

DUPLICAÇÃO DO SUJEITO EM DADOS DE ESCRITA DE ALUNOS DA REDE PÚBLICA DE ENSINO DO RECIFE

Patrycia Siqueira Campos Ferreira¹
Claudia Roberta Tavares Silva²

RESUMO: Neste artigo centramos a atenção no uso dos sujeitos duplicados (SDs) (ex.: *O menino, ele brinca.*) em textos narrativos produzidos por alunos recifenses do Ensino Fundamental e Médio de duas escolas: uma estadual e outra municipal. Para tanto, tomamos como hipótese norteadora que os dados de aquisição da fala do PB, por serem pouco produzidos SDs, segundo observam Gonçalves (2004) e Grolla (2000) *apud* Costa, Duarte e Silva (2004), serão pouco encontrados nos anos iniciais, mais produtivos nos anos intermediários, vindo a diminuir bastante nos anos finais, por influência do avançar da escolarização. Para a realização da análise, selecionamos 10 textos do gênero estória de cada ano do Ensino Fundamental I de uma escola municipal, e do Ensino Fundamental II e Ensino Médio de uma escola estadual. A análise embasa-se no modelo de Princípios e Parâmetros (CHOMSKY, 1981, 1986) e tem como objetivo central discutir a duplicação dos sujeitos em dados escritos, levando em conta o possível distanciamento entre o conhecimento linguístico que o aluno traz para a escola e o que ele aprende. Após a análise dos dados, os resultados apontam para a pouca frequência de SDs nos anos iniciais, indo ao encontro do que é observado em alguns estudos realizados no campo da aquisição da linguagem (GONÇALVES, 2004; GROLLA, 2000 *apud* COSTA, DUARTE, SILVA, 2004).

PALAVRAS-CHAVE: Escrita; Português Brasileiro; Sujeito Duplicado

RESUMEN: En este artículo centramos la atención en el uso de los sujetos duplicados (SDs) (por ejemplo, *el niño, él juega.*) en textos narrativos producidos por alumnos recifenses de la enseñanza fundamental y media de dos escuelas: una estatal y otra municipal. Para ello, tomamos como hipótesis orientadora que los datos de adquisición del habla del PB, por ser poco producidos SDs, según observan Gonçalves (2004) y Grolla (2000) *apud* Costa, Duarte y Silva (2004), esperamos encontrar pocos en los años iniciales, siendo productivos en los años intermediarios, viniendo a declinar en los años finales por influencia del avance de la escolarización. Seleccionamos por lo tanto 10 textos del género historia de cada año de la enseñanza fundamental I de una escuela municipal, y de la enseñanza fundamental II y media de una escuela estatal. El análisis se basa en el modelo de Principios y Parámetros (CHOMSKY, 1981, 1986), teniendo como objetivo central discutir la duplicación de los sujetos en datos escritos, teniendo en cuenta el posible distanciamento entre el conocimiento lingüístico que el alumno trae para la enseñanza escuela y lo que él aprende. En el presente trabajo se analizan los resultados obtenidos en el análisis de los datos, los resultados apuntan a una poca frecuencia de SDs en los años iniciales, yendo al encuentro de lo que se observa en algunos estudios realizados en el campo de la adquisición del lenguaje (GONÇALVES, 2004; GROLLA, 2000 *apud* COSTA, DUARTE, SILVA, 2004).

PALABRAS-CLAVE: Escritura; Portugués Brasileño; Sujeto Duplicado

¹ Graduanda do Curso de Licenciatura em Letras Português/ Espanhol da Universidade Federal Rural de Pernambuco.

² Professora Associada 2 do Departamento de Letras da Universidade Federal Rural de Pernambuco.

1. Introdução

Pesquisas linguísticas têm chamado a atenção para as diferenças entre a gramática da fala e a “gramática” da escrita no Brasil. Tarallo (1996, p. 70) observa que “[o] perfil da nossa gramática brasileira (no sentido da gramática normativa) tem sido ditado pela tradição portuguesa e só esse fato torna o vácuo entre língua oral e escrita muito mais profundo no Brasil do que em Portugal”. Esse fato tem levado grande parte dos docentes a deixarem de lado uma eficaz reflexão sobre o processo de reorganização por que a gramática do português brasileiro (PB) vem passando (GALVES, 2001). Essa reorganização é explicada por muitos pesquisadores pelo processo de enfraquecimento da morfologia de flexão verbal. Em seu estudo diacrônico, Duarte (1995) observa, a partir dos resultados quantitativos, um aumento acentuado de sujeitos pré-verbais plenos em PB na gramática adulta (1845 (20%), 1882 (23%), 1918 (25%), 1937 (46%), 1955 (50%), 1975 (67%) e 1992 (74%)), ao contrário do português europeu (PE), estando o sujeito nulo (ou seja, aquele não realizado foneticamente) submetido a restrições típicas de línguas de sujeito nulo parcial (HOLMBERG, NAYUDU; SHEEHAN, 2009; KATO; DUARTE, 2014a, 2014b).

Algumas das evidências dessa reorganização gramatical podem ser verificadas, a saber: a) sujeitos plenos no PE recebem interpretação contrastiva em relação a outros sujeitos pragmaticamente possíveis no domínio do discurso, ao contrário do PB (GONÇALVES, 1994; SILVA, 2004) e b) construções com sujeitos duplicados (SDs) são muito produtivas em PB, ao contrário do PE, quer em sentenças principais (1a), quer em encaixadas (1b):

- (1)a. **A Clarinha_i** *ela_i* cozinha que é uma maravilha. (?PE OKPB)
b. Eu acho que **o povo brasileiro_i** *ele_i* tem uma grave doença. (*PE OKPB)
(COSTA; GALVES, 2002)

Ao analisar os SDs em PB, Silva (2004) argumenta que, na gramática do falante adulto, esses sujeitos apresentam propriedades distintas dos sujeitos que são duplicados em outras línguas como o francês, a saber:

A) pode não haver pausa entre o sujeito e o pronome que o duplica em PB (2), ao contrário do francês (3):

- (2) **A Clarinha_i** *ela_i* cozinha que é uma maravilha. (DUARTE, 2000, p. 28)

- (3) **Kester_i**, *il_i* dit qu’*il_i* aime bien les poissons.
“Kester, ele diz que ele gosta muito de peixe.” (DE CAT, 2003, p. 11)

B) sintagmas nominais quantificados podem ocorrer em estruturas com SDs em PB (4), diferentemente do francês (5):

- (4)a. **Toda criança** *ela* aprende rápido a gostar de Coca-Cola.
b. **Todo homem** *ele* odeia se sentir fraco. (BRITTO, 2000, p. 202)

- (5) ***Tout homme**, *il* est mortel.
“Todo homem, ele é mortal.”

(DE CAT, 2003, p. 9)

Outra diferença apontada por Costa, Duarte e Silva (2004, p. 140) entre o PB e o francês é que “[...] a construção com redobro é estatisticamente irrelevante em dados disponíveis de aquisição do PB – é inexistente no *corpus* de Gonçalves (2004), e Grolla (2000) só as atesta na criança que estudou a partir dos 3;1, contrariamente ao que acontece no *corpus* do francês de De Cat (2000)”. A partir disso, a hipótese norteadora deste estudo é a seguinte: já que crianças brasileiras só começam a produzir estruturas com duplicação do sujeito tardiamente, esperamos que, nos anos iniciais de escolarização, ainda produzam menos e, nos anos intermediários, venham a produzir mais, vindo a serem muito pouco produzidos nos anos finais. Em outras palavras, a escolarização exerceria influência, sobretudo, nos anos finais para o decréscimo acentuado de SDs na escrita. Observem-se, a seguir, alguns dados extraídos do *corpus* da pesquisa que evidenciam SDs em diferentes níveis de escolarização:

- (6)a. “**o bebê** *ele* trocou seus super poderes com um ser humano que passava por ali quando aconteceu a fusão.” (Ensino Fundamental I - 5º ano)
b. “**Lilian** *ela* como tem o poder de voar *ela* voa muito” (Ensino Fundamental II – 6º ano)
c. “**A família** *ela* voltou para o hotel em que estavam hospedados e logo quando chegaram lá” (Ensino Médio - 1º Ano)

Nos contextos acima, os sintagmas nominais (SNs) em negrito são retomados por um pronome resumptivo (pronome lembrete) que se encontra em itálico. Portanto, trata-se de uma *Construção com Deslocamento à Esquerda do Sujeito* pouco frequente em outras línguas românicas, como o PE e o espanhol, línguas de sujeito nulo consistente. Em (6), portanto, os sujeitos encontram-se em uma posição de tópico marcado (DUARTE, 2003), estando o pronome resumptivo no domínio intra-sentencial e o sintagma nominal em uma posição de tópico deslocado à esquerda.

Em sua pesquisa, Salles (2004) conclui que SDs: (a) são muito frequentes na língua falada, (b) ocorrem exclusivamente com sujeitos de segunda e terceira pessoa, (c) não dependem de um corte entonacional (marcado graficamente pela vírgula) e (d) não estão associados a efeitos de definitude por poderem ocorrer com sintagmas indefinidos (7a) e quantificados (7b):

- (7)a. **Uma criança**, *ela* gosta de brinquedo.
b. **Toda criança**, *ela* gosta de brinquedo.

Com base no exposto até o momento, é nosso objetivo central neste artigo discutirmos, sob o enfoque da Teoria gerativa a partir do Modelo do Princípios e Parâmetros (CHOMSKY, 1981, 1986), os SDs em dados de escrita de alunos recifenses, levando em conta o possível distanciamento entre o conhecimento gramatical que o aluno traz para a escola (aquisição) e o que ele aprende (aprendizagem), no sentido de que aquele é reflexo do que ocorre na língua falada, ao passo que este é aprendido e terá reflexos na língua escrita. Em outras palavras, nos dados de aquisição da fala do PB, por serem pouco produzidos SDs, segundo observam Gonçalves (2004) e Grolla (2000) *apud* Costa, Duarte e Silva (2004), esperamos encontrar poucos em dados de escrita nos anos

iniciais, sendo produtivos nos níveis intermediários e vindo a declinarem nos níveis finais por influência do avançar da escolarização.

Ademais, são objetivos específicos: a) verificar que fatores de ordem (extra)linguística podem favorecer o uso de SDs nos dados escritos submetidos à análise e b) analisar, a partir dos dados selecionados, se há uma possível assimetria entre aquisição e aprendizagem, tomando por base a hipótese norteadora deste estudo. Para tanto, serão analisados dados extraídos de narrações produzidas por alunos de escolas da rede estadual e municipal da cidade do Recife-PE.

Em linhas gerais, o estudo aqui proposto torna-se relevante na medida em que as pesquisas, até onde temos verificado, centram a atenção nos SDs, sobretudo, na língua falada, não oferecendo uma análise adicional pautada na comparação entre o conhecimento gramatical adquirido naturalmente pela criança e o conhecimento que aprende via escolarização. Além disso, trata-se de um estudo pioneiro no estado de Pernambuco, contribuindo para os estudos linguísticos sobre o PB no que se refere, em específico, ao Parâmetro do Sujeito Nulo e às estratégias de preenchimento da posição pré-verbal do sujeito na gramática dessa língua.

Vale referirmos ainda que o artigo encontra-se assim estruturado: na seção 2, discorreremos sobre o aporte teórico que norteará a análise; na seção 3, centraremos nossa atenção nos procedimentos metodológicos necessários à realização dessa análise; na seção 4, será realizada a análise dos dados e, na seção 5, apresentaremos as considerações finais.

2. Referencial teórico

2.1 Teoria gerativa: pressupostos básicos

A partir do pressuposto de que todo ser humano já nasce com um dispositivo genético específico para a linguagem, a teoria gerativa, proposta por Noam Chomsky na década de 50, inscreve-se como uma teoria mentalista. Para os estudiosos da ciência cognitiva, a capacidade que todo falante de uma língua natural tem para produzir sentenças e julgá-las como (a)gramaticais só é possível porque há estruturas cerebrais que permitem que a linguagem se processe de maneira natural na mente humana. Nesse contexto, a linguística gerativa tem um papel importante no intuito de compreender, por exemplo, como o conhecimento linguístico internalizado é estruturado na mente humana. Portanto, são questões básicas dessa teoria:

- (i) O que constitui o conhecimento da língua? (NATUREZA)
- (ii) Como é adquirido o conhecimento da língua? (ORIGEM)
- (iii) Como é usado o conhecimento da língua? (USO)

(CHOMSKY, 1986, p. 23)

Estando a teoria enquadrada, portanto, em uma perspectiva cognitiva, assume-se a existência de uma Faculdade da Linguagem (FL), ou seja, de um órgão mental responsável pela linguagem, inerente ao ser humano, através da qual é possível ao indivíduo produzir e compreender infinitas estruturas frásicas e fazer intuições, por exemplo, sobre o conhecimento gramatical de natureza fonológica, sintática e semântica de sua língua materna. Segundo Chomsky (1986, p. 22), a FL é “uma componente particular da mente humana.”, o que torna a sua teoria de base mentalista.

Sobre as estruturas fráscas de uma língua, é característico da linguagem humana fazer associação entre *som* e *significado*. Dessa forma, a linguagem é concebida como um sistema organizado de acordo com princípios e regras que possibilitam a formação de expressões linguísticas ordenadas. Nesse contexto, a relação entre som e significado gera o par (π, λ) em que π é a forma fonética e “ λ ”, a forma conceitual-intencional, considerados sistemas de interface: o primeiro, o sensorio-motor (a forma fonética (FF)) e o segundo, o conceitual-intencional (a forma lógica (FL)).

Vale referirmos ainda que, associada à FL, encontra-se a Gramática Universal (GU) considerada o estágio inicial geneticamente determinado e necessário à aquisição de uma língua natural: “[...] UG [Universal Grammar] is taken to be a characterization of the child’s pre-linguistic initial state³” (CHOMSKY, 1981, p. 7). Na GU já estão previstas as propriedades invariantes (princípios) e as que distinguem as línguas naturais (os parâmetros). Sobre os parâmetros, apresentam valores binários (+) e (-), sendo o papel da criança fixar um dos valores para a aquisição de uma língua natural. Após fixar o valor paramétrico, a criança, por volta dos 5 anos de idade, já possui uma gramática nuclear de sua língua, ou seja, um conhecimento gramatical internalizado, a saber, a língua-I (interna, individual, intensional). Logo, a língua-I, objeto de estudo da teoria gerativa, distingue-se da língua-E (extensional) que se enquadra no âmbito da performance por voltar-se ao “uso efectivo da língua em situações concretas” (CHOMSKY, 1965, p. 84).

É importante relembrarmos que, para a aquisição da língua-I, valores paramétricos foram fixados pela criança no processo de aquisição. Para tanto, sua exposição aos dados linguísticos primários (*input*) de sua língua é crucial para que essa aquisição ocorra, valendo referirmos que, ao contrário da perspectiva behaviorista, não se trata de um processo embasado em estímulo-resposta, pois o falante-ouvinte é capaz de produzir e entender sentenças nunca antes ouvidas, o que evidencia um dos princípios fundantes da linguagem humana e tão defendido pela teoria gerativa: a criatividade. Nessa linha de raciocínio, Miotto, Figueiredo Silva e Lopes (2005, p. 18) assumem que “só os seres humanos são capazes de combinar itens de um conjunto de elementos segundo certos princípios básicos, que são em número finito, de modo a gerar um número infinito de sentenças novas [...]”.

2.2 Sobre o Modelo de Princípios e Parâmetros

Conforme já mencionado na seção anterior, a GU é composta de princípios e parâmetros, sendo aqueles relacionados às propriedades universais que estão presentes nas línguas, e estes, às diferenças gramaticais observáveis entre elas cujos valores já estão dados inatamente e precisam ser fixados pela criança no processo de aquisição. É sob esse enfoque que Chomsky, na década de 80, propõe o chamado Modelo de Princípios e Parâmetros (P&P) (CHOMSKY, 1981, 1986).

A partir da comparação no âmbito da sintaxe entre línguas de diversas famílias linguísticas, Chomsky observa que há um princípio que as une, a saber: a posição sujeito é sempre projetada em contextos finitos (o chamado Princípio de Projeção Estendido (EPP, do inglês *Extended Projection Principle*)), no entanto, essa posição deve ser preenchida ou por sujeitos realizados foneticamente (sujeitos plenos) nas chamadas línguas de sujeito não-nulo (*línguas não-pro-drop*) como o inglês e o francês, ou pode ser preenchida por sujeitos não-realizados foneticamente (sujeitos nulos) nas chamadas

³ “[...] A GU [Gramática Universal] é considerada uma caracterização do estágio inicial pré-linguístico da criança.” – Tradução nossa

línguas de sujeito nulo (*línguas pro-drop*), como o catalão e o italiano. A partir disso, é proposto na teoria o chamado Parâmetro do Sujeito Nulo (PSN) cujos valores (+) e (-) definirão se a língua é pro-drop ou não-pro-drop, respectivamente. Portanto, em línguas como o italiano (It.) e o catalão (Cat.), o valor positivo é fixado pela criança em fase de aquisição, ao passo que, no inglês (Ing.) e no francês (Fr.), o valor é negativo. Observem-se os exemplos extraídos de Rizzi (1997, p. 271-272) que evidenciam essa diferença em diferentes contextos frásicos:

- (8) a. ___ parla.(It.) (“fala”)
 b. ___ parla. (Cat.) (“fala”)
 c. * ___ parle. (Fr.) (“fala”)
 d. * ___ speaks. (Eng.) (“fala”)

- (9) a. ___ piove..(It.) (“chove”)
 b. ___ plou. (Cat.) (“chove”)
 c. * ___ pleut. (Fr.) (“chove”)
 d. * ___ rained. (Eng.) (“chove”)

- (10)a. *Ciòpiove.(It.) (chove”)
 b. *Aixòplou. (Cat.) (chove”)
 c. Il pleut/ça pleut. (Fr.) (chove”)
 d. It is raining. (Eng.) (“está chovendo”)

- (11)a. ___ ha telefonato Gianni.(It.) (“Telefonou João.”)
 b. ___ ha telefonat en Joan. (Cat.) (“Telefonou João.”)
 c. * ___ a téléphoné Jean. (Fr.) (“Telefonou João.”)
 d. * ___ telephoned John. (Eng.) (“Telefonou João.”)

- (12)a. Chi credi che ___ telefonerà? (It.) (“Quem achas que telefonará?”)
 b. Qui creus que ___ telefonará? (Cat.) (“Quem achas que telefonará?”)
 c. *Qui crois-tu que ___ téléphonera?(Fr.) (“Quem achas que telefonará?”)
 d. *Who do you think that ___ will telephone. (Eng.) (“Quem achas que telefonará?”)

- (13)a. It.: parl-o, parl-i, parl-a, parl-iamo, parl-ate, parl-ano
 “falo, falas, fala, falamos, faláis, falam”
 b. Cat.: parl-o, parl-es, parl-a, parl-em, parl-eu, parl-em
 c. Fr.: parl, parl-õ, parl-é
 d. Eng.: speak, speak-s

Chomsky (1981, 1986) argumenta que o PSN está associado a um conjunto de propriedades, a saber:

1. Sujeitos nulos referenciais ocorrem em línguas pro-drop (8a) e (8b);
2. Sujeitos nulos expletivos ocorrem em línguas pro-drop (9a) e (9b);

3. Sujeitos pospostos ocorrem com diferentes tipos de verbos em línguas pro-drop, deixando a posição pré-verbal de sujeito vazia (11a) e (11b);
4. Constituintes interrogativos com função de sujeito podem ser movidos nas interrogativas, deixando a posição de sujeito da encaixada vazia (12a) e (12b);
5. Línguas pro-drop apresentam morfologia de flexão verbal rica, sendo as desinências número-pessoais de cada pessoa do paradigma pronominal responsáveis pela legitimação e licenciamento dos sujeitos nulos (*Generalização de Taraldsen, 1978*) (13a) e (13b).

Pesquisas iniciais realizadas sobre o PSN no PB apontavam que essa língua estava passando por um processo de mudança, deixando de ser uma língua pro-drop para uma língua não-pro-drop, o que levou Duarte (1995) a denominar de Perda do Princípio Evite Pronome. Esse processo de mudança está associado, sobretudo, ao enfraquecimento da morfologia de flexão verbal por apresentar um decréscimo de seis desinências número-pessoais para quatro, decorrência das formas zero na segunda pessoa do singular (*tu*) e do plural (*vós*), conforme evidenciado no paradigma do PB a seguir:

Quadro 1: Paradigma de flexão verbal do português brasileiro e do português europeu extraído de Galves (2001, p. 103)⁴

Português Brasileiro	Português Europeu
(eu) canto	(eu) canto
	(tu) cantas
você canta ele canta	você canta ele canta
(nós) cantamos	(nós) cantamos
vocês cantam eles cantam	vocês cantam eles cantam

Para Duarte (2000), uma das evidências desse enfraquecimento decorre do surgimento de diversas estratégias de preenchimento da posição pré-verbal por sujeitos plenos, destacando-se, dentre elas, as construções com SDs apresentadas na introdução deste artigo tão comuns na língua falada dos brasileiros, uma situação observada com grande frequência em uma língua não-pro-drop como o francês:

(14) **Madame**_i, dist il froidement, **elle**_i se recommande tres humblement a vostre bonne grace et **pro**_i dis que **pro**_j l'avrez_j briefement.

“**A senhora**, disse ele friamente, **ela** se posicionou muito humildemente a seu bom grado e disse que o teria brevemente.”

(DUARTE, 2000, p. 30)

⁴ Galves (2001, p. 124) também verifica que alguns dialetos do PB “mostram contraste apenas entre a primeira pessoa do singular e toda as outras: eu canto/ você, nós, eles canta”. Ademais, é necessário dizermos que o paradigma do PE é baseado no dialeto de Lisboa.

No exemplo acima, o constituinte *Madame* é retomado pelo pronome lembrete *elle* na posição de sujeito. Mesmo se tratando de uma língua não-pro-drop, é possível que seja encontrado sujeito nulo (*pro*) em francês em contextos de coordenação quando o sujeito é correferente com o referente já dado no domínio do discurso, nesse caso, *Madame*. Por isso, a utilização do mesmo índice referencial i (pro_i). Nesse exemplo, observamos que a adjacência entre o constituinte *Madame* e o pronome *elle* é quebrada, o que implica considerarmos que o primeiro encontra-se em uma posição de tópico e o segundo, no domínio intrafrásico, tratando-se, portanto, de uma construção de tópico marcado. Seguindo essa perspectiva de análise, Pontes (1986), em sua pesquisa pioneira sobre o tópico no PB, considera que essa língua é de orientação para o tópico devido, por exemplo, à grande frequência de construções com SDs.

Não obstante, estudos como o de Silva (2004) e Costa, Duarte e Silva (2004) mostram evidências de que as construções com SDs no PB diferem em alguns aspectos das construções com SDs em francês, o que os leva a aventarem a hipótese de que essas construções não decorrem necessariamente do PSN, mas do enfraquecimento da morfologia de flexão verbal, sobretudo, do *déficit* morfológico da segunda pessoa. Um dos *locus* de diferença entre essas línguas é que, enquanto em PB são muito frequentes SDs de terceira pessoa, em francês predominam a primeira e segunda pessoas (COSTA, DUARTE; SILVA, 2004).

Com o avanço dos estudos sobre o PSN no PB, é sobejamente assumido hoje que o PB não está deixando de ser uma língua pro-drop, mas permanece sendo um outro sistema de língua pro-drop (RODRIGUES, 2004; HOLMBERG, 2010; HOLMBERG; NAYUDU; SHEEHAN, 2009; BUTHERS, 2009; ALEXANDRE; DUARTE; SANTOS, 2013; KATO; DUARTE, 2014a, 2014b; SAAB, 2016) que difere, por exemplo, do sistema pro-drop do PE, do italiano e do grego, perspectiva adotada neste artigo.

A ideia aqui assumida tem sua gênese na proposta de hierarquias paramétricas, de Roberts e Holmberg (2010) e Roberts (2012), segundo a qual os parâmetros continuam com valores binários, não sendo possível pensarmos em uma língua que tenha fixado, ao mesmo tempo, o valor (+) e (-) para um dado parâmetro em decorrência de um processo de mudança tal como foi proposto por alguns estudiosos para o PB ao considerá-la uma língua semi-pro-drop (SILVA, 2004).

Segundo Holmberg (2005), Holmberg, Nayudu e Sheehan (2009), Roberts e Holmberg (2010), há quatro sistemas de sujeito nulo, a saber:

- a) *línguas de sujeito nulo consistente*, como o espanhol e o PE: sujeitos nulos são produzidos em decorrência da riqueza da morfologia de flexão verbal (TARALDSEN, 1978);
- b) *línguas de sujeito nulo radical*, como o chinês: sujeitos nulos são produzidos sob restrições discursivas (HUANG, 1984, 1989);
- c) *línguas de sujeito nulo parcial*, como o hebraico (BORER, 1986), o finlandês e o PB (HOLMBERG, 2005): sujeitos nulos estão submetidos a contextos estruturais específicos;
- d) *línguas de sujeito nulo expletivo*, como o alemão: apenas os expletivos são nulos (ROBERTS; HOLMBERG, 2010).

Segundo a classificação acima, o PB enquadra-se no sistema das línguas de sujeito nulo parcial. Vejam-se as seguintes evidências apresentadas em Kato e Duarte (2014b, p. 8-10) não verificadas nos outros sistemas pro-drop:

a) sujeito nulo em sentenças encaixadas em variação com sujeitos pronominais:

(15)a. O João_i disse que $\emptyset_{i/*j}$ comprou um carro ontem.

b. O João_i disse que *ele*_{i/j} comprou um carro novo.

b) sujeito nulo genérico em variação com sujeitos genéricos pronominais:

(16)a. \emptyset_{gen} não pode fumar aqui.

b. *A gente* não pode fumar aqui.

c. *Você* não pode fumar aqui.

c) nulo expletivo em variação com construções pessoais:

(17)a. \emptyset_{expl} está entrando água por essas janelas.

b. *Essas janelas*_i estão entrando água [t]_j.

Diante do exposto, assumimos neste artigo que o PB é uma língua de sujeito nulo parcial e que construções com SDs nessa língua não têm a ver necessariamente com o PSN, mas com o enfraquecimento da morfologia de flexão verbal, tal como defendido por Silva (2004) e Costa, Duarte e Silva (2004). Além disso, sendo a duplicação do sujeito um fenômeno frequente na gramática da fala dos brasileiros (SALLES, 2004), acreditamos que o aumento da escolarização possa ser responsável por uma assimetria entre essa gramática e a “gramática” da escrita no sentido de nesta SDs serem poucos frequentes.

3. Procedimentos metodológicos

Neste estudo foram utilizados o método estatístico para obtenção dos resultados quantitativos e o método comparativo para estabelecer uma análise entre os níveis de escolarização das escolas selecionadas. Ademais, para a coleta dos dados escritos, baseamo-nos nos textos que já haviam sido coletados em 2011 pelo aluno do Programa de Bolsas de Iniciação Científica (PIBIC), Edimar Gonçalves de Barros Filho, quando estava vinculado ao Curso de Licenciatura em Letras da Universidade Federal Rural de Pernambuco (UFRPE)/ Campus Dois Irmãos.

O texto produzido pelos alunos foi predominantemente narrativo, sendo o gênero estória, e essa atividade de produção textual foi aplicada pelos professores das turmas. Além disso, houve a preocupação de uniformizar a temática das narrações em todas as turmas e, para tanto, foi dado o início da estória para que os alunos dessem continuidade. Acreditamos que esse tipo de texto estimulou bastante a imaginação dos alunos, envolvendo-os de tal forma com o enredo de uma estória voltada a uma família de super-heróis, que os levou a monitorarem-se menos quanto à escrita de seus textos. Essa situação é apropriada para testarmos nossa hipótese norteadora de que, nos anos finais,

haja menos SDs pelo fato de os alunos estarem mais conscientes do uso linguístico previsto para a modalidade escrita requerida no ambiente escolar. Com base nos textos coletados, foram selecionadas 10 narrações de cada ano de escolarização, totalizando 110 textos.

Feita a seleção das histórias, procedemos à montagem do *corpus* que é composto por frases declarativas finitas, contendo SDs. Além disso, os dados foram analisados consoante às seguintes variáveis linguísticas e extralinguísticas: a) variável dependente (presença *versus* ausência de SDs), b) Tipo de representação do sujeito (sujeito duplicado com sintagma nominal, sujeito duplicado com quantificadores/ indefinidos e não se aplica), c) traço de pessoa/número do sujeito (primeira pessoa do singular e do plural, segunda pessoa do singular e do plural, terceira pessoa do singular e do plural) e anos de escolarização (2º, 3º, 4º e 5º ano do Ensino Fundamental I (rede municipal); 6º, 7º, 8º e 9º ano do Ensino Fundamental II e 1º, 2º e 3º ano do Ensino Médio (rede estadual)).

Selecionadas as variáveis e os fatores que as compõem, os dados foram codificados e submetidos a tratamento estatístico para posterior análise linguístico-comparativa cujos resultados obtidos encontram-se na próxima seção.

4. A análise dos dados

Ao analisar os SDs em PB, Moraes de Castilho (2001, p. 61) verifica que o sujeito duplo seria um tipo específico de tópico, já que o tópico é formado por uma estrutura frásica ou um nome movido geralmente à esquerda das orações. Sobre as propriedades do tópico, Li e Thompson (1976) elencam as seguintes:

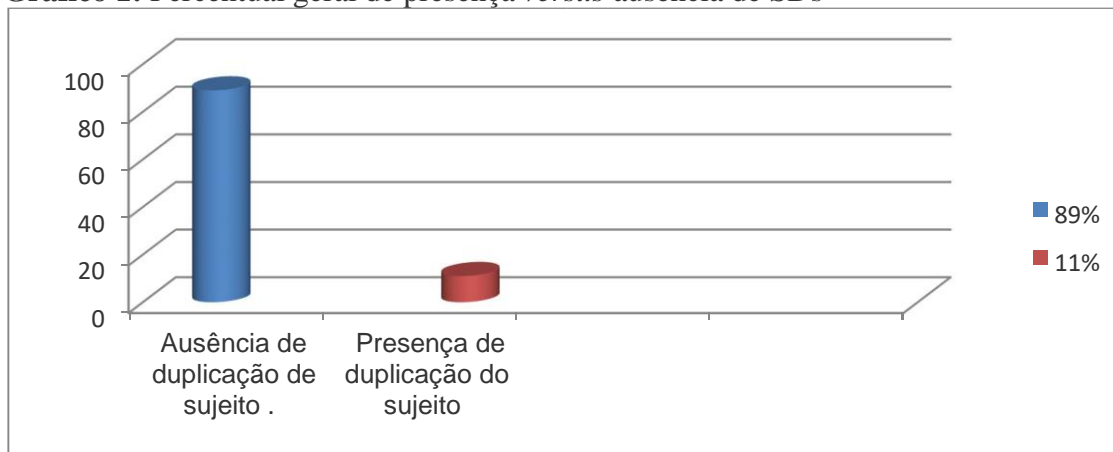
- a) é definido;
- b) existe de forma independente em relação ao verbo;
- c) não é determinado pela transitividade;
- d) não se apresenta em posição pós-verbal.

Não obstante, diversos estudos realizados sobre os SDs no PB, verificam que nem sempre as propriedades acima são observadas, pois podem ocorrer SDs com sujeitos indefinidos e quantificados, por exemplo (BRITTO, 2000; DUARTE, 2000; SILVA, 2004). É nesse sentido que os SDs em PB distinguem-se dos SDs em francês, segundo observam Costa, Duarte e Silva (2004) (cf. evidências já apresentadas na introdução deste artigo).

Ademais, apesar de estudos linguísticos apontarem que SDs sejam frequentes na língua falada, não encontramos, até onde temos verificado, um estudo como o aqui proposto que contemple o comportamento desses sujeitos em dados escritos produzidos em diferentes anos de escolarização.

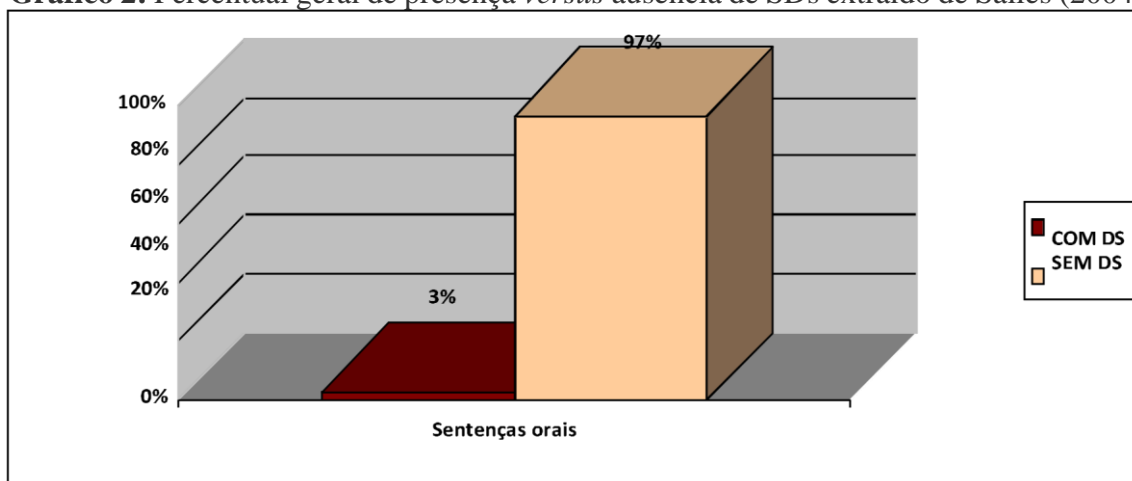
Centrando nossa atenção a partir de agora nos resultados referentes às variáveis selecionadas, verificamos que a variável dependente revela que SDs são muito pouco produzidos nos dados de escrita analisados (gráfico 1), à semelhança dos dados orais obtidos por Salles (2004) (gráfico 2) por razões que serão expostas a seguir:

Gráfico 1: Percentual geral de presença *versus* ausência de SDs



Fonte: Autoras deste trabalho

Gráfico 2: Percentual geral de presença *versus* ausência de SDs extraído de Salles (2004)



Fonte: Salles (2004, p. 68)

Observamos o baixo percentual de uso de SDs tanto em nossa pesquisa (11%) quanto na pesquisa de Salles (3%), um resultado inesperado, haja vista que estudos linguísticos apontam para o fato de esses sujeitos serem mais frequentes na gramática da fala (DUARTE, 2000, KATO; DUARTE, 2014b).

Salles (2004), em sua pesquisa explica que o SD ocorre com frequência em frases que expressam juízos categóricos, ou seja, estruturas que são organizadas a partir da relação sujeito-predicado. Logo, a frase é formada por dois atos separados, o sujeito e a afirmação ou negação expressa pelo predicado sobre o sujeito. Entretanto, no seu trabalho, o baixo número de SDs ocorreu por conta do gênero entrevista, em que foi frequente o uso de frases com juízos téticos, isto é, de construções que, a partir de um ato cognitivo, no qual, é apresentado um evento, estado, objeto ou entidade, não apresentaram a estrutura S(ujeito)-V(erbo).

Diferentemente de nossa pesquisa, em que o gênero utilizado é uma estória com o predomínio do texto narrativo favorecendo a produção de juízos categóricos, verificamos também o baixo quantitativo de SDs. Dessa forma, comparando os nossos resultados com

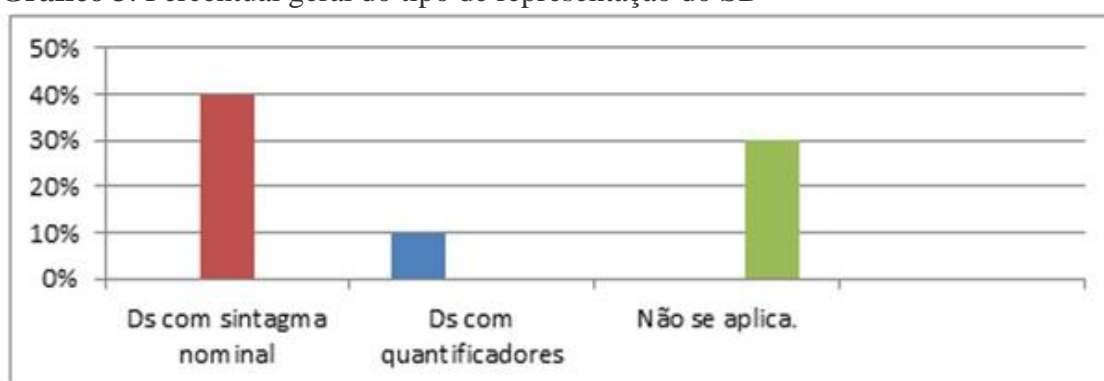
os de Salles (2004), podemos dizer que, independentemente do tipo de gênero textual, SDs ainda são pouco frequentes.

No que se refere à variável *Tipo de representação do sujeito*, encontra-se composta pela Categoria 1 (sujeito duplicado com sintagma nominal (18), Categoria 2 (sujeito duplicado com quantificadores/indefinidos (19) e Categoria 3 (não se aplica); nesse último caso, estão incluídos, por exemplo, pronome duplicado por outro pronome (20), elemento pronominal duplicado por um nome (21) e nome duplicado por outro nome (22):

- (18) “[...] **o filho mais velho** *ele* é muito curioso (Ensino Fundamental II- 8º ano)
- (19) “**Um homem bomba** *ele* se explodiu” (1º ano do Ensino Médio)
- (20) “**Ele** *ele* usou seus super poderes “ (Ensino Fundamental II - 6º ano)
- (21) “**Ela Lilian** pegou o bebe” (Ensino Fundamental II- 6º ano)
- (22) “**o bebe** *o bebe* com seus poderes sacultu o homem até que ele vomitou (Ensino Fundamental II – 6º ano)

No gráfico, a seguir, verificamos que SDs são muito comuns na *Categoria 1* (40%): sintagmas nominais plenos tendem a ser duplicados por um pronome resumptivo em posição sujeito, um resultado que vai ao encontro do que já foi verificado em outras pesquisas linguísticas no âmbito da língua falada (SILVA, 2013).

Gráfico 3: Percentual geral do tipo de representação do SD



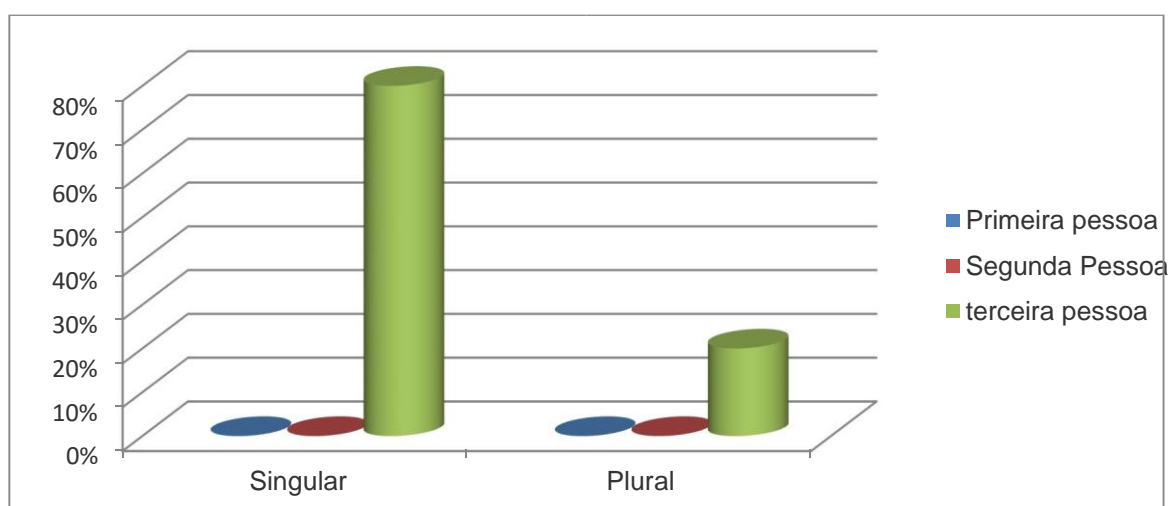
Fonte: Autoras deste trabalho

Quanto à variável *Traço de pessoa/número do sujeito*, observemos exemplos extraídos do *corpus* da pesquisa:

- (23) “O bebê *ele* trocou seus super poderes com um ser humano que passava por ali quando aconteceu a fusão.” (Ensino Fundamental II - 5º ano)
- (24) “Os heróis *eles* não tinham percebido mas o bebê estava crescendo (...)” (Ensino Fundamental II – 8º ano)

No gráfico 4, fica evidente que só ocorrem SDs de terceira pessoa do singular e do plural, indo ao encontro do que Costa, Duarte e Silva (2004) observam para os dados da língua falada do PB.

Gráfico 4: Percentual geral do traço pessoa/número do sujeito duplicado



Fonte: Autoras deste trabalho

Uma possível explicação para a existência exclusiva da terceira pessoa nos dados escritos analisados deve-se ao tipo de gênero textual produzido: uma estória cujo enfoque é sobre um referente diferente de quem a narra.

Verificando os resultados obtidos por Salles (2004) na tabela 1, observamos que a segunda pessoa do singular (peso relativo: 0,68) e a terceira pessoa do plural (peso relativo: 0,62) e do singular (peso relativo: 0,56) favorecem o uso de SDs na língua falada por razões relacionadas ao tipo de gênero: a entrevista. Nesse gênero, é comum que o entrevistado não só se refira a uma terceira pessoa, como também faça uso do pronome *you* com referência definida quando dirigido ao entrevistador ou com referência genérica:

Tabela 1: Variável Pessoas do Sujeito

Pessoa	PR em relação ao input
2ª pessoa do singular (you)	0,68
3ª pessoa do plural	0,62
3ª pessoa do singular	0,56
1ª pessoa do plural – nós	0,48
1ª pessoa do singular	0,45
1ª pessoa do plural – a gente	0,21

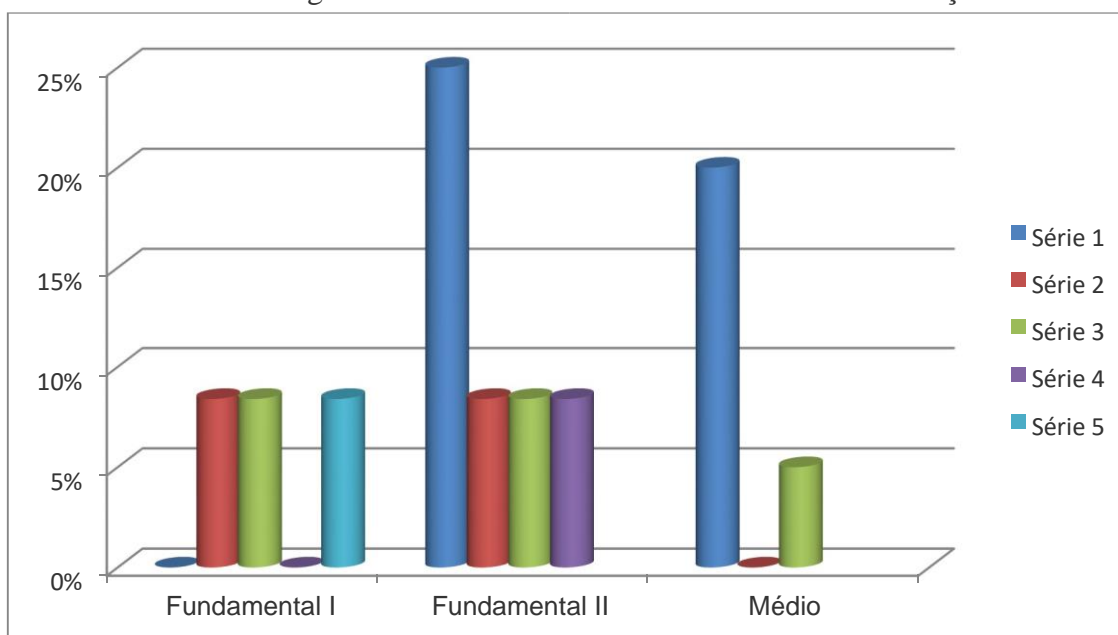
Fonte: Salles (2004, p. 78).

Em linhas gerais, ao contrário do que se observa em francês, tanto nos dados obtidos neste artigo quanto nos dados de Salles (2004), são pouco produtivos SDs de primeira pessoa no singular e no plural, corroborando a pesquisa de Costa, Duarte e Silva (2004).

Observando agora a variável extralinguística *anos de escolarização* cujos fatores são: a) Ensino Fundamental I (2º ao 5º ano) de uma escola municipal, b) Ensino

Fundamental II (6º ao 9º ano) de uma escola estadual e c) Ensino Médio (1º ao 3º ano) de uma escola estadual, foram obtidos os seguintes resultados⁵:

Gráfico 10: Percentual geral de SDs de acordo com os anos de escolarização



Fonte: Autoras deste trabalho

A partir dos resultados apresentados, é possível verificarmos que:

- no Ensino Fundamental I, o primeiro e o quarto anos não apresentam SDs. Em relação à percentagem total dos SDs encontrados, esse nível de ensino tem uma incidência de 8,4% nas demais séries;
- no Ensino Fundamental II, o sexto ano que reflete bem a série intermediária é o que apresenta maior índice de SDs (25%), ao passo que as demais séries seguem com 8,4%;
- no Ensino Médio, há 20% de SDs no primeiro ano provavelmente pelo fato de ser esse um nível de ingresso de adolescentes que, em busca de autoafirmação, usam a língua como um espaço identitário, podendo a escrita refletir um pouco a língua falada desses sujeitos; no segundo ano, não ocorrem SDs, e, no terceiro ano, há apenas 5%. Esse resultado reflete uma mudança decrescente em relação à incidência de SDs que julgamos ser resultante da escolarização: os alunos procuram escrever de forma mais próxima à norma culta quanto mais o nível de escolaridade aumenta.

⁵ Os gráficos apresentam à direita uma legenda em que a série 1 representa a primeira série de cada módulo de ensino, isto é, a série 1 representa o 1º ano do ensino fundamental I, o sexto ano do ensino fundamental II e o 1º ano do ensino médio. A série 2 refere-se ao 2º ano do fundamental I, ao 7º ano do fundamental II e ao 2º ano do ensino médio. A série 3 refere-se ao 3º ano do fundamental I, ao 8º ano do fundamental II e ao 3º ano do ensino médio. A série 4 refere-se ao 4º ano do fundamental I e ao 9º ano do fundamental II, e a série 5 refere-se ao 5º ano do fundamental I.

Os resultados apontam, portanto, para a nossa hipótese norteadora de que, nos anos iniciais, existe pouca frequência dos SDs em decorrência do fato de serem de aquisição tardia na fala, sendo mais frequentes nos anos intermediários e vindo a declinarem nos anos finais, distinguindo-se assim do que ocorre na gramática da fala do falante adulto.

Em seu estudo, Salles (2004) verificou que, na língua falada, falantes com nível médio (M) (0,53) e universitários (U) (0,48) são os que mais produzem SDs, um resultado que vai de encontro ao da língua escrita aqui obtido:

Tabela 2: Variáveis sociolinguísticas extraídas de Salles (2004)

Sexo	PR	Nível de escolaridade	PR	Faixa etária	PR
M	0,48	F	0,46	15 a 25	0,40
F	0,51	M	0,53	26 a 35	0,56
		U	0,49	36 a 49	0,60
				50 ou +	0,42

Fonte: Salles (2004, p. 85)

A partir dos resultados apresentados nesta seção, ficou evidenciado que a duplicação do sujeito em dados de escrita de alunos recifenses da rede pública ocorre principalmente com sintagma nominal de terceira pessoa do singular e o pronome é o elemento que o redobra com maior frequência. Em relação à variável extralinguística, SDs são mais frequentes nas séries intermediárias, um resultado que não se verifica na língua falada quando observado o nível médio (cf. Tabela 2 acima).

Em linhas gerais, observamos que a assimetria entre a gramática da fala e a “gramática” da escrita não decorre de fatores intrínsecos à língua, mas de fatores extrínsecos a ela. Embora SDs sejam frequentes na língua falada nos anos finais de escolarização (SALLES, 2004), não são produtivos na língua escrita, pois a escolarização atua como uma espécie de “freio” às inovações encontradas na língua falada, o que culmina, por conseguinte, na dicotomia aquisição (gramática da fala) *versus* aprendizagem (“gramática” da escrita).

5. Considerações finais

A análise aqui desenvolvida revela que SDs não são tão frequentes nos dados de escrita, sendo fatores linguísticos condicionantes para o seu uso: a) sujeitos de terceira pessoa e b) sintagmas nominais não-quantificados à semelhança do que foi encontrado em Salles (2004).

Ademais, a hipótese norteadora deste estudo corrobora-se no sentido de que, sendo os SDs de aquisição tardia em PB (GONÇALVES, 2004 *apud* COSTA, DUARTE, SILVA, 2004), ao contrário do que ocorre em francês, as séries iniciais apresentam poucos SDs. Não obstante, nas séries intermediárias, observamos um uso acentuado desses sujeitos que vai entrando em decréscimo no ensino médio. Esse decréscimo decorre, portanto, da instrução formal, ou seja, da aprendizagem e não de um processo

natural de aquisição, o que culmina na assimetria entre a gramática da fala e a “gramática” da escrita.

Referências

- ALEXANDRE, N.; DUARTE, I.; SANTOS, A. L. Infinitivos pessoais: uma viagem transatlântica. In: MOURA, M. D.; SIBALDO, M. A. (Org.). **Para a história do português brasileiro**. Maceió: EDUFAL, 2013, p. 17-47.
- BORER, H. **I-subjects**. *Linguistic Inquiry*, v. 17, n. 3, p. 375-416, 1986.
- BRITTO, H. Syntactic codification of categorical and thetic judgments in Brazilian Portuguese. In: KATO, M. A.; NEGRÃO, E. V. (Eds.). **Brazilian Portuguese and the null subject parameter**. Madrid: Iberoamericana, 2000. p. 195-222.
- BUTHERS, C. M. **Emergência da ordem [XP V (DP)] no PB contemporâneo e o parâmetro do sujeito nulo: uma abordagem minimalista**. Dissertação (Mestrado em Linguística – Universidade Federal de Minas Gerais, Belo Horizonte, 2009).
- CHOMSKY, N. **Aspects of the theory of Syntax**. Cambridge, Mass: The MIT Press, 1965.
- _____. **Lectures on government and binding**. Dordrecht: Foris Publications, 1981.
- _____. **Knowledge of language: its nature, origin and use**. London: Praeger Publishers, 1986.
- COSTA, J.; GALVES, C. External subjects in two varieties of Portuguese evidence for a non-unified analysis. In: Beyssade, Claire. et al. **Romance languages and linguistic theory 2000**, Utrecht, 30 November-2 December. v. 232. Amsterdam; Philadelphia: John Benjamins Publishing Company, 2002. p. 109-125.
- COSTA, J.; DUARTE, I.; SILVA, C. **Construções de redobro em português brasileiro: sujeitos tópicos vs. Soletração do traço de pessoa**. *Leitura*, n. 33, 2004, p. 135-145.
- DE CAT, C. **French dislocation without movement**. A minimalist account. Ms., 2003.
- DUARTE, I. Relações gramaticais, esquemas relacionais e ordem de palavras. In: MATEUS, M. H. M. et al. **Gramática da Língua Portuguesa**. 5. ed. rev. aum. Lisboa: Caminho, 2003, p. 275-321.
- DUARTE, M. E. L. **A perda do princípio ‘Evite Pronome’ no português brasileiro**. Tese (Doutorado em Linguística) – Universidade Estadual de Campinas, 1995.
- _____. The loss of the ‘avoid pronoun’ principle in Brazilian Portuguese. In: KATO, M. A.; NEGRÃO, E. V. (Eds.). **Brazilian Portuguese and the null subject parameter**. Madrid: Iberoamericana, 2000. p. 17-36.
- GALVES, C. C. **Ensaio sobre as gramáticas do português**. Campinas, SP: Ed. da UNICAMP, 2001.
- GONÇALVES, F. **Riqueza morfológica e aquisição da sintaxe em português europeu e brasileiro**. 2004. Tese (Doutoramento) – Universidade de Évora, Évora. 2004.
- GONÇALVES, M. F. H. S. L. **Para uma redefinição do parâmetro do sujeito nulo**. Dissertação (Mestrado em Linguística Portuguesa Descritiva) – Faculdade de Letras da Universidade de Lisboa, Lisboa, 1994.
- KATO, M. A.; DUARTE, M. E. L. **A variação entre construções finitas pessoais e impessoais no português brasileiro**. Web-Revista SOCIODIALETO. v. 4, 12: 153177. 2014a.
- _____. 2014b. **Restrições na distribuição de sujeitos nulos no português brasileiro**. *Revista Veredas*. v. 18, 1, p. 1-22. 2014b.

- HOLMBERG, A. **Is there a little *pro*?** Evidence from Finnish. *Linguistic Inquiry*, v. 36, p. 533-564, 2005.
- HOLMBERG, A. Null Subject Parameter, In: BIBERAUER, T.; HOLMBERG, A.; ROBERTS, I.; SHEEHAN, M. (Org.) **Parametric Variation: Null Subjects in Minimalist Theory**. Cambridge: Cambridge University Press, 2010. p. 88-124.
- _____.; NAYUDU, A.; SHEEHAN, M. **Three Partial null-subject languages:** a comparison of Brazilian Portuguese, Finnish and Marathi. *Studia Linguistica*, n. 63, p. 59-97, 2009.
- HUANG, J. C. T. **On the Distribution and Reference of Empty Pronouns**. *Linguistic Inquiry*, n. 15, p. 531-574, 1984.
- _____. Pro-drop in Chinese: a generalized control theory. In: JAEGGLI, O.; SAFIR, K. J. **The null subject parameter**. Dordrecht, London : Kluwer Academic Publishers. p. 185-214, 1989.
- LI, C.; THOMPSON, S. Subject and topic: a new typology of language. In: LI, C.(Ed.). **Subject and Topic**. London, New York: Academic Press, 1976. p. 457-489. MIOTO, C.; FIGUEIREDO SILVA, M. C.; LOPES, R. E. V. **Novo Manual de Sintaxe**. 2. ed. Florianópolis: Insular, 2005.
- MORAES DE CASTILHO, C. M. Seria quatrocentista o português implantado no Brasil. In: SILVA, R. V.M. (Org.) **Para a história do português brasileiro**. v. 2. São Paulo: Humanitas, 2001. (Primeiros estudos).
- RIZZI, L. **A parametric approach to comparative syntax properties of the pronominal system**. HAEGEMAN, L. (Ed.). *The New Comparative Syntax*. London, New York: Longman, 1997. p. 268-285.
- ROBERTS, I. **On the nature of syntactic parameters:** a programme for research. In: GALVES, C. et al. (Ed.). *Parameter Theory and Linguistic Change*. Oxford: Oxford University Press, 2012. p. 319-334.
- _____. HOLMBERG, A. Introduction: parameters in minimalist theory. In: BIBERAUER, T.; HOLMBERG, A.; ROBERTS, I.; SHEEHAN, M. **Parametric Variation: Null Subjects in Minimalist Theory**. Cambridge: Cambridge University Press, 2010. p. 1-57.
- RODRIGUES, C. **Impoverished morphology and A-movement out of Case domains**. Tese de Doutorado. Maryland: University of Maryland, 2004.
- SAAB, A. On the notion of partial (non-) pro-drop in Romance. In: KATO, M. A.; ORDÓÑEZ, F. (Ed.). **The morphosyntax of Portuguese and Spanish in Latin America**. Oxford, Oxford University Press, 2016. p. 49-77.
- SALLES, A. A. F. **O fenômeno “sujeito duplo” em PB**. Dissertação (Mestrado em Estudos da Linguagem), Universidade Estadual de Londrina, Londrina. 2004.
- SILVA, C. R. T. **A natureza de AGR e suas implicações na ordem VS: um estudo comparativo entre o português brasileiro e o português europeu**. Tese (Doutorado em Linguística) – Universidade Federal de Alagoas, Maceió. 2004.
- _____. Comportamento e natureza dos sujeitos duplicados no crioulo caboverdiano e no português falado em comunidades quilombolas. In: MOURA, M. D.; SIBALDO, M. A. (Org.). **Para a História do Português Brasileiro: Sintaxe comparativa entre o português brasileiro e línguas crioulas de base lexical portuguesa**. Maceió: EDUFAL, 2013, p. 167-206.
- TARALDSEN, K. T. **On the NIC, vacuos application and the that-trace filter**. Indiana University Linguistics Club, Bloomington, 1978.
- TARALLO, F. Diagnosticando uma gramática brasileira: o português d'aquém e d'além mar ao final do século XIX. In: ROBERTS, I.; KATO, M. A (Org.). **Português**

brasileiro: uma viagem diacrônica. 2. ed. São Paulo: Ed. da UNICAMP, 1996. p. 69106.

