados de pesquisas prévias, que indicam uma correlação positiva entre a alternância da realização de padrões de concordância e fatores tais como grau de escolarização dos provedores de *input* linguístico – no caso, as mães – o que, por sua vez, teria um papel relevante no processo de aquisição da linguagem das crianças. Nesse sentido, o papel do *input* na produção infantil será analisado a partir dos dados dos *corpora* A e C, por configurarem registros semelhantes de interação entre crianças e seus principais cuidadores. Já a inclusão de dados de crianças residentes na área rural foi motivada

²⁵As gravações constituem um banco de dados transcrito e armazenado eletronicamente pela equipe do NEALP (Núcleo de Estudos em Aquisição da Linguagem e Psicolinguística) da Faculdade de Letras da UFJF.

em virtude do nosso interesse por explorar a variável origem do falante na produção dos padrões de concordância nominal no PB. Informamos ainda que não foi possível configurar um *corpus* semelhante aos A e C com crianças da zona rural pelo fato de não ter sido possível contatar famílias dispostas a realizar as gravações. De toda forma, consideramos válida a inclusão de um *corpus* “alternativo” – formado pelas produções de crianças em interação com a professora – para contar com, pelo menos, alguns dados da fala infantil nesse grupo específico. No quadro (2) abaixo, apresentamos uma síntese da composição dos *corpora* analisados:

Quadro 2 - Síntese da composição dos *corpora* analisados

Participantes	Corpus A	Corpus B	Corpus C
Crianças residentes em zona urbana	Menino A– 3 anos de idade (3;4 a 3;8) Menina B– 4 anos de idade (4;7 a 5) Menina C– 5 anos de idade (5;4 a 5;9) Menina D– 6 anos de idade (5;9 a 6;2)		Menino E– 3 anos de idade (2;11 a 3;1) Menino F– 5 anos de idade (5;9 a 6)
Crianças residentes em zona rural		Turma de 4 anos de idade Turma de 5 anos de idade Turma de 6 anos de idade	
Adulto em interação – Ensino superior (ES) ou Ensino médio (EM)	Mãe menino A – ES Mãe menina B – ES Mãe menina C – ES Mãe menina D – ES	Professora	Mãe menino E – EM Mãe menino F – EM

Fonte: Elaborado pela autora (2021)

Buscando prover aos participantes com condições favoráveis à coleta de dados, no primeiro *corpus* analisado (coletado por MOLINA, 2018), assim como no *corpus* C (coletados por nós para a presente pesquisa), as gravações foram feitas pelas próprias mães das crianças analisadas e no ambiente de sua escolha, a saber, suas próprias residências. No caso do *corpus* B, as gravações foram feitas pela professora responsável em contexto do ambiente escolar.

Visto que o nosso objetivo nesta etapa da pesquisa era analisar a produção de concordância nominal variável pelas crianças e, por se tratar de um fenômeno amplamente observado em contextos informais de modalidade oral da língua,

pensamos que tais *corpora* seriam significativos para nosso estudo. Trata-se de gravações da interação cotidiana entre mães e seus filhos em momentos lúdicos, o que, por sua vez, costuma despertar nas crianças um maior envolvimento no discurso e, conseqüentemente, pode fornecer uma maior quantidade de dados coletados.

Vale salientar ainda que a idade das crianças participantes dos *corpora* analisados nesta etapa da pesquisa variava entre os 3 e 6 anos. A escolha de Molina (2018) pela idade inicial de 3 anos justifica-se em virtude de resultados de estudo longitudinal realizado por Ferrari-Neto (2003), no qual, a partir de dados coletados de crianças com média de idade de dois anos, o autor defende que crianças em tal faixa etária já possuem certo domínio do sistema flexional de número em nomes, mas ainda não parecem dominar as relações de concordância entre os elementos que compõem o DP. Todas as gravações em áudio foram transcritas para sua posterior análise e consulta.

3.2.2 Caracterização dos dados

De posse das transcrições das gravações, demos início ao processo de caracterização dos dados, ou seja, identificamos e tabulamos todas as ocorrências relevantes para análise, a saber, todas as estruturas nominais plurais que envolvem concordância identificável (i.e. sintagmas formados por pelo menos dois elementos) na produção das crianças e dos adultos que compõem as amostras. Apenas no caso do *corpus* B não foi considerada a produção do adulto em interação por não se tratar da mãe ou cuidador principal das crianças.

Vale lembrar que, segundo as gramáticas de cunho normativo, a concordância nominal pode ser caracterizada como um processo de harmonização, em termos formais, de gênero e número, que se estabelece entre determinantes, modificadores e adjetivos e um substantivo núcleo de um NP. Contudo, sabemos que tal fenômeno envolve não só as relações entre os itens flexionáveis do sintagma nominal propriamente dito, mas também outras estruturas, tais como predicativos, participios passivos, dentre outras. Nesse sentido, para a primeira etapa de análise de nossa pesquisa, consideramos relevantes as ocorrências de concordância de número existente entre os elementos nas configurações especificadas a seguir, podendo a concordância ser marcada em todos os itens (concordância redundante) ou em pelo

menos, um deles (concordância não redundante), como observado nos exemplos abaixo, extraídos dos *corpora* aqui analisados²⁶:

- a) Determinante²⁷ + nome:
 (Redundante) “Ele foi lá pegar *as galinhas... os pintinhos...*” (Criança C)
 (Não redundante) “... o tucano... *os papagaio...*” (Criança A)
- b) (Determinante)²⁸ + Possessivo + nome:
 (Redundante) “Mãe, você quer ver *meus trabalinhos?*” (Criança B)
 (Não redundante) “*A minhas colega* vão ficar tudo com inveja.”(Criança C)
- c) (Determinante) + Numeral + nome:
 (Redundante) “Eu só tenho *dois saltos...*” (Criança D)
 (Não redundante) “Ela tem *dois brinco?*” (Criança C)
- d) Contração da preposição *De* + determinante + nome:
 (Redundante) “Eu tenho giz de cera... e canetinhas *das princesas*”. (Criança C)
 (Não redundante) “o dia *das bruxa.*” (Criança C)
- e) Contração da preposição *Em* + determinante + nome:
 (Redundante) “*Tônas nuvens.*” (Criança D)
 (Não redundante) “as rodinhas ‘caixa’ (encaixa) só *nas rodinha...*” (Criança A)
- f) Contração da preposição *Para* + determinante + nome:

²⁶ Importante destacar que na análise da produção dos sintagmas nominais/determinantes, não foi considerado seu *status* no interior do sintagma (i.e. argumento interno, externo ou adjunto). Assim sendo, os sintagmas analisados ora correspondem a sujeitos, ora a objetos, predicativos, adjuntos preposicionados, etc.

²⁷ Dentre os elementos pertencentes à categoria Determinante, foram considerados: artigos definidos (Ex. *o* livro) e indefinidos (Ex. *um* livro), bem como pronomes demonstrativos (Ex. *este/aquele* menino).

²⁸ A categoria D é indicada entre parênteses nas configurações em que a mesma pode ou não ser morfofonologicamente realizada.

(Redundante) “A gente vai dar presente *pras mães*.” (Criança C1)

(Não redundante) “...postar no *facebook pros meu coleguinha* ver...” (Criança C)²⁹

g) Determinante + nome + adjetivo:

(Redundante) “*As pecinhas pequenas* vão ficar comigo...” (Criança D)

(Não redundante) “*Os pés vermelho*...” (Criança A, 2)

h) Predicativos:

(Redundante) “... e aí, *eles* estão *alegres*...” (Criança C)

(Não redundante) “*Essas três* aqui é muito *fácil*.”

Importante destacar que, por entendermos a concordância como um fenômeno de natureza morfossintática, optamos por incluir no escopo de nossa análise somente as estruturas que apresentam marcas formais de plural em, pelo menos, um dos itens que compõem o sintagma em análise. Nesse sentido, as estruturas que apresentam uma leitura semântica de plural sem marcas morfológicas, tais como a observada em (4) foram desconsideradas.

(4) “muita pecinha.” (Criança D)

Além disso, optamos por excluir da análise, seguindo o exemplo de Scherre (1988), dados que ocorrem em situação de neutralização de morfema, como em “os seus”, “*as minhas são essas*” em virtude de compartilharmos com tal autora a ideia de que tais ambientes inviabilizam uma percepção acurada da presença ou ausência de marca de plural. Gostaríamos de ressaltar também que, expressões do tipo “Lojas Americanas” e “Estados Unidos”, bem como estruturas produzidas em forma de canção foram igualmente excluídas da análise. A razão para a exclusão decorre do fato de aquelas configurarem-se como estruturas já cristalizadas na língua e, no segundo caso, ao cantar uma canção, uma criança não está necessariamente produzindo um enunciado tal e como o faria de forma espontânea, ao contrário, está reproduzindo uma estrutura previamente memorizada. Em outros termos, tal produção

²⁹Nesse exemplo, vale esclarecer que, apesar da presença do pronome possessivo, o ponto fundamental que queremos destacar é a alternância da realização de concordância no nome, quando combinado com a contração.

não provê evidências muito informativas acerca dos padrões de realização de concordância pela criança. Importante lembrar que ocorrências com estrutura N+Adj também foram excluídas da análise em virtude de nosso foco serem DPs contendo Ds foneticamente realizados.

A partir do exposto até aqui, passemos à apresentação da análise dos dados de produção das crianças e dos adultos – no caso, especificamente as mães – em interação com elas.

3.3.3 Análise da produção

Como pode ser observado na tabela (1) abaixo, a alternância na marcação de plural foi verificada em todos os grupos considerados, tanto nas crianças quanto nos adultos. Também foi observado que a produção dos dois padrões de concordância parece variar em função do nível de escolaridade e da origem do falante. No caso das crianças, aquelas residentes na zona urbana cujas mães possuem ensino superior utilizaram a marcação redundante de forma mais sistemática (89% das ocorrências) do que as crianças de zona urbana cujas mães possuem ensino médio (83%) e do que as crianças residentes na zona rural (63%).

Tabela 1 - Distribuição dos sintagmas plurais em função da forma de concordância aplicada (redundante X não redundante) em cada *corpus*

Crianças	Padrão de concordância No total de ocorrências e %		Qui quadrado
	Redundante	Não Redundante	
<i>Corpus</i>			
<i>Corpus A</i> ³⁰ Crianças zona urbana	275 (89%)	35 (11%)	$p < 0.01$
<i>Corpus B</i> Crianças zona rural	29 (63%)	17 (37%)	$p = 0.07$
<i>Corpus C</i> Crianças zona urbana	122 (83%)	25 (17%)	$p < 0.01$
Adultos	Padrão de concordância No total de ocorrências e %		Qui quadrado
<i>Corpus</i>			
Adultos <i>Corpus A</i>	434	35	

³⁰ A título de informação, disponibilizamos nos Apêndices, a tabela (29) na qual se pode observar a distribuição de sintagmas plurais em função da faixa etária analisada em cada *corpus*.

Ensino Superior	(93%)	(7%)	$p < 0.01$
Adulto Corpus B – prof.	---	---	---
Adultos Corpus C	195	57	$p < 0.01$
Ensino Médio	(77%)	(23%)	

Fonte: Elaborado pela autora (2021)

A comparação entre as proporções de produção redundante e não redundante em cada *corpus* revelou diferenças significativas nos dados das crianças do *corpus* A ($\chi^2 = 429.08$ (2), $p < 0.01$) e do *corpus* C ($\chi^2 = 64.007$ (1), $p < 0.01$) e o mesmo foi observado para os adultos – *corpus* A ($\chi^2 = 339.45$ (1), $p < 0.01$) e *corpus* C ($\chi^2 = 75.57$ (1), $p < 0.01$). No entanto, essa diferença não foi estatisticamente significativa nos dados que compõem a amostra de fala infantil do *corpus* B (B ($\chi^2 = 3.13$ (1), $p = 0.07$)).

Ao compararmos entre si as proporções de produção não redundante dos adultos dos corpora A e C, o resultado de um teste de Qui-quadrado indicou uma diferença significativa ($\chi^2 = 32.479$ (1), $p < 0.01$), com um número de produção não redundante significativamente maior no grupo de mães com menor escolaridade. Nesse sentido, os dados dos adultos contribuem para reforçar a hipótese bastante difundida na literatura acerca da relação existente entre o nível de escolarização e a realização dos padrões de concordância nominal no PB (BRAGA, 1977, BRANDÃO; VIEIRA, 2012; CARVALHO, 1997; DIAS; FERNANDES, 2000, SCHERRE, 1988). Em outros termos, por estarmos analisando a produção de adultos com níveis de escolaridade diferentes, era esperada uma produção maior de plural redundante por parte dos adultos com nível de escolaridade superior em comparação aos adultos com ensino médio. Tal comportamento pode ser resultado de uma maior exposição dos falantes com nível de escolaridade superior às formas ditas de prestígio, reforçadas pelo ensino formal, o que, conseqüentemente, levaria a um aumento do uso das formas de concordância redundante (BRANDÃO, 2013).

No que diz respeito à produção das crianças, de modo geral, parece ir na mesma direção que o *input* recebido de seus cuidadores. Além disso, o resultado de um teste de Qui-quadrado comparando as proporções de produção não redundante das crianças entre os três *corpora* analisados indica que há diferença significativa entre os *corpora* tomados em conjunto ($\chi^2 = 21.67$ (2), $p < 0.01$), assim como na comparação entre pares envolvendo os *corpora* A versus B ($\chi^2 = 19.97$ (1), $p < 0.01$) e B versus C ($\chi^2 = 64.00$ (1), $p < 0.01$). No entanto, não se observou uma diferença significativa na comparação entre os corpora A e C ($\chi^2 = 2.71$ (1), $p < 0.09$). Em outras palavras, as

crianças residentes na zona urbana apresentaram um comportamento linguístico semelhante, a despeito das diferenças de nível de escolaridade dos seus cuidadores primários. As crianças residentes na zona rural, por sua vez, apresentaram um maior uso de formas não redundantes, quando comparadas às crianças da zona urbana. Tal diferença observada nos dados em função da variável origem do falante é compatível com a visão reportada em diversos estudos, tais como o conduzido por Ribeiro (2013), de que falantes residentes da zona urbana tendem a aplicar o padrão redundante de concordância com frequência significativamente maior do que residentes das zonas rurais.

Uma questão interessante que surge a partir da análise dos dados do *corpus* B considerado isoladamente, diz respeito ao fato de que o resultado registrado no âmbito nominal vai na contramão dos achados reportados por Molina (2018) que, a partir dessa mesma amostra, investigou a produção de morfemas verbais de terceira pessoa do plural. Em outros termos, nos dados registrados investigando a concordância no âmbito nominal – ainda que com proporções menores quando comparado aos *corpora* A e C –, observamos a predominância do uso de marcações explícitas de plural. Em contrapartida, no que tange à concordância verbal, Molina (2018) observou a predominância do uso de marcação não redundante. Das 33 ocorrências de formas verbais produzidas em contexto de sujeito de 3ª pessoa do plural, foram registradas 10 ocorrências (30%) com marcação redundante de plural, ao passo que a produção de marcação não redundante alcançou um percentual bem maior (70%), com 23 ocorrências.

Uma possível explicação para a diferença observada na produção das crianças que compõem a amostra do *corpus* B, quando comparada a marcação da concordância nominal e verbal, está baseada na proposta formulada por Reis (2020) segundo a qual a variação encontrada nos dados infantis quando considerados os domínios nominal e verbal, corresponderia a processos distintos. De acordo com a autora, apenas a alternância na marcação de plural no sintagma nominal, mas não no verbal, estaria de fato relacionada ao uso de variantes distintas. Afinal, os resultados reportados por Reis sugerem que a variação entre a marcação /não marcação de plural nos elementos que compõem o sintagma nominal seria adquirida naturalmente pela criança. Em contrapartida, os resultados referentes à variação na realização de concordância nos verbos parecem refletir um estágio no processo de aquisição da gramática da língua-

alvo, uma vez que a produção do padrão de marcação redundante tende a aumentar proporcionalmente à idade da criança.

Esse dado observado por nós no *Corpus B* é interessante ainda porque parece entrar em conflito com o que tem sido apontado na literatura para falantes adultos, tal como apresentado em Oushiro (2015, p.245), cujos resultados, obtidos a partir da produção de adultos paulistanos, sugerem a existência de uma correlação entre concordância nominal e concordância verbal de terceira pessoa do plural:

(...) os falantes tendem a favorecer ou desfavorecer as variantes não padrão simultaneamente, o que significa que o uso de uma variante prediz o uso da outra com alto grau de confiabilidade – se um falante tende a empregar CN-Ø, é alta a probabilidade de que também tenda a empregar 3PP-Ø.

No entanto, uma análise mais detalhada sugere que nossos dados não necessariamente contrariam os de Oushiro (2015), uma vez que nossa análise se baseia em valores brutos (mais ou menos de 50% de ocorrência de uma variante), ao passo que a análise de covariação conduzida por Oushiro (2015) compara padrões relativos, ou seja, o objetivo da análise nesse caso era verificar se os falantes que tendem a empregar a variante não padrão de uma variável também tendem a empregar a variante não padrão de outra variável (no caso CN e 3PP). Lembrando ainda que, quando se fala em “tender a usar uma variante” é levada em consideração a proporção da amostra como um todo. Sendo assim, um falante que produz 20% de concordância não padrão, por exemplo, pode estar acima da média da amostra.

Outro ponto importante a ser levantado é que, embora uma comparação superficial entre os dados de Molina (2018) para a concordância verbal e os dados reportados por nós para a concordância nominal possa sugerir que ambos os processos ocorreriam de forma “dissociada” (aparentemente ocorrendo mais marcação não redundante no domínio verbal do que no nominal), a diferença entre os dois padrões de concordância nominal não foi estatisticamente significativo. Esse resultado conjugado com as observações de Molina (2018) parece indicar que não há de fato predominância do padrão de concordância redundante no grupo considerado.

Finalmente, essa aparente “discrepância” entre os resultados obtidos para a concordância nominal e verbal a partir de um mesmo conjunto de dados reforça quão difícil é estabelecer comparações verdadeiramente legítimas entre os resultados reportados por estudos diferentes, seja pelo fato de que as análises estatísticas

conduzidas não são idênticas, seja porque o conjunto completo dos dados e análises precisa ser levado em consideração e muitas vezes nem todas as informações pertinentes se encontram disponíveis.

Após essa caracterização inicial dos dados, passamos a analisar todas as ocorrências com concordância nominal plural (redundante e não redundante) em função de um dos fatores linguísticos frequentemente arrolados na literatura como condicionadores dessa alternância na realização dos padrões de concordância, a saber, SF dos itens nominais.

A tabela (2) a seguir apresenta a distribuição dos sintagmas plurais quando considerada a relação entre saliência dos nomes (mais saliente X menos saliente) e marcação da concordância no sintagma (redundante X não redundante). Para a classificação dos nomes como mais ou menos salientes, consideramos aqui a escala hierárquica de SF proposta por Scherre (1988, p.75-76), que se baseia na dimensão processos morfofonológicos de formação de plural. Nesse sentido, analisamos os itens nominais a partir de um contraste binário: itens [-salientes] X itens [+salientes]. Como itens [-salientes], consideramos nomes cuja forma plural se dá simplesmente pelo acréscimo do morfema –S, os chamados plurais regulares. Já os itens [+salientes] são representados por nomes que correspondem aos demais níveis da escala de Scherre (1988), ou seja, cuja forma plural se dá por meio de mudança silábica, tais como ocorre em *papel-papeis*, *coração-corações*, ou por acréscimo de –ES, tais como *mar-mares* ou *vez-vezes*, ou ainda os casos de plural duplo, tais como *jogo-jogos*.

Observando a relação entre marcação redundante e não redundante quando considerados nomes mais salientes, por um lado, e menos salientes, por outro, é possível verificar que os percentuais absolutos de concordância não redundante são maiores quando associados a nomes menos salientes nos *corpora* A e C, tanto para os adultos quanto para as crianças – ainda que os modelos de regressão logística conduzidos não tenham indicado diferença significativa entre as proporções comparadas (concordância não redundante com nomes mais salientes X concordância não redundante com itens menos salientes), como sugerem os valores de *p* dispostos na tabela abaixo. No caso do *corpus* B (crianças residentes na zona rural), o fato de praticamente não haver ocorrência de nomes mais salientes (1 no total da amostra), não nos permite extrair qualquer conclusão. Vale lembrar que pesquisas prévias sobre esse aspecto (BRAGA, 1977; SCHERRE, 1978) sugerem que os efeitos de SF sofreriam influência de variáveis de natureza social, como nível de escolaridade e

origem do falante. Todavia, os dados aqui reportados não se mostram suficientes para trazer novas evidências nesse sentido.

Tabela 2 - Distribuição dos sintagmas plurais em função do padrão de concordância aplicada (redundante X não redundante) e da SF do nome em cada *corpus*

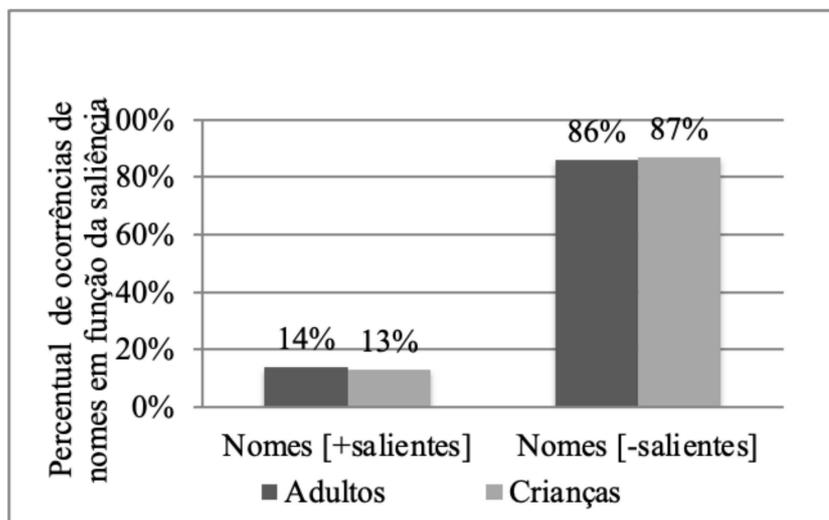
Grupo Crianças	Saliência Fônica	Padrões de concordância		Regressão logística	Total de ocorrências
		Concordância redundante	Concordância não redundante		
<i>Corpus A</i> ³¹ zona urbana	+saliente	40 (97%)	01 (2%)	$p=0.09$	41
	-saliente	236 (88%)	33 (12%)		269
<i>Corpus B</i> zona rural	+saliente	0 (0%)	01 (100%)	--	01
	-saliente	29 (64%)	16 (36%)		45
<i>Corpus C</i> zona urbana	+saliente	21 (88%)	03 (12%)	$p=0.5$	24
	-saliente	101 (82%)	22 (18%)		123
Grupo Adultos	Saliência Fônica	Padrões de concordância		Regressão logística	Total de ocorrências
		Concordância redundante	Concordância não redundante		
<i>Corpus A</i> Ensino Superior	+saliente	66 (100%)	0 (0%)	$p=0.9$	66
	-saliente	368 (91%)	35 (9%)		403
<i>Corpus B</i> – prof.	+saliente	---	---	---	---
	-saliente	---	---	---	---
<i>Corpus C</i> Ens. Médio	+saliente	36 (86%)	06 (14%)	$p=0.1$	42
	-saliente	159 (76%)	51 (24%)		210

Fonte: Elaborado pela autora (2021)

É necessário, contudo, salientar que, na interpretação dos resultados sobre SF apresentados anteriormente, seja levado em consideração o fato de que, nos três *corpora* analisados, sintagmas contendo nomes tidos como menos salientes (“plural regular”, formado pelo acréscimo de –S) foram predominantes (representando por volta de 86% das ocorrências dos adultos e em torno de 87% das ocorrências das crianças (cf. gráfico 1). Os resultados de testes de Qui-quadrado indicam que tal predominância foi significativa, com percentual estatisticamente maior de produção de nomes menos salientes em comparação aos itens mais salientes, tanto nos dados de crianças ($\chi^2 = 264.74$ (1), $p < 0.01$) quanto nos dados de seus cuidadores ($\chi^2 = 350.49$ (1), $p < 0.01$).

³¹Vale destacar que a tabela (30) contendo a distribuição dos sintagmas plurais em função da forma de concordância aplicada (redundante X não redundante) e da SF do nome em cada faixa etária encontra-se disponível nos Apêndices.

Gráfico 1 - Ocorrências de nomes em sintagmas plurais em função da SF nos dados naturalísticos (adultos e crianças)



Fonte: Elaborado pela autora (2021)

Algumas considerações podem ser tecidas a respeito dos dados apresentados. Em primeiro lugar, itens menos salientes parecem ser mais comuns no *input* recebido pelas crianças por possivelmente constituírem um inventário maior na língua de modo geral. Nesse sentido, a maior ocorrência de itens menos salientes na produção infantil representaria um reflexo do padrão encontrado na produção dos adultos. Além disso, o comportamento de crianças e cuidadores frente à influência da SF em suas produções, ou seja, o uso predominante de itens [-salientes] poderia decorrer do fato das mães estarem interagindo com suas crianças pequenas e por essa razão, as mães estariam ajustando/adaptando sua fala às habilidades linguísticas das crianças— uma das características da Fala Dirigida à Criança (ou FDC)³².

Nesse sentido, dada a pequena proporção de nomes mais salientes produzidos no material analisado, não foi possível estabelecer uma comparação direta

³² Estudos sugerem que a FDC apresenta várias características que podem desempenhar um papel relevante no processo de aquisição da linguagem. Dentre elas, é comum encontrar na FDC a ocorrência de frases curtas, repetitivas, além de – como apontado por Name e Sosa (2020) – diversas modificações sintáticas, lexicais e prosódicas quando comparada à fala entre adultos, tais como, número elevado de perguntas; repetição de palavras, diminutivos, dentre outros. Em outros termos, diríamos que a FDC auxiliaria as crianças a prestarem atenção ao código linguístico, uma vez que se adapta às habilidades linguísticas ainda em desenvolvimento.

entre nomes mais e menos salientes e nem obter informações mais detalhadas sobre a efetiva atuação das diferentes dimensões da SF, a saber, processos morfológicos de formação de plural, número de sílabas e padrão de acentuação dos nomes. No intuito de obter novos dados para uma discussão mais aprofundada desses pontos, foram conduzidas no âmbito desta tese duas atividades experimentais que reportamos no Capítulo 4.

Ainda no que se refere aos dados naturalísticos, foram conduzidas análises em função do tipo de estrutura em que os nomes ocorrem. Para fins de tratamento estatístico, foram consideradas as seguintes configurações:³³

- (i) Determinante + Nome (D+N) como em *os meninos*;
- (ii) Determinante + Nome + Adjetivo (D+N+Adj³⁴) *os piratas pequenininhos*;
- (iii) Numeral + Nome (Num+N) *três anos* e
- (iv) Numeral + Nome + Adjetivo (Num+N+Adj) *quatro flores rosas*.

Na tabela (3) abaixo, apresentamos a distribuição dos sintagmas plurais produzidos considerando a relação entre os tipos de estrutura do sintagma arrolados acima e a marcação da concordância no sintagma (redundante x não redundante). De modo geral, com exceção dos dados infantis dos *corpora* B e C frente à estrutura D+N+Adj, bem como dos dados infantis do *corpus* B frente à estrutura D+N, e dados infantis e adultos do *corpus* C frente à estrutura Num+N+Adj, em que observamos uma proporção maior de concordância não redundante, nota-se, em todos os *corpora* – tanto para os adultos, quanto para as crianças – uma predominância do uso de marcação explícita, com diferença estatisticamente significativa entre as proporções de produções redundantes em comparação à sua contraparte não redundante.

Foi observada ainda a produção mais frequente de estruturas compostas por dois elementos³⁵ quando comparadas às estruturas formadas por três itens. Os resultados dos testes de proporções indicam uma diferença estatisticamente

³³Vale destacar que a escolha dessas quatro configurações justifica-se em função do reduzido número de ocorrências de algumas das estruturas elencadas anteriormente (cf. Seção 3.2.2), fato que inviabilizou a condução de análises estatísticas para todos os padrões inicialmente encontrados.

³⁴“Adjetivo” é abreviado como “Adj” no corpo do texto e como “A” nas tabelas e figuras.

³⁵ Cabe esclarecer que estamos considerando o número de elementos efetivamente morfofonologicamente realizados.

significativa entre as proporções desses dois tipos de sintagmas, tanto nos dados das crianças – que produziram 295 (97%) sintagmas de dois elementos e apenas 09 (3%) sintagmas com três elementos ($\chi^2 = 265.32, (1), p < 0.01$) – quanto nos dados dos adultos, nos quais foram registradas 395 (96%) sintagmas de dois elementos e 17 (4%) sintagmas com três elementos ($\chi^2 = 344.97, (1), p < 0.01$).

Tabela 3 - Distribuição dos sintagmas plurais em função do tipo de estrutura e concordância em cada *corpus*

Grupo Crianças	Tipo de estrutura	Padrões de concordância		Regressão logística	Total de ocorrências
<i>Corpus</i>		Concordância redundante	Concordância não redundante		
<i>Corpus A</i> Crianças zona urbana	D+N	104 (89%)	13 (11%)	$p < 0.01$	117
	D+N+A	04 (100%)	0 (0%)	---	04
	Num+N	74 (94%)	5 (6%)	$p < 0.01$	79
	Num+N+A	0 (0%)	0 (0%)	---	0
<i>Corpus B</i> Crianças zona rural	D+N	05 (45%)	06 (55%)	$p = 1$	11
	D+N+A	0 (0%)	1 (100%)	---	01
	Num+N	06 (75%)	02 (25%)	---	08
	Num+N+A	0 (0%)	0 (0%)	---	0
<i>Corpus C</i> Crianças zona urbana	D+N	42 (86%)	07 (14%)	$p < 0.01$	49
	D+N+A	0 (0%)	2 (100%)	---	02
	Num+N	30 (97%)	01 (3%)	$p < 0.01$	31
	Num+N+A	0 (0%)	2 (100%)	---	02
Grupo Adultos	Tipo de estrutura	Padrões de concordância		Regressão logística	Total de ocorrências
<i>Corpus</i>		Concordância redundante	Concordância não redundante		
Adultos Corpus A Ensino Superior	D+N	132 (87%)	20 (13%)	$p < 0.01$	152
	D+N+A	09 (100%)	0 (0%)	---	09
	Num+N	107 (99%)	01 (1%)	$p < 0.01$	108
	Num+N+A	02 (100%)	0 (0%)	---	02
Adulto Corpus B – prof.	D+N	---	---	---	---
	D+N+A	---	---	---	---
	Num+N	---	---	---	---
	Num+N+A	---	---	---	---
Adultos Corpus C Ensino Médio	D+N	66 (70%)	28 (30%)	$p < 0.01$	94
	D+N+A	02 (67%)	01 (33%)	---	03
	Num+N	37 (90%)	04 (10%)	$p < 0.01$	41
	Num+N+A	01 (33%)	02 (67%)	---	03

Fonte: Elaborado pela autora (2021)

A partir desse levantamento dos dados, conduzimos alguns testes de Qui-quadrado, no intuito de investigar se a produção de adultos e crianças poderia apresentar um padrão mais ou menos sistemático a depender da configuração do sintagma, seja comparando o tipo de núcleo do sintagma (i.e. contrastando Determinante X Numeral), seja comparando as estruturas quanto ao número de elementos morfofonologicamente realizados (i.e. dois X três elementos). Sendo assim, para as análises, foram considerados contrastes entre:

- (i) todas as estruturas;
- (ii) estruturas encabeçadas por D x Num (com dois elementos);
- (iii) estruturas encabeçadas por D x Num (com três elementos);
- (iv) estruturas com dois x três elementos (D+N x D+N+Adj ou Num+N x Num+N+Adj).

No que diz respeito aos dados das crianças, os resultados dos testes de Qui-quadrado indicaram que não houve diferença significativa nos contrastes acima. No entanto, para os grupos de adultos, os resultados indicam uma diferença significativa entre as proporções de produções não redundantes a depender do núcleo do sintagma, com um número significativamente maior de sintagmas plurais não redundantes quando o sintagma era encabeçado por determinante como em *Os alunos* em comparação a numeral, tais como *Dois alunos*, tanto para o *corpus* A ($\chi^2 = 11.13$, (1), $p < 0.01$), quanto para o *corpus* C ($\chi^2 = 5.2747$, (1), $p = 0.02$). Dito de outra maneira, pelo menos no que tange às estruturas curtas (com dois elementos), nos *corpora* analisados a presença do numeral parece favorecer a marcação redundante quando comparado com o determinante. Nesse sentido, os resultados sugerem que o numeral nesse tipo de estrutura exerce alguma influência na realização dos padrões de concordância. Entretanto, faz-se necessário esclarecer que tais resultados dizem respeito apenas aos contrastes estabelecidos entre estruturas formadas por dois elementos morfofonologicamente realizados (D+N X Num+N). Afinal, o número muito reduzido de ocorrências registradas de sintagmas maiores, por ambos os grupos de adultos analisados, não nos permitiu extrair conclusões mais decisivas sobre as tendências observadas. Contudo, posteriormente foi possível estabelecer uma comparação entre estruturas de dois e três elementos, a saber, Num+N x Num+N+Adj, a partir dos dados coletados na tarefa experimental conduzida no âmbito da presente pesquisa, cujos

resultados serão reportados no próximo capítulo.

No tocante aos sintagmas encabeçados por numerais, despertou-nos o interesse por investigar se a constituição morfofonológica dos numerais (terminados ou não em –S) poderia desempenhar algum papel na marcação de concordância, como resultado de uma possível falsa analogia entre tal marca e o morfema de plural -S dos nomes.

A análise de dados conduzida por meio de um modelo de regressão logística, indica que a forma superficial dos numerais se mostrou relevante quando da aplicação dos padrões de concordância nos sintagmas formados por Numeral + N (seguido ou não de adjetivo) somente para as crianças, com um número maior de sintagmas não redundantes para os casos com numerais terminados em –S, os “falsos plurais”, como pode ser visto na tabela (4), abaixo.

Tabela 4 - Distribuição dos sintagmas plurais em função dos processos de formação morfológica dos numerais (Num + N X Num-S + N) por crianças e adultos – dados naturalísticos

Grupo	Morfofonologia dos numerais	Padrões de concordância		Regressão logística	Total de ocorrências
		Concordância redundante	Concordância não redundante		
Crianças	Num+N (+Adj) Ex. <i>quatro patas</i>	53 (98%)	01 (2%)	$p < 0.049$	54
	Num-S +N (+Adj) Ex. <i>três dente (mole)</i>	57 (86%)	9 (14%)		66
Adultos	Num+N (+Adj) Ex. <i>quatro passarinhos</i>	68 (100%)	0 (0%)	---	68
	Num-S +N (+Adj) Ex. <i>duas partes (iguais)</i>	79 (92%)	07 (8%)		86

Fonte: Elaborado pela autora (2021)

Adicionalmente, foi conduzida uma análise dos dados apresentados na tabela (3) em função dos padrões de marcação de plural internos ao sintagma, visando a investigar quais padrões tendem a ser mais frequentes nas produções dos adultos e crianças que compõem os *corpora*. Para tal, replicamos os dados da tabela (3) na tabela (5), na qual dispomos os dados de maneira mais detalhada no que concerne às proporções de cada padrão específico produzido. Na tabela a seguir, o símbolo “S” indica a marcação morfofonologicamente realizada de plural, enquanto o símbolo “Ø” expressa a omissão da marca em um determinado item e “Num” indica a

presença de um numeral. Por exemplo, num sintagma do tipo D+N+Adj, a indicação S $\emptyset\emptyset$ corresponde a uma estrutura na qual apenas D foi morfologicamente marcado para o plural (Ex. *As mecha \emptyset rosa \emptyset*). De maneira semelhante, num sintagma do tipo Num+N+Adj, a indicação NumSS corresponde à estrutura em que o N e Adj foram morfologicamente marcados para o plural (Ex. *Duas partes iguais*).

Tabela 5 - Distribuição dos sintagmas plurais em função dos padrões de marcação internos ao sintagma e dos tipos de estrutura analisados em cada *corpus*

Grupo Crianças	Tipo de estrutura	Padrões de concordância		Total de ocorrências
<i>Corpus</i>		Concordância redundante	Concordância não redundante	
Corpus A Crianças zona urbana	D+N	SS 104 (89%)	SØ 13 (11%)	117
	D+N+A	SSS 04 (100%)	0 (0%)	04
	Num+N	NumS 74 (94%)	NumØ 5 (6%)	79
	Num+N+A	0 (0%)	0 (0%)	0
Corpus B Crianças zona rural	D+N	SS 05 (45%)	SØ 06 (55%)	11
	D+N+A	0 (0%)	SØØ 1 (100%)	01
	Num+N	NumS 06 (75%)	NumØ 02 (25%)	08
	Num+N+A	0 (0%)	0 (0%)	0
Corpus C Crianças zona urbana	D+N	SS 42 (86%)	SØ 07 (14%)	49
	D+N+A	0 (0%)	SSØ 1(50%) SØS 1(50%)	02
	Num+N	NumS 30 (97%)	NumØ 01 (3%)	31
	Num+N+A	0 (0%)	NumØØ 1(50%) NumØS 1(50%)	02
Grupo Adultos	Tipo de estrutura	Padrões de concordância		Total de ocorrências
<i>Corpus</i>		Concordância redundante	Concordância não redundante	
Adultos Corpus A Ensino Superior	D+N	SS 132 (87%)	SØ 20 (13%)	152
	D+N+A	SSS 09 (100%)	0 (0%)	09
	Num+N	NumS 107 (99%)	NumØ 01 (1%)	108
	Num+N+A	NumSS 02 (100%)	0 (0%)	02
Adulto Corpus B – prof.	D+N	---	---	---
	D+N+A	---	---	---
	Num+N	---	---	---
	Num+N+A	---	---	---
Adultos Corpus C Ensino Médio	D+N	SS 66 (70%)	SØ 28 (30%)	94
	D+N+A	SSS 02 (67%)	SSØ 01 (33%)	03
	Num+N	NumS 37 (90%)	NumØ 04 (10%)	41
	Num+N+A	NumSS 01 (33%)	NumØØ 02 (67%)	03

Fonte: Elaborado pela autora (2021)

Como se pode ver na tabela acima, quando considerados os sintagmas curtos (formados por dois elementos), observamos um contraste entre marcação totalmente redundante (SS) e marcação mínima ($S\emptyset$), ou seja, os padrões considerados gramaticais na língua, com marcação em todos os elementos que compõem o sintagma, ou somente no determinante, tanto nos dados das crianças, quanto dos adultos. Além disso, a análise dos dados dos adultos nos fornece indícios da existência de uma correlação positiva entre grau de escolarização e a alternância das formas de concordância, visto que os cuidadores do *corpus* A (com ensino superior) apresentaram maior produção de sintagmas plurais com marcação redundante em comparação com os registros das mães do *corpus* C (escolaridade média).

Já em relação aos sintagmas com maior de número de elementos morfofonologicamente realizados, são os dados das crianças, que nos proporcionam resultados compatíveis com a ideia de que existe uma correlação positiva entre grau de escolarização e origem do falante e a alternância das formas de concordância, uma vez que as crianças do *corpus* A (residentes da zona urbana, cujas mães possuem ensino superior) produziram o padrão redundante de forma categórica, diferente do que notamos nos dados do *corpus* B (residentes da zona rural) e do *corpus* C (residentes da zona urbana, mães com ensino médio) que, por sua vez, só apresentaram produções com padrões não redundantes. A única ocorrência registrada no *corpus* B apresentava marcação mínima ($S\emptyset\emptyset$), ou seja, com marcação explícita somente no determinante. Já nos dados do *corpus* C, observamos uma alternância entre os padrões não redundantes ($SS\emptyset$ e $S\emptyset S$). Contudo, importante destacar que em todos esses padrões não redundantes, o elemento que ocupa a primeira posição do sintagma foi marcado de forma categórica, em 100% das ocorrências, reforçando a visão mais difundida na literatura de que a primeira posição do sintagma tende ser marcada explicitamente no plural de forma sistemática.

Cumpramos destacar, no entanto, que foi registrado um número muito reduzido de sintagmas maiores (formados por três elementos), tanto nos dados infantis quanto adultos. Essa e outras limitações para a condução de análises mais detalhadas, em função do escasso número (ou da ausência) de determinados tipos de ocorrência nos *corpora*, reforçaram a relevância de buscarmos outras fontes de informação. Nesse sentido, no âmbito da presente pesquisa, buscamos aliar os registros naturalísticos a dados experimentalmente obtidos.

Para finalizar a análise dos *corpora* aqui considerados, e antes de reportar os experimentos conduzidos, gostaríamos de trazer algumas considerações a partir de dados específicos coletados. Alguns dos dados registrados merecem nossa atenção por sugerirem que as crianças informantes – até mesmo as mais novas, com 3 anos de idade – produziram um conjunto de formas ausentes no *input* recebido e que poderiam ser vistas como agramaticais do ponto de vista da gramática adulta, mas que trazem fortes indícios do (re)conhecimento precoce do traço formal de número, principalmente no concernente ao uso de morfema –S como indicador de plural na língua em aquisição, como pode ser observados nos exemplos (5-8) a seguir:

(5) Concordância formal não redundante mas com concordância ideológica:

Ex: “A gente tá junta**S**.” (Criança D, *Corpus A*)

“A gente vi ficar sozinho**S**”. (Criança A, *Corpus A*)

(6) Numeral com marca de plural:

Ex: “os quatro**S**” (Criança D, *Corpus A*)

“Dos setes girassóis do Van Gogh” (Criança F, *Corpus C*)

(7) Concordância agramatical no singular, indicando já reconhecer o morfema –S como marca de plural:

Ex: “Um lápi” [em vez de um lápis](Criança E, *Corpus C*)

(8) Outros exemplos:

Ex: “Eu achei dois ôvos” (Criança F, *Corpus C*)

“ OØ meuØjogos”

Tendo-se em vista os resultados das análises conduzidas a partir dos *corpora* selecionados, podemos concluir que os dados de fala espontânea indicam uma correlação entre a produção das crianças e o *input* fornecido pelos seus cuidadores: em outras palavras, a variação encontrada na produção adulta também é observada na produção infantil. Os dados também se mostram compatíveis com a atuação de fatores de natureza sociolinguística – tais como nível de escolaridade e procedência geográfica do falante – na alternância dos padrões de concordância. Embora a

variação na marcação tenha sido observada em todos os grupos considerados, a concordância redundante ocorreu de forma mais sistemática entre os falantes adultos com ensino superior e residentes na zona urbana – e, no caso das crianças, quando as mães possuíam um nível de escolaridade maior. Embora não tenha sido registrado um efeito estatisticamente significativo, e com as devidas ressalvas em função do número não uniforme de ocorrências de nomes mais e menos salientes, os dados dos *corpora* A e C sugerem que haveria alguma diferença em termos da realização da concordância em função das propriedades fônicas dos itens nominais. Não foi possível, no entanto, estender essa afirmação para os dados do *corpus* B, em virtude da escassa ocorrência de itens de itens nominais comumente caracterizados como mais salientes.

Vale lembrar ainda que são inegáveis as contribuições das pesquisas conduzidas no campo da sociolinguística na investigação das regras de concordância variável de número no PB. Entretanto, tais estudos apresentam algumas limitações derivadas da natureza da própria metodologia empregada, que se baseia, fundamentalmente, em dados de produção espontânea. Dentre as limitações observadas na nossa própria investigação, podemos citar os casos em que não se é possível chegar a uma conclusão precisa, em virtude do número reduzido de ocorrências de uma dada estrutura. Os dados coletados nas três amostras ora investigadas reforçam tal ideia, como por exemplo, no que diz respeito ao (proporcionalmente) pequeno número de sintagmas plurais produzidos, uma vez que, o percentual de sintagmas plurais representou em média, apenas 12% de todos os sintagmas nominais com algum tipo de concordância identificável, tanto para as crianças quanto para os adultos. Além disso, como foi apontado ao longo das análises, determinadas configurações ocorreram muito raramente nos *corpora* coletados. Outra questão importante observada nos dados naturalísticos é o fato de que itens usualmente tidos como menos salientes (o chamado “plural regular”, caracterizado em função dos processos morfofonológicos de formação de plural) parecem ser os mais frequentes na produção espontânea. Nesse sentido, pelo fato de não contarmos com uma amostra suficientemente diversificada quanto aos tipos de nomes em função da SF, não foi possível traçar um quadro detalhado envolvendo os diferentes níveis e dimensões de SF delimitados na literatura.

Com isso, no intuito de buscar superar essas e outras restrições e lembrando que as diferentes dimensões da SF não têm sido igualmente investigadas na literatura,

foi desenvolvida uma atividade utilizada para a condução de dois experimentos com crianças de 6 anos – no início da alfabetização – e um grupo de controle formado por adultos. O objetivo geral desses experimentos foi explorar as duas dimensões da SF menos investigadas até o momento, quais sejam, *padrão de acentuação* e *número de sílabas* dos nomes e ampliar resultados previamente obtidos no que tange à dimensão *processos morfológicos de formação de plural* (AZALIM, 2016; AZALIM et al, 2018). No próximo capítulo, apresentamos a pesquisa experimental conduzida no âmbito desta tese.

4 CONCORDÂNCIA NOMINAL NA PRODUÇÃO INFANTIL: ABORDAGEM EXPERIMENTAL

Como vimos ao longo dos capítulos anteriores, o fenômeno da concordância variável no PB tem sido tema de numerosos estudos, fundamentalmente no âmbito da sociolinguística variacionista (MENDES; OUSHIRO, 2015). No entanto, nota-se que poucos trabalhos se dedicam a analisar a produção infantil (GOMES et al, 2011). Em número mais reduzido ainda, encontram-se disponíveis na literatura alguns poucos trabalhos investigando o fenômeno da concordância variável no âmbito do DP na aquisição da linguagem a partir de uma abordagem experimental, seja a partir de dados de compreensão (CASTRO; FERRARI NETO, 2007; CORRÊA; AUGUSTO; FERRARI NETO, 2005; FERRARI NETO, 2003, 2008; REIS, 2020), bem como de produção (FERRARI NETO, 2003; ROZA; CASAGRANDE, 2015; REIS, 2020). Embora tais estudos com foco na aquisição apresentem contribuições relevantes para as pesquisas concernentes ao fenômeno da concordância variável no PB, ainda há muito o que se investigar. Nesse sentido, no presente capítulo reportamos os resultados de um estudo experimental, conduzido por meio da técnica de produção semieliciada a partir de imagens, cujo objetivo principal foi avaliar a produção da concordância nominal variável de número por crianças adquirindo o PB. Mais especificamente, buscamos analisar os possíveis efeitos das três dimensões da variável SF – *padrão de acentuação e número de sílabas* no Experimento 1 e *processos morfofonológicos de formação do plural* no Experimento 2 – na produção da marcação de plural no âmbito nominal.

Antes de passarmos para a descrição do estudo experimental propriamente dito, apresentaremos algumas considerações importantes a respeito do método experimental de pesquisa, tais como o contraste entre tarefas do tipo *on-line* e *off-line*, além de apresentar a técnica experimental utilizada no âmbito desta tese.

4.1 O MÉTODO EXPERIMENTAL

A psicolinguística é uma área de estudos que, atualmente³⁶, pode ser localizada na interseção entre psicologia cognitiva e linguística, e cujo interesse central consiste

³⁶ Vale destacar que, na sua origem, a psicolinguística surgiu e se desenvolveu – a partir da segunda metade do século XX – no contexto da psicologia cognitiva, de forma autônoma e independente da teoria linguística.

em investigar como o conhecimento linguístico é adquirido, representado na memória e utilizado pelos falantes. Em outras palavras, trata-se do estudo de processos mentais envolvidos na produção e compreensão, bem como na aquisição da linguagem. Para atingir tal objetivo, a psicolinguística utiliza-se de métodos e técnicas experimentais herdados da psicologia cognitiva como instrumento de pesquisa.

No presente trabalho, especificamente, em virtude de visarmos a estabelecer um diálogo entre sociolinguística e psicolinguística para o estudo da concordância nominal variável, buscamos articular dados naturalísticos e dados obtidos com base em metodologia experimental. A proposta para tal articulação se deu em função de compartilharmos com Eisenbeiss (2010) a ideia de que, apesar das diversas vantagens apresentadas pela abordagem naturalística, tais como sua versatilidade e validade ecológica, bem como o fato de fornecer um excelente ponto de partida para pesquisas sobre os mais variados fenômenos, a falta de controle do pesquisador sobre aquilo que é efetivamente produzido pelos informantes pode acarretar na coleta de amostras incomparáveis entre si, assim como dificultar o estudo de fenômenos que ocorrem com uma baixa frequência, dentre outros. Nesse sentido, a escolha pela utilização da abordagem experimental justifica-se por configurar-se como uma ferramenta eficiente, no sentido de permitir uma manipulação sistemática de uma ou mais variáveis no intuito de verificar em que medida as mesmas influenciam no comportamento linguístico dos falantes. Além disso, consideramos que, aliados aos dados naturalísticos, os resultados experimentais possam fornecer evidências mais robustas a respeito do nosso objeto de estudo.

Ao optar pelo uso da abordagem experimental, o pesquisador precisa estar atento a determinados detalhes para atingir os objetivos almejados, a começar pela escolha da técnica experimental mais apropriada para ajudá-lo a encontrar respostas relevantes para as questões levantadas em seu estudo. Sendo assim, ele tem à sua disposição uma gama de paradigmas e técnicas experimentais que podem ser utilizadas isoladamente ou combinadas. Como exemplo, podemos citar os paradigmas de produção, compreensão ou julgamento, dentre os mais utilizados.

Como já adiantado, a psicolinguística objetiva compreender os processos mentais que subjazem à aquisição, à compreensão e à produção da linguagem. No entanto, em virtude da complexidade que envolve a caracterização do que sejam e de como ocorrem os processos mentais que subjazem ao processamento linguístico – principalmente, a inacessibilidade aos processos mentais *in loco* – a psicolinguística

emprega diferentes procedimentos para a coleta de dados e o faz utilizando-se de tarefas experimentais que fornecem informações que podem ser de natureza *off-line* ou *on-line*.

Por experimento *off-line* entende-se toda tarefa que fornece dados relativos às reações dos participantes na interpretação dos estímulos – sejam escritos ou acústicos – após já ter havido uma integração entre os níveis linguísticos envolvidos (fonológico, morfológico, semântico, sintático, pragmático). Em outras palavras, tarefas *off-line* fornecem dados após o processamento linguístico já ter acontecido. As chamadas tarefas *on-line*, por sua vez, permitem um acesso – ainda que indireto – aos processos mentais durante o curso do processamento linguístico.

No que tange à pesquisa experimental conduzida no âmbito desta tese, optou-se pela técnica *off-line* de produção semieliciada a partir de imagens – descrita mais detalhadamente mais adiante –, por ser esta adequada em função dos objetivos traçados – no sentido de preservar uma maior naturalidade da situação comunicativa – e pelo fato de possibilitar a coleta de um número significativo de sintagmas nominais plurais.

Além do paradigma e/ou técnica experimental escolhido, outro aspecto relevante a ser considerado pelo pesquisador na fase de elaboração de um experimento diz respeito à definição das variáveis independentes – relativas aos fatores investigados – e dependentes – isto é, relativas às respostas comportamentais e/ou neurofisiológicas que podem ser tomadas como indicativas da demanda cognitiva envolvida na realização da tarefa experimental pelos participantes. No caso do estudo aqui reportado, no qual utilizou-se a técnica da produção semieliciada, tomamos como variáveis independentes o *número de sílabas no nome alvo* a ser produzido (duas/ três sílabas) e o *padrão de acentuação* (nomes paroxítonos/ oxítonos) no Experimento 1 e o *tipo de processo morfofonológico de formação de plural* (formação do plural de nomes terminados ortograficamente em –R/ formação do plural de nomes terminados ortograficamente em –L/ formação do plural dos nomes regulares) no Experimento 2. Em ambos os casos, as variáveis independentes delimitadas foram investigadas com base na variável dependente *número de sintagmas determinantes/nominais plurais redundantes e não redundantes produzidos* em cada condição.

Como vimos no capítulo anterior, os dados naturalísticos mostram que nomes do chamado plural regular apresentam alternância na marcação: ora aparecem com

marca morfológica, ora sem. No entanto, faz-se necessário esclarecer que esses nomes são tidos como menos salientes em função dos processos morfofonológicos envolvidos, mas não necessariamente são semelhantes em termos das outras duas dimensões. Nesse sentido, as duas dimensões da SF exploradas no experimento 1 dizem respeito a nomes menos salientes em termos da dimensão *processos*, mas que apresentam diferenças em função de sílabas e acentuação. Afinal, entendemos que seja relevante aprofundarmos a compreensão acerca do denominado “plural regular” que, por sua vez, inclui um conjunto expressivo de possibilidades que usualmente não são discriminadas na literatura. Além disso, como também pôde ser observado nos dados naturalísticos, o plural regular representa o tipo de marcação mais frequente na língua e – conseqüentemente – no *input* recebido pelas crianças, fato que reforça a importância de uma melhor caracterização das possíveis diferenças existentes entre os itens dessa classe.

Além disso, é importante lembrar que as escalas de SF disponíveis na literatura estão baseadas nos *processos morfológicos de formação de plural*, sendo que esse aspecto configura-se como apenas uma das dimensões da SF levantadas na literatura sociolinguística. Assim sendo, em conjunto, os experimentos 1 e 2 buscaram obter dados em quantidade expressiva para permitir analisar os possíveis efeitos das 3 dimensões comumente associadas à SF.

Por termos optado pelo estudo da produção infantil, procuramos considerar aspectos metodológicos que mais se adequassem ao paradigma escolhido – técnica de produção semielicada –, e ao público alvo – i.e. crianças adquirindo o PB. Nesse sentido, buscamos: (i) aliar o nível de dificuldade da tarefa à idade dos participantes – nem tão fácil para deixá-los entediados, nem tão difícil que pudesse desestimulá-los; (ii) pensar numa tarefa que fosse prazerosa para as crianças a ponto de manter o interesse delas em continuar a participar da tarefa, dentre outros, que serão explicitados na seção 4.2.1, que apresenta o procedimento experimental empregado.

Os dados experimentalmente obtidos foram gravados em áudio, transcritos e tabulados para posterior análise por meio de testes de inferência estatística escolhidos em função das medidas específicas consideradas, bem como do *design* experimental. Faz-se necessário ressaltar que o nível de significância estipulado para todas as análises a serem conduzidas foi de $p \leq .05$, tal como estabelecido em pesquisas da área das ciências cognitivas.

Partindo do que foi exposto até aqui, passemos agora à seção em que será introduzida a técnica utilizada neste estudo. Em seguida, reportaremos os dois experimentos conduzidos.

4.1.1 Produção eliciada

Segundo Eisenbeiss (2010), a coleta e análise de dados naturalísticos têm desempenhado um papel crucial na pesquisa em aquisição da linguagem. No entanto, longe de consistir em um método perfeito, ele apresenta determinadas limitações, acarretando, muitas vezes, a dificuldade em se atingirem resultados mais precisos, mais acurados acerca do fenômeno sob investigação. É no intuito de sanar tais limitações que, desde os anos de 1950, a pesquisa em aquisição tem ganhado reforços com a implementação da abordagem experimental para coleta e análise de dados de produção como uma forma de complementar os achados obtidos com base no método naturalístico. Eisenbeiss (2010) apresenta e discute os que, em sua opinião, configuram-se como os três métodos mais relevantes para a coleta de dados de produção linguística, a saber: (i) o método naturalístico; (ii) experimentos de produção e (iii) produção semicontrolada. Em virtude de já termos apresentado o método naturalístico anteriormente, nesta seção, iremos nos ater à descrição de alguns dos mais difundidos métodos experimentais usados no estudo de produção linguística.

O uso de métodos experimentais de produção viabiliza ao pesquisador uma manipulação sistemática de variáveis que se consegue a partir do uso de procedimentos e estímulos padronizados. Entre as vantagens do uso de tais procedimentos e estímulos, podemos citar: (i) evitar prover acidentalmente as crianças com modelos linguísticos ou *feedback* que poderiam influenciar o seu comportamento, bem como (ii) fornecer um conjunto maior de informações que permitam ao pesquisador determinar as intenções comunicativas das crianças, e ainda (iii) facilitar a posterior interpretação dos dados. Segundo Ambridge e Rowland (2013), os métodos experimentais de produção variam num *continuum* desde métodos de produção mais abrangentes até outros mais específicos e estruturados, a depender do fenômeno que se almeja investigar. Para ilustrar, retomamos os exemplos disponibilizados pelos referidos autores que variam desde contextos experimentais em que os participantes são expostos a uma animação, imagem, ou vídeo e, em seguida, devem responder a perguntas neutras – mais abrangentes – como por exemplo, ‘O que

está acontecendo?’, até situações em que o pesquisador almeja obter um controle maior sobre as construções linguísticas que o participante poderia usar na sua resposta, tal como ocorre em estudos sobre o tempo pretérito – no caso, Passado Simples do inglês - , em que o experimentador elicia a produção por meio da utilização de um roteiro do tipo ilustrado em (9):

(9) *The bunny likes to run. Look, there he is running. Every day he runs. So yesterday he...*

(O coelho gosta de correr. Olhe, ali ele está correndo. Todo dia ele corre. Então ontem ele...)

Além desses, podemos mencionar ainda o paradigma da repetição – ou imitação eliciada – usado em substituição à produção eliciada quando há necessidade de se restringir o enunciado alvo de maneira ainda mais precisa, como nos casos em que se espera que a criança produza um determinado pronome em detrimento a um DP pleno. Tal paradigma foi originalmente proposto por Slobin e Welsh (1973) como uma ferramenta que viabilizaria ao pesquisador aprofundar sua compreensão acerca do processo de aquisição da linguagem. Neste tipo de tarefa, os participantes têm que repetir/imitar os estímulos experimentais e, posteriormente, o pesquisador compara a produção aos estímulos aos quais os participantes foram expostos. De acordo com Eisenbeiss (2010), tal comparação pode indicar um caminho para se compreender o conhecimento gramatical da criança, uma vez que, por exemplo, em casos em que a criança é exposta a estímulos longos, ela não seria capaz de simplesmente memorizá-lo por completo e realizar a repetição, sendo, então, necessário empregar seu conhecimento gramatical para reformulá-los e produzi-los. (cf. AZALIM, 2016; AZALIM et al, 2018; HENRIQUE, 2016; SLOBIN; WELSH, 1973; para exemplos de estudos utilizando tal técnica).

Já em Eisenbeiss (2009), defende-se que embora tanto o método naturalístico quanto o experimental de produção tenham sido amplamente utilizados em pesquisas sobre aquisição, um terceiro método – a produção semicontrolada – tem ganhado progressivamente espaço e apresentado benefícios para a área. Em outros termos, entre a completa liberdade de expressão observada nos estudos naturalísticos e o controle sistemático de procedimentos e variáveis na imitação eliciada, encontramos as técnicas de produção eliciada semicontrolada, em que se mantém a situação

comunicativa o mais natural possível – como acontece nos estudos naturalísticos –, mas também são utilizadas técnicas de entrevistas, vídeos ou jogos para eliciar a produção de amostras significativamente ricas e comparáveis entre si, além de proporcionar um maior controle do pesquisador sobre as situações registradas.

Além disso, diferentemente do que ocorre nos estudos naturalísticos, em que a interferência do investigador é evitada ao máximo, na produção semicontrolada, os pesquisadores criam contextos semicontrolados com funções comunicativas para viabilizar a produção de dados linguísticos relevantes. Tais técnicas, segundo Eisenbeiss (2009), podem ser usadas para a investigação da maioria dos fenômenos que são examinados em pesquisas de cunho naturalístico, mas são particularmente úteis para se estudar fenômenos pouco frequentes na fala espontânea ou ainda para estudos translinguísticos acerca de aspectos semânticos e morfossintáticos.

Em contrapartida, a autora defende que os métodos experimentais para o estudo da produção não são recomendados quando se almeja investigar dados de *input* nem tampouco frequência. Afinal, neste tipo de método, tanto as amostras de *input* quanto a frequência de distribuição seja de palavras, formas ou mesmo construções observadas na produção infantil não costumam ser muito expressivas. Com isso, podemos compreender que os métodos experimentais de produção não substituem o método naturalístico. Ao contrário, melhores resultados e análises podem ser obtidos a partir de uma articulação entre ambos. Tendo isso em mente, a presente tese busca aliar ambas as abordagens no intuito de conseguir traçar um quadro detalhado do fenômeno investigado.

Eisenbeiss (2009) destaca três tipos básicos de tarefas semicontroladas, a depender do objetivo pretendido com o estudo, a saber: (i) técnicas de amplo espectro (*broad-spectrum*); (ii) técnicas focadas na produção do sentido (*meaning-focused elicitation techniques*); e (iii) técnicas focadas na produção da forma (*form-focused elicitation techniques*). O primeiro tipo é geralmente usado para se obterem amostras grandes o suficiente para viabilizar ao pesquisador uma comparação entre a produção de falantes de diferentes idades, *backgrounds* linguísticos e culturais. De acordo com Eisenbeiss (2009), um dos exemplos mais famosos de ferramentas desse tipo usadas na pesquisa sobre a linguagem infantil seria o livro *Frog, where are you?* (MAYER, 1969), que conta por meio de imagens a história de um garoto que encontra um sapo, o leva para casa e faz dele seu animal de estimação.

Com o segundo tipo de tarefas semicontroladas arrolado acima, objetiva-se investigar um domínio semântico em particular no qual construções diferentes são usadas seja no âmbito de uma mesma língua ou entre línguas. Já o terceiro tipo de tarefas, tem foco na forma, isto é, objetiva investigar formas ou construções específicas, em particular aquelas que não costumam ocorrer com muita frequência na produção cotidiana. Nesse sentido, a autora defende que jogos são excelentes ferramentas para eliciar uma produção significativa de tais estruturas. Como veremos, a tarefa utilizada no estudo experimental aqui reportado, corresponde ao terceiro tipo arrolado.

Importante destacar que, independentemente do tipo de tarefa escolhida, a produção infantil tende a ser mais satisfatória quando a atividade apresenta um sentido comunicativo, ou seja, se a mesma fornecer à criança um motivo “real” para falar. Para atingir tal objetivo, o pesquisador pode optar por atividades de diferentes tipos: tarefas do tipo falante/ouvinte (*speaker/listener*), tarefas do tipo *director-matcher* e tarefas do tipo *co-player*.

Nas denominadas *tarefas falante/ouvinte*, os participantes são solicitados a fornecer informações para alguém que não tem acesso às mesmas, como por exemplo, relatar eventos do cotidiano ou descrever suas casas. No entanto, Eisenbeiss (2010) argumenta que tal modalidade pode apresentar problemas, no sentido de que a falta de estímulos pode dificultar a determinação dos significados pretendidos, bem como uma baixa produção em falantes que apresentem habilidades cognitivas ou linguísticas limitadas. Por essa razão, muitas tarefas do tipo *falante/ouvinte* envolvem estímulos visuais ou verbais, tais como histórias para serem recontadas, descrições de imagens ou vídeos, entre outros.

Nas tarefas do tipo *director-matcher*, a informação não é simplesmente transmitida passivamente ao ouvinte, mas os “diretores” conduzem os outros participantes de maneira que estes consigam encontrar um determinado estímulo dentro de um conjunto ou seguir as instruções dos diretores. Por exemplo, os participantes têm que seguir descrições de rotas através de cenários de brinquedo, etc.

Já nas tarefas do tipo *co-player*, os participantes devem trocar informações e coordenar ações para atingirem objetivos, seja em atividades cotidianas, como cozinhar, ou mesmo em jogos. Esse último tipo de tarefa, permite explorar a produção infantil de maneira interativa, ou seja, a troca de informações para se atingir objetivos e foi a escolhida na elaboração da tarefa experimental empregada no presente estudo.

Importante destacar que as técnicas de produção eliciada têm como finalidade favorecer produções referentes a determinados aspectos de eventos, tais como, ações, direções, localizações, bem como participantes de eventos, objetos e suas propriedades, entre outros. De acordo com Eisenbeiss (2009), falantes de todas as idades são mais suscetíveis a descreverem tais aspectos de eventos quando são expostos a contrastes.

Seguindo tal raciocínio, nos experimentos descritos neste capítulo, utilizou-se a técnica da produção eliciada a partir de imagens por pensarmos que esta seria um excelente recurso para se avaliar o comportamento linguístico de crianças adquirindo o PB frente aos dois padrões gerais de concordância de número – redundante e não redundante. Além disso, visamos a garantir que a tarefa tivesse um caráter comunicativo por compartilharmos com Ambridge e Rowland (2013) a ideia de que as crianças dificilmente respondem adequadamente questões as quais considerem desnecessárias ou sem sentido. Para tal, desenvolvemos um contexto de teste na forma de um jogo em que a criança pudesse interagir com um outro participante da mesma faixa etária. Nesta variação do paradigma da produção eliciada, elaboramos uma adaptação do jogo *Cara a Cara* – que consiste em uma atividade interativa em duplas na qual os participantes devem identificar, por meio de perguntas SIM/NÃO uma carta específica dentro de um conjunto de imagens semelhantes. No caso da tarefa desenvolvida, a identificação da carta alvo depende crucialmente da produção de sintagmas plurais contendo numerais e adjetivos, além do nome em questão.

Após apresentar as características gerais do paradigma experimental escolhido, a seguir, reportamos de forma detalhada os experimentos conduzidos.

4.2 EXPERIMENTO 1: INVESTIGANDO A SALIÊNCIA FÔNICA NA PRODUÇÃO DA CONCORDÂNCIA NOMINAL VARIÁVEL – PADRÃO DE ACENTUAÇÃO E NÚMERO DE SÍLABAS

No Capítulo 2, ficou estabelecido que, dentre os fatores linguísticos arrolados como condicionadores do fenômeno em estudo, interessa-nos em particular o princípio da SF, segundo o qual, formas fonologicamente mais salientes são mais perceptíveis acusticamente e, conseqüentemente, mais propensas a serem marcadas no plural do que formas menos salientes.

Embora, de acordo com a visão mais difundida na literatura sociolinguística, a SF teria um papel de destaque no favorecimento/desfavorecimento da marcação redundante, tanto no âmbito verbal quanto nominal, essa caracterização não é consensual. A literatura traz pontos de vista divididos e as divergências encontradas recobrem vários aspectos, a começar pelo fato de que a própria caracterização da SF não é unânime entre os estudiosos (BRAGA, 1977; GUY, 1981; SCHERRE, 1988). Justifica-se assim a relevância de se investigar tal variável de maneira mais aprofundada, levando em consideração as diferentes dimensões a ela comumente associadas. Especificamente, no caso da nossa própria pesquisa, além de buscarmos uma compreensão mais refinada do fenômeno por meio da articulação de dados naturalísticos e experimentais, um diferencial importante diz respeito ao fato de analisarmos tanto a produção adulta quanto infantil.

O experimento que reportamos a seguir teve como objetivo principal investigar os possíveis efeitos de duas das dimensões específicas associadas à SF na realização da concordância nominal variável por crianças adquirindo o PB, quais sejam: o padrão de acentuação e o número de sílabas do item nominal. Vale ressaltar que a escolha de tais dimensões toma como base a proposta de Scherre (1988), segundo a qual o fator SF deve ser investigado levando em conta não somente os processos morfofonológicos de formação de plural, como também a tonicidade do item no singular e o número de sílabas dos itens lexicais na forma singular. Nesse sentido, no Experimento 1 buscamos investigar possíveis contrastes no interior da classe de nomes tidos como [-salientes] pela literatura sociolinguística no que tange aos processos de formação de plural, ou seja, nomes correspondentes ao denominado “plural regular”. Além disso, visamos a desenvolver uma atividade experimental que permitisse coletar um número expressivo de dados relativos à produção de sintagmas nominais plurais de modo a viabilizar uma análise mais acurada de alguns dos padrões previamente investigados na produção espontânea. Tendo-se em vista tais objetivos, adaptamos uma tarefa experimental cuja versão inicial foi elaborada por uma equipe do grupo de pesquisa do NEALP com base na técnica de produção eliciada descrita na seção anterior e conduzida em uma versão piloto com participantes adultos. Após essa experiência inicial com adultos, a tarefa foi ajustada de modo a tornar a atividade mais aprazível para ser realizada pelas crianças.

A atividade experimental teve como objetivos específicos primários:

- (i) Fornecer dados relativos à produção de sintagmas nominais plurais por crianças adquirindo o PB e falantes adultos;
- (ii) Investigar experimentalmente em que medida as crianças e adultos se mostram sensíveis às dimensões *padrão de acentuação* e *número de sílabas* no que concerne a nomes tidos como menos salientes (em termos dos processos morfofonológicos de formação de plural) pela literatura sociolinguística, no intuito de obtermos subsídios para propor um refinamento da caracterização da variável SF;
- (iii) Comparar os resultados experimentalmente obtidos àqueles levantados a partir da análise da produção espontânea de crianças e adultos, bem como àqueles veiculados por pesquisas prévias.

Além disso, tivemos como objetivos secundários:

- (iii) Coletar dados que permitissem avaliar outros pontos relevantes – tais como a configuração estrutural do sintagma no qual os nomes ocorrem e os padrões de marcação internos ao sintagma –, no intuito obtermos subsídios adicionais para embasar uma formalização do fenômeno da concordância nominal variável no PB (cf. Capítulo 5);
- (iv) Investigar possíveis efeitos da constituição morfofonológica dos numerais na realização dos padrões de concordância, no intuito de verificar se a terminação em -S, encontrada em numerais como *dois*, *três* e *seis*, por exemplo, poderia ser vista como resultado de uma possível falsa analogia entre tal marca e o morfema de plural -S dos nomes.

4.2.1 Método

Em função das variáveis independentes delimitadas para o Experimento 1 (*número de sílabas no nome alvo* e *padrão de acentuação*), foi concebida uma tarefa de produção semielicada por imagens. No concernente à variável *número de sílabas no nome alvo*, foram considerados nomes dissílabos e trissílabos. No que diz respeito à variável *padrão de acentuação*, estabeleceu-se o contraste entre nomes oxítonos e paroxítonos. Dadas essas variáveis e seus correspondentes níveis, temos um experimento com *design* fatorial 2X2 – *número de sílabas no nome* X *padrão de*

acentuação, perfazendo, assim, um total de quatro condições experimentais, cujos exemplos encontram-se dispostos no quadro (3) a seguir.

Como variável dependente, foi considerado o número de sintagmas plurais redundantes e não redundantes produzidos pelos participantes em cada condição.

Quadro 3 - Condições experimentais do Experimento 1.

Condição 1: nome dissílabo oxítono
maçã, batom, gambá, sofá, boné
Condição 2: nome dissílabo paroxítono
livro, gato, garfo, cobra, chave
Condição 3: nome trissílabo oxítono
jacaré, urubu, picolé, sutiã, canguru
Condição 4: nome trissílabo paroxítono
estrela, sapato, panela, cachorro, banana

Fonte: Elaborado pela autora (2021)

Como ponto de partida para nossa pesquisa experimental, com base no observado nos dados naturalísticos – quando considerados os valores absolutos de ocorrência de nomes mais e menos salientes – crianças adquirindo o PB, em linhas gerais, parecem refletir o comportamento linguístico adulto mais difundido na literatura sociolinguística e que se mostra compatível com a atuação do princípio da SF³⁷. Lembramos, no entanto, que – além de não ter sido observada diferença estatisticamente significativa entre as ocorrências relevantes – também não foi possível conduzir uma análise refinada e completa do fenômeno a partir dos dados de produção espontânea, pelo fato de os mesmos não fornecerem uma amostra suficientemente diversificada das diversas configurações possíveis. Assim sendo, as observações realizadas com base nos dados espontâneos ficaram restritas ao contraste mais simples entre nomes tidos como + e – salientes (i.e. plurais regulares).

³⁷Nos dados naturalísticos, os percentuais absolutos de concordância não redundante foram maiores quando associados a nomes menos salientes. Contudo, os modelos de regressão logística conduzidos não indicaram diferença significativa entre as proporções comparadas (concordância não redundante com nomes mais salientes x concordância não redundante com itens menos salientes).

Assumindo como pressuposto a efetiva atuação do princípio da SF, tal como defendido por boa parte da literatura sociolinguística – embora não corroborado de forma consistente pelos nossos dados naturalísticos – quanto às duas dimensões da SF investigadas no Experimento 1, a hipótese que se levanta em relação ao *padrão de acentuação* é a de que – pelo fato de a sílaba acentuada coincidir com a marcação morfofonológica de plural – itens oxítonos favorecem a aplicação da forma redundante quando comparados com itens paroxítonos. No que diz respeito à dimensão *número de sílabas*, a hipótese é que a presença de um volume maior de material fônico nos nomes trissílabos pode favorecer a marcação redundante em comparação com nomes dissílabos. A partir dessas hipóteses e, considerando as duas dimensões de modo independente, são levantadas as seguintes previsões experimentais:

- i) No que diz respeito à variável *padrão de acentuação*, espera-se registrar um número significativamente maior de produções com marcação redundante em nomes oxítonos quando comparados aos nomes paroxítonos;
- ii) Já em relação à variável *número de sílabas*, espera-se encontrar diferença significativa entre o número de ocorrências redundantes e não redundantes a depender do número de sílabas dos nomes alvo produzidos. É esperado que itens trissílabos concentrem um número de produções com marcação redundante significativamente maior do que itens dissílabos.

No entanto, em função dos resultados discrepantes apontados na literatura quanto à possível relevância dessas dimensões da SF quando consideradas isoladamente (cf. CARVALHO, 1997; GUY, 1981; SCHERRE, 1988), aventamos aqui também a hipótese de que *número de sílabas* e *padrão de acentuação* (e conseqüentemente, também a dimensão de *processos morfológicos* que será investigada no Experimento 2) sejam fatores que atuam de forma conjunta outorgando um maior ou menor destaque perceptual a um dado item. Nesse sentido, postulamos a ideia de um *continuum* ou gradiente de SF no qual cada dimensão pode contribuir para aumentar, diminuir ou neutralizar o realce perceptual dos itens.

De acordo com essa terceira hipótese, são levantadas as seguintes previsões específicas para o Experimento 1:

- i) Nomes dissílabos paroxítonos (Ex. *livro*) devem concentrar um número de produções não redundantes significativamente maior quando comparados com os nomes trissílabos oxítonos (Ex. *picolé*). Isto é, prevê-se um contraste envolvendo a interação entre as duas variáveis independentes estabelecidas;
- ii) No caso dos nomes dissílabos oxítonos (Ex. *maçã*) e dos trissílabos paroxítonos (Ex. *estrela*), não são esperadas diferenças significativas quando comparadas ambas as condições, pelo fato de uma das dimensões envolvidas possivelmente “neutralizar” a atuação da outra;

Adicionalmente, incluímos uma previsão não diretamente relacionada às variáveis independentes delimitadas, mas relativa aos grupos a serem avaliados: em função do observado nos dados naturalísticos, não são esperadas diferenças significativas na produção de sintagmas plurais redundantes e não redundantes quando comparados os dados dos participantes adultos (grupo controle) e das crianças.

No intuito de testar todas as hipóteses e previsões experimentais estabelecidas, foram conduzidas análises específicas que serão reportadas nas seções correspondentes.

Participantes

Os participantes desta tarefa compreendem 20 crianças (sendo 10 meninos e 10 meninas), todos com 6 anos de idade completos, cursando o 1º Ano do Ensino Fundamental de duas escolas da rede privada de ensino em Juiz de Fora. Os participantes foram testados nas instalações das escolas em que estudam e participaram do experimento de forma voluntária. A faixa etária foi escolhida por dois motivos: (i) as crianças estão iniciando a fase de alfabetização que, como já mencionado, pode ser decisiva para aproximar a produção infantil do uso predominante da forma redundante de concordância; (ii) pela natureza da tarefa proposta (necessidade de contar elementos de um conjunto e identificar diferentes formas e cores), optou-se por começar a aplicação da tarefa com crianças mais velhas

(6 anos é a idade mais avançada considerada nos dados naturalísticos analisados), antes de cogitar a possibilidade de testar crianças mais novas. Como veremos adiante, a tarefa desenvolvida se mostrou bastante adequada para a faixa etária escolhida, embora possivelmente fosse complexa demais para crianças mais novas. Como grupo controle, participaram 21 adultos estudantes universitários, com idades entre 18 e 26 anos, residentes na mesma cidade.

Materiais

Para a realização da tarefa, foram utilizados dois computadores portáteis, cartões impressos em papel *couché* contendo as imagens-alvo e um gravador SONY, modelo PCM-D50 de alta resolução para captar a produção dos participantes. As imagens utilizadas para eliciar a produção de sintagmas foram projetadas em cada computador por meio do uso do programa *Power Point*.

Os estímulos experimentais consistiam de um conjunto de 20 nomes alvo (sendo 5 por condição – e 9 imagens distratoras, mas que também correspondiam a nomes que seguiam os padrões investigados, tais como *boneca, janela, lata, caju*). A produção dos nomes distratores não era obrigatória para a correta execução da tarefa, mas, eventualmente, os participantes poderiam produzir sintagmas plurais contendo esses nomes. Nesse sentido, o uso desses itens adicionais contribuiu para obtermos um volume maior de dados. Na seção de análise, esse ponto será retomado quando apresentarmos a distinção entre os resultados referentes aos dados brutos e os resultados contendo apenas os nomes alvo. A ordem de apresentação das pranchas contendo os estímulos visuais foi semialeatorizada manualmente.

Para a construção dos estímulos experimentais, foram escolhidos dez itens nominais dissílabos, sendo cinco oxítonos (*sofá, boné, gambá, batom e maçã*) e cinco paroxítonos (*livro, garfo, gato, cobra e chave*). Dentre os nomes trissílabos, cinco apresentam padrão oxítono (*canguru, urubu, jacaré, sutiã e picolé*) e cinco paroxítonos (*estrela, sapato, cachorro, panela e banana*). A escolha dos itens foi determinada tanto pela facilidade para sua representação e identificação visual, quanto pelo fato de serem itens relativamente comuns, muito provavelmente conhecidos pelas crianças.

A aplicação piloto do experimento com um grupo de adultos permitiu identificar algumas questões a serem revisadas com vista a alcançarmos mais êxito na

versão a ser apresentada para crianças. Foram introduzidas modificações nas cores das imagens para facilitar a identificação das figuras e, com o intuito de tornar a atividade prazerosa e menos cansativa para as crianças, optamos por deixar cada prancha com apenas seis conjuntos de imagens – originalmente o experimento foi construído com oito conjuntos em cada prancha. Além disso, a aplicação inicial do experimento com um grupo de adultos também permitiu ajustar o tipo de instruções a serem apresentadas para os participantes, de modo a facilitar a dinâmica do jogo e incentivar a produção do tipo de estrutura linguística que buscávamos ser eliciada.

Procedimento

A tarefa experimental foi concebida como uma adaptação do jogo *Cara a Cara*, nos moldes de uma tarefa interativa semiestruturada realizada por uma dupla de participantes, cujo objetivo seria identificar, por meio de perguntas do tipo SIM/NÃO, uma determinada carta dentro de um conjunto de imagens semelhantes dispostas em pranchas. Cada participante recebia dez pranchas (contendo seis conjuntos de imagens) e dez cartões alvo (sobre os quais deveria responder perguntas do participante oponente). As imagens impressas nas pranchas correspondiam a nomes do PB tipicamente caracterizados como menos salientes quando considerados os processos morfofonológicos de formação de plural, embora possam apresentar possíveis contrastes a depender de outras dimensões relativas à SF, quais sejam: *número de sílabas* (dissílabas vs. trissílabas) e *padrão de acentuação* (oxítone vs. paroxítone). Vale destacar que dispúnhamos das pranchas impressas em papel – tal como na versão com adultos – e também salvas no computador. Mas, no momento da realização da tarefa, julgamos que, em virtude do grande interesse das crianças por jogos eletrônicos, seria mais atrativo para elas se as pranchas fossem exibidas na tela do computador.

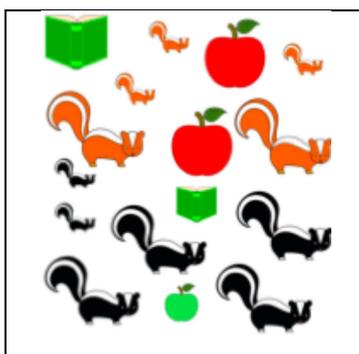
Para a realização da tarefa, os participantes foram instruídos oralmente acerca das regras do jogo, a saber:

- (i) que cada prancha correspondia a uma rodada do jogo e, para cada prancha/ rodada, os participantes receberiam um cartão-alvo sobre o qual deveriam responder perguntas feitas pelo seu oponente;
- (ii) que os participantes só poderiam fazer perguntas cujas respostas fossem “SIM” ou “NÃO”;

- (iii) que não seria permitido perguntar sobre o número total de elementos dispostos em cada conjunto (Ex. “A sua carta tem doze objetos no total?”).

Visando a eliciar a produção de sintagmas nominais plurais, os participantes eram orientados a serem o mais específicos possível em suas perguntas, utilizando todas as informações visuais disponíveis (número de elementos de cada tipo – por exemplo, *cachorro, estrela, boné* –, assim como a cor³⁸ e o tamanho). Ambos os participantes de cada dupla perguntavam e respondiam em todas as rodadas. Tendo recebido as instruções, os participantes realizavam um pequeno treinamento para se habituarem ao procedimento e terem a oportunidade de sanar quaisquer dúvidas que surgissem durante esse momento. Ao final do treinamento, os participantes iniciavam a rodada de perguntas para descobrirem qual conjunto de imagens na sua prancha correspondia ao cartão-alvo que estava com o outro participante. Para prover o leitor com uma compreensão melhor do procedimento da tarefa experimental, disponibilizamos abaixo a figura (1) que ilustra um dos cartões-alvo e a figura (2) que, por sua vez, apresenta uma das pranchas utilizadas no experimento.

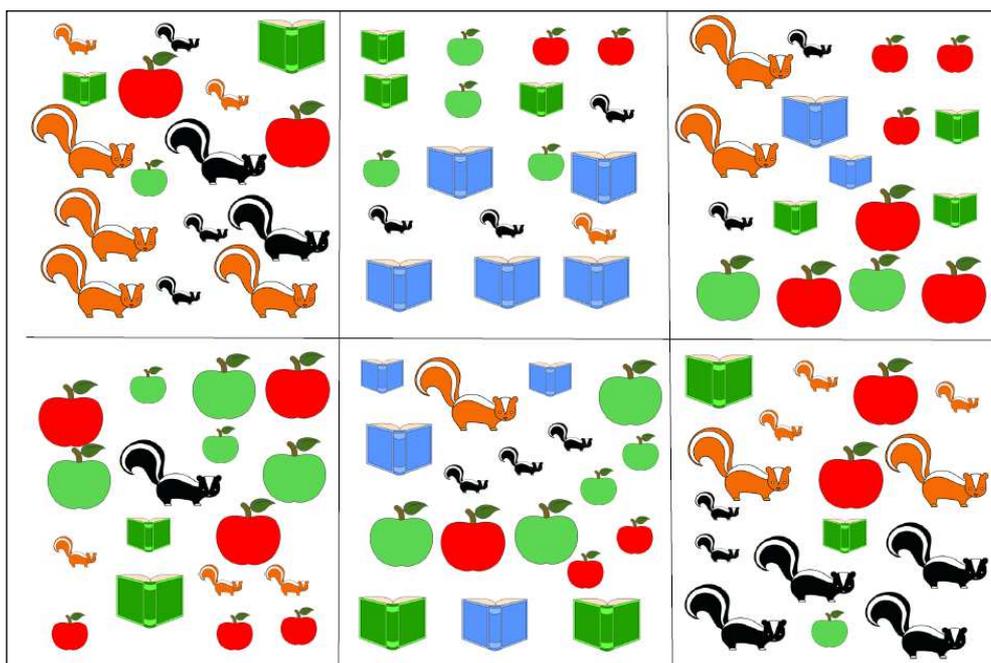
Figura 1 - Exemplo de cartão alvo utilizado no experimento – Sintagma para identificação do cartão (nome alvo gambá): Quatro gambás pretos grandes



Fonte: Elaborado pela autora (2021)

³⁸ Importante destacar que a atividade envolvia a produção de adjetivos com plural regular (*vermelho, verde, amarelo*), bem como um irregular (*azul*). No entanto, não houve propriamente um controle entre os tipos de adjetivos, tendo-se em vista que, no concernente às cores escolhidas (i.e. os mais facilmente identificáveis por parte dos participantes), há mais adjetivos regulares do que irregulares na língua.

Figura 2 - Exemplo de prancha utilizada no experimento



Fonte: Elaborado pela autora (2021)

Os estímulos visuais foram construídos de modo que, a cada rodada do jogo, cada membro da dupla fosse obrigado a produzir pelo menos um sintagma no plural. Caso eles omitissem informações, não seria possível completar a identificação da imagem-alvo, já que os conjuntos de elementos apresentavam diversas propriedades comuns. Todas as sessões experimentais foram gravadas em áudio e integralmente transcritas para posterior análise. Em média, cada sessão durou em torno de 35 minutos.

4.2.2 Resultados e discussão

Os dados coletados foram transcritos, compilados e analisados estatisticamente por meio do programa R³⁹ (R Core Team, versão 4.0.0, 2020). Foram conduzidos três conjuntos de análises com objetivos e abordagens distintas que especificamos a seguir:

³⁹ O R é uma linguagem de programação comumente usada não somente em computação gráfica e testes estatísticos, como também na compilação e anotação de *corpora*, elaboração de listas de frequências, entre outras tarefas. Em Oushiro (2017), encontra-se disponível um tutorial de introdução à estatística para o uso do R. Vale ressaltar que as análises estatísticas foram conduzidas com base nos modelos propostos no referido curso.

- 1) **Análise dos dados brutos:** inclui todos os sintagmas nominais plurais produzidos pelos participantes durante a tarefa experimental. Esses dados incluem tanto sintagmas contendo os itens nominais experimentais (i.e. nomes alvo), quanto distratores (mas que seguem um padrão semelhante aos dos nomes experimentais). No caso dos distratores, a produção dos mesmos não era obrigatória para a execução da tarefa. Dessa forma, embora os dados brutos forneçam o maior universo de ocorrências, a quantidade e o tipo de item produzido não foram de fato controlados. No entanto, em virtude de apresentar resultados muito semelhantes aos obtidos na análise não atomística, optamos por apresentar detalhadamente os resultados obtidos a partir desta segunda análise.⁴⁰
- 2) **Análise não atomística dos sintagmas plurais contendo apenas nomes alvo** (i.e. experimentais): considera todos os sintagmas plurais contendo nomes alvo independentemente do item nominal em si apresentar ou não marcação morfofonológica de plural. Os sintagmas são separados em redundantes e não redundantes. Em outras palavras, nessa análise produções do tipo *Dois livros azul* e *Dois livro azul*, pertencem ao mesmo conjunto de dados. Nesse sentido, a análise não atomística pode não ser suficientemente informativa quanto aos efeitos da SF na marcação de plural nos nomes;
- 3) **Análise atomística dos sintagmas contendo apenas nomes alvo:** os sintagmas não redundantes são divididos em função da presença ou não de marcação no nome. Essa análise é a única que focaliza estritamente os efeitos das propriedades fônicas dos itens nominais na marcação de plural presente (ou não) no próprio item nominal. Embora essa abordagem restrinja o universo dos dados considerados, quando contrastada com as análises em (1) e (2), a mesma permite verificar os possíveis efeitos da SF com maior precisão.

⁴⁰ A título de informação, apresentamos aqui os resultados da primeira análise conduzida, que considera todos os sintagmas nominais produzidos pelos participantes durante a tarefa 1. De um total de 411 sintagmas plurais produzidos pelo grupo de crianças, 223 ocorrências (54%) apresentaram marcação redundante, enquanto 188 (46%) foram não redundantes. Já no grupo controle (adultos), foram registradas 333 (86%) ocorrências de sintagmas com marcação redundante e 52 (14%) com marcação não redundante, perfazendo um total de 385 sintagmas plurais produzidos. Os resultados de outras análises conduzidas a partir dos dados brutos podem ser encontradas nos Apêndices.

Adicionalmente, na seção 4.4 deste capítulo apresentamos duas análises que extrapolam os objetivos específicos primários do Experimento 1, mas fornecem resultados relevantes para embasar a discussão teórica que tecemos no Capítulo 5 e dialogam com algumas das reflexões levantadas a partir dos dados longitudinais. Para facilitar a leitura, organizamos as diferentes análises em seções separadas.

4.2.2.1 Análise 1 – SF dos nomes e marcação de plural – Dados brutos

Essa primeira análise foi de caráter bastante exploratório, sendo análoga à conduzida com os dados longitudinais. Ou seja, foram considerados todos os sintagmas plurais produzidos, independente de se tratarem de produções contendo nomes experimentais ou distratores. Esses últimos – em virtude de terem seu número de ocorrências não controlado em função de suas propriedades fonológicas –, podem ter acarretado uma tendência a uma maior produção de itens de um determinado tipo em detrimento de outro o que, por sua vez, pode ter levado os resultados numa direção não esperada, tal como observado em função da variável padrão de acentuação nos dados dos adultos. Importante lembrar também que essa primeira análise foi conduzida seguindo os moldes comumente utilizados nas pesquisas de cunho sociolinguístico, em que se costuma avaliar as dimensões associadas à SF separadamente. Todavia, como visto no capítulo 2, os resultados arrolados nos estudos sociolinguísticos não são unânimes no tocante ao papel dessas dimensões na alternância de uso dos padrões de concordância nominal.

Contudo, essa primeira análise se mostrou relevante ao fornecer um panorama dos dados quanto aos índices de variação registrada nos dois grupos e revelou uma discrepância inesperada quando comparada a produção de adultos e crianças no que tange à proporção de sintagmas redundantes não redundantes – que retomaremos detalhadamente mais adiante.

Após essa análise inicial e no intuito de obtermos um quadro mais preciso dos possíveis efeitos das variáveis investigadas, partimos para mais duas análises considerando somente os nomes alvo. Como já adiantado, tais análises foram conduzidas nos moldes de duas abordagens diferentes: (i) uma abordagem não atomística, que levou em consideração se o sintagma apresentava uma marcação totalmente redundante *vs.* uma marcação em pelo menos um dos itens flexionáveis,

independente do padrão de marcação interno do sintagma e (ii) uma abordagem atomística, que buscou observar a presença ou ausência da marcação de plural no nome, considerando apenas o conjunto de sintagmas não redundantes. Passemos, então, à análise não atomística conduzida com os nomes alvo.

4.2.2.2 Análise 2 – SF dos nomes e marcação de plural – Análise não atomística – nomes alvo

Inicialmente, na análise não atomística considerando apenas os nomes alvo, observamos na tabela (6) um panorama geral semelhante ao que emergiu da análise dos dados brutos: percentual equivalente de ocorrências de sintagmas redundantes (54%) e não redundantes (46%) na produção das crianças. Já no grupo controle, mantêm-se a predominância de sintagmas com marcação redundante (84%). Ao compararmos o número de sintagmas plurais redundantes e não redundantes produzidos por ambos os grupos, por meio de um teste de proporções, novamente não houve diferença estatisticamente significativa no caso das crianças ($\chi^2= 2.01(1)$, $p=0.1$). No entanto, registrou-se um efeito para o grupo de adultos ($\chi^2=131.12(1)$, $p<0.01$), com percentual significativamente maior de sintagmas plurais redundantes.

Tabela 6 - Distribuição dos sintagmas plurais em função da padrão de concordância aplicado (redundante X não redundante) no grupo experimental (crianças) e controle (adultos) – Experimento 1 – nomes alvo

Grupo	Padrão de concordância No total de ocorrências e %		Qui quadrado
	Redundante	Não Redundante	
Crianças (Grupo experimental)	168 (54%)	142 (46%)	$p=0.1$
Adultos (Grupo controle)	243 (84%)	47 (16%)	$p<0.01$

Fonte: Elaborado pela autora (2021)

Em virtude desses resultados, foi aplicado um modelo de regressão logística para comparar a produção de sintagmas não redundantes de adultos (16%) e crianças (46%). A tabela de coeficientes (7) abaixo, indica que a diferença entre as proporções é significativa e a estimativa de -1.4748 para os adultos reforça o resultado

apresentado nas proporções que revela um desfavorecimento do padrão não redundante na produção dos adultos em relação à das crianças.

Tabela 7 - Tabela de coeficientes extraída do modelo de regressão logística para ocorrências de sintagmas plurais não redundantes em função do grupo analisado (Crianças X Adultos) – Experimento 1 – nomes alvo

	Estimate	Std. Error	Z	P
Crianças (Intercept)	-0.1681	0.1140	-1.475	0.14
Adultos	-1.4748	0.1959	-7.527	<0.01

Fonte: elaborado pela autora (2021)

Uma possível interpretação para a diferença entre as produções de crianças e adultos está relacionada com a visão – bastante difundida na literatura – sobre o papel do grau de escolarização na realização da concordância variável no PB: quanto maior o grau de escolarização do falante (representado aqui pelo grupo de adultos), maior a tendência ao uso de marcação redundante de plural. No entanto, é importante lembrar que esses resultados se distanciam do observado nos dados naturalísticos, nos quais a produção adulta e infantil apresentou proporções semelhantes de sintagmas redundantes e não redundantes. Assim sendo, o nível de escolaridade dos participantes não parece ser o fator decisivo para explicar as diferenças entre crianças e adultos encontradas nos dados experimentais.

Uma segunda interpretação está relacionada com aspectos próprios da situação comunicativa na qual os dados foram coletados. Nos dados longitudinais, temos a interação adulto-criança, em contextos cotidianos que, embora possam apresentar um teor lúdico em alguns casos, não se limitam à participação em jogos e brincadeiras. Já nos dados experimentais, a produção linguística está associada a um contexto inteiramente lúdico, com interações adulto-adulto e criança-criança. Nos adultos, estima-se que o nível de monitoramento da fala seja semelhante em ambos tipos de contextos de interação – já que mesmo quando o interlocutor é uma criança, o adulto está ciente de que sua fala está sendo gravada e, além disso, podemos imaginar que o cuidador busque inconscientemente apresentar uma “fala modelo” para seu interlocutor infantil. Já no caso das crianças, o contexto de interação criança-criança apresenta características da denominada “fala entre pares” (*peer talk*) fato que pode ser chave para compreender as diferenças observadas entre os dados coletados (dados

de interação criança – crianças nos dados experimentais X dados de interação criança-adulto nos dados naturalísticos). Esse ponto será retomado e aprofundado na seção 4.5.

O modelo de regressão logística foi utilizado também para analisarmos os dados em função de cada uma das duas dimensões da SF investigadas. A tabela (8) a seguir apresenta a distribuição dos sintagmas plurais em função da variável *padrão de acentuação*.

Tabela 8 - Distribuição dos sintagmas plurais em função do padrão de acentuação (oxítonos X paroxítonos) no grupo experimental (crianças) e controle (adultos) – Experimento 1 - nomes alvo

Grupo	Padrão de acentuação	Padrões de concordância		Regressão logística	Total de ocorrências
		Concordância redundante	Concordância não redundante		
Crianças	Nomes oxítonos Ex. <i>maçã</i>	90 (54%)	76 (46%)	$p=0.9$	166
	Nomes paroxítonos Ex. <i>gato</i>	78 (54%)	66 (46%)		144
Adultos (controle)	Nomes oxítonos Ex. <i>maçã</i>	102 (76%)	33 (24%)	$p < 0.01$	135
	Nomes paroxítonos Ex. <i>gato</i>	141 (91%)	14 (9%)		155

Fonte: Elaborado pela autora (2021)

No grupo das crianças, mais uma vez não se observou efeito de *padrão de acentuação*, com médias estatisticamente equivalentes entre as condições [itens oxítonos/ não-redundantes] vs. [itens paroxítonos/ não-redundantes]. Em termos percentuais, observou-se que a estimativa de ocorrências de sintagmas plurais não redundantes é cerca de 46% tanto em itens oxítonos quanto paroxítonos (cf. Gráfico de efeitos (14) nos Apêndices).

Em contrapartida, assim como na análise dos dados brutos, no grupo controle foi registrado um efeito de *padrão de acentuação*, como se vê na tabela (9), com percentual de ocorrências de sintagmas plurais não redundantes significativamente maior nas condições com itens oxítonos. A tabela indica ainda que a estimativa de -1.1812 para adultos sustenta o resultado das proporções que revela um

desfavorecimento do padrão não redundante em sintagmas com itens paroxítonos em relação aos oxítonos.

Tabela 9- Tabela de coeficientes extraída do modelo de regressão logística para ocorrências de sintagmas plurais não redundantes em função da variável padrão de acentuação (Adultos) – Experimento 1 - nomes alvo

	Estimate	Std. Error	z	p
Oxítone (Intercept)	-1.1285	0.2003	-5.635	<0.01
Paroxítone	-1.1812	0.3444	-3.430	<0.01

Fonte: Elaborado pela autora (2021)

No que concerne à segunda dimensão associada à SF considerada, os resultados do modelo de regressão logística indicam que não houve efeito de *número de sílabas* para o grupo de crianças e nem para o grupo controle, com percentuais semelhantes de sintagmas plurais não redundantes associados a itens dissílabos e trissílabos, como disposto na tabela (10).

Tabela 10 - Distribuição dos sintagmas plurais em função do número de sílabas (dissílabos X trissílabos) no grupo experimental (crianças) e controle (adultos) – Experimento 1 – nomes alvo

Grupo	Número de sílabas	Padrões de concordância		Regressão logística	Total de ocorrências
		Concordância redundante	Concordância não redundante		
Crianças	Nomes dissílabos Ex. <i>caju</i>	87 (55%)	72 (45%)	$p < 0.8$	159
	Nomes trissílabos Ex. <i>urubu</i>	81 (54%)	70 (46%)		151
Adultos (controle)	Nomes dissílabos Ex. <i>caju</i>	133 (86%)	21 (14%)	$p < 0.2$	154
	Nomes trissílabos Ex. <i>urubu</i>	110 (81%)	26 (19%)		136

Fonte: Elaborado pela autora (2021)

As análises conduzidas até aqui consideram as dimensões *padrão de acentuação* e *número de sílabas* de forma independente, tal como comumente encontrado nas pesquisas de cunho descritivo. Seguindo o mesmo raciocínio adotado

em estudos sociolinguísticos prévios, levantamos as previsões de que nomes oxítonos, bem como os nomes trissílabos, concentrariam mais marcas redundantes de plural. No entanto, os resultados obtidos não se mostram compatíveis com esse primeiro conjunto de previsões. Pelo contrário, a análise não atomística considerando as duas variáveis de forma isolada (análise univariada) revelou apenas um efeito de *padrão de acentuação* para o grupo de controle, mas na direção oposta ao previsto inicialmente (i.e. itens oxítonos favorecendo a marcação não redundante).

Em virtude de estarmos trabalhando com mais de uma variável independente (padrão de acentuação e número de sílabas), decidimos por realizar análises multivariadas, no intuito de verificarmos se o comportamento das variáveis estudadas se mantém ou não em face uma da outra, e o que se observou foi o mesmo resultado que havíamos visto nas análises univariadas. (cf. Tabela de coeficientes (34), nos Apêndices). Também conduzimos um modelo com interação para testar se as variáveis interagem entre si. E o resultado do modelo indica que a interação entre padrão de acentuação e número de sílabas não é significativa, nem para crianças, nem para os adultos. (cf. Tabela de coeficientes (35), nos Apêndices)

Nosso próximo passo, foi conduzir uma nova série de análises pautada agora no segundo conjunto de hipóteses e previsões levantadas, de acordo com o qual, as duas dimensões da SF atuariam, não de forma independente, mas conjunta, afetando os padrões de marcação de plural. Assim sendo, a tabela (11) apresenta a distribuição dos sintagmas plurais em função das condições experimentais, resultado da combinação das duas variáveis independentes, correlatas às dimensões da SF investigadas.

Tabela 11 - Distribuição dos sintagmas plurais em função das condições experimentais no grupo experimental (crianças) e controle (adultos) –

Experimento 1 - nomes alvo

Grupo	Condição experimental	Padrões de concordância		Total de ocorrências
		Concordância redundante	Concordância não redundante	
Criança	Nomes			
	Paroxítonos dissílabos Ex. <i>livro</i>	40 (53%)	36 (47%)	76
	Oxítonos Dissílabos Ex. <i>batom</i>	47 (57%)	36 (43%)	83
	Paroxítonos trissílabos Ex. <i>sapato</i>	38 (56%)	30 (44%)	68
Adultos (controle)	Oxítonos trissílabos Ex. <i>picolé</i>	43 (52%)	40 (48%)	83
	Paroxítonos dissílabos Ex. <i>livro</i>	70 (92%)	06 (8%)	76
	Oxítonos Dissílabos Ex. <i>batom</i>	63 (81%)	15 (19%)	78
	Paroxítonos trissílabos Ex. <i>sapato</i>	71 (90%)	08 (10%)	79
	Oxítonos trissílabos Ex. <i>picolé</i>	39 (68%)	18 (32%)	57

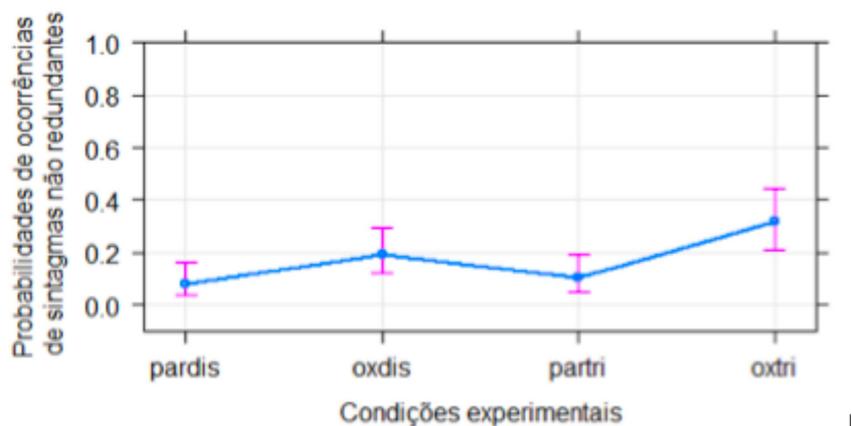
Fonte: Elaborado pela autora (2021)

Quanto à produção de sintagmas não redundantes nos dados dos adultos, os resultados de um modelo de regressão logística revelam diferença significativa entre nomes dissílabos oxítonos⁴¹ e trissílabos oxítonos em relação aos itens dissílabos paroxítonos, considerados aqui como valor de referência. Esses resultados indicam que itens trissílabos oxítonos concentraram o maior número de sintagmas não redundantes e itens dissílabos paroxítonos o menor, i.e. exatamente o oposto do esperado. Lembramos que, de acordo com a nossa hipótese de uma atuação conjunta das duas dimensões da SF investigadas, as previsões experimentais levantadas postulam o gradiente na marcação de plural apresentado em (9) a seguir. Essa previsão não é, contudo, sustentada pelos resultados reportados até aqui, como visto no gráfico (2).

(9)

+ Marcação não redundante ←-----→ + Marcação redundante
 Nomes ParDis> Nomes OxDis≈ Nomes ParTri> Nomes OxTri

Gráfico 2 - Gráfico de efeitos das condições experimentais na distribuição dos sintagmas não redundantes no grupo controle (adultos) – Experimento 1- nomes alvo



Fonte: Elaborado pela autora (2021)

⁴¹ As condições experimentais estão representadas nas tabelas de coeficientes e nos gráficos de efeitos com as seguintes nomenclaturas: Pardis (Paroxítonos dissílabos), Oxdis (Oxítonos dissílabos), Partri (Paroxítonos trissílabos) e Oxtri (Oxítonos trissílabos).

No que diz respeito aos dados das crianças, não foram registrados efeitos significativos quando contrastadas as condições experimentais (cf. Gráfico de efeitos (15) nos Apêndices).

Vale lembrar que na análise não atomística reportada até aqui, investigamos os possíveis efeitos de SF no sintagma considerado globalmente, i.e. sem entrar no mérito de o nome em si apresentar ou não marcação. No entanto, pensamos ser necessário investigar estritamente os efeitos das propriedades fônicas na realização da marca de plural nos itens nominais propriamente ditos. Para tal, conduzimos a análise atomística, a qual apresentaremos a seguir.

4.2.2.3 Análise 3 – SF dos nomes e marcação de plural – Análise atomística – nomes alvo

Na abordagem atomística – tomando novamente como universo de análise apenas os nomes alvo –, consideramos unicamente os sintagmas não redundantes produzidos, os quais foram divididos em função da presença ou não de marcação explícita de plural no nome. Sendo assim, nesta análise, sintagmas como *dois jacarés azul* e *dois jacaré azul* que, na análise não atomística seriam classificados da mesma maneira (i.e. como não redundantes), agora, pertencem a grupos diferentes. Em virtude de nosso objetivo primário ser a investigação dos efeitos da SF nos nomes, pensamos que tal abordagem possa ser mais informativa no sentido de permitir observar em que medida características fonológicas do item nominal acarretam ou não marcação no próprio item.

Como pode ser visto na tabela (12) abaixo, de um total de 142 sintagmas não redundantes (com 2, 3, 4 ou 5 palavras) produzidos pelo grupo de crianças, 54 ocorrências (38%) apresentaram marcação de plural no nome, enquanto 88 (62%) apresentaram nomes não marcados. O grupo controle (adultos), por sua vez, produziu 47 sintagmas não redundantes, dos quais 36 (77%) tinham nomes marcados para plural e 11 (23%) não apresentaram marcação de plural no nome. É possível observar que a diferença entre a produção de nomes marcados e não marcados foi estatisticamente significativa em ambos os grupos e que as tendências foram opostas: crianças produziram mais nomes não marcados ($\chi^2 = 7.669 (1), p < 0.01$) e adultos mais nomes marcados para o plural ($\chi^2 = 12.255 (1), p < 0.01$).

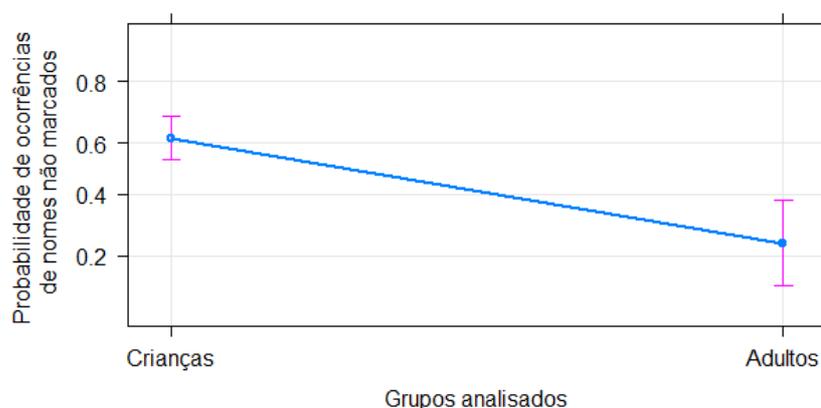
Tabela 12 - Distribuição de sintagmas plurais não redundantes em função da marcação de plural no nome (nome marcado X nome não marcado) no grupo experimental (crianças) e controle (adultos) – Experimento 1 – nomes alvo

Grupo	Marcação de plural no nome No total de ocorrências e %		Qui quadrado
	Nome marcado Ex. <i>Três maçãs</i>	Nome não marcado Ex. <i>Três maçã</i>	
Crianças (Grupo experimental)	54 (38%)	88 (62%)	$p < 0.01$
Adultos (Grupo controle)	36 (77%)	11 (23%)	$p < 0.01$

Fonte: Elaborado pela autora (2021)

Além disso, o resultado do modelo de regressão logística conduzido revelou que a diferença entre as proporções de produção de nomes não marcados pelas crianças (62%) e pelos adultos (23%) foi estatisticamente significativa, como pode ser inferido em virtude da não sobreposição dos intervalos de confiança no gráfico (3), com um percentual significativamente maior de ocorrências de nomes não marcados nos dados produzidos pelas crianças.

Gráfico 3 - Gráfico de efeitos da produção de nomes não marcados em sintagmas não redundantes em função do grupo analisado (Crianças X Adultos) - Experimento 1 – nomes alvo



Fonte: Elaborado pela autora (2021)

A distribuição da marcação de plural no nome em sintagmas não redundantes foi analisada também por meio de modelos de regressão logística em função de cada uma das duas dimensões da SF investigadas – *padrão de acentuação e número de sílabas* –, bem como das condições experimentais (combinando as duas dimensões). O resultado dos modelos conduzidos em função do *padrão de acentuação*, registrou uma diferença significativa para as crianças, mas não para os adultos. Nos dados infantis, nomes oxítonos apresentaram *contra* as previsões iniciais, menos marcação explícita de plural do que os nomes paroxítonos.

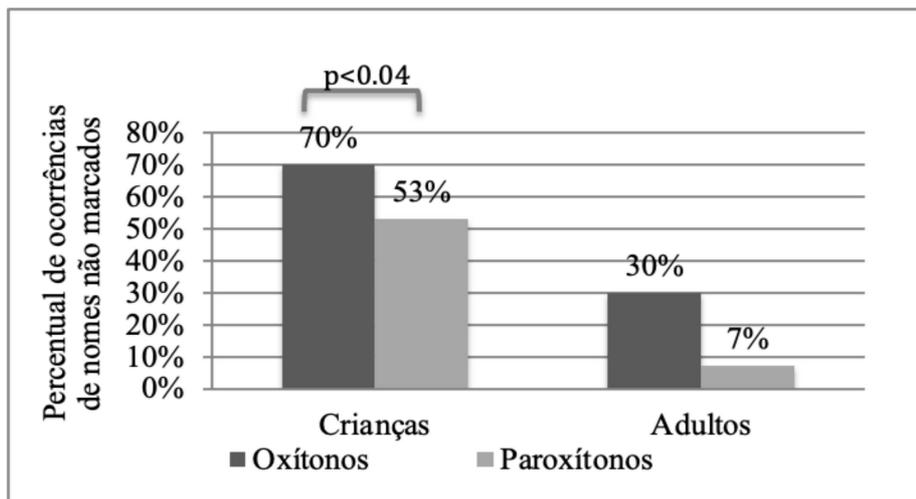
Tabela 13 - Distribuição de marcação de plural no nome em sintagmas não redundantes em função do padrão de acentuação (oxítonos X paroxítonos) no grupo experimental (crianças) e controle (adultos) – Experimento 1

Grupo	Padrão de acentuação	Marcação de plural no nome No total de ocorrências e %		Regressão logística	Total de ocorrências
		Nome marcado	Nome não marcado		
Crianças	Nomes oxítonos Ex. maçã	23 (30%)	53 (70%)	$p=0.04$	76
	Nomes paroxítonos Ex. gato	31 (47%)	35 (53%)		66
Adultos (controle)	Nomes oxítonos Ex. maçã	23 (70%)	10 (30%)	$p < 0.1$	33
	Nomes paroxítonos Ex. gato	13 (93%)	01 (7%)		14

Fonte: Elaborado pela autora (2021)

Interessante notar, como se vê na tabela (13) e no gráfico (4), que, apesar das diferenças em termos percentuais entre dados infantis e adultos, observamos que em ambos os grupos, houve uma maior incidência de nomes oxítonos não marcados.

Gráfico 4 - Percentual de nomes não marcados em sintagmas não redundantes em função de padrão de acentuação – no grupo experimental (crianças) e controle (adultos) – Experimento 1 – nomes alvo



Fonte: Elaborado pela autora (2021)

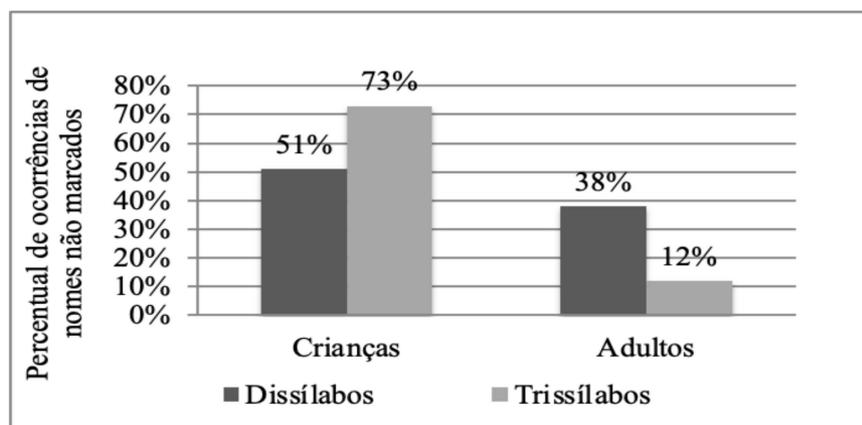
Quanto à segunda dimensão associada à SF considerada, um modelo de regressão logística indica que, para o grupo das crianças, houve efeito de *número de sílabas*, com um maior percentual de nomes não marcados associados a itens trissílabos, contrariando a visão mais difundida na literatura sociolinguística e nossas previsões iniciais. Para o grupo controle, também houve efeito de *número de sílabas*. Contudo, como disposto na tabela (14) e no gráfico (5), os dados dos adultos apresentaram um maior percentual de nomes não marcados associados a itens dissílabos.

Tabela 14 - Distribuição de marcação de plural no nome em sintagmas não redundantes em função do número de sílabas (dissílabos x trissílabos) no grupo experimental (crianças) e controle (adultos) –Experimento 1 - nomes alvo

Grupo	Número de sílabas	Marcação de plural no nome No total de ocorrências e %		Regressão logística	Total de ocorrências
		Nome marcado	Nome não marcado		
Crianças	dissílabos Ex. <i>garfo</i>	35 (49%)	37 (51%)	$p < 0.01$	72
	trissílabos Ex. <i>estrela</i>	19 (27%)	51 (73%)		70
Adultos (controle)	dissílabos Ex. <i>garfo</i>	13 (62%)	08 (38%)	$p < 0.04$	21
	Nomes trissílabos Ex. <i>estrela</i>	23 (88%)	03 (12%)		26

Fonte: Elaborado pela autora(2021)

Gráfico 5 - Percentual de nomes não marcados em sintagmas não redundantes em função de número de sílabas - no grupo experimental (crianças) e controle (adultos) – Experimento 1 - nomes alvo



Fonte: Elaborado pela autora (2021)

A distribuição de marcação de plural no nome em sintagmas não redundantes, apresentada na tabela (15), foi analisada também em função das condições experimentais (combinando as duas dimensões da SF investigadas).

Tabela 15 - Distribuição de marcação de plural no nome em sintagmas não redundantes em função das condições experimentais no grupo experimental (crianças) e controle (adultos) – Experimento 1 - nomes alvo

Grupo	Condição experimental	Marcação de plural no nome		Total de ocorrências
		No total de ocorrências e %		
	Nomes	Nome marcado	Nome não marcado	
Crianças	Paroxítonos dissílabos Ex. <i>livro</i>	19 (53%)	17 (47%)	36
	Oxítonos Dissílabos Ex. <i>caju</i>	16 (44%)	20 (56%)	36
	Paroxítonos trissílabos Ex. <i>estrela</i>	12 (40%)	18 (60%)	30
	Oxítonos trissílabos Ex. <i>jacaré</i>	07 (17%)	33 (83%)	40
Adultos (controle)	Paroxítonos dissílabos Ex. <i>livro</i>	05 (83%)	01 (17%)	06
	Oxítonos Dissílabos Ex. <i>caju</i>	08 (53%)	07 (47%)	15
	Paroxítonos trissílabos Ex. <i>estrela</i>	08 (100%)	0 (0%)	08
	Oxítonos trissílabos Ex. <i>jacaré</i>	15 (83%)	03 (17%)	18

Fonte: Elaborado pela autora (2021)

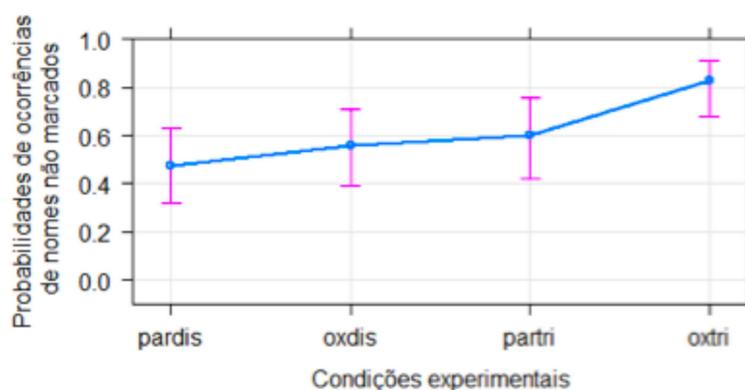
Ao conduzirmos um modelo de regressão logística nos dados das crianças, foi registrada uma diferença significativa entre nomes paroxítonos dissílabos e oxítonos trissílabos, como indicado na tabela de coeficientes (16), e no gráfico de efeitos (6), com um número de significativamente maior de nomes não marcados para o padrão oxítono trissílabo, isto é, novamente *contra* as previsões. Já a análise dos dados dos adultos, não registrou nenhuma diferença significativa entre as condições experimentais (cf. Gráfico 16, nos Apêndices).

Tabela 16 - Tabela de coeficientes extraída do modelo de regressão logística para ocorrências de não marcação no nome em função das condições experimentais (Crianças) – Experimento 1 – nomes alvo

Crianças				
	Estimate	Std. Error	Z	p
Pardis (Intercept)	-0.1112	0.3338	-0.333	0.7
Oxdis	0.3344	0.4732	-0.707	0.4
Partri	0.5167	0.5003	1.033	0.3
Oxtri	1.6618	0.5335	3.115	<0.01

Fonte: Elaborado pela autora (2021)

Gráfico 6 - Gráfico de efeitos das condições experimentais na distribuição de nomes não marcados no grupo experimental (Crianças) – Experimento 1 - nomes alvo



Fonte: Elaborado pela autora (2021)

4.2.2.4 Síntese dos principais resultados do Experimento 1

O primeiro experimento conduzido investigou duas das três dimensões da SF delimitadas pela literatura sociolinguística e seus possíveis efeitos na realização dos padrões de concordância no âmbito nominal: o *padrão de acentuação* e o *número de sílabas* dos nomes. A tarefa experimental utilizada permitiu coletar um número expressivo de sintagmas plurais (411 sintagmas nos dados infantis e 385 na produção adulta, considerando os dados brutos), possibilitando ainda a obtenção de dados que recobrem as diferentes características dos nomes que se pretendia explorar.

Os quadros 4 e 5 a seguir, esquematizam os principais resultados relativos à influência da SF, obtidos em função dos três conjuntos de análises conduzidos (dados brutos, análise não atomística e análise atomística) nas duas abordagens propostas (atuação das variáveis isoladamente e combinadas entre si). Em cinza, destacamos as análises que não revelaram efeitos significativos, em vermelho os resultados que contrariam as previsões iniciais e em verde aqueles que estão de acordo com tais previsões:

Quadro 4 - Síntese dos resultados do Experimento 1 obtidos considerando as variáveis independentes separadamente (padrão de acentuação e número de sílabas)

Grupo	Análise 1 (dados brutos)	Análise 2 (nomes alvo – abordagem não atomística)	Análise 3 (nomes alvo – abordagem atomística)
Crianças	Nenhum efeito significativo.	Nenhum efeito significativo.	Efeito de <i>padrão de acentuação</i> : <i>contra</i> as previsões itens oxítonos + marcação não redundante. ----- Efeito de <i>número de sílabas</i> : <i>contra</i> as previsões nomes trissílabos + marcação não redundante
Adultos	Efeito de <i>padrão de acentuação</i> : <i>contra</i> as previsões itens oxítonos + marcação não redundante.	Efeito de <i>padrão de acentuação</i> : <i>contra</i> as previsões itens oxítonos + marcação não redundante.	Efeito de <i>número de sílabas</i> : na direção das previsões nomes trissílabos + marcação redundante

Fonte: Elaborado pela autora (2021)

Quadro 5 - Síntese dos resultados do Experimento 1 obtidos considerando as condições experimentais (cruzamento das variáveis padrão de acentuação e número de sílabas)

Grupo	Análise 1 (dados brutos)	Análise 2 (nomes alvo – abordagem não atomística)	Análise 3 (nomes alvo – abordagem atomística)
Crianças	Análise não conduzida ⁴² .	Nenhum efeito significativo.	<i>Contra</i> as previsões nomes oxítonos trissílabos concentram o maior número de nomes não marcados
Adultos	Análise não conduzida.	<i>Contra</i> as previsões, nomes <i>oxítonos trissílabos</i> concentraram um número significativamente maior de sintagmas não redundantes.	Nenhum efeito significativo.

Fonte: Elaborado pela autora (2021)

Quanto ao delimitado em função das hipóteses e previsões iniciais, os resultados obtidos:

- (i) Apontam diferenças estatisticamente significativas em todas as análises conduzidas no que tange à proporção de sintagmas plurais redundantes e não redundantes produzida por crianças e adultos. Enquanto esses repetiram a prevalência de sintagmas plurais redundantes – previamente observada nos dados naturalísticos –, as crianças produziram sintagmas redundantes e não redundantes em percentuais semelhantes. Ainda quanto ao comportamento diferenciado em função do grupo, a única análise (atomística, considerando às variáveis separadamente) que revelou efeitos para ambas as amostras apontou um comportamento oposto quando considerados ambos os grupos. Nesse caso, foi observado que enquanto

⁴² Considerando que nos dados brutos – em função da presença de itens distratores – não há uniformidade entre as produções associadas a cada condição experimental, optamos por não conduzir a análise que envolve o cruzamento das duas variáveis para esse universo de ocorrências.

nos dados infantis o resultado foi *contra* as previsões (nomes trissílabos concentraram um número maior de marcações não redundantes), os dados adultos foram na direção inicialmente esperada (nomes trissílabos favorecendo a marcação explícita de plural);

- (ii) Não fornecem evidências robustas compatíveis com a atuação de cada variável investigada de forma isolada. Lembrando que, apenas a análise atomística apontou um resultado na direção esperada quando considerados os dados de produção adulta para *número de sílabas*, com nomes trissílabos favorecendo a marcação redundante. Os demais resultados, ora não revelam efeitos estatisticamente significativos, ora apontam na direção oposta à prevista.
- (iii) Também não se mostram consistentes com as previsões estabelecidas em função da hipótese da existência de um *continuum* ou gradiente de SF (i.e. considerando a ocorrência de sintagmas redundantes e não redundantes em cada condição experimental). Novamente, contrariando as previsões, nomes oxítonos trissílabos concentraram um número maior de ocorrências não redundantes na análise não atomística para os dados adultos e na análise atomística para os dados das crianças.

Tomados em conjunto, os resultados não confirmam os achados mais difundidos nas pesquisas conduzidas no âmbito da sociolinguística. Embora os resultados apontem alguns efeitos relacionados às propriedades fônicas dos itens nominais investigadas, certamente tais efeitos não parecem ocorrer na direção prevista com base nas hipóteses aventadas. Apenas o efeito de *número de sílabas* observado nos dados dos adultos – na análise atomística para as variáveis isoladas – se mostrou compatível com a hipótese segundo a qual uma maior quantidade de material fônico poderia favorecer a marcação redundante. No entanto, se observamos o resultado da análise não atomística que considera as condições experimentais para esse mesmo grupo, verificamos que nomes trissílabos associados ao padrão de acentuação oxítono geraram uma proporção significativamente maior de sintagmas não redundantes. Isso sugere que o *número de sílabas* isoladamente não seria suficiente para favorecer a marcação dos nomes do plural regular.

Por outro lado, o fato de nomes oxítonos trissílabos – tidos como os mais salientes dentre os nomes do plural regular em função das nossas hipóteses iniciais –

terem concentrado o maior número de sintagmas não redundantes em ambos os grupos (na análise não atomística para os adultos e na atomística para as crianças) é um resultado que precisa ser melhor explicado. Por que motivo nomes com três sílabas nos quais a marca de plural recai na sílaba tônica seriam menos marcados? Uma possível explicação diz respeito à própria classe de nomes aqui investigada. Em que medida os nomes pertencentes ao chamado plural regular apresentam de fato diferenças de saliência entre si? Em outras palavras, existe de fato uma distinção perceptual entre itens cujo plural é formado apenas pelo acréscimo de -S? A atuação das duas dimensões investigadas no Experimento 1 pode ser considerada quando dissociada da terceira dimensão – frequentemente também a mais investigada – da SF, qual seja, os processos morfofonológicos de formação de plural?

A ideia de que não haja de fato distinções de SF associadas aos itens nominais regulares não se mostra, no entanto, uma hipótese satisfatória ou suficiente, já que a mesma não permite fornecer uma explicação para os efeitos encontrados. Além disso, essa ideia implica no fato de que as três dimensões da SF precisariam necessariamente atuar em conjunto para outorgar realce fônico a um dado item.

Tomando como ponto de partida os dados longitudinais, que não apontaram diferenças estatisticamente significativas entre itens salientes e não salientes, consideramos que o Experimento 1 permitiu estabelecer as bases para uma investigação mais refinada dos efeitos da variável SF na produção dos padrões de concordância variável. No entanto, como vimos, os resultados obtidos estão longe de esgotar os nossos questionamentos a respeito do assunto. Por essa razão, um segundo experimento foi conduzido no âmbito desta tese.

4.3 EXPERIMENTO 2: INVESTIGANDO A SF NA PRODUÇÃO DA CONCORDÂNCIA NOMINAL VARIÁVEL – PROCESSOS MORFOFONOLÓGICOS DE FORMAÇÃO DE PLURAL

O segundo experimento conduzido teve como objetivo principal investigar a terceira das dimensões da SF: os *processos morfofonológicos de formação de plural* nos nomes do PB. Para tal, decidimos pela manutenção da técnica experimental utilizada no Experimento 1, ou seja, a técnica de produção semielicida a partir de imagens. A nossa escolha justifica-se em virtude de esta ter se mostrado produtiva

para eliciar sintagmas plurais, possibilitando a coleta de um número expressivo de dados, quando comparado ao número de ocorrências registrado na produção naturalística, não só no que tange à produção de sintagmas plurais, como também à produção dos tipos de nomes produzidos (em termos das dimensões da SF investigadas).

Nesse sentido, nossos objetivos específicos primários nesse segundo experimento foram:

- (i) Analisar a produção de sintagmas redundantes e não redundantes por crianças adquirindo o PB e falantes adultos;
- (ii) Investigar experimentalmente em que medida as crianças e adultos se mostram sensíveis à dimensão *processos morfofonológicos de formação de plural*;
- (iii) Comparar os resultados experimentalmente obtidos na presente pesquisa àqueles levantados a partir da análise da produção espontânea de crianças e adultos, bem como por pesquisas prévias.

Além disso, delimitamos os mesmos objetivos secundários levantados para o Experimento 1 e reiterados a seguir:

- (i) Coletar dados que permitissem avaliar outros pontos que pensamos relevantes para a compreensão da concordância nominal variável – tais como tipo de estrutura em que o nome ocorre e os padrões de marcação internos ao sintagma –, para que a partir de sua análise, pudéssemos fornecer uma possível formalização do fenômeno da concordância nominal variável no PB;
- (ii) Investigar possíveis efeitos da constituição morfofonológica dos numerais na realização dos padrões de concordância, no intuito de verificar se a terminação em -S, encontrada em numerais como *dois*, *três* e *seis*, por exemplo, poderia ser vista como resultado de uma possível falsa analogia entre tal marca e o morfema de plural -S dos nomes.

Tomando como ponto de partida os resultados de estudos descritivos prévios (SCHERRE, 1988; TABOSA, 2016; dentre outros), bem como nossos próprios dados naturalísticos que sugerem haver alguma diferença entre itens mais e menos salientes (embora não tenham sido obtidos efeitos estatisticamente nas análises conduzidas),

levantamos como hipótese geral inicial a efetiva atuação da variável SF. Especificamente, no que tange à *dimensão processos morfofonológicos de formação de plural*, a hipótese levantada é compatível com a atuação do princípio da SF, segundo o qual, quanto maior a diferenciação de material fônico na oposição singular/plural, maior a saliência do item e, conseqüentemente, maior a tendência em se preservar a marca morfológica redundante de plural. Assim sendo, processos de formação de plural que gerem os contrastes mais nítidos entre as formas singular e plural de um dado item devem favorecer a marcação redundante.

4.3.1 Método

Na tarefa de produção eliciada a partir de imagens adaptada para esse segundo experimento, foi considerada como variável independente o *tipo de processo morfofonológico de formação de plural*, com três níveis associados a ela: (i) formação do plural de nomes terminados ortograficamente em -R, (ii) formação do plural de nomes terminados ortograficamente em -L; e (iii) formação do plural dos nomes regulares. A nossa escolha toma como base a escala hierárquica de SF proposta por Scherre (1988), segundo a qual os itens nominais são ranqueados de acordo o maior ou menor grau de diferenciação de material fônico na oposição singular/plural. Mais à frente, apresentamos de forma detalhada a motivação para a escolha desses três processos de formação de plural específicos. Dessa variável e seus correspondentes níveis, temos um total de três condições experimentais, cujos exemplos podemos observar no quadro (6) abaixo. Vale lembrar que como variável dependente, assim como no Experimento 1, foi considerado o número de sintagmas plurais redundantes e não redundantes produzidos em cada condição.

Quadro 6 - Condições experimentais do Experimento 2

Condição 1: processo de formação de plural de nomes terminados ortograficamente em -R
colar, mulher, colher, cobertor, flor
Condição 2: processo de formação de plural de nomes terminados ortograficamente em -L
jornal, anel, pincel, caracol, anzol
Condição 3: processo de formação de plural de nomes regulares
gambá, sofá, maçã, jacaré, caju

Fonte: Elaborado pela autora (2021)

Foram levantadas as seguintes previsões:

- (i) Seguindo os moldes do Experimento 1, ou seja, tomando como base o observado nos dados naturalísticos, é esperado um número de produções redundantes e não redundantes semelhante para adultos e crianças;

No que tange à variável *tipos de processos*, a partir da hipótese estabelecida e das condições experimentais delineadas, foi levantada a seguinte previsão:

- ii) Tomando como base a escala proposta por Scherre (1988), em que itens em –L seriam mais salientes que itens em –R que por sua vez são mais salientes do que os nomes regulares, são esperadas diferenças significativas entre as proporções de sintagmas não redundantes quando comparados os três *tipos de processos*;⁴³

Contudo, ao tomarmos itens terminados ortograficamente em –L ou –R como itens [+salientes] e itens com plural regular como [-salientes], nos moldes de pesquisas prévias, tais como Oushiro e Guy (2015), Oushiro (2016), Azalim (2016) e Azalim et al (2018), levantamos a seguinte previsão:

- iii) não são esperadas diferenças significativas entre os dois primeiros *tipos de processos*, mas sim entre eles e os plurais regulares quando comparadas as produções de sintagmas não redundantes produzidas.

⁴³Pensamos que há outros fatores – tais como padrão de acentuação e número de sílabas –, por exemplo, que poderiam ter algum efeito nos resultados da dimensão processos de formação de plural. Tomemos por exemplo, como apontado em Azalim et al (2018), as palavras *mulher* e *papel*, ambas oxítonas. Na formação do plural de *mulher*, temos um acréscimo de sílaba (*mu-lher _mu-lhe-res*) e um deslocamento da sílaba tônica. Já no nome *papel*, na forma plural o número de sílabas, bem como o padrão de acentuação da palavra são mantidos. Nesse sentido, em *mulher* o contraste no número de sílabas entre as formas singular e plural poderia ser considerado um fator de realce da saliência da forma plural (i.e., ele seria claramente distinto da forma singular). Por outro lado, o deslocamento do acento na forma plural poderia reduzir tal saliência. Já em *papel*, a marcação de plural coincide com a sílaba tônica, o que, conseqüentemente o tornaria perceptualmente mais saliente. Seguindo esse raciocínio, era de se esperar que o padrão silábico dos itens fosse, então, relevante na análise do contraste de material fônico entre formas singulares e plurais, uma vez que tal padrão permitiria prever quais seriam as mudanças implementadas em cada caso. No entanto, como vimos no Experimento 1, em que essas duas dimensões por si só não apresentaram efeitos consistentes nos resultados, decidimos por manter aqui o estabelecido pela escala de saliência proposta por Scherre (1988).

*Participantes*⁴⁴

Os participantes desta tarefa compreendem 16 crianças (sendo 04 meninos e 12 meninas), todos com 6 anos de idade completos, cursando o 1º Ano do Ensino Fundamental de escolas da rede privada de ensino em Juiz de Fora. Parte do grupo foi testado nas instalações das escolas em que as crianças estudam, enquanto o restante foi testado na residência de algumas delas. Vale destacar que todas as crianças participaram do experimento de forma voluntária. A faixa etária foi escolhida seguindo os moldes estabelecidos para o Experimento 1 no intuito de posteriormente compararmos os dados de ambos os experimentos de forma mais equilibrada. Como grupo controle, participaram 20 adultos estudantes universitários (sendo 12 homens e 08 mulheres) com idades entre 20 e 31 anos. Vale ressaltar que tanto as crianças, quanto os adultos, só participaram de apenas um experimento, ou seja, somente do experimento 1 ou apenas do experimento 2.

Materiais

Assim como no Experimento 1, as imagens utilizadas para eliciar a produção dos sintagmas foram projetadas pelo programa *Power point*. Foram utilizados dois computadores portáteis, cartões contendo as imagens-alvo impressos em papel *couché* e um gravador SONY, modelo PCM-D50 de alta resolução para captar a produção dos participantes. Os estímulos experimentais consistiam de um conjunto de 15 nomes experimentais – sendo 5 por condição – e algumas imagens distratoras. A ordem de apresentação dos estímulos foi semialeatorizada manualmente.

Seguindo os moldes estabelecidos por nós em estudo prévio (AZALIM, 2016; AZALIM et al, 2018), a escolha dos itens nominais para compor os estímulos experimentais foi realizada em função da dimensão *processos morfofonológicos de formação de plural* do fator SF, tomando como base a escala hierárquica proposta por

⁴⁴ A aplicação do Experimento 2 teve início em fevereiro de 2020 e, pouco tempo depois, foi interrompida em função da pandemia de COVID-19. Só foi possível retomar a pesquisa experimental em 2021 e, mesmo assim, de forma muito gradual e vagarosa. Por esse motivo, embora não fosse essa nossa intenção original, o grupo de crianças acabou ficando um pouco menor que no Experimento 1.

Scherre (1988). No entanto, dentre os seis níveis propostos pela autora⁴⁵, decidimos por priorizar, apenas três, a saber: itens nominais cuja forma singular termina ortograficamente em –L, itens cuja forma singular termina ortograficamente em –R e itens com plural regular. A escolha por nomes terminados em –L e –R justifica-se por demonstrarem uma oposição bem evidente no material fônico quando comparadas aos itens de plural regular no que tange aos processos morfofonológicos de formação do plural e, ao mesmo tempo, ocuparem posições diferenciadas na escala proposta por Scherre (1988). Dessa forma, incluímos na nossa tarefa experimental itens correspondentes à parte superior, média e inferior da hierarquia proposta, garantindo que eventuais efeitos escalares derivados dos processos de formação do plural possam ser verificados.

Os demais graus da escala hierárquica de SF propostos por Scherre (1988) não foram considerados no presente experimento. A exclusão de itens nominais com plurais duplos, por exemplo, justifica-se em função de seu número bastante reduzido na língua, além das dificuldades para sua representação e identificação gráfica, fatos que poderiam dificultar a preparação do material experimental. Já os itens nominais terminados em –ÃO, também não compõem a lista de estímulos experimentais em virtude de termos levado em consideração o levantado por Scherre (1988) e reforçado por Lopes e Scherre (2012), quanto à tendência dos falantes à não pluralização desses itens, devido a uma certa insegurança quando da escolha da forma correta dentre as três postuladas pela gramática tradicional. Ademais, não podemos nos esquecer de que avaliar experimentalmente todos os níveis da escala de SF proposta por Scherre (1988) seria inviável, pelo grande volume de itens que seria necessário incluir, além da dificuldade para manter os mesmos controles para todos os itens. Nesse sentido, visando a controlar melhor o número de condições, optamos por avaliar apenas os três níveis mencionados previamente. Todos os itens compartilhavam o mesmo padrão de

⁴⁵Hierarquia de saliência proposta por Scherre (1988, p.75-7):

Plural duplo (papelzinho/papeizinhos);
 Acréscimo de -S e mudança silábica em palavras terminadas em -L (quartel/quarteis);
 Acréscimo de -S e mudança silábica em palavras terminadas em –ÃO (seleção/ seleções);
 Acréscimo de -ES em palavras terminadas em -R (cor/cores);
 Acréscimo de -ES em palavras terminadas em -S (vez/vezes; freguês/fregueses);
 Acréscimo de -S em palavras terminadas em vogal ou vogal + nasal, os chamados plurais regulares (irmão/ irmãos; lei/ leis).

acentuação, ou seja, todos eram oxítonos. Escolhemos esse padrão silábico para os nomes utilizados nos estímulos experimentais por ser um dos dois padrões mais frequentes (= com maior número de itens) na língua, junto com os dissílabos paroxítonos.

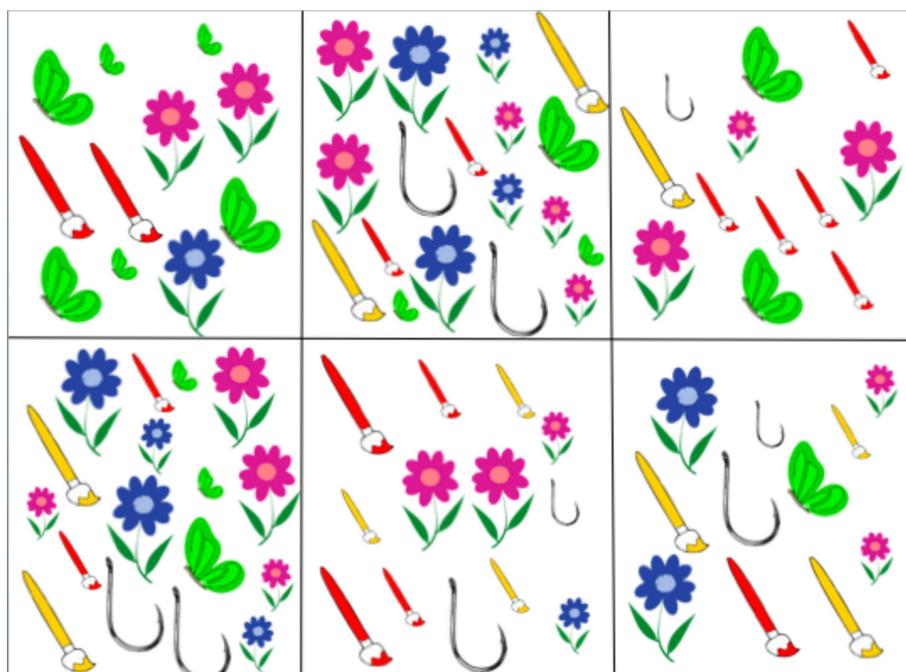
Com isso, para a construção dos estímulos experimentais, escolhemos cinco itens nominais com terminação em -R (*cobertor, colar, colher, flor, mulher*), cinco terminados ortograficamente em -L (*anel, anzol, caracol, jornal, pincel*) e cinco itens de plural regular (*caju, gambá, jacaré, maçã, sofá*). Como pode ser observado, quatro itens de cada grupo eram dissílabos e um era trissílabo. A motivação para essa variação se justifica apenas na facilidade para a representação e identificação dos itens visuais para os nossos participantes, principalmente as crianças. Vale destacar que, no Experimento 2, o padrão de acentuação e o número de sílabas não foram tomados como variáveis, sendo apenas controlados na elaboração dos estímulos. Assim como no Experimento 1, outros nomes distratores foram incluídos no material produzido. No entanto, diferentemente do Experimento 1, no Experimento 2, os nomes distratores não correspondiam aos mesmos padrões dos nomes alvo, sendo formados por itens paroxítonos, podendo ter duas, três ou quatro sílabas. Isso se motiva pelo fato de que a dimensão investigada restringia bastante o universo de nomes passíveis de serem representados por imagens e facilmente identificados pelas crianças. Por conta disso, não foi conduzida uma análise detalhada dos dados brutos nos moldes do que foi previamente apresentado para o Experimento 1.

Procedimento

A tarefa experimental foi concebida e conduzida nos mesmos moldes do Experimento 1 (cf. seção 4.2.1)

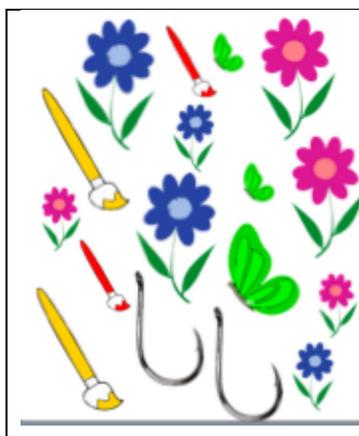
A diferença principal no que tange ao procedimento e materiais entre os Experimentos 1 e 2 consiste nas imagens que constavam nas pranchas e cartões utilizados (cf. figuras (3-4) a seguir) em função da apresentação de um novo conjunto de nomes alvo.

Figura 3 - Exemplo de prancha utilizada no Experimento 2



Fonte: Elaborado pela autora (2021)

Figura 4- Exemplo de cartão alvo utilizado no Experimento 2 – Sintagma-alvo para identificação do cartão: *Dois flores rosas pequenas*



Fonte: Elaborado pela autora (2021)

Assim como no primeiro experimento, cada sessão experimental foi gravada em áudio para posterior análise. Todas as sessões foram transcritas integralmente. Em média, cada sessão experimental com adultos durou aproximadamente 25 minutos e com as crianças, 37 minutos.

4.3.2 Resultados e discussão

Novamente, os dados obtidos foram transcritos, compilados e analisados estatisticamente por meio do programa R (R *Core Team*, versão 4.0.0, de 2020). Para o Experimento 2, foram conduzidos dois dos conjuntos de análises com objetivos e abordagens distintas – já especificados na seção 4.2.2 – a saber, análise não atomística dos sintagmas plurais contendo apenas nomes alvo e análise atomística dos sintagmas contendo apenas nomes alvo. Como já adiantado, para esse segundo experimento, não foi conduzida uma análise de dados brutos⁴⁶ mais detalhada em virtude de que nesse experimento, os nomes distratores não apresentavam as mesmas características fônicas dos nomes alvo. Passemos, então, aos resultados obtidos a partir da análise não atomística conduzida com os nomes alvo.

4.3.2.1 Análise 1 – SF dos nomes e marcação de plural – Análise não atomística – nomes alvo

Ao conduzirmos a análise não atomística considerando apenas os nomes alvo, observamos que dos 261 sintagmas plurais produzidos pelas crianças, em 107 (41%) observou-se uma marcação redundante, enquanto 154 (59%) apresentaram plural não redundante. Em contrapartida, no grupo controle (adultos), manteve-se a

⁴⁶A título de informação, apresentamos a primeira análise que considera todos os sintagmas nominais plurais produzidos pelos participantes durante a tarefa. Como disposto da tabela (A), de um total de 354 sintagmas plurais produzidos pelas crianças, foram registradas 159 ocorrências (45%) de sintagmas plurais redundantes e 195 (55%) ocorrências de sintagmas não redundantes. O grupo dos adultos produziu um total de 287 sintagmas plurais, dos quais 228 (79%) apresentaram marcação redundante, e 59 (21%) com marcação não redundante.

Tabela A - Distribuição dos sintagmas plurais em função da regra de concordância aplicada (redundante X não redundante) no grupo experimental (crianças) e controle (adultos) – Experimento 2 – dados brutos

Grupo	Padrão de concordância No total de ocorrências e %		Qui quadrado
	Redundante <i>Ex. dois bonés</i>	Não Redundante <i>Ex. cinco caracol</i>	
Crianças (Grupo experimental)	159 (45%)	195 (55%)	$p=0.06$
Adultos (Grupo controle)	228 (79%)	59 (21%)	$p<0.01$

Fonte: Elaborado pela autora (2021)

predominância de produção de sintagmas com marcação redundante (74%), em detrimento dos sintagmas não redundantes (26%), como pode ser visto na tabela (17) abaixo:

Tabela 17 - Distribuição dos sintagmas plurais em função do padrão de concordância aplicado (redundante X não redundante) no grupo experimental (crianças) e controle (adultos) – Experimento 2 – nomes alvo

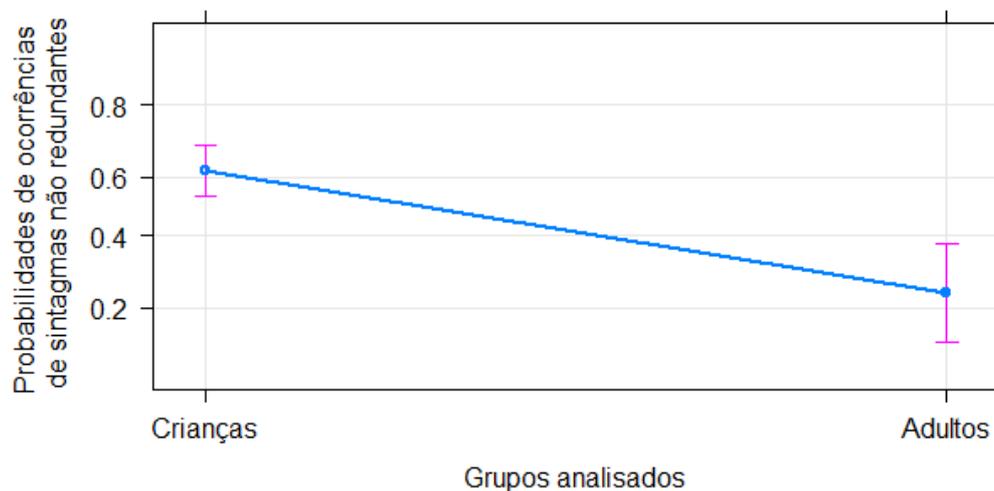
Grupo	Padrão de concordância No total de ocorrências e %		Qui quadrado
	Redundante	Não Redundante	
Crianças (Grupo experimental)	107 (41%)	154 (59%)	$p < 0.01$
Adultos (Grupo controle)	154 (73%)	57 (27%)	$p < 0.01$

Fonte: Elaborado pela autora (2021)

Novamente, o tratamento estatístico revelou diferença significativa entre as proporções de sintagmas plurais redundantes e não redundantes produzidos pelas crianças ($\chi^2 = 8.1073(1)$, $p = 0.1$), bem como pelos adultos ($\chi^2 = 43.678(1)$, $p < 0.01$). Neste caso, ao contrário do observado nos dados dos adultos, as crianças demonstraram preferência pelo padrão não redundante de concordância.

Em virtude da diferença de proporções de sintagmas não redundantes entre os dados das crianças (59%) e dos adultos (27%), foi aplicado um modelo de regressão logística com o objetivo de verificar se tal diferença seria estatisticamente significativa. O resultado do modelo, bem como a não sobreposição dos intervalos de confiança observados no gráfico de efeitos (7), indicam que a diferença é significativa.

Gráfico 7 - Gráfico de efeitos da produção de sintagmas plurais não redundantes em função do grupo analisado (Crianças X Adultos) – Experimento 2 – nomes alvo



Fonte: Elaborado pela autora (2021)

O próximo passo foi analisar os dados por meio de um modelo de regressão logística em função da dimensão *tipos de processos morfofonológicos de formação de plural* investigada. A tabela (18) abaixo apresenta os dados registrados em termos absolutos e percentuais no que diz respeito à distribuição de sintagmas plurais redundantes e não redundantes em função dos diferentes níveis da dimensão considerados.

Tabela 18 - Distribuição dos sintagmas plurais em função dos tipos de processos de formação de plural (-R X -L X plural regular) no grupo experimental (crianças) e controle (adultos) Experimento 2 – nomes alvo

Grupo	Tipos de processos de formação de plural	Padrões de concordância No total de ocorrências e %		Total de ocorrências
		Concordância redundante	Concordância não redundante	
Crianças	Nomes			
	Plural regular <i>Ex. dois cajus</i>	35 (39%)	58 (61%)	93
	Terminados em -R <i>Ex. Duas colheres</i>	32 (40%)	49 (60%)	81
	Terminados em -L <i>Ex. três anzois</i>	40 (46%)	47 (54%)	87

Adultos (controle)	Plural regular	53 (77%)	16 (23%)	69
	Terminados em -R	50 (67%)	25 (33%)	75
	Terminados em -L	51 (76%)	16 (24%)	67

Fonte: Elaborado pela autora (2021)

Os resultados dos modelos de regressão conduzidos indicam que, tanto no grupo das crianças, quanto dos adultos, não houve efeito de *tipo de processo*, comparando os itens terminados ortograficamente em -R e -L, em relação aos itens de plural regular, tomados aqui como nosso valor de referência. Afinal, apesar da estimativa de ocorrências de sintagmas plurais não redundantes diferir frente aos níveis da dimensão *processos*, tal diferença não chega a ser significativa, como se pode observar na tabela de coeficientes (19), abaixo.

Tabela 19 - Tabela de coeficientes extraída do modelo de regressão logística para ocorrências de sintagmas plurais não redundantes em função dos tipos de processos de formação de plural (Crianças e Adultos) –

Crianças				
	Estimate	Std. Error	z	p
Intercept (Plural regular)	0.50509	0.21404	2.360	0.01
-R	-0.07901	0.31220	-0.253	0.80
-L	-0.34383	0.30346	-1.133	0.25
Adultos				
	Estimate	Std. Error	z	p
Intercept (Plural regular)	-1.19770	0.28525	-4.199	2.68e-05
-R	0.50456	0.37599	1.342	0.18
-L	0.038471	0.40432	0.095	0.92

Fonte: Elaborado pela autora (2021)

Vale lembrar que, assim como no Experimento 1, nas análises 1 e 2 investigamos os efeitos da SF na marcação de plural no sintagma considerado globalmente. No entanto, em virtude de nosso interesse por investigar os efeitos das propriedades fônicas dos itens nominais na marcação de plural presente (ou não) no próprio item nominal, pensamos ser necessário conduzirmos também a análise atomística, a qual apresentaremos a seguir.

4.3.2.2 Análise 2 – SF dos nomes e marcação de plural – Análise atomística – nomes alvo

Na análise atomística, como já estabelecido para o Experimento 1, foram considerados apenas os sintagmas não redundantes, os quais foram divididos em função da presença ou não de marcação no nome. Com isso, dispostos na tabela (20), observamos um total de 154 sintagmas não redundantes produzidos pelo grupo de crianças, dos quais 58 (38%) apresentaram marcação de plural no nome, enquanto 96 (62%) apresentaram nomes não marcados. O grupo controle, por sua vez, produziu 57 sintagmas não redundantes, dos quais 54 (95%) tinham nomes marcados para plural, e apenas 03 (5%) não apresentaram marcação de plural no nome. O resultado dos testes de Qui-quadrado indicam uma diferença significativa entre as proporções de sintagmas com nomes marcados X não marcados para plural produzidos, tanto nos dados das crianças ($\chi^2 = 8.8896$ (1), $p < 0.01$), quanto nos dados dos adultos ($\chi^2 = 43.86$ (1), $p < 0.01$).

Tabela 20 - Distribuição dos sintagmas plurais não redundantes em função da marcação ou não marcação de plural no nome (nome marcado X nome não marcado) no grupo experimental (Crianças) e controle (Adultos)

Grupo	Marcação de plural no nome No total de ocorrências e %		Qui quadrado
	Nome marcado	Nome não marcado	
Crianças (Grupo experimental)	58 (38%)	96 (62%)	$p < 0.01$
Adultos (Grupo controle)	54 (95%)	03 (5%)	$p < 0.01$

Fonte: Elaborado pela autora (2021)

O resultado de um teste de Qui-quadrado confirmou que a diferença entre as proporções de sintagmas não redundantes produzidas por ambos os grupos é significativa ($\chi^2 = 52.146$ (1), $p < 0.01$), com um percentual maior de ocorrências de nomes não marcados na produção das crianças.

Em seguida, analisamos, por meio de modelos de regressão logística, a distribuição de marcação de plural no nome em sintagmas não redundantes em função da dimensão *processos morfofonológicos de formação de plural* – a qual pode ser

observada na tabela (21), abaixo. Os resultados dos modelos não registraram nenhum efeito significativo para as crianças, nem para os adultos.

Tabela 21 - Distribuição de marcação de plural no nome em sintagmas não redundantes em função dos tipos de processos de formação de plural (-R X -L X plural regular) no grupo experimental (Crianças) e controle (Adultos)

Grupo	Tipos de processos de formação de plural	Marcação de plural no nome		Total de ocorrências
		No total de ocorrências e %		
	Nomes	Nome marcado	Nome não marcado	
Crianças	Plural regular	17 (29%)	41 (71%)	58
	Terminados em -R	22 (45%)	27 (55%)	49
	Terminados em -L	19 (40%)	28 (60%)	47
Adultos (controle)	Plural regular	14 (88%)	02 (12%)	16
	Terminados em -R	25 (100%)	0 (0%)	25
	Terminados em -L	15 (94%)	01(6%)	16

Fonte: Elaborado pela autora (2021)

Em resumo, as análises conduzidas a partir dos dados coletados neste segundo experimento nos permitiram observar que:

- a) Como observado no Experimento 1, novamente o grupo de adultos apresentou um comportamento semelhante no que tange a uma produção de sintagmas redundantes significativamente maior do que não redundantes.
- b) No caso das crianças, desta vez, a produção revelou uma quantidade significativamente maior de sintagmas não redundantes.
- c) De modo geral, os resultados do Experimento 2, assim como os resultados do primeiro experimento, distanciam-se do observado nos dados naturalísticos que sugeriam um comportamento uniforme entre os grupos;
- d) Não houve efeitos da variável *tipos de processos morfofonológicos de formação de plural* na alternância dos padrões de concordância quando contrastados os três tipos de processos investigados. Em outras palavras, contraas previsões, nomes terminados ortograficamente em –

L, em –R e itens regulares apresentaram proporções semelhantes de marcação não redundante em ambos os grupos considerados.

Considerados em conjunto, os dados obtidos em ambas as tarefas experimentais não se mostram compatíveis com a hipótese de que a SF tenha uma influência clara na alternância dos padrões de concordância nominal de número no PB – pelo menos ao se considerar a variável nos moldes mais difundidos no âmbito da sociolinguística. De fato, surpreendentemente, a dimensão *processos morfofonológicos de formação de plural*, certamente o aspecto mais “pacífico” da SF, foi a dimensão que de forma mais consistente não se mostrou relevante no nosso estudo experimental.

Cumprido destacar que os resultados não alinhados às previsões aqui reportados espelham de certa forma o que vem sendo reportado na literatura sociolinguística quanto às dimensões da SF investigadas nos experimentos. Como vimos no Capítulo 2 desta tese, os estudos oscilam entre resultados que indicam a atuação de apenas um dos fatores e resultados que não apontam efeitos, chegando até mesmo a questionar um possível efeito da SF de modo mais geral. Scherre (1988), por exemplo, não obteve efeito de número de sílabas, enquanto Andrade (2003) e Tabosa (2016) informam que a tonicidade não se mostrou relevante em seus dados. Já Carvalho (1997), observa uma correlação positiva com a concordância variável, mas com destaque para os processos morfofonológicos. Campos (2015), por sua vez, bem como Azalim et al (2018), defendem que a SF não exerceria uma influência significativa sobre o comportamento linguístico dos falantes do PB com maior nível de escolaridade. Campos (2015) e Azalim et al (2018) ainda reforçam a hipótese defendida por Guy (1981), Scherre (1988) e Lopes e Scherre (2012) de que a tonicidade pode ser mais relevante para a realização da forma de concordância redundante do que os processos de formação de plural.

Já dentre os trabalhos mais recentes que apontam para a relevância do papel da SF na marcação no âmbito nominal, podemos citar Oushiro (2016), Oushiro e Guy (2015), Ribeiro (2013), Roza e Casagrande (2015). No entanto, vale ressaltar que esses últimos levaram em consideração somente os processos de formação de plural em sua análise da SF.

A pesquisa de base experimental conduzida por Reis (2020), por sua vez, traz resultados compatíveis com o reportado na presente tese no que tange a uma ausência de efeitos de SF – para os participantes adultos – e de efeitos “às avessas” no caso das

crianças.

Quando considerada a concordância verbal, também encontramos resultados divergentes. No que diz respeito às pesquisas de cunho descritivo, de um lado, temos autores como Nicolau (1984, 1995), que sugerem que os processos de formação de plural não teriam um papel relevante na realização da concordância redundante de número. Por outro lado, Molina (2018) reporta o contrário, com dados que sugerem inclusive que a SF – em termos da dimensão processos – teria uma influência mais clara na produção das crianças do que na dos adultos. A pesquisa de cunho experimental conduzida por Scher (2021), por sua vez, não identificou efeitos de SF na marcação da concordância no verbo.

Considerando esses estudos em conjunto, observa-se que a SF ainda é uma questão problemática e que maiores esclarecimentos acerca de seu papel na concordância variável ainda são necessários. Como colocado por Scher (2021), os achados reportados na literatura parecem demandar uma revisão das definições e classificações atualmente disponíveis.

Em síntese, os resultados – incluindo os nossos próprios – quanto à atuação das variáveis *número de sílabas*, *padrão de acentuação* e *processos morfofonológicos de formação de plural*, se mostram, no mínimo, instáveis e de escassa replicabilidade entre os diferentes estudos.

Embora o estudo da SF tenha sido o ponto de partida para os dois experimentos elaborados para nessa pesquisa, os dados obtidos fornecem uma fonte de informações bastante rica sobre outros aspectos para além dessa variável em particular. Assim sendo, foram levantadas outras questões importantes relativas a aspectos que parecem se mostrar relevantes para compreender a alternância de padrões de concordância nominal, tais como o contexto morfossintático, a posição linear do item dentro do sintagma e a presença de numerais. Diante disso, foram conduzidas duas análises complementares dos dados coletados nos Experimentos 1 e 2 que reportamos a seguir.

4.4 ANÁLISES COMPLEMENTARES DOS DADOS EXPERIMENTAIS: INVESTIGANDO A CONFIGURAÇÃO ESTRUTURAL DO SINTAGMA E A FONOLOGIA DOS NUMERAIS

Como vimos anteriormente, a utilização da técnica de produção semi-eliciada a partir de imagens por meio de uma tarefa interativa entre duplas de participantes se revelou bastante produtiva e permitiu coletar um número expressivo de dados. Assim sendo, o material obtido se configura como uma valiosa fonte de informações sobre o assunto mais amplo da nossa pesquisa, qual seja, a concordância nominal variável no PB. Nesse sentido, consideramos pertinente ir além do que foi inicialmente traçado como objetivos primários para os Experimentos 1 e 2 – i.e. investigar as três dimensões da SF– e explorar outros pontos relevantes para a discussão teórica que será traçada no Capítulo 5.

A seguir, reportamos os resultados das análises conduzidas a fim de verificarmos em que medida a configuração dos sintagmas, bem como a fonologia dos numerais interfere na alternância de padrões de concordância nominal no PB.

4.4.1 Configuração dos sintagmas e marcação interna de plural

Os dados brutos⁴⁷ obtidos nos Experimentos 1 e 2 foram agrupados em função da estrutura dos sintagmas plurais efetivamente produzidos. Para fins de análise, foram consideradas as estruturas que apresentavam as seguintes configurações:

- (i) Num(eral) + N(ome) como em *Cinco pinceis*;
- (ii) Num(eral) + N(ome) + A(djetivo), tais como em *Cinco pinceis vermelhos*
e
- (iii) Num(eral) + N(ome) + A(djetivo)1 + A(djetivo)2, como em *Cinco pinceis vermelhos grandes*.

⁴⁷ Cabe esclarecer que para as duas próximas análises reportadas, que investigam a produção de sintagmas plurais redundantes e não redundantes em função (i) do número de elementos nos sintagmas e marcação interna de plural; e (ii) da morfofonologia dos numerais, foram considerados todos os sintagmas plurais encabeçados por numeral produzidos no conjunto de dados brutos, uma vez que para essas análises, não estamos considerando o nível de saliência dos itens nominais que compõem os sintagmas.

Como se pode ver, desta vez, diferentemente da análise conduzida a partir dos dados naturalísticos, os sintagmas analisados são todos iniciados por Numeral, em virtude da não ocorrência de sintagmas iniciados por determinante durante a tarefa experimental⁴⁸. A tabela (22) abaixo apresenta a distribuição dos sintagmas plurais quando considerado o *tipo de estrutura analisada* (Num + N X Num + N + A X Num + N + A1 + A2) e marcação da concordância no sintagma (redundante X não redundante) no estudo experimental.

Tabela 22 - Distribuição dos sintagmas plurais em função do tipo de estrutura analisada (Num + N X Num + N + A X Num + N + A1 + A2) no grupo experimental (crianças) e controle (adultos) – dados brutos

Grupo	Tipos de estrutura	Padrões de concordância No total de ocorrências e %		Total de ocorrências
		Concordância redundante	Concordância não redundante	
Crianças	Num+N	186 (74%)	63 (26%)	249
	Num+N+A	138 (35%)	253 (65%)	391
	Num+N+A1+A2	58 (47%)	66 (53%)	124
Adultos (controle)	Num+N	109 (99%)	01 (1%)	110
	Num+N+A	283 (84%)	54 (16%)	337
	Num+N+A1+A2	114 (70%)	49 (30%)	163⁴⁹

Fonte: Elaborado pela autora (2021)

⁴⁸ Vale frisar que o uso de um D definido envolveria a retomada de um referente previamente introduzido no discurso. Já um D indefinido requereria a apresentação de um referente pela primeira vez. Ambas as opções se mostram pragmaticamente inadequadas considerando a tarefa apresentada para os participantes. A elaboração da atividade experimental utilizada colocou em evidência a enorme dificuldade de se criar uma situação comunicativa na qual o uso de DPs plurais fosse natural e pragmaticamente feliz.

⁴⁹ A diferença observada entre os valores dispostos na nota (40), na tabela (A) e na tabela (22) reside no fato de que na tabela (22) acima, foram considerados apenas os sintagmas encabeçados por numerais.

O teste de Qui-quadrado conduzido aponta para uma predominância estatisticamente significativa do uso de marcação explícita de plural nas estruturas analisadas, seja no caso de estruturas Num+N na produção infantil ($\chi^2=57.831$, (1), $p < 0.01$), seja no caso da produção adulta em todas as estruturas encabeçadas por numeral– Num+N ($\chi^2 = 104.08$, (1), $p < 0.01$), Num+N+Adj ($\chi^2= 154.26$, (1), $p < 0.01$) e Num+N+Adj1+Adj2 ($\chi^2 = 25.129$, (1), $p < 0.01$). Tal resultado se mostra bastante relevante no sentido de que ele parece apontar para o fato de que não somente a informação de número veiculada pelo determinante é diferente daquele veiculada pelo numeral, como também que aquela veiculada por D parece ser mais confiável na garantia da leitura plural da estrutura. As únicas estruturas que vão na direção oposta são Num+N+Adj e Num+N+Adj1+Adj2 (células destacadas na tabela), ambas na produção infantil. No entanto, vale destacar que o percentual de sintagmas não redundantes foi significativamente maior apenas na estrutura Num+N+Adj ($\chi^2=33.238$ (1), $p < 0.01$).

A partir desse levantamento, analisamos, por meio de um modelo de regressão logística, os dados em função do tipo de estrutura investigado para testar se a aparente diferença observada entre as estimativas de ocorrências não redundantes entre os tipos de estruturas, seria estatisticamente significativa. O resultado dos modelos de regressão conduzidos para adultos e crianças indicaram que para ambos os grupos, a diferença foi significativa do ponto de vista estatístico, com uma produção significativamente maior de ocorrências não redundantes em sintagmas com três e quatro elementos em quando comparados aos sintagmas formados por dois elementos, como indica a tabela de coeficientes (23) abaixo.

Tabela 23 - Tabela de coeficientes extraída do modelo de regressão logística para ocorrências de sintagmas plurais não redundantes em função dos tipos de estrutura - Experimento 2 – nomes alvo

Crianças				
	Estimate	Std. Error	z	p
Intercept (Num+N)	-1.0826	0.1458	-7.427	<0.01
Num+N+A	1.6887	0.1801	9.375	<0.01
Num+N+A1+A2	1.2118	0.2316	5.232	<0.01
Adultos				
	Estimate	Std. Error	z	p
Intercept (Num+N)	-4.691	1.005	-4.670	<0.01

Num+N+A	3.035	1.015	2.989	<0.01
Num+N+A1+A2	3.847	1.019	3.775	<0.01

Fonte: Elaborado pela autora (2021)

Nesse sentido, os resultados de ambos os grupos nos sugerem que o tipo de configuração do sintagma parece exercer algum efeito na realização dos padrões de concordância. Ou seja, quando se estabelece uma comparação em termos de número de elementos morfológicamente realizados, percebe-se que os sintagmas formados por até dois elementos exibem uma consistência maior no que se refere à marcação de concordância em relação aos sintagmas formados por três elementos. Em outros termos, o que se observa é que quanto maior é a estrutura em termos de elementos realizados, mais propensa ela se mostra a apresentar variabilidade na marcação.

Diante desse resultado, relacionado à configuração de cada tipo de sintagma, conduziu-se ainda uma análise em função dos padrões específicos de marcação no interior de cada estrutura. Como pode ser visto na tabela (24) abaixo, semelhante ao que foi observado nos dados naturalísticos, o padrão redundante se mostrou predominante nas produções de ambos os grupos, com exceção – como visto na análise anterior – das estruturas formadas por Num+N+Adj e por Num+N+Adj1+Adj2 produzidas pelas crianças, em que se observa o padrão inverso.

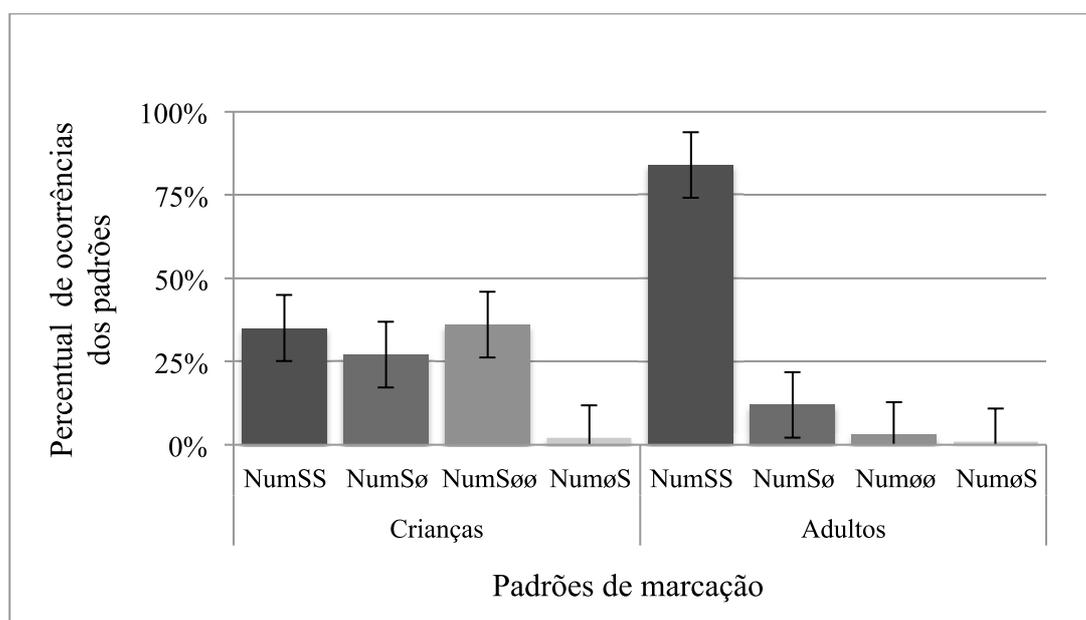
Tabela 24 - Distribuição dos sintagmas plurais em função dos padrões de marcação internos ao sintagma nas duas tarefas experimentais – dados brutos

Grupo	Tipos de estrutura	Padrões de concordância					Total de ocorrências
		No total de ocorrências e %					
		Concordância redundante	Concordância não redundante				
Crianças	Num+N	NumS 186 (74%)	NumØ 63(26%)				249
	Num+N+A	NumSS 138 (35%)	NumSØ 105 (27%)	NumØØ 141 (36%)	NumØS 07 (2%)		391
	Num+N+A1 +A2	NumSSS 58 (47%)	NumSSØ 15 (12%)	NumSØØ 27 (22%)	NumØØØ 27 (22%)	NumSØS 02 (2%)	NumØSS 02 (2%)
Adultos (controle)	Num+N	NumS 109 (99%)	NumØ 01 (1%)				110
	Num+N+A	NumSS 283 (84%)	NumSØ 41 (12%)	NumØØ 10 (3%)	NumØS 03 (1%)		337
	Num+N+A1 +A2	NumSSS 114 (70%)	NumSSØ 31 (19%)	NumSØØ 16 (10%)	NumØØØ 02 (1%)	--	

Fonte: Elaborado pela autora (2021)

No que tange aos sintagmas maiores– de 3 elementos –, ao observarmos a tabela (24) acima, e o gráfico (8) constatamos divergências na comparação entre crianças e adultos em relação aos padrões produzidos.

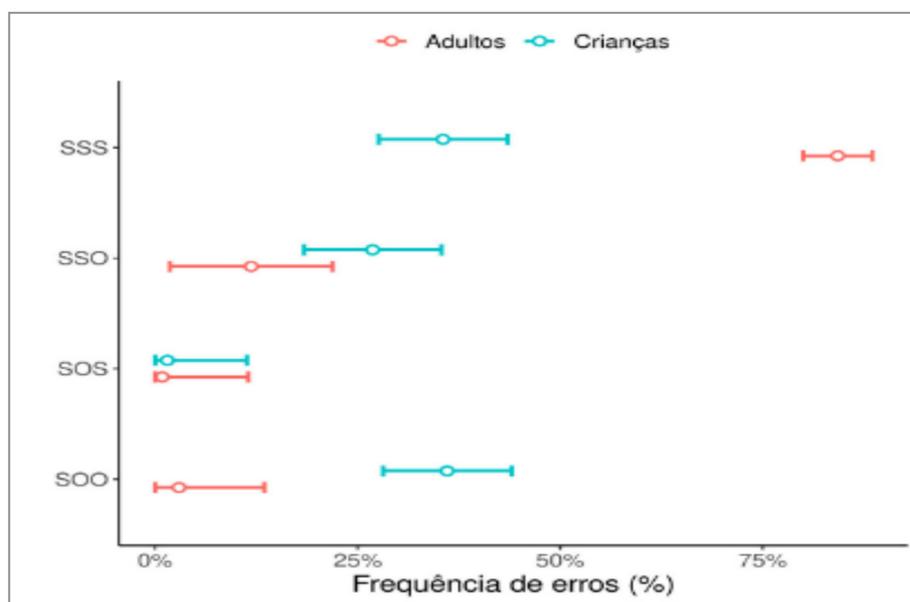
Gráfico 8 - Distribuição de padrões de marcação internos registrados nas estruturas Num+N+Adj nas duas atividades experimentais– dados brutos



Fonte: Elaborado pela autora (2021)

Em virtude dos resultados diferenciados para os dois grupos, aplicamos um teste de Intervalos de confiança no intuito de verificar se a aparente diferença entre as proporções dos padrões não redundantes em ambos os grupos é significativa em termos estatísticos (cf. Gráfico 9).

Gráfico 9- Resultado do Teste de Intervalo de Confiança conduzido em função dos padrões de marcação internos na estrutura (Crianças e Adultos) – dados brutos – Experimentos 1 e 2



Fonte: Elaborado pela autora (2021)

Os resultados do teste sugerem que as crianças apresentam um contraste equilibrado entre marcação total (35%) e marcação mínima (36%), indicando uma polarização: ora o sintagma apresenta marca em todos os elementos, ora não apresenta nenhuma marca. Nesse sentido, só parece haver diferença significativa entre as proporções de estruturas do tipo NumØS por um lado, e todos os outros padrões, por outro. Uma possível interpretação para tal comportamento diz respeito ao fato de que as crianças ainda estão adquirindo os padrões de marcação licenciados na língua. Sendo assim, faz sentido apresentarem de forma mais consistente os dois polos do contraste, ou seja, partem do contraste mais simplificado (tudo ou nada) e aos poucos vão refinando e assimilando os demais padrões – o que se explica pelo percentual intermediário de marcação representada pela estrutura NumSØ (27%). No entanto, apesar das divergências, é fundamental notar que a estrutura com padrão NumØS (marcação na terceira posição), tal como em *quatro livro azuis*, apresenta uma proporção de ocorrências muito baixa (2%) e (1%) de ocorrências nos dados das crianças e adultos, respectivamente. Nesse sentido, enquanto não obtivermos evidências em outra direção, a melhor interpretação que podemos tecer para justificar esses dados seria interpretar tais produções como possíveis lapsos de fala em ambos

os grupos. Importante ressaltar que consideramos que o primeiro elemento do sintagma (Numeral) não entra nesse “jogo” de marcação X não marcação uma vez que apresenta essa peculiaridade de carregar de forma intrínseca uma carga semântica de pluralidade. Na próxima seção, no entanto, discutiremos melhor as possíveis relações do numeral com a marcação interna do sintagma em uma perspectiva formalista de análise.

Os dados dos adultos, por outro lado, sugerem uma hierarquização diferente entre os padrões identificados, indicando que a lógica seguida pelos adultos seria diferente daquela usada pelas crianças. Enquanto a “regra” adotada pelas crianças parece ser um contraste entre marcação total X marcação mínima, os dados dos adultos sugerem que eles seguem uma gradação que vai da marcação total, seguida de marcação só no nome e, finalmente, a marcação mínima (na qual a pluralidade é representada apenas pelo numeral). Outro aspecto interessante observado nos dados dos adultos é que, como visto no gráfico (9), parece haver diferença significativa somente no contraste entre as proporções da estrutura do tipo NumSS (redundante) por um lado, e todos os padrões não redundantes, por outro. A sobreposição dos intervalos de confiança sugere que não houve diferença significativa ao compararmos os padrões não redundantes entre si.

A seguir, fornecemos algumas análises específicas sobre os numerais e sua possível influência na alternância de padrões de marcação da concordância nominal.

4.4.2 Fonologia dos numerais e marcação de plural no sintagma

Em virtude das produções coletadas nos Experimentos 1 e 2 consistirem em sua maioria de sintagmas formados por Numeral + N + (Adj/s), despertou-nos o interesse por investigar – assim como feito com os dados naturalísticos – se a constituição fonológica dos numerais teria algum efeito na realização dos padrões de concordância. Em outras palavras, se a terminação em -S, encontrada em numerais como *dois*, *três* e *seis* (diferente de *quatro* e *cinco*, por exemplo), poderia desempenhar algum papel na marcação de concordância, como resultado de uma possível falsa analogia entre tal marca e o morfema de plural -S dos nomes. Se assim for, poderíamos ter um indício da influência da variável linguística conhecida na literatura como paralelismo formal. Tal fator, prevê que marcas levam a marcas e zeros conduzem a zeros. Nesse sentido, objetivou-se verificar se haveria um número

maior de sintagmas redundantes quando da presença dos numerais “falsos plurais” do que com aqueles não terminados em -S, tais como *quatro*, *cinco* e *sete*.

Ao analisarmos os dados brutos em função dos processos de formação fonológica dos numerais por meio de um modelo de regressão logística, observou-se que os dados experimentais reforçam os resultados obtidos nos dados naturalísticos. Ou seja, que a forma superficial dos numerais não se mostrou relevante quando da aplicação dos padrões de concordância nos sintagmas formados por Numeral + N (seguido ou não de adjetivo) para os adultos, que apresentaram percentuais de produções não redundantes equivalentes para ambos os tipos de processos de formação morfológica dos numerais. No caso das crianças, o contraste entre os dois tipos de numerais registrou uma diferença próxima do nível de significância, com um número maior de sintagmas não redundantes para os “falsos plurais”, como pode ser visto na tabela (25), abaixo:

Tabela 25 - Distribuição dos sintagmas plurais em função dos processos de formação morfológica dos numerais (Num + N X Num-S + N) no grupo experimental (crianças) e controle (adultos) – dados brutos

Grupo	Morfofonologia dos numerais	Padrões de concordância		Total de ocorrências
		Concordância redundante	Concordância não redundante	
Crianças	Num+N (+A) Ex. <i>cinco balas (azuis)</i>	130 (55%)	106 (45%)	236
	Num-S +N (+A) Ex. <i>dois livros (verdes)</i>	252 (48%)	276 (52%)	528
Adultos	Num+N (+A) Ex. <i>quatro gambás (laranjas)</i>	91 (83%)	18 (17%)	109
	Num-S +N (+A) Ex. <i>três bonecas (vermelhas)</i>	416 (83%)	87 (17%)	503

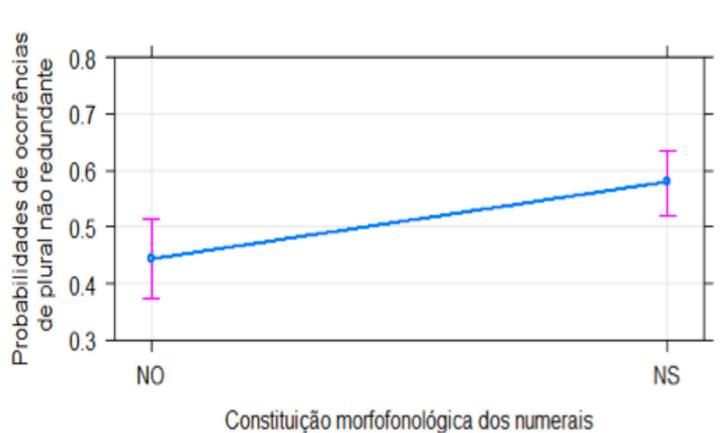
Fonte: Elaborado pela autora (2021)

Os dados sugerem que, pelo menos para os adultos – semelhante ao reportado por Scherre (1988) e Andrade (2003) entre outros –, o paralelismo formal não se mostrou relevante na realização dos padrões de concordância neste tipo de sintagma. Nesse sentido, esse resultado parece mostrar que, ainda que haja coincidência fonológica entre o -S que compõe o final desses numerais e o -S que efetivamente

marca plural nos nomes, os adultos parecem perceber o diferente estatuto morfológico desses elementos que, respectivamente, ora são parte da raiz e ora são morfemas independentes dela.

O resultado obtido na análise dos dados infantis sugere uma possível sensibilidade maior por parte das crianças – pelo menos nessa faixa etária – frente a essa marcação de pluralidade. Nesse sentido, quando o numeral apresenta uma “falsa analogia” de plural, os dados sugerem que as crianças interpretam o –S do numeral como um morfema de plural, fato que acarreta uma maior omissão do plural no nome (52%). Por outro lado, quando o numeral não é “transparente” fonologicamente (ou seja, quando não termina em –S), as crianças produziram significativamente mais marcas nos nomes. Assim, uma aparente morfologia de plural no numeral licenciaria a omissão das marcas no nome, de forma semelhante ao que ocorre quando um determinante plural encabeça o sintagma. A diferença observada quando comparados os numerais falsamente transparentes e aqueles não transparentes é compatível com uma sensibilidade por parte das crianças à informação de natureza morfológica e com o fato de que as mesmas já identificaram as principais restrições que operam no sistema de concordância da língua. Quando ocorre a falsa analogia, as crianças parecem “confiar” em que há informação suficiente de pluralidade expressa no numeral. Tal interpretação encontra sustentação no gráfico de efeitos (10) abaixo, no qual a não sobreposição dos intervalos de confiança nos permite inferir que há diferença significativa entre as produções não redundantes em função da morfofonologia dos numerais.

Gráfico 10 - Gráfico de efeitos em função da morfofonologia dos numerais na distribuição de sintagmas plurais não redundantes – grupo experimental - (crianças) – Experimentos 1 e 2



Fonte: Elaborado pela autora (2021)

Dessa forma, percebemos que, embora semanticamente todos os numerais – exceto por *zero* e *hum*–, sejam intrinsecamente plurais, essa pode não ser uma informação explícita e/ou robusta o suficiente para as crianças da faixa etária avaliada.

Na próxima seção, retomamos um ponto que foi apenas discutido superficialmente ao longo das seções anteriores, qual seja, as diferenças observadas nas produções de adultos e crianças quando comparados os dados naturalísticos e experimentais.

4.5 DIFERENÇAS ENTRE A PRODUÇÃO ADULTA E INFANTIL: DADOS NATURALÍSTICOS X EXPERIMENTAIS

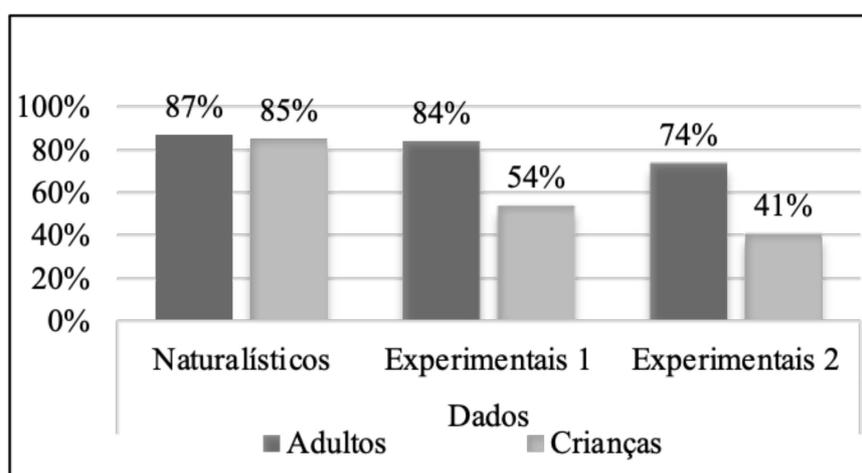
A partir do que foi exposto até aqui, pudemos notar que a articulação de dados de naturezas distintas se mostrou fundamental para a nossa pesquisa. Primeiramente, em virtude de que buscamos investigar aqui aspectos para os quais apenas dados naturalísticos ou experimentais poderiam ser de fato informativos. Por exemplo, de um lado, temos os dados experimentais que fornecem um panorama bastante detalhado no tocante às dimensões da SF – com número controlado de possíveis ocorrências entre os níveis das variáveis manipuladas – mas que não fornecem todos

os tipos de estruturas que pretendíamos analisar. Em contrapartida, os dados longitudinais trazem um panorama muito mais rico no que diz respeito à diversidade de tipos de estruturas, mas são limitados para a investigação da SF, não fornecendo uma amostra suficientemente diversificada quanto aos tipos de nomes produzidos.

Além disso, como pode ser visto no gráfico (11) abaixo, a comparação entre dados naturalísticos e experimentais da produção de crianças e adultos, nos permitiu observar, por exemplo, que nos dados de produção espontânea, o comportamento das crianças parece refletir de forma fiel o padrão encontrado na fala adulta, com uma produção significativamente maior de sintagmas plurais redundantes para ambos os grupos.

Em contrapartida, os resultados encontrados em ambos os experimentos sugerem que, embora a alternância dos padrões de concordância tenha ocorrido tanto na produção das crianças quanto dos adultos, foram registradas diferenças importantes entre os dois grupos. Enquanto os adultos demonstraram uma preferência pela marcação redundante ($p < 0.01$) – tal como observado nos dados naturalísticos –, as crianças produziram um número equivalente de sintagmas redundantes e não redundantes ($p = 0.09$). Ao observarmos especificamente os resultados do Experimento 2, encontramos o mesmo contraste reportado por Reis (2020): uma maior produção de plural redundante nos dados dos adultos, ao passo que as crianças produziram um maior percentual de marcação não redundante, como visto no gráfico (11) abaixo.

Gráfico 11- Panorama geral da produção de sintagmas plurais redundantes nos dados naturalísticos e experimentais



Fonte: Elaborado pela autora (2021)

Como vimos, uma possível interpretação das diferenças encontradas entre adultos e crianças estaria associada a questões interacionais. Cekaite et al (2014), defendem a ideia de que as habilidades dialógicas das crianças pequenas estejam relacionadas aos seus contextos culturais e situacionais de uso. Sendo assim, era de se esperar diferenças na produção quando comparadas interações do tipo “adulto-criança” – tal como ocorre nos dados naturalísticos – e “criança-criança” – como acontece nos dados experimentais. Nesse sentido, as crianças analisadas nos *corpora* apresentam comportamento compatível com a ideia defendida por Nelson (2014) segundo a qual as crianças mais jovens, na busca por estabelecer seus próprios padrões de linguagem, tendem a contar com a orientação de adultos, muitas vezes, inclusive, refletindo o seu comportamento, tal como observado na análise dos dados naturalísticos. Isso porque, de acordo com Cekaite et al (2014) e Grøver Aukrust (2004), dentre outros, na interação com os adultos, estes tendem a assumir a maior responsabilidade na conversa, além de apoiar o fluxo da conversa ao iniciar os assuntos por meio de perguntas, fornecendo interpretações e desafiando as crianças a observarem a relevância de determinados tópicos. Nesse sentido, é comum observarmos em dados obtidos em contextos de interação “adulto – criança”, aspectos centrais desse tipo de interação, tais como uma assimetria em termos de poder, habilidades e conhecimento que, por sua vez, tendem a conduzir a fala da criança na direção de um modelo normativo presente na fala adulta. (CEKAITE ET AL, 2014). Isso porque a interação do tipo “adulto-criança” pode oferecer suporte conversacional, bem como servir de modelo para uma participação comunicativa semelhante à fala adulta, viabilizando um desempenho conversacional apropriado, tanto em termos culturais, quanto cognitivos. No entanto, concordamos com Blum-Kulka e Snow (2004) na ideia de que tal suporte pode gerar restrições ao comportamento mais espontâneo das crianças, uma vez que a superioridade dos adultos pode intimidar os falantes mais novos de expressarem suas ideias.

Por outro lado, nossos dados experimentais fornecem indícios de que a fala entre pares (*peer talk*) possa ser vista, como apontado por Cekaite et al (2014), como uma oportunidade de desenvolvimento das habilidades pragmáticas das crianças em sua língua. Os dados sustentam ainda a ideia de que a fala entre pares apresente a vantagem de ser caracterizada por uma estrutura de participação relativamente igualitária e colaborativa entre as partes – característica raramente presente na interação adulto-criança. Em virtude dessas características, compartilhamos com

Blum-Kulka e Snow (2004) a ideia de que a fala entre pares possa proporcionar às crianças boas oportunidades de aprendizado mútuo, não somente em termos linguísticos, uma vez que pode fornecer a ambos os participantes a oportunidade de discussão e trocas recíprocas, promovendo assim, tipos de interação propícios ao desenvolvimento cognitivo.

Em resumo, podemos concluir que os dados infantis obtidos experimentalmente no âmbito da presente tese apresentam características próprias da fala entre pares que, por sua vez, pode ser relevante para compreendermos as diferenças observadas quando comparados os padrões de produção adulta e infantil.

4.6 SÍNTESE DOS PRINCIPAIS ACHADOS EXPERIMENTAIS

Neste capítulo, reportamos o estudo experimental conduzido no âmbito desta tese. Em ambos os experimentos desenvolvidos, utilizou-se a técnica de produção semielicada a partir de imagens com o objetivo primário de investigar as três dimensões da SF destacadas na literatura sociolinguística, quais sejam: o padrão de acentuação, o número de sílabas e os processos morfofonológicos de formação de plural dos nomes. Considerados em conjunto, os resultados obtidos não fornecem evidências robustas que permitam defender o papel da SF⁵⁰ dos itens nominais na alternância dos padrões de concordância plural no PB.

Como objetivos secundários do estudo experimental, propusemos explorar outros pontos que pensamos serem relevantes para a discussão teórica que conduzimos no Capítulo 5: a configuração e estrutura interna do sintagma, as características fonológicas dos numerais e suas relações com os padrões de marcação de plural. No quadro (7), retomamos os aspectos investigados, juntamente com seus respectivos resultados.

⁵⁰ Vale mencionar que *animacidade* associada aos nomes não foi um fator controlado quando da construção dos estímulos experimentais. Contudo, ao longo de nossas investigações, optamos por analisar os dados obtidos experimentalmente em função de tal fator. O resultado de tal análise parece contrariar a visão reportada na literatura (SCHERRE, 1988; BRANDÃO, 2009, 2013), segundo a qual nomes com o traço [+animado] tenderiam a ser mais marcados morfofonologicamente no plural do que nomes [-animados]. Nos nossos dados, no entanto, não foi registrada diferença significativa entre a produção de sintagmas redundantes e não redundantes em função de animacidade dos nomes.

Quadro 7 - Síntese com os principais achados experimentais

Aspecto investigado	Resultados
Saliência fônica	<p>A SF dos nomes não se mostrou relevante como condicionadora da alternância de padrões de concordância nominal, ao analisarmos os dados seguindo os moldes mais frequentemente encontrados nas pesquisas de cunho descritivo.</p> <p>Nossos resultados, no entanto, espelham em boa medida a diversidade de achados encontrada na literatura sociolinguística e sugerem que a variável SF – caso de fato seja relevante – pode ser um condicionador multifatorial (envolvendo aspectos sociais e linguísticos) de muito difícil apreensão .</p>
Configuração dos sintagmas	<p>Quanto à configuração dos sintagmas em termos de número de elementos morfológicamente realizados, os dados sugerem que o número de elementos que compõem o sintagma seja um fator relevante para a ocorrência de variação na marcação de número nos sintagmas nominais. Pelo menos no que tange a sintagmas encabeçados por numeral, nossos dados são compatíveis com a ideia de que uma estrutura menor (formada por dois elementos) parece favorecer o uso do padrão redundante em comparação a estruturas de três ou quatro elementos.</p>
Marcação interna de plural	<p>Sobre os padrões de marcação internos ao sintagma observados nos dados, observamos diferenças entre a produção infantil e a adulta. Os dados das crianças sugerem um contraste entre marcação total X marcação mínima, ao passo que os dados dos adultos sugerem uma hierarquização diferente dos padrões que segue uma gradação que vai da marcação total, seguida de marcação só no nome e, finalmente de marcação mínima (pluralidade representada apenas pelo numeral). Tais comportamentos sugerem que crianças e adultos seguem uma lógica diferente frente aos padrões de marcação internos.</p>
Morfologia dos numerais X Concordância	<p>Ainda que haja coincidência fonológica entre o –S que compõem o final de alguns numerais e o –S que marca plural nos nomes, os adultos parecem perceber o diferente estatuto morfológico desses elementos que, respectivamente, ora são parte da raiz e ora são morfemas independentes dela. Já para as crianças – pelo menos na idade investigada – a semântica do numeral em si não parece estar completamente resolvida quanto se mostra nos dados dos adultos. Quando o numeral apresenta uma “falsa analogia”, com uma marcação de plural os dados sugerem que as crianças parecem confiar nessa marcação, ou seja, elas interpretam o –S nos numerais como uma marca de plural, gerando, assim, uma maior omissão de marca de plural no nome.</p>

Fonte: Elaborado pela autora (2021)

Nossa pesquisa visa a aprofundar a compreensão acerca da alternância nos padrões de concordância nominal no PB, a partir de um diálogo entre a sociolinguística, a psicolinguística e teoria linguística. Nesse sentido, nosso trabalho busca ainda articular os resultados naturalísticos e experimentais obtidos à discussão teórica de cunho formalista proposta na literatura para a concordância no âmbito do DP na tentativa de propormos uma análise que dê conta do fenômeno de variação identificado no PB. Para tal, no Capítulo 5 faremos uma apresentação de algumas propostas teóricas relevantes para nossa investigação. Antes, porém, julgamos oportuno apresentar alguns conceitos relevantes ao arcabouço teórico que fundamenta nossa pesquisa, assim como a caracterização de diferentes abordagens propostas no âmbito do PM (CHOMKY, 1993; 1995 e trabalhos subsequentes) para o estudo da concordância nominal.

5 INTERPRETAÇÕES TEÓRICAS DOS FATOS EMPÍRICOS: A CONCORDÂNCIA NOMINAL NO PB SOB UMA ABORDAGEM FORMAL

Segundo Acuña-Fariña (2009), o fenômeno da concordância tem sido um desafio tanto para as teorias de gramática quanto para teorias de compreensão e produção. Para Corbett (2006, p.3), tal desafio justifica-se, por exemplo, em virtude de a concordância configurar-se como "o principal problema de interface entre morfologia e sintaxe", de modo que "tal fenômeno parece ser particularmente difícil de ser investigado a partir de qualquer um desses componentes." Nesse mesmo sentido, Acuña-Fariña (2009) ainda argumenta que a natureza, fonte e atuação da concordância exibem uma mistura de motivações formais e semânticas. O referido autor afirma que há modelos de gramática, bem como de processamento que consideram a concordância como um fenômeno essencialmente sintático, ou seja, totalmente independente dos outros níveis de representação linguística. Por outro lado, há também os que mantêm a visão de que a concordância consiste em um processo semântico por natureza, ou ainda os que a veem como um fenômeno que une aspectos de ambas as naturezas.

No que tange à nossa pesquisa, interessa-nos, em particular, a noção veiculada pelo PM (CHOMSKY, 1993, 1995 e trabalhos subsequentes), de que primitivos menores do que os itens lexicais, mais especificamente, os traços é que estejam envolvidos nas operações gramaticais, entre elas, as relações de concordância. Nesse contexto, a informação de *número*⁵¹, por exemplo, é definida como um traço que pode apresentar diferentes valores (singular, plural, dual, paucal etc.) a depender da língua analisada. Além disso, interessa-nos ainda a forma como a concordância nominal de número se manifesta, o que também é um aspecto distintivo interessante entre as línguas. Tem-se, por exemplo, casos como ocorre na variedade usualmente caracterizada como "padrão" do PB e em outras línguas como o espanhol e o italiano, em que a presença de marcas morfológicas de plural é compartilhada entre D e N, bem como entre os outros itens que mantêm uma relação estrutural apropriada com o elemento que desencadeia a operação de concordância. Por outro lado, há casos, como

⁵¹Vale lembrar também que, além de *número*, outros traços estão envolvidos em relações de concordância, como por exemplo, os traços de *gênero* e *pessoa*, cujo funcionamento, no entanto, extrapolam o escopo desta pesquisa.

certas variedades do PB que nos interessam particularmente, em que a marca explícita de plural pode ser manifesta somente em D.

A partir dos padrões empíricos identificados nos dados coletados no âmbito desta tese e reportados nos capítulos anteriores, o objetivo do presente capítulo consiste em desenvolver uma análise formal do fenômeno da concordância nominal de número. Mais especificamente, buscamos compreender se há sistematicidade na possibilidade de variação entre concordância redundante e não redundante no PB. Antes disso, no entanto, faremos uma breve apresentação relativa a algumas noções importantes dentro do contexto minimalista da teoria gerativa, tais como o modelo de língua assumido, bem como as noções envolvidas na caracterização do estabelecimento das relações de concordância. Em seguida, trataremos da Teoria dos *4M*, modelo de classificação de morfemas que, aliado às noções do PM, traz uma perspectiva interessante a partir da qual baseamos a análise formal que vamos propor para os nossos dados.

5.1 PROGRAMA MINIMALISTA: NOÇÕES FUNDAMENTAIS

O desenvolvimento de abordagens de base gerativista tem como tarefa fundamental a tentativa de explicar como as crianças adquirem sua competência gramatical apesar da “precariedade” de dados de *input* que recebem ao longo do processo de aquisição. De maneira geral, a teoria gerativa assume que os falantes adultos de uma língua natural têm internalizado, em sua mente/cérebro, um léxico e um conjunto de regras ou de propriedades responsáveis pela combinação de itens lexicais – i.e. uma gramática – que permite construir um número ilimitado de expressões linguísticas. O processo de aquisição da linguagem seria, nesse sentido, influenciado pelos dados linguísticos primários (*PLD*, do inglês *Primary Linguistic Data*) a que tal falante é exposto quando criança. No entanto, segundo a visão gerativista, a informação gramatical adquirida por meio desses dados aos quais a criança tem acesso, não seria suficiente para explicar a competência linguística que um indivíduo adulto apresenta. Por essa razão, assume-se que os seres humanos nascem com a chamada Faculdade da Linguagem, ou seja, nascem biologicamente dotados com uma capacidade inata dedicada à aquisição da linguagem – uma propriedade da mente/cérebro que seria específica da espécie humana.

Nessa perspectiva, o PM é um desenvolvimento atual da teoria gerativa que surge da percepção do sucesso da abordagem desenvolvida no modelo de *Government and Binding* (GB) (CHOMSKY, 1981 e trabalhos subsequentes), bem como em decorrência de uma preocupação relacionada à extensa quantidade de primitivos metodológicos adotados na GB. Nesse sentido, o PM assume como diretriz de base a possibilidade de construção de um aparato teórico mínimo e otimizado, com vistas a se formular uma teoria de gramática mais enxuta, mais econômica e, portanto, mais atraente do ponto de vista científico. Um exemplo bastante relevante desse enxugamento teórico pode ser visto, por exemplo, nos diferentes níveis da derivação propostos no modelo de gramática, visto que, no contexto de GB, são formulados quatro níveis de representação, a saber: (i) Estrutura Profunda (*Deep Structure*); (ii) Estrutura Superficial (*Surface Structure*); (iii) Forma Fonética (PF, do inglês *Phonetic Form*); e, (iv) a Forma Lógica (LF, do inglês *Logical Form*). No PM, por sua vez, propõe-se a abolição da distinção entre DS e SS por se tratarem de entidades intrateóricas não empiricamente motivadas pelos fatos linguísticos. Diante disso, somente os dois últimos níveis de representação, i.e., PF e LF são mantidos na abordagem minimalista, em virtude de serem os únicos que fazem interface com os sistemas de desempenho. Tais sistemas de desempenho são constituídos pelos sistemas articulatório-perceptual (A-P) e conceitual-intencional (C-I), os quais estabelecem interfaces com os níveis PF e LF, respectivamente. Assumindo-se ainda que PF e LF são os únicos níveis de interface linguística, eles são também os responsáveis por prover instruções para os sistemas de desempenho.

Vale ressaltar também que, na visão minimalista, observa-se uma nova forma de conceber a Faculdade da Linguagem. Nesse sentido, Hauser, Chomsky e Fitch (2002) estabelecem uma diferenciação entre Faculdade da Linguagem em um sentido amplo (FLB, do inglês, *Faculty of Language in the Broad Sense*) e Faculdade da Linguagem em um sentido estrito (FLN, *Faculty of Language in the Narrow Sense*). A FLB envolveria um sistema computacional interno (a FLN) e o conjunto dos sistemas de desempenho com os quais a FLN faz interface: (i) o sistema sensório-motor, também chamado articulatório-perceptual e (ii) o sistema conceitual-intencional ou sistemas de pensamento. A FLN, por sua vez, compreende apenas um sistema computacional universal, fundamentalmente centrado na operação de *Merge* (concatenação). Embora os autores não discutam com profundidade a natureza e aquisição do léxico propriamente dito, Hauser, Chomsky e Fitch (2002) reconhecem

que a capacidade de adquirir um léxico compartilhado e arbitrário é central para a capacidade linguística em si e, embora crianças possam usar mecanismos gerais para adquirir palavras, a taxa em que as crianças constroem o léxico é tão massivamente expressiva e diferente de tudo que se poderia ter como correlato em outras espécies que se deve considerar a possibilidade de um mecanismo evolutivo desenvolvido de forma independente nos humanos.

Em linhas gerais, no âmbito do PM, ao léxico caberia a tarefa de especificar os itens que entram no sistema computacional e suas propriedades idiossincráticas, sendo uma espécie de “dicionário mental” composto pelo conjunto de itens lexicais formados por traços que, por sua vez, seriam peças fundamentais na construção de sentenças. Dessa forma, os traços que compõem os itens do léxico desempenham importante papel dentro da abordagem minimalista, uma vez que são considerados as unidades atômicas da gramática e, por conseguinte, configuram-se em objetos de estudo em numerosas pesquisas. Em outras palavras, seriam os elementos primitivos da gramática, podendo ser de natureza fonológica, semântica ou formal (ADGER, 2002). Sendo assim, os traços podem ter realidade nas interfaces, ou seja, podem ser interpretados em LF, bem como ter realização fonológica em PF.

Mais especificamente, enquanto traços fonológicos e semânticos informam as interfaces, mas não desengatilham operações sintáticas, os traços formais, por sua vez, desempenham um papel importante na sintaxe, em virtude não somente de serem os blocos básicos de construção da estrutura sintática da língua, como também por serem responsáveis pela especificação das possíveis relações de combinação entre os elementos no interior da estrutura sintática. Nesse contexto, Adger e Svenonius (2012) ressaltam o papel crucial desempenhado pelos traços dentro da abordagem minimalista, construindo uma analogia segundo a qual um traço de [plural], por exemplo, representaria para a teoria o mesmo que o uso do H por um químico representaria para o elemento hidrogênio no mundo real. Para defenderem tal ideia, os autores lançam mão da seguinte argumentação em Adger e Svenonius (2012, p.27):

From this perspective it is crucial to say what the possible feature structures are such that the properties of the features allow them to enter into relationships with other features, analogously to saying what the properties of atoms are such that they can enter into relationships with other atoms [...] From this viewpoint, the constraints on the feature theory are substantive and amount to constraining the theory itself.

Ainda no que diz respeito aos traços formais, Kenedy (2013, p.137) defende que esses têm a função “de orientar o Sistema Computacional acerca das relações sintáticas que um dado item lexical deve estabelecer com outros itens no interior da sentença em que esteja inserido”. Dentre os traços formais, temos, por exemplo, os traços de categoria (N, V, D, C, etc.) e ainda outros traços como *Caso*, *Wh* (associado a itens interrogativos), EPP⁵² (do inglês, *Extended Projection Principle* – Princípio da Projeção Estendida), bem como os chamados traços-*phi*, aos quais cabe a tarefa de apresentar informações sobre gênero, número e pessoa e que, por estarem envolvidos em relações de concordância, são especialmente relevantes para as discussões deste capítulo. Sobre os traços-*phi*, Polinsky (2015), ressalta que, embora as línguas naturais apresentem variação em termos fonológicos, morfológicos e lexicais, o conjunto de traços que são combinados em decorrência da operação de concordância é bastante uniforme translinguisticamente – i.e. envolvendo os traços de pessoa, número e gênero.

No concernente à caracterização dos traços, mais especificamente, Chomsky (1999) assume uma distinção entre traços interpretáveis de traços não-interpretáveis determinada no léxico. Por traços interpretáveis entendem-se as dimensões valoradas de um dado traço, ao passo que os não-interpretáveis são aqueles traços que só adquirem seu valor no curso da derivação sintática. Esses últimos, são de extrema importância para a sintaxe, uma vez que são eles que desencadeiam as operações necessárias para que a derivação possa convergir. Isso porque os traços não valorados não são legíveis nas interfaces, fato que ficou mapeado no contexto minimalista através do Princípio da Interpretabilidade Plena (*Full Interpretation*), o qual estabelece que as informações disponibilizadas aos níveis de interface devem ser totalmente legíveis para tais níveis, sendo que informações não legíveis devem ser eliminadas durante a derivação sintática.

Mais especificamente, como um dos recursos para a eliminação dos traços não interpretáveis envolvidos na concordância é postulada, no âmbito do PM, a operação *Agree*, que, em linhas gerais, consiste no estabelecimento de uma relação entre os

⁵² Em linhas gerais, o traço EPP está relacionado à necessidade de se preencher uma posição de especificador, como o núcleo T (*Tense*), que através desse traço atrai um DP para o seu especificador.

traços interpretáveis e os não interpretáveis em um domínio estrutural apropriado como veremos de forma mais detalhada na próxima seção.

5.2 CONCORDÂNCIA NA TEORIA GERATIVA

A concepção do mecanismo de concordância e, conseqüentemente, sua formalização na teoria, sofreu transformações ao longo da história da abordagem gerativista, embora tenha se mantido quase inalterada a concepção de número como traço formal – concepção que assumimos também em nossa pesquisa. Além disso, compartilhamos com Polinsky (2015), a ideia de que, para se discutir a concordância na teoria gerativa, em particular, a partir da perspectiva do PM, é preciso ter em mente duas noções básicas, a saber, as noções de categoria funcional e de traços-*phi*, mencionados anteriormente. Por categorias funcionais entendem-se os elementos que constituem classes fechadas e compreendem itens como (D)eterminantes, (T)empo, (C)omplementizadores, entre outros. Tais elementos carecem de um ‘conteúdo descritivo’, ou seja, não atuam no estabelecimento de relações temáticas, possuindo significado puramente gramatical. Dessa forma, os núcleos funcionais são predominantemente constituídos por traços formais, que desencadeiam operações no interior da estrutura sintática. Ademais, são itens, geralmente, dependentes fonológica e morfologicamente. Com Baker (2008), entendemos que são, mais especificamente, os núcleos funcionais que estão envolvidos nas relações de concordância.

Sobre os traços-*phi*, há um outro aspecto que merece nossa atenção, a saber, o *locus* da interpretabilidade do traço de número no âmbito do DP, ponto que tem sido alvo de debate na literatura minimalista. Se, por um lado, é estável a ideia de que a especificação do gênero gramatical é uma propriedade arbitrária e inerente aos Ns, por outro lado, não é possível fazer uma afirmação tão clara a respeito do traço de número. Nesse sentido, poderíamos dizer, por exemplo, nos moldes da abordagem de valoração, a partir de autores como Hornstein, Nunes, e Grohmann (2005) e Polinsky (2015), dentre outros, que o traço de gênero seria uma propriedade [+interpretável] no N, mas [-interpretável] nos outros elementos que se relacionam a ele dentro do sintagma relevante, como o D e Adj. Por outro lado, em relação ao traço de número, há propostas variadas na literatura a respeito do seu *locus* da interpretabilidade. Tem-se, por exemplo, a proposta de Longobardi (1994), anterior ao sistema de valoração, mas representativa da ideia de que o traço de número seria interpretado em D, como

podemos observar na citação extraída de Longobardi (1994, p.620):

The crucial nature of the category D for argumenthood seems to be independently manifested by the fact that certain designation properties typical of arguments, such as the semantic import of grammatical number, lie precisely in the D position. Actually, we have already observed that an empty D (therefore, one morphologically unspecified for number) may yield semantic indeterminacy between singular and plural designation despite the plurality of the head noun.

Em contrapartida, Chomsky (1999) defende que o traço de número seja [+interpretável] no N, ainda que possa ter um reflexo em outros itens do sintagma, como modificadores e D. A justificativa do autor se dá, entre outros fatores, em virtude da marcação morfológica de número em tal sintagma ser expressa nos itens nominais no inglês.

No entanto, outras propostas acerca do *locus* de interpretabilidade do traço de número no âmbito do DP podem ser encontradas na literatura. Carstens (2000), seguindo trabalho de Ritter (1991) – portanto, anterior às propostas de Chomsky (1999) e Longobardi (1994) – argumenta que o *locus* da interpretabilidade do traço de número seria um núcleo funcional Num, alocado entre D e N, que seria responsável justamente pela especificação da informação de número, questão diretamente relevante para os propósitos deste capítulo e que trataremos mais detalhadamente na seção que trata da concordância no DP.

Quanto a estudos conduzidos sobre o PB, mais especificamente, é possível encontrar representantes das três visões apresentadas acima, dentre os quais, podemos citar, Magalhães (2004) que, seguindo Longobardi (1994), propõe que número seja [+interpretável] em D, associando a interpretabilidade do traço de número ao núcleo em que seria obrigatória a marca explícita de plural na língua. Lopes (2004), por sua vez, compartilha com Chomsky (1999) a visão de que seja o N o responsável pela interpretabilidade do traço de número. Já no que diz respeito aos estudos acerca do PB que seguem as ideias de Ritter (1991) e Carstens (2000) entre outros, de que o *locus* da interpretabilidade do traço de número seria um núcleo funcional Num, alocado entre D e N, podemos citar pesquisas tais como as desenvolvidas por Ferrari Neto (2003), Augusto, Corrêa e Ferrari Neto (2006) e Simioni (2006, 2007).

Contudo, autores como Hornstein, Nunes e Grohmann(2005) argumentam que, no concernente a número, não é possível afirmar com exatidão de que tal traço

seja [+interpretável] no N, tal como defendido por Chomsky (1999), Lopes (2004), ou no determinante (LONGOBARDI, 1994; MAGALHÃES, 2004), entre outros. Para defenderem sua posição, Hornstein, Nunes e Grohmann(2005) se baseiam no fato de que as línguas, na verdade, variam no que tange à marcação explícita de número entre os elementos que compõem o DP. Como vemos nos exemplos extraídos de Hornstein, Nunes, e Grohmann(2005, p.283), (10) mostra a marcação de número expressa somente no N em inglês, ao passo que, como se pode observar no português brasileiro coloquial, reproduzido no exemplo (11) abaixo, o plural pode ser expresso apenas em D, podendo ser omitido nos demais itens do DP, aqui representados pelo N *livro*.

(10) the books

(11) os livro

the-PL book-SG

Além da discussão a respeito do *locus* da interpretabilidade, há ainda algumas diferentes análises que visam a explicar o mecanismo da concordância no âmbito do PM, tais como: (i) a visão baseada na noção de *checagem de traços* (CHOMSKY, 1995); (ii) a proposta pautada na ideia de *valoração de traços* (CHOMSKY, 1999, 2001) e (iii) a ideia proposta por Frampton e Guttman (2000) segundo a qual, a concordância seria melhor compreendida a partir da perspectiva de *compartilhamento de traços* – a qual nos interessa em particular, por pensarmos que tal visão abre perspectivas interessantes para a análise dos nossos dados. Antes de apresentarmos a operação de *Agree* a partir dessa perspectiva, pensamos ser oportuno fazer uma breve apresentação da visão baseada na noção de valoração (CHOMSKY, (1999, 2001).

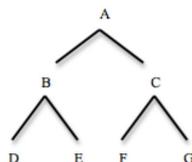
Chomsky (1999, 2001) defende que os traços não interpretáveis entrariam na derivação sem valor especificado e ganhariam efetivamente valor durante a computação sintática, a partir do pareamento de traços interpretáveis de mesma dimensão. Nesse sentido, no modelo de valoração de traços, a operação *Agree* prevê que um núcleo com traços-*phi* não valorados – e, conseqüentemente, não interpretáveis na visão do autor – atua como uma sonda que parte em busca de um constituinte alvo c-comandado⁵³ presente na mesma fase⁵⁴, contendo traços valorados

⁵³Chomsky (1995, p.31) defende que a noção de c-comando consiste em um conceito fundamental que se aplica aos módulos de gramática. Segundo o autor, dizemos que α c-

– e, conseqüentemente, interpretáveis – da mesma dimensão para valorar os seus. Frampton e Guttman (2000), por sua vez, propõem que a concordância envolva *compartilhamento de traços*, ou seja, a operação de *Agree* mapeia uma relação entre um traço não valorado que atua como sonda e um traço de um alvo c-comandado, resultando em duas instâncias de um mesmo objeto sintático. Nesse contexto, havendo identidade entre traços interpretáveis ou não-interpretáveis, eles se juntam em uma cadeia de traços compartilhados, que são simultaneamente valorados se um dos traços combinados for valorado. Sendo assim, os referidos autores chamam a atenção para um detalhe importante da proposta: traços não interpretáveis [u] se juntam entre si da mesma maneira que um traço interpretável [i] se junta com um [u]; ou seja, a operação *Agree* consiste em compartilhamento de traços, independente do seu valor. Em outros termos, *Agree* é cega aos valores dos traços. Nesse sentido, substitui-se a ideia defendida em Chomsky (1999, 2001), de que o alvo carregue necessariamente um traço valorado, atribuindo seu valor a uma contraparte não-valorada – no caso, a sonda. Vale lembrar ainda que, no modelo de compartilhamento de traços de Frampton e Guttman (2000), a operação *Agree* entre dois traços não valorados é licenciada, contanto que ocorra uma operação de *Agree* em um nível subsequente da derivação que efetivamente atribua valor ao traço compartilhado. Voltaremos a tratar desse modelo com mais detalhes mais adiante, uma vez que o compartilhamento de traços parece ser mais compatível com a abordagem de concordância que vamos implementar.

A partir do que foi exposto até aqui, vale destacar que, na nossa proposta, partimos da concepção de número como traço formal relevante na sintaxe. Além

comanda β se α não domina β e se cada γ que domina α também domina β . Sendo assim, na estrutura hipotética a seguir, B domina D e E, C domina F e G e A domina todas as outras categorias (nós). Como consequência, B c-comanda C, F e G, C c-comanda B, D e E, D c-comanda E e vice-versa, e ainda F c-comanda G e vice-versa.



⁵⁴ Em termos gerais, por fase entende-se um objeto sintático cujas partes, mais especificamente, o complemento de seu núcleo podem ser inspecionadas para fins de convergência nas interfaces (HORNSTEIN; NUNES; GROHMANN, 2004).

disso, propomos que a variabilidade na realização morfofonológica do traço de número possa ser uma pista da sua valoração e, nessa perspectiva, assumimos, como Magalhães (2004), que D seria o *locus* da interpretabilidade do traço de número no PB. A novidade do nosso sistema, no entanto, é a associação entre a valoração propriamente dita e o desenvolvimento de uma perspectiva mais morfológica. Mais especificamente, na nossa visão, a valoração do traço de número estaria atrelada à tipologia de morfemas proposta na Teoria dos *4M*, que será apresentada mais adiante. No tocante aos mecanismos de concordância apresentados, implementamos nosso sistema com base na ideia de compartilhamento de traços, proposta por Frampton e Guttman (2000), Pesetsky e Torrego (2007), uma vez que nossa estrutura prevê a relação, no âmbito do DP, entre elementos igualmente não valorados para o traço de número, como Ns e Adjs, até a entrada de uma camada mais alta na qual haja a manifestação valorada do traço de número. Isso posto, passemos para as propostas para o estudo da concordância no âmbito do DP, considerando também propostas que tentam dar conta da concordância variável.

5.3 PROPOSTAS PARA O ESTUDO DA CONCORDÂNCIA NO DP

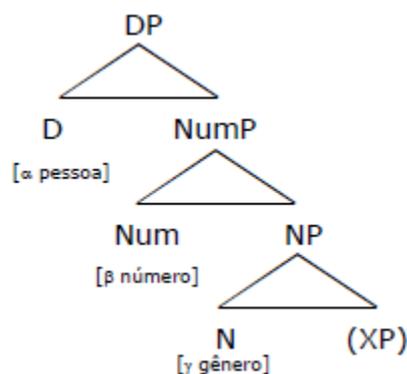
Quando se trata da concordância no interior do DP, duas questões principais surgem como relevantes: (i) a discussão a respeito de qual seria o mecanismo responsável pelo licenciamento dessa concordância e, mais especificamente, se haveria distinção desse mecanismo para aquele que regula a concordância no nível sentencial e (ii) a discussão a respeito do lugar da valoração de cada traço relevante – como gênero, número e pessoa – em relação às projeções que compõem o DP.

Quanto ao primeiro ponto, segundo Chomsky (1999), a concordância no interior do DP – a qual em nota de rodapé ele chama de *Concord* –, seria realizada através da checagem de traços entre D e N, numa relação *in situ*, ou seja, envolvendo apenas *Merge*, diferentemente da concordância no âmbito da sentença. Sobre o segundo ponto, por sua vez, Chomsky (1999) defende, como já foi salientado, que a concordância no interior do DP seria regida pelos traços contidos em N.

Na visão de Carstens (2000), contudo, os traços que integram o DP estariam, na verdade, espalhados entre os diferentes núcleos que compõem esse sintagma. Mais especificamente, seguindo autores como Ritter (1991), Carstens (2000) defende a

existência de uma categoria funcional Num e, portanto, de uma projeção NumP, situada estruturalmente entre DP e NP. De acordo com esta concepção, D carregaria traços de pessoa [+interpretável] e traços de número e de gênero [-interpretável]. A categoria Num, por sua vez, apresentaria traço de número [+interpretável], enquanto o traço de gênero seria interpretável em N. Os traços de número e gênero [-interpretável] de D seriam os fatores desencadeadores do alçamento de N, num primeiro estágio para Num, onde aconteceria a checagem de seu traço de número não interpretável, em seguida, para D, para checar seu traço não interpretável de pessoa. Nesse sentido, percebe-se que Carstens (2000), embora utilize o mecanismo de *Concord* – proposto por Chomsky (1999) – para a concordância no âmbito do DP, diferentemente deste, assume que tal mecanismo envolveria movimento⁵⁵: o alçamento de N para Num e, em seguida, para D. A figura (5) apresenta os traços e categorias funcionais vinculados à concordância no âmbito do DP propostos por Carstens (2000, p.328):

Figura 5 - Traços e categorias funcionais vinculados à concordância no âmbito do DP, segundo Carstens (2000).



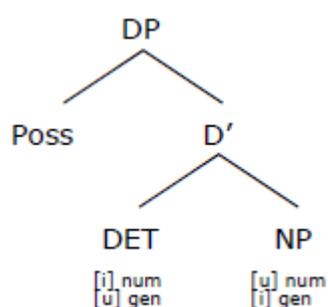
Fonte: Carstens (2000, p.328)

Para o PB, mais especificamente, trabalhos como os de Ferrari-Neto (2003), além de Augusto, Corrêa e Ferrari-Neto (2006) e Simioni (2007), seguem o sistema de Carstens (2000), assumindo a existência de uma categoria funcional Num e, portanto, de uma projeção NumP, situada estruturalmente entre DP e NP. Em

⁵⁵ Tal movimento pode acontecer na sintaxe aberta ou no componente coberto a depender da língua sob investigação.

contrapartida, Magalhães (2004), contrariando Carstens (2000), não assume a existência de uma categoria funcional independente para número – Num –, nem uma projeção NumP. Magalhães (2004) diverge de Carstens (2000) também ao defender que seja utilizado o mesmo mecanismo de *Agree* tanto para o nível sentencial como para as relações no interior do DP. Afinal, para Magalhães (2004), primeiramente, dá-se a valoração dos traços-*phi* dos elementos concordantes no interior do DP e, em seguida, tais elementos tem seu traço de *Caso* valorado por uma sonda externa ao DP. A figura (6) ilustra a estrutura de DP defendida por Magalhães (2004, p. 159):

Figura 6 - Estrutura interna do DP proposta por Magalhães (2004)

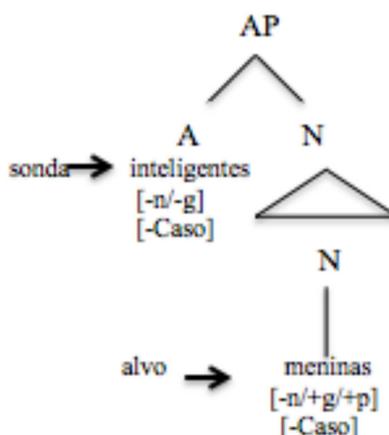


Fonte: Magalhães (2004, p.159)

Em termos da distribuição de traços no DP, a autora assume que o determinante (D) possui traços de número [+interpretável] – seguindo a proposta de Longobardi (1994) – e de gênero [-interpretável]. Em contrapartida, o N possui traços de gênero [+interpretável] e de número [-interpretável]. Para os demais elementos que possam compor o DP, tais como os Adjs, tanto o traço de número quanto o de gênero seriam [-interpretável].

No que diz respeito especificamente à posição do Adj na estrutura interna do DP, Magalhães (2004, p.[13]) assume como estrutura básica para a concordância baseada na relação sonda-alvo em DPs contendo Adjs, a mesma assumida por Abney (1987) e Cinque (1994), como se pode observar na figura (7) a seguir:

Figura 7 - Estrutura básica para a relação sonda-alvo em DPs contendo Adjs assumida por Magalhães (2004, p[13]), com base em Abney (1987) e Cinque (1994)



Fonte: Magalhães (2004, p.[13])

Como visto acima, o N (*meninas*) tem traço de número [-interpretável] e de gênero [+interpretável], enquanto o Adj (*inteligentes*), por sua vez, possui ambos os traços [-interpretável].

Mais especificamente acerca do licenciamento dos Adjs, em Cinque (1994) assume-se que a posição básica dos sintagmas adjetivais (APs) dentro do sintagma nominal seria, apesar das aparentes diferenças, a mesma nas línguas românicas e germânicas, ou seja, à esquerda de N, e que sua posição de superfície diferente seria resultado da subida de N para um núcleo funcional localizado entre N e D, em línguas românicas como o português, tal como ocorre em *As meninas inteligentes*, mas não nas germânicas, como é o caso do inglês, *The intelligent girls*, representadas nos esquemas (12) e (13), respectivamente, extraídos de Cinque (1994, p.21):

(12) [D..[AP Y [AP N]]] (línguas românicas)



(13) [D..[AP Y [AP N]]] (línguas germânicas)

Nesse sentido, para Cinque (1994), as estruturas de DPs nas línguas românicas e germânicas são mais semelhantes entre si do que se possa imaginar e que o movimento de N à esquerda, seria um parâmetro relevante da gramática das línguas.

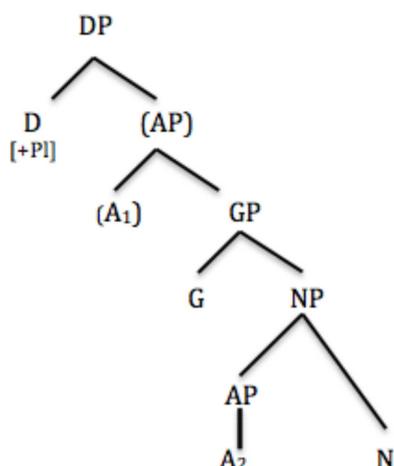
A estrutura sintática assumida por Magalhães (2004) e apresentada anteriormente na figura (7), segue a estrutura proposta por Cinque (1994), ou seja, pressupõe Adjs alocados à esquerda de N. No entanto, tal estrutura não condiz com a estrutura de superfície apresentada no exemplo (15), que por sua vez, apresenta a ordem canônica observada em PB para o âmbito do NP, ou seja, N + Adj (*as meninas inteligentes*). Dessa forma, o sistema precisa postular a subida de N para algum núcleo funcional localizado acima, cuja natureza, no entanto, não é explorada pela autora. Além disso, Magalhães (2004) não fornece explicações para as diferentes posições que os Adjs podem apresentar em relação ao N no âmbito do DP. A nosso ver, tal explicação seria fundamental, afinal sua análise baseia-se em dados do PB, que, como língua românica, permite em alguns casos que os Adjs sejam alocados antes (14) ou depois do N (15). No entanto, faz-se necessário tecer algumas considerações mais aprofundadas sobre essa questão.

Em Menuzzi (1994), por exemplo, encontramos uma discussão acerca das diferentes posições que os adjetivos podem ocupar no âmbito do DP e como tal diferença pode influenciar na concordância. Em linhas gerais, o referido autor defende que o PB configura-se como um exemplo de sistema não uniforme de concordância de traços. Nesse tipo de sistema, os núcleos concordantes não flexionam de maneira uniforme para um determinado traço. Segundo Menuzzi, isso explicaria o surgimento de diferentes padrões de concordância com adjetivos a depender da especificação de traços de elementos da categoria D ou N. Nesse sentido, o referido autor – sustentado por Brito e Lopes (2016, p.266) – argumenta que no PB, a marcação de plural é sensível a diferentes posições dos adjetivos dentro do DP. A partir do estudo conduzido, Menuzzi (1994, p.131) chega a algumas generalizações sobre a questão no PB:

- (i) se um D e/ou um Adj prenominal estiver presente, o N não é marcado no plural, tal como em *os novos alunoø*;
- (ii) se somente o N estiver presente, este deve ser marcado explicitamente no plural, como ocorre em *alunos novos começam as aulaø amanhã*.

Essas duas primeiras generalizações podem ser representadas pela seguinte figura (8), extraída de Menuzzi (1994, p.131):

Figura 8 - Configuração de DPs no PB segundo Menuzzi (1994, p.131)

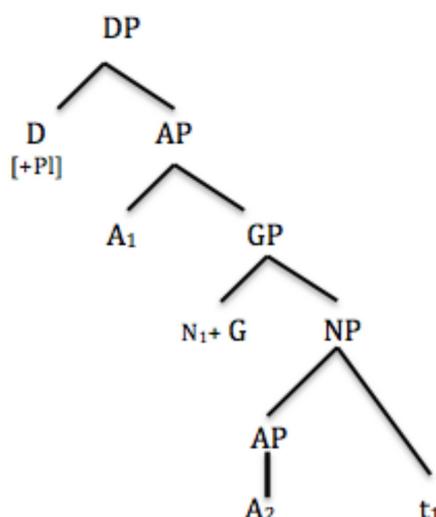


Fonte: Menuzzi (1994, p.131)

Ainda de acordo com Menuzzi, se o DP é especificado para [+PI] e D está presente, nenhum outro núcleo precisa apresentar [+PI]. Entretanto, caso não haja um D lexical para expressar [+PI], então outro núcleo lexical deverá fazê-lo. Caso N seja o único núcleo lexical disponível, este deverá se mover de G para D para expressar [+PI]. Contudo, há de se atentar para a possibilidade da presença de um Adj preposto ao N. Afinal, neste caso, a marca explícita de plural será realizada no Adj e não em N. Isso porque, como visto na figura (8), os Adjs pronominais estão em A1, ou seja, o núcleo mais próximo para expressar [+PI] quando da ausência de um D morfologicamente realizado. Sendo assim, A1 bloqueia o movimento de N para D, o que em outra situação, seria exigido. Os Adjs pospostos, por outro lado, não bloqueiam o movimento de N para D.

As outras duas generalizações propostas por Menuzzi, a saber (iii) Adjs prepostos ao N concordam com D em número, *os novos alunø*; e (iv) Adjs pospostos a N concordam com este em número, como se vê em *alunos novos começam as aulas amanhã* são mais bem compreendidas ao observarmos a figura (9) abaixo, extraída de Menuzzi (1994, p. 132):

Figura 9 - Configuração de DPs no PB segundo Menuzzi (1994, p.132)



Fonte: Menuzzi (1994, p.132)

De acordo com Menuzzi, A1 concorda com D, visto que A1 é c-comandado apenas por D, e não por N. Em contrapartida, podemos observar que A2 é c-comandado por ambos (D e N). Mas, vale lembrar que em virtude de N exercer a relação de c-comando mais próxima a A2, este está limitado pela restrição de minimalidade na acessibilidade de traços a concordar com N e não com D. Com isso, podemos entender melhor a proposta de Menuzzi para explicar porque a concordância tende a ser feita na esquerda do sintagma no PB. Diferentemente, nas línguas germânicas, como o inglês, Adjs são geralmente alocados à esquerda do N, como se observa no exemplo (16):

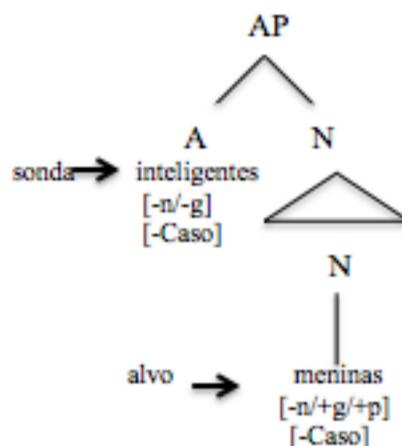
(14) As lindas meninas.

(15) As meninas lindas.

(16) The beautiful girls.

No que concerne, por sua vez, às relações de *Agree* explicitadas na análise de Magalhães (2004), observamos que na figura (7), reproduzida na figura (10), o N *meninas* entra na derivação com seus traços de número e Caso não valorados, ao passo que o Adj *inteligentes* entra na derivação com todos os seus traços não valorados.

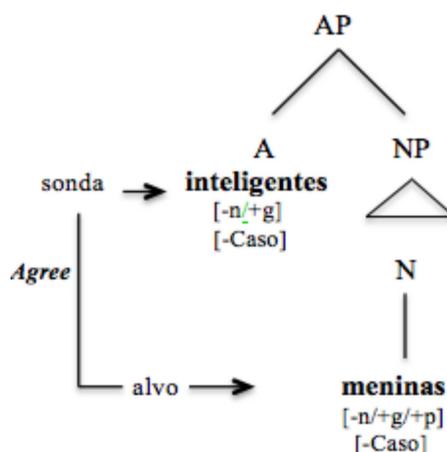
Figura 10- Explicando a operação de Agree no sistema de Magalhães (passo 1)



Fonte: Magalhães (2004, p.[12])

Sendo assim, o Adj *inteligentes* funciona como uma sonda que encontra no N *meninas* o alvo que contém a versão interpretável para valorar seu traço de gênero. Entretanto, o Adj continua ativo para o sistema visto que o seu traço de número, assim como acontece com o N, é [-interpretável], como visto na figura (11) abaixo.

Figura 11- Explicando a operação de Agree no sistema de Magalhães (passo 2)

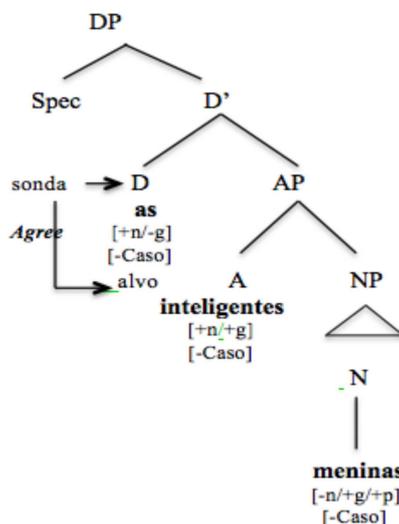


Fonte: Magalhães (2004, p.[12])

Em seguida, como esquematizado na figura (12), D entra na derivação e, em virtude de ter um traço de gênero para ser valorado, atua como sonda e entra em

relação de concordância com o primeiro alvo que encontra, ou seja, o Adj *inteligentes*, valorando os traços de número deste.

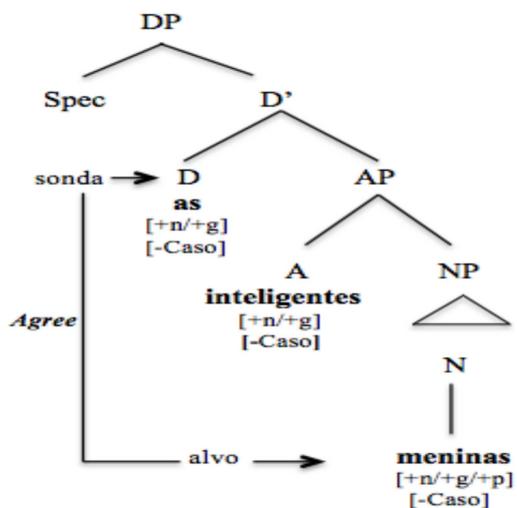
Figura 12 - Explicando a operação de Agree no sistema de Magalhães (passo 3)



Fonte: Adaptado de Magalhães (2004, p.[13])

No entanto, segundo o sistema da autora, apesar de o traço de gênero em Adj já estar valorado na derivação, o traço de gênero de D continua sem valor e, por essa razão, D continua a sondagem, encontrando N *meninas*. A relação de *Agree*, então, se efetua, resultando na valoração do traço de número de N e do traço de gênero de D, como apresentado na figura (13).

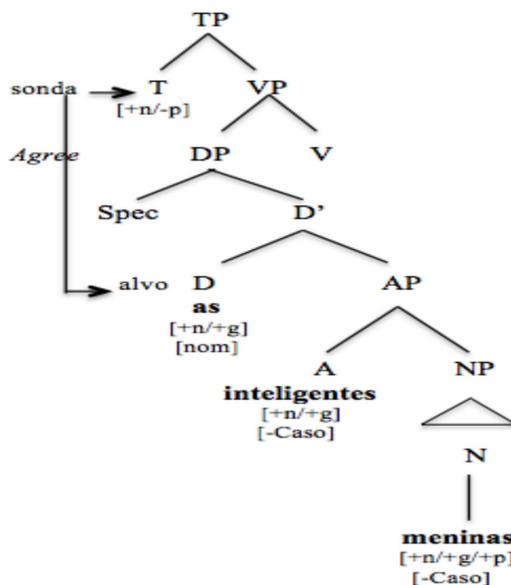
Figura 13 - Explicando a operação de Agree no sistema de Magalhães (passo 4)



Fonte: Adaptado de Magalhães (2004, p.[13])

T entra na derivação e sonda D e ao entrar na relação de *Agree* com este, T tem seu traço de número valorado e valora o traço de Caso de D, como visto na figura (14):

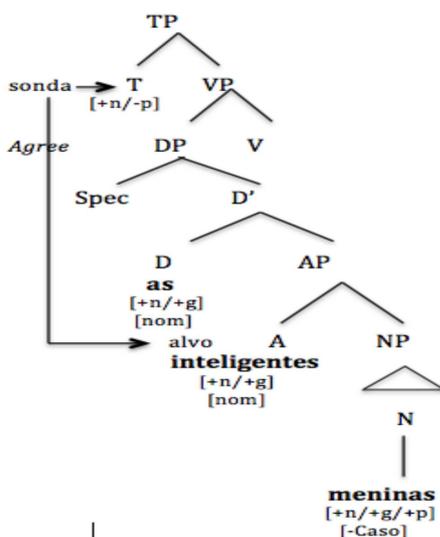
Figura 14 - Explicando a operação de Agree no sistema de Magalhães (passo 5)



Fonte: Adaptado de Magalhães (2004, p.[14])

Contudo, T ainda precisa valorar o seu traço de pessoa. Com isso, T continua a sondagem e encontra o Adj (cf. Figura 15), com o qual estabelece uma operação de *Agree* e valora o traço de Caso de A.

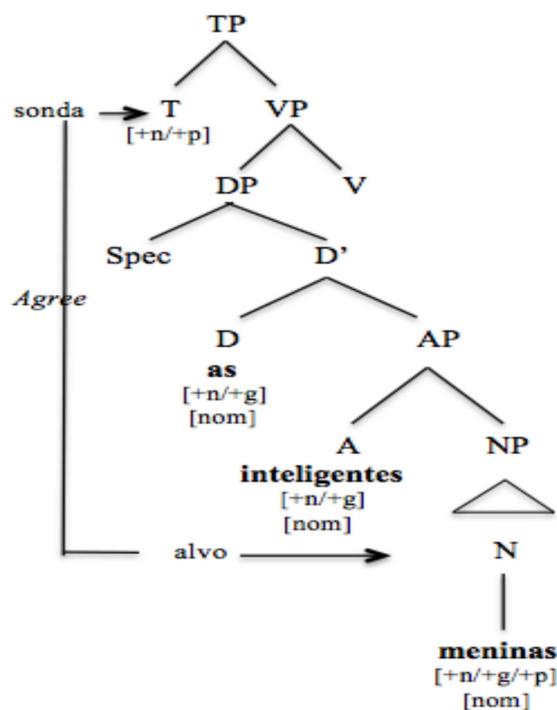
Figura 15- Explicando a operação de Agree no sistema de Magalhães (passo 6)



Fonte: Adaptado de Magalhães (2004, p.[14])

Todavia, T só terá seu traço de pessoa valorado ao se combinar com N, como se pode ver na figura (16), uma vez que Adj é defectivo para o traço de pessoa:

Figura 16 - Explicando a operação de Agree no sistema de Magalhães (passo 7)



Fonte: Adaptado de Magalhães (2004, p.[15])

Entretanto, vale ressaltar que tal análise, como bem apontado por Simioni (2007) apresenta algumas questões importantes. Para além dos aspectos relativos ao ordenamento de D+ N+ Adj, como comentado acima, Simioni (2007) ressalta que a valoração de gênero de D poderia resultar da relação de *Agree* entre D e o Adj como mostrado na figura (12), acima. No entanto, para Magalhães, o traço de gênero de D não pode ser valorado na relação de *Agree* entre D e Adj, visto que o Adj possui traço de gênero [-interpretável]. Sendo assim, Simioni argumenta que, nesse ponto da derivação (cf. Figura 12), o traço de gênero do Adj já se encontra valorado – em virtude da relação de *Agree* efetuada anteriormente entre o Adj *inteligentes* e o N *meninas* (cf. Figura 11) – podendo, portanto, atribuir valor ao traço de mesma natureza de D. Sendo assim, não seria necessário que D sondasse o N *meninas* como se vê na figura (13). Mas, não havendo *Agree* entre D e *meninas*, esse último ficaria

com o seu traço de número não valorado, ou seja, ilegível em LF e, conseqüentemente, acarretaria o fracasso da derivação.

Além disso, segundo Danon (2010), apesar de o estatuto do traço de pessoa ainda ser tema de debate na literatura, parece haver suporte empírico para rejeitar a ideia de que tal traço seja valorado e interpretado em N. Por exemplo, o mesmo nome pode licenciar DPs de primeira, segunda e terceira pessoa.

(17) *we/you/these linguists*

(DANON, 2010, p.7)

Se esse raciocínio estiver no caminho correto, não fica claro no sistema de Magalhães como o traço de Caso poderia chegar até N, uma vez que a não valoração exatamente do traço de pessoa em T é a motivação para o *Agree* entre esse núcleo e T.

Apesar das questões apontadas sobre o trabalho de Magalhães (2004), há de se atribuir à autora o caráter inovador de sua proposta para o avanço nos estudos sobre o fenômeno da concordância nominal variável no PB, ao propor que o *locus* da interpretabilidade do traço de número estaria na categoria D. A autora justifica sua defesa ao destacar que no PB, a omissão de marca explícita de plural nos demais elementos do DP não compromete a leitura plural do sintagma, tampouco compromete sua gramaticalidade.

Por outro lado, autores como Costa e Figueiredo Silva (2006), apresentam uma outra hipótese para explicar a ocorrência de concordância não redundante no PB. Seguindo pressupostos da Morfologia Distribuída, propostos por Halle e Marantz (1993) e Embick e Noyer (2001), Costa e Figueiredo Silva (2006), consideram que tal fenômeno – observado no PB, mas não no português europeu (PE) – estaria relacionado ao fato de o morfema associado à pluralidade ser do tipo *singleton* no PB – isto é, tal morfema se adjunge exclusivamente ao elemento que recebe a interpretação semântica relativa à informação de número. Em virtude de D ser o núcleo que liga o DP à sua interpretação em LF, é permitido que o morfema de plural seja realizado somente em tal núcleo, podendo ser omitido nas demais categorias no componente aberto. Quanto ao mesmo fenômeno no PE, os autores consideram que o morfema associado à pluralidade em tal língua seja do tipo dissociado, ou seja, a marca explícita de plural é atribuída pós-sintaticamente aos núcleos, independentemente de ter havido uma configuração ou operação sintática específica,

cabendo ao componente morfológico e à PF a tarefa de inseri-la em *Spell-Out* e atribuí-la a seus nós terminais.

No entanto, Silva e Magalhães (2014) apresentam evidências empíricas, a partir de dados extraídos de pesquisas conduzidas por autores como Naro e Scherre (2007) e Brandão e Vieira (2012), que contrariam a proposta de Costa e Figueiredo Silva (2006) de que o uso de padrão redundante no PE seja sempre categórico. Nesse sentido, podemos dizer que no PE, assim como ocorre no PB, é possível encontrar os dois tipos de morfema associado a pluralidade, ou seja, *singleton* e *dissociado*. Para o caso específico do PB, Silva e Magalhães sugerem que a alternância entre os padrões de concordância observados na língua decorra de uma competição entre gramáticas, em virtude de encontrarmos na produção de um mesmo indivíduo, ocorrências de ambos os padrões de concordância, sustentando a ideia defendida por Lighfoot (1999, p.92):

In that case, apparent optionality would be a function of coexisting grammars. Rather than allowing one grammar to generate forms a and b optionally, we would agree that a person has access to two grammars, one of which generates form a, the other form b; the speaker has the option at any given time on using one or other of the grammars.

Não podemos deixar de mencionar outra proposta de análise que pretende dar conta da variação dos padrões de concordância no DP no PB. Neste caso, a hipótese consiste em considerar tal fenômeno como consequência do enfraquecimento do paradigma morfológico da flexão nominal de número, como resultado do caráter instável – em termos fonéticos – do morfema –S, observado durante o processo de evolução da língua, ou seja, do latim até o estágio atual do PB, salientado em estudos como o de Nunes (1963 *apud* FERRARI-NETO, 2008), os quais destacam a forte tendência dos falantes da língua portuguesa à diácope. Tal tendência é também salientada por Guy (1981) que chama atenção para o padrão de estrutura silábica CV (consoante + vogal) do português que, por sua vez, favorece o apagamento das sibilantes finais no PB – especialmente quando precedidas de palavras iniciadas por outra consoante.

Além das propostas já apresentadas, destacamos também aquela aventada por Lopes (2001), quando da elaboração de sua tese de doutorado, de que o fenômeno da variação da concordância de número no sintagma nominal no PB possa ser melhor compreendido com base nos pressupostos da Teoria dos *4M* (do inglês, *4M Model*)

elaborados por Myers-Scotton e Jake (2000a), que será apresentada a seguir, uma vez que é central também para o sistema que desenvolvemos.

Antes, porém, e para encerrar esta seção, no quadro (8), esquematizamos uma síntese dos conceitos relevantes propostos no contexto do PM, tais como mecanismos de concordância e *locus* de interpretabilidade do traço de número no âmbito do DP apresentados até aqui, bem como das pesquisas conduzidas acerca do PB que seguem os mesmos passos:

Quadro 8 - Quadro contrastivo em função dos mecanismos propostos para concordância e do locus da interpretabilidade do traço de número apresentados nos trabalhos resenhados

Autor	Mecanismos de concordância	<i>Locus</i> da interpretabilidade do traço de número no âmbito do DP	Pesquisas acerca do PB que seguem tais passos
Chomsky (1999)	<i>Agree</i> (no nível sentencial). X <i>Concord</i> (concordância <i>in situ</i> através da operação de Merge)	N (segundo o PM, o qual estabelece que categorias lexicais possuem traços [+interpretável].	Lopes (2004), ao assumir que o traço de número é interpretável no N.
Carstens (2000)	<i>Concord</i> (no DP, mas contrariando proposta de Chomsky, tal mecanismo deflagra alçamento de N a Num e a D)	Num (seguindo autores como Ritter, 1991).	Ferrari Neto (2003), Augusto, Corrêa e Ferrari Neto (2006), Simioni (2006, 2007), ao assumirem a categoria Num
Magalhães (2004)	<i>Agree</i> (operação única de concordância para TP e DP), seguindo análises de Szabolci (1984), Abney (1987)	D (contrariando Chomsky), e seguindo análises de Abney(1987) e Longobardi (1994)	Lopes (2004), ao assumir a mesma operação de concordância no âmbito do DP e para o nível oracional.

Fonte: Adaptado de Azalim (2016, p.57)

5.3.1 Uma alternativa de análise para a concordância nominal variável: a Teoria dos 4M

Como foi apresentado nas seções anteriores, é possível identificar na literatura estudos que propõem análises para dar conta da variação dos padrões de concordância no interior do DP encontrada no PB. Na presente seção, discutiremos uma outra proposta, apresentada inicialmente em Lopes (2001) e aprofundada em Lopes (2005, 2015) que se baseia na teoria dos 4M, proposta por Myers-Scotton e Jake (2000a), para explicar o fenômeno da variação da concordância de número no âmbito do sintagma nominal do PB. Primeiramente, fornecemos um breve panorama da teoria dos 4M em si.

5.3.1.1 A Teoria dos 4M

A Teoria dos 4M foi apresentada no artigo de Myers-Scotton e Jake (2000a), no qual as referidas autoras apresentam evidências empíricas obtidas em estudos sobre *codeswitching* (doravante CS)⁵⁶, afasia de Broca⁵⁷ e aquisição de segunda língua, para sustentar um novo modelo de classificação de morfemas conhecido como *Modelo 4M* (do inglês, *4-M model*). Tal modelo enfatiza a noção de que

⁵⁶ Segundo Gumperz (1982, p.59 *apud* MOUTINHO, 2013, p.44), o *codeswitching* compreende a alternância entre códigos – línguas ou dialetos – ou seja, “a justaposição dentro da mesma interação discursiva de passagens pertencentes a dois sistemas ou subsistemas gramaticais distintos.” No entanto, vale lembrar que a literatura disponibiliza diversos trabalhos que tem se dedicado a descrever e explicar o fenômeno de CS. Para uma seleção de alguns dos trabalhos mais relevantes acerca desse tema, indicamos o estudo de Wakasa (2004).

⁵⁷ A afasia é um distúrbio de linguagem que afeta a capacidade de comunicação do falante. A afasia de Broca, em particular, decorre em função de lesão cerebral na região frontal do hemisfério esquerdo (denominada área de Broca). Um indivíduo acometido por tal tipo de afasia costuma apresentar uma produção de fala não fluente, mal-articulada, ainda que sua compreensão pareça estar relativamente preservada.

lemas⁵⁸ subjacentes a diferentes tipos de morfemas tornam-se relevantes em diferentes níveis da produção linguística. Nessa perspectiva, a classificação dos morfemas com base em um conjunto de distinções abstratas explicaria as distribuições dos tipos de morfemas. Em outros termos, de acordo com Myers-Scotton e Jake (2017, p.341), a proposta é um modelo acerca da natureza dos tipos de morfema em geral, bem como sobre a maneira como tais tipos se distinguem, são caracterizados e participam no contexto de um modelo de produção da linguagem.

É importante ressaltar que, segundo as próprias autoras, o modelo proposto investiga principalmente como os morfemas são acessados em um nível abstrato da produção linguística. Com isso, Myers-Scotton e Jake (2000a) propõem seu modelo dos *4M*, que consiste em uma classificação de quatro tipos de morfemas. A primeira distinção relevante nessa classificação engloba a separação entre morfemas de conteúdo (Ex.. Ns, verbos, Adjs, etc.) e morfemas sistêmicos ou gramaticais. Essa segunda classe, por sua vez é subdividida em dois tipos de morfemas: (i) morfemas sistêmicos precoces (Ex. determinantes, como o artigo definido *the* no inglês, o plural *-s* nos nomes em inglês e (ii) morfemas sistêmicos tardios. Esses últimos, por sua vez, se subdividem em: (i) morfemas tardios ponte (Ex. marcadores possessivos) e morfemas sistêmicos tardios '*outsider*' (Ex. afixos de concordância sujeito-verbo ou marcadores de Caso). Os quatro tipos de morfemas serão apresentados em maiores detalhes mais adiante.

Myers-Scotton e Jake (2000a) argumentam ainda que as distinções abstratas nas quais o modelo *4M* se baseia refletem a maneira universal como o léxico mental se organiza. Além disso, a noção de morfema seria uma forma de reconhecer que palavras possuem estrutura interna e que, por passarem por processos gramaticais, suas partes podem ocorrer em diferentes combinações e ainda serem reconhecidos como mapeamentos da estrutura conceitual.

⁵⁸ Segundo Myers-Scotton e Jake (2000a,p.1054-5), lemas “não são nem palavras, nem morfemas, antes são conjuntos de direcionamentos para a realização de palavras simples e complexas. [...] lemas são entradas específicas da linguagem no léxico mental e mediam as intenções no nível conceitual e a produção de estruturas gramaticais, incluindo estruturas de superfície”.

Nesse sentido, a classificação proposta no modelo dos *4M* se dá de acordo com os traços gramaticais abstratos que, segundo as autoras seriam inerentes à organização da linguagem, a saber, papéis temáticos, projeções máximas e coindexação de elementos. Dessa forma, a classificação se dá em decorrência da confluência de dois conjuntos de duas distinções abstratas:

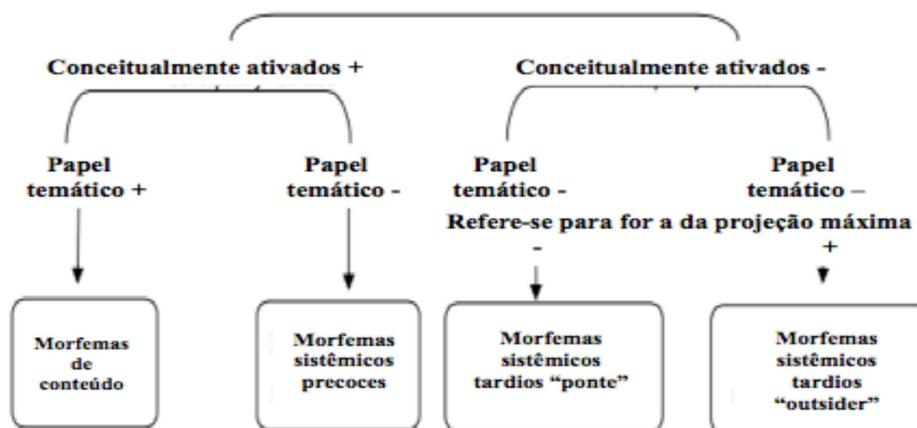
- (i) a primeira distinção entre [+/- conceitualmente ativado] se refere aos mecanismos pelos quais os morfemas são acessados, ou seja, primitivos que sustentam morfemas ativados no nível conceitual⁵⁹ ou no nível do lema possuem o traço [+ conceitualmente ativado], mas os que são ativados no nível do formulador⁶⁰ não possuem tal traço;
- (ii) o segundo conjunto de distinções diz respeito a duas propriedades da construção de estruturas sintagmáticas. A primeira consiste na distinção [+/- papel temático], ou seja, se o elemento atribui/recebe papel temático ou não. A segunda propriedade, por sua vez, se trata da distinção [+/- referir-se para fora de sua projeção máxima], que indica se os morfemas relacionam elementos dentro de um mesmo sintagma ou pertencentes a sintagmas diferentes.

Com isso, o modelo dos *4M*, postula os seguintes tipos de morfemas dispostos na figura (17):

⁵⁹Os lemas subjacentes aos morfemas, apesar de estarem presentes no léxico mental, tornam-se relevantes em diferentes níveis do processo de produção linguística. Os lemas eleitos diretamente no nível conceptual subjazem aos morfemas de conteúdo. Já os lemas que sustentam os morfemas sistêmicos tornam-se relevantes em dois níveis: os morfemas sistêmicos precoces no nível do lema e os morfemas sistêmicos tardios no nível do formulador.

⁶⁰Segundo Jake e Myers-Scotton (2009), no que se refere ao modelo de produção linguística, o formulador consiste no nível em que ocorrem os procedimentos de construção de estruturas específicas de língua. O formulador recebe direcionamentos enviados pelos lemas sobre como combinar unidades menores para formar constituintes maiores, hierarquicamente estruturados. Em outras palavras, ao formulador cabe a tarefa de ativar os procedimentos morfossintáticos que resultarão nos enunciados do nível de superfície.

Figura 17 - Classificação de morfemas



Fonte: Myers-Scotton (2002, p. 73 *apud* NAMBA, 2004, p.4)

Como observamos na figura (17) acima, os morfemas de conteúdo (do inglês, *content morphemes*) são aqueles que atribuem/recebem papel temático e são ativados no nível do lema. Dessa forma, eles são diretamente eleitos de acordo com a intenção do falante, sendo responsáveis por transmitirem o conteúdo semântico/pragmático. Quanto aos seus papéis sintáticos, os morfemas de conteúdo incluem substantivos – que recebem papel temático –, além de verbos, Adjs, preposições e advérbios – que, por sua vez, geralmente também estão envolvidos na atribuição papel temático.

No entanto, vale destacar que Myers-Scotton e Jake (2017, p.343) atentam para o fato de que a classificação de morfemas não é isomórfica ao *status* de categoria lexical. O traço [+/- papel temático], por exemplo, não define nenhum grupo específico de categorias lexicais. Para ilustrar esse ponto, Myers-Scotton e Jake (2000a) argumenta que, na categoria lexical dos pronomes, podemos encontrar alguns membros que são morfemas de conteúdo enquanto outros não o são. Ao observarmos os exemplos (18-20) abaixo, extraídos de Myers-Scotton e Jake (2000a, p.1059), podemos compreender melhor tal distinção. Enquanto pronomes pessoais em inglês, como *he* em (18), e *it* em (20) são morfemas de conteúdo em virtude de receberem papéis temáticos, pronomes expletivos – como é o caso de *it*⁶¹ em (19) –, se comportam de maneira diferente. Isso porque em sentenças como (19), *it* não recebe

⁶¹Importante lembrar que *it* também pode funcionar como pronome pessoal e, portanto, neste caso, é um morfema de conteúdo quando usado em sentenças como (20) *I have a dog. It is three years old* (Eu tenho um cachorro. Ele tem três anos de idade).

papel temático, o que pode ser evidenciado no fato de que, nesse contexto, o pronome não pode ser alvo de uma interrogação.

(18) a. I said that he went. (Eu disse que ele foi.)

b. I'm sorry. I didn't hear you. Who went? (Me desculpe. Não te ouvi. Quem foi?)

(19) a. I said it's raining. (Eu disse que está chovendo).

b. I'm sorry. I didn't hear you. *What's / who's raining? (Me desculpe. Não te ouvi. O que/ quem está chovendo?)

Na classificação das autoras, há também os morfemas sistêmicos⁶² (do inglês, *system morphemes*). Tais morfemas, dentre os quais podemos citar os itens funcionais e as flexões, distinguem-se dos morfemas de conteúdo por não atribuírem nem receberem papel temático. No entanto, os mesmos expressam relações entre morfemas de conteúdo e são essenciais na construção das estruturas gramaticais. Segundo Myers-Scotton e Jake (2000b), os morfemas de conteúdo e os morfemas sistêmicos se distinguem também por outros dois aspectos: (i) o estatuto que diz respeito à ativação conceptual (ou o nível em que os morfemas são ativados) e (ii) a maneira como suas formas participam da construção de constituintes maiores.

Os morfemas sistêmicos, por sua vez, podem ser subdivididos em dois tipos:

- (i) morfemas sistêmicos precoces (do inglês, *early system morphemes*) e
- (ii) morfemas sistêmicos tardios (*late system morphemes*).

No tocante aos morfemas de conteúdo, é possível compreender seu papel num modelo de produção da linguagem, uma vez que eles estão relacionados às intenções do falante, dizem respeito a entidades e predicados que ativam lemas subjacentes aos morfemas de conteúdo que constroem a estrutura temática. No entanto, conceitos semânticos e pragmáticos, tais como definitude e pluralidade, representados em (21) e

⁶²Jake e Myers-Scotton (2009, p.213) argumentam que o termo 'morfema sistêmico' foi introduzido por oferecer uma maneira mais abrangente e precisa de distinguir elementos do que fazê-lo em termos de "elementos de classe aberta X fechada" ou "elementos funcionais X elementos lexicais". Afinal, como defendem as referidas autoras, nem todos os elementos de classe fechada representam o mesmo tipo gramatical, bem como ocorre com os afixos, visto que nem todos se comportam da mesma maneira.

(22), respectivamente, entre outros, também se fazem necessários para codificar as intenções do falante. A tarefa de expressar tais conceitos fica a cargo de outro tipo de morfemas que também são conceitualmente relevantes e são denominados morfemas sistêmicos precoces.

(21) *I found **the** book that you lost yesterday.* (Eu encontrei o livro que você perdeu ontem.)

(22) *The books.* (os livros)

Morfemas sistêmicos precoces recebem este nome em virtude de serem ativados no nível do lema e compartilham com os morfemas de conteúdo o traço [+conceitualmente ativado]. Tal traço transmite a ideia de que morfemas sistêmicos precoces se unem a morfemas de conteúdo para expressar o feixe de traços semânticos e pragmáticos necessários para satisfazer as intenções do falante. Apesar de não atribuírem nem receberem papel temático, os morfemas sistêmicos precoces contribuem para o mapeamento da estrutura conceitual para o lema, assim como o fazem os morfemas de conteúdo.

Importante destacar que o traço [+conceitualmente ativado] também indica que, exceto pela forma fonológica expressa no nível de superfície, a informação crucial necessária para a forma dos morfemas se encontra disponível já no nível no lema. Além disso, tal traço distingue os morfemas de conteúdo e sistêmicos precoces dos morfemas sistêmicos tardios. Estes últimos recebem esse nome em função de serem acessados mais tarde no processo de produção linguística. No modelo *4M*, os morfemas sistêmicos tardios são subdivididos em dois tipos: “ponte” e “outsider”.

De maneira geral, a informação contida nos morfemas sistêmicos tardios é estritamente gramatical e não conceitual. Nesse sentido, esses dois tipos de morfemas sistêmicos tardios não são eleitos para completar um feixe de traços semânticos e pragmáticos de seus núcleos. Ao contrário, são atribuídos estruturalmente para indicar relações entre elementos quando da construção de um constituinte maior. Dizemos que a forma dos morfemas sistêmicos tardios é estruturalmente atribuída em virtude não só do papel que tais morfemas desempenham, como também do fato de que toda a informação morfossintática necessária para sua projeção não se encontra disponível até que direcionamentos sejam enviados para o formulador que, por sua vez, irá formar constituintes maiores.

Como vimos anteriormente, os morfemas sistêmicos tardios se dividem em dois tipos: os morfemas “ponte” e “outsider”. O morfema “ponte”, como o próprio nome indica, tem a tarefa de estabelecer relações entre elementos de construções específicas. Tal tipo de morfema não só desempenha uma função associativa numa oração, indicando a relação entre dois sintagmas – normalmente NPs –, como também os integra em um constituinte maior. O marcador possessivo *deem casa de mamãe*, por exemplo, consiste em um morfema sistêmico tardio “ponte”, uma vez que a preposição *de* é estruturalmente exigida pela gramática do português para integrar o NP *deem mamãe* ao NP *deem casa* no intuito de estabelecer uma relação de posse entre eles.

Em contrapartida, os morfemas “outsider” distinguem-se dos morfemas “ponte” em virtude de sua forma depender de informação gramatical fora do sintagma em que ocorrem. Em outras palavras, tal informação só é viabilizada quando o formulador envia instruções para o nível posicional/superfície para que as projeções máximas sejam unidas em um constituinte maior. Para ilustrar, Myers-Scotton e Jake (2000a) mencionam a flexão –s na terceira pessoa do singular do verbo no presente em inglês, cuja forma do morfema depende da coindexação do verbo com o NP sujeito, ou seja, um elemento que se encontra fora da projeção máxima do verbo. Com isso, o –s na terceira pessoa do singular dos verbos em inglês é considerado um morfema sistêmico tardio do tipo “outsider”.

Em resumo, os dois conjuntos de distinções abstratas, um com base no conceito e outro baseado nas propriedades de construção de estruturas sintagmáticas, resultam nos quatro tipos de morfemas postulados no modelo dos 4M esquematizados no quadro (9) abaixo:

Quadro 9- Classificação de morfemas segundo modelo dos 4M (MYERS-SCOTTON; JAKE, 2000a) e suas principais propriedades

Classificação de morfemas segundo modelo dos 4M de Myers-Scotton e Jake (2000a)			
Morfemas de conteúdo	Morfemas sistêmicos precoces	Morfemas sistêmicos tardios	
		“ponte”	“outsider”

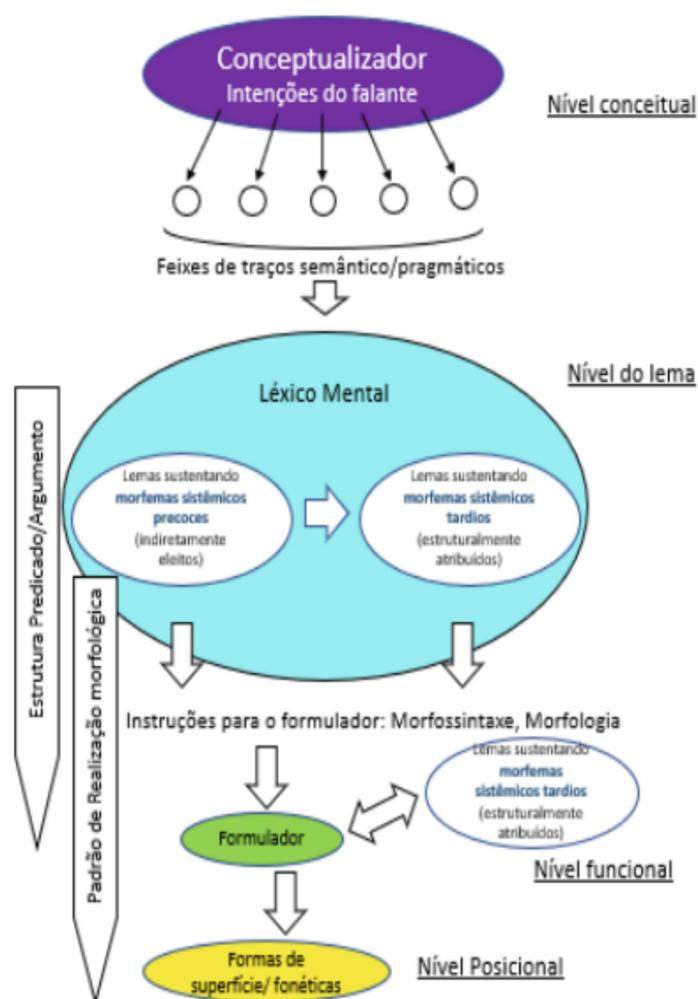
<ul style="list-style-type: none"> - ativados no nível do lema; - atribuem ou recebem papel temático; - “diretamente eleitos” por um feixe de traços semânticos/pragmáticos, mapeando a estrutura conceitual para o lema. - Ex. Ns, Adjs, Vs 	<ul style="list-style-type: none"> - ativados no nível do lema; - não atribuem ou recebem papel temático; - “indiretamente eleitos” porque os morfemas de conteúdo “apontam” para eles; - podem estar no mesmo lema que os morfemas de conteúdo ou não. - Ex. Ds, -S plural 	<ul style="list-style-type: none"> - ativado no nível do formulador; - não atribuem nem recebem papel temático; - não contribuem para as estruturas conceituais, apenas integram os morfemas de conteúdo em um constituinte maior; - dependem da informação dentro de sua projeção máxima. - Ex. marcadores possessivos tais como <i>de</i> em português 	<ul style="list-style-type: none"> - ativado no nível do formulador; - não atribuem nem recebem papel temático; - estruturalmente atribuídos no nível posicional/ de superfície; - sua forma depende de informação gramatical encontrada fora de sua projeção máxima. - Ex. -s na 3ª pes. Singular nos Vs em inglês
--	--	---	--

Fonte: Elaborado pela autora (2021)

A classificação apresentada anteriormente nos ajuda a compreender o processo de produção da linguagem proposto por Myers-Scotton e Jake (2000a, p.1055). Nesse modelo, o processo inicia-se com uma estrutura conceitual concebida como um nível pré-linguístico. A intenção de fala, por sua vez, ativa feixes de traços semânticos/pragmáticos específicos da linguagem. Tais traços selecionam lemas – entradas específicas da língua – no léxico mental. Nesse sentido, os lemas seriam os intermediários entre as intenções no nível conceitual e a produção de estruturas gramaticais. Os lemas diretamente eleitos no nível conceitual subjazem aos morfemas de conteúdo. No entanto, tais morfemas não são suficientes para realizar as intenções dos falantes e indicar as relações entre as estruturas lexicais abstratas e as formas de superfície. Para tal, os morfemas de conteúdo são acompanhados dos morfemas sistêmicos. O formulador (no nível funcional) recebe direcionamentos dos lemas sobre como formar constituintes maiores, hierarquicamente estruturados. Tais constituintes, então, são realizados no nível de superfície.

Sob o ponto de vista do modelo dos *4M*, Myers-Scotton e Jake (2000a, p.1060) argumentam que morfemas não são primitivos no léxico. Nessa visão, as entradas de lemas sustentam morfemas que são realizados somente nas formas de superfície, mas via mecanismos disponíveis em diferentes níveis de organização gramatical. Aliás, é a forma como diferentes mecanismos interagem que explica a diversidade observada nos morfemas de superfície. Nesse sentido, como observamos na figura (18) abaixo, temos o modelo de nível abstrato (do inglês, *Abstract Level model*), proposto por Myers-Scotton (2002, *apud* MYERS-SCOTTON; JAKE, 2017). O princípio central subjacente a tal modelo consiste na ideia de que a estrutura lexical abstrata pode ser descrita em três níveis. Tais níveis abrangeriam todas as estruturas linguísticas, ainda que a forma como seriam realizadas sejam específicas de cada língua. Sendo assim, o primeiro nível, denominado estrutura conceitual-lexical, diz respeito aos significados semânticos e pragmáticos. O segundo nível, denominado estrutura predicado-argumento, consiste nos princípios e processos que regem a estrutura de uma oração. Por fim, o terceiro nível – padrões de realização morfológicos –, refere-se a como as arquiteturas linguísticas são realizadas nas orações de nível de superfície no nível posicional. De acordo com Myers-Scotton e Jake (2017, p.349), quando tais níveis tornam-se salientes, eles resultam nas realizações no nível de superfície que se iniciam com os processos cognitivos com objetivos comunicativos.

Figura 18 - Diagrama do processo de produção linguística: ativação do lema e estrutura lexical abstrata, extraída de Namba (2004)



Fonte: Namba (2004, p.4)

Em Myers-Scotton e Jake (2017), ao revisitarem o modelo, as autoras propõem que o fator crucial que permite prever a distribuição de morfemas consiste no nível abstrato em que os tipos de morfemas são eleitos num modelo de produção linguística. Em outras palavras, o foco recai sobre a maneira como os morfemas são eleitos ao longo do processo de produção da linguagem e ainda em como tal eleição difere em função dos tipos de morfemas. Isto é, o objetivo principal na proposta mais recente é mostrar como a realização efetiva dos tipos de morfemas poderia estar ligada a um modelo de produção da linguagem.

Sendo assim, do ponto de vista empírico, as autoras pretendem explicar, por exemplo, como a frequência de um determinado tipo de morfema em casos de CS e

outros fenômenos de contato depende em parte do fato do morfema ser eleito no nível conceitual ou envolvido no nível do formulador, no qual o morfema é estruturalmente atribuído.

Para tal, as autoras lançam mão de dados de CS espanhol-ínglês, reportados em Jake, Myers-Scotton e Gross (2002, p.80), nos quais se observa uma predominância de Ds em espanhol no caso de DPs mistos, fato que, segundo as autoras, se dá em virtude de diferenças no nível abstrato em que os Ds são eleitos. Em outros termos, nos DPs mistos em que o D pertence a uma língua diferente da língua em que se encontra o N, a estrutura mais comumente observada é aquela formada por um D em espanhol, seguido de um N em inglês. A explicação apresentada para esse fenômeno considera que, uma vez que o D em espanhol possui um conjunto de traços-*phi* maior do que em inglês, informações sobre gênero e número devem estar disponíveis para compor a estrutura de outros elementos na oração. Nesse sentido, tais traços se mostram salientes já no nível da estrutura lexical-conceitual em tal língua. No entanto, os traços-*phi* gramaticais do inglês, incluindo pessoa e número, mas não gênero, podem estar salientes já no nível conceitual, porém, o complexo D total somente se torna saliente quando a concordância estruturalmente atribuída (AGR) é ativada no nível do formulador.

Faz-se necessário destacar também que, em espanhol, assim como em muitas línguas românicas, tais como o PB, os traços-*phi* abstratos que subjazem à informação de gênero e número nem sempre são marcados de maneira inequívoca nos Ns, mas são mapeados na realização de superfície dos elementos que coocorrem com os Ns, tais como os Ds. Nesse sentido, teríamos uma evidência para a proposta de que o -S de plural no PB não se configura como um único tipo de morfema, mas sim morfemas diferentes a depender do item que acompanha – como será discutido na seção a seguir.

O estudo conduzido por Jake, Myers-Scotton e Gross (2002) nos interessa em particular por abrir perspectivas para correlações entre tipos de morfemas e propostas no âmbito do PM, que serão exploradas mais adiante neste capítulo. Antes, porém, na próxima seção, abordaremos o tema da concordância nominal variável no PB com base na teoria *4M*. Isso porque tal teoria também se faz bastante relevante para a presente pesquisa em virtude de compartilharmos a ideia de que os diferentes tipos de morfemas postulados podem trazer aspectos relevantes para a compreensão das

diferentes realizações da concordância entre os elementos flexionáveis no âmbito do DP registradas nos *corpora* analisados nesta tese.

5.3.1.1.1 A Teoria dos 4M e a concordância de número variável no PB

No que se refere às pesquisas que versam acerca do fenômeno da concordância de número variável no âmbito do DP no PB, são poucos os estudos – pelo menos que tenhamos conhecimento –, que investigam o assunto com base nos pressupostos da Teoria dos 4M.

Dentre os raros trabalhos reportados sobre esse tema, tem-se por exemplo, o estudo de Lopes (2001, 2005, 2015), que parte dos pressupostos da teoria dos 4M, para tentar explicar o fenômeno da concordância nominal variável observada no PB. Segundo Lopes (2015, p.68), “o favorecimento ou desfavorecimento da concordância verbal e nominal de gênero e de número se dá em virtude da ordem de aquisição dos morfemas”. Essa visão se mostra compatível com a proposta de Myers-Scotton e Jake (2000a), segundo a qual, os morfemas de conteúdo seriam os primeiros a serem adquiridos pelo falante e os últimos a serem alvo de variação⁶³. Assim como os morfemas de conteúdo, os morfemas sistêmicos precoces – que, ocorrem junto aos Ns, os quais em português, são representados por exemplo, pelos artigos e outros morfemas sistêmicos precoces –, seriam acessados no mesmo nível que os morfemas de conteúdo, ou seja, no nível do lema e, por conseguinte, seriam adquiridos conjuntamente a eles (LOPES, 2015, p.69).

Além disso, Lopes (2001, 2005, 2015) defende que a definitude e a marca de número expressa nos Ds são inseridas no nível do lema. Por essa razão, tais morfemas seriam menos suscetíveis a sofrer qualquer tipo de variação nas línguas por serem ativados conceptualmente. Tal afirmação é sustentada por dados que indicam que a posição no âmbito do sintagma nominal que mais parece favorecer a marcação explícita de plural é à esquerda adjacente ao núcleo do sintagma – posição ocupada

⁶³ Quanto à questão da variação, Myers-Scotton e Jake (2000a) reportam uma pesquisa conduzida com afásicos de Broca a qual registrou que mais morfemas sistêmicos tardios estavam incorretos ou ausentes na produção dos participantes, quando comparados aos morfemas sistêmicos precoces ou aos morfemas de conteúdo. Segundo Myers-Scotton e Jake (2000a, p.1092), a vulnerabilidade observada nesse tipo de morfemas se dá em virtude de os mesmos não serem ativados conceptualmente, ou seja, pelo fato de que sua não ocorrência não comprometeria a transmissão das intenções do falante.

por morfemas sistêmicos precoces (artigos, demonstrativos) em português –, tanto no concernente a gênero quanto a número. Dessa forma, em linhas gerais, com base em dados de fala registrados em Salvador (BA), Lopes (2001) propõe a existência de uma relação entre a variação da concordância no sintagma nominal e a classe/posição relativa de cada elemento no sintagma – semelhante ao que observamos em Menuzzi (1994, p.130). Mais especificamente, nos dados analisados, observou-se que:

- (i) a posição à esquerda do núcleo nominal foi a maior favorecedora de marcas explícitas de plural, com destaque para a posição adjacente ao núcleo;
- (ii) os núcleos, por sua vez, apresentaram mais marcas de plural quando ocupavam a 1ª posição do sintagma, e
- (iii) a posição à direita do núcleo sintagmático se mostrou altamente desfavorecedora de marcas explícitas de plural.

A partir dos dados reportados em sua pesquisa, Lopes (2001, 2005, 2015) sugere que a marcação de plural e de gênero nos elementos flexionáveis que compõem o sintagma nominal no PB envolve dois tipos de morfemas, a saber, (i) um morfema sistêmico precoce no que tange aos elementos à esquerda adjacente⁶⁴ aos núcleos e nos elementos nucleares quando ocuparem a primeira posição do sintagma e (ii) um morfema sistêmico tardio, quando considerada a marca de plural e de gênero *expressa* nos elementos alocados nas demais posições relativas do sintagma nominal – elementos nucleares ocupando a segunda, terceira ou quarta posição do sintagma e elementos não nucleares à direita do núcleo -, ou seja, nos elementos que estabelecem uma relação puramente gramatical com os demais elementos que compõem o sintagma e por isso, serem mais suscetíveis à variação observada nos dados do PB.

Nesse contexto, podemos observar diferenças entre a proposta inicial feita por Myers-Scotton e Jake (2000a) na teoria dos *4M* e a análise postulada por Lopes (2001, 2005, 2015) acerca dos morfemas de plural. Enquanto na proposta inicial, Myers-Scotton e Jake (2000a) defendem que o morfema -S de plural no inglês seja considerado morfema sistêmico precoce, Lopes propõe uma subdivisão de tal morfema ao observar um comportamento diferenciado dos falantes quando da realização das marcas de plural em diferentes elementos no âmbito do DP no PB.

⁶⁴No entanto, vale destacar que, com exceção de artigos e possessivos, a autora não deixa claro quais classes ocupam tal posição nos dados analisados em sua pesquisa.

Sendo assim, a proposta de Lopes (2001) se mostra muito relevante para a nossa pesquisa, pois parece ser a teoria que melhor parece dar conta dos dados coletamos no âmbito desta tese, como será reportado na próxima seção.

No entanto, faz-se necessário destacar alguns aspectos importantes a serem questionados na proposta de Lopes (2001, 2005, 2015). Primeiramente, tendo-se em vista que sob uma das perspectivas propostas para aquisição da linguagem⁶⁵, as categorias lexicais são adquiridas antes de D, a hipótese que relaciona a classificação de morfemas com a ideia de que sistêmicos precoces são adquiridos conjuntamente aos morfemas de conteúdo, tal como apontado em Lopes (2001, p.103, 2015, p.69) é, no mínimo, problemática. A própria teoria dos 4M, na qual Lopes (2001,2005, 2015) busca argumentos para explicar o fenômeno da concordância nominal variável no PB, baseia-se – como visto anteriormente –, em resultados de pesquisas em aquisição de segunda língua (WEI, 1996 apud WEI, 2000) os quais sugerem que os morfemas de conteúdo seriam adquiridos antes dos morfemas sistêmicos precoces, e que estes seriam adquiridos antes dos morfemas sistêmicos tardios. Como veremos mais adiante, nossos dados também fornecem indícios que sustentam essa hipótese.

Além disso, a hipótese levantada por Lopes de que o traço de gênero em D seria do tipo sistêmico precoce se mostra igualmente problemática, uma vez que contraria a visão praticamente consensual na literatura – tal como apontado por Hornstein, Nunes e Grohmann (2004), Polinsky (2015) – de que a informação de gênero seja um traço inerente dos Ns e, em larga medida, arbitrária.

Antes, porém, apresentamos, no quadro (10) abaixo, uma síntese das diferentes propostas para a compreensão do fenômeno da concordância variável no PB, formuladas a a partir de modelos formais e resenhadas até aqui.

⁶⁵ Importante destacar que está sendo assumido aqui como “aquisição” o momento quando efetivamente já é observada na produção de criança a presença significativa de Ns, o que não costuma ser observado na mesma época no caso de Ds ou elementos de outra categoria.

Quadro 10 - Algumas hipóteses para explicar a ocorrência de concordância não redundante no PB

Autor	Hipótese
<p>Nunes (1963 <i>apud</i> FERRARI NETO, 2003)</p>	<p>A concordância variável estaria relacionada ao enfraquecimento do paradigma morfológico da flexão nominal de número, em decorrência do caráter instável – em termos fonéticos – do morfema –S, observado durante o processo de evolução da língua, ou seja, do latim até o estágio atual do PB.</p>
<p>Lopes (2001)</p>	<p>Tomando como ponto de partida a classificação de morfemas aventada pela Teoria dos <i>4M</i> (MYERS-SCOTTON; JAKE, 2000a), Lopes relaciona a variação observada no PB aos tipos de morfemas envolvidos e propõe que as marcas de plural nos elementos flexionáveis do sintagma nominal podem ser de dois tipos:</p> <ul style="list-style-type: none"> a) morfema sistêmico precoce (em elementos à esquerda adjacente ao núcleo e nos elementos nucleares em 1^a posição); b) morfema sistêmico tardio (a marca de plural nas outras posições relativas do sintagma nominal), que estabelece uma relação puramente gramatical entre elementos do sintagma.
<p>Magalhães (2004)</p>	<p>Propõe que o <i>locus</i> da interpretabilidade do traço de número seria nos elementos da categoria D e justifica sua defesa ao destacar que no PB, a omissão de marca explícita de plural nos demais elementos do DP não compromete a leitura plural do sintagma, tampouco compromete a sua gramaticalidade.</p>
<p>Costa e Figueiredo Silva (2006)</p>	<p>Com base na abordagem da Morfologia Distribuída, os autores propõem que o fenômeno estaria relacionado ao fato de o morfema associado com a pluralidade ser do tipo <i>singleton</i> no PB. Isto é, tal morfema se adjunge exclusivamente ao elemento que recebe a interpretação semântica relativa à informação de número. Em virtude de D ser o núcleo que liga o DP à sua interpretação em LF, é licenciado que o morfema de plural seja realizado somente em tal núcleo, podendo ser omitido nas demais categorias no componente aberto da sintaxe. Os autores propõem ainda que, diferentemente do PB, o PE apresentaria o morfema de plural como dissociado, sendo, portanto, obrigatório em todos os itens do DP.</p>
<p>Silva e Magalhães (2014)</p>	<p>Magalhães e Silva sugerem que a alternância entre os padrões de concordância observados no PB decorre de uma competição entre gramáticas, em virtude de encontrarmos na produção de um mesmo indivíduo, ocorrências de ambos os padrões de concordância.</p>

Fonte: Elaborado pela autora (2021)

5.3.1.2 Nova proposta de análise para a concordância nominal variável

Como mencionado anteriormente, a proposta de Lopes (2001), segundo a qual, a concordância nominal variável estaria relacionada ao previsto pela Teoria dos 4M, parece abrir interessantes perspectivas para dar conta do fenômeno da variação no âmbito do DP observado em nossos dados. Importante ressaltar que, diferente da abordagem de Myers-Scotton e Jake (2000a), a qual sugere que o morfema de plural consiste em um morfema sistêmico precoce, responsável por expressar conceitos tais como pluralidade, compartilhamos com Lopes (2001,2005,2015) a ideia de que o morfema de plural – tal como se observa nos dados coletados no âmbito desta tese –, corresponde a tipos de morfemas diferentes a depender do elemento em que é expresso. Nesse sentido, em sintagmas do tipo *As meninas inteligentes*, propomos que o morfema de plural em D é do tipo sistêmico precoce, porque entendemos ser no D que o morfema de plural ancora a semântica de pluralidade cuja interpretação irá se espalhar pelo sintagma, a despeito da presença ou ausência de realização morfofonológica desse traço em elementos para além de D (cf. Tabela (26)). Trata-se, nessa linha de raciocínio, de um elemento eleito precocemente, ou seja, ainda no nível do lema, o qual está associado às intenções do falante. Sendo assim, por possuírem o traço [+conceitualmente ativado], a informação crucial necessária (i.e. pluralidade) para a forma de superfície dos morfemas sistêmicos precoces já se encontra disponível no nível do lema. Em outros termos, compartilhando a proposta de Myers-Scotton e Jake (2000a), o morfema de plural no D no PB corresponderia a um morfema sistêmico precoce, responsável por veicular a mensagem que o falante quer expressar.

Por outro lado, tal como apontado por Lopes (2001), os morfemas de plural expressos nas categorias lexicais aqui analisadas (i.e. Ns e Adjs pospostos ao N), seriam classificados como morfemas sistêmicos tardios. A justificativa para tal classificação se dá em virtude da possibilidade observada no PB da omissão de marcas explícitas de plural em tais elementos. Em termos dos níveis presentes em um modelo de produção linguística, os morfemas de plural expressos nos Ns e Adjs, parecem obedecer a regras exclusivamente estruturais do sistema linguístico e, nesse sentido, não parecem ser essenciais para veicular as intenções do falante, no caso, de expressar pluralidade. Nesse sentido, tais morfemas seriam ativados mais tardiamente, a saber, no nível do formulador, isto é, são estruturalmente atribuídos e, como tal, sua

omissão não compromete a interpretação plural do sintagma, nem tampouco a gramaticalidade da sentença. Em outros termos, como morfemas sistêmicos tardios, são atribuídos estruturalmente para indicar relações entre elementos do sintagma, ou seja, por não serem eleitos para completar um feixe de traços semânticos de seus núcleos, os mesmos não contribuem para o mapeamento da estrutura conceitual para o lema. Nesse contexto, casos de DPs não redundantes tais como *os menino*∅, observados em nossos dados, são considerados gramaticais porque o morfema necessário para a interpretabilidade do traço de número da expressão linguística está plenamente realizado.

Dessa forma, empiricamente, observamos que nossos dados naturalísticos são compatíveis com a ideia defendida não só por Lopes (2001, 2005, 2015), como também por Guy (1981), Fernandes (1996), Carvalho (1997), Scherre (1988, 1994), Andrade (2003), Silva (2011), Ribeiro (2013), Costa e Teixeira (2017) e outros – e que contribui para sustentar a proposta de que: (i) o morfema de plural no PB recebe classificações diferentes a depender do elemento que acompanha e (ii) existe uma relação entre a variação da concordância no âmbito do sintagma nominal e a posição de cada elemento que compõe tal sintagma. A tabela (26) abaixo apresenta evidências de que a posição⁶⁶ do elemento dentro do sintagma nominal desempenha um papel relevante quando da realização dos padrões de concordância nos dados naturalísticos – produzidos tanto por crianças quanto por adultos – aqui analisados.

⁶⁶Optamos por substituir a expressão “posição relativa” por “posição” pois, Lopes (2001) trata da posição dos elementos dentro do sintagma nominal em relação ao núcleo sintagmático que, para a referida autora, trata-se do N. Nossa proposta, por outro lado, assume que o D tem a função de encabeçar o sintagma, tendo o N como seu complemento.

Tabela 26 - Ocorrências de marcação explícita de plural considerando a variável posição dos elementos dentro do DP – dados naturalísticos

Ocorrências de marcação explícita de plural considerando a variável <i>posição dos elementos dentro do DP</i> (dados naturalísticos)			
Participantes	Crianças	Adultos	Total
Posição morfológicamente marcada			
1ª posição (à esquerda do N) <i>Ex. As bonecas</i>	139/139 (100%)	197/197 (100%)	336/336 (100%)
2ª posição (N) <i>Ex. Aqueles bonés</i>	116/139 (83%)	158/197 (80%)	274/336 (82%)
3ª posição (à direita do N – Adj) <i>Ex. os piratas pequenininhos</i>	05/07 (71%)	6/6 (100%)	11/13 (85%)

Fonte: Elaborado pela autora (2021)

Como podemos ver no quadro acima, houve uma tendência pela marcação de plural explícita no elemento que ocupa a primeira posição do sintagma – nos nossos dados naturalísticos sempre ocupada por um elemento da categoria D. Observa-se também que tal tendência diminui de forma considerável no elemento que ocupa a 2ª posição – aqui representado por um N. Já a 3ª posição – ocupada por um Adj, apresenta um aumento percentual de marcação explícita nos dados dos adultos – o que, em princípio, parece contrariar os resultados encontrados por Lopes (2001) e outros. No entanto, vale destacar que o número de ocorrências de sintagmas com três elementos foi bastante reduzido. Sendo assim, nossos dados naturalísticos não fornecem informação suficiente para que conclusões mais precisas sobre tal comportamento possam ser aventadas.⁶⁷

⁶⁷A partir desses dados, levantamos algumas questões que consideramos importantes no que diz respeito à marcação explícita entre os itens que compõem o sintagma:

O percentual de marcação explícita de plural no Adj estaria condicionada à posição linear ocupada por esse elemento dentro do sintagma - anteposta ou posposta ao N? Ou seja, será que a taxa de marcação explícita em tal elemento teria sido maior se o Adj estivesse anteposto ao N?

Como pudemos ver ao longo deste trabalho, nossos dados não nos permitem tecer conclusões precisas a esse respeito. Mas, se para essa discussão, tomássemos como base os dados de Lopes (2001, 2005, 2015), poderíamos hipotetizar que o Adj preposto ao N – *pobres homens* - apresentaria um percentual maior de marcação explícita, quando comparado ao Adj posposto

No intuito de sanar tal limitação, recorreremos aos dados experimentalmente obtidos e representados na tabela (27) abaixo):

Tabela 27 - Ocorrências de marcação explícita de plural considerando a variável posição dos elementos dentro do sintagma – dados experimentais

Ocorrências de marcação explícita de plural considerando a variável <i>posição dos elementos dentro do sintagma</i> – dados experimentais			
Participantes	Crianças	Adultos	Total
Posição morfológicamente marcada			
2^a posição (N) Ex. <i>Quatro bonés</i>	533/764 (70%)	649/672 (97%)	1182/1436 (82%)
3^a posição (à direita de N – Adj1) Ex. <i>Dois cajus amarelos</i>	218/392 (56%)	444/513 (87%)	662/905 (73%)
4^a posição (à direita de N – Adj2) Ex. <i>Dois cajus amarelos grandes</i>	61/123 (50%)	114/163 (70%)	175/286 (61%)

Fonte: Elaborado pela autora (2021)

Os dados experimentais – que incluem um número significativamente maior de ocorrências analisáveis – se mostraram mais compatíveis com a proposta arrolada acima acerca da possível relação entre as posições dos elementos dentro do sintagma e a marcação explícita de plural, com uma maior tendência a marcar a pluralidade no elemento que ocupa a segunda⁶⁸ posição do sintagma – nos nossos dados experimentais sempre ocupada por um N – diminuindo de forma gradativa, sendo que a 4^a posição – à direita do N – configura-se, em linhas gerais, como a que menos concentra marcas de plural.

Nesse sentido, consideramos que nossos dados fornecem evidências compatíveis com a hipótese de que lemas subjacentes a diferentes tipos de morfemas se tornam salientes em níveis distintos ao longo do processo de produção linguística e que esse fato pode explicar suas diferentes distribuições no nível de superfície. Em

ao N – *homens pobre-*, ao passo que o D posposto ao N tenderia a apresentar um percentual menor de marcação quando comparado ao D preposto ao N.

⁶⁸Nessa análise dos dados experimentais, estamos considerando os elementos que ocupam a 2^a, 3^a e possivelmente, 4^a posição do sintagma. Os dados de 1^a posição não foram considerados em virtude de que esta posição era ocupada por um numeral que, por sua vez, não apresenta traço formal de número.

outros termos, no que concerne à nossa pesquisa especificamente, o morfema –S de plural manifesto em D, por ser do tipo sistêmico precoce, torna-se saliente ainda no nível do lema, ou seja, está relacionado à intenção do falante e, por conseguinte, se torna mais propenso a ser expresso explicitamente do que o –S em N ou em Adj que, por sua vez, em virtude de ser do tipo sistêmico tardio é ativado posteriormente e relacionado exclusivamente à estrutura gramatical. Sendo assim, a eventual omissão de sua realização morfofonológica não prejudica a leitura plural do sintagma.

Lopes (2005) chama a atenção ainda para um outro aspecto importante destacado pelo modelo dos *4M*, que também é relevante para nosso estudo. Mas, desta vez, diz respeito especificamente à aquisição da linguagem. Neste contexto, o modelo pressupõe que a aquisição dos diferentes tipos de morfemas ocorra em estágios distintos, a saber:

- 1) Na aquisição das línguas, os morfemas de conteúdo configuram-se como o primeiro tipo de morfema a ser adquirido;
- 2) Já no tocante aos morfemas sistêmicos, os precoces são adquiridos antes dos tardios;
- 3) No que diz respeito aos morfemas sistêmicos tardios, os morfemas “ponte” são adquiridos antes dos morfemas do tipo “*outsider*”.

Tais pressuposições são compatíveis com parte dos dados coletados em nossa pesquisa. Ao compararmos dados coletados no estudo longitudinal, com crianças residentes na zona rural de Juiz de Fora, observamos diferenças percentuais entre o uso de morfemas associados à marcação de número no âmbito do DP (os morfemas sistêmicos precoces no caso de D e tardios “ponte” no caso dos N e Adj) e também em relação aos morfemas de plural associados à concordância verbal (os denominados “*outsiders*”). Tais diferenças vão na direção dos pressupostos levantados por Lopes (2005), segundo os quais morfemas sistêmicos precoces, em virtude de serem conceitualmente ativados no nível do lema, seriam não somente aprendidos antes dos morfemas tardios, como também menos propensos à variação (MYERS-SCOTTON; JAKE, 2000a, p.1091; LOPES, 2015, p.70). Da mesma maneira, entre os morfemas sistêmicos tardios, os morfemas do tipo “ponte” – nos nossos dados representados pelo morfema de plural em itens das categorias N e Adj –, que estabelecem uma relação no interior do DP, seriam adquiridos antes dos morfemas tardios “*outsiders*” –

representados nos dados pelo morfema de plural expresso no verbo. Os dados naturalísticos da produção infantil de DPs com três elementos obtidos nos *corpora* analisados na presente pesquisa, estão dispostos na tabela (28) abaixo, que sistematiza os tipos de morfemas descritos e os respectivos percentuais de marcação explícita de plural:

Tabela 28 - Percentual de marcação explícita de plural considerando o tipo de morfema associado aos afixos de plural no PB – crianças Corpus B

Percentual de marcação explícita de plural considerando o tipo de morfema associado aos afixos de plural no PB				
	Morfema sistêmico precoce (no D)	Morfema sistêmico tardio ponte (no N)	Morfema sistêmico tardio ponte (no Adj)	Morfema sistêmico tardio “outsider” (no sintagma verbal)
Exemplo	OS	jacarezinhoS	felizeS	partiram.
Percentual	100%	71%	71%	(30%)

Fonte: Elaborado pela autora (2021)

Vale esclarecer que a análise apresentada acima configura-se como um ponto de partida para explicarmos o fenômeno da concordância nominal variável observada no PB. No entanto, não nos parece por si só suficientemente completa. Neste sentido, para darmos prosseguimento à nossa proposta de trabalho, na próxima seção faremos uma apresentação de propostas mais específicas a partir de alguns aspectos envolvidos na interface entre a classificação de morfemas e PM que pensamos ser de grande relevância para a nossa pesquisa.

5.3.2 Interfaces do modelo de classificação de morfemas com Minimalismo e desdobramentos para a presente pesquisa

Como vimos na seção anterior, os dados coletados em nossa pesquisa são compatíveis com a interpretação teórica que parte dos postulados da Teoria dos *4M* para explicar o fenômeno da concordância variável de número no PB. Nessa seção, apresentamos nossa proposta segundo a qual, a concordância nominal não redundante observada no PB é licenciada pelos tipos de morfema – seguindo a ideia geral de

Lopes (2001,2005, 2015) –, bem como pelas relações que eles estabelecem no sistema de *Agree* (FRAMPTON; GUTTMAN, 2000; PESETSKY; TORREGO, 2007).

Antes, porém, faz-se necessário destacar que a Teoria dos 4M e o PM não são teorias totalmente compatíveis, uma vez que são teorias de naturezas diferentes. Em outros termos, enquanto PM consiste em um modelo de computação linguística, ou seja, de capacidade gerativa, a Teoria dos 4M, por sua vez, configura-se como um modelo de acesso dos morfemas, i.e. uma teoria de processamento. No entanto, pensamos ser bastante relevante propor um diálogo dessa natureza, levando em conta, por exemplo, a relevância dos sistemas de Interfaces e desempenho no PM.

Feita essa ressalva, o primeiro aspecto que gostaríamos de destacar em nossa análise diz respeito ao próprio mecanismo de *Agree*, uma vez que alguns autores defendem a existência de mecanismos de concordância distintos para o âmbito do sintagma nominal e para o nível sentencial. No entanto, em consonância com as propostas arroladas em Magalhães (2004), Baker (2008) e Danon (2010), entre outros, defendemos aqui um mecanismo de concordância único para dentro ou para fora do âmbito do DP. Isso equivale a dizer que as restrições impostas pelo sistema para que *Agree* possa acontecer são as mesmas dentro e fora do DP, a saber:

- (i) deve haver uma relação de c-comando entre sonda e alvo;
- (ii) não pode haver um elemento interveniente entre sonda e alvo em termos de traços-*phi* disponíveis;
- (iii) sonda e alvo precisam estar no mesmo domínio de fase;
- (iv) o alvo deve estar ativo para o sistema computacional, isto é, deve ainda apresentar um traço não valorado (no caso da concordância verbal, assume-se que seria o traço de *Caso*).

Um outro aspecto que aproxima o nosso estudo ao desenvolvido por Baker (2008, p.44), diz respeito à ideia de que a definição de quais elementos são valorados no sistema não deve ser arbitrária, uma vez que tal ausência de sistematicidade falharia no reconhecimento de padrões:

[...] the idea that probe heads can be arbitrarily specified as having certain feature slots as opposed to others. Stipulating what unvalued features a given head has on a case-by-case basis does not capture

the systematic differences in how verbs, adjectives, and nouns behave with respect to agreement.

Em termos gerais, no sistema de *Agree*, a questão do que seja valorado ou não valorado dentro do PM pode ser considerada um tanto arbitrária. Em contrapartida, Baker (2008) defende um sistema de valoração motivada pela natureza das categorias lexicais. Segundo autor, o N, por exemplo, possui traço de número valorado em função de ser um elemento que tem um índice referencial. Para Baker (2008, p.31) somente os elementos que têm índice referencial – segundo ele, os Ns –, podem carregar traços-*phi* intrínsecos, ou seja, traços pré-especificados para pessoa, número e gênero. Nessa perspectiva, os traços de gênero e número chegariam no D e no Adj via *Agree*. No que se refere especificamente ao sistema aqui proposto, nossa análise segue uma direção diferente em virtude de se basear em dados de variação do PB, como fez, por exemplo, Magalhães (2004), que indicam que o traço de número estaria em D. Além disso, a nossa proposta, é desenvolvida a partir do modelo de classificação de morfemas do modelo dos *4M*, ou seja, o que é valorado, e o que é não valorado está correlacionado com os diferentes tipos de morfema que o –S de plural no PB pode representar. Nesse contexto, acreditamos que, quando aliamos a noção de valoração à ideia de diferentes tipos de morfema, a valoração deixa de ser vista como algo aleatório ou arbitrário e passa a ter uma motivação natural. Em síntese, no sistema que estamos propondo, a valoração, assim como em Baker (2008), também não é arbitrária e, apesar de não seguirmos a mesma direção de Baker (2008), compartilhamos com o referido autor a ideia de que a valoração tenha uma motivação.

No que diz respeito à direcionalidade do mecanismo de *Agree*, é importante ressaltar que há autores que defendem que a direcionalidade proposta em Chomsky (2000, 2001) – de acordo com o qual a sonda, em uma relação de c-comando, sempre ocupa um lugar hierarquicamente acima do alvo na estrutura – precisaria de ajustes. Entre eles, por exemplo, as propostas de Adger (2002), Baker (2008) e Danon (2010) são representativas dessa questão. Em linhas gerais, tais autores hipotetizam que desde que haja a relação de c-comando, elementos podem sondar alvos mais acima na estrutura. Embora as evidências de uma bidirecionalidade no *Agree* levantadas por esses e outros autores sejam interessantes, o sistema que propomos a seguir parece funcionar de forma adequada com uma abordagem mais canônica nos moldes da

direcionalidade proposta Chomsky (2000, 2001). Uma vez, então, que nossos dados não se mostram como evidência contra a direcionalidade defendida por Chomsky (2000, 2001), estamos assumindo que *Agree* aconteça de cima para baixo⁶⁹, ou seja, com a sonda c-comandando o alvo.

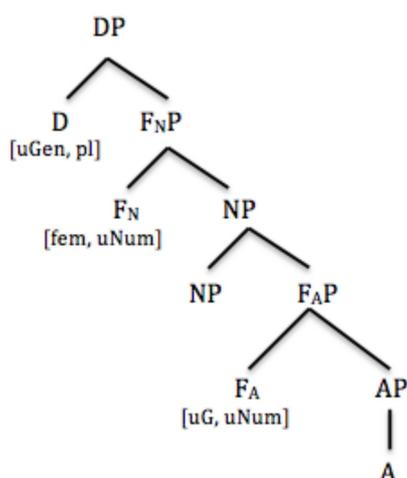
Além disso, é importante enfatizar que, com Danon (2010), propomos que nenhum dos núcleos que compõem o sintagma DP entram na derivação sintática com um conjunto completo de traços-*phi*. (DANON, 2010, p.7-8).

[...] identifying all nominal phi-features with a single head is a problematic approach. It seems that despite the great amount of disagreement regarding the exact number and nature of functional categories within the noun phrase, most studies on noun phrase structure since the early 1990s point toward a ‘distributed’ view of phi-features. I will therefore assume that phi-features are not all generated on the same head within the DP; specially, neither N nor D enters the derivation with a full set of valued phi-features.

Passemos, então, a observar a estrutura de DP que propomos na figura (19), a partir da qual podemos motivar as assunções teóricas assumidas na presente pesquisa.

Figura 19 - Estrutura de DP proposta na presente pesquisa

Ex. As bonecas vermelhas



Fonte: Elaborado pela autora (2021)

⁶⁹ Acreditamos que em face de novos dados e novas informações que precisem ser acomodadas, nosso sistema ainda funcionaria bem, caso fosse adotada uma direcionalidade ao contrário do sistema proposto em Chomsky (2000).

Sobre a posição do Adj, é importante apontar que, em nossos dados, em virtude de tal elemento ser do tipo atributivo, assumimos que ele ocupa uma posição de adjunto, seguindo propostas tais como a de Adger (2002, p.227) e Baker (2008, p. 49). Vale destacar que, do ponto de vista teórico, especialmente no âmbito do PM, em que as distinções entre os níveis de projeção – mínimo, intermediário e máximo – são abandonadas como noções primitivas, não há um consenso na literatura acerca do licenciamento formal desses elementos na estrutura sintática. Na proposta de Adger (2002), por exemplo, adjuntos são elementos anexados à estrutura da sentença, mas não por meio de checagem de traços. No entanto, uma vez que *Merge* é entendido no Minimalismo como uma operação essencialmente desencadeada por traços e o Adj atributivo não entra na estrutura para satisfazer tais traços, assumimos como Adger (2002) que tal tipo de Adj desempenhe o papel de modificador. Sendo assim, o adjunto não satisfaz nenhum traço de c-seleção, propriedade que o impede de projetar seu rótulo na estrutura sintática.

O mecanismo de inserção do adjunto, segundo Adger (2002, p.89), é mapeado através de um subtipo de *Merge*, a que o autor se refere como *Adjoin* e cuja principal propriedade seria uma inserção não desencadeada por traços:

So adjuncts are not incorporated into the sentence by Merge. We will assume, instead, that there is another basic operation, which we will call Adjoin, and which, unlike Merge, does not need to be triggered. Adjoin is somewhat like the version of Merge that were objected above, which inserted one syntactic object into another, although it is far more restricted. Adjoin inserts a phrasal object into another phrasal object at its outermost level. It does not create a new object, it expands one of the old ones by stretching its outermost layer into two parts and inserting the adjoined object between them.

Na busca de uma formalização para o licenciamento de adjuntos no PM, Chomsky (2000, 2004) propõe uma diferença entre os tipos de *Merge* que licenciam argumentos e adjuntos, a saber, *Set-Merge* e *Pair-Merge*, respectivamente. De acordo com Chomsky (2004, p. 117-118):

For structure building, we have so far assumed only the free symmetrical operation Merge, yielding syntactic objects that are sets, all binary: call them simple. The relations that come "free" (contain, c-command, etc.) are defined on simple structures. But it is an empirical fact that there is also an asymmetric operation of

adjunction, which takes two objects α and β and forms the ordered pair, α adjoined to β . Set-merge and pair-merge are descendants of substitution and adjunction in earlier theories. Given the basic properties of adjunction, we might intuitively think of α as attached to β on a separate plane, with β retaining all its properties on the "primary plane," the simple structure.

Focando especificamente na operação que concatena adjuntos, o *Pair-Merge*, trata-se, segundo Chomsky (2004), de uma operação assimétrica que toma dois objetos e retorna um par ordenado, o que, basicamente significa que não há a alteração de rótulo na estrutura sintática a partir dessa anexação.

Nesse mesmo sentido, Adger (2002) destaca que a evidência de que adjuntos nunca sejam núcleos de suas projeções seria exatamente o fato de que um constituinte ao qual um adjunto é incorporado se comporta, em termos distribucionais, da mesma maneira que se comportaria na ausência de um adjunto. Dessa forma, em termos do licenciamento do Adj, nossa proposta segue Adger (2002) e Baker (2008) e contraria aquela apresentada em Magalhães (2004), uma vez que na nossa estrutura o Adj atributivo se comporta como adjunto e, como tal, não projeta seu rótulo na estrutura sintática.

No que concerne à estrutura interna do DP, vale destacar que, apesar de haver uma visão bastante consensual na literatura sobre a existência de um conjunto de projeções funcionais acima do núcleo lexical N na estrutura sintática, ainda há muita discussão acerca de quantas e quais sejam. No concernente a tais projeções funcionais, como pode ser visto em nossa estrutura, compartilhamos com Baker (2008) a ideia de que a concordância não envolve traços de projeções lexicais e sim, traços de projeções funcionais. Tais projeções funcionais, na verdade, constituem, nos termos de Grimshaw (1991), as respectivas projeções estendidas das categorias lexicais. Sob essa perspectiva, podemos dizer que não são os traços de N que concordam com os traços de Adj e sim, que são os traços da projeção funcional de N que concordam com os traços da projeção funcional de Adj. Afinal, seguimos Baker (2008) na proposta de que traços de pessoa, número e gênero não são especificados nas categorias lexicais. Ao contrário, são especificadas nas funcionais – na nossa estrutura, representados por F_A ⁷⁰, F_N , e D. Nesse sentido, acreditamos estar

⁷⁰ Quando Baker (2008) propõe as projeções funcionais F_A e F_N , ele demonstra ter alguns *insights* do que seriam tais projeções, mas o referido autor acaba por não propor uma definição específica para as mesmas.

assumindo uma postura teórica mais simétrica, uma vez que o núcleo D, comumente assumido como um elemento envolvido em relações de *Agree*, é um núcleo de natureza funcional. É esse núcleo, por exemplo, que propomos ser o *locus* do traço de número no PB – tal como defendido por Magalhães (2004) –, o que equivale a dizer que tal elemento carrega um traço de número valorado desde o início da derivação, por força da própria formação da Numeração a partir da qual a sentença será construída, o que, conseqüentemente, o torna interpretável em ambas as interfaces: em LF em PF. Em LF tal elemento desencadeia a interpretação de singular/plural, enquanto em PF ele será obrigatoriamente o traço pronunciado na interface⁷¹. Essa abordagem teórica é, então, nos moldes do modelo dos *4M*, compatível com a ideia de que, o morfema –S de plural no núcleo D seja do tipo sistêmico precoce, responsável por veicular as intenções do falante.

F_N , por sua vez, seria a projeção funcional associada à categoria N, a qual possui traço de gênero valorado e traço de número não valorado – ideia que compartilhamos com Carstens (2000), Magalhães (2004) e Danon (2010), dentre outros. Assim como acontece em variantes não padrão no PB, em Danon (2010), observam-se exemplos de línguas – como o havaiano e o valão⁷² – em que não há morfologia de plural nos Ns. Já F_A seria a projeção funcional que carrega informações do Adj e entra na derivação não valorado para gênero e para número – uma visão bem consensual na literatura. No que diz respeito à classificação de morfemas, o morfema –S de plural nessas projeções – F_N e F_A – é do tipo sistêmico tardio e como tal, atribuído estruturalmente, ou seja, associado à estrutura linguística, tornando-se relevantes mais tardiamente no modelo de produção linguística.

Para articular a valoração de traços no interior do DP, consideramos que o sistema de Frampton e Guttman (2000) e Pesetsky e Torrego (2007) – que defende que o mecanismo de concordância envolve compartilhamento de traços – se mostra o mais interessante para explicarmos as relações de concordância. Com isso, observando nossa estrutura, assumimos que F_A entra na derivação com o traço de número e de gênero não valorados. F_N , por sua vez, entra na derivação valorado para gênero, mas não valorado para número, como mostra a figura (20a). Por essa razão,

⁷¹Vale ressaltar que fica como questão a ser explorada, o estatuto dos nomes nus, que não apresentam D fonologicamente realizado.

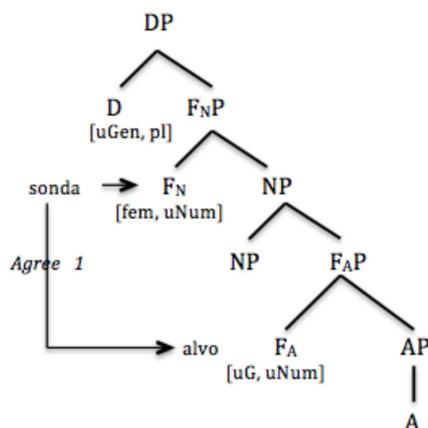
⁷²Valão é uma língua românica falada em grande parte da Valônia (Bélgica) e, em menor escala, na capital Bruxelas e algumas aldeias perto de Givet, norte da França.

F_N , precisa sondar um elemento abaixo na estrutura para valorá-lo e encontra F_A . Dessa forma, F_A e F_N entram em relação de *Agree*, F_A tem seu traço de gênero valorado e ambos (F_A e F_N) compartilham o traço de número ainda não valorado (figura 20b). Isso é possível porque, como visto anteriormente, na abordagem de compartilhamento, o alvo da operação de *Agree* não carrega necessariamente a versão valorada de um traço, ou seja – no caso da nossa estrutura –, o traço não valorado da sonda F_N pode se unir ao traço também não valorado do alvo F_A . Isso porque, nessa perspectiva, *Agree* é cega para o valor dos traços. Em seguida, D entra na derivação e, por ter um traço de gênero não valorado, D sonda o F_N para valorá-lo (figura 20b). Dessa segunda operação de *Agree*, obtém-se a valoração do traço de número compartilhado entre F_N e F_A , bem como do traço de gênero de D (figura 20c):

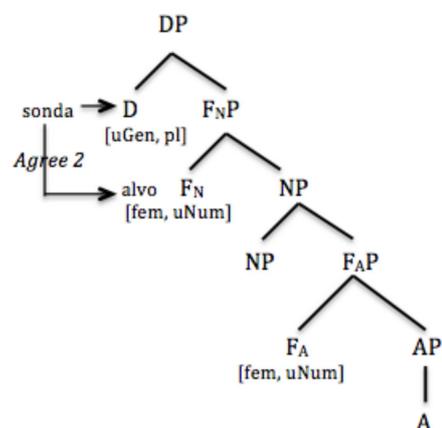
Figura 20 - Explicando a operação de *Agree* no nosso sistema

Ex. As bonecas vermelhas

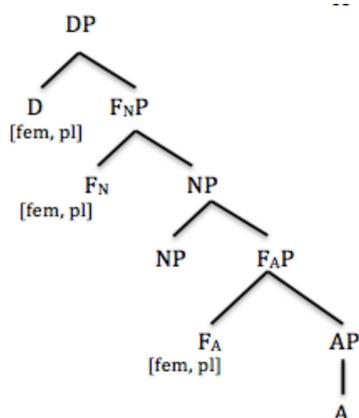
a)



b)



c)



Fonte: Elaborado pela autora (2021)

Dessa forma, notamos que nossa proposta se mostra diferente da de Magalhães (2004) quanto ao número de operações de *Agree*. No sistema de Magalhães (2004), nota-se que a valoração dos traços se dá por meio de três operações de *Agree*: (i) entre N e Adj; (ii) entre Adj e D e (iii) entre D e N. No nosso sistema, por sua vez, conseguimos resolver o problema da valoração com apenas duas operações: (i) entre o F_N e F_A e (ii) entre D e F_N – como pode ser visto na estrutura representada na figura (20) acima. Sendo assim, acreditamos que nossa proposta se mostra mais coerente com a premissa do PM de se buscar um sistema mais econômico no que tange ao número de operações sintáticas ocorridas ao longo da derivação⁷³.

Notamos ainda que nosso sistema sustenta a ideia proposta em Danon (2010) de que o modelo de valoração de traços de Chomsky (2000) se mostra incompatível com a visão de distribuição de traços-*phi* interna ao D. Afinal, de acordo com Danon (2010), nem D nem N entram na derivação com o conjunto de traços-*phi* completo. No entanto, podemos perceber na nossa estrutura também que, apesar de os traços-*phi* se originarem de núcleos diferentes dentro do sintagma, todos são coletados no nível da projeção máxima, ou seja, o DP. Vale lembrar que isso é possível em virtude de se adotar a proposta de compartilhamento de traços, segundo a qual, não há deleção de traços após a valoração. Tal fato é relevante, uma vez que, no sistema de valoração de traços proposto por Chomsky (2000), ao se estabelecer a operação de concordância entre D e N, os traços interpretáveis de N valoram e deletam os traços não interpretáveis de número e gênero de D. Nessa perspectiva, o DP nunca vai ser completo em termos de traços, o que, conforme detectado por Danon (2010) impossibilitaria a operação de *Agree* entre T e D. Nesse contexto, observamos a relevância da proposta de Pesetsky e Torrego (2007) de que se deva fazer uma dissociação entre as noções de valoração e interpretabilidade. De acordo com tal

⁷³ Além dos traços que estamos tratando no nosso sistema de *Agree*, pensamos que os traços de Caso e pessoa também devam ser considerados, já que eles, de alguma maneira, fazem parte da projeção nominal. O traço de Caso, por exemplo, tem relação com a questão da visibilidade. No entanto, a Teoria de Caso consiste em um tópico de muita discussão depois da virada minimalista. O traço de pessoa, por sua vez, existe no DP, mas parece que ele não está envolvido nas relações de *Agree* no interior desse sintagma. Para Danon (2010), por exemplo, tal traço seria do núcleo D e, só estaria envolvido nas relações de *Agree* entre o DP e o núcleo T. Nesse sentido, em virtude de estarmos enfocando a concordância no âmbito do sintagma nominal e tais traços serem relevantes para os casos de concordância entre DP e outros núcleos fora do sintagma, optamos por nos restringir a explorar as propriedades dos traços envolvidos na relações de *Agree* dentro do DP.

visão, para o sistema computacional basta que ocorra a valoração dos traços que vieram do léxico não especificados, sem se preocupar com sua interpretabilidade. Em outros termos, para a sintaxe, é relevante que os traços sejam valorados antes de *Spell-Out*. Afinal, segundo essa visão, a interpretabilidade seria determinada fora do sistema computacional e, como tal, teria a ver com relações de condições de interface. Sendo assim, a distinção entre interpretável e não interpretável seria relevante, por exemplo, para o fato de um traço de um determinado item lexical contribuir para a interpretação semântica do referido item, tal como acontece com os traços de pessoa e número no DP.

Passando, então, para os dados nos quais o numeral está presente, gostaríamos de tecer algumas considerações a respeito desse elemento. Primeiramente, vale destacar que, ao buscarmos por pesquisas acerca dos numerais, notamos que a sintaxe e a semântica de tais elementos consistem em questões ainda em debate na literatura. No que se refere à estrutura sintática do numeral, por exemplo, vemos em autores como Ionin e Matushansky (2018) que se trata de uma questão bastante instável na literatura. Há autores que defendem que os numerais podem ser núcleo de uma projeção funcional, enquanto outros argumentam que ele deva, na verdade, ocupar a posição de especificador de alguma projeção funcional ou ainda os que defendem que o numeral possa ser um adjunto de um NP.

Para ilustrar a complexidade da semântica dos numerais, retomamos os dados infantis analisados nesta pesquisa – tanto naturalísticos quanto experimentais – acerca do possível efeito da constituição morfofonológica dos numerais (terminados ou não em –S) para a realização dos padrões de concordância. Como visto nos capítulos 3 e 4, os dados infantis apresentaram diferença entre os tipos de numerais, com um número maior de sintagmas não redundantes para os numerais terminados em –S - os “falsos plurais” -, sugerindo que existe uma sensibilidade à marcação de pluralidade, bem como que a semântica em si do numeral ainda não parece estar tão resolvida para as crianças – pelo menos na idade investigada (6 anos) – quanto se mostra para os adultos. Tal comportamento é compreensível uma vez que a aquisição completa dos numerais, como defendido por MARCILESE (2011), consiste em uma tarefa árdua e complexa, em virtude de vários fatores que envolvem tanto propriedades de natureza semântica, sintática, lógica, quanto a própria estrutura morfológica dos numerais, que é diferente de outros itens. Dentre os aspectos elencados/reportados por Marcilese (2011, p.170) envolvidos no processo de aquisição dos numerais, podemos citar que

os numerais não são adquiridos como palavras isoladas – tal como se dá com outros itens lexicais, e sim, como parte de uma sequência hierarquicamente organizada. Em outros termos, a aquisição dos numerais se vincula diretamente com a aquisição da sequência de contagem. Além disso, a identificação das propriedades sintáticas ou semânticas dos elementos que compõem a sequência de contagem, como também a compreensão das propriedades específicas (e não linguísticas) da sequência, também são aspectos envolvidos no processo de aquisição dos numerais.

Por outro lado, em virtude dessa sensibilidade à pluralidade, parece que, pelo menos para as crianças, inicialmente, há uma tentativa de mapear o numeral de uma forma parecida com o determinante, ou seja, tentando estabelecer uma analogia entre tais elementos. Com isso, poderíamos sugerir que a criança está tratando número no numeral como um traço formal, o que faz sentido, uma vez que número é um traço formal no português. Isso porque nesse momento de seu desenvolvimento, ela já demonstra ter percebido que vários elementos são marcados para número por meio do acréscimo de morfema –S. Talvez isso explique ocorrências como as observadas em nossos dados naturalísticos das crianças: *os quatros* e *dos setes girassóis do Van Gogh*. Em ambas as ocorrências, temos a evidência de que em algum momento, as crianças parecem interpretar o –S dos numerais como morfema de número. Por outro lado, não encontramos esse tipo de ocorrência nos dados dos adultos. Por que isso acontece? Talvez por ser um traço semanticamente interpretável, parecido com um traço formal, a criança esteja tratando da mesma maneira. Em contrapartida, os adultos já têm noção – ainda que inconscientemente – dessa distinção entre traços semânticos e formais. Contudo, faz-se necessário lembrar que testamos adultos com alto nível de escolaridade. Talvez adultos com escolaridade menor poderiam apresentar comportamento semelhante ao das crianças.

Vale destacar que o funcionamento dos traços-*phi* no numeral é outra questão bastante complexa. No PB, por exemplo, os numerais possuem poucas marcas de gênero e não possuem marca de número. No entanto, naturalmente ligados à noção de contagem, possuem uma semântica ligada à noção de número, o que faz com que autores, como apontado em Ionin e Matushansky (2018), tratem o numeral como uma categoria mais lexical, como, por exemplo, nome, adjetivo ou até mesmo verbo, ou ainda autores como Lima (1996), que defendem a possibilidade de o numeral pertencer a uma categoria gramatical à parte.

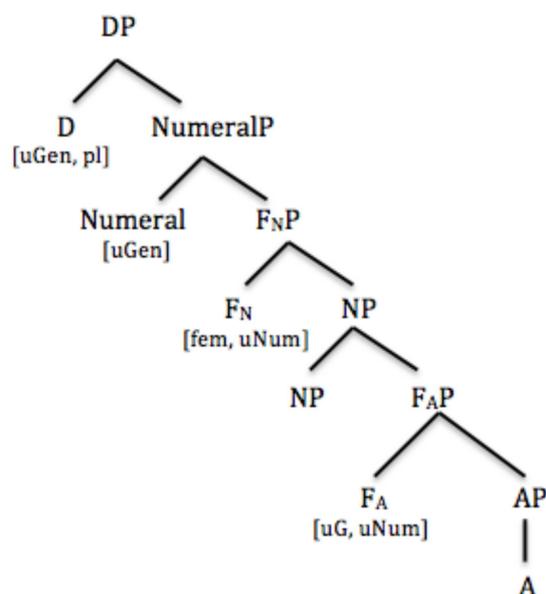
No caso específico de nossa pesquisa, entendemos que a informação de

número propriamente veiculada no numeral seja uma questão lexical, de conhecimento de mundo e não fruto de uma marcação semântica via traços formais, tal como acontece com o D.

Sendo assim, compartilhamos com Danon (2010) a hipótese de que o traço de número não está no numeral, mas numa projeção funcional acima dele na estrutura. Afinal, entendemos que a interpretação de número nos numerais é uma questão de natureza semântica e não formal. Por essa razão, como pode ser visto na figura (21) abaixo, o numeral não está envolvido na relação de *Agree* para número em nosso sistema. Com isso, assumimos a hipótese de o traço de número ser valorado em D, mesmo para as estruturas em que D não é fonologicamente realizado na interface. Para esses casos, defendemos a existência de um D nulo. Com isso, não só ressaltamos a relevância do papel de D para essas relações de *Agree* dentro do DP, como também propomos que sejam as mesmas para estruturas contendo ou não o numeral.

Figura 21- Estrutura de DP com Numeral proposta na presente pesquisa

Ex. Duas bonecas vermelhas

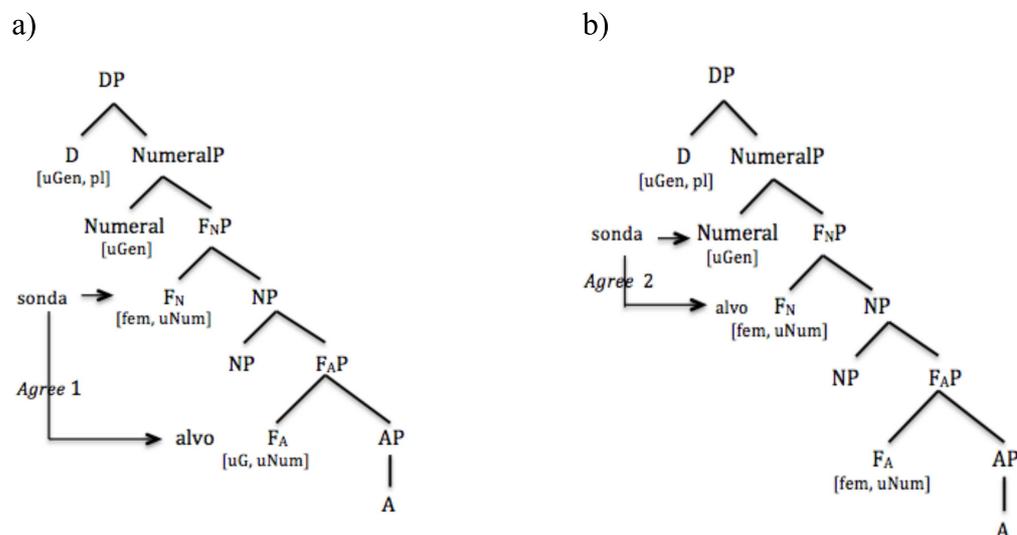


Fonte: Elaborado pela autora (2021)

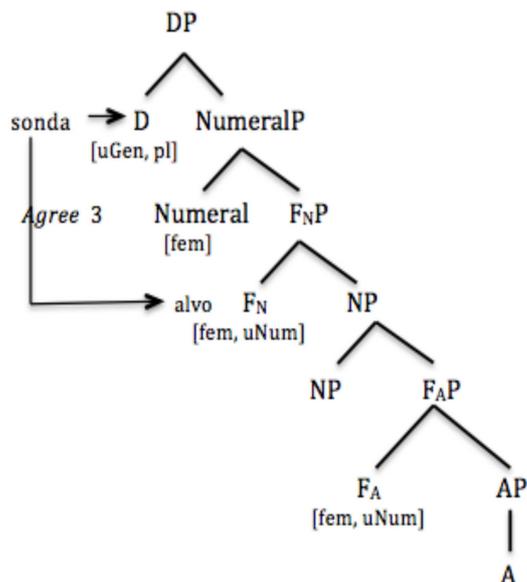
No que diz respeito à derivação de uma estrutura com Numeral, de maneira semelhante à estrutura sem esse elemento, assumimos que F_A entra na derivação com o traço de número e de gênero não valorados. F_N , por sua vez, entra na derivação valorado para gênero, mas não valorado para número, como mostra a figura (22a). Por essa razão, F_N , precisa sondar um elemento abaixo na estrutura para valorá-lo e encontra F_A . Como resultado da operação de *Agree* que se estabelece entre F_N e F_A , dá-se a valoração do traço de gênero de F_A e ambos (F_A e F_N) passam a compartilhar o traço de número ainda não valorado, figura (22b). A seguir, o Numeral entra na derivação com um traço e gênero não valorado e, por essa razão, sonda F_N para valorá-lo. (c.f figura 22b). Dessa segunda operação de *Agree*, O Numeral tem seu traço de gênero valorado pelo F_N . No entanto, F_N continua ativo na estrutura, visto que ainda possui o traço de número não valorado, como visto na figura (22c). Na mesma figura, observamos que D entra na derivação valorado pra número, mas não valorado para gênero. E é esse traço que faz que com D sonde um elemento abaixo para valorá-lo. Nessa sondagem, D encontra F_N , uma vez que o Numeral não está mais ativo na estrutura não apresentar outros traços-phi não valorados. Sendo assim, dessa terceira operação de *Agree* – entre D e F_N - obtem-se a valoração do traço de número compartilhado entre F_N e F_A , bem como do traço de gênero de D (figura 22d).

Figura 22 - Explicando a operação de *Agree* no nosso sistema nas estruturas com Numeral

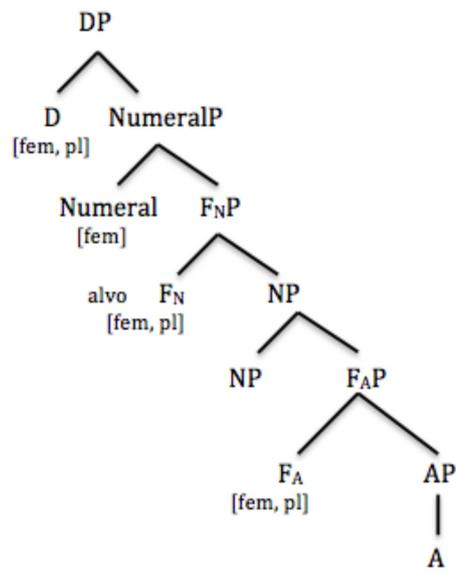
Ex. *Duas bonecas vermelhas*



c)



d)



Fonte: Elaborado pela autora (2021)

Como visto na seção 3.3.3, ao estabelecermos uma comparação entre os tipos sintagmas, ou seja, contrastando D fonologicamente realizado X Numeral, observamos nos dados dos adultos – pelo menos nas estruturas com dois elementos –, que a produção de sintagmas plurais não redundantes foi significativamente maior quando o sintagma tinha D fonologicamente realizado. Nesse sentido, podemos dizer que nossos resultados naturalísticos fornecem evidências empíricas que sustentam a hipótese de que a informação de número de D, sendo um traço formal, seja mais confiável do que a informação de número no numeral, o que permite que a marcação explícita de plural somente em D veicule a leitura plural de todo o sintagma em uma quantidade maior de dados.

Em resumo, a categoria numeral é bastante complexa, e requer uma discussão mais aprofundada do que propomos aqui, no entanto deixamos essa questão em aberto uma vez que extrapola o escopo da presente tese. Tendo-se em vista as considerações tecidas até o momento, passemos para a interpretação da variação morfofonológica observadas nos nossos dados.

5.3.2.1 Interpretando a variação na marcação morfofonológica observada nos nossos dados

Nesta seção, interessa-nos, em particular, abrir as discussões para a interface de PF, buscando perspectivas para compreender as possibilidades de não realização fonológica do traço de número nos plurais não redundantes. O nosso ponto de partida para a hipótese que desenvolvemos é o seguinte:

- (i) A valoração é uma questão relevante no componente sintático, mas não necessariamente nas interfaces;
- (ii) A interpretabilidade é uma questão relevante não no componente sintático, mas, nas interfaces.

Em termos gerais, propomos, então, que o traço de plural não valorado, ou melhor, que ganha seu valor no curso da derivação pode ter interpretabilidade variável em PF. Quando for interpretável em PF, teremos, conseqüentemente, a realização do plural redundante. Em contrapartida, quando o traço não for interpretável em PF, teremos a omissão da marca explícita de plural em, pelo menos, um dos elementos que compõem o sintagma.

Para discutir esse ponto, retomamos no quadro (11) os padrões internos de marcação do sintagma encontrados nas estruturas formadas por três elementos morfológicamente realizados nos dados experimentais:

Quadro 11-Padrões de marcação internos ao sintagma encontrados nas estruturas formadas por três elementos nos dados experimentais

Padrão	Configuração	Exemplo
NumSS	Marcação redundante	Três bonés verdes
NumSØ	Marcação não redundante: N marcado	Cinco jacaréØ marromØ
NumØØ	Marcação totalmente não redundante	Quatro caracolØ verdeØ
NumØS	Marcação não redundante: Adj marcado	Cinco gambáØ azuis

Fonte: Elaborado pela autora (2021)

Considerando os referidos dados e agora sob a perspectiva dos modelos teóricos aqui assumidos, entendemos que:

- (i) no caso de uma produção com marcação redundante, como em “Três bonés verdes”, houve o compartilhamento de traços na sintaxe e em seguida, PF efetuou a realização morfofonológica de tais traços na interface;
- (ii) apesar da presença do numeral – cuja semântica garante sua interpretabilidade em termos de número nas interfaces – as ocorrências de marcação não redundante com apenas o N marcado para plural poderiam ocorrer por causa da operação de *Agree* direta que se estabelece entre D e F_N em função da relação de localidade entre tais núcleos na estrutura – ainda que o D, nessa perspectiva, seja nulo. Além disso, hipotetizamos que o elemento que está diretamente envolvido nas relações de *Agree* – no caso F_N – apresenta uma tendência maior a ter a marcação morfológica de plural quando comparado aos outros elementos que são valorados por cadeia – no caso, o F_A . Sendo assim, faz sentido que nesta projeção se encontre o maior percentual de omissão da marca explícita de plural.
- (iii) os casos de marcação totalmente não redundante também fazem sentido no nosso sistema porque assumimos que o numeral seja lexicalmente codificado para a informação de número em virtude de ocupar uma projeção funcional que é lida na interface lógica, ou seja, a própria semântica do numeral garante essa interpretabilidade de número;
- (iv) as ocorrências com marcação não redundante, com marcação de plural apenas no Adj, apresentaram um percentual tão baixo que, sob o ponto de vista de processamento, poderiam ser interpretadas como lapsos de fala e não como realizações licenciadas pela gramática da língua.

Importante lembrar que a proposta de compartilhamento de traços pode ser vista como um desenvolvimento do sistema de *Agree* de Chomsky (2000, 2001), uma vez que, no compartilhamento, as restrições impostas pela operação de *Agree* - apresentadas anteriormente - são respeitadas, com a única exceção para o fato de que

após a valoração, os traços não interpretáveis não são deletados, como na proposta de Chomsky (2000, 2001). Isso porque a interpretabilidade é uma questão relevante não no componente sintático, mas, nas interfaces.

A partir daí, como explicar o fenômeno da concordância nominal de número variável observado nos nossos dados? Hipotetizamos que a interface de PF licencia que os traços de número não interpretáveis sejam ignorados, a despeito da valoração, uma vez que a valoração é relevante para a sintaxe propriamente dita. Em outros termos, poderíamos dizer traços que tenham recebido valoração via *Agree* durante a derivação sintática sejam invisíveis em PF que, conseqüentemente, não irá realizá-los fonologicamente. Para tal, de alguma maneira, PF deve distinguir a projeção que vem do léxico já valorada pra número – no caso, D – daquelas que tiveram seu traço de número valorado via *Agree* – tais como, o F_N e o F_A . Nesse sentido, pode-se dizer que para a concordância de número no âmbito do DP, seria obrigatório marcar o plural na projeção que veio do léxico com o traço de número valorado – o D –, podendo-se omitir a marcação explícita nas projeções em que o traço de número é valorado durante a derivação, como acontece com tal traço no F_N e no F_A . Dessa forma, pode-se dizer que em *os menino bonito*∅, o traço de plural de *menino* e de *bonito* estão sintaticamente presentes, uma vez que o *Agree* ocorre na sintaxe. No entanto, a não realização fonológica do traço de número nesses elementos acontece porque, em virtude da valoração de tal traço ter ocorrido via *Agree*, PF não é obrigada a pronunciá-lo.

Em outros termos, entendemos que *Agree* não está atrelado à realização morfofonológica. Afinal, na sintaxe, há compartilhamento de traços, os elementos concordam em si, ficando para o ramo de PF a tarefa de decidir se vai pronunciar aquele traço ou não, ou seja, parece que tal decisão não tenha uma relação direta com a questão do compartilhamento de traços na derivação. Nessa perspectiva, pensamos que *Agree* seja um processo sintático e não pós-sintático. Caso contrário, esperaríamos encontrar em nossos dados realizações de plural muito sistemáticas, isto é, não só em D ou no numeral, mas também em todas as projeções funcionais dentro do sintagma – e nossos dados apresentam evidências empíricas que sugerem que a produção de sintagmas redundantes não é categórica no PB – reforçando a visão mais difundida na literatura sociolinguística. Além disso, os resultados relativos às dimensões da SF não revelaram nenhum efeito claro dessa variável na marcação. Em

outros termos, não observamos uma atuação consistente de nenhuma das dimensões da saliência investigadas. A depender da dimensão, inclusive, nossos dados experimentais vão na contramão daquilo tradicionalmente proposto para essa variável na literatura sociolinguística e também das nossas previsões iniciais. Interessante lembrar que, em estudos anteriores, a SF se mostrou relevante na produção de falantes de classe média e alta, mas não para falantes de classe socioeconômica mais baixa. No entanto, nossos resultados também não permitem sustentar a atuação da variável entre os falantes de classe média investigados. Diante dos resultados obtidos, possíveis padrões que poderiam fornecer informações sobre a interface entre morfologia e fonologia, não foram encontrados. Assim sendo, pensamos que a explanação levantada até aqui parece mais plausível diante dos dados analisados e das ferramentas teóricas disponíveis. A partir do que foi exposto até aqui, no quadro (12) a seguir, sintetizamos as interpretações teóricas tecidas para nossos dados empíricos.

Quadro 11 - Quadro síntese com as interpretações teóricas para nossos resultados

Autor	Assunção teórica assumida
Chomsky (1995)	- Concepção de número como traço formal
Frampton e Guttman (2000)	- Mecanismo de concordância envolve compartilhamento de traços; - <i>Agree</i> é cega para o valor dos traços;
Lopes (2001)	- Concordância nominal variável no PB estaria relacionada à Teoria dos 4M. Nesse contexto, Lopes (2001) propõe que as marcas de plural nos elementos flexionáveis do sintagma nominal podem ser de dois tipos: morfema sistêmico precoce e morfema sistêmico tardio
Pesetsky e Torrego (2007)	- dissociação entre as noções de valoração e interpretabilidade: valoração relevante para a sintaxe; interpretabilidade associada às interfaces;
Baker (2008)	- Adjetivo atributivo ocupa posição de adjunto na estrutura, contrariando (Magalhães (2004) - Mesmo mecanismo de concordância pro âmbito do DP e pro nível oracional, contrariando Chomsky (1999) - Traços envolvidos nas relações de <i>Agree</i> seriam traços de projeções funcionais; - Valoração dos traços é motivada, não arbitrária como no sistema de Chomsky (2000)

Danon (2010)	-Nenhum elemento no sintagma nominal entra na derivação com o conjunto de traços-phi completo.
Azalim (2021)	- Os morfemas sistêmicos precoces, tal como proposto no modelo dos 4M, são valorados para número no sistema de <i>Agree</i> e, conseqüentemente, têm sua pronúncia garantida na interface; - Morfema sistêmicos tardios, nos moldes do modelo 4M, seriam não valorados para número no sistema de <i>Agree</i> . Tais elementos podem, portanto, ser omitidos na interface fonológica.

Fonte: Elaborado pela autora (2021)

Dito em termos de nossas suposições acerca da Teoria dos *4M*, PF seria “obrigada” a pronunciar morfemas do tipo sistêmico precoce, como é o caso do morfema –S de plural em D. Entretanto, no que diz respeito aos morfemas do tipo sistêmico tardio, aqui representado pelo morfema –S de plural no F_N e F_A , fica a critério da visibilidade desses traços em PF se tais elementos serão ou não realizados.

Contudo, nossos dados, bem como nosso sistema formal – por limitação do próprio quadro teórico –, não conseguem prever exatamente quando o falante vai pronunciar ou não a marcação explícita de plural. A nossa intuição para explicar fenômenos de natureza morfofonológica, tais como o da concordância de número variável no PB, é a de que tais fenômenos poderiam ter implicações pragmáticas, tais como estar inserido em um discurso pouco monitorado, ou num ambiente menos formal, viabilizando à PF “ignorar” os traços com mais facilidade ou, considerando o raciocínio inverso, diminuindo o nível de monitoramento necessário para a realização redundante de traços. Para integrar esse tipo de visão em um modelo formal, seria necessário mapear algum tipo de acesso ao contexto discursivo influenciando a inserção de fonologia no ramo de PF, o que, no entanto, envolveria questões que extrapolam o escopo desse trabalho.

É interessante ressaltar, no entanto, que, apesar de não ser possível mapear formalmente quando haverá inserção no morfema de número ou não, é possível, pelo menos, garantir não somente que pode haver a realização fonológica em todas as projeções funcionais ou não, como também que, no caso de ocorrência de marcação não redundante, tal como ocorre em *os meninoØbonitoØ*, não se trata de um fenômeno aleatório. Ao contrário, para tal padrão, existem algumas motivações, tal como explicitado ao longo do capítulo.

6 CONSIDERAÇÕES FINAIS

Esta tese investigou a variação observada na realização da concordância nominal no PB e buscou propor uma formalização para o fenômeno a partir de dados de produção de adultos e de crianças. Para tal, foi estabelecido um diálogo entre a sociolinguística variacionista, a psicolinguística experimental – em particular, no que tange à aquisição da linguagem – e teoria linguística de base formalista.

Nesse sentido, a tese tem como objetivo geral contribuir para a discussão mais ampla sobre o tema investigado a partir do estudo de uma série de questões relacionadas à alternância de padrões de marcação observada no PB. Sendo assim, as questões iniciais que nortearam nossa investigação e os resultados obtidos das análises dos dados naturalísticos foram:

- (a) Em que medida os padrões de concordância nominal variável no PB (redundante e não redundante) descritos no âmbito da sociolinguística variacionista, a partir de dados de fala de adultos, são atestados na produção de crianças adquirindo o PB?

Para essa questão, assumimos como hipótese de trabalho a ideia defendida em estudos, tais como os conduzidos por Capellari e Zilles (2002) e Lamprecht (1997) de que o *input* recebido pelas crianças afeta diretamente a sua produção linguística. Os dados de produção espontânea coletados e analisados no âmbito da tese se mostram compatíveis com tal hipótese, uma vez que o comportamento das crianças parece refletir de maneira fiel o padrão registrado na fala adulta, com uma produção significativamente maior de sintagmas plurais redundantes para ambos os grupos. Os dados experimentais, por sua vez, apontam diferenças importantes nos padrões de marcação de plural quando comparadas as produções de adultos e crianças. Considerados em conjunto, os dados aqui reportados indicam que a variação na marcação de plural no âmbito nominal ocorre na fala adulta e infantil em todos os grupos sociais investigados (delimitados em função de nível de escolaridade e origem geográfica) e que aspectos interacionais parecem ser relevantes para entender e explicar a maior ou menor prevalência das variantes redundante e não redundante.

- (b) Em que medida a produção espontânea infantil se mostra compatível com o *input* recebido a partir da fala de seus cuidadores primários, quando consideradas amostras de diferentes grupos socioeconômicos?

Diversos estudos identificam o nível socioeconômico dos falantes como um

fator extralinguístico condicionador da marcação variável de número no PB (ANDRADE, 2003; BRANDÃO, 2013; JAKUBOW, 2018; MARTINS, 2010; OUSHIRO, 2015; SCHERRE, 1978, 1988, dentre outros). No que se refere, especificamente ao nosso estudo, a análise dos dados naturalísticos revelou que o nível socioeconômico, aqui investigado em função do nível de escolaridade das mães das crianças que compõem os *corpora*, parece influenciar a alternância dos padrões de concordância na fala das crianças. Tal afirmação justifica-se uma vez que, embora todos os grupos analisados tenham apresentado variações na marcação de plural, a concordância redundante ocorreu de forma mais sistemática entre as crianças residentes da zona urbana, cujas mães possuíam nível de escolaridade superior.

- (c) A variável SF pode ser considerada como um preditor da alternância dos padrões de concordância nominal na produção espontânea adulta e infantil no conjunto de dados analisados na presente tese?

Sobre essa questão, tomamos como ponto de partida a visão mais difundida na literatura da sociolinguística, que defende a relevância do princípio da SF (cf. seção 2.1) como um fator importante para explicar a alternância de padrões de concordância. No entanto, nos dados naturalísticos, observou-se que, embora os percentuais gerais de ocorrência parecem ir na direção do previsto pelo princípio da SF, a análise estatística não revelou diferença significativa entre os nomes em função dessa variável. Além disso, os dados de fala espontânea não forneceram uma amostra suficientemente diversificada no que diz respeito aos diferentes níveis e dimensões da SF associada aos nomes, motivo pelo qual não foi possível traçar um quadro mais detalhado sobre a atuação dessa variável. Sendo assim, o que se pode concluir é que a SF, a depender de como é conceitualizada, não parece ser a variável mais importante para explicar a alternância de padrões de concordância nominal, o que não significa dizer que não tenha seu papel.

- (d) Há evidências de interação entre os fatores SF dos itens nominais e origem do falante (zona urbana x zona rural) na alternância de formas de concordância nominal nos *corpora* analisados?

Para essa questão, tomamos como ponto de partida resultados obtidos em pesquisas prévias (BRAGA, 1977; SCHERRE, 1978), os quais sugerem que os efeitos de saliência sofreriam influência de variáveis de natureza social, como nível de escolaridade e origem do falante. Contudo, os dados naturalísticos aqui analisados não se mostram suficientes para trazer novas evidências nesse sentido. Vale destacar, no

entanto, que no caso da presente pesquisa os dados que permitem contrastar falantes da zona urbana e rural correspondem unicamente à produção infantil, já que não contamos com uma amostra da fala de adultos residentes na zona rural.

Estudos desenvolvidos a partir do método naturalístico permitem traçar um quadro geral do fenômeno sob investigação e podem fornecer um ponto de partida sólido para diversos tipos de pesquisa. No entanto, o método naturalístico também apresenta restrições importantes. Dentre as principais limitações do método enfrentadas na nossa própria pesquisa, destacamos a dificuldade em se atingirem resultados mais precisos acerca do fenômeno sob investigação, como pudemos perceber nos dados, principalmente – mas não apenas – no que se refere à distinta produção dos tipos de nomes em função das dimensões da SF. No intuito de sanar essas e outras limitações, partimos para a condução de dois experimentos de produção eliciada cujo objetivo geral foi explorar as três dimensões da SF delimitadas pela literatura sociolinguística. No Experimento 1, exploramos as duas dimensões menos investigadas até o momento, quais sejam, *padrão de acentuação e número de sílabas* dos nomes. Já no Experimento 2, buscamos ampliar resultados previamente obtidos no que tange à dimensão *processos morfológicos de formação de plural* (AZALIM, 2016; AZALIM et al, 2018).

No que tange ao estudo experimental, nossas perguntas iniciais de pesquisa e seus respectivos resultados foram os seguintes:

- (e) A produção de sintagmas nominais redundantes e não redundantes se mostra semelhante para adultos e crianças, tal como observado nos dados naturalísticos?

A análise dos dados experimentais revelou uma maior variação na produção infantil em relação aos dados naturalísticos quando comparada com a fala dos adultos. Nas duas tarefas experimentais conduzidas, independentemente das dimensões da SF investigadas em cada uma, as crianças produziram um número semelhante de sintagmas redundantes e não redundantes, diferentemente dos adultos que mantiveram a tendência pelo padrão redundante. Como já apontado, uma possível interpretação para essas diferenças estaria associada a questões interacionais. Nesse sentido, quando comparados os dados naturalísticos e experimentais encontramos interações do tipo “adulto-criança”, no primeiro caso, e interações do tipo “criança-criança”, no segundo.

Assim sendo, os dados infantis apresentam evidências condizentes com a

hipótese de que questões relativas ao tipo de interação, tais como o paralelismo discursivo – quando o interlocutor utiliza determinada variante –, bem como o contexto de interação ou ainda o *status* do interlocutor, mostram-se relevantes para a produção infantil na faixa etária investigada. No entanto, acreditamos que tais questões ainda precisam ser mais bem investigadas como desdobramentos futuros da presente pesquisa.

Assumindo como pressuposto inicial a efetiva atuação do princípio da SF, tal como defendido por boa parte da literatura sociolinguística – embora não corroborado de forma consistente pelos nossos dados naturalísticos – quanto às dimensões da saliência investigadas, as questões que nortearam nossa pesquisa experimental, e seus respectivos resultados foram:

- (f) Quando considerados os itens do plural regular, em relação ao *padrão de acentuação* – pelo fato de a sílaba acentuada coincidir com a marcação morfofonológica de plural – itens oxítonos favoreceriam a aplicação da forma redundante quando comparados com itens paroxítonos?
- (g) Novamente considerando os nomes do plural regular, no que diz respeito à dimensão *número de sílabas*, a presença de um volume maior de material fônico nos nomes trissílabos poderia favorecer a marcação redundante em comparação com nomes dissílabos?
- (h) Quanto aos *processos de formação de plural*, aqueles que geram os contrastes mais nítidos entre as formas singular e plural de um dado item nominal favorecem a marcação redundante quando comparados ao plural regular em -S?

Como reportado no Capítulo 4, embora os resultados experimentais tenham apontado alguns efeitos significativos, os mesmos não fornecem evidências robustas compatíveis com a atuação de cada variável investigada de forma isolada. Em alguns casos, por exemplo, como observado em função do *padrão de acentuação*, os resultados ora não revelaram efeitos estatisticamente significativos, ora apontaram na direção oposta à prevista.

Considerando nossos próprios dados, aliados a resultados conflitantes reportados na literatura acerca da possível relevância das dimensões da saliência quando consideradas isoladamente (cf. CARVALHO, 1997; GUY, 1981; SCHERRE, 1988), levantamos a seguinte questão:

- (i) O número de sílabas aliado ao padrão de acentuação poderiam ser fatores que atuariam de forma conjunta outorgando um maior ou menor destaque perceptual a um dado item?

Observamos, no entanto, que os resultados do Experimento 1 também não se mostraram consistentes para sustentar a hipótese da existência de um *continuum* ou gradiente de saliência (i.e. considerando a ocorrência de sintagmas redundantes e não redundantes em cada condição experimental).

Em resumo, tomados em conjunto, os resultados obtidos nos Experimentos 1 e 2 dialogam com o reportado pelos nossos estudos prévios (AZALIM, 2016 e AZALIM et al, 2018), bem como com resultados obtidos por meio de análise de *corpora* desenvolvidos por Campos (2015) os quais não se mostram compatíveis com a hipótese de que a SF tenha uma influência clara na alternância dos padrões de concordância nominal de número no PB – pelo menos ao se considerar a variável nos moldes mais difundidos no âmbito da sociolinguística.

Dessa forma, de um lado, temos a visão bastante defendida na literatura, a saber, a relevância da SF como um dos principais condicionadores da alternância de padrões de concordância. Por outro lado, vemos os resultados conflitantes da SF que a própria literatura reporta – como os que discutimos ao longo da tese –, com trabalhos indicando algum efeito de SF – ainda que de uma maneira não muito clara (RIBEIRO, 2013; ROZA; CASAGRANDE, 2015) –, outros que apontam efeitos mais nítidos – embora não explorem todas as dimensões (OUSHIRO, 2016; OUSHIRO; GUY, 2015)–, e ainda aqueles que questionam o possível efeito de tal fator de modo mais geral (NICOLAU, 1984; REIS, 2020; SCHER, 2021).

Diante dos resultados obtidos e do quadro geral que emerge a partir da análise da literatura, surge como questão importante, o que seria exatamente a SF. Uma possível resposta envolve a hipótese de que a saliência consiste, na verdade, em um acúmulo, uma sobreposição de informações e não propriamente em dimensões isoladas. Nesse sentido, apenas certos nomes, com combinações específicas de propriedades seriam de fato “salientes” e, conseqüentemente, fariam alguma diferença para a marcação. Com isso, uma questão que permanece em aberto para futuras pesquisas é em que medida a SF é uma propriedade que depende de uma combinação de fatores.

Apesar de a SF ter sido tomada como ponto de partida para nossas investigações empíricas no âmbito desta tese, os dados coletados se revelaram uma

fonte de informações bastante rica sobre outros aspectos para além dessa variável em particular. Assim sendo, foram levantadas outras questões que parecem ser importantes para compreender a alternância de padrões de concordância nominal, tais como a configuração dos sintagmas, os padrões de marcação internos ao sintagma e a presença de numerais. Nesse contexto, as seguintes perguntas foram levantadas:

- (j) A configuração da estrutura na qual os nomes ocorrem teria algum papel na alternância na marcação da concordância?

Quanto à configuração em termos de número de elementos morfológicamente realizados, os dados sugerem que o número de elementos que compõem o sintagma seja um fator relevante para a ocorrência de variação na marcação de número nos sintagmas nominais. Pelo menos no que tange a sintagmas encabeçados por numeral, os resultados são compatíveis com a ideia de que estruturas mais curtas (formadas por dois elementos morfofonologicamente realizados, tais como Num+N) parecem favorecer o padrão redundante em comparação a estruturas de três (Num+N+Adj) ou quatro (Num+N+Adj+Adj) elementos⁷⁴.

- (k) A constituição morfofonológica dos numerais teria algum efeito na realização dos padrões de concordância nominal?

No que tange à produção infantil na faixa etária investigada, os resultados sugerem uma diferença em relação ao padrão adulto. Isso porque quando o numeral apresenta uma “falsa analogia” de plural, as crianças parecem interpretar o –S como um morfema de plural, omitindo a marcação nome com mais frequência, de forma semelhante como ocorre no PB no caso dos determinantes (Ex. *Os papagaio*∅– *Dois papagaio*∅). Já nos numerais em que tal analogia não é possível (*quatro*, *sete*), a marcação redundante ocorre de forma mais consistente no nome. Em outros termos, as crianças parecem se guiar mais pela sistematicidade da morfologia – que pode ser observada, por exemplo, nos determinantes – do que pela semântica de pluralidade intrinsecamente associada aos numerais.

- (l) A partir dos resultados empíricos reportados, qual seria a nossa proposta de análise formal para a concordância nominal variável no PB?

⁷⁴ Vale lembrar que ao compararmos sintagmas em função da configuração, estamos considerando o número de elementos morfofonologicamente realizados.

Em termos teóricos, nossos resultados são compatíveis com uma interpretação que parte da premissa da Teoria dos 4M (MYERS-SCOTTON; JAKE, 2000a) para explicar o fenômeno da concordância variável de número no PB. Nessa perspectiva, propomos que a concordância não redundante é licenciada pelo tipo de morfema envolvido, sua constituição de traços formais, bem como pelas relações que eles estabelecem no sistema de *Agree* (FRAMPTON; GUTMANN, 2000; PESETSKY; TORREGO, 2007). Mais especificamente, propomos que o morfema de número em D seja do tipo sistêmico precoce e o morfema de número em N e em Adj seja do tipo sistêmico tardio. Nos moldes do sistema de *Agree*, propomos que os primeiros sejam valorados e interpretáveis e que os últimos correspondam a traços não valorados e não interpretáveis. A partir desse raciocínio, com base no sistema de Frampton e Guttman (2000), Pesetsky e Torrego (2007), propomos que o mecanismo de concordância envolve, na verdade, compartilhamento de traços. Isso possibilita, não só que se estabeleça uma relação de *Agree* entre uma sonda e um alvo igualmente não valorados para um determinado traço, como também uma dissociação entre as noções de valoração e interpretabilidade, uma vez que, segundo a perspectiva assumida em Pesetsky e Torrego (2007), para o sistema computacional basta que ocorra a valoração dos traços que vieram do léxico não especificados. Já a interpretabilidade seria determinada fora do sistema computacional e, como tal, teria a ver com relações de condições de interface. E é nesse contexto das interfaces, que propomos que a variação de padrões de concordância nominal de número estaria atrelada ao critério de visibilidade desse traço em PF. Sendo assim, morfemas do tipo sistêmico precoce – morfema –S de plural em D – por entrarem na derivação com o valor já especificado no léxico, são visíveis para PF e que, garante sua realização morfológica. Já os morfema sistêmico tardio – aqui representado pelo morfema -S de plural no F_N e F_A – são valorados ao longo da derivação o que, por sua vez, pode torná-los invisíveis para PF e, se assim for, não serão morfológicamente realizados.

Vale destacar que os questionamentos levantados ao longo desta tese não se esgotam aqui. No entanto, acreditamos que esse trabalho possa fornecer novos subsídios para o estudo da concordância nominal variável, bem como apresentar algumas das vantagens da articulação entre as abordagens naturalística, experimental e a teoria formalista para uma melhor compreensão do fenômeno em questão. Esperamos assim ter contribuído para alcançar, pelo menos de forma parcial, os

objetivos levantados por Nelson (2014, p.250) ao afirmar que “we need to explain the causes of diversity as much as we need explanation for regularities.”

7 REFERÊNCIAS BIBLIOGRÁFICAS

- ABNEY, S. **The English Noun Phrase in its Sentential Aspect**. 1987. Tese (Doutorado em Linguística) – MIT, Indiana University, 1981.
- ACUÑA-FARIÑA, J. C. The linguistics and psycholinguistics of agreement: A tutorial overview. **Science Direct**. *Lingua* 119, p.389-424, 2009.
- ADGER, D. **Core Syntax: A Minimalist Approach**. Oxford: Oxford University Press, 2002.
- ADGER, D; SVENONIOUS, P. 2012. Aldridge, Edith. 2008. Generative approaches to ergativity. **Language and Linguistic Compass: Syntax and Morphology**2, 966–995.
- ALMEIDA, E. M. **A variação da concordância nominal num dialeto rural**. 1997. Dissertação (Mestrado em Letras) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1997.
- AMBRIDGE, B.; ROWLAND, C. F. Experimental methods in studying child language acquisition. **WIREsCognSci**2013.doi: 10.1002/wcs.1215, 2013.
- ANDRADE, L. M. **Rupturas e contínuos da Concordância Nominal de número em textos orais de Informantes de Tubarão (SC) e São Borja (RS)**. 2003. Dissertação (Mestrado em Ciências da Linguagem) - Faculdade de Letras, Universidade do Sul de Santa Catarina USSC, Tubarão, 2003.
- AUGUSTO, M.; Corrêa, L.S.; Ferrari Neto, J. Explorando o DP: a presença da categoria NumP. **Revista de Estudos da Linguagem**, Belo Horizonte, v. 14, n. 2, p. 245-275, 2006.
- AZALIM, C. **Variação e processamento linguístico: um estudo experimental sobre a concordância nominal variável no PB**. Dissertação (Mestrado em Linguística) – Faculdade de Letras, Universidade Federal de Juiz de Fora, Juiz de Fora, 2016.
- AZALIM, C., et al. Concordância nominal variável de número e saliência fônica: um estudo experimental. **DELTA**. Documentação de Estudos em Linguística Teórica e Aplicada, v. 34, p. 513-545, 2018.
- BAKER, M.C. **The Syntax Of Agreement and Concord**. Cambridge: Cambridge University Press.2008.
- BECHARA, E. **Moderna gramática portuguesa** – 38ª ed. rev. e ampl., Rio de Janeiro: Nova Fronteira, 2015.
- BLUM-KULKA, S; SNOW,C. The potential of peer talk. **Discourse Studies: Thematic Issue: Peer Talk and Pragmatic Development**. London, Thousand Oaks, CA and New Delhi , Vol 6(3), p.291-306.2004.

BRAGA, M. L. **A concordância de número no sintagma nominal no Triângulo Mineiro**. 1977. Dissertação. (Mestrado em Língua Portuguesa) – Departamento de Letras. PUC-Rio, Rio de Janeiro, 1977.

BRAGA, Maria Luiza. & SCHERRE, Maria Marta Pereira. A concordância de número no SN na área urbana do Rio de Janeiro. In: ENCONTRO NACIONAL DE LINGÜÍSTICA, 1o, 1976. Anais [...] Rio de Janeiro, PUC, 1976. p.464-77.

BRANDÃO, S. F.; VIEIRA, S. R. (Org). **Ensino de gramática-descrição e uso**. 2.ed., São Paulo: Contexto, 2013.

BRANDÃO, S. F. Patterns of plural agreement within the Noun Phrase. **Journal of Portuguese Linguistics**, v. 12, n. 2, p. 51–100, 2013.

BRANDÃO, S. F.; VIEIRA, S. R. Concordância nominal e verbal: contribuições para o debate sobre o estatuto da variação em três variedades urbanas do português. **Alfa**, São Paulo, v. 53, n. 3, p. 1035-1064, 2012.

BRITO, A.M.; LOPES, R.E.V. The Structure of DPs. In: W.LeoWetzels; MENUZZI, S; COSTA, J. (Org.) **The Handbook of Portuguese Linguistics**. 1a. ed. Oxford: Blackwell, 2016, v., p. 254-274.

CAMPBELL-KIBLER, K. The nature of sociolinguistic perception. **Language Variation and Change**, Cambridge University Press. v. 21, p. 135-156, 2009.

CAMPBELL-KIBLER, K. (2010) New Directions in Sociolinguistic Cognition, **University of Pennsylvania Working Papers in Linguistics**. Vol. 15 :Iss. 2 , Article 5. Available at: <https://repository.upenn.edu/pwpl/vol15/iss2/5>.

CAMPOS, O.G.L.A. de S. Concordância nominal. In: RODRIGUES, A.; ALVES, I. M. (Org.). **A construção morfológica da palavra. Gramática do português culto falado no Brasil**. São Paulo: Editora Contexto, 2015.

CAPELLARI, E.T.; ZILLES, A. M. S. A marcação de plural na linguagem infantil - estudo longitudinal. **Revista da ABRALIN**, v. 01, p. 185-218, 2002

CARSTENS, V. Concord in minimalist theory. **Linguistic Inquiry**, Cambridge, v. 31, n. 2, The MIT Press. p.319-355, 2000.

CARVALHO, R. C. **A concordância de número no sintagma nominal na fala urbana de Rio Branco**. 1997. Dissertação (Mestrado em Lingüística) - Instituto de Estudos da Linguagem – Universidade Estadual de, Campinas, 1997.

CASTRO, A.; FERRARI NETO, J. Um estudo contrastivo do PE e do PB com relação à identificação de informação de número no DP. **Letras de Hoje**, Porto Alegre, v. 42, n.1, p.65-76, 2007.

CEKAITE, A.; BLUM-KULKA, S.; GRØVER, V; TEUBAL, E. Children’s peer talk and learning: uniting discursive, social, and cultural facets of peer interactions: editor’s introduction. In: Ed. Cekaite, A.; Blum-Kulka, S.; Grøver, V; Teubal, E. **Children’s Peer Talk: Learning from Each Other**. Nova Iorque: Cambridge University Press, 2014. p.237-250.

CHAVES, R.G. Princípio de saliência fônica: isso não soa bem. **Letrônica**, Porto Alegre, v. 7, n. 2, p. 522-550, jul./dez., 2014.

CHAVES, R. G. 2017. A redução/desnasalização de ditongos nasais átonos e a marcação explícita da CVP6: um estudo de correlação. Tese de Doutorado, Universidade Federal de Santa Catarina, Florianópolis. Inédita.

CHOMSKY, N. Lectures on Government and Binding: The Pisa Lectures. Holanda: Foris Publications, 1981.

CHOMSKY, N. A Minimalist program for linguistic theory, in: HALE, K. e S. J. KEYSER, S. J. (orgs.), **The view from building 20, essays in honor of Sylvain Bromberger**. Cambridge, MIT Press, pp. 1-52. 1993.

CHOMSKY, N. **The Minimalist Program**. Cambridge, Mass: The MIT Press, 1995.

CHOMSKY, N. Derivation by phase. **MIT Occasional Papers in Linguistics**, 18, 1999.

CHOMSKY, N. Minimalist inquiries: The framework. In ROGER M. et al. (eds.) **Step by step: Essays on minimalist syntax in honor of Howard Lasnik**, 89-155. Cambridge: MIT Press. 2000.

CHOMSKY, N. Beyond explanatory adequacy. In: **The cartography of syntactic structures.**, Structures and beyond, ed. by Belletti, A. Vol. 3. Oxford: Oxford University Press, 2004.

CINQUE, G. 1994. On the evidence for partial N-movement in the romance DP. In **Paths Towards Universal Grammar**, Cinque G.; Koster, J. Pollock, J.-Y. ; Rizzi, J.-Y.; Zanuttini, R.(Eds), 85–110. Washington DC: Georgetown University Press, 1994.

CORBETT, G. **Agreement**. C.U.P., Cambridge, UK. 2006

CORRÊA, L. M. S.; AUGUSTO, M. R. A.; FERRARI NETO, J. The Early Processing of Number Agreement in the DP: Evidence from the Acquisition of Brazilian Portuguese. In: **30 th Annual Boston University Conference on Language Development (BUCLD 30)**, 2006, Boston. BUCLD 30. Somerville, Mass: Cascadilla Press, 2005. v. 2.

COSTA, J.; FIGUEREIDO SILVA, M. C. Nominal and verbal agreement in Portuguese: an argument for Distributed Morphology. In: COSTA, J.; FIGUEIREDO SILVA, M. C.(Eds.) **Studies on Agreement**. Amsterdam/Philadelphia: John Benjamins, 2006.

COSTA, C. S. S. M ; TEIXEIRA, V.P. Um estudo sociolinguístico da concordância nominal de número em uma comunidade de Caxias - Ma. **Revista do GELNE** , Natal, v. 19, p. 18, 2017.

DANON, G. Agreement and DP-Internal Feature Distribution. **Syntax**, v.14, n.4, p.297-317, 2011.

DIAS, J.F. V.; FERNANDES, M.A. A inter-relação da concordância nominal e da concordância nos predicativos/ participios passivos, sob enfoque da teoria da variação e mudança linguística. **Organon . Porto Alegre**, v.14 (28-29). p. 115-131, 2000.

EINSENBEISS, S. Contrast is the name of the game: Contrast-based semi-structured elicitation techniques for studies on children's language acquisition. **Essex Research Reports in Linguistics. Colchester**, v.57, n.7, p.1-27, 2009.

EINSENBEISS, S. Production Methods In: **Language acquisition research**. Experimental Methods in Language Research. Blom.E.; Unsworth, S. (eds.). Amsterdam: John Benjamin Publishing, p.11-33, 2010.

ELY, R.; GLEASON, B. Socialization across contexts. Em P. Fletcher; B.Macwhinney (Orgs.), **The handbook of child language**, p. 251-270. Oxford, U.K.: Blackwell. 1996.

EMBICK, D.; NOYER, R. Movement operations after syntax. **Linguistic Inquiry**, v. 32, n. 4, p. 555-595, 2001.

FERNANDES, M. **Concordância nominal na região sul**.1996. Dissertação (Mestrado em Letras/Linguística) – Curso de Pós-Graduação em Letras/Linguística. Universidade Federal de Santa Catarina, Florianópolis, 1996.

FERRARI-NETO, J. **Reconhecimento do número gramatical e processamento da concordância de número no sintagma determinante na aquisição do português brasileiro**. 2003. Dissertação (Mestrado em Estudos da Linguagem) - Departamento de Letras PUC-Rio, Rio de Janeiro, 2003.

FERRARI-NETO, J. **Aquisição de Número Gramatical no Português Brasileiro: Processamento de Informação de Interface e Concordância**. 2008. Tese (Doutorado em Letras – Estudos da Linguagem) - Departamento de Letras. Pontifícia Universidade Católica- Rio, Rio de Janeiro, 2008.

FIELD, J. **Psycholinguistics**. The Key concepts. Oxon: Routledge, 2004.

FORTES, F. da S. Sintaxe greco-romana: Prisciano de Cesareia e Apolônio Díscolo na história do pensamento gramatical antigo. 2012. Tese (Doutorado em Linguística – Estudos Clássicos) – Instituto de Estudos da Linguagem. Universidade Estadual de Campinas, Campinas. 2012.

FRAMPTON, J. & S. GUTMANN. (2000) **Agreement is feature sharing**. Northeastern University, ms.

FREITAG, R. M. K.; LIMA, G. O. S. Sociolinguística. 2010. (Desenvolvimento de material didático ou instrucional - Ensino a distância (Graduação em Letras/Português)).

GOMES, C.A.; BENAYON, A.R.; VIEIRA, M.C.P. Variação e aquisição da flexão nominal e da flexão verbal. **Gragoatá: Niterói**, v.30, n.1, pp. 39-54. 2006.

GRIMSHAW, J. Extended Projections. Ms. Brandeis University. 1991.

GRØVER AUKRUST, V. Explanatory discourse in young second language learners' peer play. **Discourse Studies: Thematic Issue: Peer Talk and Pragmatic Development**. London, Thousand Oaks, CA and New Delhi , v.6, n.3, p.393-412.

GUY, G.R. **Linguistic Variation in Brazilian portuguese: Aspects of the Phonology, Syntax, and Language History**. 1981. Tese (Doutorado em Linguística) - University of Pennsylvania, Philadelphia, 1981.

HALLE, M.; A. MARANTZ.. Distributed Morphology and the Pieces of Inflection In: K. Hale and S. J. Keiser (eds.) **The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger**, MIT Press, Cambridge, MA p. 111-176,1993.

HAUSER, M.; CHOMSKY, N.; FITCH, W. T. The Faculty of language: what is it, who has it, and how did it evolve? **Science**, v. 298, p.1569-1579, 2002.

HENRIQUE, K. S. **Variação linguística e processamento: investigando o papel da distância entre sujeito e verbo na realização da concordância verbal variável no PB**. 2016 -Dissertação (Mestrado em Linguística) – Faculdade de Letras, Universidade Federal de Juiz de Fora, Juiz de Fora, 2016.

HORNSTEIN, N., NUNES J. e GROHMANN, K.K. **Understanding Minimalism** Cambridge Textbooks in Linguistics. Cambridge: Cambridge University Press, 2005.

IONIN , T.; MATUSHANSKY. **Cardinals: The Syntax and Semantics of Cardinal-Containing Expressions**. London: MIT Press, 2018.

JAKE, J., MYERS-SCOTTON, C., & GROSS, S. A minimalist approach to codeswitching work: Adding the Matrix Language. **Bilingualism, Language and Cognition**,v.5, p.61–91, 2002.

JAKE, J., MYERS-SCOTTON, C. Which language? Participation potentials across lexical categories in codeswitching. In. L. Isurin, D. Winford, & K. de Bot (Eds.), **Multidisciplinary approaches to code switching**. Amsterdam, the Netherlands / Philadelphia, PA: Benjamins Publishing, p.207–242, 2009.

JAKUBÓW, A.P.S.P. **Language acquisition based on variable input: the case of number agreement in Brazilian Portuguese**. Tese (Doutorado em Estudos da Linguagem). Departamento de Letras, Pontifícia Universidade Católica do Rio de Janeiro, Rido de Janeiro, 2018.

KAJORNBOON, A. B. Using interviews as research instruments. Language Institute, Chulalongkorn University, Bangkok, Thailand. Disponível em:<http://hsmi.psu.ac.th/upload/forum/Annabelinterviewguide.pdf>

KENEDY, E. **Curso básico de linguística gerativa**. São Paulo: Contexto, 2013.

LABOV, W. **Padrões Sociolingüísticos**. Trad. Marcos Bagno; Marta Scherre; Caroline Cardoso. São Paulo: Parábola, 2008 [1972].

LABOV, W.; ASH, S.; RAVINDRANATH, M.; WELDON, T.; BARANOWSKI, M.; NAGY, N. Properties of the sociolinguistic monitor. **Journal of Sociolinguistics**, v.15, p. 431-463, 2011.

LAMPRECHT, R. R. Aquisição da morfologia do plural por crianças bilíngues Português - Alemão. In: I ENCONTRO DO CELSUL, v.1, 1997, Florianópolis, Anais [...] Florianópolis: Universidade Federal de Santa Catarina, 1997. p.107-117.

LEMLE, Miriam. & NARO, Anthony Julius. *Competências básicas do português*. Relatório final de pesquisa apresentado às instituições patrocinadoras Fundação Movimento Brasileiro (MOBRAL) e Fundação Ford, Rio, 1977. 151p.

LIEVEN, E. V. M. Crosslinguistic and crosscultural aspects of language addressed to children. In C. Gallaway & B. J. Richards (Eds.), **Input and interaction in language acquisition**. Cambridge University Press. P. 56-73. 1994.

LIGHTFOOT, D. *The development of language: acquisition, change and evolution*. Massachusetts: Blackwell Publishers, 1999.

LIGHTFOOT, D.; WESTERGAARD, M. Language acquisition and language change: Interrelationships. **Language and Linguistics Compass** 1(5). p. 396–416. 2007.

LIMA, M.C. Numeral: Uma classe à parte? **Rev. de Letras** - vol. 18 - n.º. 2 - jul/dez 1996.

LIPSKI, J. The many faces of Spanish /s/-weakening: (re) alignment and ambisyllabicity. In: Gutierrez-Rexach, J.; Martí nez-Gil, F. (Eds.), **Advances in Hispanic linguistics**. Somerville, MA : Cascadilla Press, , pp. 198–213. 1999.

LONGOBARDI, G. Reference and proper names: a theory of movement in syntax and LF. **Linguistic Inquiry**, v. 25, n.4, The MIT Press, p. 609-665, 1994.

LOPES, N. da S. Concordância nominal, contexto lingüístico e sociedade. 2001. Tese (Doutorado em Letras)- Instituto de Letras, Universidade Federal da Bahia, Salvador, 2001. Inédito.

LOPES, N. da S. Aquisição da concordância no português: uma explicação com base na teoria dos 4M. **Papia**, Brasília, v. 15, p. 72-79, 2005.

LOPES, N. da S. O mecanismo da variação da concordância no português: observações quanto a marcas nos verbos e nos nomes. **Estudos da Língua(gem)** (Online) , v. 13, p. 59-72, 2015.

LOPES, R. E. V. . Estágios no processo de aquisição de número no DP do português brasileiro. **Letras de Hoje** , Porto Alegre, v. 39, n.3, p. 157-171, 2004.

LOPES, L. O. J.; SCHERRE, M. M. P. Influência da saliência fônica no processo de concordância nominal no português falado na zona rural de Santa Leopoldina-ES. In: **II Conel- Congresso Nacional de Estudos Linguísticos**, 2014, Vitória. Anais [...]. Vitória: 2014, p. 67-81.

- LORANDI, A. Aquisição da variação: a interface entre aquisição da linguagem e variação linguística. **Alfa**, São Paulo, 57 (1). p.133-162. 2013.
- LUCCHESI, D. Um balanço crítico das análises da variação na concordância nominal de número no português brasileiro. In: Apresentação no Congresso Internacional da ABRALIN, 2019, Maceió, 2019.
- MAGALHÃES, T.M.V. A valoração de traços de concordância dentro do DP. **DELTA**, vol.20, n.1, p.149-170, 2004.
- MARCILESE, M. **Sobre o papel da língua no desenvolvimento de habilidades cognitivas superiores**: representação, recursividade e cognição numérica. 2011. Tese (Doutorado em Estudos da Linguagem) – Faculdade de Letras, Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), Rio de Janeiro, 2011.
- MARCILESE, M et al. Processamento da concordância variável no PB em uma perspectiva experimental. **Revista Linguística** / Revista do Programa de Pós-Graduação em Linguística da Universidade Federal do Rio de Janeiro. v. 11, n. 1, p. 118-134, junho de 2015.
- MARCILESE, M et al. Efeitos de distância linear e marcação no processamento da concordância verbal no PB? Linear distance and markedness effects in variable subject-verb agreement processing in BP. **Revista de Estudos da Linguagem**, v.25, p.1291-1325. 2017.
- MARTELOTTA, M.E. Conceitos de gramática. In: MARTELOTTA, M.E. (org) **Manual de Linguística**. 2ª ed.- São Paulo: Contexto, 2011.
- MARTINS, F. S. Uma abordagem sociolinguística da concordância nominal de número no falar dos habitantes do município amazonense de Benjamin Constant. **Working papers em linguística (online)**, Florianópolis, v. 11, p. 45-56, 2010.
- MAYER, M. **Frog, Where are You?** New York NY: Dial Book for Young Readers, 1969.
- MENDES, R. B.; OUSHIRO, L. Variable Number Agreement in Brazilian Portuguese: An Overview. **Language and Linguistics Compass**. John Wiley & Sons Ltd. 9/9, p.358–368, 2015.
- MENUZZI, S. (1994). Adjectival positions inside DP. In C. Cremers and R. Bok-Benema (eds.), **Linguistics in the Netherlands**, Amsterdam: John Benjamins, vol. 11, p. 127–138.
- MILLER, K.; SCHMITT, C. Effects of variable input in the acquisition of plural in two dialects of Spanish. **Lingua** v.120, n.5. 1178–1193, 2010.
- MILLER, K.; SCHMITT, C. Variable Input and the Acquisition of Plural Morphology. **Language Acquisition**, v.19, n.3, 223-261, 2012.
- MOLINA, D. S. L. **Aquisição da linguagem e variação linguística: um estudo sobre a flexão verbal variável na aquisição do PB**. 2018. Tese (Doutorado em

Linguística) – Faculdade de Letras. Universidade Federal de Juiz de Fora, Juiz de Fora. 2018

MOLLICA, M.C. Fundamentação teórica: conceituação e delimitação in: MOLLICA, M. C. e BRAGA, M.L. (orgs.). **Introdução à sociolinguística**: o tratamento da variação. São Paulo: Contexto, 2003.

MOUTINHO, A.P.D.A; PEREIRA, M.G.D. O code-switching na perspectiva da intercompreensão: interações em chat plurilíngue no projeto Galanet. VEREDAS Online, Juiz de Fora, v.19, n.1, p. 176-197, 2015.

MYERS-SCOTTON, C.; JAKE, J. L. Four types of morpheme: evidence from aphasia, code switching, and second-language acquisition. In KLEIN, Wolfgang et alii (ed.). **Linguistics**: an interdisciplinary journal of the language sciences. v.38, n.6,p. 1053-1100, 2000a.

MYERS-SCOTTON, C. e JAKE, J. L. Testing the 4-M model: an introduction. TRAVIS, C. **Linguistics**. International journal of bilingualism. v.4, p. 1-8, 2000b.

MYERS-SCOTTON, C. e JAKE, J. L. Revisiting the 4-M model: Codeswitching and morpheme election at the abstract level. **International Journal of Bilingualism** 2017, v. 21, n.3, p. 340–366. 2017

NAMBA, K. An overview of Myers-Scotton’s Matrix Language Frame Model. **Senri International School (SIS) Educational Research Bulletin**,v.9p.1-10, 2004.

NAME, C.; SOSA, J. M. . Cadê o amor da mamãe? As interrogativas na Fala Dirigida à Criança adquirindo o PB. **VEREDAS - REVISTA DE ESTUDOS LINGÜÍSTICOS**, Juiz de Fora, v. 24, p. 72-93, 2020.

NARO, A. J. The social and structural dimensions of a syntactic change. **Language**, v. 57, p.63-98, 1981.

NARO, A. J.; GÖRSKI; FERNANDES, E. Change without change. **Language Variation and Change**, v.11, p. 197-211, 1999.

NARO, A. J.; SCHERRE, M. M P. (Org.). **Origens do Português Brasileiro**. São Paulo: Parábola, 2007.

NELSON, K. (2014) What, when, and how do children learn from talking with peers? In: Ed. Cekaite, A.; Blum-Kulka, S.; Grøver, V; Teubal, E. **Children’s Peer Talk: Learning from Each Other**. Nova Iorque: Cambridge University Press, 2014. p. 237-250.

NICOLAU, E. M. das D. **A Ausência de Concordância Verbal em Português: uma abordagem sociolinguística**.1984. Dissertação (Mestrado em Linguística) – Faculdade de Letras , Universidade Federal de Minas Gerais, Belo Horizonte, 1984.

NICOLAU, E. M. das D. A influência da constituição morfológica da forma verbal na ausência de concordância em Português. **Revista de Estudos da Linguagem**. Belo Horizonte: UFMG, ano IV, ,v. 2, n. 3, p. 41-67, 1995.

OUSHIRO, L. **Identidade na pluralidade: avaliação, produção e percepção**

linguística na cidade de São Paulo. 2015. Tese (Doutorado em Letras) –Faculdade de Filosofia, Letras e Ciências Humanas. Universidade de São Paulo, São Paulo, 2015.

OUSHIRO, L. Social and structural constraints in lectal cohesion. **Lingua** (Haarlem. Print) v. 172-3, p. 116-130, 2016.

OUSHIRO, L. **Introdução à Estatística para Linguistas**,v.1.0.1 , dez/2017. Disponível em. Licença Creative Commons 4.0 Atribuição – Não comercial. Acesso em 17/04/2019.

OUSHIRO, L. ; GUY, G. R. The Effect of Salience on Co-variation in Brazilian Portuguese, **University of Pennsylvania Working Papers in Linguistics**: v.21 :Iss. 2 , Article 18. 2015. Disponível em: <https://repository.upenn.edu/pwpl/vol21/iss2/18> Acesso em: 21/10/2021.

PAGOTTO, E. G. Norma e condescendência; ciência e pureza. **Línguas e instrumentos linguísticos**, Campinas, n. 2, p. 49-68, jul./dez. 1998.

PESETSKY, D; TORREGO, E. The syntax of valuation and interpretability of features. In: KARIMI, S. et al. (Org.). **Phrasal and clausal architecture: Syntactic derivation and interpretation**. Amsterdam: John Benjamins, p. 262-294, 2007.

POLINSKY, M. Agreement in Archi from a Minimalist Perspective. In: O. Bond, G. Corbett, M. Chumakina & D. Brown, eds., **Archi. Complexities of Agreement in Cross Theoretical Perspective**. Oxford: Oxford University Press, 2016. pp. 184–232.

REIS, M. M. **Aquisição da variação e mudanças na(s) gramática(s) das crianças : um olhar sobre a Concordância Variável no PB**. 2020. Dissertação (Mestrado em Linguística) – Instituto de Estudos da Linguagem - Universidade Estadual de Campinas, Campinas, 2020.

RIBEIRO, P.R.O. **O perfil sociolinguístico de Oliveira Fortes- MG: A concordância nominal e verbal**.2013. Dissertação (Mestrado em Linguística). Faculdade de Letras , Universidade Federal de Juiz de Fora, Juiz de Fora, 2013.

RITTER, E. 1991. Two functional categories in noun phrases: Evidence from modern Hebrew. In **Syntax and Semantics 25: Perspectives on phrase structure**, ed. Susan Rothstein. New York: Academic Press, p. 37-62. 1991.

ROBERTS, J. Child language variation. In: CHAMBERS, J.; SCHILLING-ESTES, N.; TRUDGILL, P. (Ed.). **Handbook of language variation & change**. Oxford: Blackwell, 2002.

ROCHA LIMA, C. H. **Gramática normativa da língua portuguesa**. 53^a ed. Rio de Janeiro: José Olympio, 2017.

RODRIGUES, E. dos S. **Processamento da Concordância de número entre sujeito e verbo na produção de sentenças**. 2006. Tese (Doutorado em Letras) - Departamento de Letras. PUC-Rio, Rio de Janeiro, 2006.

ROZA, A.; CASAGRANDE, S. A aquisição da concordância de número no sintagma nominal: dados de produção e imitação eliciada. Disponível em: www.rd.uffs.edu.br/handle/prefix/295. 2015

SALOMÃO, M. H. **A variação de pluralidade nas estruturas predicativas da variedade falada na região de São José do Rio Preto**. Dissertação (Mestrado em Estudos Linguísticos) – Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista, São José do Rio Preto, 2010.

SCHER, L.S. **A saliência fônica e o processamento da concordância verbal variável no PB**. Dissertação (Mestrado em Linguística) – Instituto de Estudos da Linguagem – Universidade Estadual de Campinas, Campinas, 2021.

SCHERRE, M. M. P. **A regra de concordância de número no sintagma nominal em português**. 1978. Dissertação (Mestrado em Língua Portuguesa) - Departamento de Letras, Pontifícia Universidade Católica- Rio, Rio de Janeiro, 1978.

SCHERRE, M. M. P. **Reanálise da concordância nominal em português**. 1988. Tese (Doutorado em Linguística) – Faculdade de Letras, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 1988.

SCHERRE, M.M.P. A concordância de número nos predicativos e participios passivos. **Organon – A variação no português do Brasil**, Instituto de Letras, UFRGS – Porto Alegre, v.18, n,5, p.52-70, 1991.

SCHERRE, M. M. P. Aspectos da concordância de número no português do Brasil. **Revista Internacional de Língua Portuguesa (RILP)** - Norma e Variação do Português. Associação das Universidades de Língua Portuguesa. 12, p.37-49. dez. de 1994.

SCHERRE, M. M. P. **Doa-se lindos filhotes de poodle** – Variação linguística, mídia e preconceito. São Paulo: Parábola, 2005.

SILVA, J. B. **Uma reflexão sobre a concordância nominal na fala capixaba e suas contribuições para o ensino de língua portuguesa**. 2011. Trabalho de Conclusão de Curso. Universidade Federal do Espírito Santo, Vitória, 2011.

SILVA, C.R.T e MAGALHAES, T.M.V. Sobre a natureza do morfema de pluralidade em português: discutindo o domínio da concordância. In: **Veredas** (UFJF. Online), v.18, p.188-210, 2014.

SIMIONI, L. Aquisição da concordância nominal de número: um estudo de caso. **Revista de Estudos da Linguagem**, Belo Horizonte, v. 14, n. 2, p. 539- 570, 2006.

SIMIONI, L. **A aquisição da concordância nominal de número no português brasileiro: um parâmetro para a concordância nominal**. 2007. Dissertação (Mestrado em Linguística) - Universidade Federal de Santa Catarina, Florianópolis, 2007.

SLOBIN, D.I.; WELSH, C.A. Elicited imitation as a research tool in developmental psycholinguistics. In: FERGUSON, G.A.; SLOBIN, D.I. (eds.) **Studies of child language development**. New York: Holt, Rinehart and Winston, p.485-497. 1973.

SQUIRES, L. It don't go both ways: Limited bidirectionality in sociolinguistic perception. **Journal of Sociolinguistics**, 17/2, 2013, p. 200–237.

SQUIRES, L. 2014. Social differences in the processing of grammatical variation. In: **Penn Working Papers in Linguistics**. 20(2) (Selected Papers from NWAV42). Disponível em: <http://repository.upenn.edu/pwpl/vol20/iss2/20>

SZABOLCSI, A. The possessor that ran away from home. **The Linguistic Review** 3, p.89 – 102. 1984.

TABOSA, M.V. **A variação na concordância nominal de número no falar do cariri cearense**. 2016. Dissertação (Mestrado em Linguística). Universidade Federal do Ceará, Fortaleza. 2016.

TARALLO, F. **A pesquisa sociolinguística**. 7. ed. São Paulo: Ática, 2002.

The functional categories in the Noun Phrase. Manuscrito anônimo disponível em: https://shodhganga.inflibnet.ac.in/bitstream/10603/174873/7/07_chapter%201.pdf

WAKASA, M.M. The Matrix Language Frame Model of Codeswitching as applied to Kiswahili – English Codeswitching in Kenya. 2004. Thesis (Master of Arts in Linguistics) – Department of Linguistics and African Languages, Faculty of Arts, University of Nairobi, 2004.

Wei, L. Types of morphemes and their implications for second language morpheme acquisition. *International Journal of Bilingualism*, 4(1), 29-43. 2000. Disponível em: <https://doi.org/10.1177/13670069000040010301>

WEINREICH, U.; LABOV, W.; HERZOG, M.. Empirical foundations for a theory in language change. In: LEHMANN, W. P.; MAKIEL, Y. (Eds). **Directions for Historical Linguistics**. Austin: University of Texas Press, 1968.

8 APÊNDICES

APÊNDICE A

Tabela 29 - Distribuição dos sintagmas plurais em função do padrão de concordância aplicado (redundante X não redundante) e faixa etária em cada corpus– dados naturalísticos

Crianças	Idade	Padrão de concordância	
		No total de ocorrências e %	
<i>Corpus</i>		Redundante	Não Redundante
<i>Corpus A</i> Crianças zona urbana	3 anos	50 (78%)	14 (22%)
	4 anos	45 (98%)	01 (2%)
	5 anos	66 (85%)	12 (15%)
	6 anos	115 (94%)	07 (6%)
<i>Corpus C</i> Crianças zona urbana	3 anos	09 (75%)	03 (25%)
	5 anos	113 (84%)	22 (16%)

Fonte: Elaborado pela autora (2021)

APÊNDICE B

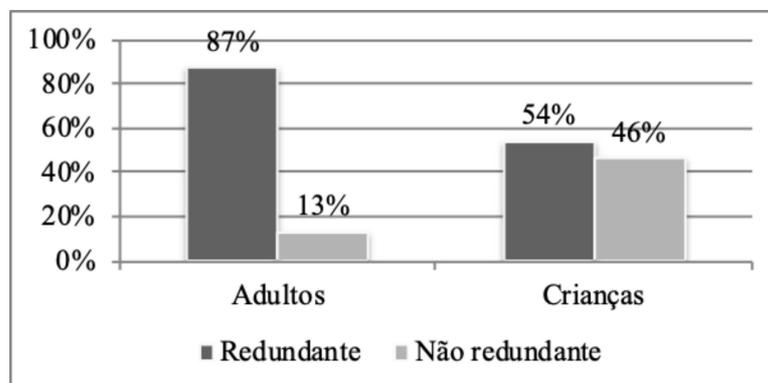
Tabela 30 - Distribuição dos sintagmas plurais em função do padrão de concordância aplicado (redundante X não redundante) e SF em cada faixa etária– dados naturalísticos

Grupo	Idade	Saliência Fônica	Padrões de concordância		Total de ocorrências
			Redundante	Não redundante	
<i>Corpus A</i> zona urbana	3 anos	+sal	12 (92%)	01 (8%)	13
		-sal	38 (75%)	13 (25%)	51
<i>Corpus A</i> zona urbana	4 anos	+sal	04 (100%)	0 (0%)	04
		-sal	41 (98%)	01 (2%)	42
	5 anos	+sal	09 (100%)	0 (0%)	09
		-sal	57 (83%)	12 (17%)	69
6 anos	+sal	15 (100%)	0 (0%)	15	
	-sal	100 (93%)	07 (7%)	107	
<i>Corpus C</i> zona urbana	3 anos	+sal	0 (0%)	01 (100%)	01
		-sal	09 (82%)	02 (18%)	11
	5 anos	+sal	21 (91%)	02 (9%)	23
		-sal	92 (82%)	20 (18%)	112

Fonte: Elaborado pela autora (2021)

APÊNDICE C

Gráfico 12 - Percentual de sintagmas plurais produzidos por crianças e adultos na primeira tarefa experimental – Experimento 1 - dados brutos



Fonte: Elaborado pela autora (2021)

APÊNDICE D

Tabela 31 - Tabela de coeficientes extraída do modelo de regressão logística para ocorrências de sintagmas plurais não redundantes em função do grupo analisado (Crianças X Adultos) – Experimento 1 – dados brutos

	Estimate	Std. Error	Z	p
Crianças (Intercept)	-0.17073	0.09901	-1.724	0.08
Adultos	-1.68617	0.17899	-9.420	<0.01

Fonte: elaborado pela autora (2021)

APÊNDICE E

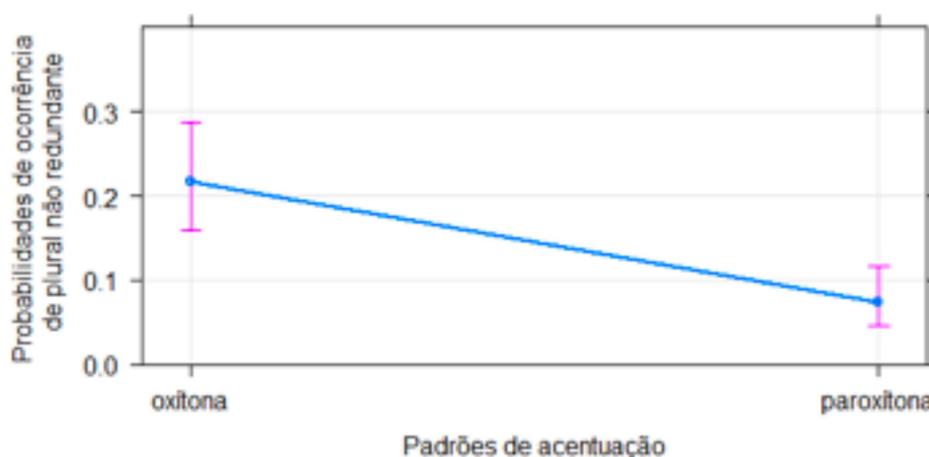
Tabela 32 – Distribuição dos sintagmas plurais em função do padrão de acentuação (oxítonos X paroxítonos) em ambos os grupos (Crianças e adultos) – Experimento 1 – dados brutos

Grupo	Padrão de acentuação	Padrões de concordância		Regressão logística	Total de ocorrências
		Concordância redundante	Concordância não redundante		
Crianças	oxítonos	102(54%)	86(46%)	$p=1$	188
	paroxítonos	121 (54%)	102(46%)		223
Adultos (controle)	oxítonos	122(78%)	35(22%)	$p< 0.01$	157
	paroxítonos	211 (93%)	17(7%)		228

Fonte: Elaborado pela autora (2021)

APÊNDICE F

Gráfico 13 - Gráfico de efeitos da variável padrão de acentuação na produção de sintagmas plurais não redundantes pelos Adultos (Grupo controle) – Experimento 1 – dados brutos



Fonte: Elaborado pela autora (2021)

APÊNDICE G

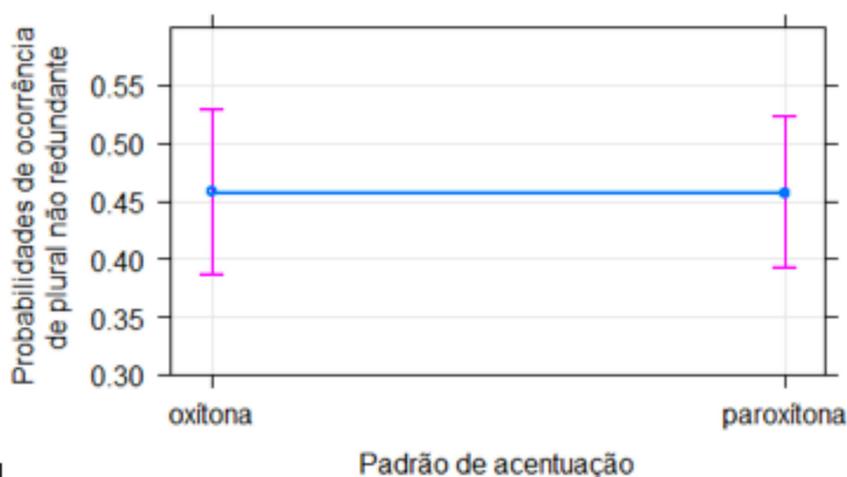
Tabela 33 - Distribuição dos sintagmas plurais em função do número de sílabas (dissílabos X trissílabos) no grupo experimental (crianças) e controle (adultos) – Experimento 1 – dados brutos

Grupo	Número de sílabas	Padrões de concordância		Regressão logística	Total de ocorrências
		Concordância redundante	Concordância não redundante		
Crianças	Nomes dissílabos	111 (52%)	102 (48%)	$p < 0.3$	213
	Nomes trissílabos	112 (57%)	86 (43%)		198
Adultos (controle)	Nomes dissílabos	181 (88%)	25 (12%)	$p < 0.4$	206
	Nomes trissílabos	152 (85%)	27 (15%)		179

Fonte: Elaborado pela autora (2021)

APÊNDICE H

Gráfico 14 - Gráfico de efeitos da produção de sintagmas plurais não redundantes em função da variável padrão de acentuação (Grupo de crianças) – Experimento 1 – nomes alvo



Fonte: Elaborado pela autora (2021)

APÊNDICE I

Tabela 34 - Tabela de coeficientes extraída do modelo de regressão logística para ocorrências de sintagmas plurais não redundantes em função das variáveis padrão de acentuação e número de sílabas (Modelo sem interação) – Crianças e Adultos - Experimento 1

Crianças				
	Estimate	Std. Error	z	p
Intercept (Oxítona, dissílaba)	-0.190785	0.193217	-0.987	0.323
PAD.ACE Paroxítona	0.003228	0.228664	0.014	0.989
NUM.SIL Trissílaba	0.043377	0.228154	0.190	0.849
Adultos				
	Estimate	Std. Error	z	p
Intercept (Oxítona, dissílaba)	-1.3709	0.2583	-5.308	<0.01
PAD.ACE Paroxítona	-1.2390	0.3484	-3.556	<0.01
NUM.SIL Trissílaba	0.5324	0.3307	1.610	<0.01

Fonte: Elaborado pela autora (2021)

APÊNDICE J

Tabela 35 - Tabela de coeficientes extraída do modelo de regressão logística para ocorrências de sintagmas plurais não redundantes em função das variáveis padrão de acentuação e número de sílabas (Modelo com interação) – Crianças e Adultos – Experimento

Crianças

	Estimate	Std. Error	z	p
Intercept (Oxítona, dissílaba)	-0.2666	0.2215	-1.204	0.22
PAD.ACEParoxítona	0.1613	0.3191	0.505	0.61
NUM.SIL Trissílaba	0.1943	0.3119	0.623	0.53
PAD.ACEParoxítona:NUM.SILTrissílaba	-0.3253	0.4580	-0.710	0.477

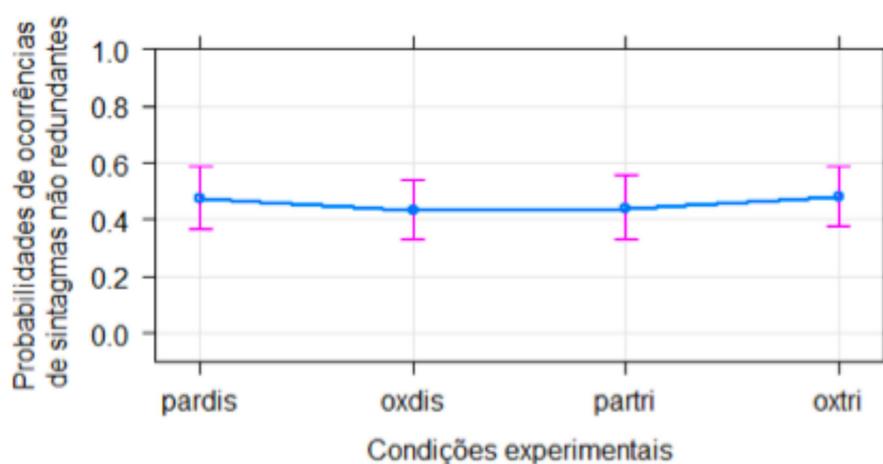
Adultos

	Estimate	Std. Error	z	p
Intercept (Oxítona, dissílaba)	-1.4351	0.2873	-4.995	<0.01
PAD.ACEParoxítona	-1.0217	0.5133	-1.990	0.04
NUM.SIL Trissílaba	0.6619	0.4046	1.636	0.10
PAD.ACEParoxítona:NUM.SILTrissílaba	-0.3884	0.6955	-0.558	0.57

Fonte: Elaborado pela autora (2021)

APÊNDICE K

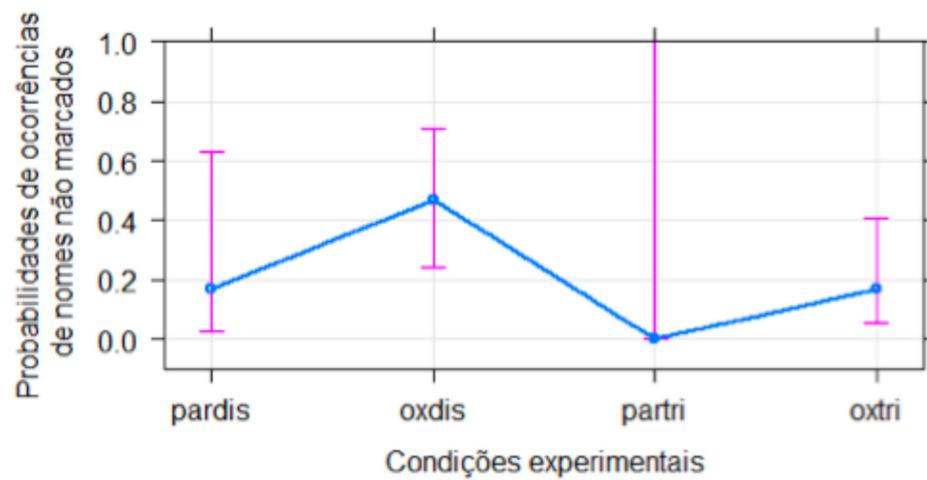
Gráfico 15 - Gráfico de efeitos da produção de sintagmas plurais não redundantes em função das condições experimentais (Grupo de crianças) – Experimento 1 – nomes alvo



Fonte: Elaborado pela autora (2021)

APÊNDICE L

Gráfico 16 - Gráfico de efeitos da produção de nomes não marcados em sintagmas plurais não redundantes em função das condições experimentais (Grupo de Adultos) – Experimento 1 – nomes alvo



Fonte: Elaborado pela autora (2021)