Pontifícia Universidade Católica
do Rio de Janeiro

Ana Paula da Silva Passos Jakubów

Language acquisition based on variable input The case of number agreement in Brazilian Portuguese

Tese de Doutorado

Thesis presented to the Programa de Pós-Graduação em Estudos da Linguagem of PUC-Rio in partial fulfillment of the requeriments for the degree of Doutor em Letras/Estudos da Linguagem.

Advisor: Profª. Letícia Maria Sicuro Corrêa

Rio de Janeiro,
September 2018

Pontifícia Universidade Católica
do Rio de Janeiro

Ana Paula da Silva Passos Jakubów

Language acquisition based on variable input The case of number agreement in Brazilian Portuguese

Thesis presented to the Programa de Pós-Graduação em Estudos da Linguagem of PUC-Rio in partial fulfillment of the requeriments for the degree of Doutor em Letras/Estudos da Linguagem. Approved by the undersigned Examination Commitee.

Profa. Letícia Maria Sicuro Corrêa
Advisor
Departamento de Letras – PUC-Rio

Profa. Cilene Aparecida Nunes Rodrigues
Departamento de Letras – PUC-Rio

Profa. Marina Rosa Ana Augusto
UERJ

Prof. Theodoros Marinis
University of Reading

Profa. Maria Eugenia Lammoglia Duarte
UFRJ

Profa. Monah Winograd
Vice Dean of Graduate Studies
Centro de Teologia e Ciências Humanas – PUC-Rio

Rio de Janeiro, September 24th, 2018

Ana Paula da Silva Passos Jakubów

Graduated in Letras/ Inglês-Literaturas at the State University of Rio de Janeiro in 2011 and obtained her Master's Degree in Linguistics from the State University of Rio de Janeiro in 2014.

Ficha Catalográfica

Jakubów, Ana Paula da Silva Passos

Language acquisition based on variable input : the case of number agreement in Brazilian Portuguese / Ana Paula da Silva Passos Jakubów ; advisor: Letícia Maria Sicuro Corrêa. – 2018.

191 f. : il. color. ; 30 cm

Tese (doutorado)–Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Letras, 2018.

Inclui bibliografia

1. Letras – Teses. 2. Variação. 3. Concordância de número. 4. Aquisição da linguagem. 5. Produção da linguagem. 6. Português brasileiro. I. Corrêa, Letícia Maria Sicuro. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Letras. III. Título.

CDD:400

To the Brazilian children.
Para as crianças brasileiras.

Acknowledgements

To the Brazilian children, specially to the *carioca* children who participated in this study. This thesis made me revisit my story and how, growing up in a low socioeconomic environment in Rio de Janeiro, I became the first academic in the family. I needed data for my thesis, but I learned much more than that. I really hope that this thesis can return to these children. And now my duty is to find a way to accomplish that.

Meu muito obrigada para as crianças brasileiras, especialmente as cariocas, participantes deste estudo. Esta tese fez com que eu relembrasse minha história e como me tornei a primeira professora doutora na família, mesmo crescendo em um ambiente de baixo nível socioeconômico. Eu precisava de dados para a minha tese, mas aprendi muito mais do que metodologia experimental e de coleta de dados. Eu realmente espero que os resultados obtidos nesta tese possam retornar de alguma maneira para essas crianças. Meu dever, a partir de agora, é descobrir como.

To the participant schools, the Secretary of Education of Rio de Janeiro, the Regional Council of Education, the parents. Without you, none of this would be possible. Special thanks to Helane, Luana, Maristela, Silvania, Marcia, Renata e Eni. You were very important to the development of this thesis and for my personal and professional development. Thanks to the patience and to the talks and advices after the sessions. Special thanks to Izabel, who supported this work since the beginning.

Agradeço às escolas participantes, à Secretaria Municipal de Educação, ao Conselho Regional de Educação e aos pais. Sem vocês, este estudo não seria possível. Agradeço especialmente à Helane, Luana, Maristela, Silvania, Marcia, Renata e Eni. Vocês foram importantíssimas para o desenvolvimento desta tese e para o meu desenvolvimento pessoal e profissional. Obrigada pela paciência e pelas conversas e conselhos ao final das sessões. Agradecimento especial para a Izabel por ter apoiado esta pesquisa desde o início.

To my parents for all the support. There will never be enough words to express my gratitude for what you have always done for me and my brothers. Thanks for

listening to me when I first asked to go to school with my brothers, for your patience when you had to listen to me reciting the first school books I memorized, for buying my first chalk board and for never interrupting my classes when I was teaching Portuguese to our dogs. Thanks for accepting my wish to attend English classes when all I wanted was to learn Spice Girls' songs. Most importantly, thanks for teaching me the value of study and for all your love and care.

Agradeço aos meus pais por todo o apoio, amor e cuidado. Jamais encontrarei palavras para expressar minha gratidão por tudo o que vocês têm feito por mim e meus irmãos. Obrigada por me ouvirem quando pedi para ir para a escola com o Bruno e o Rômulo, pela paciência por me ouvirem recitando várias vezes os livros paradidáticos que memorizava, por comprarem meu primeiro quadro de giz e por nunca interromperem minhas aulas de português para nossos cachorros. Obrigada por aceitarem meu pedido para fazer aulas de inglês quando eu só queria cantar as músicas das Spice Girls. E o mais importante de tudo, obrigada por acompanharem cada passo da minha vida escolar e por me ensinarem o valor do estudo.

To my dear husband. I still can't understand how you handled me for all these years! Or maybe I do. Thanks for listening carefully about my research topic, for helping me organizing all the data in countless worksheets in Excel. For your eyes full of pride whenever I tried to explain anything about Linguistics to you. For supporting me in all the ways possible and for taking so good care of me. Thanks for being there, always holding my hand.

To my brothers. You were the first reference I had for school and studies. Seeing you getting ready to go to school or doing your homework made me want that too. Thanks for all the support when I thought I could not finish this work.

To Gabriela. When I started this doctorate, you were born. You still have no idea, but I could not carry on with this thesis if it weren't for your smiles, your first words, for you! Thanks for bringing so much joy and love to the development of this work.

To my supervisor, Letícia. Thank you so much for not letting me give up. I know I started a bit lost in this process but I needed to fix some things before I could be

back on tracks. Thanks for all the patience, for teaching me so much and for kindly pushing me when you knew I could do better. Thanks for all the talks, all the advices and all the learning. For sure, I am taking a bit of you with me in my professional practice.

To the committee. For carefully reading this work and for all the contribution to this research. Theodoros Marinis and Cilene Rodrigues, thank you for your comments during the qualifying exam anad, again in the final exam. Marina Augusto and Maria Eugênia Duarte, thank you for bringing new insight into this work. You all greatly contributed to the development of this research and to my academic career. What a pleasure and an honor to have a discussion with this committee!

To Marina. For introducing Linguistics in my life and for making me a passionate for Language Acquisition. You were there in Linguistics I and now you see me becoming a doctor. My gratitude for all your teaching throughout these years.

To all the professors at the *Departamento de Letras* of PUC-Rio. Special thanks to Erica who were always supporting my academic development and thanks to Cilene, for sharing ideas, laughs, coffees and syntactic trees.

To the *Departamento de Letras* of PUC-Rio for always kindly answering my questions and doubts, especially to Chiquinha.
Meus agradecimentos ao Departamento de Letras da PUC-Rio por gentilmente atender a todas as minhas dúvidas, especialmente à Chiquinha.

To Eunice and Norma. Without your professional support, I would not be able to write these acknowledgements. Thank you very much for the emotional and professional awakening along these years.
Obrigada, Eunice e Norma. Sem o apoio profissional de vocês, eu não poderia estar escrevendo estes agradecimentos. Obrigada pelo auxílio ao longo destes anos para o meu despertar pessoal e profissional.

To my friends Cita, Roberta, Nicole, Mariana, Clara, Dani, Jessica, Noelle, Luiza, João, Raiane, Igor, Monica, Leslie, Sabrina and Eneida. You made this path much smoother. Thanks for sharing a bit of your lives with me, for all the laughter, all the support and for believing in me when I doubted myself. Special thanks to Maria Isabel: even so far away, you were so close during all this process sharing each moment and kindly reviewing my English. And a very special thanks to Odete and Paula and, most recently, baby Olivia: you were of great importance in all this process, since undergraduate years.

To all the ones involved in selecting and acquiring material for PUC library. Throughout these years, I could have access to material of great importance to develop my research.

To CNPq, for the financial support which allowed me to go to different places around the world to present this work, to enrich my academic development and meet amazing people among students and professors.

Abstract

Jakubów, Ana Paula da Silva Passos; Corrêa, Letícia Maria Sicuro (Advisor). **Language acquisition based on variable input: the case of number agreement in Brazilian Portuguese**. Rio de Janeiro, 2018. 192 p. Tese de Doutorado - Departamento de Letras, Pontifícia Universidade Católica do Rio de Janeiro.

This thesis investigates how children growing up in Rio de Janeiro city (Rio) deal with variable input regarding the morphophonological expression of number agreement in Brazilian Portuguese (BP). Plural in BP may be expressed only in the determiner (non-redundant, non-standard variety) or in all agreeing elements (redundant, standard variety). The level of variation is influenced by social factors such as level of education and socioeconomic status (SES): the more educated the speaker is, the more redundancy is morphologically expressed (NARO, 1981; NARO; SCHERRE, 2015; SCHERRE; NARO, 1998). In Rio, these varieties co-exist, given that people from different SES interact on a daily basis. It is proposed that exposure to variable input yields underspecification (ADGER, 2006; ADGER; SMITH, 2010). In the case of number agreement in BP, it is suggested that variation results in underspecification (see ROORYCK, 1994) of morphophonological features in the lexicon pertaining to plural agreement redundancy. We hypothesize that there is a level of bilingualism regulated by SES in number agreement in BP: gradual specification of morphophonological information is dependent on social factors, resulting in a sort of bilingualism. An elicited production task was carried out with preschoolers, from both private (Priv) and public (Pub) schools in Rio's suburban area, in order to verify whether preschoolers exhibit preference for any of the morphophonological expressions of number agreement; type of school is taken as a social variable (ALVES; SOARES; XAVIER, 2014). Results show that preschoolers exhibit a considerable level of variation. However, they differ in terms of production of non-standard varieties, being Pub more likely to produce non-standard forms than Priv. It is argued that during morphophonological encoding in language production, the access to morphemes stored in a *Pool of Variants* (ADGER, 2007) becomes subject to frequency (LEVELT, 1999) and influence of social factors. Obligatory plural marking in the determiner is explained both from the perspective of language knowledge and its representation and from an online

computational model for language production (CORRÊA; AUGUSTO, 2007, 2011), assuming a two-phase DP (PICALLO, 2017) which is phase-based transferred (CHESI, 2007) to morphophonological encoding. The effect of schooling/ literacy in number agreement variation is also verified with Priv and Pub 6th graders. Results show an effect of type of school/ SES in which Priv produces more standard responses than Pub. Additionally, an effect of overall academic performance was obtained among Pub 6th graders from the same school: Pub A (above-average academic performance) are less subject to variation than Pub B (below-average academic performance). Overall, 6th graders' results show that academic performance and SES interact in a gradient *continuum* of number agreement redundancy production: Priv > Pub A > Pub B. Furthermore, a test of assessment of linguistic abilities regarding number agreement, MABILIN II (CORRÊA, 2000), verified whether Pub B 6th graders are able to process grammatical information pertaining to number, given that their performance was similar to Pub preschoolers, despite their age difference. Preschoolers' and 6th graders' results are discussed in terms of bi/multilingual-like contexts: a) BP speakers acquire both varieties simultaneously, as in Bilingual First Language Acquisition (MEISEL, 1994); b) Good academic performance may enhance metalinguistic awareness allowing for code-switching depending on the level of proficiency in one of the varieties (CRAIG; WASHINGTON, 2004); c) good academic performance and higher SES may result in a sort of passive bilingualism or passive bidialectalism regarding the non-standard forms (see CORNIPS, 2014). Finally, preschoolers' and 6th graders' results are discussed in terms of a multilingual *continuum*, characterized as underspecification as variation within one grammar in one extreme to access to different, independent specified grammars in another extreme. More broadly, this thesis is inserted in an interdisciplinary field, combining sociolinguistics, psycholinguistics and theoretical formal linguistics.

Keywords

Language acquisition, variation, number agreement; speech production; Brazilian Portuguese

Resumo

Jakubów, Ana Paula da Silva Passos; Corrêa, Letícia Maria Sicuro (Advisor). **Language acquisition based on variable input: the case of number agreement in Brazilian Portuguese**. Rio de Janeiro, 2018. 192 p. Tese de Doutorado - Departamento de Letras, Pontifícia Universidade Católica do Rio de Janeiro.

Esta tese investiga como as crianças que vivem no Rio de Janeiro (Rio) lidam com *input* variável no que concerne às marcas morfofonológicas de concordância de número em Português Brasileiro (PB). Em PB, o plural pode ser expresso apenas no determinante (D) (não-redundante, variedade não-padrão) ou em todos os elementos que concordam com D (redundante, variedade padrão). O grau de variação é influenciado pelo nível de escolaridade e pelo nível socioeconômico (NSE) do falante: quanto mais escolarizado o falante é, mais marcas de plural são produzidas (NARO, 1981; NARO; SCHERRE, 2015; SCHERRE; NARO, 1998). O Rio é um ambiente de contato entre variedades, já que pessoas de diferentes NSE interagem cotidianamente. Sugere-se que *input* variável resulta em aquisição e representação de traços morfofonológicos pertinentes à expressão da redundância da concordância de número plural subespecificados (*cf.* ROORYCK, 1994) no léxico. Assume-se que a gradual especificação dessas informações depende de fatores sociais, o que pode acarretar uma espécie de bilinguismo. Um experimento de produção induzida por repetição foi conduzido com crianças em idade pré-escolar de escolas pública e particular, no qual possíveis combinações de concordância foram manipuladas a fim de verificar se há diferença entre a performance dos dois grupos, uma vez que tipo de escola é tomado como variável social (ALVES; SOARES; XAVIER, 2014). Os resultados demonstram que ambos os grupos apresentam variação, mas diferem na produção das variedades não-padrão: Pub apresenta respostas não-padrão com mais frequência que Priv. No momento da codificação morfofonológica, na produção da linguagem, o acesso aos morfemas armazenados em uma espécie de *Pool of Variants* (ADGER, 2007), sofreria influência de fatores sociais e de frequência (LEVELT, 1999). A marca de plural obrigatória no determinante é explicada pelo viés sintático e pelo viés do processamento via um modelo de computação online para a produção da linguagem (CORRÊA; AUGUSTO, 2007, 2011), assumindo transferência para as interfaces

em fases (CHESI, 2007) e que o DP em PB possui duas fases (PICALLO, 2017). O efeito de escolaridade na variação da concordância de número é verificado com dados de crianças do 6º ano de escolas pública e particular também no subúrbio do Rio. Os resultados mostram um efeito de tipo de escola/SES, em que Priv produz respostas padrão mais frequentemente que Pub. Além disso, foi verificado um efeito de desempenho escolar com os alunos de 6º ano de uma mesma escola pública: Pub A (bom desempenho acadêmico) é menos sujeito à variação do que Pub B (desempenho escolar regular). De maneira geral, os resultados do 6º ano indicam interação entre NSE e desempenho escolar em um *continuum* gradiente de produção de concordância de número: Priv > Pub A > Pub B. Adicionalmente, um teste de habilidades linguísticas (MABILIN II (CORRÊA, 2000)) buscou verificar habilidades dos alunos de 6º ano em extrair informação de número gramaticalmente relevante na compreensão e na produção do PB pelo grupo Pub B, visto que este grupo do 6º ano apresentou performance similar ao grupo de pré-escolares da escola pública, apesar da diferença de idade entre os dois grupos. Os resultados dos experimentos são discutidos em termos de um contexto bi/multilíngue: a) falantes de PB adquirem as duas variedades simultaneamente, como na aquisição bilíngue simultânea (MEISEL, 1994); b) um bom desempenho escolar pode melhorar a consciência metalinguística do falante, permitindo *code-switching* entre as variedades, dependendo do nível de proficiência em cada uma delas (CRAIG; WASHINGTON, 2004); c) um bom desempenho escolar e NSE alto podem resultar em uma espécie de bilinguismo passivo ou bidialetalismo passivo em relação às formas não-redundantes (see CORNIPS, 2014). Por fim, os resultados são discutidos em termos de um *continuum* multilíngue, caracterizado, em um extremo, como uma gramática subespecificada e, em outro, duas gramáticas separadas, acessadas de maneira independente. Esta tese, portanto, está inserida em um campo interdisciplinar com vistas a conciliar sociolinguística, psicolinguística e teorias linguísticas formais.

Palavras-chave

Variação, concordância de número, aquisição da linguagem, produção da linguagem, português brasileiro

Table of Contents

1
Introduction

Brazilian Portuguese (BP) presents variable plural number agreement in the noun phrase, in third person subject-verb agreement and in predicative structures as attested by variationist studies in sociolinguistics. Variable plural number agreement in Rio de Janeiro city (Rio)[1] has been extensively investigated. The linguistic and social factors affecting it have been substantially characterized (BRANDÃO, 2013, 2015, 2016; BRANDÃO; VIEIRA, 2012; GUY, 1981; NARO, 1981; SCHERRE, 1978, 1991, 1994, SCHERRE; NARO, 2006, 2010, 1998; VIEIRA; BRANDÃO, 2014). Spontaneous production of adults shows that the variation in plural agreement marking is related to speakers' socioeconomic status (SES) and educational level, two social factors that are interconnected in the Brazilian society (NARO, 1981): the higher the speakers' educational level is, the greater the tendency to express plural marking redundantly in the DP (determiner phrase) and as a result of subject-verb agreement, represented in this thesis by TP (tense phrase)[2]. Alternatively, the lower the speakers' educational level is, the greater the tendency to express plural only in the D(eterminer), exhibiting variation of plural marking in agreeing elements. The redundant type of agreement (1) is the standard variety in written and spoken BP[3] and the non-redundant type of agreement figures as the non-standard spoken BP, a socially stigmatized variety (2 - 5):

(1) Redundant (SCHERRE and NARO, 2006):

a-s *coisa-s* *tão* *muito car-a-s,* *né*

[1] Variable number agreement occurs all over Brazil (GUIMARÃES; SILVA, 2016; LUCCHESI, 2006, 2012; MENON, 1995; MONTE, 2015; PEREIRA; ARAÚJO, 2017 (for a review on studies on variation in four of the five regions in Brazil (in portuguese)); SCHERRE; NARO, 1998). Studies on variable number agreement in urban centers are more prominent in Rio de Janeiro (see TARALLO, 1991). Before that, variable number agreement was investigated mostly in rural areas of Brazil (see SCHERRE, 1978).

[2] The verb raises to T in Brazilian Portuguese.

[3] *Variety* is the term adopted here based on Labov (1964, 1997, 2008 [1972]), who establishes that phonological *variants* of a given *variable*, such as the possible sounds (variants) for /-r/ (variable) in the English *variety* spoken in New York. Thus, *variety* covers concepts of language, dialects, sociolects and other types of linguistic manifestation. Bailey's (1973) uses the term *lect* as an umbrella term.

ART.DEF.FEM-PL thing.FEM-PL be.3.PRS.SG very expensive-FEM-PL right
'Things are very expensive, right?'

(2) Non-redundant nominal agreement (SCHERRE and NARO, 2006):

Essa-s estrada-s nova-ø
DEM.FEM-PL road.FEM-PL new.FEM-SG
'These new roads'

(3) Non-redundant nominal and subject-verb agreement (SCHERRE and NARO, 2006):

Esse-s cara-ø hoje só qué-ø
DEM.MASC-PL guy.MASC-SG today only want.3.PRS-SG
curtí mesmo, né.
enjoy-INF indeed right
'Nowadays, these guys just want to have some fun, you see'

(4) Non-redundant subject-verb agreement (ALMEIDA, 2010):

A-s dúvida-s faz parte ...
ART.DEF.FEM-PL doubt.FEM-PL make.3.SG.PRS part
'(The) Doubts are part of the process…'
A-s pessoa-s não toma-ø atitude
ART.DEF.FEM-PL people.FEM-PL not take-SG attitude
'(The) People don't have initiative...'
A-s atitude-s ruin-s afeta-ø...
ART.DEF.FEM-PL attitude.FEM-PL bad-PL affect.3.PRS-SG
'(The) Negative acts affect…'

(5) Non-redundant nominal, subject-verb and predicative agreement (SCHERRE and NARO, 2006):

que a-s coisa-ø tá car-a-ø,
COMP ART.DEF.FEM-PL thing.FEM-SG be.3.PRS.SG expensive-FEM-SG
'because things are expensive'

Based on adults' speech production, we assume that Brazilian children are inevitably exposed to variable input. This thesis aims to characterize (i) the

representation of linguistic knowledge during language acquisition based on exposure to variable input and (ii) the outcome of such representation in terms of language production. The focus of this thesis is on the variation in the morphophonological expression of plural number agreement markings (redundant and non-redundant) in BP.[4] Spoken and written BP exhibit lexical, structural and morphophonological differences; in some cases, depending on socio-cultural-economic factors there is a large gap between both modalities (BAGNO, 2006, 2007; BAGNO; RANGEL, 2005; DUARTE, 2013; DUARTE; SERRA, 2015; KATO, 1990, 2005, 2012). Number agreement is one of the linguistic phenomena that may distinguish these modalities. The extent to which variation in number agreement in BP can be compared to bi/multilingual-like contexts as far as language acquisition is concerned is discussed in this thesis on the basis of a multilingual *continuum*. We assume that BP exhibits variety contact within and across speakers (see WEINREICH, 1970). We suggest that Brazilian children may stand in the middle of a multilingual *continuum*, resembling bilinguals and/or bidialectals exposed to closely-related varieties (CORNIPS, 2014; GROHMANN, 2014; GROHMANN *et al.*, 2016; GROHMANN; LEIVADA, 2012; LEIVADA *et al.*, 2017). Furthermore, the idea of a *continuum* is associated with sociolinguistic studies on Portuguese varieties from former colonies (BRANDÃO; VIEIRA, 2012; GONÇALVES, 2009). Additionally, we propose that children's performance varies not only in terms of SES but also in terms of literacy and schooling, in consonance with the *Socio-Syntax of Development Hypothesis* (GROHMANN and LEIVADA, 2012). We suggest that such multilingual *continuum* has a direct impact on the representation of linguistic knowledge.

This thesis discusses possibilities for the representation of variation in terms of a single grammar with variation (ADGER, 2006; ADGER; SMITH, 2005, 2010; HENRY, 1998; NEVINS; PARROTT, 2010) or multiple grammars (ROEPER, 1999; YANG, 2002) and argues that these configurations vary along the path of linguistic development. We suggest that exposure to variable input entails acquisition of underspecified morphophonological features in the lexicon (ADGER, 2007, 2014, ADGER; SMITH, 2005, 2010, CORRÊA, 2009, 2011,

[4] This thesis focuses on nominal and subject-verb agreement. There are also predicative structures in BP that exhibit variable phenomena in gender and number agreement (see CARVALHO, 2018 and references in it).

2014). Assuming a lexicon with representation of phonological, semantic and φ-features, namely number, person and gender (CHOMSKY, 1995), we propose that morphophonological features are also represented in the lexicon and that these ones, not the syntactic ones, are underspecified in BP, in the sense of having variable, open values (ROORYCK, 1994). Such underspecified morphophonological representation has different outcomes in language production. A model of language production is provided in which the morphophonological underspecification results in competition of insertion of morphemes during morphophonological encoding, being, thus, subject to frequency effects (LEVELT, 1999). The gradual specification of morphophonological features pertaining to each variety, represented as grammars, is dependent on environmental and social factors. Depending on the pressure these factors impose, the result may be different degrees of bi/multilingualism.

We assume that, in language acquisition, the extraction of grammatical information of number, mainly number agreement, is realized on the basis of processing at the interfaces (CORRÊA, 2009, 2014). Distributional properties and morphophonological as well as phonotactic patterns are identified and processed by children at early ages (CORRÊA and NAME, 2003; CORRÊA, AUGUSTO, FERRARI NETO, 2006; CORRÊA, 2009, 2014). Thus, we suggest that morphophonology is instrumental to the mapping of agreement relations. Mapping of morphophonology leads to the representation of morphosyntactic information which, in turn, are represented in the lexicon as formal features, so morphophonology is assumed to work as possible externalization of *Agree* (CHOMSKY, 1995), being subject to parametric differences between languages; differences that children would have to map in the input.

In terms of language production, we assume that morphophonology may work as indexes for boundaries of phase-based transfer to different components in an online computational model (CORRÊA; AUGUSTO, 2007, 2011). Furthermore, we assume that morphophonology is the *locus* of cross-linguistic variation (as well as intra-speaker and cross-speaker variation) as it has been recently suggested (BOECKX, 2011; BOECKX; LEIVADA, 2014).

Therefore, this thesis offers an interdisciplinary investigation which aims to reconcile sociolinguistics, psycholinguistics, language acquisition and theoretical formal linguistics. This investigation is part of recent developments carried out in

LAPAL/PUC-Rio concerning a procedural account for language acquisition in terms of feature specification, such as number and gender (BAGETTI, 2009; BAGETTI; CORRÊA, 2013; CORRÊA; NAME, 2003; FERRARI NETO, 2003, 2008; NAME, 2002). Even though this thesis is concerned with typical language development, an account of the acquisition of variable number agreement may also shed light on possible manifestations in atypical language development, such as SLI (Specific Language Impairment)[5] in contexts in which there is exposure to variable input (GROHMANN and KAMBANAROS, 2016; KAMBANAROS; GROHMANN and MICHAELIDES, 2013). Linguistically impaired children may maintain underspecification for longer time than typically developing children. Additionally, an account of the acquisition and processing of variable number agreement could differentiate speakers exposed to closely-related varieties from any type of disturbance in speech production (WILSON, 2012).

As stated before, the main goal of this thesis is to characterize language acquisition based on variable input, more specifically, based on variation of the morphophonological expression of plural agreement redundancy in BP. The specific aims of this thesis are:

- To analyze possible preferences in relation to the morphophonological expression of number agreement by preschoolers and 6th graders from different SES;
- To investigate the role of literacy in 6th graders' preferences for varieties;
- To compare experimental results obtained in this research with spontaneous speech results in the sociolinguistic literature;
- To provide a description of the access to the representation of variation and its realization in speech production;
- To evaluate whether BP presents multiple grammars (ROEPER, 1999; YANG, 2002) or one grammar with variation (ADGER, 2006; HENRY, 1998) in relation to number agreement redundancy;
- To assess the impact of literacy in linguistic knowledge in terms of grammatical representation and speech production;

[5] Recent work of Bishop (2017) has been proposed a change in the denomination of SLI to DLD (Developmental Language Disorder).

- To situate BP varieties of number agreement redundancy within a multilingual *continuum* ranging from monolingual to bilingual settings;
- To characterize the findings of intra-speaker variation within a broader theoretical account of a minimalist language architecture;
- To present possible scenarios to the difficulties that children with atypical language development may exhibit due to exposure to variable input.

Studies on the processing of variable phenomena in BP are scarce, though this is an extensively explored topic in sociolinguistic approaches to BP. Few psycholinguistic investigations were carried out with adults (AZALIM, 2016; HENRIQUE, 2016; MARCILESE *et al.*, 2015, 2017) and children (MOLINA; MARCILESE; NAME, 2017). However, a proposal for language acquisition for number agreement variation in BP in terms of a multilingual gradient *continuum* (see GROHMANN; KAMBANAROS, 2016) from a formal perspective is unknown.

This study is organized as follows: Chapter 2 presents a characterization of variable number agreement in BP together with an overview of Rio de Janeiro social dynamics as a context of bi/multilingual-like environment for varieties in contact. Chapter 3 presents the design of the elicited production experiment as well as preschoolers' results. Chapter 4 presents a discussion of preschoolers' results by combining formal linguistic theories for language knowledge representation with processing models and provides a characterization for the representation of number agreement variation in preschoolers' grammars as well as their production. In chapter 5, the effect of literacy in number agreement variation in BP is verified in 6th graders production by presenting the results obtained; it also presents the results of a test of linguistic abilities assessment, MABILIN and discusses the impact of SES and academic performance in 6th graders' results. Chapter 6 brings together preschoolers' and 6th graders' results in order to present a characterization of the linguistic knowledge pertaining to number agreement in BP as well as a model of language production that accounts for number agreement variation, considering the influence of social factors such as SES and academic performance. Chapter 7 discusses the overall results in terms of a multilingual *continuum* for number agreement variation in BP. Finally, chapter 8 presents the conclusions and possible further developments of the present thesis.

2
Variation in number agreement in BP

2.1
Coming to terms

The terms *optionality* and *variation*, are used indistinguishably in the literature as the following excerpts show:

a. "The label 'optional' fails to convey any information as to how the elements of the structural description of a rule favor or constrain its operation. Rather, use of this label implies that all such information is foreign to the competence of the native speaker" (CEDERGREN; SANKOFF, 1974, p. 333);

b. "The term optionality is used in a variety of ways in the literature, some quite inconsistent, and it is not always clear in what sense of the term a particular item is optional." (GUY, 2013, p. 1149);

c. "Because all variation seems to enable more than one option, both optionality and variation are terms applied indistinguishably. Actually, there is no apparent consensus in the area [...]"[6] (SCHWINDT, 2014, p. 24);

d. "This variation is also termed 'optionality' or 'variability' and refers to the performance data of the individual L2 speaker."(TSIMPLI, 2006, p. 387);

e. "The existence of optionality is well attested in natural languages. It can be defined as the coexistence of two variants of a given construction with identical Logical Form (LF) representations within the same grammatical system." (PAP, 2000, p. 173);

f. "As our goal is to understand the mechanisms of variation (or, optionality) in a single individual's internal grammar, [...]" (NEVINS; PARROTT, 2010, p. 1138).

[6] Original in Portuguese: "Porque toda a variação parece carregar consigo uma gama de opções, os termos variação e opcionalidade são muitas vezes empregados indistintamente. Não se trata, contudo, de um consenso na literatura da área, [...]" (SCHWINDT, 2014, p. 24)

However, it is possible to grasp some distinctions between these terms. On one hand, in variationist sociolinguistics, *variation* is manifested when a linguistic phenomenon is related to extra-linguistic factors such as social issues, level of education of the speaker, gender, regional area, age and so forth. That is, a linguistic variable is an element that is subject to external pressures (WEINREICH; LABOV; HERZOG, 1968). In this perspective, language is a heterogeneous system and *variation* is a natural consequence of such heterogeneity, *i.e.*, *variation* is part of the system and it may point to a possible change in language (LUCCHESI, 2012: 794). *Variation* is, thus, the object of investigation of sociolinguistics.

On the other hand, the term *optionality* seems to refer to optional rules and it is very often associated to the *Optional Infinitive* proposed by Wexler (1998). Wexler suggests that, around 2 years of age, children produce non-finite as well as finite verbs, which is interpreted as they having *optionality* in grammar. Hyams (2001) argues that *optionality* emerges from the semantic-morphology interface in which meaning is mapped onto morphological information. Parodi and Tsimpli (2005) provide two ways of viewing *optionality*: both options may be equivalent in terms of economy (cost of derivation), named *unconstrained* or *real optionality*, or both options are equivalent in meaning but differ in terms of discourse constraints, named *constrained* or *apparent optionality*. In relation to the *loci* of *optionality*, other accounts are provided. Van Kampen (2004) argues that optionality is present in the lexicon as stored default and/or inflected forms, distinguishable by featural combination. This proposal is in line with Adger's (2006) assumption that variants can be distinguished due to their featural specification in the lexicon. Sorace and Keller (2005) suggest that interpretable features may be underspecified in the interfaces between syntax and other cognitive domains, causing optionality in native and non-native speakers. Richards (2008) indicates that *optionality* may be associated to imperfect mapping between syntactic material and PF.

Based on these concepts, the term *optionality* is more often used in the generative framework, in terms of representation of conflicting options in grammar, whereas *variation* is preferably used in the variationist sociolinguistics framework, associated to social factors. Therefore, *optionality* seems to be related to intra-speaker variation, whereas *variation* refers to systematic variation in a speech community.

From the point of view of intra-speaker grammar representation, exposure to inconsistent input yields optionality in order to cope with variability. From the point of view of language production, optionality in grammar results in an output with variation. In this thesis, we adopt the term *intra-speaker variation* as an alternative to *optional/ optionality* to refer to the representation of grammatical information and the term *cross-speaker variation* will be used for the sociolinguistic type of variation in the speech community level. Whenever further clarification is necessary, a note will inform the specific meaning intended.

Therefore, this thesis aims to integrate both intra-speaker variation, from a linguistic and a psycholinguistic perspective, and cross-speaker variation, from a sociolinguistic perspective, situating different types of variation in a multilingual *continuum*. Thus, the integration between psycholinguistics and sociolinguistics becomes a tool for understanding the dynamics of language representation and processing as far as variable number agreement in BP is concerned (see WEINREICH, 1970 for psychological and socio-cultural aspects of languages in contact).

2.2
Number in BP

Regular plural count nouns in BP are formed with */-s/* addition or by means of allomorphs (*mulher* ('woman')*/ mulheres* ('women'); *papel* ('paper')/ *papéis* ('papers'); *balão* ('balloon')/ *balões* ('balloons')). Regarding verbs in BP, a single morpheme expresses number and person subject-verb agreement. However, spoken BP presents variable paradigms. Duarte (1995) indicates that the loss of functional variability in the pronominal paradigm in BP affected the verbal inflectional paradigm. The nominative pronominal paradigm went through changes in BP: the verbal form that accompanied the 2nd person pronoun *tu* (2.SG/ 'you') has changed to 3rd person verbal form (GALVES, 2001; MENON, 1995).[7] Additionally, *você* (2.SG/ 'you') has become an alternative for 2nd person singular, requiring 3rd person

[7] In the South of Brazil, however, the pronoun *tu* (you.2.SG) is still used with 2nd person verbal inflection: *tu vais* (you.2.SG go.2.SG) *vs. tu vai* (you.2.SG go.3.SG).

verbal inflection. The 2nd person plural pronoun *vós* was replaced by *vocês* (2.PL/ 'you') as mentioned by Duarte (1995) and triggers 3rd person plural verbal inflection.[8] Therefore, both 2nd and 3rd person pronouns trigger 3rd person verbal inflection in singular and plural in most varieties. Also, *a gente* ('we') entered the pronominal paradigm in BP as an alternative to *nós* ('we') (DUARTE, 1995, 2018; LOPES, 2003). Table 1, below, shows three paradigms of subject-verb agreement with pronouns in BP in the present tense proposed by Duarte (1993 [2018]) in a diachronic study based on plays written by *carioca*[9] authors in Rio de Janeiro. The table illustrates the progressive simplification of the person/number inflectional paradigm related to the pronominal paradigm:

Person/ Number	Nominative Pronoun	Paradigm 1 19th century	Paradigm 2 20th century/1	Paradigm 3 20th century/2
1st SG.	*eu*	*cant-o*	*cant-o*	*cant-o*
2nd SG.	*tu*	*cant-a-s*	*cant-a-s*	*cant-a(-s)*
	*você**	-	*cant-a-ø*	*cant-a-ø*
3rd SG.	*ele/ela*	*cant-a-ø*	*cant-a-ø*	*cant-a-ø*
1st PL.	*nós*	*cant-a-mos*	*cant-a-mos*	*cant-a-mos*
	*a gente**	-	*cant-a-ø*	*cant-a-ø*
2nd PL.	*vós*	*cant-a-is*	-	-
	*vocês**	*cant-a-m*	*cant-a-m*	*cant-a(-m)*
3rd PL.	*eles/elas*	*cant-a-m*	*cant-a-m*	*cant-a(-m)*

Table 1: Evolution of pronominal and inflectional paradigms in BP of the verb *cantar* ('to sing') (adapted from DUARTE, 2018, p. 84 (originally published in 1993)) **Você* and *a gente* are 1st person plural forms in the discourse but 3rd person singular in the grammar.

[8] *Vós*, though, is still present in religious texts and prayers, being, thus, a pronoun known by most BP speakers.

[9] *Carioca* is the term used, from the 18th century on, to refer to people who were born in Rio de Janeiro city. It is based on the words from the indigenous language *Tupi-guarani*, spoken by natives of Brazilian territory: *kara'iwa* ('white man') e *oka* ('house'), thus 'white man's house'. Source: https://super.abril.com.br/mundo-estranho/qual-e-a-origem-de-termos-como-carioca-e-gaucho/ accessed on 22/07/2018. The term is used to differentiate those who were born in Rio de Janeiro city from those born in Rio de Janeiro state, but not the capital; these are called *fluminense*.

The first paradigm is similar to EP, correspondent to the privileged, standard form of BP. The pronoun *a gente* ('we' – semantically plural and morphosyntactically singular) is mostly restricted to spoken BP and *vós* ('thy') is restricted to archaic and/or religious texts. Paradigms 2 and 3 also allow *tu ama* (2.SG love.SG/ 'you love') instead of *você ama* (2.SG love.SG/ 'you love') (see footnote 7), which may be socially stigmatized in some regions of Brazil. It is also possible in spoken BP to hear *nós/a gente amamo* (1.PL/1.SG love.PL/ 'we love') with omission of final *-s* in *amamos*; again, it is socially stigmatized. Phonetic adjustments are also possible in the past tense such as in *eles comeram* ([komerɐ̃w̃]) (3.PL ate.3.PST.PL/ 'they ate') *vs. eles comeru* ([komerʊ]) (3PL ate.3.PST.PL/ 'they ate'), in which the final *-m* is omitted and the vowel is replaced.

As observed in Table 1, BP presents great variability in relation to its inflectional system in the verbal paradigm.[10] Gonçalves (2006) contrasted the inflectional paradigm of pronouns and verbal forms in spontaneous speech production of Brazilian and Portuguese children and observed a heterogeneous scenario for BP acquisition, whereas children acquiring EP exhibit relative homogeneity in their production.

Costa and Figueiredo Silva (2006) contrast the behavior of plural agreement markings in BP and in EP from a formal point of view. In relation to DP-internal agreement they observe that, in EP and in one of BP varieties, plural agreement marking is obligatory in all agreeing elements: nouns, determiners, quantifiers, adjectives, possessive pronouns and demonstrative pronouns. However, their analysis seems to point to categorical paradigms in which speakers produce only one of the forms, which does not seem to be the case for most speakers of BP. In other non-standard varieties of BP, an interesting behavior is observed: plural marking appears in D but not in the noun (examples from COSTA; FIGUEIREDO SILVA, 2006, p. 28):

(6) EP/ standard BP:

O-s / este-s / algun-s / un-s

ART.DEF.MAS-PL/ DEM.MASC-PL/ INDF.MASC-PL/ ART.INDF.MASC-PL

[10] The impoverishment observed in the inflectional paradigm in BP has been raising issues related to the status of null subject in BP. Duarte (1995) explores this topic showing that BP is becoming partially pro-drop, with increasing use of lexical subjects, moving further from EP, still considered a pro-drop language.

livro-s *muito bonito-s*
book.MASC-PL very pretty.MASC-PL
'The/these/some/ books very pretty'

(7) *EP/ non-standard BP:
O-s */ este-s* */ algun-s* */ un-s*
ART.DEF.MAS-PL/ DEM.MASC-PL/ INDF.MASC-PL/ ART.INDF.MASC-PL
livro-ø *muito bonito-ø*
book.MASC-SG very pretty.MASC-SG
'The/these/some/ books very pretty'

There is, thus, a difference between BP and EP in DP-internal agreement. Additionally, in non-standard varieties of BP, pre-nominal elements must receive plural marking, whereas plural marking in post-nominal elements may vary:

(8) Quantifiers:
a. ***Todo-s*** ***o-s*** | *aluno-ø/-s* *sai-u/saí-ram*
QF.MASC-PL ART.DEF.MASC-PL student.MASC-SG/PL leave.3.PST-SG/PL
'All the students left'
b. ****Todo-s*** *o-ø* | *aluno-ø/-s* *sai-u/saí-ram*
QF.MASC-PL ART.DEF.MASC-PL student.MASC-SG/PL leave.3.PST-SG/PL
'All the students left'
c. ***O-s*** | *aluno-ø/-s* *todo-ø/*-s (*tudo*[11])
ART.DEF.MASC-PL student.MASC-PL QTF.MASC-PL
sai-u/saí-ram
leave.3.PST-SG/PL
'All the students left'

(9) Adjectives:
a. ***Os*** *único-ø/-s* | *aluno-ø/-s*
ART.DEF.MASC-PL unique.MASC-SG/PL student.MASC-PL
que *saiu/saíram*

[11] Godoy and Cançado (2006) discuss the phenomenon of *tudo* ('the whole', meaning 'all the things') as a quantifier in spoken BP, instead of *todo/todos* (every.SG/PL or all.SG/PL). The authors suggest that *tudo* has lost its restrictive nature as a complex expression, becoming a pure quantifier in spoken BP, as a result of a possible language change.

COMP leave.3.PST-SG/PL

'The only students who left'

In example (8), the quantifier occupies different positions, however, when only the quantifier is marked, but not the determiner, the sentence is ungrammatical (8b); when the quantifier appears in postnominal position, the plural marking is optional (8c). Adjectives are generally postnominal in BP, with few of them allowed to appear pre-nominally (PRIM, 2015); however, even when adjectives appear in pre-nominal position, plural number is optionally expressed (9a). Numerals in BP, cardinals and ordinals, are inflected for gender, but only the latter is inflected for number, thus, ordinals are optionally marked for plural (10): [12]

(10) *A-s* *primeira-ø/-s* *cinco* *pessoa-ø/-s* *que*
ART.FEM-PL first.FEM-SG/PL five person.FEM-SG/PL COMP
cheg-ou/ cheg-aram
arrived.3.PST-SG/PL
'The first five people that arrived'

(11) *(A-s)* *duas* *pessoa-ø/-s* *cheg-ou/ cheg-aram*
(ART.FEM-PL) two.FEM person.FEM-SG/PL arrived.3.PST-SG/PL
'Two people arrived'

Based on these examples few observations can be made: a) demonstratives are in complimentary distribution with articles; both are heads of D; b) possessive pronouns proceed demonstratives and articles, and necessarily receive plural marking; c) prenominal adjectives are optionally marked for plural; d) numerals are always prenominal and are invariant for plural marking but inflected for gender; d) ordinals are variable in number and gender; e) articles can be omitted when co-occurring with cardinals, which function as determiners. Hence, we conclude that

[12] It is debated in the literature whether numerals have a noun-like nature, standing between adjectives and nouns (CORBETT, 1978; HURFORD, 1998; IONIN; MATUSHANSKY, 2006a; MATUSHANSKY; IONIN, 2015; ROTHSTEIN, 2010; VERKUYL, 1981). For Hurford (1998), numerals are simple lexical numerals. Dal Pozzo (2007) argues that, in Finnish, numerals belong to the tripartitioned class divided into quantity nouns, quantity adjectives and quantifiers. In syntactic terms, Ghomeshi (2003) atributes a Cardinality Phrase (CardP) for numerals in Persian in order to differentiate them from Num, also assumed by the author to be the head of grammatical number inflection. Shlonsky (2004) also assumes a Card# phrase. However, she argues that this discussion is an open issue in the literature.

determiners and possessives are the *locus* of the morphophonological expression of plural in BP and cardinals may replace determiners. Though they are not grammatically inflected for number – they convey the semantic idea of number - only for gender.

Additionally, Portuguese exhibits an interesting behavior concerning possessives: in the presence of indefinite articles[13], possessives necessarily appear in postnominal position (12); with definite articles, possessives appear necessarily in prenominal position (13). Notice, however, that the use of the article is obligatory in EP but optional in BP (13b) (examples from CASTRO, 2009)[14]:

(12) Indefinite article and possessives:

a. Fui tomar café com um amigo meu
go.1.PST.SG take.INF coffee with INDF.MASC.SG friend POSS.1.SG
ontem.
yesterday
'I went for a coffee with a friend of mine yesterday'

b. **Fui tomar café com um meu amigo*
go.1.PST.SG take.INF coffee with INDF.MASC.SG POSS.1.SG friend
ontem.
yesterday
'I went for a coffee with my friend yesterday'

(13) Definite article and possessives

[13] Indefinite articles are morphologically inflected for plural number and singular indefinite articles are homophonous to numeral 'one'. Ghomeshi (2003) proposes that indefinite articles present a quantifier nature in Persian.

[14] Castro (2006, 2009) assumes that definite articles exhibit an expletive nature when used with possessive pronouns in Portuguese. According to her, prenominal possessives are weak forms occupying D and postnominal possessives occupy XP and are strong pronouns, which can be prosodically stressed. Floripi (2008) and Magalhães (2011) reinforce the relation between possessives and definiteness proposed by Castro; to Floripi, the definite article in D is used whenever the possessive is not able to express definiteness; to Magalhães, possessives in BP may show variable definiteness features; whenever they do not exhibit definiteness features, the presence of a definite article is necessary. Rinke (2010) shows a different argument to account for some distributional properties of articles and possessive pronouns in European Portuguese (EP). The definite article marks that a nominal phrase is definite. Thus, definite articles accompany possessives in Portuguese because possessives are not inherently definite, although there is a tendency to interpret them as definite; what turns them definite-like is the presence of the definite article. (RINKE, 2010, p. 131).

a. *Fui tomar café com o meu amigo*
go.1.PST.SG take.INF coffee with DEF.MASC.SG POSS.1.SG friend
ontem.
yesterday
'I went for a coffee with my friend yesterday'

b. *Fui tomar café com meu amigo ontem.* (BP/ *EP)
go.1.PST.SG take.INF coffee with POSS.1.SG friend yesterday
'I went for a coffee with my friend yesterday'

c. **Fui tomar café com o amigo meu*
go.1.PST.SG take.INF coffee with DEF.MASC.SG friend POSS.1.SG
ontem.
yesterday
'I went for a coffee with my friend yesterday'

Therefore, the presence of a possessive pronoun turns D optional in BP, hence the optionality of plural markings in D whenever it co-occurs with possessives. In sum, we have the following picture in BP, supposing that elements in the rows precede the elements in the columns:

	Article	**Demonstrative pronoun**	**Possessive pronoun**
Article	Plural	Ungrammatical	Possessive plural
Demonstrative pronoun	Ungrammatical	Plural	Possessive plural
Possessive pronoun	Ungrammatical	Ungrammatical	Plural

Table 2: Matrix for number agreement markers among pre-nominal elements

Numerals are not represented in the matrix above given that they are not inflected for number. Notice that, in isolation, these elements are necessarily marked for plural, however, the possessive pronoun is the only element which can

co-occur with a determiner (both articles and demonstratives), being, then, necessarily marked for plural.

(14) ***A-ø*** ***minha-s*** | *três* *primeira-**s*** *tia-**s***[15]
ART.DEF.FEM-SG POSS.1.FEM-PL three first.FEM-PL aunt.FEM-PL
'My first three aunts'
(SCHERRE, 1978, p. 83)

(15) *Mas* ***o-ø*** ***meu-s*** | *filho-**s*** *num ficar-am*
But ART.DEF.MASC-SG POSS.1.MASC-PL son-PL not stay.3.PST-PL
órfão-s
orphan.MASC-PL
'But my sons were not orphans'
(SCHERRE, 1991, p. 57)

(16) ***O-ø*** ***meus*** | *livro-ø*
ART.DEF.MASC-SG POSS.1.MASC-PL book-SG
'My books'
(COSTA and FIGUEIREDO SILVA, 2006: 29)

(17) *Não* *vi* ***o-ø*** ***seus*** | *(dois)*
Not see.1.PST.SG ART.DEF.MASC-SG POSS.2.MASC-PL (two)
carro-ø *novo-ø*
car.MASC.-SG. new.MASC-SG
'I haven't seen your (two) new cars'
(PEREIRA, 2016, p. 221)

(18) ***O-ø*** ***meu-s*** | *filho-ø*
ART.DEF.MASC-SG POSS.1.MASC-PL son-SG
'My sons'
(CASTRO, 2009, p. 13)

[15] Data with plural marker only in the possessive pronoun are produced by the same speaker in Scherre (1978). It is not clear how frequent this type of number morphological marker is frequent in BP. Oushiro (2015: 408) finds in her sociolinguistic investigation that speakers from São Paulo show that articles are always marked for plural whereas prenominal possessives are 98,4% marked for plural and, in post-nominal position, possessives are 93,9% marked.

(19) *Vou arrumar* ***o-ø*** ***teu-s*** **|** *quarto-s.*
go.1.FUT.SG clean.INF ART.DEF.MASC-SG POSS.2.MASC-PL room.MASC-PL
'I am going to clean your rooms'
(spontaneous production of a 3;8-year-old girl)

(20) *Coloquei café* ***na-ø*** ***tua-s*** **|** *xícara-ø.*
Put.1.PST.SG coffee in.FEM-SG. POSS.2.FEM-PL. cup.FEM-SG
'I poured coffee in your cups'
(spontaneous production of a 3;8-year-old girl)

(21) Pega **o-ø** **teu-s** | brinquedo-ø agora.
take.IMP ART.DEF.MASC-SG POSS.2.MASC-PL toy.MASC-SG now
'Pick up your toys now'
(spontaneous production directed to a 3-year-old boy)

These examples show that grammatical number is not encoded in N in BP, nor it is semantically interpretable in N, but in prenominal elements, as argued by Pereira (2016). It seems plausible to assume, then, that grammatical number in BP is encoded in the layers above N. Danon argues that this seems to be the case in other languages such as Finnish and Wallon: "there seems to be independent empirical evidence that grammatical number is not necessarily a feature that originates from the noun." (DANON, 2011, p. 300). In Finnish nominals, adjectives are singular when they appear between the numeral and the noun, but plural when they appear between the demonstrative and the numeral. In BP, Pereira (2016) argues that:

> With respect to the DP-hierarchy, I have argued that cardinals divide BP DPs into two domains in that phrases preceding NumP are marked with the plural morpheme, while phrases following it are unmarked. Prenominal possessives precede cardinals and hence must be marked; whereas postnominal possessives follows cardinals and hence must be unmarked. (PEREIRA, 2016, p. 226)

In this thesis, we provide an analysis along the lines of Pereira (2016) but with adjustments following a division within the DP in terms of functional and lexical layers, based on Picallo (2017). Such analysis is presented in the experimental results in the chapter regarding the representation of linguistic knowledge the BP speakers have in terms of number agreement within the DP. Following a characterization of how plural number is marked in BP and that it is necessarily encoded in prenominal elements, we present the social factors that influence whether plural number agreement is expressed redundantly or not in BP.

2.2.1
Variable number agreement in BP: studies in variationist sociolinguistics in Rio

Variation in number agreement in BP is largely explored within the variationist framework. According to Duarte and Varejão (2013) variable number agreement is "certainly one of the best described variable phenomena in BP and the only morpho-syntactic feature which is subject to strong social stigma." (DUARTE; VAREJÃO, 2013, p. 109).

Scherre (1978) investigated number agreement within the NP in the speech production of ten adults, from 16 to 23 years old and mapped the linguistic and social factors that influence variation. Morphological processes, most notably phonic saliency, formal situations and higher educational level favor plural marking to be expressed redundantly, while less formal situations and lower educational level favor plural marking only in D. Scherre (1978, p. 75) also verifies a hierarchy of elements and their linear position which favors plural marking within the NP: articles, indefinite pronouns, possessive pronouns, nouns and adjectives, respectively. In relation to number agreement within the linear position of elements in the NP and between different groups, Scherre finds the following results:

Position of elements within the NP		Highly educated speakers	Moderately educated speakers
0	Freq.	98,34 %	97,65%

1	Prob.	0,92	0,98
	Freq.	67,71%	30,46%
	Prob.	0,30	0,45
2	Freq.	65,64%	20,19%
	Prob.	0,26	0,31
3	Freq.	68,25%	2,56%
	Prob.	0,29	0,06
4	Freq.	87,50%	0,00%
	Prob.	0,58	-
Input	Prob.	0,94	0,91

Examples:

a$_0$ *minhas*$_1$ *três*$_2$ *primeiras*$_3$ *tias*$_4$

ART.DEF.FEM.SG POSS.1.FEM.PL three first.FEM.PL aunt.FEM.PL

'My first three aunts'

Table 3: Distribution and Probabilities of occurrences of plural marking within elements in the NP across groups (Source: (adapted from table 9 in SCHERRE, 1978, p. 83)

Observing subject-verb agreement, Scherre and Naro (1998) show that phonic saliency regulates plural marking in subject-verb agreement according to schooling years:

Level of phonic Saliency[16]		Schooling years		
		9-11	1 - 8	illiterates
1	*Eles conhece*	52	30	15

[16] Phonic saliency is understood by the authors as the difference in phonic material between singular and plural forms. In this thesis, we understand phonic saliency as related to the stress of the syllable. Azalim *et al.* (2018) show, by means of experimental results, that syllable stress seems to be more relevant than phonic saliency in terms of quantity of phonic material. The definition of phonic saliency is discussed in Chaves (2014) who argue for a reformulation of the concept, which tends to be interpreted as the contrast between singular and plural forms. It is interesting to highlight that, in the nominal domain, plural marking is not phonic salient, it does not change the stress of the syllable: regular nouns - *telefone-Ø/(s)* ('phone-SG/(PL)); irregular nouns - *papel /papéis* ('paper.SG/ paper.PL'). In few nouns, there is change in pronunciation: *ovo* [ovu] ('egg.SG')/ *ovos* [ɔvʊs] ('egg.PL') (see AZALIM *et al.*, 2018 for an experimental study on phonic saliency and plural in nominals in BP – they find differences regarding agreement as a function of saliency in response time in number agreement of non-salient nouns, rather than in salient plural nouns, the contrary of what is expected in the sociolinguistic literature).

	3.PL know.3.PRS.SG *Eles conhecem* 3.PL know.3.PRS.PL 'They know'			
2	*Eles ganha* 3.PL win.3.PRS.SG *Eles ganham* 3.PL win.3.PRS.PL 'They win'	77	54	30
3	*Eles diz* 3.PL say.3.PRS.SG *Eles dizem* 3.PL say.3.PRS.PL 'They say'	82	58	36
4	*Os filho tá* ART.DEF.MASC.PL son.SG be.3.PRS.SG 'The sons are' *Eles tão* 3.PL be.3.PRS.PL 'They are'	94	71	65
5	*Bateu dois senhores* knock.3.PST.SG two elderly.PL 'Two elderlies knocked' *Eles bateru* 3.PL knock.3.PST.PL 'They knocked'	87	78	73
6	*Veio aqueles cara* DEM.MASC.PL guy.SG 'Those guys came' *Vieram os ladrões* come.3.PRS.PL ART.DEF.MASC.PL thief.PL 'The thieves came'	90	85	80
	Total	81	78	48

Table 4: Distribution (%) of plural marking in subject-verb agreement in relation to phonic saliency level according to years of schooling (Source: SCHERRE; NARO, 1998)

As seen in the table above, phonic saliency distinguishes the groups: illiterates vary from 15% to 80% of plural marking production, moderately educated speakers vary from 30% to 85% and highly educated speakers from 52% to 90% of plural marking production in 3rd person subject-verb agreement. Within the regular verbs

in 1, 2 and 5, the case in 5 seems to be the context in which phonic saliency is more relevant for all the groups.

Additionally, there are phonetic and phonological processes related to plural /-s/ spoken in Rio which distinguishes it from other regions of Brazil, being pronounced as /-ʃ/, rather than /-s/ as in other regions of Brazil. Furthermore, plural /-s/ may have different phonetic realizations depending on the phonetic/phonological context it appears, as shown by Cristófaro-Silva (2012):

(22) a. *moças sentimentais* > *moça*[s]*entimentais*
girl.PL sentimental.PL
'sentimental girls'

b. *moças zuretas* > *moça[z]uretas*
girl.PL crazy.PL
'crazy girls'

c. *moças chatas* > *moça[ʃ]atas*
girl.PL boring.PL
'boring girls'

d. *moças geniosas* > *moça[ʒ]eniosas*
girl.PL bad-tempered.PL
'bad-tempered girls'

Cristófaro-Silva shows that plural nouns ending in /-s/ can be realized as [s,z,ʃ,ʒ]. For instance, in *os bolo* (ART.DEF.MASC.PL cake.SG/ 'the cakes'), the /-s/ in the determiner is realized as [ʒ]. In relation to verbs, especially 3rd person plural, which exhibit high degrees of variation, different phonological aspects are also verified: *cantaram* (sing.3.PST.PL/ 'they sing'), for example, can be pronounced as [kantarʊ] ou [kantarɐ̃w̃] (CHAVES, 2016, p. 184).

Without focusing necessarily on phonic saliency, Vieira and Brandão (2014, p. 101) illustrate the distribution of variation in number agreement across different varieties in Portuguese, considering different structures:

Variety		Type of agreement			
		Nominal	Subject-verb		
			3.PL[17]	1.PL *a gente vai*	1.PL *Nós vamos*
EP	Oeiras	99,9	99,1	82	100
	Cacém	99,9	99,2		
	Funchal	99,7	94,7		
BP	Copacabana	92,4	88,1	99	90
	Nova Iguaçu	91,1	78,2		

Table 5: Distribution (%) of number agreement in different structures in different varieties of Portuguese (Adapted from: Table 8 in VIEIRA and BRANDÃO, 2014, p.101)

Concerning especially Rio de Janeiro state (RJ), Vieira and Brandão (2014) show that there is variation across different SES, given the contrast between a high SES area (Copacabana) and a low SES area in the state (Nova Iguaçu). As the figures in table 5 show, EP presents more consistency in number agreement production even in more isolated areas, such as Funchal in Madeira Island, whereas variation is present in BP in a single state territory. Moreover, when comparing Table 5 with Table 6, presenting the same data analyzed in relation to educational level in BP, it can be attested that the effect of schooling in the morphological marking of plural agreement is more noticeable in the higher SES urban area:

Neighborhood in RJ state	Schooling Level		
	Fundamental (5 to 8 years)	Intermediate (9 – 11 years)	Higher (12 to 15 years)
Copacabana	72,9%	89,0%	97,8%
Nova Iguaçu	72,9%	67,1%	89,9%

Table 6: Third person plural agreement marking in BP (Adapted from: Table 8 in VIEIRA; BAZENGA, 2013, p. 25)

[17] It is not clear whether the authors report 3rd person pronouns or 3rd person NPs. They sum the results into 3rd person plural verbal marking and offer the following example: *todos os homens **usam** camisa* ('all the mean wear shirt') ((example (8) from Vieira & Brandão (2014, p. 92)).

It is important to highlight, given table 6, that SES and educational level are intersected in the Brazilian reality (NARO, 1981) and that social stigma is projected over the non-standard forms of number agreement (BAGNO, 2007; BAGNO; RANGEL, 2005). Vieira (2015) proposes, based on extensive data samples, a *continuum* for number agreement variation in BP related to educational level, specifically in RJ:

+ agreement ◄		► - agreement
RJ Urban area		
(Highly Educated)	**(Moderately Educated)**	**(Illiterate)**
97,8% (Vieira and Bazenga, 2013)	73% (Scherre and Naro,2006) 89% (Vieira and Bazenga, 2013)	48% (Naro,1981)

Table 7: *Continuum* for 3rd person plural agreement in verbs (subject-verb agreement) in BP varieties according to educational level (adapted from VIEIRA, 2015, p. 114 Figure 2)

Sociolinguistic investigations demonstrate that variable number agreement seems to be part of the linguistic knowledge of BP speakers (SCHERRE, 1994), differently from EP. The broad investigation carried out by sociolinguists clearly shows how variable adult speech is as far as number agreement in BP is concerned and it also shows to what extent this variation may be socially influenced. Brandão (2013) proposes that the rule of number agreement in BP is semicategorical for higher SES speakers and variable for lower SES speakers, while in EP, the rule is categorical: "in EP, the input for the agreement is robust, regardless of the education level of the individual. In BP, where the rule is variable, the input is differentiated depending on the socio-economic-cultural class to which the child's family belongs […]" (BRANDÃO, 2013, p. 95).

Therefore, Brazilian children are exposed to variable input regarding number agreement. Basically, BP experiences language contact, in terms of varieties of number agreement (see WEINREICH, 1970). In Rio de Janeiro city (Rio), varieties of number agreement coexist given its social dynamics, presented in the following section.

2.2.2
Varieties in contact in Rio

Rio is one of the municipalities and the capital of Rio de Janeiro State (RJ):

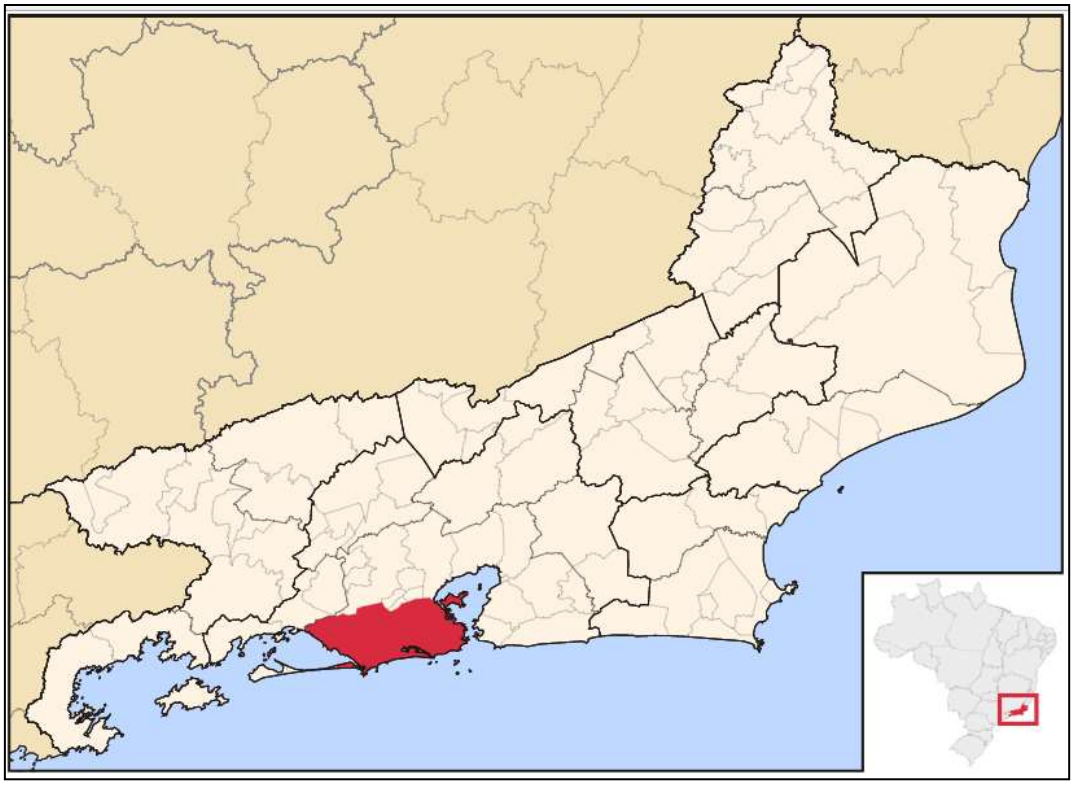


Figure 1: Rio de Janeiro city and its location in Rio de Janeiro State and in Brazil (Source: Raphael Lorenzeto de Abreu - Image: RiodeJaneiro MesoMicroMunicip.svg, own work, CC BY 2.5, https://commons.wikimedia.org/w/index.php?curid=843576)

In order to understand Rio's social background that influences linguistic phenomena, this section brings a very brief history of the city. The social inequality is remnant of the colonial period. Escaping from Napoleonic invasion in Portuguese lands, the Portuguese royal family moved to Brazil in 1808 and settled in Rio[18],

[18] The former capital was Salvador, Bahia. However, the capital transfer to Rio was due to geographical strategic position to build ports and proximity to Minas Gerais where precious metals were mined and transported to Portugal.

which became the capital of the Portuguese Empire. To receive the Portuguese royals, a great plan of Rio's urbanization was implemented in order to turn a precarious colonial settlement into a European-like city. The slavery period brought to Brazil different cultures and languages, transforming the cities into a great space for language contact between Europeans and Africans.[19] Most slaves that arrived in the coast of Brazil, in Rio and Bahia for instance, were from different ethnic groups. The Portuguese deliberately mixed the groups of slaves, speakers of different languages, in order to avoid rebellion and close contact among them (NEGRÃO; VIOTTI, 2012). Lucchesi (2009) proposes that most African slaves in Brazil were speakers of Bantu languages. Additionally, very little is known about the contact with indigenous languages spoken by natives of Brazil. In 1888, with the abolition of slavery in Brazil, lack of job opportunities led former slaves to settle in slope regions.

At the end of the 19th century, after its urban revitalization, Rio became an important center and its population grew, expanding to southern and northern regions of the city. Industries developed towards the north of the city and concentrated a great number of workers (PINHEIRO-FILHO, 2010). Therefore, Rio started experiencing a sociospatial segregation that is felt until nowadays (see OTTONI, 2008).[20] Jacob *et al.* (2015)[21] describe Rio as follows: "[…] side by side with its image of a tropical paradise, there is a city with sharp social contrasts." Nowadays, Rio concentrates around 6 million inhabitants.[22]

Rio is divided in four zones which delimit spatial and social areas (BECKER; COSTA, 2016): south and west coasts concentrate the highest levels of education and income and suburban west and north are characterized by lower income and educational level. However, there is urban mobility in which people from suburban west and north work on a daily basis in the south and west coasts. Besides that, the central financial/business area in Rio concentrates workers from varied SES and

[19] The slaves who arrived in the Southeast of Brazil, the region in which Rio de Janeiro is located, received mostly slaves from Central-Western Africa. Available at: http://www.slavevoyages.org/voyage/search Accessed on: 27/08/2018.

[20] Rio was an important city for the development of the country until 1960's when the capital was transferred to Brasília, central region of Brazil.

[21] In 2015, a complete map of the living conditions in RJ was published by Jacob *et al.* (2015)The e-book is available at: http://www.editora.vrc.puc-rio.br/media/ebook_mapping_living_conditions_in_rio_metropolitan_area.pdf Accessed on 07/08/2017

[22] Information available at: https://cidades.ibge.gov.br/brasil/rj/rio-de-janeiro/panorama Accessed on 04/06/2018.

level of education. Besides urban mobility, social mobility increased among the lower SES families according to data collected between 2013 and 2014 for the National Research per House Sampling (PNAD 2014 - *Pesquisa Nacional por Amostra de Domicílios*)[23] carried out by the Brazilian Institute of Geography and Statistics (IBGE – *Instituto Brasileiro de Geografia e Estatística*).

Also in the 19th century, there was a rapid growth of subnormal agglomerates (*favelas*) as the result of unorganized growth and urbanization by illegal occupation of slope regions and massive migration from other regions of Brazil to Rio and São Paulo, bringing along different linguistic varieties (see Jacob *et al.*, 2015). The 2010's *census* carried out by IBGE reports that Rio concentrates one of the greatest number of subnormal agglomerates in the country, 1.332, just behind São Paulo (2.087).[24] The social contrasts in the city are revealed by income *per capita*.[25] Jacob *et al.* (2015) characterize Rio's metropolitan area income according to their social class:

Continuum of income-class						
Upper class		Middle class		Low class		
Very-high	High	Middle-High	Middle	Middle-low	Low	Very-Low
14,9%		46%		39,1%		

Table 8: Distribution (%) of the population in each income-class forming a *continuum* (Source: adapted from JACOB; HEES; WANIES, 2015, p. 16)

This brief socioeconomic description of Rio reveals a city with massive contrasts of income and level of education. Higher level of education is generally

[23] IBGE, 2014. Supplement of PNAD 2014. *Mobilidade Sócio-ocupacional* 2014 (Socio-ocupational mobility). Available at: https://agenciadenoticias.ibge.gov.br/media/com_mediaibge/arquivos/331e3fd38ba3dce6411dfe876b4c0f76.pdf. Accessed on: 06/02/2018.

[24] IBGE, 2010 Census: 11.4 million Brazilians (6.0%) live in subnormal agglomerates websitehttps://censo2010.ibge.gov.br/en/noticias-censo.html?busca=1andid=1andidnoticia=2057andt=2010-census-11-4-million-brazilians-6-0-live-in-subnormal-agglomeratesandview=noticia Accessed on 22/03/2018.

[25] Brazilian Economic classification criteria (Brazilian Criteria) defines different labels of socioeconomic statuss according to monthly income (in R$): A – 20.272,56; B1 – 8.695,88; B2 – 4.427, 36; C1 – 2.409,01; C2 – 1.446, 24; D - E 639,78. English version available at: www.abep.org (ABEP - *Associação Brasileira de Empresas de Pesquisa*/ 'Brazilian Market Research Association, 2014). Accessed on 16/12/2017.

associated to the use of the privileged form, standard BP, which requires plural number agreement to be expressed redundantly. Non-standard spoken BP is socially stigmatized and related to low educational level; this variety allows for number marking to be expressed only in D. The sharp socioeconomic reality related to differential linguistic use, more specifically number agreement, led to many sociolinguistic researches on this phenomenon.

As presented in the previous section, Naro (1981) highlights the fact that SES influences the production of number agreement. Lower SES leads to a tendency to rely on phonic saliency. Additionally, female and higher SES speakers tend to favor the use of the standard BP variety: "Education is an important component of social class, a concept difficult to measure objectively, especially in Brazilian society, where social frontiers are not so rigidly delimited. In Brazil, most researchers try to measure social class indirectly by using the number of years a speaker has spent in school.". (SCHERRE; NARO, 2014, p. 183).

Rio's social configuration favors the coexistence of standard and non-standard varieties of number agreement. Varieties in contact, or in coexistence, result in a context of *diglossia* as proposed by Fergusson (1959) and may be compared to contexts of bilingualism (FISHMAN, 1967). These contexts are explored in the next section.

2.3
Language contact and multilingualism: *Diglossia*, Bidialectalism, Bilingualism, Monolingualism

Weinreich (1970, p. 1) proposes that language contact emerges whenever two or more languages are used alternatively by the "same persons". Weinreich calls *language* any instance of dialect, languages or varieties. In short, languages may be in contact within a speaker and/or across speakers, both resulting in interference and linguistic variation.

In a multilingual globalized world where social and spatial mobility is intense, especially in large urban centers, defining *monolingualism* is a complex task. Given frequent interaction between people from different backgrounds, language contact

is well-spread in urban centers. Barnes (1954) proposes the idea of networks to account for the dynamics of social interactions: a person in contact with a group of people, relatives or friends, who, in turn, are in contact with other friends, who are also in contact with other groups of people and so on. Mitchell (1974, p. 292) highlights that social networks connect people in terms of communication lines that form a diagram of social network; this configuration builds a channel through which information is spread. Information and communication are achieved through language. Thus, whenever there is social interaction, there is, necessarily language interaction and exchange of linguistic information. Without mentioning the term *network* specifically, Sankoff and Labov (1979, p. 202) state that "every speaker is a member of many nested and intersecting speech community". The concept of social network and social interaction is embedded in their view. Milroy and Milroy (1985, p. 347) present a view in which language is used as a diffusor. They argue that a speaker's linguistic innovation may diffuse into a community by means of contact of the innovator with different groups of people. Along the same lines, Bortoni-Ricardo (2011) proposes that social networks are good indexes for diffusion of linguistic behaviors in the sense that they influence speakers' linguistic choices.

From these social networks, different linguistic features emerge, configurating dialects, languages, sociolects, idiolects, variants, varieties, *diasystems* (see BAILEY, 1973; WEINREICH; LABOV; HERZOG, 1968). Chambers and Trudgill (1980) propose that all speakers are bidialectals to some extent. Hence, every speaker is exposed to language variation to some degree. This issue is also raised by Roeper (1999) who suggests that all speakers exhibit a degree of discrete bilingualism (see also DĄBROWSKA, 2012; HENRY, 1998; MUFWENE, 2008; SMITH; DURHAM; RICHARDS, 2013 for intra-speaker variation and different grammatical representations within the same speaker).

The constant interaction among speakers in a communicative-social network creates many different linguistic contexts of multilingualism. Thus, how to distinguish contexts of language contact? The literature has been mainly explored differences between bilingualism, bidialectalism and monolingualism.

Baratz (1969) finds interference of non-standard variety in standard variety in a sentence repetition task carried out with Afro-American children. Kennedy (1973) suggests that bidialectals must control two varieties, pretty much in the way

bilinguals do with different languages. Akere (1980) finds code-switching between local varieties: Ijebo dialects and Lagos urban dialects in Yoruba speech communities. Bokamba (1988) also finds code-mixing in Bantu languages, between Kiswahili-English and Lingala-French varieties. More recently, there has been a discussion on how to develop code-switching abilities among dialect speakers of General English and Afro-American English (AAE) (CRAIG; WASHINGTON, 2004; WHEELER; SWORDS, 2004)

In terms of language processing, recent studies have been investigating the differences and similarities among different contexts of multilingualism. Papapavlou and Philli (2009) compared 43 monolinguals, 40 bilinguals and 40 bidialectal children from kindergarten, aged 4-5, and first graders, aged 5-7. In phonological awareness tasks, bilinguals and bidialectals performed better than monolinguals. They concluded that bidialectal children may be dealing with varieties in the same level that bilinguals must cope with two different languages given the similar performance presented by both groups. Sumner and Samuel (2009) suggested that experience with dialects may allow storage for separate phonological representation for each variety: "one can have a dialect in one domain (word recognition) and not another (production); or at one time (immediate processing) but not another (long-term processing)" (SUMNER; SAMUEL, 2009, p. 499). Thus, they claimed that dialects should be characterized in terms of representation, production and perception. In a lexical retrieval task, Kambanaros *et al.* (2013) found no distinct performance between bidialectal children speakers of Cypriot Greek/ Standard Modern Greek and Standard Modern Greek monolinguals. Cornips (2014) reveals that bidialectals from Limburg, the Netherlands, are more aware of grammatical gender in elicited production tasks than bilinguals; there is a significant development in the use of neuter grammatical gender in Dutch for 4-6 year-olds bidialectals, whereas bilinguals do not show development in this aspect at this age range, meaning that bidialectals differ from bilinguals in terms of linguistic development. In relation to bidialectals, Cornips highlights that there is not a clear-cut distinction between monolinguals and bidialectals; rather, another category emerges "in-between" those two, the passive bidialectals. Additionally, Grohmann and Leivada (2012) highlight the role of schooling in bidialectalism, what they term *Socio-Syntax of Development Hypothesis*, in which competing social motivations influence the morphosyntactic

development in language acquisition (GROHMANN; LEIVADA, 2012). Regarding executive control in multilingual contexts, there seems to be advantages regardless of how closely-related the varieties are. This was suggested by Antoniou *et al.* (2016) who compared monolinguals, bilinguals, and bidialectals performance on executive control tasks.

Evidences suggesting that bidialectalism seems to differ from monolingualism and bilingualism poses a challenge to multiple grammars representation. Bidialectalism boundaries seem to be blurred (CORNIPS, 2014; LEIVADA *et al.*, 2017). Language proximity, as argued by Grohmann (2014; see also TSIMPLI, 2014), plays a major role, in that, the more similar the systems are, the harder to distinguish them (see also CORNIPS, 2014). On the contrary, Weinreich presented the language proximity problem as: "The greater the differences between the systems, the greater, *i.e.*, the more numerous the mutually exclusive forms and patterns in each, the greater is the learning problem and the potential area of interference" (WEINREICH, 1970, p. 1).[26]

Much has been debated on how such multilingual contexts may differ. In bilingualism, grammars are understood to be separately and independently represented (*cf.* BIALYSTOK, 2001; DE HOUWER, 2009 for a review). For bidialectalism, Grohmann and Leivada (2012) and Leivada *et al.* (2017) propose a dialect design: 1) blurred boundaries between variants; 2) dialect *continua* with intermediate possibilities and 3) social pressures influencing speakers' perception towards their own production. Regardless of the fact that two different languages or varieties are easier or harder to identify, there seems to be a *continuum* of language settings that may range from monolinguals to bilinguals with bidialectals standing in the middle of this *continuum*. Grohmann (2014) proposes a gradience in multilingualism in terms of *comparative linguality*.

The present thesis intends to apply the notion of gradience multilingualism to number agreement variation in BP, arguing for different multilingual-like settings within BP. In relation to point 3 in the design proposed by Leivada *et al.* (2017), psycholinguistic investigations have been showing how social pressures influence

[26] It is still an empirical issue to be investigated: whether perception of linguistic patterns is actually facilitated when varieties are structurally similar. This sort of psycholinguistic investigation could shed light on theoretical accounts of linguistic representation. Thanks to prof. Jason Rothman for productive talks and comments on this matter in between coffee-breaks and presentations during the *Child Language Symposium* 2018 in Reading, UK.

speakers' perception and production. Some of these investigations are discussed in the next section.

2.4
Social influence on language use and processing: possible constraints on variable input

Variation has long been neglected by Psycholinguistics according to Boland *et al* (2016). Recently, more studies on processing of variable phenomena have been carried out considering the speakers' social characteristics such as SES and level of education.

Squires (2014) investigates the influence of social experience on adults' processing of agreement in English. A self-paced reading task was conducted with 112 Americans divided into groups: social class (high and low), gender (male and female) and ethnic group (white, Afro-American). Results show shorter latency in the standard variety, intermediate latency for the non-standard one and longer latency for unusual sentences. Squires suggests: "My results seem to mirror the well-known 'achievement gap' present in American education: students with lower socioeconomic status and who are in racial minority groups tend to have lower reading/literacy abilities […]" (SQUIRES, 2014, p. 186). However, her results also point to the fact that non-standard forms have longer latency even for low-social class groups: "It could be that my measure of class does not get at real class/status differences, or that even people who have experience/ knowledge with the form are surprised to see it in print, in a university setting." (SQUIRES, 2014, p. 187). Hence, lower-class participants are also sensitive to non-standard forms. Notice, however, that all participants in Squire's experiment were undergraduates; they all somehow had contact with standard forms in university academic settings.

A similar study with reaction time and the contrast between standard and non-standard forms was carried out in BP. We highlight that studies on variable number agreement in BP from a psycholinguistic perspective are quite recent. Costa (2013) investigated variable agreement with the so-called meteorological verbs, such as 'to rain', with university students, thus, high educational level speakers. This type of

verb is inflected in 3rd person singular as a structure without subject. It has been recently observed that, when occurring in relative clauses, meteorological verbs may vary in terms of number agreement, standard (22) and non-standard (23) (2013)):

(23) Standard*:*

Sei que há lugares em que chove-ø muito
know.1.PRS;SG. COMP have.SG place.PL in COMP rain-SG very
'I know that there are places in which it rains a lot'

(24) Non-standard*:*

...sei que há lugares que chove-m muito...
know.1.PRS;SG. COMP have.SG place.PL COMP rain-PL very
'I know that there are places in which it rains a lot'

Costa's results show that speakers with high educational level exhibit variable verb agreement in non-standard relative clauses. However, the occurrences of plural meteorological verbs in non-standard relative clauses are more frequent than the occurrences of plural meteorological verbs in standard relative clauses. Costa argues that the first case may be considered as the nuclear grammar of the speaker, a sort of innovative grammar that allows plural meteorological verbs, while plural meteorological verbs in standard relative clauses constitute agreement errors in speech production. Therefore, according to Costa, speakers with high educational level are *diglossic* and are able to manage both varieties.

Contrasting the performance of adult speakers' from different educational levels, Azalim (2016) investigates the processing of number agreement within the DP by two groups: university students and adults who are completing the basic level of schooling (*Educação de Jovens e Adultos* (EJA) - Youth and Adult Literacy). In a comprehension experiment, using self-paced listening technique, she verified whether university students accept the non-standard, non-redundant variety of number agreement in BP. Results reveal that latency was longer in non-redundant agreement conditions. Azalim suggests that participants with high level of schooling take longer time to process non-redundant agreement, which is interpreted as heavier processing cost of non-standard BP variety by this group of

speakers. However, plural marking only in D proved to be sufficient to trigger plural reading for both groups.

In an experiment of elicited production by repetition, Azalim exposed participants to stimuli manipulated according to phonic saliency in the DP and redundancy in the number agreement (*o-ø bone-ø / os bonés* (/-s/) ('the cap'); *o varal/ os varais* (/-is/) ('the clothesline')). The latency between the presentation of the auditory stimulus and the onset of the repetition was the dependent variable. Latencies were longer in non-redundant agreement in both groups. Phonic saliency was relevant only for EJA students, who presented shorter latency in this condition. Actually, the stress in the syllable receiving plural morpheme proved to be relevant for this group, rather than phonic saliency in terms of the contrast of phonic material between singular and plural forms. In relation to target repetition, both groups presented similar performance with higher target responses for the redundant agreement condition. Phonic saliency was not relevant for neither group in the production induced by repetition task, probably due to the nature of the experimental situation, which increased speech register awareness of the participants, revealing preference for the standard, redundant agreement. Analyzing the type of alternative responses in relation to the target stimuli during repetition, Azalim verified an overall tendency for the university students to correct the sentences by adding plural to the nouns, while *EJA* students seemed to fluctuate between the number markings.

Studying variable number agreement in BP in the verbal domain, Henrique (2016) also compared the performance between *EJA* students and university students and observed an effect of linear distance in subject-verb agreement in both groups when applying a self-paced reading experiment using the *maze task* technique[27]: long linear distance favors non-redundant agreement. Observing the results of an elicited production task based on repetition, university students corrected the stimuli to redundant agreement, matching the number of the subject and the verb; few instances of incomplete answers were detected. Though, for *EJA* students, the greatest number of non-target responses were actually incomplete or not repeated, especially in the long linear distance condition. Henrique suggested that the performance difference attested between both groups may be related to

[27] Participants have to choose one option between the two segments presented simultaneously in the computer screen in order to continue the task.

difficulty in keeping information in the working memory and/or difference of years of schooling.

Overall, Azalim's and Henrique's investigations point to the fact that the processing of agreement in BP, in production and in comprehension, seem to differ; in consonance with results obtained by Almeida (2016) for number and gender agreement in production errors.

Marcilese *et al.* (2017) compared standard and non-standard varieties of subject-verb agreement in BP. The longest latency was verified for the plural subject + singular verb condition which may reflect the cost of processing to account for values that do not belong to the participants' grammars (MARCILESE *et al.*, 2017, p. 27). An alternative explanation provided for the long latency in this condition has to do with the computational implementation during processing. Marcilese *et al.* outlined that these results converge with other results in the literature suggesting longer latency for non-redundant agreement.

As presented, Brazilian adult speakers, regardless of their schooling level, accept redundant and non-redundant agreement as two possible functional forms in BP (AZALIM, 2016; COSTA, 2013; HENRIQUE, 2016; MARCILESE *et al.*, 2015, 2017). However, non-redundant agreement seems to be costly, also regardless of speakers' schooling level. In the DP domain, the distinction between redundant and non-redundant number agreement seems to be sharper (AZALIM, 2016) than the distinction between redundant and non-redundant subject-verb number agreement - differences in this type of number agreement are subtler (HENRIQUE, 2016). Additionally, speakers with longer years of schooling tend to produce utterances corresponding to the standard variety of number agreement (AZALIM, 2016; COSTA, 2013).

As presented in the aforementioned investigations from processing perspectives, many language-internal and language-external factors seem to influence language processing based on variable input. Studies on language acquisition based on variable input in BP are very scarce despite innumerous investigations on variation in number agreement by sociolinguists. Even in sociolinguistics, studies on the acquisition of variable input are quite rare (see GOMES *et al.*, 2011). Molina (in prep.)[28] finds, based on spontaneous speech

[28] I am grateful to Daniele Molina for providing her findings for the ongoing development of her doctoral thesis. These data were orally presented during ABRALIN (*Associação Brasileira de*

production data of 4 children from Juiz de Fora, Minas Gerais, aged 3;4 to 6;2, the presence of variation in singular and plural morphophonological markings, although their highly educated mothers produce mostly redundant 3rd person plural agreement. These children exhibit fluctuation in terms of number agreement in the verb both when the subject DP is redundant and when the DP is non-redundant. Adults, in turn, exhibit clear preference for the standard agreement when the subject DP is redundant, but such clear difference is not detected when the DP is non-redundant. In Molina's results, children and adults present the same performance for verb agreement with non-redundant subject DP, but with redundant subject DP, adults exhibit clear preference for redundant agreement, while, for children, both number marking seem to compete. In an experiment concerning 3rd person subject-verb agreement comprehension carried out with 5-to-6-year-olds, Molina *et al.* (2017) found that exposure to variable input does not affect children's comprehension, though children from higher SES perform better than children from lower SES. Additionally, Molina investigates the spontaneous speech production of 5-to-6-year-olds from urban and rural areas of Juiz de Fora, in Minas Gerais and finds that the production of non-redundant 3rd person number agreement in BP is more frequent for children in the rural area than for children in the urban area.

Roberts (2008), based on Labov (2008 [1972]), proposes that social influence on linguistic variation and language use are more prominent in adolescence. In the initial path of language acquisition, children are mapping probabilities and measuring frequencies in the input. Such input analysis has a direct effect in children's production. In the case of BP, highly-educated adults tend to produce number agreement in all agreeing elements of the structure at higher rates than low-educated adults, who exhibit 48% of number marking, at chance level. Studies on language acquisition based on variable input show that children are sensitive to variation in the input, producing the rates of variation they are exposed to (GOMES *et al.*, 2011; SMITH; DURHAM; RICHARDS, 2013).

Therefore, in this thesis we assume that children acquiring BP are inevitably exposed to variable input concerning number agreement (BRANDÃO, 2013). We also assume the coexistence of two varieties of number agreement in BP, redundant and non-redundant, a *diglossic*, bidialectal or even bilingual environment for

Linguística) International congress in March, 2017 at Universidade Federal Fluminense (UFF) in Niterói, Rio de Janeiro, Brazil.

language acquisition. Exposure to variable input yields direct impact on speakers' linguistic knowledge and, consequently, on their production. Studies on speech perception and psycholinguistics show that children acquiring Portuguese are sensitive to number marking only in D. However, Brazilian children, as opposed to Portuguese children, are less sensitive to variation in plural marking in the noun when the determiner is plural (CASTRO *et al.*, 2009; CASTRO; FERRARI-NETO, 2007; CORRÊA; AUGUSTO; FERRARI-NETO, 2006; FERRARI NETO, 2008). In the next section, the procedural model for language acquisition, on which this thesis is based, is presented.

2.5
A procedural model for language acquisition: the case of agreement

Studies on speech perception demonstrate that children acquire language by means of processing of prosodic cues and phonotactic and morphophonological properties in the input (BAGETTI; CORRÊA, 2009; CHRISTOPHE *et al.*, 1997; JUSCZYK, 1997; LI; SHI; HUA, 2010; SHI, 2014; SHI; MARQUIS; GAUTHIER, 2006; SHI; MELANÇON, 2010). In short, children's task when acquiring a language is to map the input they receive and process grammatical information in a way that it can be represented and stored in memory.

Corrêa (2002, 2009, 2014) combines phonological bootstrapping and minimalist assumptions (CHOMSKY, 1995) in order to account for grammar identification during language acquisition. Initially, stress patterns detected by children and their position in the speech string would enable the representation of a feature related to word order. Afterwards, children must identify morphophonological properties of functional elements in the speech string. These closed-class elements have short morphophonological forms and appear in limited number, thus, they are very frequent in speech (determiners, affixes, connectives, some prepositions and classifiers). In this account, the perception of phonotactic patterns in PF would enable the establishment of groups of words grammatically mapped as functional categories. In short, phonetic/phonological information

processed in PF would be mapped onto morphosyntactic representations in grammar, that is, morphosyntactic features. Corrêa indicates that, after closed-class items are identified and represented in children's grammar as features of functional categories, a minimally specified lexicon emerges, that is, a lexicon with underspecified formal features. The presence of features that minimally distinguish closed and open-class items – functional and lexical categories, respectively - would be sufficient to trigger the computational system and the *Merge* operation. Minimally specified functional category features would enable the projection of hierarchical structures:

> The mapping of open class elements onto entities would contribute to the definition of a nominal domain headed by an underspecified determiner – an underspecified determiner phrase (uDP). An underspecified tense domain (uTP) would be delimited in relation to the possible presence of functional elements such as auxiliaries and verbal affixes. The conflation of cues pertaining to the boundaries of hierarchically related prosodic units would (in the presence of other cues) enable the maximal unit for parsing to be delimited (an underspecified complementizer phrase [uCP]). In sum, parsing would start with a minimal number of formal features long before the grammar is fully identified, i.e., before the full specification of the formal features of functional elements and their properties. Once its operation starts, C_{HL} would become instrumental to further analysis of the linguistic stimuli and to the full identification of the grammar. (CORRÊA, 2009, p. 40)

Under this view, the initialization of the computational system occurs by the end of children's first year of life. Hence, in this approach, grammar identification is achieved by processing linguistic material and representing formal features.

The identification of formal features is crucial for language acquisition. These features have two counterparts: interpretable and uninterpretable. Formal interpretable features represent semantic, (morpho)phonological and morphosyntactic information. According to Chomsky (1995), phonological features are represented, but we argue, in this thesis, for morphophonological features representations for reasons that will be clarified further. Uninterpretable formal features assume an important and exclusive role for computation; φ-features person, gender and number are, to some extent, independent from the semantic contents they convey for the purposes of syntactic computation. Their interpretable counterpart, which is directly related to conceptual/intentional content, are not

visible by the computational system, but rather to LF interface; uninterpretable φ-features[29] may have a visible outcome in PF such as inflections of agreement which indicate linguistic relations between lexical items.

In relation to agreement, Corrêa proposes that morphophonology indicates to children that there is an agreement relation between elements. For instance, in terms of gender, the processing of morphophonological marking related to gender in the determiner in BP leads to the postulation of relations of agreement between elements D and N (CORRÊA and NAME, (2003). Corrêa, Augusto and Ferrari (2006) propose the same type of analysis for number: identification of systematic variation in suffixes of the determiner in BP indicating number leads to the identification of number agreement relations between D and N. In this sense, morphophonological patterns are paramount to the acquisition of agreement relations in BP.

The left-to-right parsing together with the early perception of word order would allow Brazilian children to conceive the first underspecified DP in the structure as the subject of the sentence.[30] Morphophonological variation in the verbal item would lead children to assume that the underspecified DP and a verb to the left of the DP in the speech string must enter an agreement relation (feature valuation in terms of *Agree*). This procedural account would ensure that the person feature detected in the verb is assigned to the subject underspecified DP. In this thesis, we extend Corrêa's analysis of formal feature (under)specification to morphophonological feature (under)specification in the lexicon in order to account for variation in number agreement in BP. Therefore, we argue that processing of the morphophonology of agreement is instrumental to the identification of φ-features.

Hence, acquisition of agreement involves processing at the interfaces: initially, children must recognize morphophonological patterns pertaining to agreement (number, gender, person – pronouns and verbs - in the case of BP); after such recognition, children need to identify the value of the feature of the controlling element (the determiner, in the case of number and gender in BP). In the case of

[29] As assumed in the previous section, uninterpretable but valued features are able to receive content in PF.

[30] Prosody or stress may indicate that this DP is actually a topic, rather than a subject for instance. These distinctions would be established in terms of information processing in PF.

variable agreement, children should also be sensitive to variation in patterns (markings expressed in D but optional in N). In sum, processing of information at the interfaces, starting from morphophonology, leads to the representation of formal and morphosyntactic features in the lexicon.

In terms of agreement, Vieira (2006) and Gomes *et al.* (2011), observed gradual increase in number inflections as children grow older. Simioni (2007) detected production of non-redundant agreement until 3;4 years of age by a child who lives with intermediately educated parents. Differently, data from other two children show an opposite pattern: predominance of redundant agreement in the DP until 2;3 and decrease of non-redundant agreement by 3;0 years old. (examples (1-3) from SIMIONI, 2007, p. 125–127):

(27) *trouxe* ***a-s*** ***faca-ø*** *pra cortar* (2;3 years old)
bring.1.PST.SG ART.DEF.FEM-PL knife-SG to cut.INF
'I brought the knives to cut'

(28) *eu boto* ***a-s*** ***coisa-s*** *aqui* (3;2 years old)
1SG put.1.PRS.SG ART.DEF.FEM-PL thing.FEM-PL here.
'I put the things here'

In terms of acquisition of verbal inflections, Wuerges (2014, p. 118) suggests it starts from the less marked/ specified forms and, around 2 years of age, children start to manipulate grammatical information of agreement. Regarding acquisition of regular and irregular verb inflections, Ferrari Neto and Lima (2015) detected a gap in comprehension and production: at the same age, around 4 years old, children produce overregularized verb forms, but are also able to judge these forms as ungrammatical. The following path of acquisition is, then, provided:

Approximate age	Morphological development of determiners, nouns and verbs in BP
9 - 12 months old	• sensibility to phonic distinctions in suffixes (Bagetti, 2009)
14-18 months old	• sensibility to morphophonological distinctions in verbal affixes (Bagetti, 2009)
1;7 years old	• production of 1st and 3rd person singular in present and past tenses verbal forms (Wuerges, 2014)
1;10 years old	• use of null Ds and numerals to express plural (Simioni, 2007);

2;0 years old	• number interpretability in D (Ferrari Neto, 2003); • acceptability of redundant and non-redundant agreement in DP (Ferrari Neto, 2008)
2;4 years old	• production of 1st and 3rd person plural in present tense verbal forms (Wuerges, 2014 and Martins, 2007)
2;1 – 2;6 years old	• production of definiteness in D (Lopes, 2004, 2006; Simioni, 2007); • production of plural morphological marking in D (Lopes, 2004, 2006; Simioni, 2007)
2;7 years old	• use of numerals to express plural with nouns decreases (Simioni, 2007); • definiteness in D increases (Simioni, 2007)
2;9 years old	• 1st and 3rd person plural in past tense verbal forms (Wuerges, 2014)
3;0 years old	• phonic saliency as relevant for agreement marking in third person verbal forms (Vieira, 2006)
4;0 years old	• production of plural marking effective in third person verbal forms (Vieira, 2006); • tendency to judge overregularized verbal forms as ungrammatical (Ferrari Neto and Lima, 2015); • remaining production of verbal overregularized forms (Ferrari Neto and Lima, 2015)
5;0 years old	• verbal plural marking patterning adults' performance (Vieira, 2006; Ferrari Neto and Lima, 2015); • countable reading assigned to singular bare Ns (Ferrari Neto and Lima, 2015)

Table 9: Acquisition of morphophonological agreement inflections in BP

Therefore, around 5 years old, children are already immersed in the morphological paradigm of number in BP (FERRARI NETO; LIMA, 2015).

Following the procedural account proposed by Corrêa (2009, 2014), children must perceive, based on frequency, phonotactic and morphophonological properties of the speech, and that plural is obligatorily expressed in pre-nominal elements in BP. In this procedural account, children perceive that the root is constant with or without the presence of a sublexical item; for instance, an article may appear with or without a number morpheme. Thus, children take the presence of a root and alterations in affixes as relevant grammatical information that should be encoded as a formal feature. Such linguistic knowledge seems to be early acquired by Brazilian children, around 2 years of age. As soon as such information is mapped, a semantic interpretation is assigned, via reference. In order to verify how children naturally acquire number agreement in the nominal and verbal domains based on

variable input in BP, an elicited production experiment was carried out with preschoolers from a suburban area in Rio. All the participants had never been formally exposed to standard/written BP and did not know how to read or write at the time the experiment was carried out.

3

Elicited production experiment: variation in number agreement redundancy prior to schooling

This chapter describes an elicited production task conducted with preschoolers from both public (Pub) and private (Priv) schools in a suburban area in Rio. The Brazilian educational system includes the public and the private sectors. The public sector is divided by administrative dependency such as municipal, state and federal government levels. According to *Censo de Educação Básica de 2013*[31] (*Census* of Elementary Education) (INEP, 2017), 82,8% of Brazilian students are enrolled in public schools, being 46,4% of this percentage enrolled in municipal schools[32], while private schools concentrate 17,2% of students. Alves, Soares and Xavier (2014) analyzed socioeconomic questionnaires of students all over Brazil and revealed the following scenario: municipal schools concentrate students from the lowest to mid-low SES and private schools concentrate students from mid-high to high SES.[33,34,35] We contrast municipal public schools and private schools in order to verify whether SES influences on the production of number agreement

[31] A Brazilian national measure for basic education (from pre-school to high-school).

[32] 35,8% in state schools and 0,6% in federal schools.

[33] Cross-group studies regarding type of school is frequently carried out in Brazil given its social inequality. Simões (2006) conducted a study with children from the south of Brazil, in Porto Alegre, comparing data before and after the literacy stage and emphasizes that literacy pedagogical practices and social environment may impact the production of number agreement. Mariano (2013) investigated the use of number agreement within the NP in essays of elementary schools students in RJ and observed a *continuum* going from the area with lowest SES showing less use of the standard variant, to the area with highest SES, in the opposite extreme, using more standard agreement.

[34] The comparison of SES and how it impacts academic performance is also carried out in Brazilian high schools as the *INSE* (*Indíce Socioeconômico* - Socioeconomic Level Index). MEC (*Ministério da Educação* - the Brazilian Ministry of Education) and the *Instituto Nacional de Estudos e Pesquisas Educacionais (Inep)* (National Institute of Educational Studies and Reseraches) reveal that in the *ENEM* (*Exame Nacional do Ensino Médio* - national exam for high school students to attend university) socioeconomic level and teachers' education is determinant for students to succeed in the exam: the higher their socioeconomic level is, the higher their chance to succeed (Source: http://portal.mec.gov.br/ultimas-noticias/418-enem-946573306/21521-nivel-social-do-aluno-e-formacao-do-professor-sao-determinantes. Accessed on 29/06/2018). However, independent research carried out by *Folha de São Paulo*, an important news vehicle in Brazil, shows that schools that concentrate students from high to very high socioeconomic status do not achieve the expected grades in *ENEM*, that is, there is variation even within the same students' profile (Source: https://www1.folha.uol.com.br/educacao/2018/06/1-a-cada-3-escolas-de-ricos-tem-nota-no-enem-abaixo-do-esperado.shtml. Accessed on 29/06/2018).

[35] *Indicador de Nível Socio-econômico das Escolas (INSE)* (Socioeconomic status index for schools) (Source: http://download.inep.gov.br/informacoes_estatisticas/indicadores_educacionais/2011_2013/nivel_socioeconomico/nota_tecnica_indicador_nivel_socioeconomico.pdf Accessed on: 30/04/2018)

redundancy by preschoolers, since SES and level of education is a social variable that influences variation in number agreement (NARO, 1981; SCHERRE; NARO, 1998, 2014 and many others).

When entering preschool, children have a naturally acquired grammatical representation that may be influenced by schooling (*cf.* PIRES; ROTHMAN, 2009). The objective of this experiment was to verify what grammar representation children bring to school, considering SES influence. The aim of this experiment was to verify:

(i) whether preschoolers acquiring BP exhibit preference for any of the morphophonological expressions of number agreement (standard/redundant or non-standard/non-redundant);
(ii) if the social variable defined as the type of school plays any role on the use of number agreement
    a. in which direction influence would take place;

The linguistic profile of the children's families from each school was assumed to be the following:[36]

a) Priv: parents are mainly speakers of the standard variety (passive bilinguals for the non-standard one) and/or simultaneous bilinguals, having the standard variety as the dominant one;
b) Pub: parents are mainly speakers of the non-standard variety (passive bilinguals for the standard one).

It is proposed that children's production reflect their grammar representation when repeating given stimuli (KOMEILI; MARSHALL, 2013; SLOBIN; WELSH, 1967; THEODOROU; KAMBANAROS; GROHMANN, 2017). It is assumed that preschoolers acquiring BP exposed to variable input regarding number agreement redundancy have already acquired obligatory number marking in the D around 2 years old and that they have already mapped agreement relations. It is hypothesized, thus, that fluctuation in preschoolers' production depends on SES. The prevision is

[36] Assumptions had to be made given the difficulty in accessing families' data and the difficulty to receive even authorization for participation back from families.

that the two groups present different performance according to SES and the frequency of one or the other variety would be influenced by external factors, as attested by sociolinguistic investigations: Priv preschoolers would show preference for the standard variety (redundant agreement) more frequently than Pub preschoolers.

3.1
Experiment Design

Number agreement in BP can be expressed in different ways as shown in examples (1-5) and repeated here:

(29) Redundant (SCHERRE and NARO, 2006):

a-s coisa-s tão muito car-a-s, né

ART.DEF.FEM-PL thing.FEM-PL be.3.PRS.SG very expensive-FEM-PL right

'Things are very expensive, right?'

(30) Non-redundant nominal agreement (SCHERRE and NARO, 2006):

Essa-s estrada-s nova-ø

DEM.FEM-PL road.FEM-PL new.FEM-SG

'These new roads'

(31) Non-redundant nominal and subject-verb agreement (SCHERRE and NARO, 2006):

Esse-s cara-ø hoje só qué-ø

DEM.MASC-PL guy.MASC-SG today only want.3.PRS-SG

curtí mesmo, né.

enjoy-INF indeed right

'Nowadays, these guys just want to have some fun, you see'

(32) Non-redundant subject-verb agreement (ALMEIDA, 2010):

A-s dúvida-s faz parte ...

ART.DEF.FEM-PL doubt.FEM-PL make.3.SG.PRS part

‘(The) Doubts are part of the process…’

A-s *pessoa-s* *não toma-ø atitude*

ART.DEF.FEM-PL people.FEM-PL not take-SG attitude

‘(The) People don’t have initiative...’

A-s *atitude-s* *ruin-s* *afeta-ø...*

ART.DEF.FEM-PL attitude.FEM-PL bad-PL affect.3.PRS-SG

‘(The) Negative acts affect…’

(33) Non-redundant nominal, subject-verb and predicative agreement (SCHERRE and NARO, 2006):

que *a-s* *coisa-ø* *tá* *car-a-ø,*

COMP ART.DEF.FEM-PL thing.FEM-SG be.3.PRS.SG expensive-FEM-SG

‘because things are expensive’

For this reason, *Redundancy in Number Agreement* was taken as the independent variable of an elicited production experiment in which preschoolers were asked to retell what a robot had told them. The robot produced sentences that varied according to possible morphophonological expressions of redundancy in number agreement. Therefore, the independent variables were: 1. *Redundancy in Plural Number Agreement* (within-subject; four levels) and 2. *Type of School* (between-subject; 2 levels), giving rise to 4 conditions:

Conditions	Examples of Stimuli
1. DP red TP red	*O-**s*** *cachorro-**s*** *encontr-**aram*** *o leão* ART.DEF.MASC-PL dog.MASC-PL find-3.PST.PL the lion ‘The dogs found the lion’
2. DP red TP n-red	*O-**s*** *cachorro-**s*** *encontr-**ou*** *o leão* ART.DEF.MASC-PL dog.MASC-PL find-3.PST.SG the lion ‘The dogs found the lion’
3. DP n-red TP red	*O-**s*** *cachorro-ø encontr-**aram*** *o leão* ART.DEF.MASC-PL dog.MASC-SG found-3.PST.PL the lion ‘The dogs found the lion’
4. DP n-red TP red	*O-**s*** *cachorro-ø encontr-**ou*** *o leão* ART.DEF.MASC-PL dog.MASC-SG found-3.PST.SG the lion ‘The dogs found the lion’

Table 10: Experimental conditions – Experiment 1

The retelling task (not a repetition one) was intended to induce the spontaneous reliance on the variety most easily accessible to the children at the moment of utterance production. In order to manipulate the social variable in terms of SES/educational level, preschoolers from two types of school participated in the task: a private school (Priv) (middle-high SES/educational level) and a public school (Pub) (middle-low SES/educational level).

The dependent variable 1 was the number of matching responses corresponding to the form of the stimuli provided by the robot. Given the linguistic profile of the schools' population, an effect of *Type of School* is expected with more responses matching to the stimuli in Pub group.

The dependent variable 2 was the number of responses matching the standard form (condition 1). Given the linguistic profile of the schools' population, an effect of *Type of School* is expected with more responses matching the standard form in Priv group.

The research questions are:

(i) Do preschoolers' acquiring BP exhibit preference for any of the morphophonological expressions of number agreement?
(ii) Does *Type of school* play any role on the production of number agreement by preschoolers?
(iii) What sort of non-standard possibility prevails in preschoolers' responses?
(iv) Is there a tendency for preschoolers to provide alternative redundant responses to the non-standard forms?

3.2
Method

3.2.1
Participants

In total, 59 preschoolers from a private school (Priv) participated in the task. However, one child did not complete the task, 3 children did not provide responses

according to what was asked and 2 children were too young (3;1 and 3;2 years old) in relation to the others. Therefore, data from 53 Priv preschoolers (26 girls and 27 boys) were considered. The age range was from 3;8 to 6;11 years old.

In relation to Pub a total of 30 preschoolers (11 girls and 19 boys) aged from 3;5 to 6;3 years old participated in the task. From this total, 4 children did not complete the task, one child could not perform the task and 3 children did not provide responses according to what was asked. Children in Pub had much more difficulty in performing the task and following the instructions. The stimuli had to be repeated quite often and still with all the repetition, some children could not perform the task. We believe that this difficulty may be attributed to social factors that economic and social inequality brings related to a varied range of factors, from nutrition to access to different motor, linguistic and socio-cultural stimuli and activities during childhood. Given the high level of difficulties to perform the experimental task with Pub, we decided, based on the 22 children who could perform the task, to pair both groups of preschoolers in terms of number in the population sample and age, forming the following groups:

Type of School	n	girls	boys	Age range	Mean age
Priv	22	13	9	3;8 - 6;3	5;5
Pub	22	11	11		

Table 11: Distribution of preschoolers per groups as a function of Type of School

Preschoolers had not yet been formally introduced to reading and writing at the time the experiment was conducted; thus, they were not formally exposed to standard/written BP, but to the standard spoken variety.

3.2.2 Material

In the linguistic stimuli, the type of subject was controlled: masculine and animated subjects, thus, determiners, the articles, were always definite, masculine, plural. The type of verb and its complement were controlled: transitive and action verbs in the simple past tense and animacy and gender of the complement were balanced. Stimuli have 8 – 11 syllables. Each child was exposed to 16 stimuli in all 4 conditions (4 stimuli per condition). Different lists contained randomized items following the Latin square design were built: 4 lists with different sentences in all 4 conditions (see the Appendix). Sentences were recorded and edited in the software *Audacity* for pitch alterations in order for the recorded voice to sound friendlier to preschoolers. *Bob*, an animated character[37] was introduced. Bob was designed to be a character from a distant place who spoke in an unusual way, therefore its robot-or-alien-like characteristics:

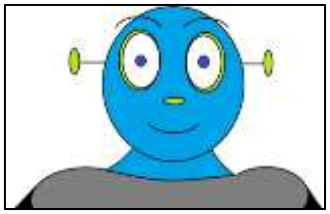


Figure 2: Bob, the animated robot presented during the task

All the procedure was performed in a *DELL Inspiron*, i3 processor laptop. The responses were recorded in a *Sony* MP3 player-recorder and/or in a Panasonic RR-US511 recorder. Whenever the school did not allow recording, participants' responses were written down. If children changed words, saying *cachorros* ('dog') instead of *gatos* ('cat') as provided in the stimuli, the response would be considered as matching the stimuli, since the plural /-s/ was produced.

3.2.3
Procedure

[37] With the use of the software *Crazy Talk 2* by *Reallusion*.

Participants performed the task individually in an available isolated room in the school. The experimenter informed the participant that Bob is a robot who lives far away and sometimes speaks in an unusual way. Children were supposed to pay attention to what Bob had said and tell the experimenter, who pretended she could not listen to anything, what had happened, according to the protocol below:

> *O Bob mora em um lugar muito longe e ele fala um pouco diferente da gente. Lá onde ele mora, acontece um monte de coisas. Ele vai te contar o que aconteceu, mas eu não vou ouvir. O Bob disse que você tem que ouvir e me contar, tá bom? Então, eu vou tampar os ouvidos e você me conta o que aconteceu lá onde o Bob mora*.
> 'Bob lives far away and he speaks in an unusual way. There, where he lives, many things happen. He is going to tell you what happened but I am not going to listen. Bob said you have to listen and tell me, ok? So, I will cover my ears and, then, you tell what happened where Bob lives.'

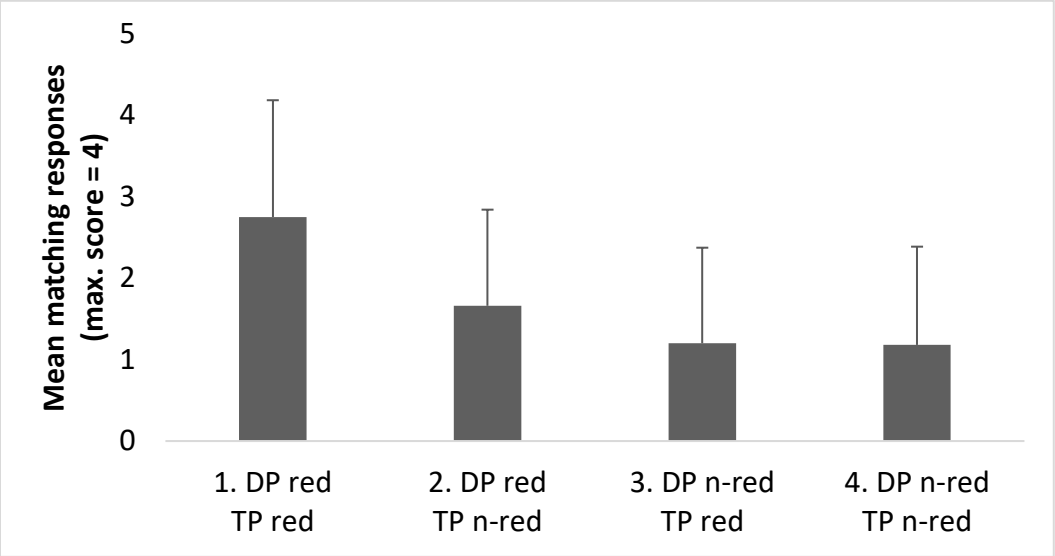
At the end of the task, the experimenter asked participants whether they thought Bob really spoke in an unusual way. All the sentences produced were recorded and transcribed and, eventually, notes were also taken. The task lasted around 5 minutes.

3.3
Results

3.3.1
Dependent variable 1: Matching responses

The data was analyzed by means of a 2-way ANOVA (design 4 X 2), in which *Redundancy in Number Agreement* is a within-subject factor (4 levels), and *Type of School* is a between-subject factor (2 levels). There was a significant main effect of

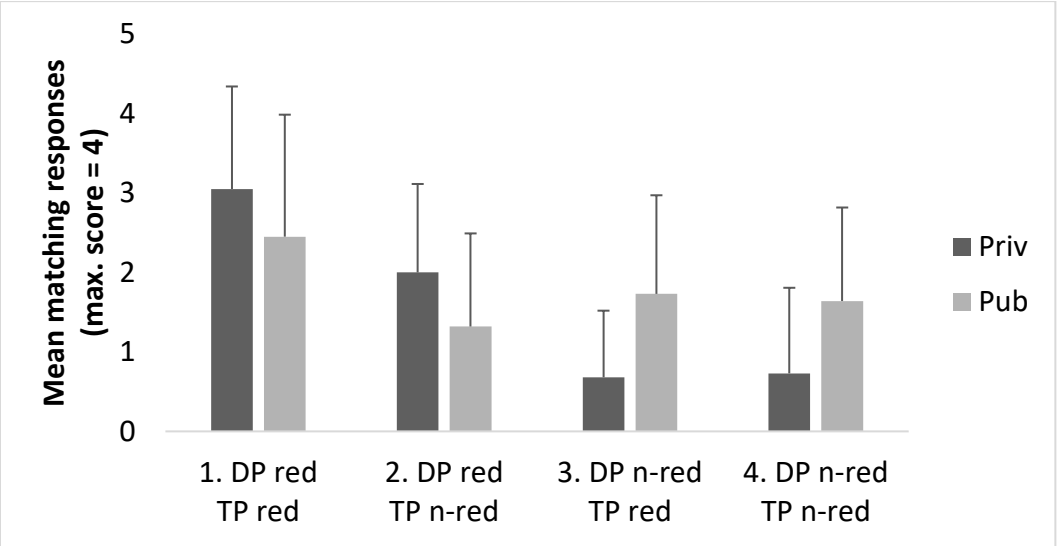
Redundancy (F (1, 42) = 21,262; p < .0001; C1: *M* = 2,75, *SD* = 1,5; C2: *M* = 1,66, *SD* = 1,2; C3: *M* = 1,2, *SD* = 1,2; C4: *M* = 1,18, *SD* = 1,2):



Graph 1: Mean matching responses as a function of *Redundancy in Number Agreement* (Error bars = *SD*)

There was no significant effect of *Type of School* for the matching responses provided (F (1, 42) = 0,845; p = .363; Priv: *M* = 1,62, *SD* = 1,1; Pub: *M* = 1,79; *SD* = 1,3), meaning that both groups repeat the stimuli at similar rates. Overall, the analysis of matching responses show that preschoolers tend to reproduce more standard forms. However, they exhibit different performance in relation to non-standard forms, mainly conditions 3 (DP red/ TP n-red) and 4 (DP n-red/ TP n-red): Pub tends to reproduce the non-standard stimuli, whereas Priv correct the stimuli.

There was a significant effect of interaction between *Redundancy* and *Type of School* (F (1, 42) = 7,745; p = .008; Priv C1: *M* = 3,05, *SD* = 1,3; C2: *M* = 2, *SD* = 1,1; C3: *M* = 0,68, *SD* = 0,8; C4: *M* = 0,73, *SD* = 1; Pub C1: *M* = 2,45, *SD* = 1,5; C2: *M* = 1,32, *SD* = 1,2; C3: *M* = 1,73, *SD* = 1,2; C4: *M* = 1,64, *SD* = 1,2):

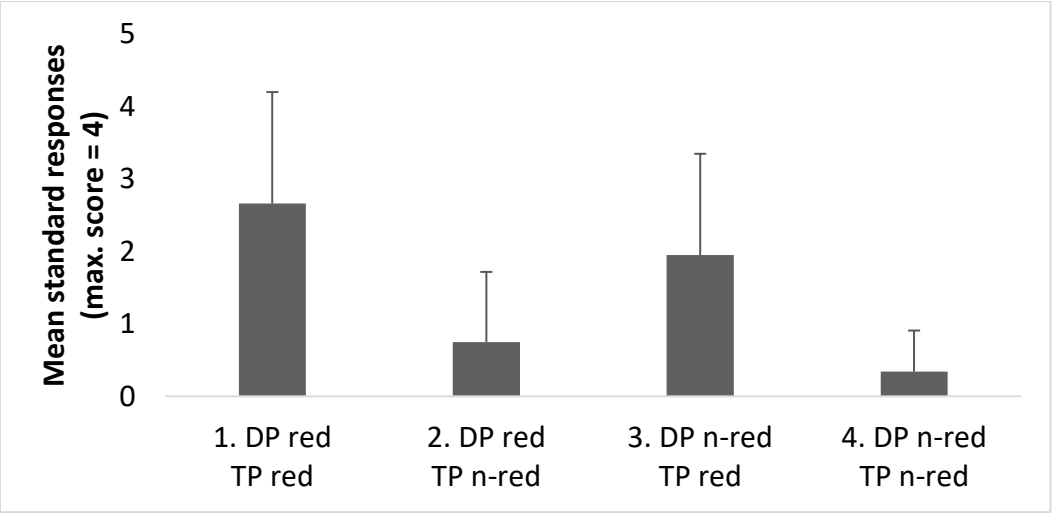


Graph 2: Mean matching responses as a function of *Redundancy in Number Agreement* per *Type of School* (Error bars = *SD*)

As the graph above shows, groups differ in relation to standard matching responses provided: Priv provided more standard responses than Pub. Additionally, the groups differ in relation to the non-standard matching responses provided: Pub provides more non-standard responses than Priv.

3.3.2
Dependent variable 2: Standard responses

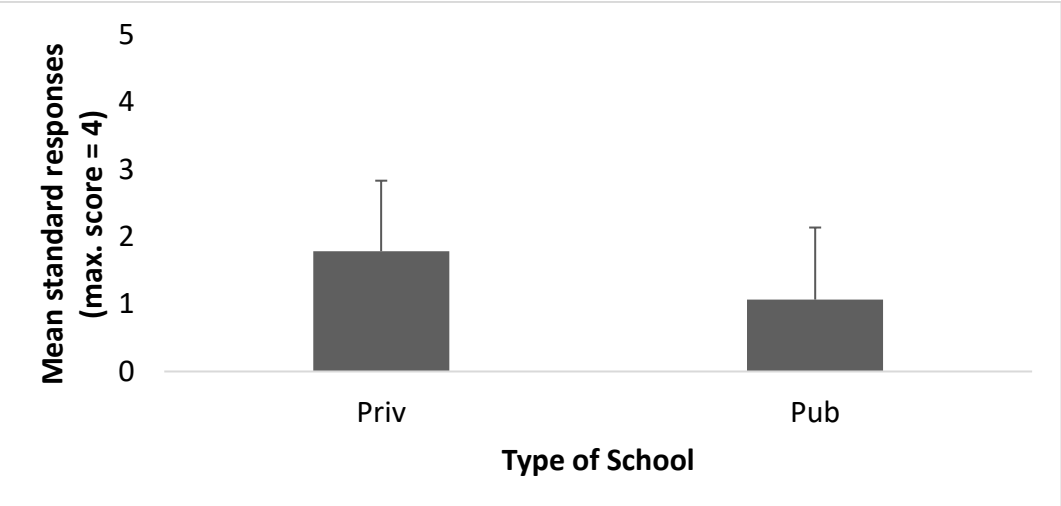
As for the second dependent variable, standard responses provided, the data was also analyzed by means of a 2-way ANOVA (design 4 X 2), in which *Redundancy* is a within-subject factor (4 levels), and *Type of School* is a between-subject factor (2 levels). There was a significant main effect of *Redundancy* (F (1, 42) = 101, 151; p < .0001; C1: *M* = 2,66, *SD* = 1,5; C2: *M* = 0,75, *SD* = 1,0; C3: *M* = 1,95, *SD* = 1,4; C4: *M* = 0,34, *SD* = 0,6):



Graph 3: Mean standard responses as a function of *Redundancy in Number Agreement* (Error bars = *SD*)

Graph 3 shows that preschoolers produce standard responses when the stimuli is standard or when the stimuli presents plural verb (condition 3: DP n-red/ TP red) (*cf.* Graph 3). This analysis suggests that plural morphology in the verb seems to be instrumental for the production of standard number agreement in BP.

Unlike the first dependent variable, there was a main effect of *Type of School* (F (1, 42) = 8,255; p = .006; Priv: *M* = 1,79, *SD* = 1; Pub: *M* = 1,07, *SD* = 1,1):



Graph 4: Mean standard responses as a function of *Type of School* (Error bars = *SD*)

Notice, however, that Priv did not even reach the half of the maximum score. Though Priv preschoolers produce more standard responses than Pub, the means are very low. This result may point to the fact that the input is extremely variable and preschoolers are reproducing such variation.

There was no effect of interaction between *Redundancy* and *Type of School* (F (1, 42) = 0,988; p = .326; Priv C1: *M* = 3,14, *SD* = 1,3; C2: *M* = 0,95, *SD* = 1,1; C3: *M* = 2,55, *SD* = 1,2; C4: *M* = 0,5, *SD* = 0,6; Pub: C1: *M* = 2,18, *SD*: 1,7; C2: *M* = 0,55, *SD* = 0,8; C3: *M* = 1,36, *SD* = 1,3; C4: *M* = 0,18, *SD* = 0,5). In the next section, we provide an analysis of the type of responses provided when preschoolers were not reproducing the stimuli.

3.3.3
Types of responses across conditions

In this section, a breakdown of the responses is provided. In Table 12, such distribution is shown:

	Types of responses														
	DP red TP red		DP red TP n-red		DP n-red TP red		DP n-red TP n-red		Singular		No D		Ungrammatical		Total valid responses
Type of School	Priv	Pub	Priv	Pub	Priv	Pub	Priv	Pub	Priv	Pub	Priv	Pub	Priv	Pub	
Condition															
1	**38**	**31**	39	31	7	10	0	3	2	2	2	1	1	4	176
2	11	9	**25**	**16**	1	1	4	13	7	10	1	1	1	1	176
3	31	19	2	1	**9**	**22**	1	1	0	0	2	2	5	7	176
4	6	3	14	13	2	2	**9**	**20**	16	11	2	0	1	0	176

Condition 1 = DP red/ TP red
Condition 2 = DP red/ TP n-red
Condition 3 = DP n-red/ TP red
Condition 4 = DP n-red/ TP n-red

Matching responses in bold.

Table 12: Distribution (%) of responses per condition and type of school - Preschoolers

As shown in the table above, in condition 1 (DP red/ TP red), children from both types of school tend to reproduce the stimuli or provide responses equivalent to condition 2 (DP red/ TP n-red). When presented with stimuli in condition 2 (DP red/ TP n-red), these preschoolers also tend to reproduce the stimuli. However, there is a difference between the groups regarding the alternative responses provided in condition 2: Priv provides standard responses, equivalent to condition 1 (DP red/ TP red), whereas Pub provides non-standard responses, equivalent to condition 4 (DP n-red/ TP n-red). Singular responses were also provided by both groups in condition 2. Concerning condition 3 (DP n-red/ TP red), Priv clearly provides standard responses (DP red/ TP red), whereas Pub shows competition between reproducing the stimuli and standard responses. As for condition 4 (DP n-red/ TP n-red), Priv tends to provide singular forms or responses equivalent to condition 2 (DP red/ TP n-red) respectively, whereas Pub tends to reproduce the stimuli, followed by responses equivalent to condition 2 (DP red/ TP n-red) and singular forms. Notice, however, that, regardless of groups, the most frequent type of non-standard response is DP red/ TP n-red, followed by DP n-red/ TP red for Pub, and Priv prefers the standard form for Priv. The non-standard responses preferred by Priv follows an order: DP red/ TP n-red > DP n-red/ TP n-red > DP n-red/ TP red. For Pub, the non-standard responses scale is: DP n-red/ TP n-red > DP red/ TP n-red > DP n-red/ TP red.

Altogether, the ANOVA results and the breakdown of type of responses show that Pub is sensitive to the stimuli provided, reproducing them more often than Priv, regardless of condition. Alternatively, Priv seems to be more sensitive to the contrast of redundancy among the conditions, trying to correct the stimuli, mainly in non-standard conditions. However, such perception does not seem to be a product of schooling or metalinguistic awareness, given that, when asked if the robot spoke in an unusual way, none of the preschoolers reported lack of plural markings. The participants actually mentioned that the robot said funny things about animals and that robots cannot have animals. Hence, it seems that children were mostly driven by semantic and pragmatic content than by morphosyntactic information. This makes the results more reliable in terms of showing the grammatical representation of number agreement they have acquired without being aware of which of the forms is the standard/not stigmatized one. In the next chapter, we propose an account for such variability for number agreement in preschoolers' production.

4
Accounting for variation in preschoolers' production: underspecification of morphophonological features in the lexicon

We assume that identifying a grammar consists in representing linguistic information in the lexicon. The literature suggests different perspectives on the accommodation of variation in grammar: a) one grammar with variation and; b) two coexistent grammars. The latter is founded on diachronic investigations: Weinreich, Labov and Herzog (1968) proposed the idea of (sub)systems, a single *diasystem* in which speakers can alternate between different functionally distinct options. Kroch (1994) proposed the existence of *doublets* which consist of two forms competing for the same meaning. In this view, whenever there are two competing options, each constitutes a different grammar.

Following Kroch's reasoning, Roeper (1999) proposed the Universal Bilingualism theory in which every speaker is, essentially, bilingual: "It [bilingualism] is present whenever two properties exist in a language that are not statable within a single grammar [...] any consistent grammar cannot have contradictory rules. Therefore, one must postulate two grammars, even if they differ only in a single rule." (ROEPER, 1999, p. 169–170). In this framework, grammars acquired by children must be compatible with the input they receive. Initially, children are provided with a Minimal Default Grammar (MDG), a set of default representations based on economy principles. These representations do not necessarily correspond to the target grammar and MDG can be eventually abandoned due to social factors. Roeper (2016) argues that Multiple Grammars theory allows speakers to establish a common rule within different grammars.

Yang (2002) also assumes different grammars in competition. He proposes a model that combines statistical learning and UG: "Upon the presentation of an input datum *s*, the child: a) selects a grammar G_i with probability p_i; b) analyzes *S* with G_i; c) if successful, reward G_i by increasing p_i, otherwise, punish G_i by decreasing p_i" (YANG, 2002). The rewarded grammar(s) receive(s) more weight than the punished ones. There may be, then, a moment of competition between grammars,

meaning that these grammars achieved the same weight: “Therefore, the variability of a speaker’s linguistic competence can be viewed as a probabilistic combination of multiple grammars” (YANG, 2002). Yang assumes that only one grammar is responsible for input analysis, sentence comprehension and production. In both proposals, Roeper’s and Yang’s, it is possible to infer that children must check possible grammars against the input they receive.

Considering BP variable number agreement, two possibilities are allowed: redundant agreement and non-redundant agreement. These two possibilities would lead to the identification of different grammars, however, none of these two seem to be abandoned by BP speakers as shown by variationist studies (SCHERRE; NARO, 1998, 2006, 2010, 2014). Instead, they seem to have distinct weights attributed, as proposed by Yang. It is not clear, however, to which extent redundant and non-redundant agreement constitute different representations of number or different options concerning number expression of agreement at the interfaces. Besides that, as discussed in the previous section, the boundaries between closely-related varieties are blurred (LEIVADA *et al.*, 2017), an aspect that posits difficulties to the concept of multiple grammars, since the limits of the competing systems are not clear. Additionally, other proposals consider that assuming multiple grammars is not plausible: if for any case of variation or optionality encountered in language, there are multiple grammars, the number of grammars would be exponentially high (NEVINS; PARROTT, 2010).

The multiple grammars approach is an attempt to explain intra-speaker variation and to account for the fact that optionality should not emerge within a single grammar, *i.e.*, grammars are not supposed to accommodate optionality. However, Henry (1998, 2005) argues against this idea and proposes that optionality should be allowed in the grammar whenever speakers face different outputs. Henry suggests that optionality occurs in the selection to form the numeration, not in the moment of the derivation. The options speakers maintain in their grammar produce different outcomes. Thus, children would be exposed to different outputs. This account emphasizes the role of the input: a child exposed to input *x* and input *y* represents a grammar with optionality. Separate grammars would be acquired only in cases of bilingualism, that is, acquisition of different languages.

Another proposal that explores intra-speaker variation within one grammar is discussed by Adger (2006, 2007, 2014, 2016) and Adger and Smith (2005, 2010).

Adger and Smith (2005) Adger and Smith analyze Buckie, a dialect from Scotland which presents *was/were* variation with 1st and 2nd person pronouns and full NPs (examples (3-4) from ADGER; SMITH, 2005, p. 154):

(34) He says 'I thocht you ***were*** a diver or somethin'.'
He said 'I thought you were a diver or something.'

(35) 'Aye, I thocht you ***was*** a scuba diver.'
'Yes, I thought you were a scuba diver.'

They argue that variation emerges from two possible semantically equivalent forms derived from the syntactic system. They suggest that the arrangement of features in syntax accounts for variation: when a lexical item [F: a][38] is combined with a lexical item [G: b], it follows that the final representation is the combination of these two with all interpretable features ([F: a]…[G: b]); but if [F: a] is merged with [*u*G: b], syntax will necessarily value *u*G and the outcome of the representation is ([F:a, ~~*u*G~~:b] … [G:b]). They argue that both representations have the same semantic interpretation given its interpretable features. However, the checked uninterpretable feature in the second representation may have an outcome in the spell-out:

> The minimalist system we have proposed above is more akin to the competing grammars/parametric variation approaches, since there is no notion of a probability tied to a particular rule (in fact the only rules, Merge, Move and Agree, are invariant and apply categorically in particular cases). In our proposal, like that of the competing grammars/parametric variation frameworks, the notion of choice of variant is not assumed to be part of the specification of the syntactic system itself, rather it is a separate mechanism that interacts with the syntax. However, our perspective differs from these approaches in that it assumes only one invariant grammatical system, containing universal mechanisms, rather than a range of systems. Each speaker, however, has a lexicon, a memorized store of pairings of syntactic features and lexical meanings, and it is the choice of lexical item that is the source of variation. This choice is influenced by various factors: ease of lexical access (perhaps related to how common the word is), questions of speaker-hearer relationships, notions of social identity, ease of processing etc. (ADGER and SMITH, 2005, p. 164)

[38] We borrow the notation in bracket [ ] from Adger (2006) to refer to features of lexical items.

Adger (2006) proposes a model, *Combinatorial Variability*, to account for the variability that emerges from the featural (under)specification of lexical items in Buckie. He argues that, depending on the way interpretable and uninterpretable features are combined in syntax, via *Agree*, the subset of features in PF yields different results, where LI means lexical item:

(36) LI_1 $\{F_1, F_2, F_3\}$...
LI_2 $\{uF_1\} \rightarrow$ PF $(LI_2) = x$
$LI_3\{uF_2\} \rightarrow$ PF $(LI_3) = x$
$LI_4\{uF_3\} \rightarrow$ PF $(LI_4) = z$

Variation arises within the combinatorial system, given that all these items have the same interpretable features but differ in PF forms: "What allows the variability is the possibility that particular lexical items may be underspecified for the uninterpretable agreement features that they contain. This underspecification is irrelevant to the semantic systems, since these features are not interpreted." (ADGER, 2006, p. 510). The lexical items must then be in an agreement-chain, that is, the set of interpretable features of a lexical item must agree with the subset of uninterpretable features of the other lexical item. Adger proposes the following algorithm:

> Seek Maximal Generalization by
> (a) Generating all n-feature LIs, where n=1.
> (b) Matching them with forms.
> (c) Rejecting all LIs where it is not the case that a feature bundle can always be mapped to a single form (Reject Optionality); that is, an LI is kept if there is a form which that LI always matches.
> (d) If a feature can always be mapped to a single form, but this creates synonymy, eliminate LIs to reduce synonymy as much as possible (Reject Synonymy).
> (e) Recursing over n=n+1, with the proviso that if a form has been successfully analysed in the n–1th step, LIs capturing it in the n^{th} step will be rejected (Minimize Lexicon).(ADGER, 2006, p. 518)

Optionality is supposed to be reduced to a minimum lexicon, otherwise, Adger proposes that, if optionality remains, it is a case of *doublets* in Kroch's terms.

Adger claims that the algorithm is able to capture possible frequency rates of output. These frequency rates, however, may be perturbed by "performance-related effects" (ADGER, 2006). These factors are explored in the concept of *Pool of Variants* (PoV).

Adger (2007) proposes that PoV would be a sort of storage of variants ($v_1...v_i...v_n...$). Thus, grammar is described as: $G \rightarrow \{v_1, ...v_i...v_n\}$(= PoV), where (G)rammar exists within a performance model: "One can conceive of the systems of use U as a choice function on the pool of variants, given a context of utterance C: $U (PoV,C) = v_i \in PoV$" (ADGER, 2007, p. 696). Factors of U (systems of use) and C (context) influence PoV which is equivalent to G (grammar) and its variants. Thus, G is responsible for representations and structural derivations, while U accommodates frequency and social factors. The interaction between PoV and U give rise to variability. Refining their analysis, Adger and Smith (2010) propose a lexical-feature based approach to variation. In this approach, variation is viewed as underspecification of the mapping between morphosyntactic features and morphological forms of functional categories. This approach seems to be compatible with minimalist principles such as interface legibility.

In a more recent view, Adger (2014) proposes that morphemes can span a sequence of functional categories. He suggests that language acquisition proceeds from matching morphemes in the primary linguistic data to extended projections. Additionally, he proposes that a functional label may contain instructions for spell-out. Basically, Adger proposes, based on Brody (2000 *apud* ADGER, 2014) that a syntactic structure, built by a sequence of labels, have correspondent morphemes that agglutinate in a morphological structure for spell-out of the sentence.

In this view, children's task during language acquisition is to map distributional morphophonological properties to a syntactic object. According to Adger (2014, p. 191), children use syntactic resources given by UG to build structures and must linearize them into morphophonological chunks. Variation emerges from the tension between what children generate in syntax and the morphemes associated to labels of functional categories. Consequently, variation vanishes or reduces when children try to match the distributional properties and frequencies encountered in the input.

For children acquiring Buckie, for instance, two possible morphological structures in terms of linearization are possible for the verb in a sentence like 'The

teachers asks them': 'ask' and 'asks'; both are kept in PoV, however only one will be pronounced. Therefore, in this view, variation emerges from the mapping between morphophonological properties and labels of syntactic nodes. Additionally, Adger highlights that the Combinatorial Variability system proposed in Adger (2006) is actually relevant for frequency and probability in language production, during lexical access: "The grammar generates a restricted range of variants depending on the syntax and then some function will apply to the Pool of Variants to select a single one in any occasion of use." (ADGER, 2014, p. 193). Adger's accounts for variation combine feature (under)specification, labelling in syntax and morphophonology. In a somewhat different direction, some approaches for variation present a strictly post-syntactic account for variation.

Based on the Distributed Morphology framework (HALLE; MARANTZ, 1993), Nevins and Parrott (2010) propose variation as a result of application of rules of Impoverishment in the features of Vocabulary Items on the way to PF. Impoverishment rules delete marked features of a specified Vocabulary Item, turning it into an *elsewhere*, less specified Vocabulary Item. In this approach, Impoverishment variable rules are discussed in Monmouthshire English. In this Welsh dialect, the following inflectional paradigm is found for the present tense (example (19) from NEVINS; PARROTT, 2010, p. 1114):

Inflectional Paradigm in Monmouthshire English				
Person	Singular	Plural	Features	Vocabulary
1	*I be/am*	*us be*	[+auth -pl]	æm
2	*thee beest*	no data	[+part -auth]	bist
3	*her is*	*they be*	[-pl]	ɪz
			elsewhere	bi

Table 13: Monmoutshire English inflectional paradigm and Vocabulary Items for each featural specification (adapted from: Nevins and Parrott (2010, p. 1144))

Markedness-driven Impoverishment rules apply to the marked feature [+author], resulting in Vocabulary Insertion of *be*. Therefore, the inflectional paradigm goes through a process of levelling by means of Impoverishment, which removes the marked feature of the most specified Vocabulary Item, allowing variability to emerge.

Notice that Nevins and Parrott (2010) propose that variation emerges from language-internal mechanisms: rules of Impoverishment, with no apparent external motivation for such rule application. Their account seems to be on a different direction of those presented before. Adger (2007) suggests that processing issues and social factors influence linguistic variation. Other proposals such as Roeper's (1999) mention social factors as possible influence on language variation. Henry (1998) also mentions the linguistic environment as a possible influence on variation. Yang (2002) considers the possibility of statistical and probability effects on the choice of a grammar.

Given that number agreement in BP is socially influenced, as it has been long attested, we believe that social factors must be considered in a characterization of language representation and language production. Although the Distributed Morphology approach such as the one offered by Nevins and Parrott (2010) seems to account for variability, we explore different architectures of language in order to evaluate the most suitable one for the purposes of the present thesis.

4.1
Combining formal linguistic theories and language production

This thesis is concerned with the identification and representation of the type of variation that seems to be part of the speakers' knowledge about language (SCHERRE, 1994). Thus, it is assumed that this sort of knowledge must be represented somehow in the lexicon in order to differentiate systematic variation from agreement errors (BOCK; MILLER, 1991; HARTSUIKER *et al.*, 2003; RODRIGUES, 2011; TUCKER; IDRISSI; ALMEIDA, 2015) or slips of the tongue (LEVELT, 1999) which seem to be restricted to utterance externalization. In BP, the frequency of agreement errors of attraction is 6,86% in educated speakers according to Rodrigues (2006) and systematic variation in illiterate's speech is at 48% in 3rd person subject-verb agreement (see Table 2).[39]

[39] Notice that agreement errors involve intervening elements between the subject and the verb; thus, the distance between subject and verb seem to contribute to agreement errors (MARCILESE et al., 2017; RODRIGUES, 2006). In number agreement variation, it is acknowledge that distance between subject and verb favors less redundancy in plural agreement, however, high levels of variation is

In the minimalist grammar architecture, mainly, the Strong Minimalist Thesis, grammar consists of: a lexicon that feeds syntax which, in turn, provides outputs to PF and LF, interfaces with Sensory-Motor (S-M) and Conceptual-Intentional (C-I) systems respectively. The computational system is optimal and must adhere to legibility conditions posed by the interfaces. In this framework, all variation and idiosyncrasies are located in the lexicon and variation emerges from experience (CHOMSKY, 1995). In this thesis, we argue for lexical representation in terms of matrices of features of different types following Chomsky (1995): semantic, phonological and formal features. The first type of features, semantic and phonological, are legible at the interfaces and the formal features can be interpretable or uninterpretable, which are computed by syntax via *Agree*. According to Chomsky (2013): "Clearly LIs [Lexical Items] are complex objects with many properties that do not derive from operations of narrow syntax, and there are many questions about where these complexities enter into a derivation (e.g., are phonological properties introduced in externalization to SM, as proposed in distributed morphology?). If φ-features can move independently of the complex in which they appear in an LI, then they are atoms of computation but are not LIs." (CHOMSKY, 2013, p. 46). Chomsky mentions the Distributed Morphology (HALLE; MARANTZ, 1993; HARLEY; NOYER, 1999; MARANTZ, 1997). Such framework is compatible with Levelt's production model in which there is a post-syntactic component for morpheme insertion, the morphophonological encoding.

In DM, phonological information and *Vocabulary Item* insertion takes place only after syntactic computation, hence *Late Insertion*. Information on form is accessed post-syntactically, on the way to PF: "it is not the case that all morphemes are represented underlying with both a meaning component and a sound component. Instead, some morphemes possess only a representation of meaning -(and syntax-) related features, and acquire phonological properties only after they have been combined into complex structures." (EMBICK, 2015, p. 6). Embick (2015) reinforces the advantage of *Late Insertion* of *Vocabulary Items* under DM by assuming that it maximizes generalizations and minimizes the number of *Vocabulary Items* in grammar. For instance, in Spanish, in a case of syncretism, second and third person plural (*habla-**n***), though encode distinct *synsem* features,

reported even when there is no element intervening between elements within the nominal phrase or between the subject and the verb (SCHERRE; NARO, 1998).

must be analyzed with the same Vocabulary Item /-***n***/, instead of two distinct ones (EMBICK, 2015, p. 113). A way of describing syncretism is assigning underspecified *Vocabulary Items*, that is, *Items* that are not fully specified, may fit into more than one distinct *synsem* features: "Underspecifying Vocabulary Items allows for syncretism to be analyzed as systematic effects, and not as instances of accidental homophony." (Embick 2015, 118).

Therefore, DM's main assumptions are *Late Insertion*, Underspecification and Syntactic hierarchical structure all the way down (HARLEY; NOYER, 1999). Additionally, DM proposes a difference between *l-morpheme*s and *f-morphemes*, roughly equivalent to open and closed-class items respectively: "[…] f-morphemes are those whose content (as defined by syntactic and semantic features made available by Universal Grammar) suffices to determine a unique phonological expression. […] In contrast, an l-morpheme is defined as one for which there is a choice in spell-out: an l-morpheme is filled by a Vocabulary Item which may denote a language-specific concept." (HARLEY; NOYER, 1999, p. 4). Such distinction between *l-morphemes* and *f-morphemes* seems to catch the nuances already verified by language processing results in which morphemes are accessed independently (ROELOFS; MEYER; LEVELT, 1998; TAFT; FORSTER, 1975). DM's morpheme distinction is also compatible with the notion that free and bound morphemes, open and closed class items respectively, are stored and accessed independently (SCHILLER; COSTA, 2006) – other morphology theories also assume that there must be a distinction between types of morphemes (see BEARD, 1987; BEARD; VOLPE, 2005 for a model of lexeme/morpheme distinction). [40]

Harley and Noyer (1999) propose that in some cases, *Vocabulary Items* may compete for insertion in f-morphemes; subset principles determine the winner, the

[40] Beard (1987) proposes that grammars must necessarily account for independence of lexical and morphological components. The basic assumption of the *Lexeme/Morpheme-Based Morphology Model* is exactly the distinction between *lexeme* and *morpheme* (BEARD, 1987, p. 31):

"(27a) Assume lexeme L, a mutually implied triplet P, F, R, *where P = a* prespecified nonnull phonological matrix, F = a feature inventory specifying lexical and syntactic categories, and R = all knowledge of the projected world reference associated with P.
(27b) A morpheme, M, is a(ny) modification of P."

Beard's model seems to be in line with processing results related to bound and free morphemes as mentioned.

most specified Item is inserted or the Item "that uniquely has the highest feature in the hierarchy is inserted" (HARLEY; NOYER, 1999, p. 5). In the case of variable number agreement in BP, for instance, f-morphemes of plural agreement are variable, they actually exhibit two options of Vocabulary Item, /-s/[41] or /-ø/, depending on the variety spoken.

In order to deal with variation in morpheme insertion regarding plural agreement in BP, we combine these two frameworks. On one hand, lexical items are characterized as matrices of features, on the other hand, morpheme insertion is conceived of as occurring post-syntactically:

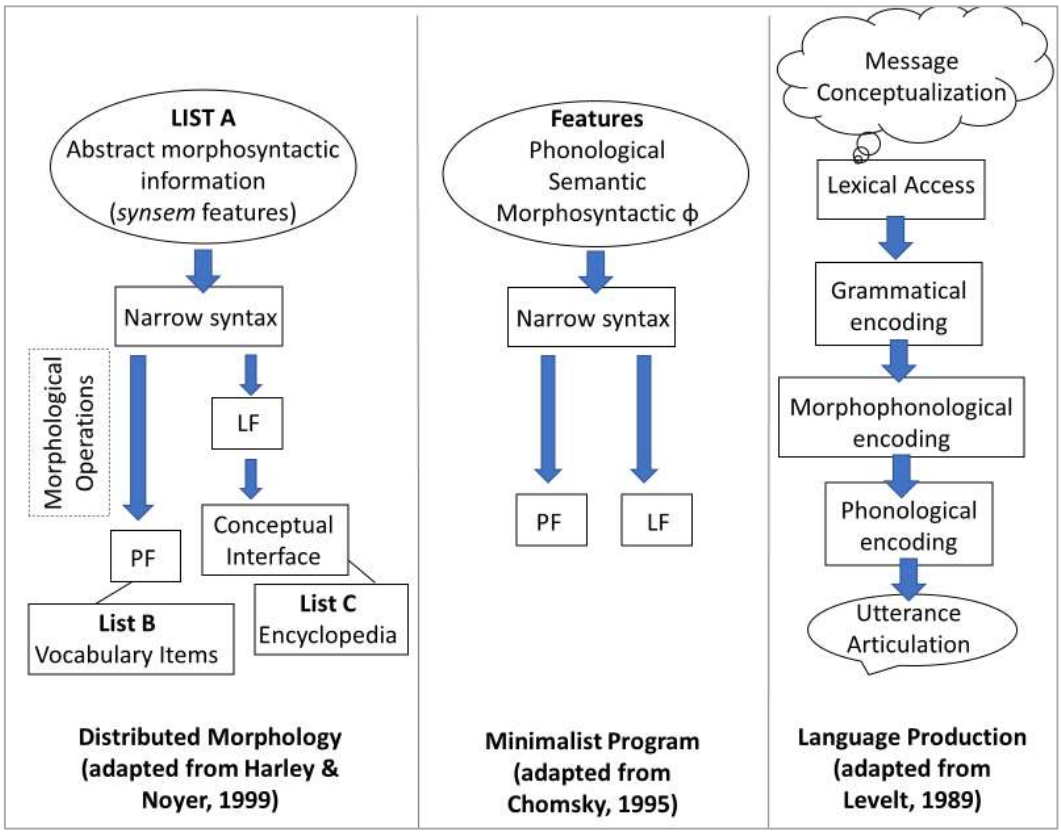


Figure 3: Comparison of Distributed Morphology and Minimalist language architecture with language production model

As mentioned before, we suggest that morphophonological – and not only phonological – features are represented in the lexicon. However, such morphophonological representations are accessed post-syntactically, in PF.

[41] /-s/ is used as a generic reference to plural morphemes.

We propose that the access to morphophonological information in the lexicon, which is exclusively post-syntactic, as suggested by DM, is actually guided by pre-syntactic/pre-lexical intentions oriented by means of C-I interaction, which is compatible with the minimalist framework.

4.2
An account for variation in number agreement in BP

For our proposal on variation, we rely on Adger (2006) and Adger and Smith (2005, 2010) that intra-speaker variation emerges from the underspecification on the mapping between forms and morphosyntactic information. Recently, Adger (2014) refined his analysis by proposing that different morphophonological forms correspondent to syntactic labels of categories are stored in PoV, which we assume to be a repository in the lexicon. Different morphophonological forms that have the same semantic interpretation are interpreted in this thesis as underspecification of morphophonological features in the lexicon. It is worth highlighting that *underspecification* here is understood in the sense of Rooryck (1994), who suggests the idea of variable underspecified features or *α-valued features*[42]: "At first sight, underspecified α-valued features simply appear to be unspecified features, but I argue that their complete absence of specification plays a role in the grammar 'variable' or α-features have no 'fixed' value, but can 'pass on' the features of the elements surrounding them" (ROORYCK, 1994: 209 – author's highlights). Nevertheless, in the current proposal, *underspecification* is a representational tool adopted to account for the possibility that redundancy of plural number agreement in BP may be expressed or not in agreeing elements. Therefore, *underspecification* implies competition in the expression of plural number agreement (± redundant). In sum, *underspecification* is considered in this thesis as a natural consequence of acquisition based on variable input. It is presumed to be represented in the lexicon,

[42] Rooryck (1994) contrasts variable underspecified features, *α-valued*, and non-variable underspecified features, 0-valued. The former is "chameleonlike" feature with no particular value set and can adapt to the featural environment; the latter is not + nor – but has a neutral value associated to it, marking the absence of a feature.

more specifically, in the morphophonological features of the lexical items. Hence, underspecification causes gradience and fluctuation in speech production.

Part of the planning of the utterance to be articulated needs to address the context, the type of interlocutor, register used and consequently, choice of variety. In Levelt's (1989, figure 1.1, p. 9) model of language production, discourse model, situation knowledge, and encyclopedic knowledge are accessed in message generation within the conceptualizer. Therefore, for those speakers who have acquired the standard variety via schooling, there can be the possibility of choosing the variety as a function of the context during message generation. In this case, consistent redundant marking at a post-syntactic level will depend on the standard variety being chosen pre-syntactically.

In sum, morphophonological information is not accessed pre-syntactically, but, the possibility of the choice of a variety, standard or non-standard, may occur pre-syntactically during message generation:

Possibility of the choice of variety during message generation	**Specification *status***	**Outcome during morphophonological encoding**
Standard	Specified (+redundant)	/-s/
Non-standard	Underspecified (±redundant)	/-s/ or /-Ø/

Table 14: Analysis for possible outcomes for variable number agreement in BP during language production

Morphophonological information is accessed in the lexicon, the equivalent of *lexemes* (*cf.* ROELOFS; MEYER; LEVELT, 1998; ROELOFS, 1998), during morphophonological encoding, that is, post-syntactically, and it is assumed that inflectional morphemes are stored separately from stems (*cf.* BORER, 1981, 1984; MARSLEN-WILSON; TYLER, 2007; STEMBERGER; MACWHINNEY, 1986; TAFT; FORSTER, 1975) in a sort of a *Pool of Variants* (ADGER, 2007). When morphophonological information of [number: + plural] is accessed in the lexicon, it necessarily forces plural marking at least one pre-nominal element, for instance, in D. For agreeing elements, as they are underspecified for the morphophonological

expression of number agreement redundancy: two morphemes are available in PoV: /-s/ or /-ø/; the insertion of one or the other morpheme is governed by the weight attributed to each variety (YANG, 2002). The weight of each variety is based on information in PoV regarding frequency rates (JESCHENIAK; LEVELT, 1994; *cf.* LEVELT, 1999). Therefore, the following production model is proposed for variation in number agreement in BP:

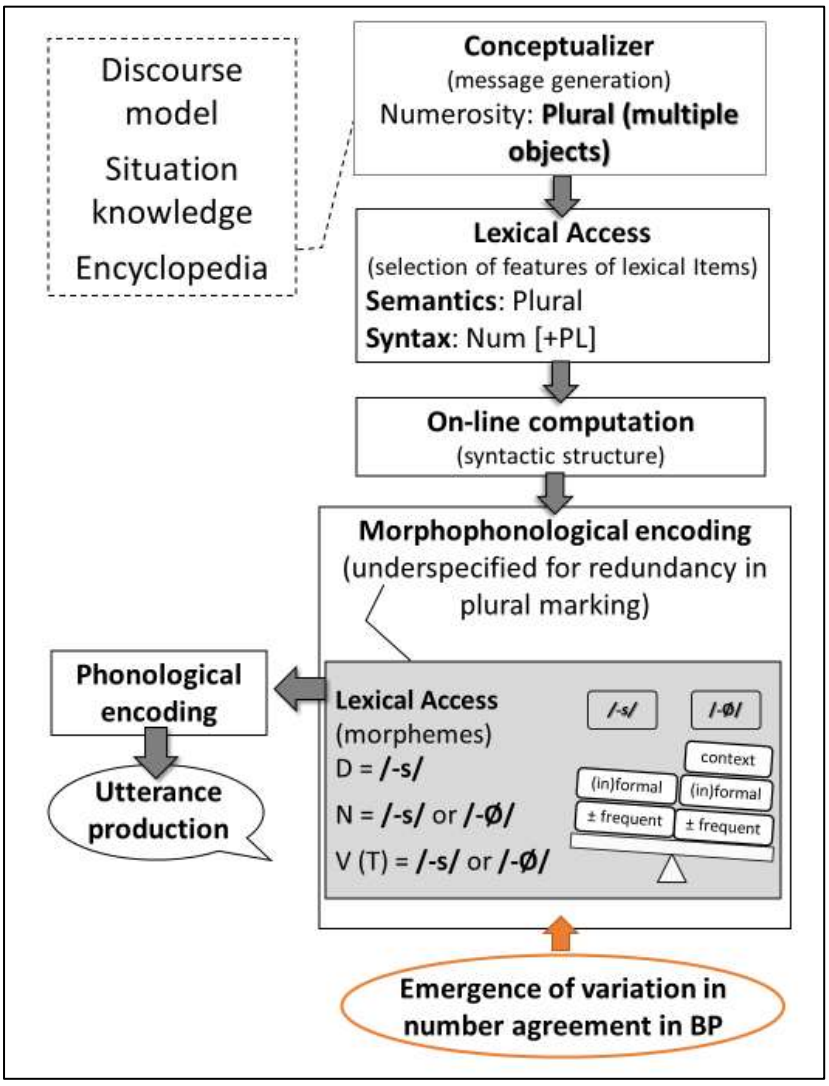


Figure 4: Emergence of number agreement variation in BP during speech production

In our analysis, we assume different types of morphemes: bound and freestanding (see LAPOINTE; DELL, 1988; SCHILLER; COSTA, 2006). Number morphemes are bound morphemes and determiners, such as personal, possessive

and demonstrative pronouns are considered freestanding functional morphemes. Hence, number inflection is necessarily bound to a freestanding item, D, for instance, during morphophonological encoding.

In sum, variation in production of number agreement in BP is assumed to emerge from: (i) morphophonological underspecified representation in the lexicon allowing competition between two or more items to be inserted; (ii) independent access to the storage of inflectional bound morphemes and freestanding functional morphemes during morphophonological encoding in speech production; (iii) influence of performance factors and linguistic experience stored as relevant information for each variety in PoV, influencing on their weight (YANG, 2002). Therefore, in the analysis presented, syntax is invariant, it does not accommodate variation. Actually, variation is posited in the way interface information interact in language processing and acquisition, which is reflected in morphophonology. This is in accordance with Chomsky (1995: 169)'s assumption that "variation must be determined by what is 'visible' to the child acquiring language, that is, by the PLD (primary linguistic data)" and it converges with recent proposals in the literature that variation is an externalization aspect pertaining phonological/ PF components (ADGER, 2014, 2016; BERWICK; CHOMSKY, 2011; BOECKX, 2016; BOECKX; LEIVADA, 2014; EMBICK, 2015; NEVINS; PARROTT, 2010).

With this characterization, we are able to account for variation as a post-syntactic phenomenon in terms of morphophonological insertion of items In sum, we propose that exposure to variable input results in underspecification of features in the lexicon. Specifically, for number variation in BP, underspecification of morphophonological features pertaining to plural agreement redundancy that are accessed post-syntactically during morphophonological encoding.

Regarding exposure to variable input, Miller and Schmitt (2010, 2012, 2014) find that children exposed to variation exhibit a sort of delay in the acquisition of some patterns: "What is clear, though, is that variation that causes ambiguity does make children take longer to acquire the ability to use variable forms in comprehension." (MILLER; SCHMITT, 2010, p. 1191). Such delay is interpreted under the presented account as a result of underspecification.

Additionally, we assume that underspecification may actually be gradually specified depending on external environmental influences. Miller and Schmitt (2010, 2012), for instance, find differences in relation to SES. They investigate

adults' and children's production of plural in Spanish by grouping participants according to age and SES in Spanish. Chilean Spanish and Mexican Spanish differ in relation to pronunciation of plural /-s/: Mexican Spanish presents redundant agreement in the DP, whereas Chilean Spanish may present redundant agreement /-s/, non-redundant agreement /ø/ or substitution from /-s/ to /-h/ (examples from MILLER; SCHMITT, 2012):

(37) Plural marking pronunciation in Chilean Spanish:
Plural: *vaca-s* - [*bakas*], [*bakah*], [*baka*]
cow-PL
'cows'
Singular: *vaca-ø* [*baka*]
cow-SG
'cow'

Production tasks (spontaneous speech, sentence repetition and picture naming) reveal that Mexican children and adults present ceiling performance; Chilean adults from working-class present less plural marking than Chilean middle-class adults. Chilean children scores mirror the adults' scores. As for comprehension, Chilean children interpret plural as 'more than one' but ignore the plural morphological marking. Hence, exposure to variable input leads to underspecification of some sort in the features represented in the lexicon.

Notice that in our experimental results, there was also a difference of SES regarding standard responses provided: Priv preschoolers, higher SES, provided more standard responses than Pub preschoolers. Additionally, we verified a tendency for Pub to repeat non-standard stimuli more often than Priv. Both Priv and Pub preschoolers presented very low means of standard responses, which shows that there is a genuine competition between standard and non-standard varieties for both groups, but at different frequency rates. Effects of frequency can only be accounted for in a speech production model and those effects appear exactly during morphophonological encoding. Thus, our results show that preschoolers are sensitive to effects of frequency, both in the stimuli provided (a sort of priming effect) and in the input they receive on a daily basis, being Pub more sensitive to frequency effects than Priv, who tends to correct the stimuli. Therefore, we believe

that our account, presented in Figure 4, explains the delay in Miller and Schmitt's data – availability of two options in PoV – and explains the results of BP preschoolers presented in the previous chapter.

Our preschoolers' result shows a high level of variation in language production and frequency effects influence. Remember that they were not exposed to written/standard BP and this is consonance with illiterates' data in Table 7, in which, number agreement production is at 48%. Thus, grammars who were never exposed to formal instruction at school show a high level of variation in number agreement.

Social and pragmatic mapping of information may not be developed in preschoolers – Labov comments that only in adolescence such social factors of language use are fully mapped – so, their production would be subject to frequency effects and the type of variation in their production would be strictly post-syntactic. That is, /-s/ or /-ø/ would be retrieved in PoV according to effects of frequency of the input.

Therefore, we assume BP, acquired spontaneously, to be an underspecified grammar, that is, with morphophonological features pertaining to plural agreement redundancy represented as underspecified. Gradual specification of varieties would be dependent on external factors such as processing, social factors and schooling. In order to verify whether schooling and introduction to written/standard BP favors a bi/multilingual environment in what concerns number agreement, the same experiment was carried out with 6th graders. In the next chapter results are presented.

5

The effect of literacy in the variable number agreement in BP: 6th graders' results

Kato (1990) highlights that literacy affects speaking and writing in different ways: in syntax, the type and length of the syntactic structure; in relation to lexical aspects, density and size of vocabulary; in morphology, tense, modals, negation and agreement. Kato makes a parallel between spoken standard form and written standard form as if spoken standard was a simulation of written standard. Additionally, she comments on the use of technology for the diffusion of standard spoken forms and the environmental factors that influence language use. For instance, children growing up in a low SES environment but with literate parents may develop linguistic and literacy skills similar to literate speakers, despite their social environment.

In order to verify the impact of literacy and SES on children's preference for morphophonological expressions of plural number agreement redundancy, the same experiment carried out with preschoolers was conducted with 6th graders from a private and a public school in Rio. The aims of this second experiment were to verify:

(i) if type of school plays any role on the use of number agreement during schooling years;

(ii) if there is influence of literacy and schooling on morphological awareness regarding the use of number agreement;

(iii) if there is any difference between preschoolers and 6th graders;
   a. when the influence of these external factors can be detected;
   b. how SES and literacy interact;
      i. which are the directions of influences of SES and literacy

As in the first experiment, the linguistic profile of the 6th graders' families from each school was assumed to be the following:

a) Priv: parents are mainly speakers of the standard variety (passive bilinguals for the non-standard one) and/or simultaneous bilinguals, having the standard variety as the dominant one;
b) Pub: parents are mainly speakers of the non-standard variety (passive bilinguals for the standard one).

It is hypothesized that literacy and schooling affect 6th graders performance. The previsions are that:

i. Priv would show preference for the standard form more frequently than Pub;
ii. 6th graders would show preference for the standard form more frequently than preschoolers.

5.1
Method

5.1.2
Participants

The same experiment was conducted with 77 6th graders from a private and a public school in the same suburban area in Rio. The groups are arranged as follows:

Type of school	n	girls	boys	Age range	Mean age
Priv	25	14	11	10;11 – 12;11	11;1
Pub	52	29	23	11;1 – 14;6	12;6

Table 15: Groups of 6th graders as a function of *Type of School*

The age difference between groups of 6th graders has to do with the fact that, in Pub, some students have failed, at least, one academic year.

5.1.3
Procedure

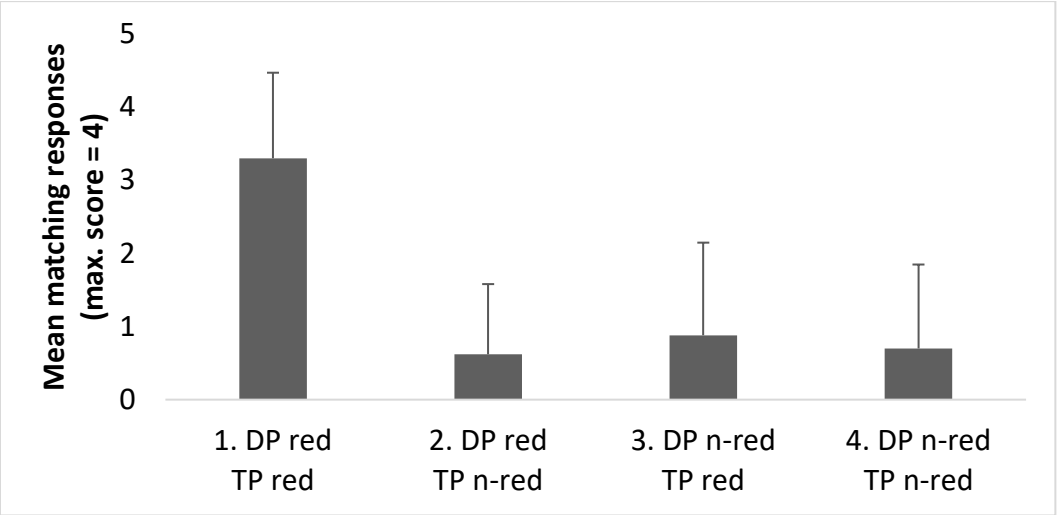
The procedure of the experimental task was slightly different with 6th graders. In between the end of the stimuli presentation and the beginning of the elicited production, the experimenter asked simple math sums such as 'how much is 5+2? (answer) Ok, now tell me what Bob has told you.' Such strategy intended to swift the attention from the memorization of the sentence to the resolution of math calculations which, in turn, provided reliability in participants' responses, in the sense that they would not simply repeat the stimuli but access the most prominent form of number agreement redundancy in their grammar.

5.2
6th graders: Results per Type of School

The data was analyzed by means of a 2-way ANOVA (design 4 X 2), in which *Redundancy in Number Agreement* is a within-subject factor (4 levels), and *Type of School* is a between-subject factor (2 levels). The dependent variables were the number of responses matching the stimuli and the number of standard responses provided.

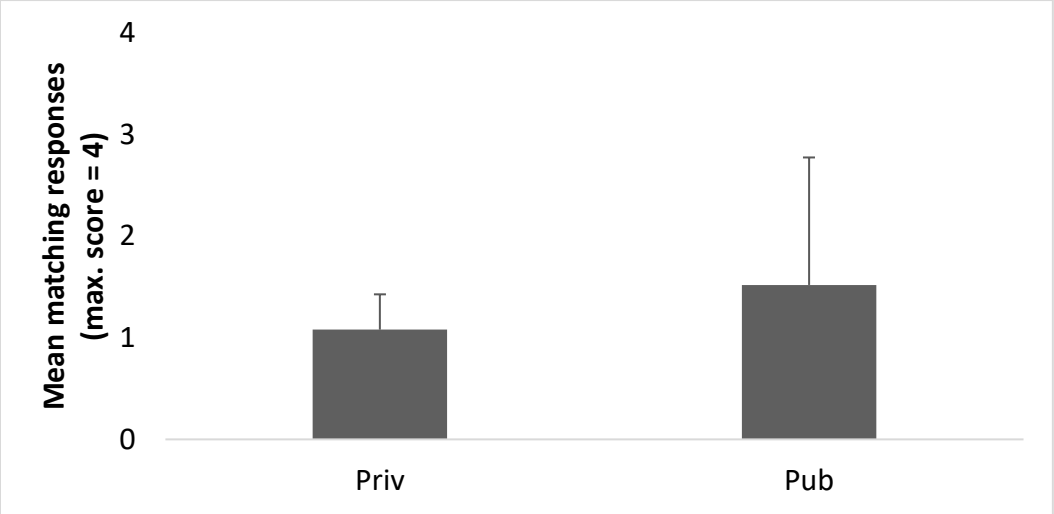
5.2.1
Dependent Variable 1: Matching responses

As for the responses that matched the stimuli provided, there was a significant main effect of *Redundancy* (F (1, 75) = 117, 829; p < .0001; C1: *M*= 3,30, *SD* = 1,2; C2: *M* = 0,62, *SD* = 1; C3: *M* = 0,88, *SD* = 1,3; C4: *M* = 0,7, *SD* = 1,1):



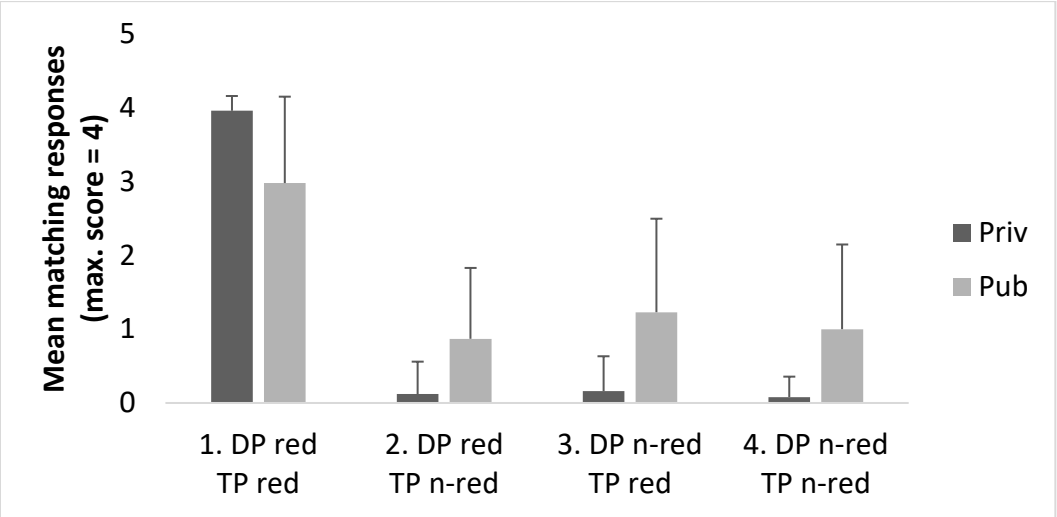
Graph 5: Mean matching responses as a function of *Redundancy in Number Agreement* (Error bars = *SD*)

Additionally, there was a main effect of *Type of School* (F (1,75) = 11,195; p = .001; Priv: *M* = 1,08, *SD* = 0,3; Pub: *M* = 1,52, *SD* = 1,3):



Graph 6: Mean matching responses as a function of *Type of School* (Error bars = *SD*)

There was also an effect of interaction between *Redundancy* and *Type of School* (F (1,75) = 14,488; p < .0001; Priv C1: *M* = 3,96, *SD* = 0,2; C2: *M* = 0,12, *SD* = 0,4; C3: *M* = 0,16, *SD* = 0,5; C4: *M* = 0,08, *SD* = 0,3; Pub C1: *M* = 2,98, *SD* = 1; C2: *M* = 0,87, *SD* = 1,2; C3: *M* = 1,23, *SD* = 1; C4: *M* = 1, *SD* = 1,1):

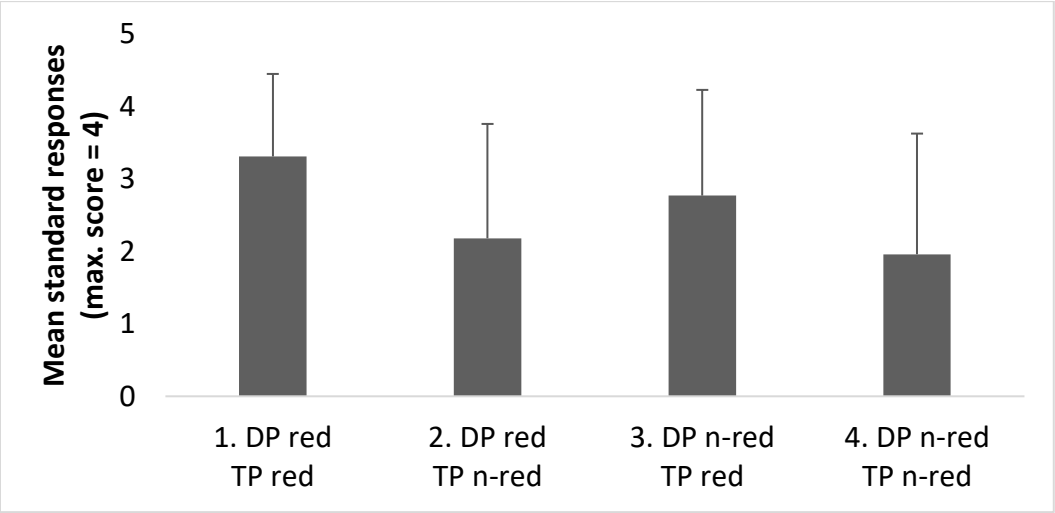


Graph 7: Mean matching responses as a function of *Redundancy in Number Agreement* per *Type of School* (Error bars = *SD*)

Altogether, the results of matching responses show the same pattern attested with preschoolers' results, but with sharper differences: both groups prefer the standard form. However, they differ in the production of non-standard forms; Pub exhibits higher means for non-standard conditions than Priv. Graph 7 shows that Priv's matching responses are concentrated in one condition, which is condition 1 (DP red/ TP red), the other 0,08 is distributed along the non-standard conditions. Pub exhibits higher means in graph 6 showing a tendency to reproduce the stimuli.

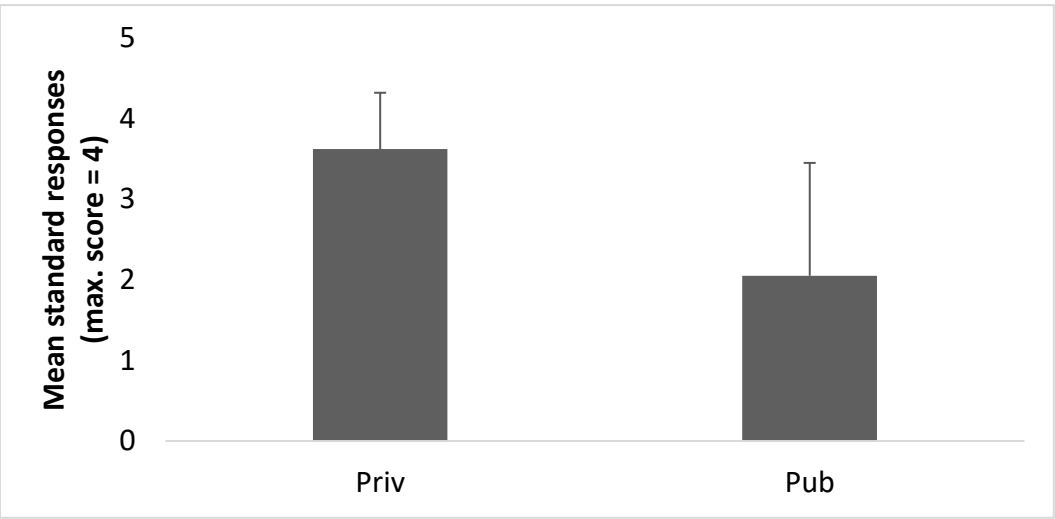
5.2.2
Dependent Variable 2: Standard responses

Regarding the standard responses provided in each condition, the data was analyzed by means of a 2-way ANOVA (design 4 X 2), in which *Redundancy* is a within-subject factor (4 levels), and *Type of School* is a between-subject factor (2 levels). There was a significant main effect of *Redundancy* ($F(1, 75) = 49{,}060$; $p < .0001$; C1: $M = 3{,}31$, $SD = 1{,}1$; C2: $M = 2{,}18$, $SD = 1{,}6$; C3: $M = 2{,}77$, $SD = 1{,}5$; C4: $M = 1{,}96$, $SD = 1{,}7$):



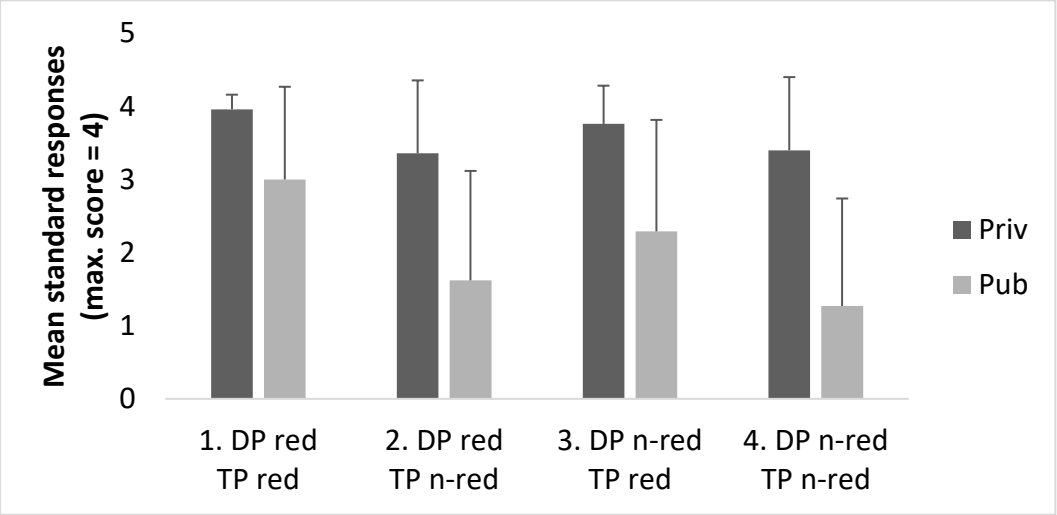
Graph 8: Mean standard responses as a function of *Redundancy in Number Agreement* (Error bars = *SD*)

Furthermore, there was a main effect of *Type of School* (F (1, 75) = 37, 373; p < .0001; Priv: *M* = 3,62, *SD* = 0,7; Pub: *M* = 2,05, *SD* = 1,4):



Graph 9: Mean standard responses as a function of *Type of School* (Error bars = *SD*)

There was also an effect of interaction between *Redundancy* and *Type of School* (F (1, 75) = 15, 306; p < .0001; Priv C1: *M* = 3,96, *SD* = 0,2; C2: *M* = 3,36, *SD* = 1; C3: *M* = 3,76, *SD* = 0,5; C4: *M* = 3,4, *SD* = 1; Pub C1: *M* = 3, *SD* = 1,3; C2: *M* = 1,62, *SD* = 1,5; C3: *M* = 2,29, *SD* = 1,5; C4: *M* = 1,27, *SD* = 1,5):



Graph 10: Mean standard responses as a function of *Redundancy in Number Agreement* per *Type of School* (Error bars = *SD*)

Notice that the analysis of standard responses shows a similar pattern to the preschoolers: conditions 1 (DP red/ TP red) and 3 (DP n-red/ TP red) trigger the higher means of standard responses (*cf.* graph 9). In graph 10, it is possible to conclude that Priv and Pub present qualitatively similar performance but quantitatively different. Such quantitative difference shows up in graph 9, in which Priv exhibits much higher means for standard responses than Pub, who stay close to the average total score.

Overall, the 6th graders' results show that the differences between groups seem to be sharpened during schooling years. Literacy seems to enhance 6th graders awareness on the use of standard forms. Given the tracking school system adopted by the public school, we decided to run another analysis within the Pub group to verify whether overall academic performance is a factor that influences the use of number agreement in BP.

5.3
6th graders: Results per Academic Performance in the public school

In order to evaluate the extent to which academic performance can alter the picture provided by the type of school, the data of the public school 6th graders were further analyzed by contrasting two groups as above-average academic performance (Pub A) and below-average academic performance (Pub B):

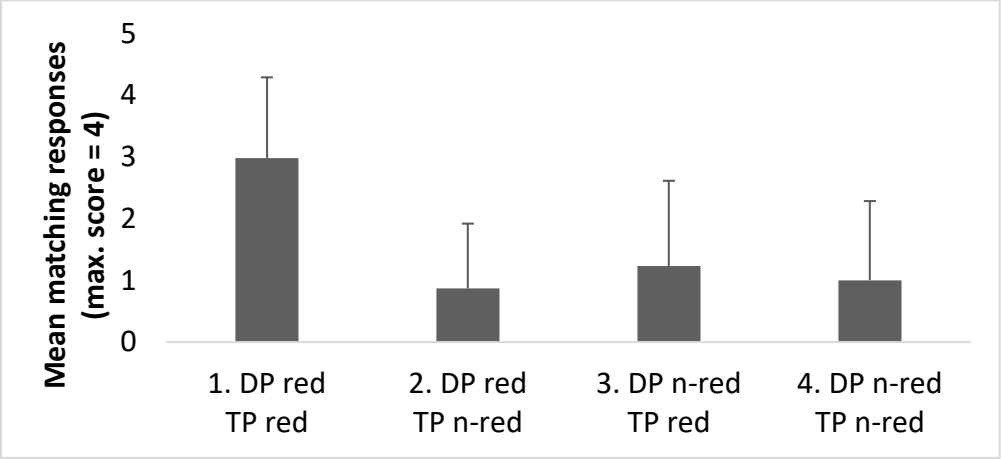
Academic Performance	n	girls	boys	Age range	Mean age
Pub A	25	17	8	11;3 – 14;6	12
Pub B	27	12	15	11;1 – 14;5	12;6

Table 16: Groups of Pub 6th graders as a function of *Academic Performance*

The data was analyzed by means of a 2-way ANOVA (design 4 X 2), in which *Redundancy* is a within-subject factor (4 levels), and *Academic Performance* is a between-subject factor (2 levels). The dependent variables were, as usual, the number of responses matching the stimuli and the number of standard responses provided.

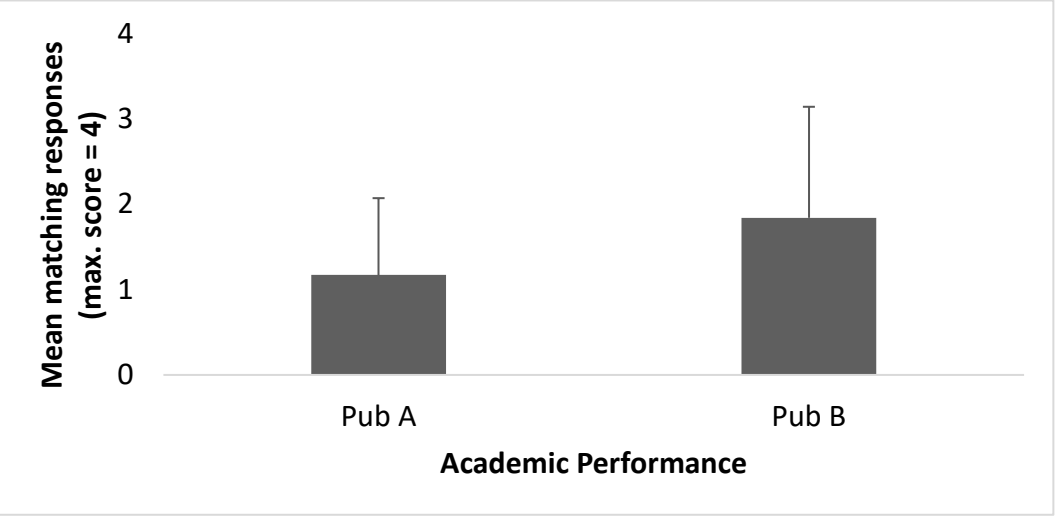
5.3.1
Dependent Variable 1: Matching responses

There was a main effect of *Redundancy* (F (1,50) = 35,571; p < .0001; C1: *M* = 2,98, *SD* = 1,3; C2: *M* = 0,87, *SD* = 1; C3: *M* = 1,23, *SD* = 1,4; C4: *M* = 1, *SD* = 1,3):



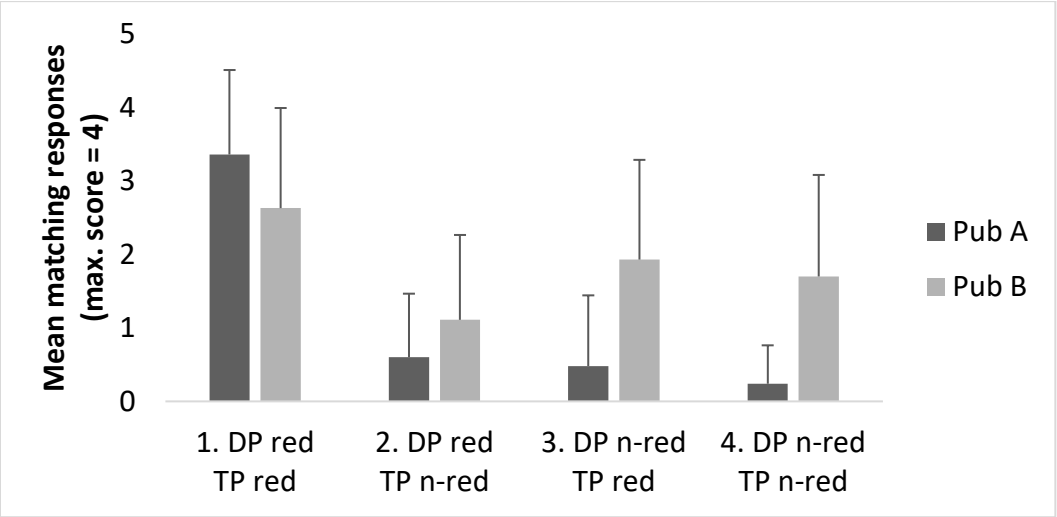
Graph 11: Mean matching responses as a function of *Redundancy in Number Agreement* (Error bars = *SD*)

A main effect of *Academic Performance* was also obtained (F (1,50) = 19,761; p < .0001; Pub A: *M* = 1,17, *SD* = 0,9; Pub B: *M* = 1,84, *SD* = 1,3):



Graph 12: Mean matching responses as a function of *Academic Performance* (Error bars = *SD*)

Additionally, the analysis showed an effect of interaction between *Redundancy* and *Academic Performance* (F (1,50) = 15,35; p < .0001; Pub A C1: *M* = 3,36, *SD* = 1,2; C2: *M* = 0,6, *SD* = 0,9; C3: *M* = 0,48, *SD* = 1; C4: *M* = 0,24, *SD* = 0,5; Pub B C1: *M* = 2,63, *SD* = 1,4; C2: *M* = 1,11, *SD* = 1,2; C3: *M* = 1,93, *SD* = 1,4; C4: *M* = 1,7, *SD* = 1,4):

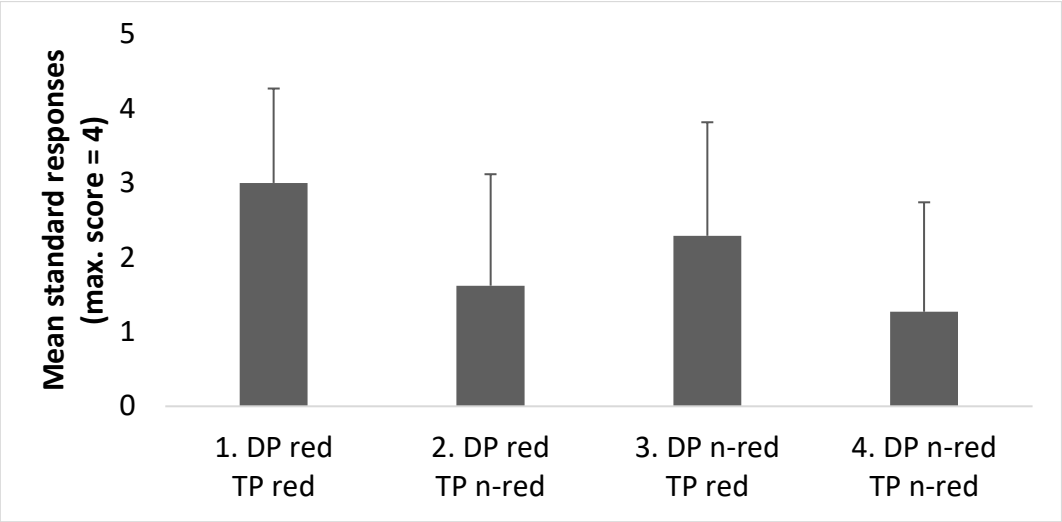


Graph 13: Mean matching responses as a function of *Redundancy in Number Agreement* per *Academic Performance* (Error bars = *SD*)

Overall, the results show that the standard condition (DP red/ TP red) is the most reproduced one for both groups as shown in graph 11. However, Pub B reproduces non-standard forms more often than Pub A, as seen in graph 12. This distribution is confirmed in graph 13 in which Pub B has higher means for matching responses across different conditions than Pub A, who presents means close to 1 in most part because of condition 1 (DP red/ TP red). Next, we present the standard responses analysis.

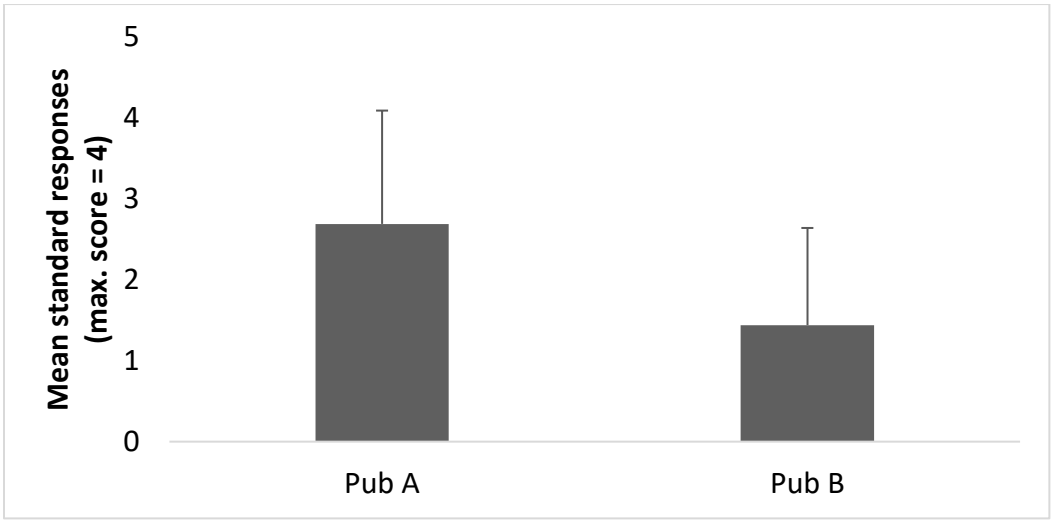
5.3.2
Dependent Variable 2: Standard responses

Again, the data was analyzed by means of a 2-way ANOVA (design 4 X 2), in which *Redundancy* is a within-subject factor (4 levels), and *Academic Performance* is a between-subject factor (2 levels). There was a main effect of *Redundancy* (F (1,50) = 80,728; p < .0001; C1: *M* = 3, *SD* = 1,3; C2: *M* = 1,62, *SD* = 1,5; C3: *M* = 2,29, *SD* = 1,5; C4: *M* = 1,27, *SD* = 1,5):



Graph 14: Mean standard responses as a function of *Redundancy in Number Agreement* (Error bars = *SD*)

There was a main effect of *Academic Performance* (F (1,50) = 17,877; p < .0001; Pub A: *M* = 2,69, *SD* = 1,4; Pub B: *M* = 1,44, *SD* = 1,2):



Graph 15: Mean standard responses as a function of *Academic Performance* (Error bars = *SD*)

There was no effect of interaction between *Redundancy* and *Academic Performance* (F (1,50) = 3,272; p = 130; Pub A C1: *M* = 3,4, *SD* = 1; C2: *M* = 2,44, *SD* = 1,6; C3: *M* = 2,92, *SD* = 1,3; C4: *M* = 2, *SD* = 1,6; Pub B C1: *M* = 2,63, *SD* = 1,4; C2: *M* = 0,85, *SD* = 1; C3: *M* = 1,7, *SD* = 1,5; C4: *M* = 0,59, *SD* = 1).

The patterns seen in the analysis of standard responses provided by Pub 6th are in parallel with preschoolers' and Priv 6th graders' results: standard responses are mostly provided in condition 1 (DP red/ TP red), which is the standard form, and in condition 3 (DP n-red/ TP red). Graph 15 shows that good academic performance enhances the production of standard forms, regardless of SES. Overall, these results show that academic performance during schooling years is also a factor that influences the production of number agreement in BP. Actually, academic performance in interaction with SES seem to give rise to a sort of *continuum* in terms of the use of standard forms by 6th graders: Priv > Pub A > Pub B. Now, the breakdown of responses of 6th graders is compared:

Type of School/ Academic Performance	Types of responses																					
	DP red TP red			DP red TP n-red			DP n-red TP red			DP n-red TP n-red			Singular			No D			Ungrammatical			
Condition	Priv	Pub A	Pub B	Priv	Pub A	Pub B	Priv	Pub A	Pub B	Priv	Pub A	Pub B	Priv	Pub A	Pub B	Priv	Pub A	Pub B	Priv	Pub A	Pub B	Total valid responses
1	**27**	**23**	**32**	0	1	1	0	2	10	0	0	1	0	3	1	0	0	0	0	0	0	308
2	27	20	7	**5**	**10**	**1**	1	2	5	1	2	10	2	4	3	0	0	0	0	0	0	308
3	31	24	15	0	0	0	**4**	**17**	**1**	0	0	1	0	0	0	0	0	0	0	4	2	308
4	28	16	5	0	6	8	2	3	3	**2**	**15**	**1**	2	5	3	0	0	0	0	0	1	308

Condition 1 = DP red/ TP red
Condition 2 = DP red/ TP n-red
Condition 3 = DP n-red/ TP red
Condition 4 = DP n-red/ TP n-red

Matching responses in bold.

Table 17: Distribution (%) of responses per condition and type of school/ academic performance – 6th graders

As shown in the table above, in conditions 1 (DP red/ TP red) and 2 (DP red/ TP n-red), both Priv and Pub A present similar performance. In condition 3 (DP n-red/ TP red), the difference is sharpened. Finally, in condition 4 (DP n-red/ TP n-red), there is a clear *continuum* of the distribution of standard responses across the groups of 6th graders. In condition 1 (DP red/ TP red), Pub B provides some responses equivalent to condition 2 (DP red/ TP n-red), unlike the other 6th graders. In condition 2 (DP red/ TP n-red), Pub B tends to produce responses equivalent to condition 4 (DP n-red/ TP n-red), whereas Pub B produces standard forms or reproduce the stimuli as presented. In condition 3 (DP n-red/ TP red), when 6th graders were not producing the standard form, they tended to reproduce the stimuli. In condition 4 (DP n-red/ TP n-red), both Pub A and Pub B provided different non-standard responses, whereas Priv tended towards the standard variety.

Priv clearly prefers the standard form, among the non-standard forms, the following preferred order appears: DP n-red/ TP red > DP red/ TP n-red > DP n-red/ TP n-red. Thus, for Priv, the *continuum* varies from standard to plural in the verb, plural in the subject and plural only in D. For Pub A, the preference is also for the standard form and the non-standard scale is: DP n-red/ TP red > DP red/ TP n-red and DP n-red/ TP n-red. Pub A exhibits competition among non-standard forms and this group's *continuum* varies from the standard form to plural in the verb and competition between plural in the subject DP or only in D. In relation to Pub B, there is also preference for the standard form and the non-standard scale is: DP n-red/ TP red > DP n-red/ TP n-red > DP red/ TP n-red. Hence, for this group, the *continuum* of number agreement varies from the standard form to plural only in the verb, plural only in D and plural in the subject DP. Results show that Priv and Pub A exhibit very similar performance with the exception that Pub A but not Priv present competition among non-standard forms. As for Pub B, the non-standard scale is qualitatively different. Additionally, this table shows that, overall, the preferred non-standard form in 6th graders' result is DP n-red/ TP red, meaning that plural marking in the verb is relevant for these groups.

Notice that academic performance and SES interact in different ways in order to regulate 6th graders' variable production of number agreement in BP. In the next section, academic performance and its relation to morphological awareness of plural number agreement in BP are discussed.

5.4
Academic performance as a boost for morphological awareness of varieties among 6th graders

Unlike the preschoolers, half of the 6th graders noticed the switch between standard and non-standard forms of number agreement in BP in the robot speech. Those who actually noticed the presence of non-standard number agreement exhibited metalinguistic morphological awareness by providing responses such as: "he drops the 's'", "he does not speak the plural", "he should speak like this (redundant) but he speaks like that (non-redundant)", "he alternates between present and past tenses"[43]; "he speaks in plural and diminutive forms"[44]; "he forgets the plural"; "he should pay attention to singular and plural forms". The chart below shows the division among 6th graders in terms of morphological awareness for each group defined as overall academic performance and type of school, taken as social group:

Social Group	Morphological Awareness
Priv (25)	84
Pub A (25)	72
Pub B (27)	29

Table 18: 6th graders – Distribution (%) of children as a function of Social Group and Morphological Awareness

There is, again, a sort of *continuum* also in terms of morphological awareness in 6th graders' results which corresponds to the quantitative results analyzed in the previous sections: Priv > Pub A > Pub B. Therefore, these results point to the fact that academic performance actually boosts metalinguistic/ morphological

[43] Though these participants misunderstood tense, number and word form, they noticed differences in grammatical features.
[44] *Cf.* note 50

awareness, that is, academic performance may help enhance the sensibility to varieties distinction in terms of standard and non-standard forms of number agreement redundancy in BP.

However, though academic performance may help enhance morphological awareness as discussed, being morphologically aware does not entail standard-variety-only production. Within Priv, 66,6% of participants who were aware of the lack of plural produced standard-only forms. For Pub A, the percentage of standard-only production drops to 27,7% for the morphologically aware 6th graders. For Pub B, this percentage drops even more, to 12,5%. An interesting fact is that one participant in Pub B exhibited morphological awareness ('he omits the plural') but did not produce any standard responses.

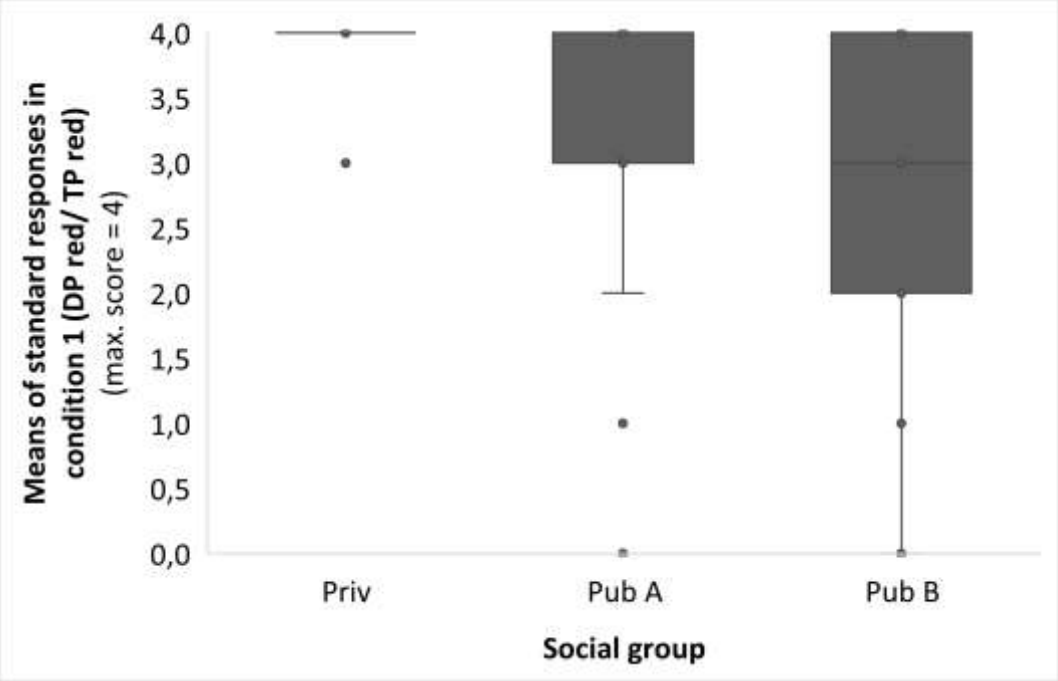
Morphological awareness is a by-product of schooling and formal exposure to grammatical rules. The production of written essays in the standard variety as required by the educational system in the 6th grade and exposure to reading and writing activities with great intensity affect the production of number agreement in BP. These results show that, within the same suburban area, and even within the same school, students' responses are subject to a great range of variability regarding redundancy in number agreement. Blurred borders between varieties affect perception and production (LEIVADA *et al.*, 2017): children may be aware of the lack of morphological redundant agreement, but they still produce it.

Thus, 6th graders seem to lay in a *continuum* of variation. Such *continuum* is also attested in so-called *creoles*. Gonçalves (2009) proposes that the Portuguese spoken in Moçambique, as in most former colonies, presents a polilectal *continuum* ranging from the standard EP-like variety to a basilectal, forming sub-varieties that present different frequency rates within and cross-speakers. Furthermore, in African American English (AAE) studies, researchers have been dealing with gradience in language production, equivalent to the notion of *continuum*, in terms of dialect density: high and low dialect users (see EDWARDS *et al.*, 2014 for a discussion on lexical comprehension in Mainstream English and African American English). In terms of number agreement in BP, Pub B would be the high dialect users, Pub A would be low dialect users and Priv would be very-low dialect users, mostly restricted to comprehension.

It seems that morphological awareness indicates dialect awareness, basically compared to sociolinguistic or metalinguistic awareness (see AYIOMAMITOU;

YIAKOUMETTI, 2017 for a discussion on Standard Modern Greek and Cypriot Greek). Interestingly, the participants who exhibit metalinguistic awareness, perform better and also exhibit improved code-switching abilities, trying to inhibit as much as possible the non-standard variety, as in Pub A case (see CRAIN and WASHINGTON, 2004; TERRY, 2012; WHEELER, 2016).

As an additional analysis, we show in graph 16, below, that few 6th graders perform differently from their peers, but similar to other social groups:



Graph 16: Outliers per social group in condition 1 (DP red/ TP red)

This graph illustrates the variability in the suburban area of Rio in which there is not a clear-cut division between SES, meaning that speakers from different social groups interact on a daily basis. For all these reasons, a characterization in terms of *continuum* of variation seems to be more appropriate for number agreement in BP.

5.5
Discussing the impact of literacy and SES on variable number agreement in BP

6th graders' results are in consonance with preschoolers' results: there is no categorical production of standard or non-standard variety in language production in BP. The categorical production in our experiment was verified intra-speaker, mainly in Priv 6th with standard forms and one participant in Pub B who produced only non-standard forms. Overall, in terms of speech community and/or cross-speaker variation, number agreement in BP is semi-categoric for Priv and variable for Pub. Hence, these results reinforce the variable nature of number agreement in BP as attested by sociolinguistic investigations concerning the speech production in RJ (BRANDÃO, 2013; FIGUEIREDO BRANDÃO, 2016; GUY, 1981; SCHERRE; NARO, 2006; NARO, 1981; NARO; SCHERRE, 2015; SCHERRE, 1978, 1991, SCHERRE; NARO, 2010, 2014; VIEIRA; BRANDÃO, 2014 among many others).

There is, thus, a striking difference between the 6th graders from the public school. How is it that two groups from the same school, in the same region, with similar social background, same teachers most of the time, exhibit such contrast in the production of morphophonological plural marking and, consequently, morphological awareness? Investigating AAE speakers, Craig and Washington (2004) found a relation between dialect awareness and overall academic performance, AAE students who could code-switch outperformed their non-switcher peers in reading and vocabulary tasks. In this thesis' results, overall academic performance and literacy seems to play a role on the production of morphophonological number agreement, however, academic performance is also dependent on a range of other factors (*cf.* KUHL, 2010, 2011). Perkins *et al.* (2013) highlight that SES, home language, home literacy environment, parenting style as well as chronic stress depending on SES influence the development of language and brain. Fernald *et al.* (2013) show that SES differences in language processing skills and vocabulary emerge from 18 months old: at this age, children from lower SES exhibit lower vocabulary scores and less efficient real-time processing than children from higher SES background. Basit *et al.* (2015) find differences in preschoolers' production and comprehension according to their SES and highlight the challenge that nurseries and preschools face when preparing children for compulsory schooling, given that children from low and very low SES tend to show some delay in language production and comprehension when assessed by *New Reynell Developmental Language Scales*.

Regarding the social reality in Brazil, or most developing countries, a cluster of factors may contribute to children's poor academic performance: lack of financial resources for education (and health as well), social issues affecting school year and children's attendance to school, such as urban violence, high number of illiterates, unemployment, restricted access to varied types of food and/ or malnutrition and many other issues that social inequality brings (*cf.* DEHAENE-LAMBERTZ; HERTZ-PANNIER; DUBOIS, 2006).

In relation to the 6th graders in this study, some observations should be presented. Students are generally organized in classes per age: Pub B (mean age: 12,2; median: 12) and Pub A (mean age: 11,4; median: 11). The age difference is somewhere close to a year. In both groups, it was reported that some students are older due to previous school year failure.[45] Another issue is the complex relation between schools and families in Brazil. Reports on Brazilian educational context carried out by UNESCO have been discussing the controversial relation between schools and families in Brazil: there are expectations from both sides and transference of roles due to the unclear boundaries and lack of cooperation between both social spheres.[46,47] There is much complaint on parents not participating in students' school life: this was a complaint for Pub B during some visits to the school.

Literacy and overall academic performance seem to enhance participants' morphological awareness due to exposure to formal instruction in the standard variety but it does not necessarily inhibit the production of non-standard forms of number agreement in BP. Rather, it seems to contribute to the improvement of code-switching ability between varieties. Therefore, literacy and overall academic performance seem to be acting directly on the *Pool of variants* (ADGER, 2007), changing the weight attributed to each variety due to frequency effects and

[45] Each 6th grader student may come from a different Pub in the neighborhood. Public schools in Brazil are divided into preschool, elementary school I (from 1st to 5th grade) and elementary school II (from 6th to 9th grade) and, finally, highschool (1st to 3rd year); 6th graders must move from their elementary school I to another school offering elementary school II. It is not an easy task to ascertain the 6th graders background education due to school changes.

[46] By parents here, it is meant the adult legally responsible for the child, not necessarily mother and father. It is common in Brazil that children are raised by grandparents, mostly grandmothers. Hence, it is not a matter of family structure in terms of mother and father but an adult who would be responsible for the child.

[47] Interação escola-família: subsídios para práticas escolares. UNESCO/MEC. 2010. Available at: http://portal.mec.gov.br/index.php?option=com_docmanandview=downloadandalias=4807-escola-familia-finalandItemid=30192

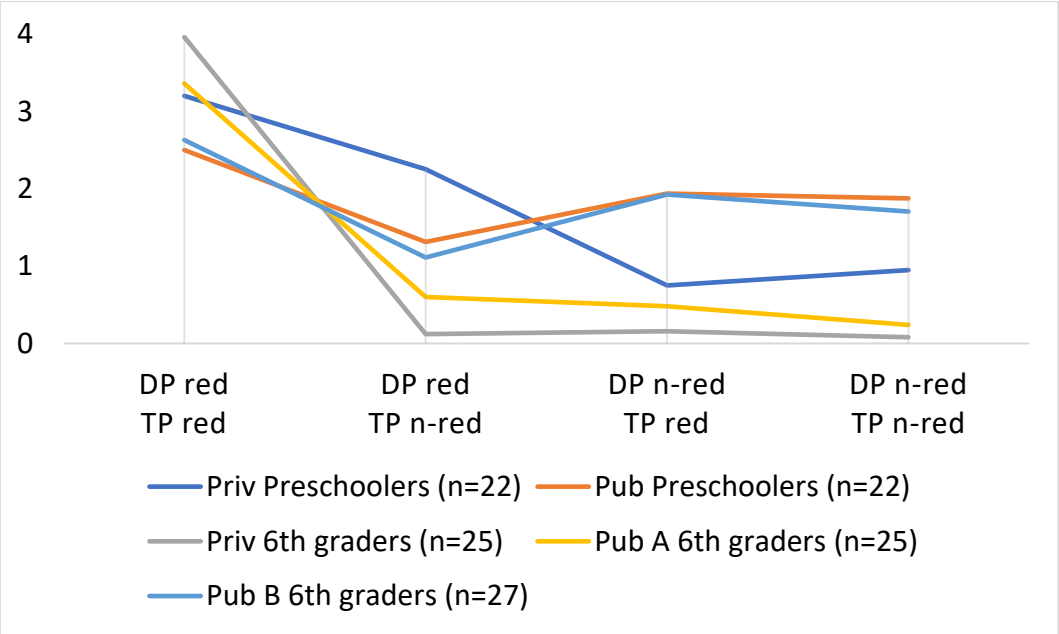
environmental factors (LEVELT, 1999; YANG, 2002; YANG; ELLMAN; LEGATE, 2015).

Therefore, a range of factors contribute to the difference observed among 6th graders. However, the question on whether SES and overall academic performance directly affects language in specific domains, such as number agreement, remains an open issue for further inquiry. As pointed out by Pace *et al.* (2017), the integration between psycholinguistic approaches for language acquisition/ development and bioecological approaches is necessary in order to understand how SES affects language development. Furthermore, Pace *et al.* (2017) highlight the difficulty in tearing apart causal relations between SES and language development.

In the results of the present thesis, language development during preschool seems to be dependent on SES. Academic performance, together with SES, seem to be decisive during schooling years. Notice, though, that, SES seems to be less relevant than overall academic performance during schooling years, given that Pub A exhibit similar responses to Priv, despite the SES difference among them. In the next section, we present an overview of the results joining both preschoolers and 6th graders.

5.6
Variation in number agreement from preschool to school

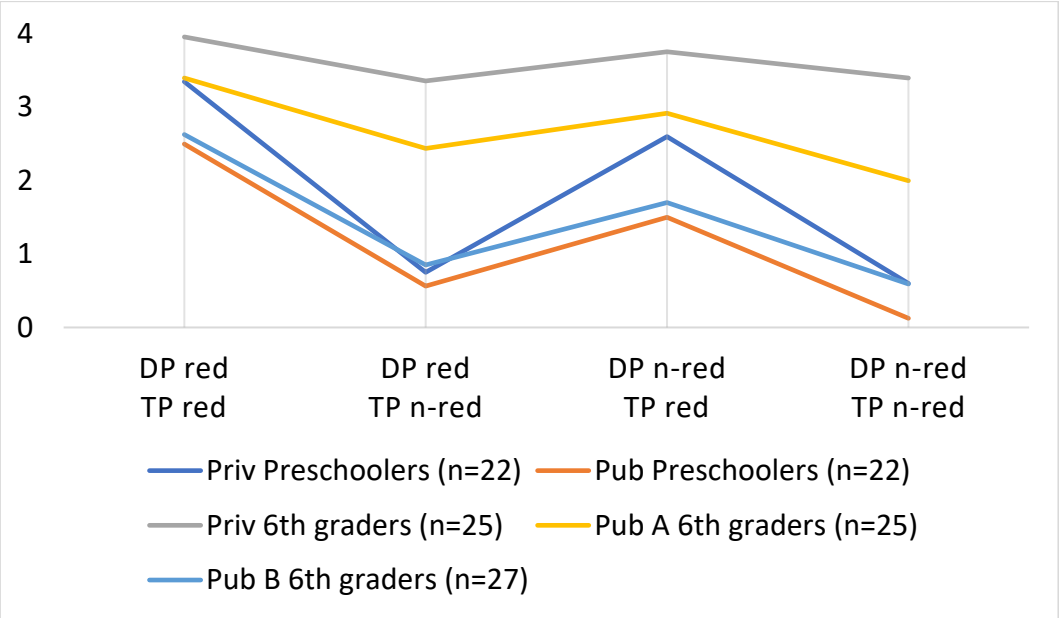
As 6th graders exhibited morphological awareness due to literacy by formal means at school, an age comparison was carried out in order to verify possible changes in participants' performance before and after literacy. When comparing data from preschoolers and 6th graders in terms of the score of matching responses in each experimental condition, a curious relation emerges:



Graph 17: Preschoolers and 6th graders comparison of matching responses across conditions of redundancy in number agreement

Notice that Priv preschoolers, the dark blue line, and Pub A, the yellow line, exhibit very similar means of matching responses provided in condition 1 (DP red/ TP red), but differ in relation to non-standard conditions. Priv 6th graders and Pub A exhibit similar qualitative results, being the difference mainly quantitative: both groups present clear preference for the redundant form. Priv preschoolers exhibit different performance from all the other groups with preference for reproducing redundant and DP redundant stimuli. Pub preschoolers and Pub B 6th graders exhibit very similar performance.

In relation to the standard responses provided, again, we see that Pub preschoolers and Pub B 6th graders exhibit very similar performance and Priv preschoolers show higher means of standard responses than Pub preschoolers and even Pub B 6th graders in conditions 1 (DP red/ TP red) and 3 (DP n-red/ TP red):



Graph 18: Preschoolers and 6th graders comparison of standard responses across conditions of redundancy in number agreement

Overall, Priv and Pub preschoolers and Pub B 6th graders exhibit the same performance qualitatively, with quantitative differences. Priv and Pub A 6th graders present similar performance: Priv 6th graders show higher means than Pub A. Because Pub B 6th graders and Pub preschoolers exhibited similar performance despite their age difference, we decided to carry out more tests with Pub B in order to verify any possible difficulty with number agreement morphology. The results are presented in the next section.

5.7
Assessing possible number morphology difficulties

Due to the observation that Pub B 6th graders exhibit similar performance to preschoolers, despite their age difference, a test on number agreement morphology was carried out with these students. The test applied is a development of a broader project concerning MABILIN (*Módulo de Avaliação de Habilidades Linguísticas* – Module of Language Abilities Assessment) a battery of tests for the evaluation of linguistic abilities in different domains (CORRÊA, 2000). MABILIN is a test that

aims to verify the extent to which children can extract grammatical relevant information from complex syntactic structures and morphology in BP. MABILIN I focuses on the assessment of complex syntactic structures such as passives, interrogatives, relative clauses. MABILIN II focuses on morphological aspects of BP. In the case of this thesis, MABILIN II for number agreement morphology for language production and comprehension was carried out with Pub B 6th graders.

5.7.1
MABILIN II – number morphology: comprehension

The aims of this test were to verify (i) whether children are able to rely on D as the main or unique source of morphological information concerning grammatical number in BP and (ii) the extent to which redundancy in the morphological expression of number in the DP facilitates children's identification of the referent of a plural DP. Additionally, it aims to verify (iii) the extent to which children are able to extract number information from nouns (with regular and irregular plural forms).

5.7.1.1
Test structure

In order to verify (i), number invariant nouns were used in singular and in plural DPs such as:

(35) *O-ø/-s* *ônibus*
ART.DEF.MASC-SG/PL bus
'The bus'

(36) *O-ø/-s* *lápis*
ART.DEF.MASC-SG/PL pencil
'The pencil'

In order to verify (ii), DPs containing number marking in D, D and N, and D, N and Adj(ective) were presented. Nouns had either intrinsic gender (masculine and

feminine, counterbalanced) or optional gender (feminine, since it is morphologically marked):

(37) *O-s* *carro-s*
ART.DEF.MASC-PL car.MASC-PL
'The cars'

(38) *A-s* *bolsa-s*
ART.DEF.FEM-PL purse.FEM-PL
'The purses'

(39) *O-s* *copo-s* *pequeno-s*
ART.DEF.MASC-PL glass.MASC-PL small.MASC-PL
'The cars'

(40) *A-s* *torre-s* *altas*
ART.DEF.FEM-PL tower.FEM -PL tall.FEM-PL
'The big towers'

(41) *O-s* *caderno-s* *amarelo-s*
ART.DEF.MASC-PL notebook.MASC-PL yellow.MASC-PL
'The yellow notebooks'

(42) *A-s* *gata-s* *branca-s*
ART.DEF.FEM-PL cat.FEM -PL white.FEM -PL
'The white cats'

In order to verify (iii), bare nouns with regular and irregular plural markings were used:

(43) *Relógio-ø/ Relógio-s*
watch-SG / watch-PL

(44) *Jornal* */ Jornais*
newspaper.SG/ newspaper.PL

It was assumed that language unimpaired children would consistently rely on D as a source of number information in the DP and that redundancy would not affect their overall performance. However, plural gender inflected nouns would be expected to be costly since two agreement relations would have to be processed, number and gender itself. No predictions were made concerning 6th graders ability to identify the number information only in regular and irregular plural nouns.

5.7.1.2
Participants

For this assessment, 21 Pub B 6th graders (age range: 14;5 – 11;11; mean age: 13;6) participated in a picture selection task.

5.7.1.3
Material

The stimuli were told by the experimenter in order to make the test more interactive. For each stimulus there was a figure of a boy or a girl pretending to say something and three pictures to which participants should point. The position of the picture matching the stimuli was altered among the slide presentation and there was always a singular, a plural and a non-matching picture:

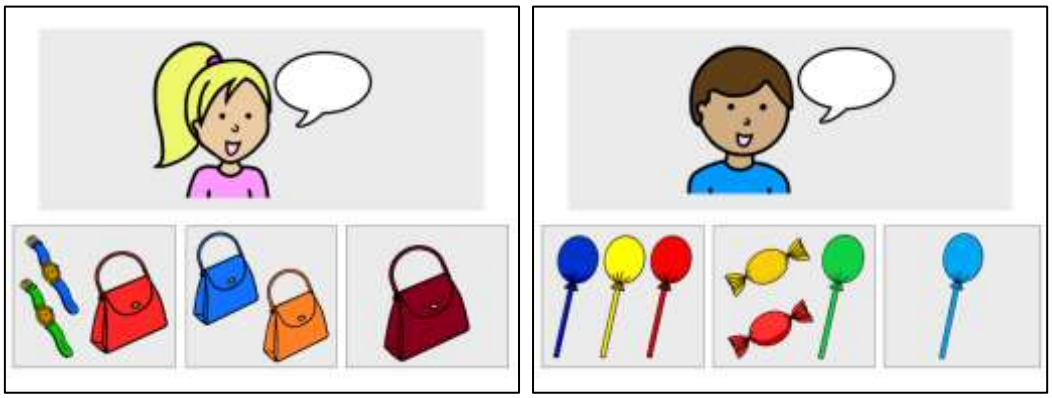


Figure 5: Example of experimental board for MABILIN II - number agreement comprehension

Pictures were shown in slide presentation mode. All the procedure was performed in a *DELL Inspiron*, i3 processor laptop. The responses were recorded in a *Sony* MP3 player-recorder and/or in a Panasonic RR-US511 recorder and written down.

5.7.1.4
Procedure

Participants were introduced to a boy, named João, and a girl, named Maria. The experimenter told the participants that they should point to the picture that matches what João or Maria said. The experimenter, then, said: *O João diz: "Eu vou chamar os alunos" Que figura você vai apontar?*/ 'I am going to call the.PL students'. Which picture are you going to point to?' Participants should point to the picture that matches what was said or tell the number related to the position of the picture: 1, 2 or 3. The dependent variable was the number of correct responses according to what was said by the experimenter, the maximum score was 3.

5.7.1.5
Results

In order to test (i), whether children are able to rely on D as the main or unique source of morphological information concerning grammatical number in BP, invariant nouns were contrasted. The independent variable was *Number* (singular /plural). A 2-tailed Paired-Sample *t*-test showed no significant effect (t (20) = 1,581; p = .130; Plural M = 2,90, SD = 0.4; Singular M = 2,57, SD = 0,8).

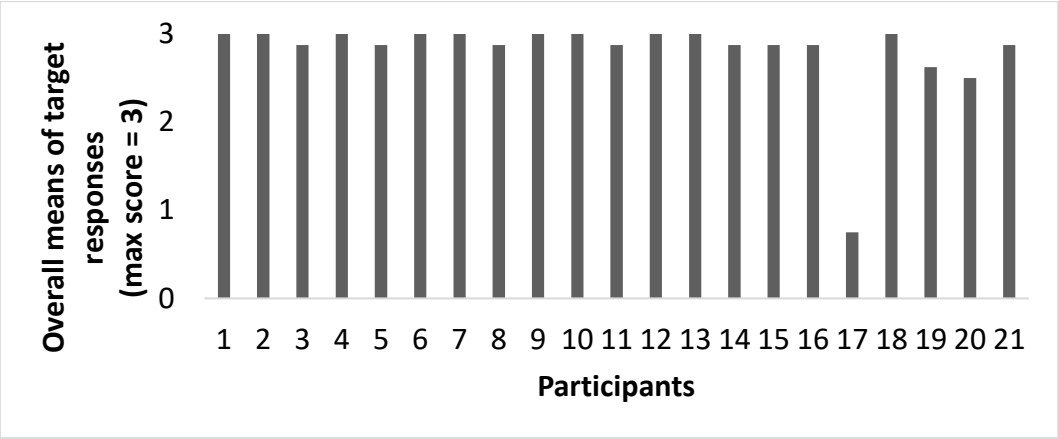
In order to test (ii), the extent to which redundancy in the morphological expression of number in the DP facilitates children's identification of the referent of a plural DP, two analyses were carried out:

a) DPs with plural invariant nouns were compared with DPs with plural variant nouns (intrinsic gender) with and without adjectives. A one-way ANOVA with *Type of Plural DP* (3 levels) as independent variable showed no significant effect (F (1, 20) = 2,105; p = .162; Plural DP with invariant nouns *M* = 2,90, *SD* = 0,4; Plural DP with variant nouns, intrinsic gender without adjective *M* = 2,86, *SD* = 0,5; Plural DP with variant nouns, intrinsic gender with adjective *M* = 2,81, *SD* = 0,7);

b) The plural DPs with intrinsic gender were further contrasted with plural DPs with optional gender with and without adjectives. A one-way ANOVA with *Type of Plural DP with variant nouns* (4 levels) as independent variable showed no significant effect (F (1, 20) = .041; p = .841; Plural DP with variant noun inherent gender and without adjective *M* = 2,86, *SD* = 0,5; Plural DP with variant noun, inherent gender and with adjective *M* = 2,81, *SD* = 0,7; Plural DP with variant noun, inflected gender without adjective *M* = 2,76, *SD* = 0,8; Plural DP with variant noun, inflected gender with adjective *M* = 2,86, *SD* = 0,7).

In order to test (iii), the extent to which children are able to extract number information from nouns (with regular and irregular plural forms), bare nouns with regular plural forms were contrasted with irregular ones. A 2-tailed Paired-Sample *t*-test showed no significant effect (*t* (20) = 0,568; p = .576; Bare noun with allomorph *M* = 2,86, *SD* = 0,5; Bare noun without allomorph *M* = 2,81, *SD* = 0,7).

Overall, MABILIN II results show that there is no possible impairment in Pub B 6th graders. These children are able to extract relevant grammatical information of number from the input and identify the correct plural referent. However, one participant in particular presented an unusual performance:



Graph 19: Participants' overall score of target responses by Pub 6thB in MABILIN II – number agreement comprehension

Notice that participant 17 exhibits a very low score, while most participants reach ceiling results. This participant tended to point mainly to pictures represented by singular objects regardless of the number morphology provided in the stimuli. Recall that the average age of this group is 13;6 years old and this particular participant is 13;3 years old. Thus, further linguistic assessment is needed with this participant in order to understand the greater picture and track possible impairments.

As shown by MABILIN results, there seems to be no problems for this group to extract grammatical information of number in BP, though their performance in Bob's experiment was similar to preschoolers' performance. The similarity between Pub B 6th graders and Pub preschoolers can be regarded as a product of exposure to variable input in number agreement in BP. We highlight that the omission of determiners in language production can be regarded as a marker of impairment in BP given that number morphology is encoded, extracted and processed in D. Thus, it is necessary to disentangle variation as a result of social variables and actual impairment, which in BP would be regarded as the recurrent omission of the determiner in language production. We emphasize that language production and comprehension should be verified for a thorough assessment of linguistic abilities.

5.7.2
MABILIN II – number morphology: production with numerals

An additional test was carried out to verify if the presence of /-s/ that is not plural morpheme, but part of the root, would trigger plural agreement. This test was based on anecdotal data from a conversation in which a person said *um real*[48], *dois reais*, *três reais*, *quatro real* (one real.SG, two real.PL, three real.PL, four real.SG). When asked why not *reais* (real.PL) with the numeral *quatro* ('four') the person answered that there is no /-s/ in *quatro* ('four') so there was no need to say *reais* (real.PL). This BP speaker used phonic assimilation to deal with number agreement. Hence, this test intends to verify if a phonological rule based on a final /-s/ realized as [ʃ, ʒ] is taken as instrumental for number agreement.

5.7.2.1
Test Structure

The test structure is quite simple and the presence of /-s/ in the numeral preceding the noun was manipulated:

Numerals with /-s/	Numerals without /-s/
Dois ('two')	*Quatro* ('four')
três ('three')	*cinco* ('five')
seis ('six')	*sete* ('seven')
	oito ('eight')

Table 19: Types of cardinal numerals in BP ending in /-s/

Each participant was exposed to 54 stimuli, 27 with a numeral ending in /-s/ and 27 with a numeral that does not end in /-s/. It aimed to verify if the presence of /-s/ in the numeral would trigger plural noun production.

[48] Brazilian currency.

5.7.2.2
Participants

For this assessment, 19 out of the 21 Pub B students (age range: 11;11 – 14,5; mean age: 13,6) that participated in the previous comprehension test also took part in the numerals test. However, 2 participants informed the experimenter that they did not want to continue in the activity.

5.7.2.3
Material

The material for this assessment consisted of one picture in the left side of the screen of a laptop and a second picture which would appear in the right side of the screen:

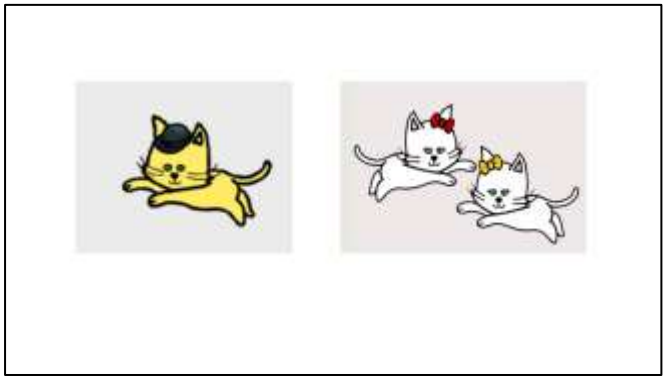


Figure 6: Example of MABILIN board – production of number agreement morphology with numerals

Pictures were shown in slide presentation mode. All the procedure was performed in a *DELL Inspiron*, i3 processor laptop. The responses were recorded in a *Sony* MP3 player-recorder and/or in a Panasonic RR-US511 recorder and transcribed by the experimenter.

5.7.2.4
Procedure

Participants should tell what appeared in the screen. As the first picture appeared in the left side of the laptop screen, the experimenter said: *Aqui tem um gato*. ('There is one cat here')[49]; the experimenter pressed a button and a second picture appeared in the right side of the screen while the left one was still displayed; as soon as the picture appeared on the right side, the experimenter said: *E aqui tem…?* ('And here there are…?'). The experimenter explained to the participant, they had to use a numeral plus the name of the object that appeared. However, some occurrences without nouns were produced.

5.7.2.5
Results

Results show that the presence of /-s/ in the numeral does not trigger a plural noun, rather, the general semantics of plurality, 'more than one', conveyed by the numeral seems to trigger plural marking in the noun:

numeral with /-*s*/	occurences	%	numeral without /-*s*/	occurences	%
plural noun	368	69,3	plural noun	331	69,2
singular noun	163	30,7	singular noun	147	30,8
total	531	100	total	478	100

Table 20: Distribution of responses provided by Pub B 6th graders in MABILIN II – number agreement with numerals

The table above shows that the frequencies of plural and singular nouns are almost the same regardless of the presence of /-s/ in the numeral. The preference is for plural nouns when numerals are greater than 1. Therefore, /-s/ that is not a plural

[49] The sentence in BP is presented with the verb *ter* ('to have') that is not inflected for number in existential sense.

morpheme does not necessarily trigger plural marking in subsequent elements. If variation was merely a PF phenomenon, and not a matter of morphophonogical encoding, it would be expected that the sound [ʃ, ʒ] in numerals ending in /-s/ would automatically trigger /-s/ plural morpheme in the subsequent noun. This prediction was not verified even in the group presenting the most variable production, Pub B; when faced with numerals, they tend to add plural /-s/ in the noun. Furthermore, /-s/ that is not a morpheme but part of the root of the word was never omitted in Pub B speech production, differently from the /-s/ plural morpheme, showing that small inflectional units are independently manipulated during language processing. This information can be useful for distinctions of word storage and access in the lexicon: if only grammatical inflectional morphemes are omitted, it supports the view that inflectional morphemes are accessed independently from the root (ROELOFS; MEYER; LEVELT, 1998; ROELOFS; BAAYEN, 2002; SCHILLER; COSTA, 2006; TAFT; FORSTER, 1975). However, further inquiry is needed. If such assumption is on the right track, inflectional morphemes must be necessarily represented in the lexicon and, in this thesis, it is argued that they are stored in the *Pool of Variants*, whenever two options compete. Thus, we argue that systematic variation must necessarily be represented in the lexicon and it may not constitute a PF-only operation. The high frequencies of systematic variation in urban BP (48% - 97,8%, according to VIEIRA, 2015) as opposed to agreement errors in language production in university students (6,86% according to RODRIGUES, 2006) also reveal that systematic variation may not be a PF-only issue. For these reasons, we argue that, in cases of systematic variation, there must be a sort of underspecified representation that displays actual variation in PF - assuming a model of linguistic knowledge – and/or in the morphophonological encoding - assuming a processing model for language production. We reinforce, in this thesis, the necessity of interdisciplinary investigations in linguistic theories and language processing.

5.7.2.6
Impacts for non-typically developing population

As Pub B 6th graders showed similar performance to Pub preschoolers despite their age gap, a module of linguistic abilities assessment, MABILIN II, verified characteristics of number morphology agreement in production and comprehension by this group.

The comprehension test revealed that, although this group exhibits high levels of variation in speech production, their comprehension regarding plural and singular is unimpaired; this group is pretty much aware of plural *vs.* singular distinction. Participant 17, though, is an exception in this group and needs further linguistic assessment. In this regard, it is important to inform speech therapists that, in contexts with high level of language variation, a thorough assessment must be carried out and, most importantly, variation in production does not entail linguistic impairment (see WILSON, 2012), unless determiner omission is attested in BP.

Another important issue that our results highlight is that control groups with typically developing children speakers of BP should be carefully assessed, given the natural variability in BP number agreement system. Children who participated in the experiments were reported not to have any cognitive or linguistic impairment. They were all considered typically developing children. Thus, we emphasize that when forming control groups in BP much attention has to be paid to social factors that influence linguistic development and speech processing. However, we highlight that such social factors have also to be considered when studying non-typically developing children. As discussed in previous section, social factors interact with cognitive and linguistic development. In sum, we suggest that social factors should also be taken as a variable of study when dealing with heterogeneous population. As we know, linguistic studies are generally carried out with university students. Such profile should not be considered as a representative sample of how a BP grammar, for instance, is represented and processed.

With these results in mind, we propose in the following sections how the linguistic knowledge regarding variation in number agreement is represented and how such knowledge is put to use during language production.

6
Accounting for linguistic knowledge and the production of variable number agreement in BP

This chapter presents a theoretical analysis regarding the representation of linguistic knowledge as well as a characterization of the production of variation in number agreement, based on the results obtained in this thesis. In the first section, an account of the linguistic knowledge in terms of acquisition and representation of plural agreement variation is presented. In the second section, a characterization of on-line language production is presented discussing its outcome concerning language variation.

6.1
The state of linguistic knowledge: obligatory plural marking in the DP in BP under a formal theoretical perspective

As shown in chapter 2, at least one prenominal element is necessarily marked for plural in the DP in BP: the determiner or the possessive (whenever it co-occurs with the determiner). Thus, there seems to be two possible domains in the DP, as suggested by Pereira (2016)[50], one for prenominal elements and another for nominal and post-nominal elements. We suggest that these two domains, can be interpreted as phases (CHOMSKY, 2001), following Picallo (2017).[51] Additionally, we argue for the necessity of the functional category Num, head of NumP.

[50] Pereira has recently developed her analysis, proposing that cardinals and silent nouns under Num function as a syntactic frontier from morphological plural marking if number in non-standard BP (see PEREIRA, 2018).

[51] Picallo claims that *n*P is a lexical phase, whereas functional elements above *nP* would constitute a functional phase.[51] It has been argued that grammatical gender is encoded in *n*, as a nominalizer (ACQUAVIVA, 2009; ATKINSON, 2015; CARVALHO, 2018; KRAMER, 2016a; LOWENSTAMM, 2007; MERCHANT, 2011; PANAGIOTIDIS, 2015; PICALLO, 2017). Picallo presents the following structure for the DP (adapted from example (7) PICALLO, 2017, p. 7):

(27) [DET … [INFL … [Num] … [*n* (Gender) … √ROOT]]]

The root N is dominated by an *n* head which encodes grammatical gender; above the 'little-n zone'. An inflectional domain (INFL) concentrates Num, responsible for grammatical number and D

6.1.2
On the necessity of Num in the DP in BP

Number is a semantic information regarding numerosity that has a formal counterpart in grammatical representation. In generative grammar, number, as well as gender and person, constitute formal φ-features (CHOMSKY, 1995). Formal φ-features are computed in syntax and the morphological expression of agreement is dependent on how each language encodes agreement (see MIYAGAWA, 2010). Agreement is, thus, taken as the possibility of morphophonological expression of the syntactic operation *Agree*. The morphophonological expression of agreement is realized in functional categories: "[…] their interaction [functional and lexical categories] imbues language with the enormous expressiveness that makes it what it is." (MIYAGAWA, 2010, p. 143).

From the point of view of language acquisition, since early ages, around 22 - 24 months old, Brazilian children are sensitive to morphophonological properties pertaining to gender and number agreement within the DP (CORRÊA; AUGUSTO; FERRARI-NETO, 2006; CORRÊA; NAME, 2003). As discussed in previous chapters, the morphophonological patterns expressed in D are instrumental for the computation of agreement between D and N in BP, as a reflex of feature valuation in syntax. Therefore, Brazilian children must rely on number information in D in order to account for agreement relations and also number semantic interpretability. We assume that number information, singular and plural, is encoded in a head under D, Num.

dominates all the nominal structure. For BP, we assume the structure presented in (27) and follow Carvalho's proposal that grammatical gender is encoded in *n* but semantically interpretable in D. Picallo (2017) also proposes that, in cases of *singular* or *pluralia tantum*, the head *n* may also bear number properties: "we could assume that inherent (or 'lexical') Number is an abstract independent category in the little-*n* space […]" (PICALLO, 2017, p. 10). Similarly, Kramer (2016b) argues for number encoding in *n* given the possibility of regular or irregular plural inflection such as *child/children* and the possibility of double plural in Amharic. In this thesis, we propose an N for the sake of simplicity, given that gender is not the main point of the discussion in this thesis. See Carvalho (2018) for a proposal in which, in BP, gender is encoded in *n*. See also Matushansky and Ionin (2015) for a proposal of number as a 'free-rider' in the sense that number is valued in the noun as a result of another feature matching. Thus, we propose that the lexical phase is constituted of N and post-nominal elements.

Given its semantic importance for mass/count distinction, bare nouns and, also, its relation with definiteness, Num has been considered a relevant head in the literature (see ALEXIADOU, 2011; CYRINO; ESPINAL, 2015; DAYAL, 2004; GRIMM, 2012; LIMA, 2014; MUNN; SCHMITT, 2005). Some proposals in the minimalist syntax have also suggested the presence of the head Num under the DP (HARLEY; RITTER, 2002; PICALLO, 2008, 2017; RITTER, 1995). Nomoto (2013, p. 156) proposes that number is interpretable in Num: "Given that determiners inflect for number in many languages, D should also have a NUMBER feature. However, it is not D but Num where the number information is encoded and interpreted. Hence, the number feature in D is uninterpretable." (author's highlights).

In BP, Num has also been assumed as an important functional category (AUGUSTO; FERRARI NETO; CORRÊA, 2006; CYRINO; ESPINAL, 2015; LOPES, 2006; PEREIRA, 2016). As for BP acquisition, Lopes (2004, 2006) and Simioni (2007) suggest that the morphological inflection of number is produced when definite articles are produced. Augusto (2007) finds that the interaction between the type of DP – bare or (in)definite – together with number marking is relevant to display different semantic interpretations; around 5 years old, the definiteness-number relation is not yet well established in BP. In terms of adults' grammaticality judgment, Ionin *et al.* (2010) found that articles are optional with plural nouns in BP and definite singulars fit into a 'well-defined-kind' restriction. Cyrino and Espinal (2015) agree with Schmitt and Munn (2002) on the presence of an empty D in BP, but claim that Num is necessary in order to instantiate different semantic interpretations; they assume that number is interpretable both in D and in Num in BP.[52]

In this study, Num is considered the head in which grammatical number, singular/ plural distinction, and numerals are encoded. Additionally, we assume along the lines of Depréz (2005 see also TRUPPI, 2014, p. 243 for a language taxonomy concerning number) the necessity of the projection NumP for semantic interpretation (see also HARLEY; RITTER, 2002). We follow Corrêa, Augusto and

[52] Cyrino and Espinal (2015) discuss the nature of bare nouns in BP by assuming an empty, null D in the lines of Schmitt and Munn (2002). However, while Schmitt and Munn propose that bare nouns lack the functional projection of Num, resulting in a bare singular, unspecified for number, Cyrino and Espinal argue for the presence of Num under DP. In this thesis, the nature of bare nouns in BP is not discussed.

Ferrari-Neto (2006)'s assumption that formal features that bear semantic import to reference have an independent projection related to a functional category: "It would allow for a distinction to be established between intrinsic and optional features during the processing of agreement in so far as the latter would project independent functional categories." (CORRÊA; AUGUSTO; FERRARI-NETO, 2006, p. 4).

Therefore, we suggest that number, singular/ plural and numerals, is encoded and interpretable in Num, above N in the DP. Num is considered an exclusive head for grammatical number and numerals based on evidence from cognitive development (FEIGENSON; DEHAENE; SPELKE, 2004; HIRAIWA, 2017; KINZLER; SPELKE, 2007; LIPTON; SPELKE, 2005; SARNECKA; KAMENSKAYA, 2008).[53] Furthermore, we propose that Num constitutes a barrier within the DP in BP, a barrier that forms two domains (PEREIRA, 2016).

6.1.3
Two domains within the DP in BP

We adopt Pereira's (2016) analysis regarding two domains in the DP in BP, however, we suggest that the frontier that divides these two domains is actually Num, the exclusive head for grammatical number, that is, singular/plural distinction

[53] Numerals in BP are not morphologically inflected for number, but share with number inflection the semantic properties of individuation and unit: *um* ('one') conveys singular interpretation (single item) and from *dois* ('two') on, the meaning is 'more than one', plural interpretation; they agree only in gender with N. In this sense, numerals and singular/plural distinction is encoded in Num. However, there are arguments showing that numerals present an adjectival, noun-like nature. On the lines of Matushansky and Ionin (2015), cardinals may have an adjectival nature, modifying the noun. Formal analysis of the syntax-semantic interface also show that numerals may provide different information to different interfaces and that the semantic notion of number may not be morphosyntactically mapped. Ionin and Matushansky (2006b) suggest that cardinals have a lexical nature in the sense that these are nominal heads which select other nominal heads as complement; hence, cardinals actually work as modifiers of nouns. Danon, investigating Finnish and Dutch, suggests that numerals have a 'noun-like' nature: "numerals that are plural-marked (as a result of merging with Num) were shown to behave like nouns, in contrast to numerals that are not plural-marked." (DANON, 2012, p. 40). The 'noun-like' nature of numerals would distinguish them from grammatical number, which we assume to be expressed by the functional head Num. Therefore, it is proposed that numerals are not realized in Num in BP but in a different head. Ghomeshi (2003) attributes a Cardinality Phrase (CardP) for numerals in Persian in order to differentiate them from Num, also assumed by the author to be the head of grammatical number inflection. Dal Pozzo (2007) argues that, in Finnish, numerals belong to the tripartitioned class divided into quantity nouns, quantity adjectives and quantifiers. However, she argues that this discussion is an open issue. Here, it is argued that numerals modify nouns along the lines of Hurford (1998) they are simple lexical numerals..

and numerals, which convey the semantic idea of quantity. We assume along the lines of Augusto, Ferrari-Neto and Corrêa (2005) that numerals and grammatical number are realized in Num, so Num would also bear a gender feature, given that numerals in BP are inflected for gender. Therefore, we suggest the following structure for the DP in BP:

(45) [D([Poss])[Num[N]([AdjP])]]

The categories from Num up are responsible for the establishment of the reference, such as definiteness and number (see QUINE, 2013 [1960] for a discussion).[54] The semantic and pragmatic imports of reference are encoded in the upper layers of the DP: definiteness, gender and number. Considering language acquisition, children acquiring BP would have in the left-most elements of the DP, semantic and morphophonological information that would enable them to represent relevant grammatical information. Morphophonology of the functional categories above Num would be instrumental for children acquiring the number system and consequently, number agreement system in BP. Alternatively, children acquiring English would be provided with such information encoded in N and not in D. Thus, morphophonology on the noun would lead English speakers to acquire and represent relevant grammatical information pertaining to reference, that is, number, gender and definiteness in the noun. Notice, then, that there seems to be a micro-parametric distinction cross-linguistically pertaining to how and where information, such as number, is encoded across languages. Children should be able to set these micro-parameters from early ages. Children acquiring BP do so around 2 years of age.

The two domains in the DP in BP - one concerning the functional categories for the establishment of the reference and another concerning lexical information about entities involved in some event, the noun - can be formally interpreted as two phases (CHOMSKY, 2001), considering Picallo's (2017)'s. In this way, the DP in BP would have the following structure:

[54] The relation between definiteness and number is also explored in Portuguese-based creoles.[54] Alexandre and Hagemeijer (2007) assume that number is parasitic on definiteness in Saint-Tome Portuguese (STP).

(46)

[D([Poss])[Num	[N]([AdjP])]]
Functional phase	Lexical phase

So, in order to set the micro-parameter pertaining to how and where information, such as number, is encoded across languages, children acquiring BP would have to perceive morphophonological alterations in the functional phase of the DP. The procedural model of language acquisition extended in this thesis acknowledges that morphophonology is instrumental for the identification and acquisition of a language and for the representation of linguistic knowledge in terms of agreement (gender (CORRÊA; NAME, 2003; NAME, 2002) and number (CORRÊA; AUGUSTO; FERRARI-NETO, 2006; FERRARI NETO, 2008). Such assumption is in consonance with current proposals that have been highlighting the role of morphophonology and have been placing it as the *locus* of cross-linguistic variation; thus, syntax does not accommodate variation (BOECKX, 2011; BOECKX; LEIVADA, 2014). Therefore, mapping of morphophonological properties in the input is relevant for basic structure building as well as grammar identification and feature representation.

In the next section we review a formal analysis for agreement and its relation with morphophonology and integrate both accounts for the representation of linguistic knowledge regarding DP-internal agreement and subject-verb agreement.

6.1.4
DP-internal and subject-verb feature agreement in BP

As mentioned before, around 2 years of age, children acquiring BP have already identified D as the *locus* of number expression and have already mapped agreement relations within the DP by means of morphophonology. We suggest that DP-internal agreement occurs on the basis of feature sharing as proposed by Danon (2011), following Framptom and Gutmann (2000, 2006) and Pesetsky and Torrego

(2007). The present analysis differs from the one presented in Augusto, Ferrari Neto and Corrêa (2006), who assume Num to have a gender feature, given the presence of cardinals, that are inflected for gender in BP. In the present analysis, cardinals are not realized under Num, rather, Num is an exclusive head for singular/plural inflection distinction (see note 50). For this reason, we assume a different mechanism of DP-internal agreement:

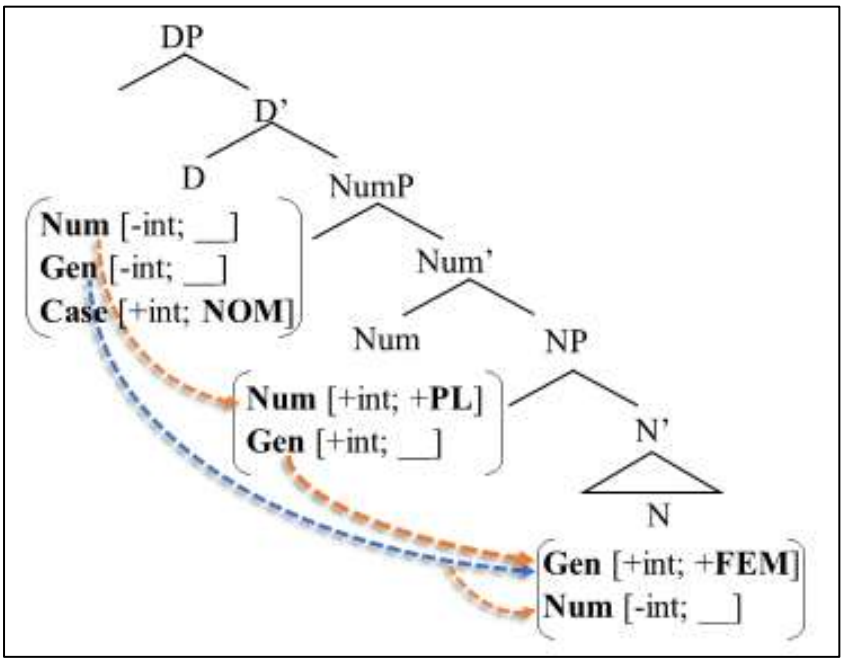


Figure 7: Analysis for DP-internal agreement in BP

D [number: -val] searches for a goal in its c-command domain and finds Num [number: +val]. In BP, articles are inflected for gender as well, so D [num: ~~-val~~; gen: -val] would also probe N [num: -val; gen: +val]. As for valuation of [number: -val] in N, we suggest, on the basis of Matushansky and Ionin's (2015) analysis, that number agreement in N is a free-rider: "number agreement on the lexical NP is contingent on an agreement relation established for another feature (number as a free-rider) and therefore conditioned by it.". Such analysis is combined with the feature sharing idea: Num probes N for gender and it can, then, share its number features with N. Interpretability of number in D [number: + int] would guarantee semantic interpretation of plural at LF in D.

For subject-verb agreement, we follow the traditional minimalist account, in which subject-verb agreement occurs on the basis of *Agree* between the DP in Spec, TP and T, in order to value Case (see Adger and Smith, 2005, p. 151-152):

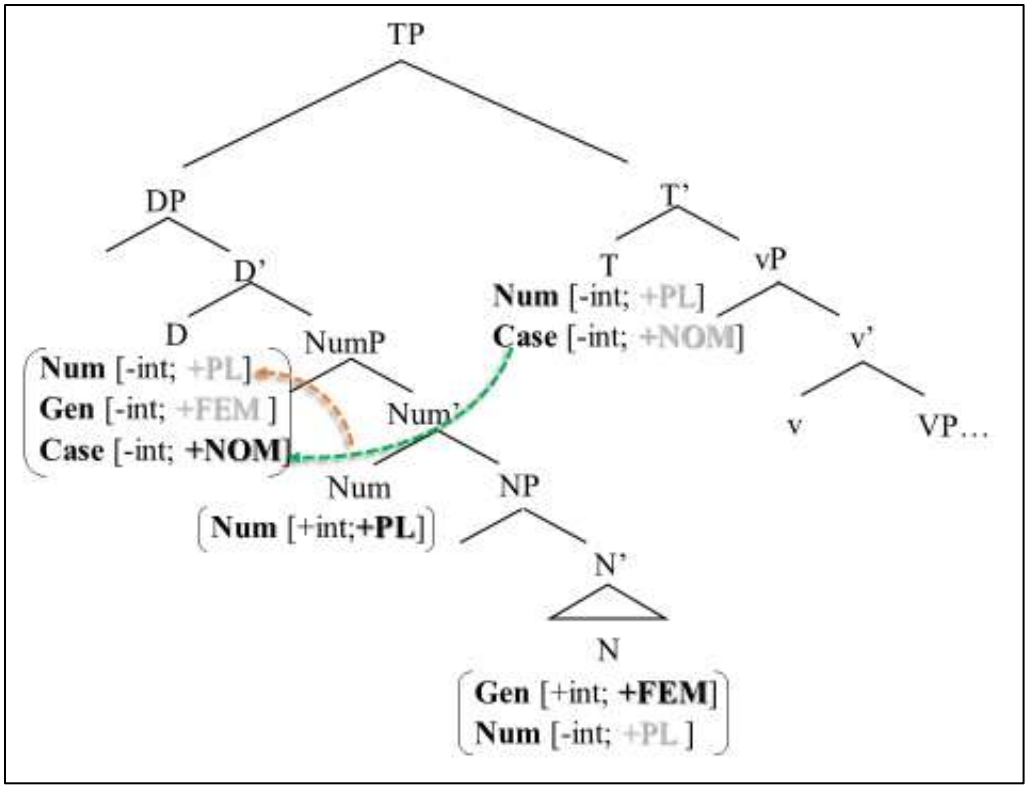


Figure 8: Proposal for subject-verb agreement in BP

T probes the DP to value Case and, in turn, number in T is valued by D; again, number in T is valued as a 'free-rider', following a feature sharing approach (PESETSKY and TORREGO, 2007) a sort of side-effect of Case valuation.

Another account for variable number agreement in the DP in Afro-Bolivian Spanish, very similar to BP, is put forth by Sessarego (2011) and Sessarego and Ferreira (2016). In their account, following Adger and Smith (2005), variation is depicted as a "by product of different lexical items entering the derivation." (SESSAREGO and FERREIRA, 2016, p. 296). Comparing Afro-Bolivian Spanish and BP variable number agreement in the DP, in which number can be expressed only in D, Sessarego and Ferreira (2016, p. 298) assume two lexical entries: 1) items lacking Num feature specification and 2) items with specified Num features – the former would surface as a default singular form and the latter would surface as plural or singular. In the present thesis, we are proposing an alternative *locus* for

variation: we do assume that variation to be a by-product of (under)specification of features in lexicon, but we propose that, for number variation in BP, underspecification is manifested specifically in morphophonological features pertaining to number.

Another account for variation in number agreement in BP is presented by Costa and Figueiredo Silva (C and FS) (2006). C and FS propose, based on DM (EMBICK; NOYER, 2001), that the plural morpheme in BP is a singleton - present in syntax and relevant for semantic interpretation -, given that it can be expressed only in D, whereas plural morpheme in EP is dissociated – inserted post-syntactically, being, thus, not relevant for semantic interpretation – that is, expressed in all agreeing elements. C and FS (2006) also suggest that Spec-head configurations would be a trigger for subject-verb agreement:

> The difference between European and Brazilian Portuguese derives from the type of morpheme associated with plurality. We contend that [plural] is a dissociated morpheme only in EP. This proposal derives the fact that all elements able to bear (subject-agreeing) plural morphology will actually carry such markers, independently of the existence of movement and spec-head configurations. In BP, on the other hand, [plural] is not a dissociated morpheme. Therefore, it will attach to the element anchoring the information concerning number. Following Enç (1991), D is the head linking the DP to its LF-interpretation. Accordingly, the plural morpheme is realized on this head, which is able to carry this marker. (COSTA; FIGUEIREDO SILVA, 2006, p. 37)

However, as sociolinguistic investigations and our results show, the production of standard and non-standard forms in BP is not categorical, rather, there is intra-speaker variation. Additionally, BP also has a number agreement variety correspondent to EP, the redundant agreement. Their analysis does not seem to account for cases in which the same speaker produces redundant (dissociated) and non-redundant (singleton) agreement. C and FS account for parametric variation, that is, cross-linguistic variation, in syntax-morphology interface.

Our analysis departs from the idea of positing two lexical entries with different specification in the lexicon, as proposed by Sessarego and Ferreira (2016). We propose that there is a single entry and variation is an exclusive product of pre and/or post-syntactic components. It also departs from C and FS's analysis to explain post-syntactic variation, given that it does not account for other types of

variable agreement present in BP, such as the sort of agreement represented as Condition 2 in the experiment (DP red/ TP n-red) and it does not account for intra-speaker variation, which is also the topic of this thesis. Instead, we offer a different analysis that would combine effects of syntactic computation in terms of *Agree* and systems of language use and processing, mainly language production. Basically, we offer an analysis that may be compatible with factors related to processing and/or online computation: the result of *Agree* in syntax may be visible at the interfaces via morphophonology.

6.1.5
A path for the acquisition of the number agreement system in BP

From early ages, children identify the morphophonological patterns pertaining to number expressed in D in BP. When they perceive that the morphemes may also appear in the noun, they start to identify agreement relations between the elements and the possibility of variation, given that sometimes plural appears in the noun, sometimes it does not. In this moment, they have already acquired that plural is necessarily expressed in D and/ or in Poss, that is, they have acquired the micro-parametric distinction in DP that plural is expressed in the functional phase of the DP.

As children grow older, they realize that the subject DP also agrees in number with the verb, so they start to extract information pertaining to number agreement and transform it into (morpho)syntactic information. Our results show that preschoolers and 6th graders provide standard responses mostly in conditions 1 (DP red/ TP red) and 3(DP n-red/TP red), thus, we assume that morphology in the verb is instrumental for acquisition of subject-verb agreement and its possibilities of variation as redundant or non-redundant.

In sum, we assume, that, initially, Brazilian children recognize and identify information pertaining to number in D, but information pertaining to number agreement redundancy is identified and recognized by means of processing of morphophonological properties in the verb, in T. There are, thus, two moments in the acquisition of the number agreement system in BP: (1) processing and

identification of number in the functional phase of the DP: in Num under D, expressed as morphophonology[55] of plural/singular in D and/or in Poss, which, in turn, leads to the identification of number agreement between pre-nominal elements and N and (2) processing and identification of subject-verb agreement by means of morphophonology in the verb.

We assume that these steps to recognize the number system and number agreement patterns in BP have already been acquired by the time children enter preschool, given that these are pre-requisites for setting micro-parameters in the number system of BP. In our experiment results, we suppose that children have already gone through this acquisition path and are actually showing how variation surfaces after the number agreement system is acquired. As shown in the results, there is a difference in standard responses provided by Pub and Pri preschoolers and differences in standard and matching responses provided by Pub and Pri 6th graders. This suggests that social factors play a role in the type of number agreement expressed in speech production. For these reasons, we believe that a model of language production, combined with the sort of knowledge of number agreement system acquired as proposed in this chapter, may shed light on the manifestation of number agreement variation in the speech of BP speakers.

6.2
An account for the production of variable number agreement

[55] Morphophonology can also be a sort of projection blocker, as proposed by Miyagawa, Wu and Koizumi (2018). The authors propose a new perspective on Problems of Projection (CHOMSKY, 2013) by assuming interference of morphological marking for projection labelling. The presence of morphological markings would block an element to be projected, and, thus, labelling the projection. Assuming Miyagawa, Wu and Koizumi (2018)'s proposal for BP, the raising of the subject DP out of *v*P to Spec, TP results in φ-feature agreement, in which the probe T, has its number feature valued by D. This φ-feature agreement is what allows the merged pair (DP, T) to be labelled as TP. Additionally, the obligatory morphological plural marking in D in BP could work as a projection blocker. Miyagawa, Wu and Koizumi (2018) propose an analysis in which morphology and syntax interact, however, further analysis is needed to verify whether the obligatory morphological marking in D in BP could work as a projection blocker. If so, morphophonology would also have an instrumental role in basic structure building, such as projections. Further research is necessary to verify this possibility.

In the previous chapters, we have seen that syntax does not accommodate variation and that children are sensitive to (morpho)phonological information in the input. Additionally, morphophonology is assumed to be a sort of trigger for language acquisition and interacts with properties of basic structure building in the syntactic component. Furthermore, we have shown that children are sensitive to variation in the input and are able to reproduce such variation. In the previous section, we proposed an account for language acquisition of number agreement that leads to the grammatical identification of BP, a sort of parameter setting for the number agreement system in BP. In this section, we aim to account for variation in number agreement redundancy in language production based on the linguistic knowledge the speaker possesses.

We rely on Levelt (1989)'s model of production and its levels: message conceptualization, grammatical encoding, morphophonological encoding, phonological encoding and articulation. We propose that variation emerges either during message conceptualization and/or during morphophonological encoding.

As for grammatical encoding, we adopt a different view from Levelt's by assuming a top-down and bottom-up hierarchical structure building based on Corrêa (2002, 2009, 2014) and Corrêa and Augusto (2007, 2011)'s integrated model for online computation (MINC – *Modelo Integrado de Computação Online/* Integrated Model of Online Computation). This model integrates the derivations of a minimalist theory of language and procedures implemented during grammatical encoding and parsing.[56]

6.2.1
An online computational model for grammatical encoding

[56] Levelt (1989) assumes a lexical-based syntax, that is, syntactic structures are formed based on lemmas retrieved from the lexicon. Other syntactic models such as the Incremental Procedural Grammar (KEMPEN; HOENKAMP, 1987) and the Tree-Adjoining Grammar (FERREIRA, 2000) as well as the syntactic frames (DELL, 1995) and others are also used in psycholinguistic models for speech processing. In this thesis, the online computational model proposed by Corrêa and Augusto (2007, 2011) is adopted, given that this model is able to account in details for number agreement in BP associated with a language acquisition model proposed by Corrêa (2009, 2014) which proposes grammar identification in terms of features representation due to processing of morphophonological information in PF interface.

For online language production, we assume, unlike standard minimalist assumptions, that computation in actual language use occurs top-down, left-to-right (CHESI, 2007, 2012; CORRÊA and AUGUSTO, 2007, 2011).

According to MINC, speakers' speech intent guides the retrieval of lexical items. The array of lexical items retrieved from the lexicon drives the directionality in structure building: intentional meanings pertaining to reference to events and entities are associated to functional categories which are projected in a top-down fashion based on the speakers' speech intent; in parallel derivational spaces, lexical categories related to conceptual systems, containing semantic content and information on argument position in the structure, are bottom-up projected. In MINC, CP, TP and DP are built online in a top-down manner, while NP, VP, AdjP, AdvP and others are bottom-up derived and attached to the structure of functional categories according to structural position information encoded in features:

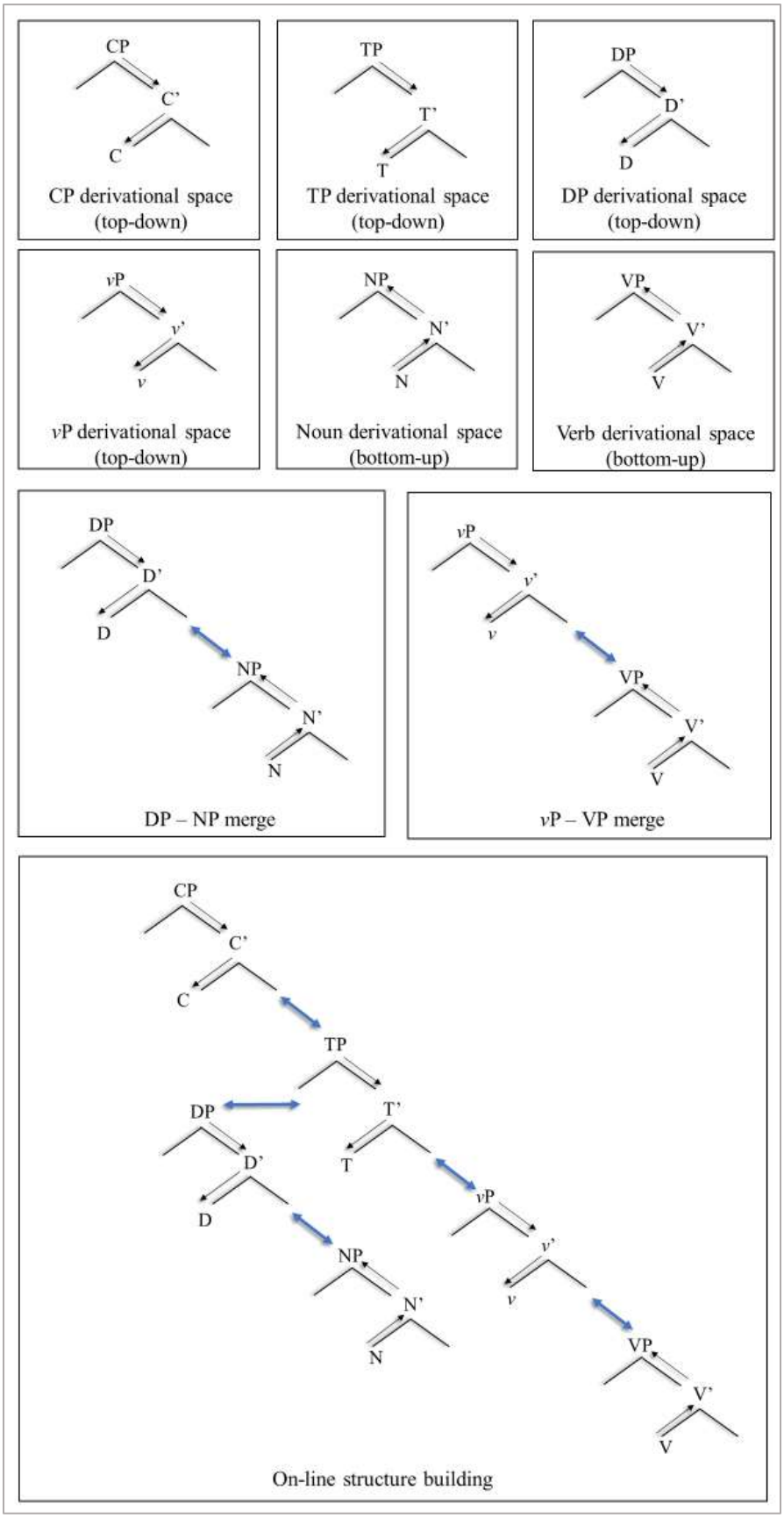


Figure 9: Online structure building in MINC (Adapted from figures 1 and 2 in CORRÊA; AUGUSTO, 2011, p. 479)

The figure above illustrates online structure building of constituents in a hierarchical structure. In language production, the DP does not have to be generated as the external argument of the VP; in actual language use, the speaker knows, due to his/her linguistic knowledge, that a subject is composed of a pronoun or a determiner and a noun, which is governed by a functional category, DP. So, properties on message conceptualization guide such structuring and ordering and morphophonology regulates functional categories to be projected. Thus, message conceptualization licenses the elements of a subject DP to be built in a parallel derivational space with valued features. Afterwards, DP and NP are merged and φ-feature agreement computation takes place, giving rise to DP internal-agreement:

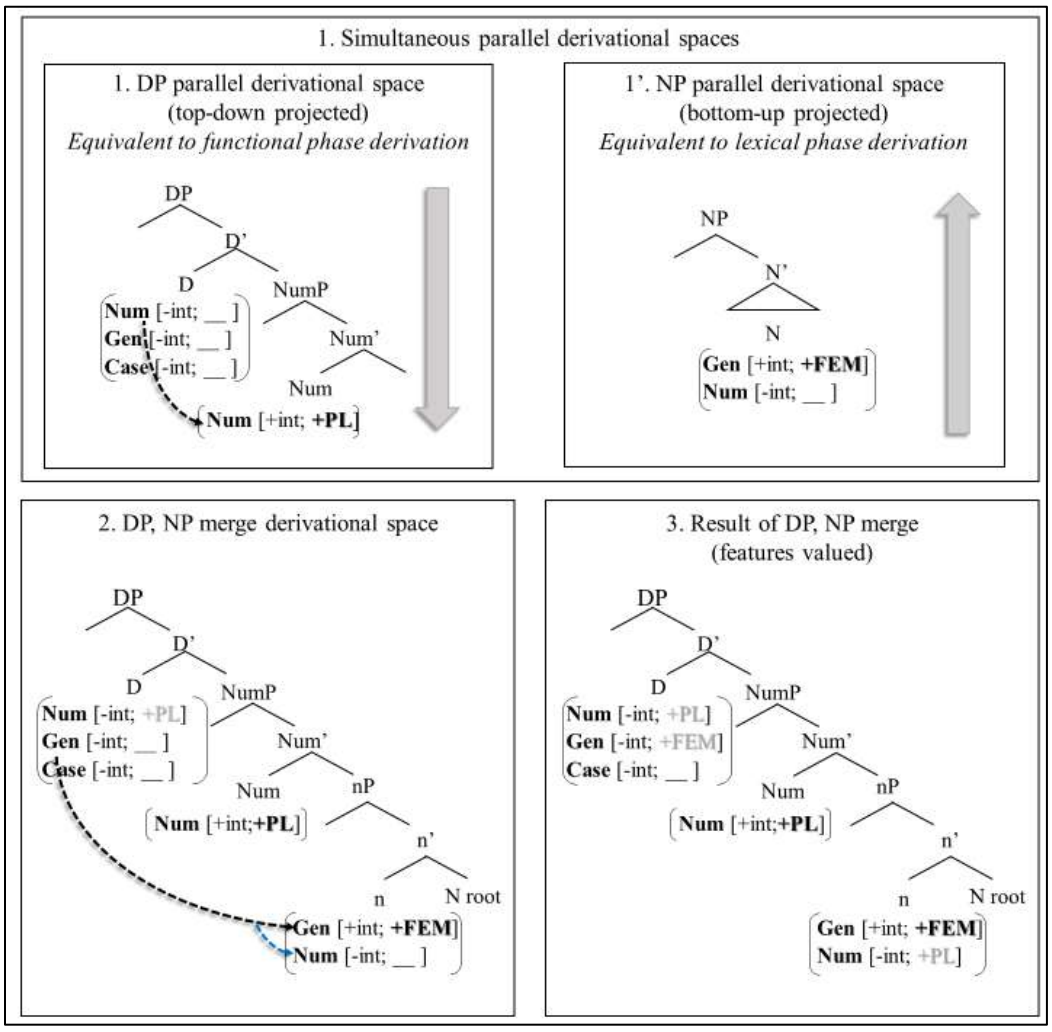


Figure 10: DP online computation

Hence, number agreement is first established between D and Num in an online model in language production - remember that D and Num form the DP domain equivalent to the functional phase, and N, equivalent to the lexical phase of the DP. When the DP is ready with features valued, it is attached to Spec TP for subject-verb agreement. We reinforce that for such computation to take place, micro-parametric properties have to be established *a priori*, during language acquisition. This sort of DP computation assumes that the speakers has already established the functional phase of the DP as the necessary *locus* for plural marking and that D (or Poss) work as a projection blocker in BP. Therefore, the linguistic knowledge in terms of BP grammar would guide such online computation.

As for subject-verb agreement, in MINC (CORRÊA and AUGUSTO, 2007, 2011), the subject DP does not move to Spec, TP as in the minimalist derivation. In a production model, the subject is built already in its position, thus, Case is valued by T right away.

With all the features valued, we rely on a top-down phase-based theory (CHESI, 2007, 2012) for externalization to the morphophonological encoding, where morphemes are inserted.[57] Chesi proposes that the notion of *phases* (CHOMSKY, 2001) may be used to guarantee the left-to-right directionality in language production. Assuming that the DP works as a processing unit equivalent to a phase (AUGUSTO; CORRÊA; FORSTER, 2012), we suggest a top-down, two-phase transfer of the DP in BP to the morphophonological component in language production:

[57] Chesi (2007, 2012) proposes a top-down model computation in actual language use, whereas Corrêa and Augusto (2007, 2011) propose a mixed model, in which functional categories are top-down projected and lexical categories are bottom-up projected. In this thesis, we argue for a mixed directionality regarding projection of categories on the lines of Corrêa and Augusto but we suggest a top-down phase-based transfer to the interfaces, based on Chesi.

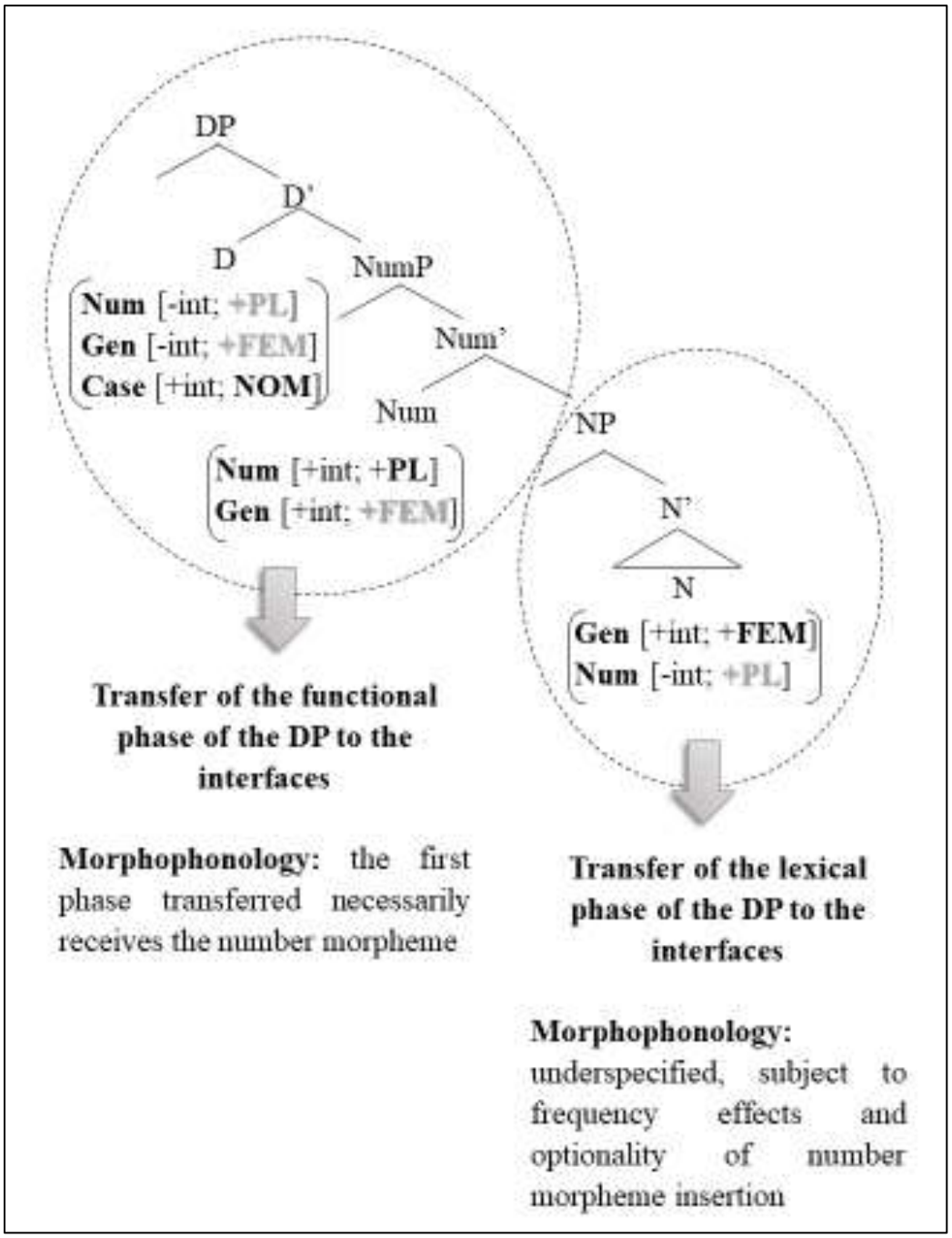


Figure 11: Online phase-based transfer of the functional phase of the DP in BP during language production

When transferred to the interfaces, the functional phase would signal to the morphophonological encoding, that a number morpheme should be inserted.

There has been much debate on how many chunks of linguistic elements should be planned prior to its articulation. Lee *et al.* (2013) propose that lexical planning is mediated by hierarchical relationships and a very high degree of look-

ahead planning, suggesting that speakers decide for a structure prior to speaking. Their results converged when participants had time to prepare the speech and when they did not have time for preparation. Corrêa, Augusto and Marcilese (2018) found similar results with relative clauses in BP in which a fully planned condition facilitates reliance on standard relative clauses, which are more demanding in terms of processing cost, whereas partially planned conditions trigger non-standard relative clauses production, less costly than standard ones.

Such brief review shows that discovering how much of the utterance has to be planned to initiate articulation is not a trivial issue and it would depend on research in prosody and phonetic/phonological aspects of language production and comprehension. As this is not the scope of this thesis, but it may well be a topic for further inquiry in order to improve an online computational model, we assume, for the purposes of this thesis, that a unit or "chunk" to be planned in advance is the size of a subject + verb +object structure, given the affirmative intonation of the sentences. This reasoning allows us to assume that the sentences used in the experimental tasks are considered a chunk, once they belong to the same intonational unit.

6.2.2
An account for language production

We base our account in Levelt's (1989) language production model, with few adjustments, mainly in grammatical encoding and morphophonological encoding. We refine the analysis for morphophonological underspecification presented in chapter 3 by suggesting an interaction between grammatical encoding, presented in the previous section, and morphophonological encoding:

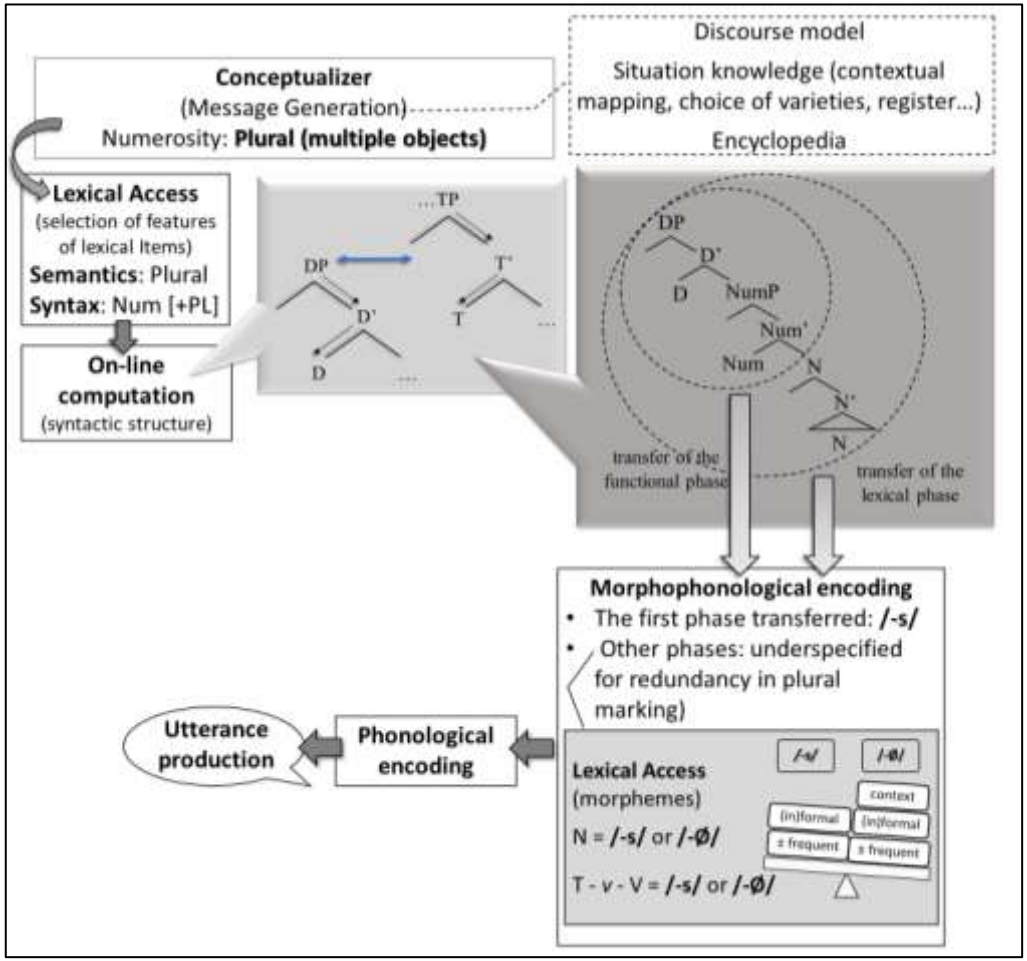


Figure 12: Production model of the DP in BP

The model in Figure 13 combines the literature reviewed so far with the results obtained. In language production the speaker initializes the externalization of an utterance from message conceptualization, in which main features of the events are mapped such as the entities involved in an event and fine-grained characteristics such as numerosity, definiteness, analysis of the context, of the interlocutor and so on. In this level, different cognitive resources are allocated in order to map different types of information. Contextual and informational mapping leads to lexical access in a way that nouns represent entities and verbs describe events, for instance. However, such an access is not to lexical items, but to features that encode the message to be conveyed; at this moment, morphophonological features are not accessed, but only the variety (standard or non-standard) is chosen. After featural selection, syntax is able to proceed with structure building in terms of functional and lexical categories, based on MINC. After structure building (CORRÊA; AUGUSTO, 2007, 2011; CORRÊA; AUGUSTO; MARCILESE, 2018), phase-based transfer takes place (CHESI, 2007, 2012) and the syntactic objects are transferred to morphophonological encoding. The functional phase of the DP necessarily receives plural morphemes and the other agreeing elements are subject to variation given its morphophonological underspecification.

If the speaker does not encode a certain variety (standard or non-standard) during message generation, underspecified morphophonological features (± redundant) will be subject to frequency and processing effects during morpheme insertion (-s or -Ø) in morphophonological encoding, since underspecification means no fixed value. If the speaker makes a choice of variety (standard) during message generation, such choice will activate a specified type of agreement (+redundant), which results in plural morpheme insertion in all agreeing elements during morphophonological encoding (see table 14). Different factors interfere on the weight (YANG, 2002) each variety possesses in a grammar, resulting in different scenarios. Based on the production model presented, we suggest different *loci* for variation in number agreement in BP.

6.3
Loci for variation: One grammar or two grammars in BP?

As discussed in chapter 3, we propose that children growing up exposed to BP must acquire that plural marking is necessarily expressed in the functional phase of the DP and that morphophonological expression of number agreement redundancy is represented as underspecified in the speakers' linguistic knowledge. Factors such as literacy and social context (SES), as our results show, would act as forces attracting or repelling the possibilities of number agreement redundancy:

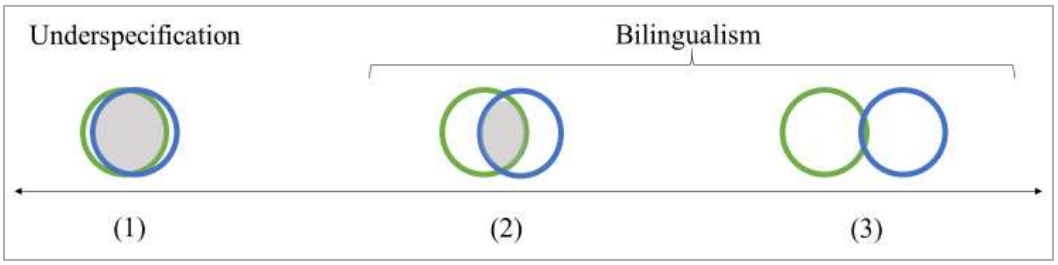


Figure 13: Proposed multilingual *continuum* in terms of grammatical representation

The context in (1) would be representative of underspecified morphophonological expression of number agreement redundancy in which redundant and non-redundant plural marking overlap; (3) would be representative of the possibility of choosing one variety or the other, redundant or non-redundant; context (2) would be representative of an in-between situation with ongoing process of variety specification. We suggest that in contexts (2) and (3), variation can be located pre-syntactically, that is during message conceptualization. Mapping of the communicative context and type of situation and interlocutors influence the choice of variety and the morphophonological content to be accessed during morphophonological encoding. Alternatively, in context (1), variation emerges post-syntactically, subject to frequency effects during morphophonological encoding. The experimental results obtained and the theoretical discussion provided shows that BP may well exhibit one underspecified grammar or different grammars regarding number agreement redundancy. Additionally, we suggest, based on the language production model presented in this thesis that variation may have two *loci*: pre-syntactic, in the choice of variety to be used according to contextual and social information and post-syntactically, when morphophonological features are

accessed and morphemes are inserted. We suggest, based on our results, that the choice of variety is regulated by the level of management of varieties in terms of mapping of contextual information a speaker possesses:

Speakers' ability to map contextual and linguistic information	**Pre-syntactic**	**Specification of features**	**Post-syntactic**
	message generation (possible choice of variety)	morphophonological features corresponding to varieties	morphophonological encoding
Non-proficient	-	± redundant	-s or -Ø
Proficient	Standard	+ redundant	-s
	non - standard	± redundant	-s or -Ø

Table 21: Possible *loci* of variation depending on speaker's proficiency of variety use

Considering the table presented above, BP could present both one grammar with variation (underspecified grammar) or two grammars (when the variety is specified pre-syntactically). Such an account allows for different dynamic configurations that may change or adapt during language development, ranging from an underspecified to (a) possibly specified grammar(s). We also propose that such configurations are not limited to those three forms; there is a *continuum* in which several different configurations could be represented (see GONÇALVES, 2009). An analysis in terms of a multilingual *continuum* is proposed in the next chapter.

7
Multilingual contexts within variable number agreement in BP

Based on chapters 2 and 3, there seems to be a language contact situation as far as varieties of number agreement in BP are concerned, given the history of the country. Language contact or variety contact, considering the possibilities of number agreement, yields variation and consequent exposure to variable input. Our main thesis is that exposure to variable input results in underspecification, particularly in the case of number agreement in BP, more precisely, underspecification of the morphophonological expression of number agreement redundancy. In underspecified grammars, there is no clear boundaries between varieties, different options are available. However, consistent exposure to standard BP, related to literacy/ schooling, and social factors, such as SES, yields a sort of bilingual context, with different degrees of dominance of each variety, distinguishing speakers in terms of a multilingual *continuum*: a) highly-educated speakers would establish redundant agreement as their main variety, non-redundant agreement would be accessible in language comprehension; thus, these speakers would be functionally similar to passive bilinguals (see CORNIPS, 2014 for the concept of "passive bidialectals") as far as non-redundant agreement is concerned; b) moderately-educated speakers would keep both varieties with different levels of dominance depending on contextual information; these speakers would exhibit code-switching abilities and switch between variable production, in which both standard and non-standard varieties are accessible, and the redundant variety, similar to L2 speakers.

Kato (2017) assumes the perspective of code-switching in BP for the case of clitics that are not naturally acquired in BP but only by means of formal instruction at school; the nuclear grammar of the speaker would consist of null objects and lexical pronouns, whereas the clitics would be located in its periphery:

> The I-language of the literate speaker of BP should contain a great component of his/her nuclear grammar, as in child language, which is represented in informal oral language as a left-cliticization process or with alternatives as lexical pronouns

> and, mainly, null objects. […] In written modality, though, the literate adult keeps a periphery that consists of an older grammar with left-cliticization.[58] (KATO, 2017, p. 149 – my translation, my highlights)

Notice that in the excerpt, Kato proposes that the nuclear grammar has different alternatives: lexical pronouns and null objects and left-cliticization in informal oral language. A grammar that contains two or more accessible options naturally acquired as proposed by Kato corresponds to the analysis presented in this thesis as underspecification, an underspecified grammar. We develop such analysis if bilingualism within variable number agreement in BP in the following pages.

The results obtained in the experiments show a *continuum* in which plural agreement redundancy is influenced by schooling, SES and academic performance in the case of 6th graders. Priv preschoolers may follow two paths after schooling: one similar to Priv or another one similar to Pub A 6th graders, given that they already show a tendency to prefer redundant forms. Pub preschoolers seem to show a path in which they are subject to frequency with a great level of variation in their production, so, they may follow any path after schooling: similar to Priv and Pub A 6th graders, improving their sensibility to redundant forms of number agreement or similar to Pub B maintaining the level of variation in their production. Considering the relation between SES, schooling/ literacy and academic performance can be illustrated as follows:

[58] Original excerpt: "*A língua-I do brasileiro letrado deve ter, portanto, um bom componente da sua gramática nuclear, como na criança, que se manifesta principalmente na língua oral informal, com um processo de cliticização à esquerda ou com alternativas via pronomes retos e, principalmente, objetos nulos [...] Na língua escrita, porém, o adulto letrado terá uma periferia que consiste de processos de uma gramática mais antiga, como, por exemplo, o direcionamento de cliticização à esquerda*" (KATO, 2017, p. 149)

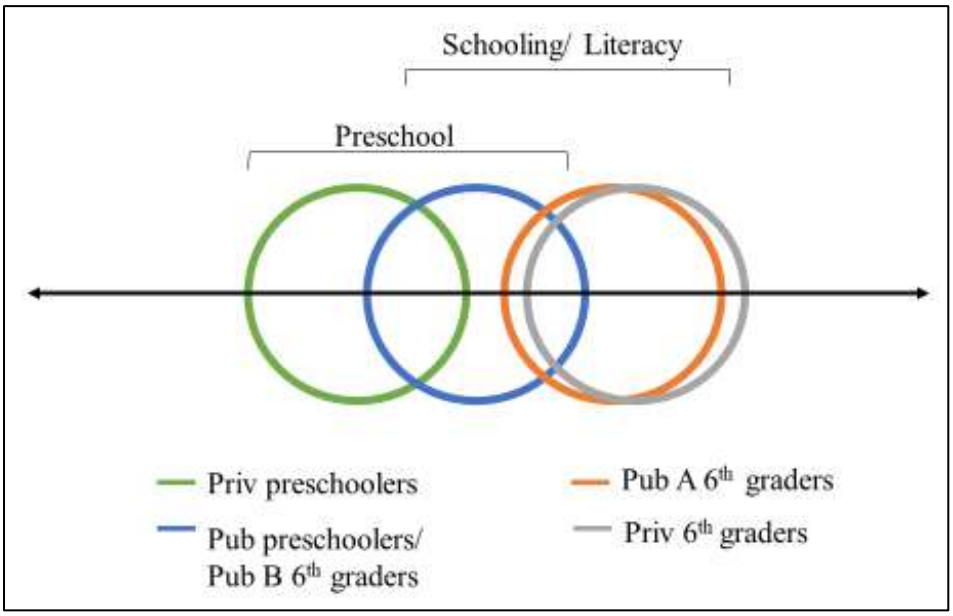


Figure 14: *Continuum* of number agreement redundancy based on preschoolers and 6th graders matching responses

Therefore, schooling and SES are intersected, as well as the forms of number agreement redundancy: preschoolers exhibit different tendencies of preference, but they all seem to acquire both standard and non-standard varieties; the frequency of non-standard forms varies according to SES. In the 6th graders' data, SES still plays a relevant role in terms of the production of non-standard varieties, but other factors, such as overall academic performance, seem to be more prominent, leading 6th graders from low SES to achieve similar results to higher SES, as in the case of Pub A. Overall academic performance seems to guarantee morphological awareness and, consequently, enhance code-switching abilities. The following matrix shows the intersection of SES and academic performance combined with 6th graders' results:

	Low SES	**High SES**
Below-average academic performance	High levels of variation; Higher rates of production of non-standard varieties	No data
Above-average academic performance	Scarce variation; Preference for the standard variety; Enhanced code-switching abilities	Extremely rare variation; Clear preference for the standard variety;

	Judgement of non-standard varieties as "wrong language"

Table 22: Matrix of SES and academic performance intersection

There is no data to support below-average academic performance in high SES speakers. This is a direct consequence of social and educational systems. Few students present below-average academic performance in high SES schools. As shown before, there are outliers in the standard variety both in Pub and Priv groups. The number of outliers is very small, so it would not be possible to form a group in order to assess their performance. Additionally, there is an increasing debate in pedagogical areas about the tracking system, its advantages and disadvantages and if it helps or segregates students or whether schools increase or decrease inequality (DUFLO; DUPAS; KREMER, 2008). In this thesis, academic performance and literacy was only observed as an effect due to a particular school that adopts the tracking system. However, it is not guaranteed that most schools adopt this system. Priv, for example, does not follow such system. Another interesting fact in this scenario is that most of the elementary school teachers in Brazil are generally from low-to-middle SES and are speakers of both varieties (BAGNO, 2006, 2007)[59]. Hence, it becomes a challenge to define what is exactly the nature of the BP being learned at school.[60] For this reason, many aspects of language contact and multilingualism may be found in varieties contact.

Therefore, it is quite complex to disentangle the boundaries among all these external factors. Figure 14 reinforces the fact that number agreement redundancy variation in the suburban area in Rio does not present clear-cut boundaries, coinciding with bidialectalism literature findings (CORNIPS, 2014; LEIVADA *ET AL.*, 2017). Another interesting fact highlighted by the results found is that

[59] Interview: https://www.brasildefato.com.br/node/5396/. See also UNESCO report (2004) on Brazilian teachers' profile available at: http://unesdoc.unesco.org/images/0013/001349/134925por.pdf (accessed on 20/03/2018) and the text for discussion on teacher's of elementary education in Brazil by IPEA (Instituto de Pesquisa em Economia Aplicada – Brazilian Institute for Research in Applied Economics) (2017): http://www.ipea.gov.br/portal/images/stories/PDFs/TDs/td_2304.pdf (acessado em 20/03/2018)

[60] Also, BP seems to be undergoing a diachronic change in terms of pronominal and verbal paradigms and plural markers. It may be a possibility that schooling, by introducing the standard written/ variety, is actually working as a pressure against language change (Lopes, 2014).

schooling actually sharpens the difference between the use of varieties. After schooling, the quantitative and qualitative differences are more prominent between groups when comparing preschoolers and 6th graders. In this sense, schooling, also considering academic performance, may improve code-switching abilities, as in the case of Pub A, or it may not have any impact on speech production, as in the case of Pub B. Thus, schooling may boost metalinguistic awareness. It is important, though, that, during schooling, Brazilian students become aware of the fact that both varieties co-exist in BP (BAGNO, 2005).

Comparing experimental data obtained in this thesis with spontaneous speech production of adults in the sociolinguistic literature, the following distribution is presented:

+ agreement ◄				► - agreement
		RJ Urban area		
97,8% - 94% *Highly Educated* (Graciosa,1991; Vieira and Bazenga, 2013)	89% - 73% *Moderately Educated* (Scherre and Naro,2006; Vieira and Bazenga, 2013)		48% *Illiterate* (Naro,1981)	
90,5% *Priv 6th*	67,0% *Pub A 6th*	44,4% *Priv Pre*	36,1% *Pub B 6th*	29,3% *Pub Pre*

Table 23: Comparison of adults' spontaneous speech results from Vieira (2015) and experimental results with preschoolers and 6th graders

It is relevant to highlight that, although Vieira (2015) does not indicate whether the subjects of 3rd person plural verbs in her data are pronouns or full DPs, the distribution of percentages are very similar to the results obtained in this thesis regarding 3rd person DP subject. Comparing the distribution, illiterates and preschoolers would exhibit similar performance, what makes sense given that both populations were not exposed to formal written standard BP at school. The striking result is that a group of 6th graders exhibit similar performance to those never exposed to written/ standard BP, Pub B. It is alarming how environmental issues such as SES and academic performance affect linguistic development. 6th graders from low SES, but good academic performance, are mainly similar to moderately-educated speakers in terms of production of plural agreement markers. Finally, 6th

graders from higher SES stand close to the highest level of agreement of moderately-educated speakers and highly-educated speakers.

Observing number agreement in the nominal domain, the following picture emerges for the groups:

Group	DP-internal agreement redundancy
Priv Pre	67,2%
Pub Pre	47,7%
Priv 6th	91,5%
Pub A 6th	76,8%
Pub B 6th	49,3%

Table 24: Distribution (%) of DP-internal agreement redundancy per group

Pub Pre and Pub B 6th present the lowest distribution of nominal agreement redundancy, close to 50%, around chance. Following them, Priv Pre and Pub 6th show distribution around 65 – 75% and, finally, Priv 6th graders exhibit the highest percentage of nominal agreement, but it does not reach ceiling, 91,5%. The next table presents subject-verb agreement redundancy in relation to the DP redundant subject:

Group	Subject-verb agreement redundancy
Priv Pre	54,4%
Pub Pre	53,5%
Priv 6th	94,0%
Pub A 6th	75,0%
Pub B 6th	60,6%

Table 25: Distribution (%) of subject-verb agreement redundancy per group

The distribution of subject-verb agreement differs from DP-internal agreement. Regarding subject-verb agreement, preschoolers exhibit similar percentages, less than 55%. The distribution of agreement redundancy in the subject-verb domain grows with 6th graders, varying from 60 – 75% in Pub 6th graders and reaching 94% in Priv 6th graders' production. Notice that for all the groups, the distribution of subject-verb agreement is higher than the percentage of DP-internal agreement redundancy, except for Priv Pre (nominal: 67,2% and

verbal: 54,4%) and for Pub A 6th (nominal: 76,8% and verbal: 75%). For Pub A 6th, the difference is bigger, around 13%, thus, for this group, DP-internal agreement seems to be slightly more relevant than subject-verb agreement, unlike the other groups.

7.2
A multilingual *continuum* in BP

In this thesis, we suggested that an underspecified grammar is the natural state of BP grammar concerning number agreement and that exposure to formal teaching and awareness of the social stigma on the '-redundant' variety and the prestige of the '+redundant' variety may alter the configuration of grammar(s). The distinction in terms of a *continuum* from underspecification to specification results in a multilingual *continuum*. This *continuum* would account for different levels of variation and also different contexts of varieties in contact. Furthermore, such *continuum* would dispense with the debate on whether variation is represented by a single grammar or by two grammars: both possibilities are entertained depending on the interaction of linguistic and environmental aspects and speakers' awareness of the use of each variety.

According to the sociolinguistic literature, BP presents different *continua*: a) a *continuum* from rural to urban (NARO, 1981); b) a *continuum* inside the urban speech (SCHERRE; NARO, 2006; VIEIRA, 2015). In this thesis, we add a third one: c) a *continuum* inside the suburban speech in Rio as far as preschoolers and 6th graders are considered, as suggested by the experimental results. Therefore, a plural number morphology scale seems to be shared among different varieties of spoken BP.

The preference for the use of one or the other variety may be altered due to schooling as a tendency to rescue grammatical aspects that are not anymore naturally acquired. Thus, schooling may mask an ongoing change, as argued by Pires, Rothman and Santos (2011). As verified in Pub A and Priv 6th graders results, training in standard written BP may change the frequency rates for each variety for each child. During school years, children would alter the configuration of their PoV,

recognizing the social value of each variety, hence, the more children are aware of social value, the more code-switching is expected. Pub A and Priv 6th graders have stronger code-switching abilities between varieties: they recognize the existence of non-standard varieties and even produce it in different rates, but they are able to inhibit them and produce the standard variety more frequently. Differently from these groups, Pub B tends to code-switch less, exhibiting a sort of levelled production in which standard and non-standard varieties compete. These different performances seem to be directly related to literacy skills and overall academic performance (see CRAIG; WASHINGTON, 2004; TERRY, 2012; WHEELER; SWORDS, 2004 for a discussion on African-American English). The following table provides both the linguistic and social *continua* regarding number agreement in BP based on the experimental results obtained:

+ agreement ←				→ - agreement
Social *continuum*				
Priv 6th	Pub A 6th	Priv Pre	Pub B 6th	Pub Pre
Linguistic *continuum*				
DP red/ TP red	DP red/ TP red	DP red/ TP red	DP red/ TP red	DP red/ TP red
DP n-red/ TP red	DP n-red/ TP red	DP red/ TP n-red	DP n-red/ TP red	DP n-red/ TP n-red
DP red/ TP n-red	DP red/ TP n-red and DP n-red/ TP n-red	DP n-red/ TP n-red	DP n-red/ TP n-red	DP red/ TP n-red
DP n-red/ TP n-red		DP n-red/ TP red	DP red/ TP n-red	DP n-red/ TP red

Table 26: Social and linguistic *continua* for number agreement in BP based on preschoolers' and 6th graders' production

The expression of number agreement in BP does not seem to involve discrete, separate grammars (CORNIPS, 2014); rather, there is a *continuum*, forming gradience in the production of morphological number agreement (see MARINS; SOARES DA SILVA; DUARTE, 2017 for a discussion of a gradient analysis in terms of a continuum in the variation of null subjects in BP). The notion of nuclear grammar and marked periphery for literate speakers, as proposed by Kato (2005), is contemplated in the cases of L2 bilinguals and different degrees of proficiency. Thus number agreement in BP constitute closely-related (see GROHMANN *et al.*, 2016; ROTHMAN, 2011; ROTHMAN; GIANCASPRO; HALLORAN, 2014; TSIMPLI, 2014; WESTERGAARD *et al.*, 2017 for discussions on typological or structural proximity), similar to a condition of *diglossia* (FERGUSON, 1959) and/or bilingualism (FISHMAN, 1967).

The production model offered in this thesis, locates variation in the storage of grammatical information in the lexicon and in the access to this information during morphophonological encoding, depending on the level of (under)specification of morphophonological features. In the following table, we summarize our findings, considering different multilingual configurations:

	Grammar *status*	*Locus* of variation	Multilingual *status*
Preschoolers	• Underspecified morphophonological features for the expression of number agreement redundancy: [± redundant]	• Post-syntactic, in the morphophonological encoding • Subject to frequency and language-external effects	• Simultaneous Bilinguals • Possibly the monolingual state of BP in terms of number agreement redundancy
Priv 6th graders	• Two grammars, being [-redundant] identified in language comprehension and [+redundant] preferred in language production	• The variety spoken comes specified from the lexicon in the moment of message planning, a pre-syntactic stage	• Passive Bilinguals
Pub A 6th graders	• Two grammars, being the underspecified [-redundant] available for comprehension and production in certain contexts, maybe family and/or informal environments and [+redundant] used in formal contexts and/or at school	• The variety spoken comes specified from the lexicon in the moment of message planning, a pre-syntactic stage depending on the context of interaction	• Proficient L2 bilinguals (greater code-switching ability) • Heritage speakers
Pub B 6th graders	• Underspecified morphophonological features for the expression of number agreement redundancy: [± redundant]	• Post-syntactic, in the morphophonological encoding • Subject to frequency and language-external effects • Eventual recognition of the necessity to use [+redundant] according to the contextual situation, but do not exhibit abilities in doing so	• Low proficient L2 bilinguals (poor code-switching ability)

Table 27: Multilingual configuration for preschoolers and 6th graders exposed to variable number agreement redundancy in BP

Considering the results obtained and the sociolinguistic literature on number agreement in BP, such phenomenon is better described in terms of discrete, gradient *continua* rather than categoric distinction between redundant and non-redundant agreement.

7.3
Impacts of a multilingual *continuum* for language ecology

Based on our results for number agreement in BP, we propose a broader hypothesis on language variation and language contact, an issue brought up by Weinreich (1970). We assume that interference and variation is expected to appear when there is no clear boundary between linguistic varieties, that is, when the varieties share a great area of intersection, which we are proposing in terms of underspecification. Linguistic varieties may be distinguished depending on environmental factors such as need of use of each variety, context of use, schooling, level of proficiency in each variety, academic achievement, family environment, frequency factors. The predictions of such assumption are: (i) if such factors act with greater intensity in speakers' environment, varieties can be easily accessed separately, improving code-switching skills, allowing the speaker to easily change varieties as the context requires; (ii) if such factors act on the linguistic varieties with less intensity, the boundaries are clearer than in an underspecified context but not clear enough to be easily accessed separately, thus, some interference or variation is expected and code-switching abilities may not be effectively implemented. These multilingual configurations are dynamic in the sense that different configurations may exist in-between these other settings and a single speaker may exhibit different configurations of multilingualism in different moments, depending on the environmental forces acting upon the linguistic varieties:

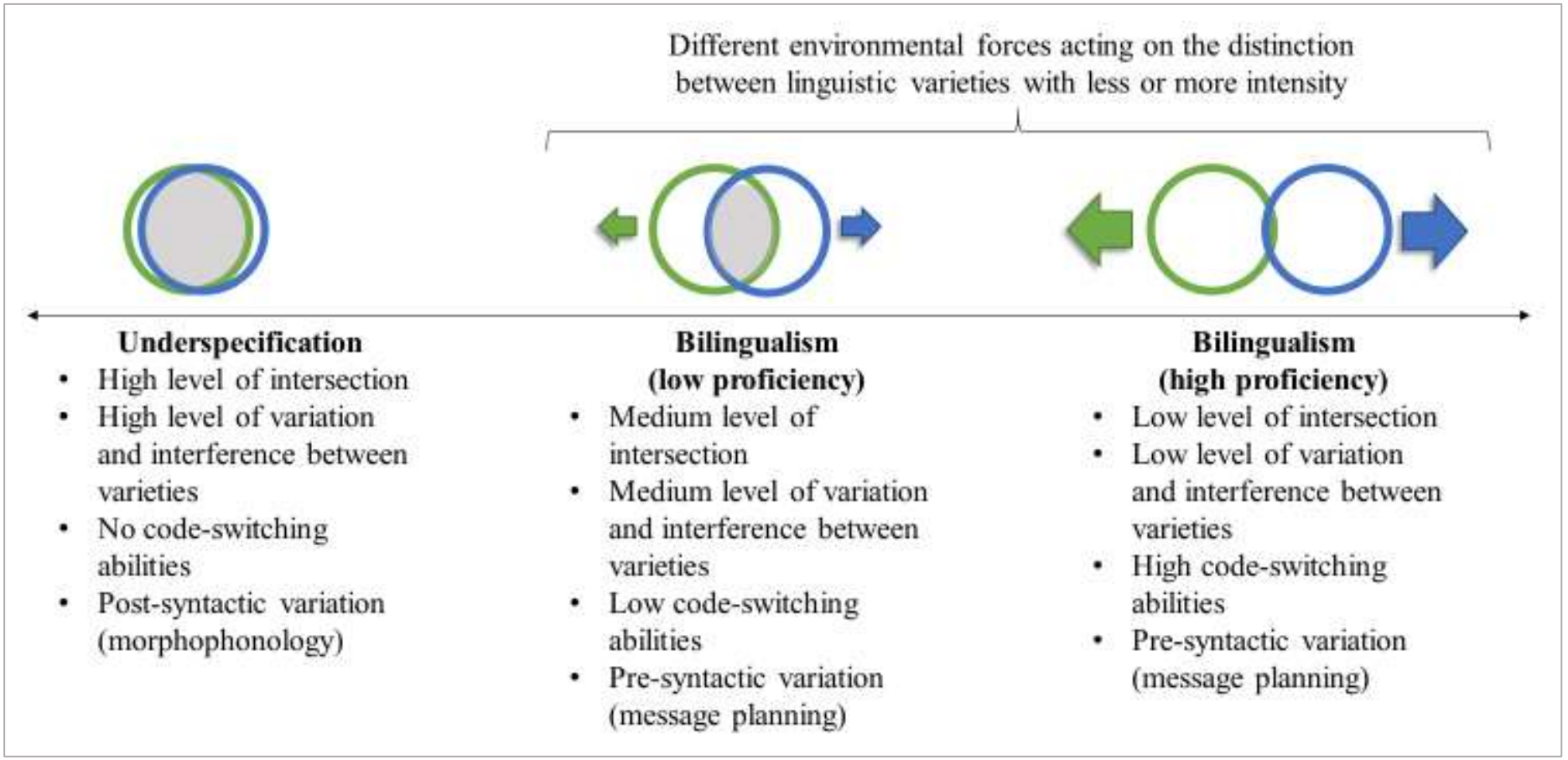


Figure 15: A proposal for a multilingual *continuum* and contexts characteristics

The possibility of combination between two grammars and one grammar with variation allows to account for the gradience in production within a multilingual *continuum*. Additionally, such combination highlights that language contact, within and across speakers may be realized in different manners, resulting in different arrangements of multilingual populations.

In sum, contexts of language contact exhibit embedded characteristics of each language contact situation, such as bilingualism, bidialectalism, monolingualism. Contact among varieties in multilingual settings influences the frequency of the varieties used by each speaker and contributes to the overall frequency in the speech community. Consequently, it contributes to the natural dynamics of the ecology of language as highlighted by Mufwene (2001, 2008).

Though Mufwene proposes the *feature pool*[61] idea for language evolution, a parallel situation may be established in case of synchronic variation that may lead to a possible change in language. The *feature pool* would represent the set of all possible preferences of all speakers in the speech community, given that these speakers interact with each other in a social network. This pool would serve as input for children to acquire the features of the target-grammar. In relation to the downwards arrows, during language acquisition, children may select any of the features in this pool in a way that their grammar may not be represented as their parents' grammars (HENRY, 1998).

In a simulation of linguistic change through generations, Kenny Smith and colleagues (2017) show that speakers reproduce the variation frequency rates in the input, however with tendency to overgeneralize one of the forms. These tendencies, in turn, are transformed into statistical tendencies that affects the production of the next generation, who would identify and exaggerate these statistical tendencies in a snowball effect as the authors highlights: "In this way, initial 'errors' in reproducing the perfectly unpredictable variation gradually snowballed to yield perfectly conditioned systems." (SMITH *et al.*, 2017, p. 4).

---

61 See also: http://mufwene.uchicago.edu/feature_pool.html/ (access on 20/06/2016)

If all these possibilities are considered, the *Feature Pool* is constantly changing and each born child would face a different configuration of the pool, depending on the speech community - in smaller proportions, depending on the family environment.[62] Such analysis could account for both language contact and variety contact.

[62] As for the right arrow in Mufwene figure, it may represent influence in the speech community due to migration; Rio, for instance, receive mostly migrants from Northeast of Brazil (45,5% of the total migrants) and from other Southeast states such as São Paulo and Minas Gerais (25,6% of the total migrants) and they would bring along their own variation rates for number agreement. The percentage of migrants was taken from Jacob *et al.* (2015) based on *2010 IBGE* (*Instituto Brasileiro de Geografia e Estatística* – Brazilian Institute for Geography and Statistics) census. Jacob *et al.* highlight the IBGE census is realized each ten years, which restricts the possibility of mapping numbers for migration in Brazilian cities, given the large gap between the mappings. Thus, there is not a fixed number of migrants per period, since mobility in great. However, they reveal that more than 80% of Rio municipality is constituted by natives. In this sense, the number of migrants in Rio is small and, nowadays, the population may be mainly formed by children of migrants who were born and raised in Rio. Though, this number is not available, given the difficulties pointed by Jacob *et al*., which limits a possible linguistic mapping in Rio city of which types of migrants brought in which type of variable phenomena and if the number of these speakers is really significant in order to impact the variable number agreement phenomena, which actually pervades all the Brazilian territory as sociolinguistic studies attest.

8
Conclusion

This thesis is concerned with the acquisition and production of variable number agreement in BP. One of the aims of this thesis was to analyze possible preferences regarding morphophonological number agreement expressions by preschoolers and 6th graders. The data gathered shows that SES is related to the production of variation of number agreement, as attested by sociolinguistic studies. Differences in production are not attested in preschool but social aspects and overall academic performance play a significant role in schooling years, distinguishing groups. The degree of (under)specification is defined by the linguistic experience influenced by SES. The elicited production results show that, both preschoolers and 6th graders from both SES, exhibit preference for the standard variety, however, SES influences on the rates of production of non-standard variety: Pub exhibit higher rates of production of non-standard variety. Thus, it is proposed that, for preschoolers, variation emerges from the underspecification of morphophonological features pertaining to redundancy in number agreement in the lexicon.

This thesis also aimed to investigate the role of literacy during schooling years regarding possible preferences for varieties. 6th graders' results show that overall academic performance may overcome SES differences. As seen in the contrast between Pub A and Pub B, the latter is able to reach similar performance to Priv, apparently due to schooling, given that Pub A and Priv belong to different SES groups. Low SES together with below-average academic performance seems to interfere on Pub B linguistic development, levelling them with Pub preschoolers. However, this does not mean a necessary impairment in these speakers; on the contrary: they do exhibit knowledge of singular and plural, however, the *status* of (under)specification of varieties does not seem to change, regardless of schooling, literacy and age, which impacts their ability to code-switch between varieties. Overall academic performance, thus, seems to allow speakers to enhance their code-switching abilities, as in the case of Pub A. 6th graders' results point to social issues of a society with high levels of social inequality in which SES and academic

performance seem to be interwoven in a complex network of socio-cultural-economic factors.

Another objective of this thesis was to compare results obtained with experimental methodology to the spontaneous speech results in the sociolinguistic literature. In this thesis, results from all groups exhibit some level of variation, especially in preschoolers' which is supposedly the representation of a naturally acquired grammar. Systematic variation of number agreement is, thus, part of the knowledge of BP speakers (SCHERRE, 1994). Our results are in consonance with adults' results gathered by Brandão (2013): number agreement variation is semi-categorical for highly-educated speakers and variable for speakers with lower educational level. In the present case, number agreement is clearly semi-categorical for Priv 6th and variable for Pub B; Pub A stands in between those two, more likely towards semi-categorical *status*.

The results also enable a discussion on the similarities between bilingualism and bidialectalism considering both contexts as part of a multilingual *continuum.* In the context of Rio, especially in its suburban area, speakers seem to be in contact with both varieties since birth, similarly to BFLA contexts, as shown in preschoolers' results. Schooling and literacy may work either as L2 acquisition, as in the case of the standard variety for Pub A or as passive bilingualism (or even heritage speakers, as in the case of non-standard varieties for Priv 6th). Based on Cornips (2014), children from Rio may be considered as "in-between", standing in the middle of a multilingual *continuum*, sharing characteristics of bilingual and bidialectal populations.

Additionally, this thesis intended to provide both a characterization of the linguist knowledge pertaining to number agreement in BP and a characterization of the access to this knowledge during actual language use. Hence, this study proposes that variation in number agreement in BP emerges due to the following conditions:

(i) Proper mapping of distributional properties of morphophonological information regarding functional categories (ADGER, 2014, CORRÊA, 2009, 2014; MIYAGAWA, 2010);
    a. Morphophonology interacts with syntax in terms of structure building: the result of *Agree* may be visible at the interfaces by means of morphophonology;

(ii) Children must represent different varieties regarding their morphophonology;
   a. Exposure to variable input leads to an underspecified representation in the lexicon (PoV), allowing for two different morphemes to be inserted during morphophonological encoding in language production, that is, post-syntactically;
   b. Different degrees of (under)specification leads to different representations; higher proficiency in the use of different varieties, increases code-switching abilities and variation may emerge during message encoding, when the variety is chosen;

(iii) External factors influence the level of (under)specification in grammar;
   a. As children enter school, SES, literacy and social interactions determine the *status* and weight (YANG, 2002) of each variety in *PoV* with effects on the speech production, more specifically, in morphophonological encoding.

In the approach advanced in this thesis, it is possible to predict possible difficulties for non-typically developing children acquiring BP, which was also an objective of this thesis. Thus, non-typically developing children with some sort of linguistic impairment could present difficulties in:

a) the mapping of distributional properties of morphophonological information regarding functional categories;
b) the processing of the functional phase within the DP which forces interaction of different sources of information such as semantics and pragmatics to establish a referent, exhibiting overall difficulties in the grammar-pragmatic interface (LONGCHAMPS; CORRÊA, 2014);
c) the access to inflectional morphemes from PoV during morphophonological encoding.

As Tomas *et al* suggest, learning morphophonological processes may pose challenges for SLI children, for instance:

> […] learning morphophonological regularities may be particularly problematic in languages other than English, where rich allomorphic variation, as well as contractions, liaison and elisions occur. In such languages problems in acquiring various morphophonological patterns might even serve as a clinical marker of SLI. Further research is required to investigate the effects of morphophonological and phonological factors on SLI children's use of grammatical morphemes crosslinguistically. (TOMAS *et al.*, 2015, p. 13)

We emphasize that, in the case of number agreement in BP, the omission of the determiner or in the elements of the functional phase of the DP, the *locus* of number marking, could be a marker for impairment. Additionally, all these difficulties could be amplified by SES and overall academic performance, just like with typically developing children. Therefore, investigation of Brazilian non-typically developing children with some sort of linguistic impairment may reveal strategies adopted by this population to deal with variable input, more specifically number agreement. [63]

From a theoretical point of view, it is expected that this material contributes to a broader research agenda of language architecture from a formal perspective. Boeckx (2011) suggests that cross-linguistic variation lies within the morphophonological surface string and that syntax is actually invariant, a view also advanced by Adger and Smith (2005). His view is not different from Chomsky's (1995) who posits that all variation and idiosyncrasies are located in the lexicon and that variation emerges from experience (2nd factor in Chomsky, 2005). Actually, even before that, Borer (1981, 1984) proposed that inflectional morphology must be stored in the lexicon and that properties of inflectional morphology guarantees interlanguage variation. Borer (1984, p. 29) proposes that "the burden of learning" is placed exactly in the inflectional system.

This thesis comes back to these aspects of the literature in an attempt to provide more fine-grained analysis of what is meant by variation in post- and pre-syntactic components. The model of production adopted in this thesis and the

[63] Linguistic aspects are likely to reveal social contrasts and may be indicators for language disorders such as DLD (Developmental Language Disorder). This issue is acknowledged by *Early Intervention Foundation* in UK recent report: to assess language as an index for children's wellbeing: *Early Intervention Foundation* Report, September, 2017. Available at: http://www.eif.org.uk/publication/language-as-a-child-wellbeing-indicator/

concept of underspecification of morphophonological features in the lexicon may account for such issues long discussed in the literature.

Under the approach for variation presented in this thesis, the notion of multiple grammars (ROEPER, 1999; YANG, 2002) and one grammar with variation may be combined in a multilingual *continuum* perspective. This thesis presented empirical results as well as theoretical accounts which seem to reinforce the importance of interfaces in the minimalist framework and the relevance of bridging the gap among the fields of theoretical linguistics, language processing, language acquisition and sociolinguistics. Additionally, it is highlighted that studies on intra-speaker variation, reserved to sociolinguistics until recently, may be approached from a formal perspective in order to contribute to accounts for language design, language architecture and, consequently, language change and language ecology (see MUFWENE, 2001, 2008). In sum, it is expected that this thesis contributes to a broader research agenda on formal linguistic theory and to a closer communication among fields of language studies formerly seen as incompatible.

This thesis also contributes to the debate on multilingualism and how closely-related varieties may occupy a space within multilingual studies, given that both contexts share language/variety contact characteristics (WEINREICH, 1970). This may point to the fact that there seems to be common pressures regulating language contact settings, whether it concerns different languages or closely-related varieties.

Furthermore, this thesis touches very briefly upon the nature-nurture debate (*cf.* DEHAENE-LAMBERTZ; HERTZ-PANNIER; DUBOIS, 2006) when bringing SES, literacy and language development together. Given the vast Brazilian territory that concentrates so much social variability, it is acknowledged that this investigation intended to highlight SES and academic performance, leaving other social factors aside. However, these factors are embedded in a greater picture of the social reality in countries that exhibit high levels of social inequality. As Pace *et al.* (2017) highlight:

> Because the majority of the available research is correlational, we do not yet know whether children from lower-SES homes begin with lower levels of processing skill or whether they develop different processing skills due to lesser levels of input, exposure, and experience. Future research that experimentally manipulates the level of input required to improve children's

> potential to acquire new vocabulary words and grammatical structures will be essential if we are to accurately characterize the relation between what children know about language and how they learn new language. (PACE *et al.*, 2017, p. 293)

In this thesis, results show a relation between SES and academic performance in a very specific linguistic phenomenon, the variable morphophonological expression of number agreement redundancy in BP. However, the boundaries of this interaction are not clear-cut; recent studies investigate the impact of SES on syntactic structure (see CORRIVEAU; KURKUL; ARUNACHALAM, 2016; HUANG; LEECH; ROWE, 2017; MUSOLINO, 2009; ROWE, 2008) but, to the moment, investigations on SES and morphophonology from the perspective of brain and cognitive development are not known. Investigations that address the interaction between language acquisition and development from a psycholinguistic standpoint with bioecological approaches may be highly informative for language architecture in terms of mental representation, language variation and language use and, ultimately, in language evolution. Moreover, it may be informative for pedagogical practices and educational policies.

As for teaching, this thesis shows that children arrive at school with different grammatical representations regarding number agreement in BP. During schooling age, children are presented to the standard/ written variety of BP and are supposed to engage on its use during academic development enlarging the gap between spoken and written BP (KATO, 1990, 2005). Also, the socially stigmatized non-standard varieties raise the issue of linguistic prejudice (BAGNO, 2007; BAGNO; RANGEL, 2005). It is relevant to enlighten students and teachers on the possibility of code-switching in native language (see WHEELER, 2016). Hence, pedagogically-oriented activities benefit from studies in the acquisition and processing of variation. As pointed out recently in (psycho)linguistics, a better understanding of how variable phenomena are processed may lead to improvement of teaching techniques for the learning process and contribute to the achievement of full literacy (cf AMARAL, 2017, p. 11; KENEDY, 2016).

Finally, it is acknowledged that social factors that influence language acquisition and development are not completely covered for many other reasons that would require a long-term research. However, there is hope that this investigation may, at least, shed some light on few issues regarding the relation

between social factors, grammar representation, language acquisition and processing, which can be achieved through an integrative approach. Such approach may help uncover the constraints on language variation from different perspectives, exploring its dynamics (TAMMINGA; MACKENZIE; EMBICK, 2016). An interdisciplinary approach may offer a better understanding of the language acquisition process under different contexts, such as multilingualism, and language disorders.

8.1
Possible further developments

As in any research, this thesis results in other questions and possible developments. In relation to multilingual settings important questions emerge: what are the linguistic aspects that favor a great area of intersection between linguistic systems as proposed in the underspecified grammar? Typological proximity (ROTHMAN, 2011) or structural proximity (WESTERGAARD *et al.*, 2017)? Furthermore, how would these internal factors correlate with or relate to external factors? Would the linguistic factors be convergent forces in the sense that contribute to the intersection area between linguistic systems and the external factors would work as diverging pressures, that is, those that make the systems separate? Furthermore, is it easier to process or acquire linguistic systems that are closely-related or completely different?

Such questions would demand a comparative investigation across different multilingual settings with different pairs of linguistic varieties being acquired, for instance, English-BP, Spanish-EP, Spanish-BP, EP-BP, Japanese-BP. Additionally, it would require an investigation both from comprehension and production and also perception. Another study could compare children acquiring such languages and adults' performance in these languages in order to verify when such factors are most relevant in language use and representation. More studies on variation with multilingual population could reveal language processing strategies that may not be visible in monolinguals (*cf.* FRICKE; KROLL; DUSSIAS, 2016).

Another question concerns specifically a model for language production and the verification of the extent of lookahead. This is a core issue in Psycholinguistics. It would be relevant to investigate the same structures with different intonation and verify which type of intonation reveals more errors of agreement, for instance. Such results could inform, from the point of view of production, how much structural planning is necessary based on the amount of errors produced.

In terms of language use it is also relevant to verify whether external factors such as SES and academic development, as proposed in this thesis, may interfere on specific *loci* of linguistic structure, such as morphophonological properties or if this influence is indirect, mediated by other systems or operations in language production. The use of different experimental techniques and brain image could be informative.

Regarding the representation of linguistic knowledge and formal theories, the functional phase of the DP should be better investigated. The relation between D, Poss and number marking in BP needs further characterization. Definiteness should be considered when analyzing this relation and it would be relevant from the point of view of language acquisition to verify how these properties are related and how they are represented in grammar. Such analysis could have a direct impact in current accounts for bare nouns in BP. Bare nouns in BP involve the omission of a definite article as well as the omission of number markings. All these properties are encoded in the functional phase of the DP, as assumed in this thesis. Further analysis with experimental data manipulating elements in terms of definiteness and number in the functional phase of the DP could shed light on the type of nominals in BP and characterize a possible parametric distinction between BP and other Romance languages that do not allow bare singulars in subject and/or object position (see SCHMITT and MUNN, 2005; CYRINO and ESPINAL, 2015 and others).

Another possible development of this thesis is to conduct the same experiment in different regions in Rio in order to comprehend a broader *spectrum* of SES and compare children's performance. A linguistic history questionnaire could also be helpful to map the environmental aspects of the speaker. Besides that, the phonic saliency factor aspect needs further experimental investigation to verify to which extent it really plays a role in language processing and if it is dependent on SES in the sense that speakers from low educational level would more often rely

on phonic saliency than speakers from high educational level. Additionally, phonic saliency itself should be better characterized (CHAVES, 2014).

These are only few questions that emerged during the elaboration of this thesis and these questions may be important for sociolinguistics and for psycholinguistics. In sum, multilingual settings and variation are a fertile area of investigation.

9

References

ACQUAVIVA, P. Roots, categories and nominal concepts. **Lingue e Linguaggio**, v. VIII, n. 1, p. 25–51, 2009.

ADGER, D. Combinatorial Variability. **Journal of Linguistics**, v. 42, n. 03, p. 503, 2006.

_____. Variability and modularity: A response to Hudson Variability and modularity : A response to. **Journal of Linguistics**, v. 43, n. 43, p. 695–700, 2007.

_____. Variability and grammatical architecture. In: **Linguistic Variation in the Minimalist Framework**. New York: Oxford University Press, 2014. p. 179–196.

_____. Language variability in syntactic theory. In: EGUREN, L.; FERNÀNDEZ-SORIANO, O.; MENDIKOETXEA, A. (Eds.). . **Rethinking Parameters**. First edit ed. New York: Oxford University Press, 2016.

ADGER, D.; SMITH, J. Variation and the minimalist program. In: CORNIPS, L.; CORRIGAN, K. P. (Eds.). . **Syntax and Variation: Reconciling the Biological and the Social**. Amsterdam/ Philadelphia: John Benjamins Publishing. p. 149–178.

_____. Variation in agreement: A lexical feature-based approach. **Lingua**, v. 120, n. 5, p. 1109–1134, 2010.

AKERE, F. Verbal Strategies in Communal Meetings: Code-switching and Status Manipulation in a Bi-dialectal Yoruba Speech Community. **Language Sciences**, v. 2, n. 1, p. 102–126, 1980.

ALEXANDRE, N.; HAGEMEIJER, T. Bare nouns and the nominal domain in Santome. In: GUÉRON, J.; BAPTISTA, M. (Eds.). . **Noun phrases in Creole languages : a multi-faceted approach**. Amsterdam/ Philadelphia: J. Benjamins Pub. Co, 2007. p. 37–60.

ALEXIADOU, A. Plural mass nouns and the morpho-syntax of number. **Proceedings of the 28th West Coast Conference on Formal Linguistics**, p. 33–41, 2011.

ALMEIDA, D. R. **Processamento da concordância de gênero e número em estruturas predicativas**. Rio de Janeiro, 2016. 188 p. Dissertação de Mestrado. Departamento de Letras. Pontíficia Universidade Católica do Rio de Janeiro.

ALVES, M. T. G.; SOARES, J. F.; XAVIER, F. P. Índice Socioeconômico das

Escolas de Educação Básica Brasileiras. **Ensaio: aval. pol. públ. Educ.**, v. 22, n. 84, p. 671–704, 2014.

AMARAL, L. **Entrevista Especial In: Psicolinguística: Estudos no BrasilSoletras Revista**Revista do Programa de Pós-Graduação em Letras e Linguística – PPLIN Faculdade de Formação de Professores / Universidade do Estado do Rio de Janeiro (UERJ), , 2017.

ANTONIOU, K. GROHMANN, K. K.; KAMBANAROS, M.; KATSOS, N. The effect of childhood bilectalism and multilingualism on executive control. **Cognition**, v. 149, p. 18–30, 2016.

ATKINSON, E. Gender features on n & the root. In: SMITH, J.; IHSANE, T. (Eds.). . **Romance Linguistics 2012: Selected papers from the 42nd Linguistic Symposium on Romance Languages (LSRL), Cedar City, Utah, 20-22 April 2012**. Amsterdam/ Philadelphia: John Benjamins Publishing Company, 2015. p. 229–244.

AUGUSTO, M. R. A. Marcação de número e genericidade : interpretação genérica na aquisição do PB. **Letras de Hoje**, v. 42, n. 1, p. 35–51, 2007.

AUGUSTO, M. R. A.; CORRÊA, L. M. S.; FORSTER, R. An argument for DPs as phases in an integrated model of on-line computation: the immediate mapping of complex DPs with relative clauses. **ReVEL**, v. special, n. 6, p. 7–26, 2012.

AUGUSTO, M. R. A.; FERRARI NETO, J.; CORRÊA, L. M. S. Explorando o DP : a presença da categoria NumP. **Revista de Estudos Linguísticos**, v. 14, n. 2, p. 245–275, 2006.

AYIOMAMITOU, I.; YIAKOUMETTI, A. Skewed sociolinguistic awareness of a native non-standard dialect: Evidence from the Cypriot Greek writing of Greek Cypriot students. **Frontiers in Psychology**, v. 8, n. NOV, p. 1–10, 2017.

AZALIM, C. **Variação e processamento linguístico: um estudo experimental sobre a concordância nominal variável no PB**. Juiz de Fora, 2016. 129 p. Dissertação de Mestrado. Programa de Pós-Graduação em Linguística - Universidade Federal de Juiz de Fora, 2016.

BAGETTI, T. **Um Estudo Experimental do Processamento na Interface Fônica e da Análise Sintática Inicial: O Papel de Elementos Funcionais na Aquisição da Linguagem**. Rio de Janeiro, 2009. 151 p. Tese de Doutorado. Departamento de Letras - Pontíficia Universidade Católica do Rio de Janeiro, 2009.

BAGETTI, T.; CORRÊA, L. M. S. **The early recognition of verb affixes: evidence from Portuguese**. Proceedings of the 35th Annual Boston University Conference on Language Development. **Anais**...2009. Available at: http://www.bu.edu/bucld/files/2011/07/BAGETTI-CORREA-BUCLD-35.pdf Access on: 18/07/2015.

_____. O reconhecimento do padrão morfofonológico dos afixos verbais na aquisição inicial do português brasileiro. **Veredas atemática**, v. 17, n. 2, p. 102–122, 2013. Available at: 2013 http://www.ufjf.br/revistaveredas/files/2014/04/2%C2%BA-ARTIGO.pdf Access on: 18/07/2015

BAGNO, M. **Nada na língua é por acaso – Por uma pedagogia da variação linguística.** São Paulo: Parábola, 2007.

_____. **Preconceito Linguístico: o que é, como se faz**. 49ª ed. São Paulo: Edições Loyola, 2007.

BAGNO, M.; RANGEL, E. D. O. Tarefas da educação lingüística no Brasil. **Revista Brasileira de Linguística Aplicada**, v. 5, n. 1, p. 63–81, 2005.

BAILEY, C.-J. N. **Variation and Linguistic Theory**. Center for Applied Linguistics, Arlington, VA. 169 p. Report. 1973. Available at: https://eric.ed.gov/?id=ED125238 Access on: 09/01/2017.

BARATZ, J. C. A bidialectal task for determining language proficiency in economically disadvantaged Negro children. **Child Development**, v. 40, n. 3, p. 889–901, 1969.

BARNES, J. A. Class and Committees in a Norwegian Island Parish. **Human Relations**, v. 7, n. 1, p. 39–58, 1954.

BASIT, T. N.; HUGHES, A.; IQBAL, Z.; COOPER, J. The influence of socio-economic status and ethnicity on speech and language development. **International Journal of Early Years Education**, v. 23, n. 1, p. 115–133, 2 jan. 2015.

BEARD, R. Morpheme order in a lexeme/morpheme-based morphology. **Lingua**, v. 72, p. 1–44, 1987.

BEARD, R.; VOLPE, M. Lexeme-Morpheme Base Morphology. In: STEKAUER, P.; LIEBER, R. (Eds.). . **Handbook of Word-Formation**. The Netherlands: Springer, 2005. p. 189–205.

BECKER, O. M. S.; COSTA, G. E. DA. **Urbanização, segregação e áreas sociais na região metropolitana do rio de janeiro**. Anais do XX Encontro Nacional de Estudos Populacionais. **Anais**...2016

BERWICK, R. C.; CHOMSKY, N. The biolinguistic program: the current state of its development. In: BOECKX, C.; DI SCIULLO, A.-M. (Eds.). . **The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty**. New York: Oxford University Press, 2011. p. 17–41.

BIALYSTOK, E. **Bilingualism in Development: Language, Literacy & Cognition**. Cambridge: Cambridge University Press, 2001.

BOCK, K.; MILLER, C. A. Broken agreement. **Cognitive Psychology**, v. 23, n. 1, p. 45–93, 1991.

BOECKX, C. Approaching Parameters from Below. In: BOECKX, C.; DI SCIULLO, A.-M. (Eds.). . **The Biolinguistic Enterprise: New Perspectives on the Evolution and Nature of the Human Language Faculty**. Oxford, UK: OUP Oxford, 2011. p. 205–221.

_____. Considerations pertaining to the nature of logodiversity. In: EGUREN, L.; FERNÀNDEZ-SORIANO, O.; MENDIKOETXEA, A. (Eds.). . **Rethinking Parameters**. First edit ed. New York: Oxford University Press, 2016.

BOECKX, C.; LEIVADA, E. On the particulars of Universal Grammar: Implications for acquisition. **Language Sciences**, v. 46, n. PB, 2014.

BOKAMBA, E. G. Code-mixing, language variation, and linguistic theory: Evidence from Bantu Languages. **Lingua**, v. 76, p. 21–62, 1988.

BOLAND, J. E.; KAAN, E.; KROFF, J. V.; WULFF, S. Psycholinguistics and variation in language processing. **Linguistics Vanguard**, v. 2, n. s1, p. 1–10, 2016.

BORER, H. **Parametric Variation in Clitic Constructions**. Cambridge, MA: Massachusetts Institute of Technology, 1981.

BORER, H. **Parametric syntax, case studies in Semitic and Romance languages**. Dordrecht, Holland: Foris Publications, 1984.

BORTONI-RICARDO, S. M. **Do campo para a cidade: estudos sociolinguísticos de migração e redes sociais**. São Paulo: Parábola Editorial, 2011.

BRANDÃO, S. F. Patterns of plural agreement within the Noun Phrase. **Journal of Portuguese Linguistics**, v. 12, n. 2, p. 51–100, 2013.

_____. **Concordância nominal em três variedades do português: resultados gerais, novas indagações**. Cuadernos de la ALFAL. **Anais**...2015

_____. **Concordância Nominal Variável em Português**. (J.-P. Chauveau, M. Barbato, I. Fernández-Ordóñez, Eds.)Actes du XXVIIe Congrès international de linguistique et de philologie romanes (Nancy, 15-20 juillet 2013). Section 8 : Linguistique variationnelle, dialectologie et sociolinguistique. Nancy, ATILF. **Anais**...Nancy: 2016

BRANDÃO, S. F.; VIEIRA, S. R. Concordância nominal e verbal: contribuições para o debate sobre o estatuto da variação em três variedades urbanas do português. **Alfa**, v. 56, n. 3, p. 1035–1064, 2012.

CARVALHO, D. DA S. On gender agreement in Brazilian Portuguese. In: DALI, M.; MATHIEU, E.; ZAREIKAR, G. (Eds.). . **Gender and noun classication**. New York: Oxford University Press, 2018.

CASTRO, A. Possessivos e artigo definido expletivo em PE e PB. **Revista Varedas**, n. 1992, p. 18, 2009.

CASTRO, A; CORRÊA, L. M. S.; AUGUSTO, M. R. A; FERRARI NETO, J.. **The interpretation of the number morphology in Portuguese**. Théorie syntaxique et acquisition (a)typique du langage: journée scientifique en hommage a Celia Jakubowicz. Poster Presentation. Paris: 2009.

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**, v. 42, n. 1, p. 65–76, 2007.

CEDERGREN, H.; SANKOFF, D. Variable Rules : Performance as a Statistical Reflection of Competence. **Language**, v. 50, n. 2, p. 333–355, 1974.

CHAMBERS, J. K.; TRUDGILL, P. **Dialectology**. First Edit ed. Cambridge: Cambridge University Press, 1980.

CHAVES, R. G. Influência de Processos Fonológicos na Marcação Explícita de CVP6. **ReVEL**, v. edição esp, n. 13, p. 181–206, 2016.

CHESI, C. **An introduction to Phase-based Minimalist Grammars: why move is Top-Down from Left-to-Right**. StiL (Studies in Linguistics/ CISCL WP). **Anais**...2007

CHESI, C. **Competence and Computation: toward a processing a processing friendly minimalist grammar**. Padova, Italy: Unipress, 2012.

_____. On Directionality of Phrase Structure Building. **Journal of Psycholinguistic Research**, v. 44, n. 1, p. 65–89, 2014.

CHESI, C.; MORO, A. The Subtle Dependency between Competence and

Performance. **MIT Working Papers in Linguistics**, v. 77, n. 50 Years Later Reflections Chomsky's Asp., p. 33–46, 2015.

CHOMSKY, N. **The Minimalist Program**. Cambridge, Massachusetts: The MIT Press, 1995.

_____. Derivation by Phase. In: KENSTOWICZ, M. (Ed.). . **Ken Hale : a life in language**. Cambridge: MIT Press, 2001. v. 18p. 1–52.

_____. Problems of projection. **Lingua**, v. 130, p. 33–49, 2013.

CHRISTOPHE, A.; GUASTI, T.; NESPOR, M.; DUPOUX, E.; VAN OOYEN, B. Reflections on Phonological Bootstrapping: Its Role for Lexical and Syntactic Acquisition. **Language and Cognitive Processes**, v. 12, n. 5–6, p. 585–612, 1997.

CORBETT, G. G. Universals in the Syntax of Cardinal Numerals. **Lingua**, v. 46, p. 355–368, 1978.

CORNIPS, L. Socio-syntax and variation in acquisition Problematizing monolingual and bidialectal acquisition. **Linguistic Variation**, v. 14, n. 1, p. 1–25, 2014.

CORRÊA, L. M. S. MABILIN Módulos de Avaliação de Habilidades Linguísticas. **Relatório Técnico Projeto FAPERJ CNE 1999**: Procedimentos diferenciados na aquisição do Português e um instrumental psicolinguístico para a avaliação de deficiências de linguagem em crianças, 2000.

_____. Explorando a relação entre língua e cognição na interface: o conceito de interpretabilidade e suas implicações para teorias do processamento e da aquisição da linguagem. **Veredas**, v. 6, n. 1, p. 113–129, 2002.

_____.

_____. Bootstrapping language acquisition from a minimalist standpoint: On the identification of φ-features in Brazilian Portuguese. **Minimalist Inquiries Into Child and Adult Language Acquisition : Case Studies Across Portuguese**, v. 35, p. 35–62, 2009.

_____. Minimalist perspectives for psycholinguistic research. **Perspectivas minimalistas en la investigación psicolingüística.**, v. 29, n. 2, p. 214–230, 2011.

_____. Interface Information and Computational Cost: An Integrated Procedural Approach to Language Acquisition with Some Implications for SLI. **New Directions in the Acquisition of Romance Languages**, 2014.

CORRÊA, L. M. S.; AUGUSTO, M. R. A. Computação Linguística no processamento on-line: soluções formais para a incorporação de uma derivação

minimalista em modelos de processamento. **Cadernos de Estudos Lingüísticos**, v. 49, n. 2, p. 167–183, 2007.

_____. Possible loci of SLI from a both linguistic and psycholinguistic perspective. **Lingua**, v. 121, n. 3, p. 476–486, 2011.

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**. (D. Bamman, T. Magnitskaia, C. Zaller, Eds.). BUCLD 30 Online Proceedings Supplement. **Anais**...Boston: 2006

CORRÊA, L. M. S.; AUGUSTO, M. R. A.; MARCILESE, M. Competing analyses and differential cost in the production of non-subject relative clauses. **Glossa: a journal of general linguistics**, v. 3, n. 1, p. 62, 18 maio 2018.

CORRÊA, L. M. S.; NAME, M. C. L. The processing of determiner-noun agreement and the identification of the gender of nouns in the early acquisition of Portuguese. **Journal of Portuguese Linguistics**, v. 2, n. 1, p. 19–43, 2003.

CORRIVEAU, K. H.; KURKUL, K.; ARUNACHALAM, S. Preschoolers' Preference for Syntactic Complexity Varies by Socioeconomic Status. **Child Development**, v. 87, n. 5, p. 1529–1537, 2016.

COSTA, I. DE O. **Verbos meteorológicos no plural em orações relativas do português brasileiro: sintaxe e processamento**. Rio de Janeiro, 2013. 196 p. Dissertação de Mestrado. Departamento de Letras - Pontíficia Universidade Católica do Rio de Janeiro.

COSTA, J.; FIGUEIREDO SILVA, M. C. Nominal and verbal agreement in Portuguese. In: **Studies on Agreement**. Padova: Unipress, 2006. p. 25–46.

CRAIG, H. K.; WASHINGTON, J. A. Grade-Related Changes in American English. **Journal of Speech, Language, and Hearing Research**, v. 47, n. April 2004, p. 450–463, 2004.

CYRINO, S.; ESPINAL, M. T. Bare Nominals in Brazilian Portuguese: more on the DP/NP analysis. **Natural Language and Linguistic Theory**, v. 33, n. 2, p. 471–521, 2015.

DĄBROWSKA, E. Different speakers, different grammars. **Linguistic Approaches to Bilingualism**, v. 2, n. 3, p. 219–253, 2012.

DANON, G. Agreement and DP-internal feature distribution. **Syntax**, v. 14, n. 4, p. 297–317, 2011.

DANON, G. Two structures for numeral-noun constructions. **Lingua**, v. 122, n. 12,

2012.

DAYAL, V. Number Marking and (in)Definiteness in Kind Terms. **Linguistics and Philosophy**, v. 27, n. 4, p. 393–450, ago. 2004.

DE HOUWER, A. **Bilingual First Language Acquisition**. Clevedon, UK. Multilingual Matters, 2009.

DEHAENE-LAMBERTZ, G.; HERTZ-PANNIER, L.; DUBOIS, J. Nature and nurture in language acquisition: anatomical and functional brain-imaging studies in infants. **Trends in Neurosciences**, v. 29, n. 7, p. 367–373, 2006.

DELL, G. S. Speaking and Misspeaking. In: **An Invitation to Cognitive Science: Language**. Cambridge: MIT Press. p. 183–208.

DEPREZ, V. Morphological number, semantic number and bare nouns. **Lingua**, v. 115, n. 6 SPEC. ISS., p. 857–883, 2005.

DUARTE, M. E. L. **A perda do princípio "evite pronome" no portguês brasileiro**. Campinas, 1995. 161p. Tese de Doutorado. Instituto de Estudos da Linguagem - Universidade Estadual de Campinas.

_____. O papel da Sociolinguística no (re)conhecimento do português brasileiro e suas implicações para o Ensino. **Revista Letra**, 2013.

_____. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. In: ROBERTS, I.; KATO, M. A. (Eds.). . **Português Brasileiro: uma viagem diacrônica**. 3rd edition. São Paulo: Editora Contexto, 2018. p. 83–103.

DUARTE, M. E. L.; SERRA, C. R. Gramática(s), Ensino de Português e "Adequação Linguística". **Matraga**, v. 22, n. 36, p. 31–55, 2015.

DUARTE, M. E. L.; VAREJÃO, F. Null subjects and agreement marks in European and Brazilian Portuguese. **Journal of Portuguese Linguistics**, v. 2, n. 12, p. 101–123, 2013.

DUFLO, E.; DUPAS, P.; KREMER, M. **Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya**. NBER Working Paper No. 14475
Issued in November 2008, Revised in December 2011. Cambridge, Massachusetts. Available at: http://www.nber.org/papers/w14475 Access on: 29/11/2017

EDWARDS, J.; GROSS, M.; CHEN, J.; MACDONALD, M. C; KAPLAN, D; BROWN, M.; SEIDENBERG, M. S. Dialect Awareness and Lexical Comprehension of Mainstream American English in African American English–

Speaking Children. **Journal of Speech Language and Hearing Research**, v. 57, n. 5, p. 1883, 1 out. 2014.

EMBICK, D. The Morpheme: A Theoretical Introduction. p. 247, 2015.

EMBICK, D.; NOYER, R. Movement operations after syntax. **Linguistic Inquiry**, v. 32, n. 4, p. 555–595, 2001.

FEIGENSON, L.; DEHAENE, S.; SPELKE, E. Core systems of number. **Trends in Cognitive Sciences**, v. 8, n. 7, p. 307–314, 2004.

FERGUSON, C. . Diglossia. **Word**, v. 15, p. 325–40, 1959.

FERNALD, A.; MARCHMAN, V. A.; WEISLEDER, A. SES differences in language processing skill and vocabulary are evident at 18 months. **Developmental Science**, v. 16, n. 2, p. 234–248, 1 mar. 2013.

FERRARI NETO, J. **Reconhecimento do número gramatical e processamento da concordância de número no sintagma determinante na aquisição do português brasileiro**. Rio de Janeiro, 2003. 112 p. Departamento de Letras - Pontíficia Universidade Católica do Rio de Janeiro.

_____. **Aquisição de número gramatical no português brasileiro: processamento de informação de interface e concordância**. Rio de Janeiro, 2008. 154 p. Tese de Doutorado. Departamento de Letras - Pontíficia Universidade Católica do Rio de Janeiro.

FERRARI NETO, J.; LIMA, M. A. F. DE. Aquisição da morfologia flexional verbal em português brasileiro: um estudo com dados de compreensão. **Revista Prolíngua**, v. 10, n. 1, p. 106–120, 2015.

FILHO, A. A. P. **O bairro da Cidade Nova no Rio de Janeiro: a organização espacial e segregação urbana na segunda metade do século XIX**. Anais XVI Encontro Nacional dos Geógrafos. **Anais**...2010

FISHMAN, J. A. Bilingualism With and Without Diglossia; Diglossia With and Without Bilingualism. **Journal of Social Issues**, v. 23, n. 2, p. 29–38, abr. 1967.

FLORIPI, S. A. **Estudo da Variação do Determinante em Sintagmas Nominais Possessivos na História do Português**. Campinas, 2008. 271 p. Instituto de Estudos da Linguagem - Universidade Estadual de Campinas.

FRAMPTON, J.; GUTMANN, S. **Agreement is Feature Sharing**. ms. Northeastern University, 2000.

_____. How sentences grow in the mind. In: BOECKX, C. (Ed.). . **Agreement Systems**. Amsterdam: John Benjamins Publishing Company, 2006. p. 121–157.

FRICKE, M.; KROLL, J. F.; DUSSIAS, P. E. Phonetic variation in bilingual speech: A lens for studying the production–comprehension link. **Journal of Memory and Language**, v. 89, n. 2, p. 110–137, ago. 2016.

GALVES, C. **Ensaio sobre as gramáticas do português**. Campinas, Brasil: Editora da Unicamp, 2001.

GHOMESHI, J. Plural marking, indefiniteness, and the noun phrase. **Studia Linguistica**, v. 57, n. 2, p. 47–74, 2003.

GOMES, C. A.; PONTES, M. C. V.; ALMEIDA, M. C. S.; ABREU, A. C. B. Variação e aquisição da flexão nominal e da flexão verbal. **Gragoatá**, v. 30, p. 39–54, 2011.

GONÇALVES, F. **Concordância verbal e determinação de parâmetros sintácticos no processo de aquisição do Português Europeu e Brasileiro**. XXI Encontro Nacional da Associação Portuguesa de Linguística. **Anais**...Lisboa: 2006. Available at: http://www.apl.org.pt/docs/21-textos-seleccionados/34_Goncalves.pdf Access on: 28/07/2014

GONÇALVES, P. **(Dados para a) história da língua portuguesa em Moçambique**, 2009.

GRIMM, S. **Number and Individuation Submitted To the Department of Linguistics**. PhD Dissertation. Stanford University, 2012.

GROHMANN, K. K. Towards Comparative Bilingualism. **Linguistic Approaches to Bilingualism**, v. 4, n. 3, p. 337–342, 2014.

GROHMANN, K. K.; KAMBANAROS, M; LEIVADA, E.; ROWE, C. A developmental approach to diglossia: Bilectalism on a gradient scale of linguality. **Poznan Studies in Contemporary Linguistics**, v. 52, n. 4, p. 629–662, 2016.

GROHMANN, K. K.; KAMBANAROS, M. The Gradience of Multilingualism in Typical and Impaired Language Development: Positioning Bilectalism within Comparative Bilingualism. **Frontiers in Psychology**, v. 7, n. February, p. 1–18, 2016.

GROHMANN, K. K.; LEIVADA, E. Interface ingredients of dialect design. In: DI SCIULLO, A. M. (Ed.). . **Towards a Biolinguistic Understanding of Grammar: Essays on interfaces**. Amsterdam/ Philadelphia: John Benjamins Publishing Company. p. 239–262.

GUIMARÃES, M. A. DE S.; SILVA, J. A. A. Variação na Concordância Nominal de Número no Português Popular de Vitória da Conquista - BA: Sócio-História do

Português do Brasil. **Tabuleiro de Letras**, v. 10, n. 2, p. 139–151, 2016.

GUY, G. R. **Linguistic Variation in Brazilian Portuguese: Aspects of the Phonology, Syntax, and Language History**. PhD Thesis. University of Pennsylvania, 1981.

_____. The cognitive coherence of sociolects: How do speakers handle multiple sociolinguistic variables? **Journal of Pragmatics**, v. 52, p. 63–71, 2013.

HALLE, M.; MARANTZ, A. Distributed Morphology and the Pieces of Inflection. **The View from Building 20: Essays in Honor of Sylvain Bromberger**, p. 111–176, 1993.

HARLEY, H.; NOYER, R. Distributed Morphology. **Glot International**, v. 4, n. 4, 1999.

HARLEY, H.; RITTER, E. **Person and Number in Pronouns: A Feature-Geometric AnalysisLanguage**Linguistic Society of America, , 2002.

HARTSUIKER, R. J.; SCHRIEFERS, H. J.; BOCK, K.; KIKSTRA, G. M.. Morphophonological influences on the construction of subject-verb agreement. **Memory & cognition**, v. 31, n. 8, p. 1316–1326, 2003.

HAUSER, M. D.; CHOMSKY, N.; FITCH, W. T. The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? **Science**, v. 298, n. 5598, p. 1569–1579, 2002.

HENRIQUE, K. DA S. **Variação linguística e processamento: investigando o papel da distância entre sujeito e verbo na realização da concordância verbal variável no PB.** Juiz de Fora, 2016. 149 p. Dissertação de Mestrado. Programa de Pós-Graduação em Linguística. Universidade Federal de Juiz de Fora.

HENRY, A. Dialect Variation , Optionality , and the Learnability Guarantee. **Linguistica atlantica**, v. 20, p. 51–71, 1998.

_____. Non-standard dialects and linguistic data. **Lingua**, v. 115, n. 11 SPEC. ISS., p. 1599–1617, 2005.

HIRAIWA, K. The faculty of language integrates the two core systems of number. **Frontiers in Psychology**, v. 8, n. MAR, 2017.

HUANG, Y. T.; LEECH, K.; ROWE, M. L. Exploring socioeconomic differences in syntactic development through the lens of real-time processing. **Cognition**, v. 159, p. 61–75, 2017.

HURFORD, J. The interaction between numerals and nouns. **Noun Phrase Structure in the Languages of Europe**, n. Section 3, p. 561–620, 1998.

HYAMS, N. **Now you hear it, now you don't: The nature of optionality in child grammars**. (A. H, Ed.)BUCLD 25 Proceedings. **Anais**...Somerville, MA: Cascadilla Press, 2001

Censo Escolar da Educação Básica 2016: Notas Estatísticas. Instituto Nacional de Estudos e Pesquisa. Brasília: Brasil. Available at: http://download.inep.gov.br/educacao_basica/censo_escolar/notas_estatisticas/2017/notas_estatisticas_censo_escolar_da_educacao_basica_2016.pdf Access on: 14/10/2017.

IONIN, T.; MATUSHANSKY, O. The Composition of Complex Cardinals. **Journal of Semantics**, v. 23, n. 4, p. 315–360, 2006.

IONIN, T.; MONTRUL, S.; SANTOS, H. The Expression of Genericity in English and Brazilian Portuguese : An Experimental Investigation. **English**, v. 2008, n. 2, p. 2007–2008, 2010.

JACOB, C. R.; HEES, D. R.; WANIES, P. **Mapping Living Conditions in Rio de Janeiro Metropolitan Area [electronic resources]**. Available at: http://www.editora.vrc.puc-rio.br/cgi/cgilua.exe/sys/start.htm?infoid=329&sid=3 Accessed on: 14/08/2017

JESCHENIAK, J. D.; LEVELT, W. J. M. Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. **Journal of Experimental Psychology: Learning, Memory, and Cognition**, v. 20, n. 4, p. 824–843, 1994.

JUSCZYK, P. W. **The discovery of spoken language**. Cambridge, MA: MIT Press, 1997.

KAMBANAROS, M.; GROHMANN, K. K.; MICHAELIDES, M. Lexical retrieval for nouns and verbs in typically developing bilectal children. **First Language**, v. 33, n. 2, p. 182–199, 2013.

KATO, M. A. **No mundo da escrita: uma perspectiva psicolingüística**. 3ª edição ed. São Paulo: Editora Ática, 1990.

_____. A gramática do letrado. **Ciências da Linguagem: trinta anos de investigação e ensino**, p. 131–145, 2005.

_____. O português são dois… ou três? In: LOBO, T. *et al.* (Eds.). . **Rosae: linguística histórica, história das línguas e outras histórias**. Salvador: EDUFBA, 2012. p. 93–108.

_____. A variação no domínio dos clíticos no português brasileiro.

Lingüística, v. 33, n. 1, 2017.

KEMPEN, G.; HOENKAMP, E. An incremental procedural grammar for sentence formulation. **Cognitive Science**, v. 11, n. 2, p. 201–258, 1987.

KENEDY, E. O status da norma culta na língua-I dos brasileiros e seu respectivo tratamento na escola: algumas contribuições de estudos formalistas à educação. In: GUESSER, S. (Ed.). . **Linguística: Pesquisa e Ensino**. Boa Vista, Roraima: Editora da Universidade Federal de Roraima, 2016. v. 2p. 185–208.

KENNEDY, D. F. **Bidialectalism vis-a-vis Bilingualism, with Specific Reference to Black English (and Application to Early Reading)**. Course paper for University of Maryland. 68 p. Available at: http://www.eric.ed.gov/contentdelivery/servlet/ERICServlet?accno=ED408852. Access on: 03/06/2017

KINZLER, K. D.; SPELKE, E. S. Core systems in human cognition. **Progress in Brain Research**, v. 164, p. 257–264, 2007.

KOMEILI, M.; MARSHALL, C. R. Sentence repetition as a measure of morphosyntax in monolingual and bilingual children. **Clinical Linguistics & Phonetics**, v. 27, n. 2, p. 152–162, 2013.

KRAMER, R. The location of gender features in the syntax. **Linguistics and Language Compass**, v. 10, n. 11, p. 661–677, 2016a.

_____. A Split Analysis of Plurality: Number in Amharic. **Linguistic Inquiry**, v. 47, n. 3, p. 527–559, 2016b.

KROCH, A. Morphosyntactic variation. **Proceedings of the thirtieth annual meeting of the Chicago Linguistics Society**, v. 2, p. 180–201, 1994.

KUHL, P. K. Brain Mechanisms in Early Language Acuisition. **Neuron**, v. 67, n. 5, p. 713–727, 2010.

_____.. Early Language Learning and Literacy: Neuroscience. **Mind Brain Education**, v. 5, n. 3, p. 128–142, 2011.

LABOV, W. Phonological Correlates of Social Stratification. **American Anthropologist**, v. 66, n. 6, p. 164–176, 1964.

_____. The Social Stratification of (r) in New York City Department Stores. In: COUPLAND N., J. A. (Ed.). . **Sociolinguistics**. London: Macmillan Education UK, 1997. p. 168–178.

_____. **Padrões sociolinguísticos (Sociolinguistic Patterns)**. Tradução de Marcos Bagno, Maria Marta Pereira Scherre, Caroline Rodrigues Cardoso. ed. São

Paulo: Parábola Editorial, 2008.

LAPOINTE, S. G.; DELL, G. S. A Synthesis of Some Recent Work in Sentence Production. In: CARLSON, G. N.; TANENHAUS, M. K. (Eds.). . **Linguistic Structure in Language Processing**. Studies in Theoretical Psycholinguistics. Dordrecht: Springer Netherlands, 1988. v. 7p. 107–156.

LEE, E.-K.; BROWN-SCHMIDT, S.; WATSON, D. G. Ways of looking ahead: Hierarchical planning in language production. **Cognition**, v. 129, n. 3, p. 544–562, dez. 2013.

LEIVADA, E.; PAPADOPOULOU, E.; KAMBANAROS, M.; GROHMANN, K. K. The influence of bilectalism and non-standardization on the perception of native grammatical variants. **Frontiers in Psychology**, v. 8, n. FEB, p. 1–11, 2017.

LEVELT, W. J. M. The Speaker as Information Processor. In: **Speaking: From Intention to Articulation**. Camdridge, MA: The MIT Press, 1989. p. 1–28.

_____. Models of word production. **Trends in Cognitive Sciences**, v. 3, n. 6, p. 223–232, 1999.

LI, A.; SHI, R.; HUA, W. Prosodic cues to noun and verb categories in infant-directed Mandarin speech. **Proceedings of the 5th International Conference on Speech Prosody (SP2010)**, p. 58–61, 2010.

LIMA, S. **The grammar of individuation and couting**. PhD Thesis. Amherst: University of Massachussetts Amherst, 2014.

LIPTON, J. S.; SPELKE, E. S. Preschool Children's Mapping of Number Words to Nonsymbolic Numerosities. **Child Development**, v. 76, n. 5, p. 978–988, 2005.

LONGCHAMPS, J. R.; CORRÊA, L. M. S. Interface gramática-pragmática e problemas de aprendizagem: definitude e relevância no processamento da referência. **Veredas atemática**, v. 18, n. 2, p. 1–19, 2014.

LOPES, C. R. DOS S. **A gramaticalização de a gente em português em tempo real de longa e de curta duração: retenção e mudança na especificação dos traços intrínsecos**. Fórum Lingüístico, Florianópolis, v. 4, n.1 (47-80), julho de 2004.

LOPES, R. E. V. Estágios no processo de aquisição de número no DP do Português Brasileiro. **Letras de Hoje**, v. 39, n. 3, p. 157–171, 2004.

_____. Bare nouns and DP number agreement in the acquisition of Brazilian Portuguese. **Selected Proceedings of the 9th Hispanic Linguistics Symposium**, p. 252–262, 2006.

LOWENSTAMM, J. Little n, ROOT, and Types of Nouns. In: HARTMANN, J.; HEGEDUS, V.; VAN RIEMSDIJK, H. (Eds.). . **Sounds of Silence: Empty Elements in Syntax and Phonology**. Oxford: Elsevier, 2007. p. 105–143.

LUCCHESI, D. Parâmetros Sociolingüísticos Do Português Brasileiro. **Revista da Abralin**, v. 5, n. 1 e 2, p. 83–112, 2006. Available at: http://www.abralin.org/revista/RV5N1_2/RV5N1_2_art4.pdf Access on: 04/11/2014.

_____. A Teoria da Variação Linguística : um balanço crítico. **Estudos Linguísticos**, v. 41, n. 2, p. 793–805, 2012.

LUCCHESI, D.; BAXTER, A.; RIBEIRO, I. **O Português Afro-Brasileiro**. Salvador, Bahia: Editora da Universidade Federal da Bahia, 2009.

MAGALHÃES, T. M. V. O uso de artigo definido diante de pronome possessivo em textos portugueses do século XVI a XIX. **Leitura**, n. 47, p. 123–143, 2011.

MARANTZ, A. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. **University of Pennsylvania Working Papers in Linguistics**, v. 4, n. 2, p. 201–225, 1997.

MARCILESE, M.; HENRIQUE, K. S.; AZALIM, C.; NAME, C. Processamento da concordância variável. **Revista LinguíStica / Revista do Programa de Pós-Graduação em Linguística da Universidade Federal do Rio de Janeiro**, v. 11, n. 1, p. 118–134, 2015.

MARCILESE, M.; RODRIGUES, E. S.; AUGUSTO, M. R. A.; HENRIQUE, K. S. Efeitos de distância linear e marcação no processamento da concordância verbal variável no PB Linear distance and markedness effects in variable subject-verb agreement processing in BP. **Revista de Estudos da Linguagem**, n. Thematic Issue: Experimental Linguistics, 2017.

MARIANO, M. P. **O fenômeno da concordância nominal em redações escolares**. Rio de Janeiro, 2013. 92 p. Dissertação de Mestrado. Faculdade de Letras - Universidade Federal do Rio de Janeiro.

MARINS, J.; SOARES DA SILVA, H.; DUARTE, M. E. L. Revisiting duarte (1995): for a gradient analysis of indeterminate subjects in brazilian portuguese. **Diadorim**, v. Especial, p. 101–126, 2017.

MARSLEN-WILSON, W. D.; TYLER, L. K. Morphology, language and the brain: the decompositional substrate for language comprehension. **Philosophical Transactions of the Royal Society B: Biological Sciences**, v. 362, n. 1481, p. 823–

836, 2007.

MATUSHANSKY, O.; IONIN, T. **When cardinals agree (a catholic point of view)**Leiden, the NetherlandsThe 48th Annual Meeting of Societas Linguistica Europaea, , 2015.

MEISEL, J. M. **Bilingual First Language Acquisition : French and German Grammatical Development**. Amsterdam: John Benjamins Pub. Co, 1994.

MENON, O. P. DA S. O Sistema Pronominal do Português do Brasil. **Letras**, n. 44, p. 91–106, 1995.

MERCHANT, J. **Where's gender? Evidence from Greek**Universiteit Groningen, Kluck's workshop, 2011.

MILLER, K. L.; SCHMITT, C. Effects of variable input in the acquisition of plural in two dialects of Spanish. **Lingua**, v. 120, n. 5, p. 1178–1193, 2010.

_____. Variable Input and the Acquisition of Plural Morphology. **Language Acquisition**, v. 19, n. December, p. 223–261, 2012.

_____. Spanish-speaking children's use of verbal inflection in comprehension. **Lingua**, v. 144, p. 40–57, 2014.

MILROY, J.; MILROY, L. Linguistic change, social network and speaker innovation. **Journal of Linguistics**, v. 21, n. 02, p. 339, 1985.

MITCHELL, J. C. Social Networks. **Annual Review of Anthropology**, p. 279–299, 1974.

MIYAGAWA, S. Why Agree? In: **Why Agree? Why Move? Unifying Agreement-Based and Discourse-Configurational Languages**. Cambridge, MA: The MIT Press, 2010. p. 1–29.

MIYAGAWA, S.; WU, D.; KOIZUMI, M. **Deriving Case Theory**. Ms. 2018.

MOLINA, D. S. L.; MARCILESE, M.; NAME, C. M. L. Ora está, ora não está: input variável e aquisição da flexão verbal de 3ª pessoa do plural no PB. **Matraga**, v. 24, n. 41, p. 288–309, 2017.

MOLINA, D. S. L. **Aquisição da linguagem e variação linguística: um estudo sobre a flexão verbal variável na aquisição do PB**. Juiz de Fora, *in preparation* (a sair). Tese de Doutorado. Programa de Pós-graduação em Linguística - Universidade Federal de Juiz de Fora.

MONTE, A. **O estatuto variável da concordância verbal de terceira pessoa do plural no PB e no PE**. Cuadernos de la ALFAL. **Anais**...2015

MUFWENE, S. S. **The Ecology of Language Evolution**. New York: Cambridge

University Press, 2001.

_____. **Language Evolution: contact, competition and change**. London/ New York: Continuum International Publishing Group, 2008.

MUNN, A.; SCHMITT, C. Number and indefinites. **Lingua**, v. 115, n. 6 SPEC. ISS., p. 821–855, 2005.

MUSOLINO, J. The logical syntax of number words: Theory, acquisition and processing. **Cognition**, v. 111, n. 1, p. 24–45, 2009.

NAME, M. C. L. **Habilidades Perceptuais e Lingüísticas no Processo de Aquisição do Sistema de Gênero no Português**. Rio de Janeiro, 2002. 169 p. Tese de Doutorado. Departamento de Letras - Pontíficia Universidade Católica do Rio de Janeiro.

NARO, A. J. The Social and Structural Dimensions of a Syntactic Change. **Language**, v. 57, n. 1, p. 63–98, 1981.

NARO, A. J.; SCHERRE, M. M. P. Drifting Toward the Standard Language: A Panel Study of Number Concord in Brazilian Portuguese. In: CACOULLOS, R. T.; DION, N.; LAPIERRE, A. (Eds.). . **Linguistic Variation: confronting fact and theory**. New York. Routledge, 2015. p. 356.

NEGRÃO, E. V.; VIOTTI, E. Em busca de uma história linguística. **Revista Estudos da Linguagem**, v. 20, n. 2, p. 309–342, 2012.

NEVINS, A.; PARROTT, J. K. Variable rules meet Impoverishment theory: Patterns of agreement leveling in English varieties. **Lingua**, v. 120, n. 5, p. 1135–1159, 2010.

OTTONI, F. C. B. **Os Jovens e as Desigualdades Sócio-Espaciais no Município do Rio de Os Jovens e as Desigualdades Sócio-Espaciais no Município do Rio de**. Anais do XVI Encontro Nacional de Estudos Populacionais. **Anais**...2008

PACE, A.; LUO, R.; HIRSH-PASEK, K.; GOLINKOFF, R. M.. Identifying Pathways Between Socioeconomic Status and Language Development. **Annual Review of Linguistics**, v. 3, n. 1, p. 285–308, 2017.

PANAGIOTIDIS, P. **(Grammatical) gender troubles and the gender of pronouns**, 2015.

PAP, S. Stable and developmental optionality in native and non-native Hungarian grammars. **Second language Research**, v. 16, n. 2, p. 173–200, 2000.

PAPAPAVLOU, A.; PHILI, A. **Childhood Bidialectism and Metalinguistic Awareness**. Selected Papers from the 18th ISTAL. **Anais**...2009

PARODI, T.; TSIMPLI, I.-M. “Real” and apparent optionality in second language grammars: finiteness and pronouns in null operator structures. **Second Language Research**, v. 21, n. 3, p. 250–285, 2005.

PEREIRA, B. K. Feature interpretability and the positions of 2nd person possessives in dialectal Brazilian Portuguese. **Filologia e Linguística Portuguesa**, v. 18, n. 2, p. 199–229, 2016.

PEREIRA, M. L. DE S.; ARAÚJO, A. A. DE. Variação na concordância verbal no português brasileiro : olhares variacionistas. **Revista de Letras**, v. 18, n. 23, p. 148–164, 2017.

PERKINS, S. C.; FINEGOOD, E. D.; SWAIN, J. E. Poverty and language development: roles of parenting and stress. **Innovations in clinical neuroscience**, v. 10, n. 4, p. 10–9, abr. 2013.

PESETSKY, D.; TORREGO, E. The Syntax of Valuation and the Interpretability of Features. In: KARIMI, S.; SAMIIAN, V.; WILKINS, W. (Eds.). . **Clausal and Phrasal Architecture: syntactic derivation and interpretation: in Honor of Joseph E. Emonds (Linguistik Aktuell / Linguistics Today)**. Amsterdam/ Philadelphia: John Benjamins Publishing Company, 2007. p. 262–294.

PICALLO, M. C. Gender and Number in Romance. **Lingue e Linguaggio**, v. 7, n. 1, p. 47–66, 2008.

_____. A note on the locus and function of formal gender. **Borealis – An International Journal of Hispanic Linguistics**, v. 6, n. 1, p. 1, 30 maio 2017.

PIRES, A.; ROTHMAN, J. Child and adult language acquisition, linguistic theory and (microparametric) variation. **Minimalist Inquiries Into Child and Adult Language Acquisition : Case Studies Across Portuguese**, v. 25, p. 5-, 2009.

PIRES, A.; ROTHMAN, J.; SANTOS, A. L. L1 acquisition across Portuguese dialects: Modular and interdisciplinary interfaces as sources of explanation. **Lingua**, v. 121, n. 4, p. 605–622, 2011.

POZZO, L. D. **The Finnish Noun Phrase**. Doctoral Thesis. Venezia: Università Ca’ Foscari di Venezia Facoltà, 2007.

PRIM, C. D. S. Os adjetivos exclusivamente pré-nominais do português. **Signum: Estudos da Linguagem**, v. 18, n. 2, p. 377, 2015.

QUINE, W. V. **Word and Object**. New Edition ed. Cambridge, Mass: The MIT Press, 2013.

RICHARDS, M. Two Kinds of Variation in a Minimalist System. **Varieties of**

Competition, p. 133–162, 2008.

RINKE, E. A combinação do artigo definido e pronome possessivo na história do português. **Estudos de Lingüística Galega**, v. 2, p. 121–139, 2010.

RITTER, E. On the syntactic category of pronouns and agreement. **Natural Language and Linguistic Theory**, v. 13, n. 3, p. 405–443, ago. 1995.

ROBERTS, J. Child Language Variation. In: **The Handbook of Language Variation and Change**. Oxford, UK: Blackwell Publishing Ltd, 2008. p. 318–333.

RODRIGUES, E. DOS S. **Processamento da concordância de número entre sujeito e verbo na produção de sentenças**. Rio de Janeiro, 2006. Tese de doutorado. 202 p. Departamento de Letras - Pontíficia Universidade Católica do Rio de Janeiro.

_____. Concordância verbal com construções partitivas – uma proposta de análise. **Veredas**, v. 15, n. 1, p. 93–107, 2011.

ROELOFS, A; MEYER, A S.; LEVELT, W. J. A case for the lemma/lexeme distinction in models of speaking: comment on Caramazza and Miozzo (1997). **Cognition**, v. 69, n. 2, p. 219–230, 1998.

ROELOFS, A. Rightward incrementality in encoding simple phrasal forms in speech production: Verb-particle combinations. **Journal of Experimental Psychology-Learning Memory and Cognition**, v. 24, n. 4, p. 904–921, 1998.

ROELOFS, A.; BAAYEN, H. Morphology by itself in planning the production of spoken words. **Psychonomic bulletin & review**, v. 9, n. 1, p. 132–138, 2002.

ROEPER, T. Universal Bilingualism. **Bilingualism: Language and Cognition**, v. 2, n. 03, p. 169–186, 1999.

_____. Multiple Grammars and the Logic of Learnability in Second Language Acquisition. **Frontiers in psychology**, v. 7, n. February, p. 14, 2016.

ROORYCK, J. On two types of underspecification: Towards a feature theory shared by syntax and phonology. **Probus**, v. 6, n. 2–3, p. 207–234, 1994.

ROTHMAN, J. L3 syntactic transfer selectivity and typological determinacy: The typological primacy model. **Second Language Research**, v. 27, n. 1, p. 107–127, 2011.

ROTHMAN, J.; GIANCASPRO, D.; HALLORAN, B. On the structural basis of non-redundant acquisition Evidence from Spanish bilingual L3 Portuguese. **Portuguese-Spanish Interfaces: Diachrony, synchrony, and contact**, p. 317–334, 2014.

ROTHSTEIN, S. Counting, Measuring And The Semantics Of Classifiers. **Baltic International Yearbook of Cognition, Logic and Communication**, v. 6, n. 1, 1 jan. 2010.

ROWE, M. L. Child-directed speech: relation to socioeconomic status, knowledge of child development and child vocabulary skill. **Journal of Child Language**, v. 35, n. 01, p. 185–205, 2008.

SANKOFF, D.; LABOV, W. On the uses of variable rules. **Language in Society**, v. 8, n. 2, p. 189–222, 1979.

SARNECKA, B. W.; KAMENSKAYA, V. G. From Grammatical Number to Exact Numbers: Early Meanings of "One," "Two," and "Three" in English, Russian and Japanese. **NIH Public Access**, v. 55, n. 2, p. 136–168, 2008.

SCHERRE, M. M. P. **A regra de concordância de número no sintagma nominal em português**. Rio de Janeiro, 1978. 174 p. Dissertação de Mestrado. Departamento de Letras - Pontifícia Universidade Católica do Rio de Janeiro.

_____. A concordância de número nos predicativos e nos particípios passivos. **Organon**, n. 18, p. 52–70, 1991.

_____. Aspectos da concordância de número no português do Brasil. **Revista Internacional de Língua Portuguesa (RILP) - Norma e Variação do Português**, p. 37–49, 1994.

SCHERRE, M. M. P.; NARO, A. J. **Sobre a concordância de número no português falado do Brasil**. (G. Ruffino, Ed.)Dialettologia, geolinguistica, sociolinguistica.(Atti del XXI Congresso Internazionale di Linguistica e Filologia Romanza). **Anais**...1998

_____. Mudança sem mudança: a concordância de número no português brasileiro. **Scripta**, v. 9, n. 18, p. 107–129, 2006.

_____. Perceptual vs. grammatical constraints and social factors in subject-verb agreement in Brazilian Portuguese. **University of Pennsylvania working papers in linguistics**, v. 16, n. 2, p. 164–171, 2010.

_____. Sociolinguistic correlates of negative evaluation: Variable concord in Rio de Janeiro. **Language Variation and Change**, v. 26, n. 2, p. 331–357, 2014.

SCHILLER, N. O.; COSTA, A. Different selection principles of freestanding and bound morphemes in language production. **Journal of Experimental Psychology: Learning Memory and Cognition**, v. 32, n. 5, p. 1201–1207, 2006.

SCHMITT, C.; MUNN, A. The syntax and semantics of bare arguments in

Brazilian Portuguese. **Linguistic Variation Yearbook**, v. 2, p. 185–216, 2002.

SCHWINDT, L. C. O acesso à morfologia por processos fonológicos variáveis e a arquitetura da gramática. **Cadernos de ESTUDOS LINGÜÍSTICOS**, v. 56, n. 1, p. 23–43, 2014.

SHI, R. Functional Morphemes and Early Language Acquisition. **Child Development Perspectives**, v. 8, n. 1, p. 6–11, mar. 2014.

SHI, R.; MARQUIS, A.; GAUTHIER, B. **Segmentation and representation of function words in preverbal French-learning infants**. (D. D. Bamman, T. Magnitskaia, C. Zaller, Eds.) Proceedings of the 30th annual Boston University conference on language development. **Anais**...Boston, MA: Cascadilla Press, 2006

SHI, R.; MELANÇON, A. Syntactic categorization in French-learning infants. **Infancy**, v. 15, n. 5, p. 517–533, 2010.

SHLONSKY, U. The form of Semitic noun phrases. **Lingua**, v. 114, n. 12, p. 1465–1526, 2004.

SILVA, T. C. Organização fonológica de marcas de plural no português brasileiro: uma abordagem multirrepresentacional. **Revista da Abralin**, v. 11, n. 1, p. 273–305, 2012.

SIMIONI, L. **A aquisição da concordância nominal de número no portguês brasileiro: um parâmetro para a concordância nominal**. Santa Catarina, 2007. 172 p. Tese de Doutorado. Programa de Pós-Graduação em Linguística - Universidade Federal de Santa Catarina.

SIMÕES, L. J. Aprendizagem da gramática do português escrito : algumas reflexões a partir da aquisição da língua falada. **Caleidoscópio**, v. 4, n. 1, p. 51–59, 2006.

SLOBIN, D. I.; WELSH, C. A. **Elicited imitation as a research tool in developmental psycholinguistics**. Report. 21 p. Available at: https://eric.ed.gov/?id=ED012892. Access on: 26/06/2016.

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.

SMITH, K.; PERFORS, A. FEHÉR, O.; SAMARA, A.; SWOBODA, K.; WONNACOTT, E.. Language learning, language use and the evolution of linguistic variation. **Philosophical Transactions of the Royal Society of London B: Biological Sciences**, v. 372, n. 1711, 2017.

SORACE, A.; KELLER, F. Gradience in linguistic data. **Lingua**. Volume 115, Issue 11, Pages 1497-1524, November 2005.

SQUIRES, L. Social Differences in the Processing of Grammatical Variation. **UPenn Working Papers in Linguistics**, v. 20, n. 2, 2014. Available at: http://repository.upenn.edu/pwpl/vol20/iss2/20 Access on: 08/12/2015

STEMBERGER, J. P.; MACWHINNEY, B. Frequency and the lexical storage of regularly inflected forms. **Memory & cognition**, v. 14, n. 1, p. 17–26, 1986.

SUMNER, M.; SAMUEL, A. G. The effect of experience on the perception and representation of dialect variants. **Journal of Memory and Language**, v. 60, n. 4, p. 487–501, 2009.

TAFT, M.; FORSTER, K. I. **Lexical storage and retrieval of prefixed wordsJournal of Verbal Learning and Verbal Behavior**, 1975.

TAMMINGA, M.; MACKENZIE, L.; EMBICK, D. The dynamics of variation in individuals. **Linguistic Variation**, v. 16, n. 2, p. 300–336, 2016.

TARALLO, F. Major sociolinguistic patterns in Brazilian Portuguese. **International Journal of the Sociology of Language**, v. 89, n. 1, p. 9–24, 1991.

TERRY, N. P. Examining Relationships Among Dialect Variation and Emergent Literacy Skills. **Communication Disorders Quarterly**, v. 33, n. 2, p. 67–77, 15 fev. 2012.

THEODOROU, E.; KAMBANAROS, M.; GROHMANN, K. K. Sentence Repetition as a Tool for Screening Morphosyntactic Abilities of Bilectal Children with SLI. **Frontiers in Psychology**, v. 8, n. December, 2017.

TOMAS, E.; DEMUTH, K.; SMITH-LOCK, K. M.; PETOCZ, P.. Phonological and morphophonological effects on grammatical development in children with specific language impairment. **International Journal of Language & Communication Disorders**, v. 50, n. 4, p. 516–528, jul. 2015.

TRUPPI, C. **Bare Nouns among and beyond Creoles**. PhD Dissertation. Humboldt Universität, 2014.

TSIMPLI, I. M. Variation in Second Language Acquisition. **Encyclopedia of Language & Linguistics**, n. 1981, p. 387–394, 2006.

_____. Early, late or very late? **Linguistic Approaches to Bilingualism**, v. 4, n. 3, p. 283–313, 2014.

TUCKER, M. A.; IDRISSI, A.; ALMEIDA, D. Representing number in the real-time processing of agreement: Self-paced reading evidence from Arabic. **Frontiers**

in Psychology, v. 6, n. MAR, 2015.

VAN KAMPEN, J. An acquisitional view on optionality. **Lingua**, v. 114, n. 9–10, p. 1133–1146, 2004.

VERKUYL, H. J. Numerals and quantifiers in X-syntax and their semantic interpretation. In: GROENENDIJK, J. A. G.; JANSSEN, T. M. V.; STOKHOF, M. B. (Eds.). . **Formal Methods in the Study of Language Volume 2**. Amsterdam: University of Amsterdam, 1981. p. 567–599.

VIEIRA, M. C. P. **A emergência do padrão flexional variável da 3ª pessoa do plural na aquisição do PB como L1**. Rio de Janeiro, 2006. 120 p. Dissertação de Mestrado em Linguística. Centro de Letras e Artes - Universidade Federal do Rio de Janeiro, 2006. Available at: http://www.letras.ufrj.br/poslinguistica/wp-content/uploads/2012/04/marcia_cristina_vieira.pdf Access on: 28/07/2014

VIEIRA, S. R. **Padrões de concordância de terceira pessoa plural no português: um balanço das contribuições do projeto 21 da ALFAL**. Cuadernos de la ALFAL. **Anais**...2015

VIEIRA, S. R.; BAZENGA, A. Patterns of third person plural verbal agreement. **Journal of Portuguese Linguistics**, v. 12, n. 2, p. 7–50, 2013.

VIEIRA, S. R.; BRANDÃO, S. F. Tipologia de regras linguísticas e estatuto das variedades/ línguas: a concordância em português. **Lingüística**, v. 30, n. 2, p. 81–112, 2014.

WEINREICH, U. **Languages in Contact: Findings and Problems**. Paris: Mounton Publishers, 1970.

WEINREICH, U.; LABOV, W.; HERZOG, M. I. Empirical Foundations for a Theory of Language Change. In: LEHMAN, W. P.; MALKIEL, Y. (Eds.). . **Directions for Historical Linguistics - A Symposium**. Austin-London: University of Texas Press, 1968. p. 95–199.

WESTERGAARD, M.; MITROFANOVA, M.; MYKHAYLYK, R.; RODINA, Y.. Crosslinguistic influence in the acquisition of a third language: The Linguistic Proximity Model. **International Journal of Bilingualism**, v. 21, n. 6, p. 666–682, 2017.

WEXLER, K. Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. **Lingua**, v. 106, n. 1–4, p. 23–79, 1998.

WHEELER, R. “So much research, so little change”: Teaching Standard English in

African American classrooms. **Annual Review of Linguistics**, v. 2, n. 1, p. 367–390, 2016.

WHEELER, R. S.; SWORDS, R. Codeswitching : Tools of Language and Culture Transform the Dialectally Diverse Classroom. **Language Arts**, v. 81, n. 6, p. 470–480, 2004.

WILSON, S. African American English : Dialect Mistaken As an Articulation Disorder. **McNair Scholars Research Journal**, v. 4, n. 1, 2012.

WUERGES, T. E. **A Aquisição da Morfologia Verbal por Crianças Falantes de Português Brasileiro e o Uso de Formas Variantes**. São Paulo, 2014. 155 p. Dissertação de Mestrado - Departamento de Linguística, Universidade de São Paulo. Available at: http://www.teses.usp.br/teses/disponiveis/8/8139/tde-28052014-125108/pt-br.php. Access on: 28/07/2014

YANG, C. D. **Knowledge and Learning in Natural Language**. New York: Oxford University Press, 2002.

YANG, C. D.; ELLMAN, A.; LEGATE, J. A. **Input and its structural description** (Á. J. Gallego, D. Ott, Eds.)**50 Years Later: Reflections on Chomsky's Aspects**. Cambridge, MA: MITWPL, 2015.

Appendix I

Stimuli for each experimental condition

1) *O-s* *cachorro(s)* *encontr(aram/ou)* *o*
ART.DEF.MASC.PL dog.MASC.PL/SG find.3.PST.PL/SG ART.DEF.MASC.SG
leão
lion.MASC.SG
'The dog(s) found the lion'

2) *O-s* *gato(s)* *assust(aram/ou)* *a*
ART.DEF.MASC-PL cat.MASC-PL/SG scare.3.PST.PL/SG ART.DEF.FEM.SG
borboleta
butterfly.FEM.SG
'The cat(s) scared the butterfly'

3) *O-s* *elefante(s)* *arrum(aram/ou)* *o*
ART.DEF.MASC-PL elephant.MASC-PL/SG organize.3.PST.PL/SG ART.DEF.MASC.SG
quarto
room.MASC.SG
'The elephant(s) organized the room'

4) *O-s* *macaco(s)* *consert(aram/ou)* *a*
ART.DEF.MASC-PL monkey.MASC-PL/SG fix.3.PST.PL/SG ART.DEF.FEM.SG
porta
door.FEM.SG
'The monkey(s) fixed the door'

5) *O-s* *cavalo(s)* *visit(aram/ou)* *o*
ART.DEF.MASC-PL horse.MASC-PL/SG visit.3.PST.PL/SG ART.DEF.MASC.SG
pinguim
penguin.MASC.SG
'The horse(s) visited the penguin'

6) *O-s* *pintinho(s)* *abraç(aram/ou)* *a*
ART.DEF.MASC-PL chick.MASC-PL/SG hug.3.PST.PL/SG ART.DEF.FEM.SG
galinha
chicken.FEM.SG
'The chick(s) hugged the chicken'

7) *O-s* *urso(s)* *quebr(aram/ou)* *o*
ART.DEF.MASC-PL bear.MASC-PL/SG break.3.PST.PL/SG ART.DEF.MASC.SG
brinquedo
toy.MASC.SG
'The bear(s) broke the toy'

8) *O-s* *esquilo(s)* *limp(aram/ou)* *a*
ART.DEF.MASC-PL squirrel.MASC-PL/SG clean.3.PST.PL/SG ART.DEF.FEM.SG
balcony
balcony.FEM.SG
'The squirrel(s) cleaned the balcony'

9) *O-s* *pato(s)* *pente(aram/ou) o*
ART.DEF.MASC-PL duck.MASC-PL/SG comb.3.PST.PL/SG ART.DEF.MASC.SG
papagaio
parrot.MASC.SG
'The duck(s) combed the parrot'

10) *O-s* *sapo(s)* *suj(aram/ou)* *a*
ART.DEF.MASC-PL frog.MASC-PL/SG make.dirty.3.PST.PL/SG ART.DEF.FEM.SG
tartaruga
turtle.FEM.SG
'The frog(s) made the turtle dirty'

11) *O-s* *coelho(s)* *varre(ram/u)* *o*
ART.DEF.MASC-PL rabbit.MASC-PL/SG sweep.3.PST.PL/SG ART.DEF.MASC.SG
jardim
garden.MASC.SG
'The rabbit(s) swept the garden'

12) *O-s* *pássaro(s)* *come(ram/u)* *a*
ART.DEF.MASC-PL bird.MASC-PL/SG eat.3.PST.PL/SG ART.DEF.FEM.SG
maçã
apple.FEM.SG
'The bird(s) ate the apple'

13) *O-s* *canguru(s)* *pul(aram/ou)* *o*
ART.DEF.MASC-PL kangoroo.MASC-PL/SG jump.3.PST.PL/SG ART.DEF.MASC.SG
jacaré
aligator.MASC.SG
'The kangaroo(s) jumped the alligator'

14) *O-s* *tigre(s)* *molh(aram/ou)* *a*
ART.DEF.MASC-PL tiger.MASC-PL/SG made.wet.3.PST.PL/SG ART.DEF.FEM.SG
coruja
owl.FEM.SG
'The tiger(s) made the owl wet'

15) *O-s* *porco(s)* *pintaram* *o*
ART.DEF.MASC-PL pig.MASC-PL/SG paint.3.PST.PL/SG ART.DEF.MASC.SG
desenho
drawing.MASC.SG
'The pig(s) painted the drawning'

16) *O-s* *lobo(s)* *empurr(aram/ou)* *a*
ART.DEF.MASC-PL wolf.MASC-PL/SG pushed.3.PST.PL/SG ART.DEF.FEM.SG
janela
window.FEM.SG
'The wolves/wolf pushed the window'

Appendix II

Approval of the Ethics Committee of the Pontifical Catholic University of Rio de Janeiro

PONTIFÍCIA UNIVERSIDADE CATÓLICA
DO RIO DE JANEIRO

Câmara de Ética em Pesquisa da PUC – Rio

PARECER DA COMISSÃO DE ÉTICA EM PESQUISA DA PUC-RIO (2017-01)

A Câmara de Ética em Pesquisa da PUC-Rio foi constituída como uma Câmara específica do Conselho de Ensino e Pesquisa conforme decisão deste órgão colegiado com atribuição de avaliar projetos de pesquisa do ponto de vista de suas implicações éticas.

Identificação:
Título: Representação e acesso a informação gramatical do DEL (Distúrbio Específico da Linguagem): sintomas compartilhados com transtornos de alta incidência em crianças em idade escolar (Departamento de Letras da PUC-Rio)
Autora: Letícia Maria Sicuro (Professora do Departamento de Letras da PUC-Rio)

Apresentação: O projeto em questão foi apreciado favoravelmente e homologado do ponto de vista ético pela Câmara de Ética em Pesquisa em reunião do Conselho de Ensino e Pesquisa da PUC-Rio em 27 de abril de 2016 (Parecer 2016-07). Ele inscreve-se no campo da Psicolinguística concebido originalmente com vistas à obtenção de resultados comportamentais da parte de crianças em idade escolar (7 a 12 anos), focalizando relações sintáticas de longa distância em estruturas passivas, interrogativas e relativas. Justifica-se, no momento, sua reapresentação pela necessidade de modificações para inserção de investigações sobre o processamento de questões morfo-sintáticas relacionadas à expressão de relações de concordância (particularmente de número) na produção e na compreensão da linguagem e seu impacto na escolarização. Assim, apresenta como objetivo adicional: caracterizar o desenvolvimento típico linguístico típico de crianças adquirindo o português brasileiro desde a idade pré-escolar, como padrão de referência para o desenvolvimento atípico. As técnicas a serem usadas na investigação no ambiente escolar com crianças a partir de dois anos e seis meses de idade serão: produção eliciada; identificação de imagens a partir de bases; manipulação de brinquedos, conforme instruções; fixação preferencial do olhar. As atividades programadas serão conduzidas em ambiente escolar. A participação das crianças será feita sem qualquer coação ou desconforto. Qualquer manifestação de cansaço, desinteresse ou aborrecimento será suficiente para que a atividade seja interrompida. A proposta enriquece a pesquisa especialmente no que tange a aquisição da linguagem típica e atípica. Os resultados a serem alcançados darão subsídios para avaliar-se como e em que medida as manifestações do DEL comprometem a expressão da concordância no português brasileiro.

Vice-Reitoria para Assuntos Acadêmicos
Rua Marquês de São Vicente, 225 - Gávea - 22453-900.
Rio de Janeiro - RJ - Tel. (021) 3527 1619 FAX (021) 3527 1132.
E-mail: vrac@puc-rio.br