



Aquisição da linguagem, variação linguística e desenvolvimento sociolinguístico

Language acquisition, linguistic variation, and sociolinguistic development

Mercedes Marcilese

Universidade Federal de Juiz de Fora (UFJF), Juiz de Fora, Minas Gerais / Brasil
mercedes.marcilese@ufjf.br

<http://orcid.org/0000-0002-9058-8367>

Cristina Name

Universidade Federal de Juiz de Fora (UFJF), Juiz de Fora, Minas Gerais / Brasil
Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) / Brasil
cristina.name@ufjf.br

<http://orcid.org/0000-0001-5625-9503>

Cristina Azalim

Universidade Federal de Juiz de Fora (UFJF), Juiz de Fora, Minas Gerais / Brasil
azalimc@gmail.com

<http://orcid.org/0000-0002-9830-8948>

Daniele Molina

Universidade Federal de Juiz de Fora (UFJF), Juiz de Fora, Minas Gerais / Brasil
daniele.molina@ufjf.br

<http://orcid.org/0000-0002-8555-3039>

Resumo: Durante a aquisição da sua língua materna, por vezes, as crianças adquirem uma gramática que difere – em algum grau – daquela apresentada pelos seus cuidadores principais. Esse fato coloca em destaque a estreita relação que existe entre aquisição da linguagem, variação e mudança linguística. A simples presença de variação nos dados primários disponíveis para a criança, no entanto, nem sempre gera uma situação de

mudança linguística. Ao mesmo tempo, a atribuição de valor social às formas linguísticas constitui uma das facetas do processo de aquisição, mas a percepção da variabilidade linguística *per se* não parece ser suficiente para que tal atribuição seja realizada pelas crianças. Adotando um conjunto de pressupostos básicos no campo da psicolinguística – que caracteriza a língua no contexto mais amplo da cognição humana –, este artigo tece uma reflexão sobre a relação entre a presença de variabilidade no *input* durante a aquisição de língua materna e o desenvolvimento sociolinguístico de bebês e crianças. Buscamos ainda ilustrar a relação entre o contato com um *input* variável e a aquisição da linguagem a partir da análise de um fenômeno de variação específico, qual seja, a concordância variável – nominal e verbal – observada no português brasileiro. Analisamos a produção linguística de crianças e seus cuidadores a fim de verificar a relação entre o *input* recebido pelos jovens aprendizes e sua própria produção. O processo de escolarização – e seu efeito normativizador na língua –, a identidade do participante e as diversas dimensões da interação comunicativa, bem como aspectos vinculados às ferramentas de coleta dos dados reportados nas pesquisas, se mostram cruciais para uma compreensão aprofundada das questões em discussão.

Palavras-chave: aquisição da linguagem; *input* variável; desenvolvimento sociolinguístico; concordância variável.

Abstract: This paper discusses the relation between variable input and sociolinguistic development in mother-tongue acquisition process. During language acquisition, children can acquire a linguistic system that differs – to some degree – from that of their primary caregivers. This fact emphasizes the close relationship that exists between language acquisition, linguistic variation and linguistic change. However, the presence of variation in primary linguistic data does not guarantee linguistic change. At this point, it is worthy to say that, although the attribution of social values to linguistic forms is a particular phase of the acquisition process, the perception of linguistic variability *per se* does not seem to be sufficient for such attribution to be made by children. By adopting the psycholinguistic assumptions, this paper aims at discussing some relevant aspects of language acquisition, linguistic variation and sociolinguistic development by analyzing the phenomenon of variable nominal and verbal number agreement in Brazilian Portuguese. Naturalistic data were analyzed in order to verify the input received by children and their own linguistic production. Formal education – and its normative effects in language –, speaker identity and the different dimensions of communicative interaction, as well as methodology issues, seem to be crucial for a deeper understanding of the topics under discussion.

Keywords: language acquisition; variable input; sociolinguistic development; variable agreement.

Recebido em 01 de setembro de 2022.

Aceito em 12 de dezembro de 2022.

1 Introdução

A aquisição da língua materna¹ constitui um dos principais exemplos da capacidade de aprendizagem indutiva dos seres humanos. A partir de um conjunto finito de sentenças geradas com base no sistema a ser adquirido, bebês e crianças pequenas se mostram capazes de extrair os padrões mais gerais que subjazem a esse sistema, isto é, capazes de identificar a gramática da língua. Talvez o ponto mais surpreendente nesse processo seja o fato de que qualquer membro da espécie com desenvolvimento típico costuma ser bem-sucedido nesse aprendizado, embora os dados disponíveis no *input* não sejam – por si só – suficientes para explicar tal sucesso (CHOMSKY, 1981; SINGLETON; NEWPORT, 2004; dentre outros). Essa combinação de sucesso a partir de dados relativamente pobres ou parciais se encontra na base da visão inatista para a aquisição da linguagem. De acordo com essa perspectiva, restrições, tendências e/ou mecanismos inatos seriam em parte responsáveis pelo bom desempenho dos jovens aprendizes.

Como em todos os campos das denominadas ciências humanas, a aquisição de língua materna se configura como um campo de investigação no qual coexistem abordagens teóricas diversas e visões epistemologicamente distintas. Em particular, as teorias diferem quanto à natureza das restrições inatas que orientariam o processo de aquisição e em que medida tais restrições seriam de domínio específico ou fariam parte de mecanismos mais gerais, sendo compartilhados em outros âmbitos do aprendizado humano (e até mesmo, por outras espécies). Mas, afinal, o que significa *adquirir uma língua*? Embora não haja uma resposta única para essa questão, veremos que muitos dos possíveis caminhos não são excludentes, mas sim, complementares.

Em primeiro lugar, vale lembrar que adquirir uma língua diz respeito a um *processo*. O ponto de partida – ou estado inicial – é caracterizado de formas variadas a depender da perspectiva teórica assumida. Já o estado “final” ou estável desse processo envolveria o conhecimento pleno de uma (ou mais) língua(s) que possibilitaria

¹ As expressões “aquisição de língua materna” e “aquisição da linguagem” remetem aqui – como será detalhado a seguir – ao processo pelo qual o bebê/a criança parte de um estado inicial e atinge um estado estável de conhecimento de uma (ou mais) língua(s)/gramática(s). Os termos “língua”, “sistema linguístico” e “gramática” são utilizados como equivalentes ao longo do artigo.

que a criança – de posse dele – fosse capaz de atingir um desempenho comunicativamente satisfatório, em contextos de interação diversos, com interlocutores e objetivos variados. Assim, adquirir uma língua inclui o desenvolvimento de habilidades de expressão e comunicação, de interação social e de organização do pensamento, bem como a constituição de uma identidade e a aquisição de conhecimento, tanto linguístico quanto não linguístico. Além desses aspectos mais amplos e gerais, o processo requer de forma crucial a identificação – espontânea, natural e sem necessidade de instrução formal – das características do sistema linguístico no qual a criança se encontra imersa, durante seus primeiros anos de vida. Será com base nessa identificação que a criança poderá interpretar e produzir enunciados linguísticos.

Numa visão “alargada”, o estudo da aquisição da linguagem pode incluir as habilidades expressivas e comunicativas não necessariamente verbais (por exemplo, aspectos vinculados à gestualidade), bem como investigar o desenvolvimento de habilidades discursivas, o aprendizado da escrita e os processos de leitura e letramento (CORRÊA, 1999). Já numa perspectiva mais restrita, pesquisas sobre a aquisição visam a identificar o que os bebês e crianças são capazes de perceber e discriminar a partir dos estímulos linguísticos disponíveis no ambiente, como ocorre a análise do material linguístico e de que maneira o conhecimento é armazenado e representado ao longo do processo de aquisição. O estudo da aquisição pode contribuir ainda para a identificação de fontes e manifestações de alteração na linguagem ao longo do desenvolvimento.

Perspectivas teóricas de orientação pragmática, sócio-interacionista ou sócio-cognitivista, com frequência demonstram um interesse maior por aspectos contemplados na visão mais ampla mencionada, enquanto abordagens desenvolvidas no contexto da psicolinguística e da linguística teórica se vinculam principalmente à visão mais restrita salientada. Vemos, no entanto, que esses diferentes pontos de vista se ocupam de aspectos distintos e constitutivos de um processo que é complexo e multifacetado.

A perspectiva aqui assumida para abordar a aquisição da linguagem está alinhada com pressupostos fundamentais no campo da psicolinguística que, seja em vertentes de orientação mais empirista ou inatista, investiga os processos mentais subjacentes à aquisição e ao uso de uma (ou mais) língua(s), considerando tanto capacidades perceptuais e cognitivas disponíveis aos seres humanos, quanto características e propriedades dos

estímulos linguísticos (CLARK, 1999; CORRÊA, 2011). Assim sendo, entendemos que a aquisição envolve um processo que parece se iniciar ainda na vida intrauterina (MAMPE et al., 2009; dentre outros) e depende de habilidades perceptuais, de identificação e de categorização (GERVAIN; WERKER, 2008; WERKER; GERVAIN, 2013), por parte do bebê/criança, e que tais habilidades interagem com sistemas de desempenho de ordem mais geral – tais como memória, processos atencionais e controle executivo – além de sofrerem os efeitos da maturação do organismo de forma mais ampla. Assumimos ainda que a linguagem interage com outros domínios cognitivos, dentre os quais, interessa na discussão que desenvolvemos aqui, em particular, a cognição social, responsável pelas representações, habilidades e capacidades necessárias para o desenvolvimento de raciocínios sociais, fundamentais para o estabelecimento e a manutenção de relações em contextos de interação.

O contexto social no qual a aquisição ocorre é um aspecto pouco explorado nas pesquisas da área de aquisição da linguagem e da psicolinguística, de modo mais geral. Isso porque mesmo nas abordagens que buscam esclarecer o papel da interação e a dimensão sociocognitiva do processo de aquisição, determinados pontos acabam sendo pouco ou não investigados. Em particular, a natureza variável da língua – presente até em contextos aparentemente mais “estáveis” de monolinguismo – e os potenciais desafios que ela representa em termos de aquisição, têm um espaço ainda reduzido na literatura (MILLER; RAMOS, 2014; SMITH, DURHAM; RICHARDS, 2013; SMITH; DURHAM; FORTUNE, 2009). Se – pelo menos em parte – a tarefa da criança envolve a identificação de regularidades que permitam chegar a generalizações sobre o sistema em aquisição, como acontece esse processo quando os dados de entrada não são categóricos, mas sim, variáveis²?

A variação linguística abrange aspectos que vão desde diferenças na qualidade da voz, na velocidade da fala e na pronúncia, nas escolhas lexicais e nas realizações morfossintáticas e discursivas – relacionadas

² De acordo com uma visão inatista, o processo de aquisição da linguagem ocorre crucialmente a partir da identificação e mapeamento de regularidades. Nesse sentido, a variação (mesmo aquela sistemática) pode ser vista como um “problema” ou “desafio” para a criança em fase de aquisição. Embora todo *input* seja potencialmente variável, o modo como essa variação se apresenta nos dados primários pode apresentar diferenças substanciais em função da comunidade de fala específica considerada.

a características individuais dos falantes, tais como sexo, idade, procedência geográfica etc. – até a variação socialmente condicionada e vinculada a fatores tais como escolaridade, nível socioeconômico, âmbito de atuação profissional etc. Podemos afirmar então que a variação presente no *input* disponível para a criança é uma questão de grau: algumas crianças podem estar inseridas em contextos com um volume maior de variabilidade do que outras, mas a variação sempre estará presente em alguma medida. Até as mais “inocentes” diferenças de pronúncia que poderiam ser vistas – de forma equivocada – como mais simples em termos de variação, podem ser correlacionadas a fatores socialmente relevantes. Da mesma forma, o uso de palavras distintas para se referir a uma mesma entidade/objeto e a presença de formas variáveis em configurações gramaticalmente relevantes que codificam informações cruciais para a interpretação dos enunciados e que muitas vezes envolvem dependências não adjacentes entre os elementos (ex. relações de concordância, envolvendo informações de número, gênero e pessoa), são comuns no *input* ao qual bebês e crianças são expostos.

Embora a existência de variação linguística no *input* seja um fato que as crianças em fase de aquisição irão enfrentar, os efeitos que essa variabilidade tem no processo e o modo como a consciência social – relativa, por exemplo, à identificação e ao reconhecimento de parceiros sociais – e a consciência sociolinguística se desenvolvem nos jovens falantes, são aspectos pouco investigados na literatura sobre aquisição da linguagem. De modo geral, a maior parte das investigações sobre a aquisição de língua materna tem como foco aspectos regulares, sistemáticos e mais ou menos categóricos dos sistemas linguísticos. Na literatura mais recente, no entanto, é possível observar um aumento significativo de pesquisas que consideram tanto a aquisição de aspectos linguísticos variáveis quanto questões vinculadas ao desenvolvimento sociolinguístico³ e relativas, por exemplo, à consciência infantil sobre a identidade do falante e os valores socialmente atribuídos às formas linguísticas (JOHNSON; WHITE, 2020).

³ O desenvolvimento sociolinguístico envolveria, dentre outros aspectos, a construção de uma consciência sociolinguística. Nesse sentido, “desenvolvimento sociolinguístico” é mais amplo do que “consciência sociolinguística” e está atrelado de forma mais direta à cognição social não estritamente linguística.

O presente artigo visa a fornecer uma reflexão sobre os dois aspectos mencionados e certamente (co)relacionados, mas cujas implicações ainda não são bem compreendidas. De um lado, a variabilidade presente no *input* linguístico e seus efeitos no processo de aquisição de língua materna (cf. Seção 2); do outro, o desenvolvimento sociolinguístico da criança que ocorre como parte e consequência natural do processo de aquisição (cf. Seção 3). Para fechar essa reflexão, na quarta seção fornecemos dados relativos a um fenômeno de variação linguística bastante investigado no português brasileiro (doravante, PB), qual seja, a concordância variável de número. Em particular, buscamos caracterizar – a partir de dados naturalísticos (MOLINA, 2018; AZALIM, 2021) – o modo como a concordância nominal e verbal se manifesta no *input* primário recebido por crianças de diferentes grupos socioeconômicos adquirindo o PB e na própria produção linguística das crianças.

2 Variabilidade no *input* e aquisição da linguagem

Ao adquirir sua língua materna, por vezes, as crianças adquirem um sistema que difere daquele utilizado pelos seus cuidadores primários⁴. Esse ponto é particularmente relevante para a compreensão da mudança linguística e sinaliza a estreita relação que existe entre aquisição, variação e mudança (LIGHTFOOT, 2010; LIGHTFOOT; WESTERGAARD, 2007). O *input* recebido pelos bebês e crianças – também caracterizado em termos de uma língua externa (CHOMSKY, 1981, 1986) – pode desencadear o desenvolvimento de uma gramática ou língua interna diferente nos jovens aprendizes.

De acordo com Cornips e colaboradores (2016), as crianças teriam um papel fundamental como agentes da mudança diacrônica que ocorreria devido a um processo de transmissão não-alvo (também denominada transmissão irregular) de formas e estruturas linguísticas entre as gerações de falantes. Os autores destacam que, nesse sentido, é notável que os estudos desenvolvidos sobre aquisição da linguagem e sobre mudança linguística tenham permanecido distantes por tanto tempo. O relativo isolamento entre esses campos de estudo parece estar

⁴ Consideramos aqui como cuidadores primários aqueles adultos que atuam no cotidiano como responsáveis principais pelos cuidados e pela maior parte das tarefas envolvendo o bebê/criança.

vinculado, principalmente, a diferenças nas questões fundamentais de pesquisa e aos métodos de coleta de dados adotados.

Lightfoot (2010) destaca que, durante a aquisição da linguagem, as crianças seriam influenciadas tanto pela biologia – em particular, pelos mecanismos gerais e específicos que guiam o processo – quanto pelo ambiente. O autor se refere aqui ao “ambiente” em termos do estímulo linguístico disponível, i.e., aquilo que as crianças ouvem ou percebem visualmente (no caso das línguas de sinais). Assim, em determinados momentos, o *input* que as crianças recebem pode mudar um pouco, configurando dessa forma novos dados primários que podem promover a configuração de novas gramáticas internas. Pequenas mudanças na língua externa poderiam ser responsáveis por desencadear mudanças de maior alcance nas línguas internas. No entanto, nem sempre uma alteração no *input* irá determinar uma mudança de fato.

A relação entre o processo de aquisição da linguagem e a mudança linguística tem sido mais investigada no contexto de transmissão irregular, em ambientes de contato linguístico. Pidgins e crioulos nas suas etapas iniciais, em particular, constituem fontes bastante relevantes para a compreensão dessas relações. O poder regularizador que os jovens aprendizes têm, no entanto, não parece operar da mesma forma em todos os contextos. Assim, enquanto nos ambientes em que de fato são verificadas lacunas e/ou inconsistências no *input*, o conhecimento desenvolvido pelas crianças parece resolver ou superar os problemas originais; quando se trata de variabilidade sistemática no interior de sistemas linguísticos já consolidados, a mudança linguística não parece ocorrer com a mesma velocidade ou com a mesma extensão.

Chomsky (1981, 1986) destaca que o *input* linguístico “normal” é naturalmente pobre, já que certas propriedades fundamentais da gramática podem ficar indeterminadas nos dados disponíveis durante o processo de aquisição da linguagem. Contudo, em circunstâncias típicas, os dados disponíveis fornecem exemplos dos tipos de estrutura que podem ser gerados pelo sistema em aquisição e possuem um grau considerável de regularidade e consistência. O que acontece então nos casos em que o *input* não apresenta suficiente consistência ou sistematicidade ou não fornece evidências da complexidade da gramática? Os exemplos mais extremos desse tipo de situação relatados na literatura correspondem a crianças que receberam pouco ou nenhum *input* durante o período crítico (LENNENBERG, 1967; cf. WERKER; HENSCH, 2015 para uma revisão atualizada). Os

casos documentados envolvem crianças criadas sem contato humano – as denominadas crianças selvagens (ITARD, 1932 apud SINGLETON; NEWPORT, 2004; LANE, 1979 apud SINGLETON; NEWPORT, 2004; ZINGG, 1940 apud SINGLETON; NEWPORT, 2004) – ou crianças em situações de negligência e maus tratos graves, como o caso de Genie (CURTISS, 1977). Contextos de privação linguística como os mencionados envolvem ainda – em maior ou menor grau – situações mais gerais de privação social e, para além das deficiências observadas no desenvolvimento linguístico, incluem efeitos psicológicos mais amplos que precisam ser levados em consideração. Assim, é importante que generalizações com base em dados dessa natureza sejam sempre tomadas com cautela.

Desconsiderados esses casos extremos, com forte privação social, existem contextos em que o *input* disponível para a criança apresenta características especiais. Crianças surdas sem contato com uma comunidade linguística “típica” (no caso, surdos sinalizadores nativos), crianças em contextos migratórios ou filhos/as de falantes de pidgin são populações especialmente relevantes para a compreensão dos limites sob os quais a aquisição pode acontecer (SINGLETON; NEWPORT, 2004).

Hudson Kam e Newport (2005, 2009) investigaram o processo de aquisição quando o *input* disponível apresenta morfemas gramaticais inconsistentes ou imprevisíveis, tal e como observado em pidgins e crioulos incipientes⁵. As autoras buscaram determinar em que medida os aprendizes apenas reproduzem o *input* disponível ou introduzem alterações, tornando o sistema mais regular enquanto adquirem a língua. Por meio de experimentos utilizando uma língua artificial, as autoras concluem que as crianças entre 5 e 7 anos – mas não os adultos – se mostraram capazes de introduzir novos padrões, distintos dos observados no *input*, regularizando aspectos inicialmente inconsistentes. Diante desses resultados, as autoras defendem que as crianças teriam um papel fundamental e único na formação de línguas crioulas, em virtude de sua capacidade regularizadora dos padrões gramaticais.

⁵ Segundo as autoras: “[t]here are several important ways in which the variable rules of natively acquired natural languages are different from those of second languages, pidgins, or young creoles [...] because the variable forms of L2 learners are errors, they do not have conditioning linguistic contexts. This is what we refer to as inconsistent or unpredictable variation” (HUDSON KAM; NEWTON, 2005, p.145).

Diferente da pesquisa citada – cujos resultados foram obtidos com base em dados de língua artificial, em contexto de experimentação –, Singleton e Newport (2004) trazem resultados compatíveis com o papel da aquisição na mudança e regularização gramatical a partir de um estudo de caso. Foram investigados os efeitos de um *input* bastante inconsistente no processo de aquisição da língua americana de sinais (*American Sign Language, ASL*) por parte de Simon, uma criança surda de 7 anos de idade. Os cuidadores primários de Simon, também surdos, tiveram contato com a ASL apenas durante a adolescência (depois dos 15 anos) e, assim como outros aprendizes tardios de uma primeira língua, apresentam um desempenho bastante diferente dos sinalizadores nativos com exposição precoce à língua. No estudo, foi comparado o desempenho de Simon, de outras crianças cujos pais são sinalizadores nativos e dos pais do próprio Simon, em uma tarefa de morfologia da ASL. Os resultados obtidos mostram diferenças significativas entre a produção de Simon e a dos seus pais. No caso dos adultos, foram observadas diversas inconsistências e erros no uso da morfologia que não estão presentes na sinalização de Simon. De modo geral, Simon apresentou um desempenho próximo ao das crianças que receberam um *input* nativo precoce de ASL, em quase todos os aspectos testados. Os resultados do estudo sugerem que, a despeito do *input* primário recebido pela criança apresentar irregularidades e inconsistências, Simon foi capaz de adquirir um sistema de regras morfológicas regular e organizado.

Qual seria o mecanismo de aquisição subjacente que possibilita a configuração de uma língua interna que supera as irregularidades, inconsistências e lacunas observadas no *input* recebido? Singleton e Newport (2004) destacam que Simon parece prestar especial atenção à consistência ou regularidade dos mapeamentos entre forma e significado. Com base num grau moderado de consistência nesse tipo de mapeamento, Simon estabelece uma relação e aumenta a frequência de uso na sua própria produção. Por outro lado, mapeamentos com baixa frequência não são adquiridos, assim como formas frequentes, mas sem consistência na relação com um significado. Assim sendo, a frequência de ocorrência das formas por si só não seria suficiente para explicar o sucesso obtido por Simon na aquisição da ASL. A capacidade de “superar” o *input* irregular observada em Simon parece estar relacionada, em boa medida, a uma das propriedades básicas das línguas humanas salientadas por Chomsky (1981): serem compostas por um sistema computacional finito que permite a formação de um número

infinito de expressões, cada uma das quais possui uma interpretação definida no sistema semântico-pragmático e nos sistemas sensório-motores (e articulatórios perceptuais), isto é, em termos simplificados, combinando pensamento e som/gestos de forma bastante produtiva.

Ainda com relação ao papel da frequência, a literatura fornece evidências compatíveis com a ideia de que no processo de generalização de regras com base em um *input* inconsistente, o tipo de frequência – considerada em função de *tokens* ou *types* – parece ser bastante relevante nas etapas iniciais da aquisição. Em particular, para crianças na faixa dos 14 meses de idade, a frequência de *types*, mas não de *tokens*, parece ser decisiva para que uma generalização seja estabelecida (KOULAGUINA; SHI, 2019).

Considerando um tipo de contexto mais frequente na aquisição da linguagem, Miller e Schmitt (2012) investigaram o efeito do *input* variável na aquisição da morfologia de plural em duas variedades do espanhol: o espanhol chileno e o espanhol falado na cidade de México. Enquanto na variedade chilena o morfema de plural é omitido em função de um processo de lenição do /s/ em final de sílaba, a marcação morfofonológica de plural se mostra bastante consistente na variedade mexicana considerada. As autoras buscaram determinar em que medida um *input* variável quanto à morfologia de plural – ora com realização explícita, ora sem – pode afetar o processo de aquisição. Para tal, foram conduzidos um experimento de produção eliciada a partir de imagens e um teste de manipulação de brinquedos (*act-out*) com crianças nas faixas etárias compreendidas entre os 3 e os 7 anos de idade. O estudo incluiu três grupos experimentais: crianças mexicanas de classe baixa, crianças chilenas de classe baixa e crianças chilenas de classe média. A distinção entre classes baixa e média (*working-class* e *middle-class*, segundo a nomenclatura utilizada pelas autoras) foi baseada na classificação dos empregos desempenhados pelos adultos responsáveis pelas crianças. Os resultados obtidos revelaram diferenças no tipo de sintagmas utilizados para descrever conjuntos plurais. Enquanto as crianças chilenas produziram plurais e singulares nus (ex. *vacas/vaca* como resposta diante de uma pergunta como *O que tem aqui?*), as crianças mexicanas preferiram o plural indefinido (ex. *unas vacas*) para descrever conjuntos contendo mais de um elemento. As crianças chilenas de ambos os grupos sociais avaliados produziram menos marcação de plural do que as crianças mexicanas, mas essa diferença não foi significativa em termos estatísticos. Enquanto as crianças chilenas de classe média

e as crianças mexicanas apresentaram uma produção sistemática – seja omitindo ou produzindo sempre a marcação de plural – as crianças chilenas do grupo de classe baixa exibiram um padrão variável de produção. O experimento de compreensão, por sua vez, revelou diferenças na habilidade de associar o morfema de plural /-s/ nos plurais indefinidos ao sentido de “mais de um”. As crianças mexicanas, diferente das chilenas, tiveram um desempenho semelhante ao dos adultos do grupo de controle. Por fim, tanto as crianças expostas ao espanhol mexicano quanto ao espanhol chileno apresentaram sistematicidade nas suas interpretações do plural indefinido, seja demonstrando uma leitura de pluralidade ou não. Tomados em conjunto, as autoras consideram que os resultados são compatíveis com a ideia de que a natureza do *input* afeta o percurso de aquisição, com as crianças chilenas – que recebem um *input* variável – demorando mais tempo para adquirir a morfologia de plural do que as crianças mexicanas.

Os casos discutidos até aqui ora envolvem situações muito atípicas de privação de estímulo linguístico, ora trazem situações de *input* irregular e assistemático, ora se associam a contrastes em função de diferenças dialetais. No entanto, van Heugten e Johnson (2017) salientam que, mesmo entre crianças em contextos “típicos” adquirindo a mesma língua, o *input* recebido – em particular nos momentos iniciais do processo de aquisição – está longe de ser uniforme. Por exemplo, a depender do ambiente linguístico no qual estão inseridas, certas crianças serão expostas a uma dada variedade fonético/fonológica (no texto original, *accent*, doravante, VFF⁶ enquanto outras terão contato simultâneo com VFFs diversas.

Van Heugten e Johnson (2017) destacam que quase não existem estudos que investiguem o papel que a exposição precoce a um *input* fonética/fonologicamente mais variado teria no processo de aquisição e

⁶ Vários dos estudos que reportaremos ao longo das próximas seções utilizam o termo em inglês *accent*, para se referir ao que tradicionalmente tem sido traduzido para o português como *sotaque*. No contexto de boa parte das pesquisas sociolinguísticas recentes, esse termo tem sido evitado, embora essa não seja uma decisão categórica da área (cf. CARREÃO, 2019; HERNÁNDEZ-CAMPOY; CUTILLAS-ESPINOSA, BRITAIN; 2020, dentre outros). No presente artigo, utilizaremos a expressão *variedade fonético/fonológica* para denotar características fônicas sistemáticas que podem ser observadas na fala nos níveis segmental e/ou suprasegmental, podendo ou não ter implicações no nível fonológico. Tais características seriam particularmente relevantes para a identificação de grupos sociais, incluindo a distinção entre “nativo/ não nativo”.

reportam resultados que sugerem que a exposição diária a mais de uma VFF tem um forte impacto no desempenho dos bebês em uma tarefa experimental de reconhecimento de palavras. Foram testados bebês nas faixas dos 12, 14 e 18 meses de idade por meio de um experimento de escuta preferencial (*Headturn Preference Procedure*). Os bebês participantes foram separados em dois grupos: (i) de baixa variabilidade (expostos de forma quase exclusiva à variedade regional – inglês com acento canadense) ou (ii) de alta variabilidade (expostos a múltiplas variedades do inglês, representando o inglês canadense por volta de 33% do *input* total). Já a proporção de *input* em inglês (independente da variedade específica) não variou entre os dois grupos, ficando estabelecido em 99% e 98%, respectivamente.

Os resultados revelaram que os bebês expostos à baixa variabilidade no ambiente cotidiano rapidamente reconheceram palavras familiares a partir dos 12 meses de idade. Os bebês expostos a um *input* com maior grau de variabilidade, por sua vez, não demonstraram realizar uma distinção sistemática entre palavras e não palavras do inglês canadense até os 18 meses de idade⁷. Essa diferença foi observada a despeito de os dois grupos serem pareados também em função do tamanho do vocabulário dos participantes e do grupo socioeconômico.

Uma possível explicação para esses resultados aventada pelas autoras considera que talvez os bebês habituados a um *input* mais variado ficariam interessados tanto nas palavras que reconhecem quanto nas que não reconhecem e, nesse sentido, um estudo baseado em preferência – do tipo empregado – não seria um bom caminho para investigar suas habilidades de reconhecimento de palavras. No entanto, as próprias autoras apontam pelo menos dois motivos pelos quais essa interpretação dos dados não parece satisfatória. O primeiro é o fato de que os bebês do grupo de alta variabilidade demonstram de fato uma preferência por palavras familiares, apenas parecem demorar mais tempo para exibir tal preferência (por volta dos 18 meses em contraste com 12 meses para os bebês do grupo de baixa variabilidade). O segundo motivo está relacionado a achados prévios com bilíngues, uma vez que seria esperado que tivessem um comportamento

⁷ As autoras não apresentam quais seriam as variantes fonéticas ou fonológicas observadas nas variedades do inglês a que os bebês são expostos. Na atividade experimental, foram apresentadas palavras familiares e palavras inventadas produzidas por uma falante nativa do inglês canadense residente na grande Toronto (Canadá).

semelhante ao dos bebês em contexto de alta variabilidade. No entanto, experimentos com bilíngues utilizando a mesma técnica revelam preferências por palavras da língua já a partir dos 11 meses de idade.

Van Heugten e Johnson (2017) propõem uma interpretação alternativa para as diferenças observadas entre os grupos de bebês com base na ideia de que a representação de palavras (ou as estratégias que permitiriam mapear o sinal sonoro a determinadas representações) seria qualitativamente diferente em bebês que recebem um *input* mais ou menos variável. Para os bebês expostos a uma maior variabilidade em termos de VFFs, essas representações seriam menos restritas/categóricas (ou talvez menos específicas ou, até mesmo, subespecificadas?) do que para os bebês expostos a uma única VFF. Esse fato não impediria que, em condições naturalísticas, os bebês fossem capazes de reconhecer uma palavra, mas na ausência de um contexto – tal e como ocorre no paradigma experimental empregado no estudo – essa tarefa seria mais difícil para os bebês do grupo de alta variabilidade. Em idades mais avançadas, no entanto, essas mesmas crianças poderiam ter vantagem em situações nas quais precisem lidar com novas VFFs.

Ainda no que tange ao reconhecimento de palavras e os efeitos da variação linguística na aquisição, Best e colaboradores (2009) destacam que o reconhecimento eficiente de palavras no *continuum* da fala depende da sensibilidade às distinções fonológicas (*phonological distinctiveness*) que envolve a percepção das diferenças críticas entre os segmentos que permitem distinguir palavras ou não palavras contendo sons semelhantes (ex. *cake* versus *take* ou *pake* – como possíveis exemplos em português, *bola* versus *gola* versus *dola*). Uma outra habilidade, menos investigada até o momento, é a denominada constância fonológica que diz respeito à percepção de que uma palavra ainda é a mesma palavra a despeito de variações fonéticas que não envolvem um contraste no nível fonológico, tal e como acontece no caso da variação individual da fala ou na variação regional (BEST et al, 2009).

Estudos sobre a habilidade de distinção fonológica revelam que, por volta dos 18 a 19 meses de idade, os bebês rejeitam trocas de pronúncia em pares mínimos do tipo *baby/vaby* de forma mais consistente do que entre os 11 e os 17 meses de idade (HALLE; BOYSSON-BARDIES, 1996 apud BEST et al, 2009; STAGER; WERKER, 1997 apud BEST et al, 2009; SWINGLEY; ASLIN, 2002 apud BEST et al, 2009). Alguns resultados apontam, no entanto, que antes dos 19 meses os bebês teriam

uma sensibilidade compatível com tal habilidade, mas apenas em determinados contextos experimentais (FENNELL; WAXMAN, 2010 apud BEST et al, 2009; SWINGLEY; ASLIN, 2002 apud BEST et al., 2009). Best e colaboradores (2009) destacam que, embora experimentos de “pronúncia incorreta” possam fornecer informações sobre as habilidades de distinção fonológica, eles não permitem a investigação da habilidade complementar relacionada à constância fonológica. O estudo da variação dialetal, por sua vez, fornece uma ótima alternativa natural para investigar esse aspecto. Isso porque a variação no nível fonético pode diferir entre falantes nativos e não nativos de uma dada variedade, mas preservando a estrutura fonológica da palavra e a identidade lexical. Os autores investigaram a constância fonológica em bebês de 15 e de 19 meses de idade por meio de dois testes de preferência de familiaridade de palavras. Um dos testes foi construído a partir da variedade vernácula dos bebês (o inglês americano de Connecticut) e o outro com base numa variedade desconhecida pelos bebês (o mesoleto jamaicano, *i.e.*, um ponto intermediário no *continuum* pós-crioulo que se distingue do inglês americano em sua realização fonética das consoantes, vogais e padrões de acentuação). Foi observado que todas as crianças se mostraram capazes de reconhecer as palavras produzidas na sua variedade nativa/vernácula, mas apenas as crianças na faixa etária superior (19 meses) reconheceram palavras familiares apresentadas na variedade desconhecida. Assim, aos 15 meses de idade, o reconhecimento de palavras parece ficar restrito às realizações familiares para a criança e presentes no *input*. Segundo os autores, a emergência conjunta da distinção e da constância fonológica por volta dos 19 meses de idade poderia ser um indicador de uma consciência fonológica crescente, já que ambas as informações são cruciais na definição da identidade de uma palavra⁸.

Os resultados obtidos com bebês de idade inferior a dois anos de idade permitem tecer algumas considerações relevantes. Se, por um lado, nas etapas iniciais da aquisição a sensibilidade perceptual dos bebês se mostra bastante refinada (ASLIN; PISONI; JUSCZYK, 1983; EIMAS et al. 1971; JUSCZYK, 1997; WERKER; GERVAIN, 2013; dentre outros), tal sensibilidade sofre mudanças importantes ao longo

⁸ A identidade de uma palavra diz respeito ao reconhecimento de uma forma como sendo a mesma unidade a despeito de eventuais variações superficiais (irrelevantes para a interpretação semântica).

do primeiro ano de vida do bebê com um progressivo “afunilamento” ou especificação das representações fonológicas. A partir dos 5 meses, bebês demonstram preferência pela VFF familiar e local quando comparada a variedades desconhecidas e não locais (BUTLER; FLOCCIA; GOSLIN; PANNETON, 2011 apud JEFFRIES, 2019). Por volta dos 12 meses de idade, os bebês parecem ter estabelecido os contrastes fonéticos com valor fonológico na língua em aquisição, mas exibem uma falta de discriminação da maioria dos contrastes que não são relevantes para a língua em aquisição (WERKER, 1994). Em outras palavras, uma “explosão” perceptual inicial parece ocorrer na direção de possibilitar a identificação dos aspectos fonológicos cruciais para a gramática em aquisição, sobretudo para a aquisição lexical. Após essa fase, o processamento do sinal acústico não parece ocorrer com o mesmo nível de detalhamento, pelo menos, até etapas posteriores nas quais – como veremos mais adiante – irá, por exemplo, ocorrer a aquisição de novas informações cruciais para a detecção e identificação de VFF distintas.

Até aqui, vimos que a natureza do *input* recebido pela criança – inclusive a maior ou menor variabilidade presente nos dados – parece ter um papel relevante no processo de aquisição da linguagem. Vimos também que a relação entre forma e sentido, bem como a frequência com a qual as unidades (em diversos níveis) ocorrem nos dados, parecem ter um peso bastante importante no processo. Não é claro, contudo, quando e de que modo as crianças começam a relacionar a variabilidade presente no *input* a valores socialmente relevantes, adicionando assim mais uma camada de sentido no processo de interpretação de enunciados (e também de codificação de sentidos na produção). Na próxima seção, abordaremos a questão do desenvolvimento sociolinguístico na aquisição.

3 Desenvolvimento sociolinguístico e aquisição da linguagem

Investigar a percepção, produção e compreensão de variedades linguísticas estigmatizadas ou com menor prestígio social pode ser caracterizado como uma tarefa árdua. Em boa medida, as dificuldades encontradas estão relacionadas a aspectos da metodologia empregada que, no âmbito da psicolinguística, envolve muitas vezes material escrito e inclui tarefas de julgamento e/ou avaliação de sentenças pelos participantes das pesquisas.

Henry (2005), ao refletir sobre as dificuldades encontradas na investigação de aspectos sintáticos em variedades linguísticas não padrão, chama a atenção para o fato de que o julgamento dos participantes pode ser (fatalmente) afetado pelo fato de esses falantes terem o conhecimento de que muitas estruturas produzidas no seu vernáculo são consideradas erradas ou não gramaticais pelos falantes da variedade com maior prestígio social. Esse tipo de (auto)avaliação pode ser ainda reforçado pelo contexto educativo no qual variedades não padrão são muitas vezes rejeitadas e desvalorizadas.

Outro ponto relevante salientado por Henry (2005) diz respeito ao fato de que coletar intuições de falantes *naïves*⁹ pode ser mais difícil do que obter julgamentos a partir de falantes com um certo treinamento nesse tipo de tarefa. Aparentemente, seria mais simples que esse segundo grupo de falantes compreendesse, por exemplo, a natureza de um julgamento de gramaticalidade e entendesse a diferença entre tal julgamento e uma avaliação prescritivista ou, até mesmo, o que seria aceitável em termos pragmáticos. Nesse sentido, um falante não *naïve* teria uma facilidade maior em focar na sintaxe desconsiderando outros aspectos não relevantes, como por exemplo, as escolhas lexicais. Inclusive, há quem considere que variedades não padrão seriam mais fluidas e que, justamente a partir dos mecanismos de padronização (ou normatização), a gramática se tornaria mais clara e haveria – como consequência – um incremento na capacidade de julgamento de gramaticalidade por parte dos falantes (CHESHIRE; STEIN, 1997 apud HENRY, 2005).

Pesquisas revelam que, durante o processamento linguístico, falantes adultos são sensíveis aos efeitos da normatização, à frequência com a qual uma determinada variante ocorre e são influenciados pelos sentidos socialmente atribuídos para formas variáveis (ZAHARCHUK; SHEVLIN; VAN HELL, 2021). Dados da atividade cerebral obtidos a partir de potenciais relacionados a eventos se mostram compatíveis com a ideia de que informações de natureza sociolinguística, relacionadas à identidade do falante, parecem ter um impacto no modo como o cérebro processa estímulos linguísticos (HANULÍKOVÁ et al., 2012). Frente a violações gramaticais presentes na produção de um falante nativo de

⁹ A expressão *falante naïve* é comumente utilizada nos estudos em psicolinguística para fazer referência a um participante sem nenhum tipo de conhecimento especializado ou treinamento prévio sobre o objeto da pesquisa.

holandês, foram registradas ondas do tipo P600 na atividade elétrica cerebral dos participantes, mas quando as mesmas violações foram produzidas por um falante estrangeiro (turco como L1 e holandês como L2), esse efeito não foi constatado. Por outro lado, sentenças contendo violações de natureza semântica provocaram efeitos N400 comparáveis quando produzidas por ambos os falantes, o que sugere que não haveria um problema geral de integração vinculado a uma variação apenas na realização fonética dos enunciados. Os autores concluem que os resultados são compatíveis com a ideia de que o P600 – tradicionalmente vinculado a processos sintáticos e morfossintáticos – seria modulado pela identidade do falante. No entanto, não é claro ainda como essa consciência sociolinguística se desenvolve nem quando isso acontece ao longo do processo de aquisição da linguagem.

No que tange à detecção de variantes fonéticas e/ou fonológicas e sua relação com a cognição social, Creel (2018) destaca que a literatura apresenta resultados mistos. Enquanto adultos são capazes de realizar inferências sociais sutis com base na VFF dos falantes, achados de estudos conduzidos com crianças apresentam resultados variados. Se, por um lado, bebês muito novos – na faixa dos 5 meses de idade – demonstram preferência por VFFs familiares (BUTLER; FLOCCIA; GOSLIN; PANNETON, 2011 apud JEFFRIES, 2019), há estudos que reportam que crianças de até 5 anos de idade teriam dificuldades em identificar VFF como diferentes dos seus. Mesmo no caso da suposta preferência por VFF familiares nos bebês, tal tendência poderia estar associada apenas a uma pura familiaridade perceptual, sem qualquer associação com uma resposta de natureza social. Assim sendo, enquanto a preferência pela própria língua parece ser de fato bastante precoce, a efetiva detecção de diferenças menos evidentes, como àquelas associadas a variantes fonéticas regionais, não parece ocorrer tão cedo na aquisição.

Diversos estudos permitem estabelecer que bebês e crianças pequenas demonstram preferências pelo próprio grupo social em detrimento de outros grupos e que tais preferências estariam baseadas, pelo menos de forma parcial, na linguagem verbal (CREEL, 2018; MOON; COOPER; FIFER, 1993; NAZZI; BERTONCINI; MEHLER, 1998; KINZLER; DUPOUX; SPELKE, 2007; KINZLER; SHUTTS; DE JESUS; SPELKE, 2009; KINZLER; SHUTTS; SPELKE, 2012; dentre outros). Bebês entre 5 e 6 meses de idade olham por mais tempo para pessoas que previamente viram falar na sua língua nativa (no

caso, inglês americano), enquanto bebês de 10 meses preferem aceitar brinquedos oferecidos por falantes nativos da sua língua materna (inglês ou francês) e crianças na faixa dos 5 anos de idade escolhem de maneira seletiva outras crianças falantes da mesma língua (francês ou inglês) como potenciais amigos (KINZLER; DUPOUX; SPELKE, 2007). Pesquisas fornecem evidências de que crianças na faixa dos 5 a 6 anos de idade, com pouca exposição à variação linguística e cultural realizam avaliações sociais com base na língua. Crianças brancas, falantes de inglês e residentes em Boston nos Estados Unidos, na faixa dos 5 anos de idade, demonstram preferência por crianças que falam inglês americano, quando comparadas com crianças com o francês como L1 falando em inglês ou com crianças francesas falando francês (KINZLER; DUPOUX; SPELKE, 2007; KINZLER; SHUTTS; DE JESUS; SPELKE, 2009). Quando as preferências com relação à VFF e à variação racial foram investigadas, crianças demonstraram preferência por indivíduos de sua VFF (caracterizada pelos autores como *native accent*) e pertencentes a raças diferentes da sua (negra ou branca), frente a outras crianças de sua raça, mas com uma realização fonética-fonológica identificada como variedade estrangeira (*foreign accent*, inglês vs. francês) (KINZLER, SHUTTS, DE JESUS; SPELKE, 2009).

Ainda em relação a VFFs, Creel (2018) investigou a sensibilidade de crianças americanas frente a variantes nativas e não nativas. Foram conduzidos dois experimentos, sendo um de identificação de imagens com rastreamento ocular e outro conduzido a partir de uma tarefa de julgamento de amizade (nos moldes da utilizada por KINZLER; DUPOUX; SPELKE, 2007). O primeiro experimento buscou avaliar em que medida crianças entre 3 e 5 anos de idade seriam capazes de identificar falantes (sobre os quais já tinham um certo conhecimento prévio), em função da VFF. Os estímulos auditivos foram produzidos por falantes nativos de inglês americano e por falantes de inglês L2 e holandês L1, ora por vozes masculinas, ora femininas. Embora tenham sido registradas mais respostas corretas para os estímulos produzidos por falantes nativos, a diferença entre ambas as variedades investigadas não foi significativa. Os resultados relativos à movimentação ocular não forneceram evidências compatíveis com a ideia de que as crianças tenham utilizado pistas relativas à VFF dos falantes na identificação visual dos objetos.

No segundo experimento, foram testadas crianças entre 3 e 7 anos de idade no intuito de se verificar um possível efeito de idade no desenvolvimento da consciência sobre diferentes variantes. Foram

utilizados os mesmos estímulos auditivos do Experimento 1, mas numa tarefa diferente que envolvia duas etapas: escolha de amizade seguida por um julgamento de procedência. Na primeira etapa – seleção do amigo – dois rostos¹⁰ eram apresentados junto com os áudios correspondentes. Um rosto produzia a VFF americana enquanto o outro produzia a variedade não nativa. As crianças indicavam suas preferências de amizade apontando para o rosto da personagem escolhida. Após completar a primeira tarefa, era apresentada a tarefa de julgamento de procedência: depois de ouvir uma única frase (um dos enunciados previamente apresentado durante a escolha de amigos), era perguntado para os participantes se o falante “é daqui” ou “não é daqui”. As crianças forneceram respostas verbais para as perguntas realizadas. Diferente da tarefa de escolha de amizade, o julgamento de procedência demandaria o conhecimento da existência de uma correlação entre geografia e variabilidade na realização fonético/fonológica. Nesse sentido, a expectativa da autora era que as respostas em ambas as tarefas fossem coerentes entre si, caso as escolhas de amizade refletissem uma falta de familiaridade com a variedade não nativa.

Os resultados revelaram um gradiente de idade, com indícios de sensibilidade às VFFs nas crianças a partir dos 5 anos de idade. De acordo com Creel (2018), tal sensibilidade parece emergir lentamente e ir aumentando com a idade da criança. Nesse sentido, a identificação das diferentes VFFs e a atribuição de valor social para as mesmas constituiriam uma aprendizagem mais demorada que dependeria de uma exposição prolongada aos padrões sonoros (no caso específico desse estudo, entendidos como características segmentais e suprasegmentais) da língua.

No intuito de fornecer evidências que possam ampliar a compreensão do modo como se desenvolve a percepção sobre o significado de indexadores sociais da variação linguística pelas crianças, Jeffries (2019) investigou as habilidades de percepção e discriminação de variantes fonéticas regionais por crianças na faixa dos 3-4 anos de idade, residentes em York, no Reino Unido. Estudos prévios com participantes na mesma faixa etária indicam que, na produção, há evidências de mudança estilística¹¹ de certas variantes como resposta ao

¹⁰ Os estímulos visuais foram compostos por fotografias de rostos infantis extraídos do site flickr.com com licença do *Creative Commons*. Os rostos foram pareados por gênero, cor de olhos, cor e estilo de cabelo e tom de pele semelhantes.

¹¹ Com base em um dialeto escocês, Smith, Durham e Fortune (2007), consideram como variantes sensíveis à mudança de estilo a alternância entre ditongo e monotongo e o uso

input dos cuidadores (SMITH; DURHAM; FORTUNE, 2007; SMITH; DURHAM; RICHARDS, 2009) e de que a interação social entre pares afeta a adoção de variantes pelas crianças.

Na sua pesquisa, Jeffries (2019) buscou investigar a consciência das crianças relativa a características da variante indicativas da distinção entre falantes do norte e do sul da Inglaterra. A tarefa experimental consistia em um jogo de pareamento de personagens em função da variante fonética. Por exemplo, o bebê urso devia ser pareado com a mamãe urso que apresentasse o mesmo conjunto de variantes que ele. Foram investigadas as seguintes condições: (i) fones distintos, mas que não trazem uma distinção fonológica, na mesma palavra (b[a]sket/ b[ɑ:]sket); (ii) mesmo fone em palavras distintas (gr[a]ss e p[a]th / gr[ɑ:]ss e p[ɑ:]th); (iii) diferentes fones em diferentes palavras (gl[a]ss e br[e:]k *versus* gl[ɑ:]ss e br[ei]k). Foi avaliado ainda em que medida as habilidades de discriminação seriam afetadas pela idade e o sexo das crianças e pelo *input* fornecido pelos cuidadores primários no que tange às diferentes variantes fonéticas regionais.

Os resultados revelaram que crianças na faixa dos 3-5 anos de idade tiveram um desempenho acima do nível da chance na tarefa de agrupar falantes em função de propriedades vinculadas às variantes regionais. Esse resultado contraria achados prévios que sugeriam que essa habilidade só emergiria por volta dos 7 anos de idade (FLOCCIA et al, 2009 apud JEFFRIES, 2019) e traz dados parcialmente compatíveis com o observado por Creel (2018). Foi observada ainda uma melhora no desempenho das crianças em função da idade, com um efeito mais acentuado na condição de mesmo fone em palavras diferentes. A análise da variável sexo revelou que, em média, as meninas tiveram um desempenho melhor do que os meninos no teste. No entanto, a diferença apenas foi significativa em termos estatísticos na condição de mesmo fone e mesma palavra. A autora sugere que as diferenças observadas podem ser, pelo menos parcialmente, motivadas pelas próprias características da tarefa: centrada em falas com vozes femininas, com imagens representando personagens femininas e conduzida por uma experimentadora.

Quanto ao papel do *input* recebido pelas crianças, os resultados indicaram que as crianças cujos cuidadores apresentam variantes de fora de York tiveram um melhor desempenho na condição experimental

de -s em contextos de 3ª pessoa do plural. Smith, Durham e Richards (2009), por sua vez, investigam a deleção de -t/-d (ex. Lift your hanØ! It's not time for breakfasØ yet).

de diferentes fones em diferentes palavras, sugerindo que a variação presente nos dados recebidos pela criança pode ter contribuído para uma categorização mais precisa entre variante local/não padrão e não local/padrão. A autora destaca, no entanto, que o número de crianças cujos cuidadores produzem variantes não locais foi reduzido na amostra e, portanto, novos estudos precisam investigar melhor essa questão.

Uma questão que se destaca ao nos aprofundarmos nos estudos sobre percepção e identificação de VFF é a dimensão metodológica. Resultados diferentes geralmente se associam a tarefas experimentais diferentes. Os achados reportados por Jeffries (2019) permitem observar, de forma mais detalhada, os efeitos da constância fonológica em distintos contextos. Em conjunto, os resultados disponíveis na literatura parecem indicar um aumento progressivo da sensibilidade a aspectos relevantes da variação linguística, com a idade como fator decisivo, embora não único. Novamente, surge o papel do *input* primário recebido pelas crianças – isto é, o papel da exposição à variação na aquisição da própria variação – e as mudanças que os processos de percepção, identificação e categorização sofrem ao longo da aquisição da linguagem, conforme a criança se desenvolve¹².

Na próxima seção, retomaremos alguns dos pontos levantados até aqui, mas, dessa vez, a partir da análise e discussão sobre um fenômeno de variação específico, qual seja, a realização variável da marcação morfofonológica de plural no português brasileiro. Em particular, buscaremos caracterizar a manifestação desse fenômeno na produção infantil e no *input* fornecido pelos cuidadores primários das crianças. Para tal, apresentaremos dados coletados por Molina (2018) ao analisar a concordância verbal e Azalim (2021) ao investigar a concordância nominal. Além dos achados reportados nesses trabalhos, acrescentamos dados adicionais que ampliam os resultados iniciais sobre a concordância verbal, realizamos uma comparação inédita entre os padrões observados nas amostras investigadas e estabelecemos um contraponto entre o observado nos corpora de Molina (2018) e Azalim (2021) e o reportado

¹² Tais processos e suas transformações durante o percurso de desenvolvimento linguístico da criança podem ser considerados tanto em um domínio mais geral (não estritamente linguístico), quanto em um domínio específico vinculado à língua e, ainda, em uma eventual especialização – aparentemente mais tardia no curso temporal da aquisição da linguagem – para aspectos pontuais, por exemplo, relativos à variação.

por Reis (2020) com base em resultados experimentais. A comparação entre esses conjuntos de dados mostra-se particularmente interessante por terem sido coletados na mesma cidade, em períodos de tempo bastante próximos e com crianças em faixas etárias semelhantes.

4 Variação linguística e aquisição: a concordância variável no PB

A concordância de número é um dos fenômenos de variação mais largamente investigados na literatura sociolinguística sobre o PB. Embora o início das pesquisas remonte à década de 1970 (BRAGA, 1977; LEMLE; NARO, 1977; SCHERRE, 1978), até hoje o assunto se mostra relevante e motiva novas investigações (cf. MENDES; OUSHIRO, 2015 para uma revisão; ARAUJO; FREITAG, 2021, MARTINS; DE OLIVEIRA, 2015; OUSHIRO; GUY, 2015; VIEIRA, 2015; dentre outros).

Os estudos conduzidos no âmbito da sociolinguística – majoritariamente conduzidos com participantes adultos – permitem estabelecer a coexistência de dois padrões principais na marcação da morfologia de plural. De um lado, uma marcação morfofonológica redundante, na qual se observa uma reiteração da informação de pluralidade em todos os elementos gramaticalmente relevantes. Do outro, uma marcação não redundante, na qual a pluralidade é codificada de forma explícita em pelo menos um dos itens – geralmente localizado mais à esquerda na estrutura –, podendo não ocorrer a realização morfofonológica nos demais elementos. Vale destacar que esses mesmos padrões têm sido atestados na produção infantil, em ocorrências como as apresentadas em (1) e (2), respectivamente.

(1.a) As pecinhas pequenas. (AZALIM, 2021)

(1.b) As minhas orelhas ficam em cima da minha cabeça.
(MOLINA, 2018)

(2.a) Os pés vermelhoØ. (AZALIM, 2021)

(2.b) Seis bonéØ verdeØ. (AZALIM, 2021)

- (2.c) Elas duas apareceØ na quarta temporada. (MOLINA, 2018)
- (2.d) Lá em casa as minhaØ salaØ sempre foiØ pequena. (MOLINA, 2018)

Como pode ser observado nos exemplos em (1) e (2), ambas as regras gerais se verificam tanto no domínio nominal – na relação estabelecida entre artigo-nome-adjetivo-predicativo, etc. – quanto sentencial, na relação entre sujeito e verbo. Com base nas pesquisas mais antigas, a marcação redundante foi inicialmente vinculada à variedade de maior prestígio social, enquanto a codificação não redundante de pluralidade foi associada a variedades linguísticas estigmatizadas. Vale salientar, no entanto, que a alternância de padrões parece ser constitutiva – em proporções variáveis – da produção de falantes de grupos sociolinguisticamente diversos. Nesse sentido, a variação em questão não constitui um fenômeno restrito a uma ou outra variedade específica no PB, embora a prevalência de um ou outro padrão possa apresentar proporções distintas a depender de fatores tanto internos quanto externos.

Apesar de se tratar de um assunto bastante investigado, os estudos têm se concentrado na população adulta. Em comparação, o número de pesquisas conduzidas com crianças ainda é pequeno (AZALIM, 2021; CAPELLARI; ZILLES, 2002; GOMES; BENAYON; VIEIRA, 2006; JAKUBÓW, 2018; LAMPRETTCH, 1997; REIS, 2020; ROZA; CASAGRANDE, 2015 dentre outros). Além disso, alguns dos trabalhos que fornecem evidências sobre os efeitos da variabilidade na realização da concordância encontrada no *input* recebido pelas crianças durante o processo de aquisição não têm, de fato, o objetivo de investigar a variação em si nem o seu processo de aquisição (CASTRO; FERRARI-NETO, 2007; CORRÊA; AUGUSTO; FERRARI NETO, 2005; SIMIONI, 2006) ou, alternativamente, apresentam algum desses pontos apenas como objetivo secundário ou derivado de outras questões em discussão (MOLINA, 2018; SIMIONI, 2006).

Dentre os estudos conduzidos no âmbito da psicolinguística com foco na relação entre o *input* variável e o processo de aquisição, destacamos aqui a pesquisa de Jakubów (2018), que investigou crianças residentes na cidade do Rio de Janeiro. A autora parte da premissa de que a alternância entre os padrões de marcação de concordância teria uma

correlação direta com o nível de escolaridade, sendo influenciada ainda pelo nível socioeconômico do falante. A autora assume que a cidade do Rio de Janeiro configura-se como um ambiente de contato entre ambos os padrões, em virtude da interação constante entre pessoas de diferentes níveis socioeconômicos. Jakubów (2018) conduziu um experimento de produção eliciada por repetição com crianças na faixa dos 3;8 a 6;3 anos de idade e com estudantes do 6º ano do ensino fundamental (na faixa dos 10;11 e 14;6 anos de idade) frequentando escolas das redes pública e particular. Os resultados reportados indicam que tanto crianças em idade pré-escolar quanto os alunos do 6º ano produziram a variante redundante da concordância com mais frequência do que a não redundante. Foi observada ainda uma influência do fator nível socioeconômico nas taxas de produção da variante não redundante, com um número significativamente maior de produções contendo esse padrão por parte dos alunos da rede pública de ensino.

Resultados experimentais como os reportados por Jakubów trazem dados relevantes para a caracterização da produção infantil – e também permitem investigar a percepção e a compreensão –, mas não fornecem evidências diretas sobre a natureza do *input* recebido pelas crianças. Nesse sentido, dados naturalísticos e experimentais podem ser considerados de forma integrada para se obter um quadro mais detalhado do processo de aquisição e, em particular, para se compreender melhor o papel do *input* nesse processo.

Com base nos *corpora* coletados por Molina (2018) e Azalim (2021), é possível analisar a marcação da morfologia de plural, tanto no domínio nominal quanto no verbal, na produção infantil e dos seus cuidadores, em situações de interação espontânea. Serão considerados aqui três *corpora* que diferem entre si em função da idade, nível de escolaridade e procedência geográfica dos participantes. Em conjunto, os *corpora* fornecem informações sobre a produção de crianças e adultos, moradores de região urbana ou rural e com ensino superior ou médio (no caso dos adultos).

O corpus A (coletado por Molina, 2018) é composto por quatro crianças (de 3, 4, 5 e 6 anos de idade, respectivamente), residentes na zona urbana de Juiz de Fora/MG, frequentadoras de escolas privadas, e suas mães, que possuem alto nível de educação formal. O corpus B (Molina, 2018) é formado por gravações pontuais em sala de aula de uma escola localizada na zona rural de Juiz de Fora, com crianças de 4, 5 e 6 anos.

Nesse caso, contamos apenas com a produção infantil, sem evidências diretas do *input* recebido pelas crianças. Finalmente, o corpus C (coletado por Azalim, 2021), é formado por duas crianças (de 3 e 5 anos de idade, respectivamente), residentes na zona urbana de Juiz de Fora, frequentando escola privada e cujas mães possuem nível médio de escolaridade formal. Esse terceiro corpus foi coletado para que fosse possível estabelecer uma comparação entre crianças residentes em zona urbana no intuito de se observar um possível efeito do nível de escolaridade do cuidador primário no processo de aquisição da criança. A Tabela 1, a seguir, apresenta a distribuição das ocorrências de sintagmas plurais (redundante X não redundante) em função do tipo de concordância (nominal e verbal) e da idade das crianças em cada corpus. A análise que apresentamos aqui se limita ao contraste geral entre os dois padrões principais de marcação de plural nos âmbitos nominal e verbal, sem adentrarmos em outros aspectos que têm sido levantados pela literatura (ex. saliência fônica dos itens, posição do sujeito, natureza do sujeito – nulo/preenchido, semântica do verbo, etc.) e investigados de forma detalhada nos estudos de Molina (2018) e Azalim (2021).

TABELA 1 – Distribuição de ocorrências de sintagmas plurais (redundantes X não redundantes), tipo de concordância (nominal e verbal) e idade em cada *corpus*

Participantes		Padrão de concordância nominal no total de ocorrências e %				Padrão de concordância verbal no total de ocorrências e %			
<i>Corpus</i>	Idade	Redundante		Não		Redundante		Não Redundante	
Corpus A Zona urbana cuidados com alto nível de escolaridade	3 anos	50	78%	14	22%	21	50%	21	50%
	Adulto	70	99%	01	1%	23	96%	1	4%
	4 anos	45	98%	01	2%	4	67%	2	33%
	Adulto	92	93%	07	7%	40	83%	8	17%
	5 anos	66	85%	12	15%	14	67%	7	33%
	Adulto	161	97%	5	3%	75	99%	1	1%
	6 anos	115	94%	7	6%	44	77%	13	23%
	Adulto	111	83%	22	17%	47	80%	12	10%
Totais gerais		710	91%	69	9%	268	80,5%	65	19,5%

	Totais crianças	276	89%	34	11%	83	66%	43	34%
	Totais adultos	434	93%	35	7%	185	89%	22	11%
Corpus B Crianças zona rural *Sem dados dos cuidadores	4 anos	6	50%	6	50%	1	14%	6	86%
	5 anos	8	50%	8	50%	2	33%	4	67%
	6 anos	15	79%	4	21%	7	35%	13	65%
	Totais crianças	29	62%	18	38%	10	30%	23	70%
Corpus C Zona urbana Cuidador com ensino médio	3 anos	9	75%	3	25%	0	-	0	-
	Adulto	74	67%	37	33%	6	40%	9	60%
	5 anos	113	84%	22	16%	21	75%	7	25%
	Adulto	121	86%	20	14%	29	71%	12	29%
	Totais gerais	317	79%	82	21%	56	67%	28	33%
	Totais crianças	122	83%	25	17%	21	75%	7	25%
	Totais adultos	195	77%	57	23%	35	63%	21	37%

Em primeiro lugar, vamos considerar os dados de produção infantil de forma isolada¹³. O corpus A (zona urbana, cuidador com alto nível educativo) revelou uma produção significativamente maior do padrão redundante na concordância nominal (89% redundante e 11% não redundante), ($\chi^2 = 429,08$ (2), $p < 0,01$). O mesmo predomínio do padrão redundante foi observado para o sintagma verbal (66% redundante e 34% não redundante) ($\chi^2 = 12,071$ (1), $p < 0,01$).

No que tange ao corpus B (crianças, zona rural), também se observa um percentual maior de ocorrências redundantes para o sintagma nominal (62% redundante e 38% não redundante), mas nesse caso a

¹³ Foram realizados testes de proporção para avaliar a existência de diferenças significativas entre as proporções das variantes investigadas (no caso, redundante X não redundante) e de variáveis nominais (aqui representadas pela concordância nominal X verbal). O teste de qui-quadrado de Pearson foi utilizado para comparar as proporções entre grupos, ou seja, duas variáveis nominais / categóricas independentes entre si. Aqui, estamos comparando as proporções de produções não redundantes nas diferentes amostras.

comparação entre o número de ocorrências redundantes e não redundantes não se mostrou estatisticamente significativa ($\chi^2 = 3,13$ (1), $p=0,07$). Além disso, contrariamente ao observado no corpus A, os resultados para a concordância verbal indicam uma predominância do padrão não redundante (com apenas 30% de marcação redundante na amostra) ($\chi^2 = 4,363$ (1), $p<0,05$).

Os resultados do corpus C (zona urbana, cuidador com ensino médio) para a produção infantil, por sua vez, se mostram compatíveis com o observado para o corpus A. Foi registrada uma predominância do padrão redundante tanto no domínio nominal quanto verbal: 83% redundante e 17% não redundante para os sintagmas nominais ($\chi^2 = 75,57$ (1), $p<0,01$) e 75% redundante e 25% não redundante para a concordância verbal ($\chi^2 = 6,035$ (1), $p<0,01$).

Embora as amostras possam ser consideradas reduzidas do ponto de vista sociolinguístico, seguindo uma abordagem padrão nos estudos sobre aquisição da linguagem – na qual é frequente o trabalho com amostras pequenas e não totalmente equilibradas –, é possível sinalizar alguns pontos relevantes. A comparação entre as produções não redundantes – tanto no domínio nominal quanto verbal – entre os três grupos indica diferenças apenas entre o corpus B e os outros dois corpora, sendo: A *versus* B ($\chi^2 = 19,97$ (1), $p<0,01$) e B *versus* C ($\chi^2 = 64,00$ (1), $p<0,01$), no domínio nominal, e A *versus* B ($\chi^2 = 12,202$ (1), $p<0,01$) e B *versus* C ($\chi^2 = 10,38$ (1), $p<0,01$), no domínio verbal. No entanto, não se observaram diferenças relevantes entre os corpora A e C entre si, no que tange à concordância nominal ($\chi^2 = 2,71$ (1), $p<0,09$) nem tampouco à verbal ($\chi^2 = 0,503$ (1), $p=0,4$). Diante disso, levantamos as seguintes considerações:

A procedência geográfica parece se configurar como um fator relevante para a alternância entre os padrões de marcação de plural na produção infantil. As principais diferenças observadas na produção infantil ocorreram em função do contraste urbano X rural.

A princípio, o nível de escolaridade dos cuidadores primários não parece ter influenciado a produção das crianças, já que a comparação entre as produções das crianças dos corpora A e C não revelou efeitos estatisticamente significativos. No entanto, como veremos a seguir, o quadro é um pouco mais complexo do que isso.

Quanto ao *input* fornecido pelos cuidadores primários, foi observado que a produção das crianças foi semelhante a do seu adulto de referência na concordância nominal. Já para a concordância verbal

a situação foi similar com exceção da criança de 3 anos do corpus A: enquanto a mãe quase não produziu o padrão não redundante de marcação no verbo, a criança apresentou uma produção “mista” com 50% de ocorrências para cada um dos padrões de marcação de plural. De modo geral, excluindo o caso anteriormente mencionado, a produção adulta e infantil parecem estar alinhadas no que tange à prevalência dos padrões gerais de marcação.

No que diz respeito à produção adulta, no corpus A foi observada uma diferença entre o número de produções redundantes e não redundantes, tanto para a concordância nominal ($\chi^2 = 337,75$ (1), $p < 0,01$) quanto verbal ($\chi^2 = 126,78$ (2), $p < 0,01$), com predomínio do padrão redundante em ambos os casos. Já no corpus C ($\chi^2 = 74,48$ (1), $p < 0,01$), embora para a concordância nominal seja atestada a mesma tendência observada no corpus A, a diferença entre produções redundantes e não redundantes no domínio verbal foi menos evidente ($\chi^2 = 3,01$ (1), $p < 0,08$).

Diferentemente do que foi observado na produção infantil, ao compararmos entre si as proporções de produção não redundante dos adultos dos corpora A e C, foi registrada uma diferença com o padrão não redundante (na concordância nominal ($\chi^2 = 32,479$ (1), $p < 0,01$), e verbal ($\chi^2 = 21,34$ (1), $p < 0,01$)) sendo produzido mais pelas 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.

A comparação da produção dos adultos traz um outro ponto relevante: mesmo quando o *input* recebido pelas crianças é diferente – como sugerido pelos contrastes encontrados na produção dos cuidadores¹⁴ –, as crianças desses dois grupos não apresentam diferenças entre si ($\chi^2 = 2,71$ (1), $p < 0,09$). Esse resultado pode indicar que: (i) as crianças não se limitam apenas a reproduzir os dados disponíveis; (ii) os efeitos da escolarização e sua ação normativizadora são capturados na produção dos adultos, mas não na fala das crianças que ainda se encontram em

¹⁴ Vale lembrar que a produção das crianças não reproduz a produção da mãe. De fato, é possível observar nas amostras diferentes tendências (ex. no grupo A, a criança de 5 anos produz menos concordância redundante do que sua mãe, enquanto a criança de 6 anos e sua mãe mostram a tendência oposta).

uma fase anterior ao início da alfabetização¹⁵. Vale destacar que, embora em linhas gerais a produção infantil se mostre compatível com a dos seus cuidadores, essa produção não é idêntica, como pode ser observado em função das diferenças encontradas na comparação dos corpora considerando crianças e adultos separadamente.

Cabe ressaltar, ainda, que os dados naturalísticos reportados não se mostram compatíveis com uma marcação categórica da concordância nem nos adultos e nem nas crianças, em nenhum dos grupos considerados. Todos os falantes apresentaram algum grau de variação entre a marcação redundante e não redundante em ambos os âmbitos nominal e verbal.

Considerando estudos recentes no âmbito da aquisição da linguagem sobre o mesmo fenômeno, é possível observar que o trabalho de Reis (2020), desenvolvido com base em metodologia experimental, traz resultados um pouco diferentes e uma interpretação que também se distancia do que discutiremos mais à frente. Os achados desse estudo são particularmente interessantes já que, embora existam diferenças metodológicas importantes, todos os dados foram coletados na mesma cidade (Juiz de Fora/MG), em períodos de tempo bastante próximos, com participantes nas mesmas faixas etárias consideradas nos nossos dados naturalísticos. Além disso, as análises dos dados também consideraram fatores socioeconômicos.

Reis (2020) conduziu um experimento de produção eliciada com crianças nas faixas dos 3, 4 e 5 anos de idade frequentando escola pública ou particular (i.e. um correlato socioeconômico da distinção estabelecida por nós em termos de nível educativo dos cuidadores primários). Durante a tarefa experimental, as crianças eram convidadas a participar

¹⁵ Como observado por um/a parecerista anônimo/a, é necessário considerar ainda possíveis efeitos da metodologia adotada: seguindo um procedimento comum nas pesquisas sobre aquisição, as gravações foram realizadas pelos próprios cuidadores, em situações do cotidiano, visando a se obter registros o mais espontâneos possíveis. Embora no caso das crianças a validade ecológica dos dados provavelmente seja bastante alta, no caso dos adultos responsáveis, existe a consciência de que a produção linguística está sendo gravada e que a mesma será analisada para fins de pesquisa. Além disso, considerando os critérios usualmente adotados no campo da sociolinguística, a metodologia de coleta não seguiu nenhum protocolo estruturado, embora todas as amostras compartilhem o mesmo escopo temporal. Como salientado por Freitag e Snichelotto (2015), escolhas metodológicas da constituição das amostras orais são fundamentais para a generalização de resultados.

de uma brincadeira na qual deviam contar para um fantoche que estava vendado onde estavam os objetos – apresentados em conjuntos de um ou mais elementos – manuseados pela pesquisadora. Assim, a criança poderia produzir enunciados do tipo *A flor vermelha está na caixa* ou *A bolinha amarela está na caixa*, tanto no singular, quanto no plural. Além das crianças, a autora incluiu também um grupo de controle formado por adultos. A seguir, apresentamos uma síntese dos resultados mais relevantes para nossa discussão.

TABELA 2 – Distribuição de ocorrências de sintagmas plurais (redundantes X não redundantes), tipo de concordância (nominal e verbal) e grupo de participantes (crianças X adultos) reportada por Reis (2020)

Participantes	Padrão e concordância nominal no total de ocorrências e %				Padrão de concordância verbal no total de ocorrências e %			
	Redundante		Não Redundante		Redundante		Não Redundante	
Crianças	467	34%	903	66%	154	38%	255	62%
Adultos	213	84%	40	16%	47	100%	0	--

Os resultados indicam que, no grupo de adultos, o padrão não redundante ocorreu apenas no domínio nominal, ao passo que não houve nenhum caso de marcação não redundante no âmbito verbal. Diferentemente, as crianças produziram ambos os padrões de concordância nos dois ambientes: 34% de concordância redundante e 66% de concordância não redundante para o sintagma nominal e 38% de concordância redundante e 62% não redundante para a relação entre sujeito e verbo. O tratamento estatístico dos dados de concordância nominal mostrou que não houve diferença significativa entre as crianças que frequentavam escola pública e privada nem em função da idade das crianças. Já a análise relativa à concordância verbal revelou um efeito de idade, sugerindo uma diminuição do uso do padrão não redundante em função do aumento da idade da criança. O tipo de escola novamente não se mostrou significativo.

Com base nas diferenças na realização dos padrões de concordância em função de idade, Reis (2020) levanta a hipótese de que a concordância de número no PB está relacionada com processos distintos de aquisição, se considerados os sintagmas nominal e verbal. Enquanto para a concordância nominal, a autora propõe que as crianças estariam adquirindo os dois padrões possíveis, no caso da concordância verbal, Reis (2020) defende que a variação observada seria de outra natureza. As crianças, inicialmente, “regularizariam negativamente” a concordância de número em suas produções linguísticas (produção de verbos sem a marca explícita de plural) e, em seguida, haveria um período em que se verificaria a competição das duas gramáticas (marcação redundante e não redundante no verbo), até que, por fim, a aquisição caminharia para uma marcação explícita do verbo (regularização positiva da concordância de número no sintagma verbal). Em outras palavras, Reis assume que a verdadeira variação ocorreria no caso da concordância nominal, apenas.

Os resultados relatados por Reis (2020) diferem substancialmente dos dados naturalísticos reportados por nós. Em primeiro lugar, a autora encontra uma clara preferência pela marcação não redundante na produção infantil, fato que apenas foi observado nos dados naturalísticos para a concordância verbal no grupo de crianças residentes na zona rural nos nossos *corpora*. Em segundo lugar, os adultos do grupo controle não produziram o padrão não redundante para a concordância verbal, sendo que nos dados naturalísticos observamos que o domínio verbal parece ser aquele em que o padrão não redundante ocorre com mais frequência em todos os grupos, tanto pelos adultos quanto pelas crianças. Essas diferenças podem ser explicadas com base na metodologia adotada em cada caso. Azalim (2020) reporta um aumento significativo da produção do padrão não redundante por crianças em situação experimental, quando comparado com o observado previamente em dados naturalísticos. Nesse estudo em particular, a tarefa empregada envolvia a interação entre duplas de crianças, trazendo à tona mais uma vez a influência que a identidade do interlocutor pode ter no processamento linguístico, no caso, na produção. Na pesquisa de Reis (2020), por sua vez, a interação era “triangulada” entre a criança, a experimentadora e um fantoche, no contexto de uma atividade lúdica que pode ter gerado um tipo de engajamento diferente do observado na interação espontânea entre mãe-criança ou professora-crianças que caracteriza os dados naturalísticos aqui reportados. Assim, diferenças na coleta dos dados podem ser, pelo menos em parte, responsáveis pelas diferenças observadas nos resultados reportados.

Outro ponto relevante está relacionado às considerações de Reis quanto às eventuais diferenças na aquisição da concordância nominal e verbal. De acordo com a autora, a produção das crianças no domínio verbal não estaria de fato relacionada a uma verdadeira variação, mas faria parte do processo de aquisição da marcação de plural na desinência verbal. Os dados naturalísticos dos falantes adultos fornecidos pelo nosso corpus C podem fornecer evidências que questionam tal interpretação. Longe de convergir na direção de uma gramática com preferência pela marcação redundante no verbo, não foi observada diferença significativa na marcação redundante e não redundante na produção dos adultos do corpus C. Até mesmo os informantes adultos do corpus A apresentaram uma taxa não desprezível de marcação não redundante no sintagma verbal. Mais uma vez, questões metodológicas somadas ao papel da escolaridade e seus efeitos de normatização podem estar relacionados às diferenças observadas. Diferentemente dos contextos de interação espontânea entre mãe-criança registrados nos dados naturalísticos, os resultados experimentais de Reis (2020) foram coletados em contexto avaliativo, o que pode ter colaborado para favorecer um maior monitoramento da fala por parte dos participantes.

Finalmente, Reis (2020) não encontra diferenças na produção infantil em função de as crianças frequentarem escola pública ou particular (um indicador vinculado a grupo socioeconômico). Embora nos nossos dados naturalísticos o indicador considerado tenha sido nível de escolaridade – e não grupo socioeconômico –, é possível tecer algumas considerações com relação a esse ponto. A ausência de efeitos de grupo socioeconômico reportada por Reis é compatível com o observado ao compararmos o nível educativo das mães das crianças que compõem os nossos corpora A e C. Por outro lado, os dados naturalísticos revelaram que essa ausência de contrastes entre a fala das crianças não se verifica no caso dos adultos. Nesse sentido, é possível imaginar que os efeitos de normatização, em boa medida vinculados ao ensino formal, somente sejam evidentes em crianças com idade mais avançada que as consideradas no presente estudo e na pesquisa de Reis (2020). Diante disso, um ponto relevante a ser aprofundado é em que momento fatores de cunho socioeconômico começam a ter de fato uma influência clara na produção dos falantes.

Ao longo desta seção, buscamos trazer dados do PB para a discussão sobre o papel da variação na aquisição. A análise de dados naturalísticos e sua posterior comparação com o reportado em estudos

conduzidos com base em metodologia experimental (em particular, REIS, 2020) corroboram dois dos pontos já destacados anteriormente. Em primeiro lugar, o fato de que pesquisas conduzidas no âmbito da aquisição da linguagem podem contribuir para uma melhor compreensão da variação linguística, para além da fala infantil. Em segundo lugar, que aspectos vinculados ao método de coleta podem ter grande impacto na caracterização dos fenômenos de variação sob investigação, tanto nas crianças quanto nos adultos.

5 Considerações finais

No presente artigo, buscamos tecer uma discussão ao redor de duas questões frequentemente tomadas como indissociáveis, mas que apresentam aspectos intrínsecos bastante relevantes. De um lado, a questão do papel do *input* variável no processo de aquisição de uma língua materna. Do outro, o desenvolvimento sociolinguístico que ocorre em algum momento – que ainda precisa ser melhor especificado – no decorrer desse mesmo processo. Buscando ilustrar alguns dos aspectos discutidos, trouxemos dados relativos a um fenômeno de variação linguística do PB, qual seja, a concordância variável de número.

No que concerne à primeira questão, nossos dados de interação espontânea, em alguma medida, relativizam o papel do *input*, já que a produção das crianças não parece de fato espelhar a produção adulta. Quanto ao desenvolvimento sociolinguístico, embora não tenha sido esse o foco, vale a pena retomar os achados de Azalim (2020). A autora observou aumento significativo da produção do padrão não redundante por crianças interagindo com outra criança em situação experimental, em comparação com a produção infantil em interações espontâneas com adulto. Tais diferenças poderiam ser, ao menos em parte, interpretadas como decorrentes da influência da identidade do interlocutor na interação (socio-)linguística.

Ao longo da revisão da literatura e das reflexões levantadas, alguns pontos surgiram de forma recorrente. Um deles se relaciona com a dimensão metodológica das pesquisas: dados coletados a partir de ferramentas distintas muitas vezes trazem resultados discrepantes, quando não, antagônicos. Outra questão que acaba chamando a atenção para aspectos metodológicos diz respeito ao impacto da escolaridade na manifestação da variação linguística. Nesse sentido, resultados discrepantes com adultos e com crianças podem estar relacionados não

apenas com o processo de aquisição da linguagem em sentido restrito, mas também com aspectos vinculados à visão alargada da aquisição, que envolve questões vinculadas à alfabetização e ao letramento.

Esse último ponto nos leva a refletir sobre as características das populações usualmente investigadas nas pesquisas sobre processamento e aquisição da linguagem. Assim como ocorre nas pesquisas com falantes adultos, que têm como alvo mais frequente as denominadas populações WEIRD (*Western, Educated, Industrialized, Rich, Democratic*), nos estudos conduzidos com crianças predominam participantes brancos, de classe média ou média alta (JOHNSON; WHITE 2020). Trabalhos que considerem outros perfis são proporcionalmente escassos. Uma maior diversificação das populações estudadas pode ser a chave para novas descobertas que irão esclarecer muitas das questões ainda obscuras na literatura da área. A questão das populações investigadas pode ser relacionada ainda com o grau de variabilidade encontrada no *input* variável. Como vimos, a exposição a um *input* mais ou menos variável parece ter efeitos importantes no processo de aquisição e, principalmente no desenvolvimento da consciência sociolinguística dos falantes. Nesse sentido, o *background* dos participantes das pesquisas no que tange ao seu maior ou menor contato com variação linguística é uma questão que merece bastante atenção.

Finalmente, questões relacionadas à interação comunicativa precisam ser observadas com cuidado já que podem estar – junto com aspectos metodológicos mais amplos – na base de alguns dos resultados conflitantes ou de difícil análise/interpretação reportados na literatura.

Declaração de autoria

O presente artigo é resultado do trabalho em parceria de suas autoras. Detalhamos a seguir as etapas nas quais cada autora se envolveu de forma mais direta. A primeira autora participou na delimitação do encaminhamento teórico, na revisão e discussão da literatura específica sobre *input* variável e aquisição da linguagem e sobre desenvolvimento sociolinguístico, na consolidação da discussão dos resultados e na redação do texto do artigo. As principais contribuições da segunda autora foram na consolidação da discussão dos resultados e na redação do texto do artigo. A terceira autora contribuiu na coleta e análise de dados relativos à concordância nominal, na revisão e discussão da literatura relativa à concordância nominal e na aplicação de testes de inferência estatística.

A contribuição da quarta autora se concentrou principalmente na coleta e análise de dados relativos à concordância verbal e na revisão e discussão da literatura relativa a concordância verbal.

Referências

ARAUJO, S. S. F.; FREITAG, R. M. K. Concordância verbal, difusão da mudança linguística no contínuo rural-urbano e mudança em curto espaço de tempo. *Working Papers em Linguística*, Florianópolis, v. 22, n.2, p. 266-294, 2021. DOI: <https://doi.org/10.5007/1984-8420.2021.e76094>

ASLIN, R. N.; PISONI, D. B.; JUSCZYK, P. W. Auditory Development and Speech Perception in Infancy. In: Haith, M.; Campos, J. (eds.). *Carmichael's Handbook of Child Psychology: Infancy and Developmental Psychobiology*. New York: Wiley. 1983. p. 573–670.

AZALIM, C. *Concordância nominal variável na produção infantil: dados naturalísticos, experimentais e caracterização formal*. 2021. 249f. Tese (Doutorado em Linguística) – Faculdade de Letras, Universidade Federal de Juiz de Fora, 2021.

BEST, C. T.; TYLER, M. D.; GOODING; T. N.; ORLANDO, C. B.; QUANN, C. A. Development of phonological constancy: toddlers' perception of native – and Jamaican – accented words. *Psychological Science*, v. 20, n. 5, p. 539-42, 2009. DOI: <https://doi.org/10.1111/j.1467-9280.2009.02327.x>

BRAGA, M. L. *A concordância de número no sintagma nominal no Triângulo Mineiro*. 1977. 88f. Dissertação. (Mestrado em Língua Portuguesa) – Departamento de Letras, Pontifícia Universidade Católica do Rio de Janeiro, 1977.

CAPELLARI, E. T.; ZILLES, A. M. S. A marcação de plural na linguagem infantil – estudo longitudinal. *Revista da ABRALIN*, [s.l.] v. 01, n. 1, p. 185-218, 2002. Disponível em: <https://revista.abralin.org/index.php/abralin/article/view/733>. Acesso em: 6 mar. 2023.

CARREÃO, V. Linguistic evaluation and variation: Accent marks in the opinion of Louveirenses. *Revista de Estudos Linguísticos*, Belo Horizonte, v. 27, n. 4, p. 1649-1670, 2019. DOI: 10.17851/2237-2083.27.4.1649-1670

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, p.65-76, 2007. Disponível em: <https://revistaseletronicas.pucrs.br/ojs/index.php/fale/article/view/671>

CHOMSKY, N. *Lectures on government and binding*. Dordrecht: Foris, 1981.

CHOMSKY, N. *Knowledge of language*. New York: Praeger, 1986.

CLARK, H. H. Psycholinguistics. In: WILSON, R.A.; KEIL, F.C. (eds.). *The MIT Encyclopedia of the cognitive sciences*. Cambridge: The MIT Press, 1999. p. 688-689.

CORNIPS, L.; SWANENBERG, J.; HEERINGA, W.; DE VRIEND, F. The relationship between first language acquisition and dialect variation: Linking resources from distinct disciplines in a CLARIN-NL project. *Lingua*, v. 178, Special Issue p. 32-45, 2016. DOI: <https://doi.org/10.1016/j.lingua.2015.11.007>

CORRÊA, L.M.S. Aquisição da linguagem: uma retrospectiva dos últimos trinta anos. *DELTA: Documentação de Estudos em Linguística Teórica e Aplicada*, São Paulo, v. 15, n. especial, p. 339-383, 1999. DOI: <https://doi.org/10.1590/S0102-44501999000300014>

CORRÊA, L.M.S. Aquisição e processamento da linguagem: uma abordagem integrada sob a ótica minimalista. *Gragoatá*, Niterói, v. 16, n. 30, p. 55-75, 2011. DOI: <https://doi.org/10.22409/gragoata.v16i30.32923>

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: 30th BUCLD. Boston: Cascadilla Press, 2005. Online Proceedings Supplement.

CREEL, S. C. Accent detection and social cognition: evidence of protracted learning. *Developmental Science*, v. 21, n. 2, 2018. DOI: <https://doi.org/10.1111/desc.12524>

CURTISS, S. *Genie: A psycholinguistic study of a modern day "wild child"*. New York: Academic Press, 1977.

EIMAS, P.; SIQUELAND, E.; JUSCZYK, P.; VIGORITO, J. Speech Perception in Infants. *Science*, v. 171, n. 3968, p. 303–306, 1971. DOI: <https://doi.org/10.1126/science.171.3968.303>

FREITAG, R. M. K.; ROST SNICHELOTTO, C. A. Análises contrastivas: estabilidade, variedade ou metodologia? *Working Papers em Linguística*, Florianópolis, v. 16, n.1, p. 157-167, 2015. DOI: <https://doi.org/10.5007/1984-8420.2015v16n1p157>

FENNELL, C.; WAXMAN, S. What paradox? Referential cues allow for infant use of phonetic detail in word learning. *Child Development*, v. 10, n. 5, p. 1376-1383, 2010. DOI: <https://doi.org/10.1111/j.1467-8624.2010.01479.x>

GAHL, S.; GARNSEY, S.M. Knowledge of Grammar, Knowledge of Usage: Syntactic Probabilities Affect Pronunciation Variation. *Language*, v. 80, n. 4, p. 748–775, 2004. DOI: <http://www.jstor.org/stable/4489780>

GERVAIN, J.; WERKER, J. F. How Infant Speech Perception Contributes to Language Acquisition. *Language and Linguistics Compass*, v. 2, n. 6, p. 1149-1170, 2008. DOI: <https://doi.org/10.1111/j.1749-818X.2008.00089.x>

GOMES, C. A.; BENAYON, A. R.; VIEIRA, M. C. P. Variação e aquisição da flexão nominal e da flexão verbal. *Gragoatá*, Niteroi, v. 30, n. 1, p. 39-54, 2006. DOI: <https://doi.org/10.22409/gragoata.v16i30.32922>

HANULÍKOVÁ, A.; VAN ALPHEN, P. M.; VAN GOCH, M. M.; WEBER, A. When one person's mistake is another's standard usage: The effect of foreign accent on syntactic processing. *Journal of Cognitive Neuroscience*, v. 24, n. 4, p. 878–887, 2012. DOI: https://doi.org/10.1162/jocn_a_00103

HENRY, A. Non-standard dialects and linguistic data. *Lingua*, v. 115, n. 11, p. 1599-1617, 2005. DOI: <https://doi.org/10.1016/j.lingua.2004.07.006>

HUDSON KAM C. L. H.; NEWPORT, E. L. Getting it right by getting it wrong: when learners change languages. *Cognitive Psychology*. v. 59, n. 1, p. 30-66, 2009. DOI: <https://doi.org/10.1016/j.cogpsych.2009.01.001>

HERNÁNDEZ-CAMPOY, J. M.; CUTILLAS-ESPINOSA, J. A.; BRITAIN, D. Variação e competência sociolinguísticas no ensino de inglês como língua estrangeira. *Entrelinguas*, Araraquara, v. 6, n. 1, p. 183-201, 2020. DOI: <https://doi.org/10.29051/el.v6i1.13379>

HUDSON KAM, C. L. H.; NEWPORT, E. L. Regularizing Unpredictable Variation: The Roles of Adult and Child Learners in Language Formation

and Change. *Language, Learning and Development*, v. 1, n. 2, p. 151-195, 2005. DOI: https://doi.org/10.1207/s15473341lld0102_3

JAKUBÓW, A. P. S. P. *Language acquisition based on variable input: the case of number agreement in Brazilian Portuguese*. 2018. Tese (Doutorado em Estudos da Linguagem) – Departamento de Letras, Pontifícia Universidade Católica do Rio de Janeiro, 2018.

JEFFRIES, E. Preschool children's categorization of speakers by regional accent. *Language Variation and Change*, v. 31, n. 3, p. 329-352, 2019. DOI: <https://doi.org/10.1017/S0954394519000176>

JOHNSON, E. K.; WHITE, K. S. Developmental sociolinguistics: Children's acquisition of language variation. *WIREs Cognitive Science*, v. 11, n. 1, 2020. DOI: <https://doi.org/10.1002/wcs.1515>

JUSCZYK, P. *The Discovery of Spoken Language*. Cambridge: MIT Press, 1997.

KINZLER, K. D.; SHUTTS, K.; SPELKE, E. S. Language-based Social Preferences among Children in South Africa. *Language Learning and Development*, v. 8, n. 3, p. 215-232, 2012. DOI: <https://doi.org/10.1080/15475441.2011.583611>

KINZLER, K. D.; DUPOUX, E.; SPELKE, E. S. The native language of social cognition. *In: PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES*, v. 104, p. 12577-12580, 2007.

KINZLER, K. D.; SHUTTS, K.; DE JESUS, J.; SPELKE, E. S. Accent trumps race in guiding children's social preferences. *Social Cognition*, v. 27, n. 4, p. 623-634, 2009. DOI: <https://doi.org/10.1521/soco.2009.27.4.623>

KOULAGUINA, E.; SHI, R. Rule generalization from inconsistent input in early infancy. *Language Acquisition*, v. 26, n. 4, p. 416-435, 2019. DOI: <https://doi.org/10.1080/10489223.2019.1572148>

LAMPRECHT, R. R. Aquisição da morfologia do plural por crianças bilíngues Português–Alemão. *In: I ENCONTRO DO CELSUL*, 1, 1997, Florianópolis, *Anais [...]*. Florianópolis: Universidade Federal de Santa Catarina, 1997. p.107-117.

LEMLE, M.; NARO, A. J. *Competências básicas do português*. 1977. Relatório final de pesquisa apresentado às instituições patrocinadoras

Fundação Movimento Brasileiro (MOBRAL) e Fundação Ford. Rio de Janeiro, 1977.

LENNENBERG, E. H. (1967). *Biological foundations of language*. Nova Iorque: John Wiley & Sons, 1967.

LIGHTFOOT, D. Language acquisition and language change. *WIREs Cognitive Science*, v. 1, n. 1, p. 677-684, 2010. DOI: <https://doi.org/10.1002/wcs.39>

LIGHTFOOT, D.; WESTERGAARD, M. R. Language acquisition and language change: inter-relationships. *Language and Linguistic Compass*, v. 1, n. 5, p. 396-416, 2007. DOI: <https://doi.org/10.1111/j.1749-818X.2007.00023.x>

MAMPE, B.; FRIEDERICI, A. D.; CHRISTOPHE, A.; WERMKE, K. Newborns' Cry Melody Is Shaped by Their Native Language. *Current Biology*, v. 19, n. 23, p. 1994-1997, 2009. DOI: <https://doi.org/10.1016/j.cub.2009.09.064>

MARTINS, F. S.; DE OLIVEIRA, H. L. C. Variação na concordância nominal de número na fala dos moradores da Costa da Lagoa e Região Central da cidade de Florianópolis (SC). *Working Papers em Linguística*, Florianópolis, v. 16, n. 1, p. 36-58, 2015. DOI: <https://doi.org/10.5007/1984-8420.2015v16n1p36>

MENDES, R. B.; OUSHIRO, L. Variable number agreement in Brazilian Portuguese: an overview. *Language and Linguistics Compass*, v. 9, n. 9, p. 358-368, 2015. DOI: <https://doi.org/10.1111/lnc3.12156>

MILLER, K.L.; RAMOS; M. Acquisition of Sociolinguistic Variation: Acoustic Characteristics of s-lenition in the Speech of Spanish-speaking Children. In: *BUCLD 38*, Boston, 2014. *Online Proceedings Supplement*.

MILLER, K. L.; SCHMITT, C. Variable Input and the Acquisition of Plural Morphology. *Language Acquisition*, v. 19, n. 3, p. 223-261, 2012. DOI: <http://dx.doi.org/10.1080/10489223.2012.685026>

MOLINA, D. S. L. *Aquisição da linguagem e variação linguística: um estudo sobre a flexão verbal variável na aquisição do PB*. 2018. 279f. Tese (Doutorado em Linguística) – Faculdade de Letras, Universidade Federal de Juiz de Fora, 2018.

- MOON, C.; COOPER, R. P.; FIFER, W. P. Two-day-olds prefer their native language. *Infant Behavior and Development*, v. 16, n. 4, p. 495–500, 1993. DOI: [https://doi.org/10.1016/0163-6383\(93\)80007-U](https://doi.org/10.1016/0163-6383(93)80007-U)
- NAZZI, T.; BERTONCINI, J.; MEHLER, J. Language discrimination by newborns: Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, v. 24, n. 3, p. 756–766, 1998. DOI: <https://doi.org/10.1037//0096-1523.24.3.756>
- OUSHIRO, L.; GUY, G. R. The Effect of Salience on Co-variation in Brazilian Portuguese. *University of Pennsylvania Working Papers in Linguistics*, v. 21, n. 2, 2015. Disponível em: <https://repository.upenn.edu/cgi/viewcontent.cgi?article=1881&context=wpwl>
- REIS, M. M. *Aquisição da variação e mudanças na(s) gramática(s) das crianças: um olhar sobre a Concordância Variável no PB*. 2020. 179f. Dissertação (Mestrado em Linguística) – Instituto de Estudos da Linguagem, Universidade Estadual de Campinas, 2020.
- 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.
- SCHERRE, M. M. P. *A regra de concordância de número no sintagma nominal em português*. 1978. 158f. Dissertação (Mestrado em Língua Portuguesa) – Departamento de Letras, Pontifícia Universidade Católica do Rio de Janeiro, 1978.
- 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. DOI: <http://dx.doi.org/10.17851/2237-2083.14.2.539-570>
- SMITH, J.; DURHAM, M.; RICHARDS, H. The social and linguistic in the acquisition of sociolinguistic norms: Caregivers, children, and variation. *Linguistics*, v. 51, n. 2, p. 285-324, 2013. DOI: <https://doi.org/10.1515/ling-2013-0012>
- SMITH, J.; DURHAM, M.; FORTUNE, L. Universal, dialect-specific pathways of acquisition: Caregivers, children and t/d deletion. *Language Variation and Change*, v. 21, n. 1, p. 69–95, 2009. DOI: <https://doi.org/10.1017/S0954394509000039>

SINGLETON, J. L.; NEWPORT, E. L. When learners surpass their models: the acquisition of American Sign Language from inconsistent input. *Cognitive Psychology*, v. 49, n. 4, p. 370-407, 2004. DOI: <https://doi.org/10.1016/j.cogpsych.2004.05.001>

SUMNER, M. The social weight of spoken words. *Trends in Cognitive Sciences*, v. 19, n. 5, p. 238-239, 2015. DOI: <http://dx.doi.org/10.1016/j.tics.2015.03.007>

VAN HEUGTEN, M.; JOHNSON, E. K. Input matters: Multi-accent language exposure affects word form recognition in infancy. *The Journal of the Acoustical Society of America*, v. 142, n.2, p.196-200, 2017. DOI: <https://doi.org/10.1121/1.4997604>

VIEIRA, S. R. (org.). *A concordância verbal em variedades do português: a interface fonética-morfossintaxe*. Rio de Janeiro: Vermelho Marinho, 2015.

WERKER, J. Cross-Language Speech Perception: Developmental Change Does Not Involve Loss. In: GOODMAN, J. C.; NUSBAUM, H. (eds.). *The Development of Speech Perception: The Transition from Speech Sounds to Spoken Words*. Cambridge: MIT Press, 1994. p. 93–120.

WERKER, J.F.; GERVAIN J. Speech perception in infancy. A foundation for language acquisition. In: ZELAZO, P. D. (ed.). *The Oxford Handbook of Developmental Psychology*. Vol. 1: Body and Mind. Oxford: Oxford Univ. Press, 2013. p. 909–25.

WERKER, J. F.; HENSCH, T. K. Critical Periods in Speech Perception: New Directions. *Annual Review of Psychology*, v. 66, n. 1, p. 173-196, 2015. DOI: <https://doi.org/10.1146/annurev-psych-010814-015104>

ZAHARCHUK, H. A.; SHEVLIN, A.; VAN HELL, J. G. Are our brains more prescriptive than our mouths? Experience with dialectal variation in syntax differentially impacts ERPs and behavior. *Brain and Language*, v. 218, p.1-24. 2021. DOI: <https://doi.org/10.1016/j.bandl.2021.104949>